Some Applications of Combinatorial Hopf Algebras to Integro-Differential Equations and Symmetric Function Identities

by

Nicholas Olson-Harris

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Combinatorics and Optimization

Waterloo, Ontario, Canada, 2024

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Nantel Bergeron
Professor, Dept. of Mathematics and Statistics,
York University

Supervisor: Karen Yeats
Professor, Dept. of Combinatorics and Optimization,
University of Waterloo

Internal Member: Olya Mandelshtam
Assistant Professor, Dept. of Combinatorics and Optimization,
University of Waterloo

Internal Member: David Wagner
Professor, Dept. of Combinatorics and Optimization,
University of Waterloo

Internal-External Member: Jason Bell
Professor, Dept. of Pure Mathematics,
University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Chapters 1, 2 and 5 were written solely by the author of this thesis under the supervision of Karen Yeats. Chapters 3 and 4 incorporate material from a paper [8] which was coauthored with Paul-Hermann Balduf, Amelia Cantwell, Kurusch Ebrahimi-Fard, Lukas Nabergall, and Karen Yeats. Section 4 of [8], which forms the basis for Sections 3.3 and 4.2 of this thesis, is primarily the work of the author with some contributions from other coauthors.

Abstract

Hopf algebras built from combinatorial objects have found application both within combinatorics and, following the work of Connes and Kreimer, in quantum field theory. Despite the apparent gulf between these areas, the types of Hopf algebras that arise are very similar. We use Hopf algebra techniques to solve two problems, one coming from quantum field theory and one from algebraic combinatorics.

(1) *Dyson–Schwinger equations* are a formulation of the equations of motion of quantum field theory. From a mathematical perspective they are integro-differential equations which have a recursive, tree-like structure. In some cases, these equations are known to have solutions which can be written as combinatorial expansions over connected chord diagrams. We give a new expansion in terms of rooted trees equipped with a kind of decomposition we call a *binary tubing*. This is similar to the chord diagram expansion, but holds in greater generality, including to systems of Dyson–Schwinger equations and to Dyson–Schwinger equations in which insertion places are distinguished by different variables in the Mellin transform. Moreover we prove these results as a direct application of a purely Hopf-algebraic theorem characterizing maps from the Connes–Kreimer Hopf algebra of rooted trees (and variants thereof) to the Hopf algebra of univariate polynomials which arise from the universal property of the former.

(2) A pair of skew Ferrers shapes are said to be *skew-equivalent* if they admit the same number of semistandard Young tableaux of each weight, or in other words if the skew Schur functions they define are equal. McNamara and van Willigenburg conjectured necessary and sufficient combinatorial conditions for this to happen but were unable to prove either direction in complete generality. Using Hopf-algebraic techniques building on a partial result of Yeats, we prove sufficiency.

Acknowledgements

I thank my supervisor Karen Yeats for giving me interesting problems to work on, for her assistance in the process of writing and editing this thesis, and and most of all just for putting up with me for over five years. I am also indebted to my other collaborators Paul Balduf, Mimi Cantwell, Kurusch Ebrahimi-Fard, and Lukas Nabergall on the binary tubings project, to which a large chunk of this thesis directly or indirectly owes its existence.

I thank my committee members Olya Mandelshtam, David Wagner, Jason Bell, and Nantel Bergeron for agreeing to read my thesis and for their patience in dealing with an endless stream of indefensibly optimistic claims as to when it would be finished.

Over the years I have had numerous interesting conversations with various people related to the work of this thesis. Many, many people contributed some mathematical insight to this thesis, whether they meant to or not. Among those not already mentioned I would like to thank in no particular order Oliver Pechenik, William Dugan, Andrew Sack, Ben Moore, Cam Marcott, Lucas Gagnon, Steph van Willigenburg, Kel Chan, Sam Yusim, Alejandro Morales, Lucia Rotheray, Dirk Kreimer, Erik Panzer, Ali Mahmoud, and Farid Aliniaeifard.

Finally, some non-mathematical contributions are also in need of acknowledgement. I thank Fernanda for believing in me, James for trying to show up to my defence, Ben for actually showing up to my defence, Santiago for reading my thesis, Kel for eventually figuring out my name, and everyone who comes to board game night for providing entertaining distractions when I needed them least. As for you, Adwait, you know what you did.

Dedication

To the memory of George Olson

Table of Contents

List of Figures

List of Symbols

Symbols which begin with a Greek or Latin letter are listed first in alphabetical order. Non-alphabetical symbols are then listed in order of appearance. Symbols which are given explicit definitions in the text are listed with the page number on which they are defined.

$b(\tau)$	number of tubes containing the original root in a binary tubing τ, 37
$\beta^k(\tau)$	kth β-vector of a binary tubing τ of an edge-decorated tree, 50
$\mathrm{cat}(n)$	nth Catalan number, $\frac{1}{n+1}\binom{2n}{n}$
$\mathfrak{ch}(H)$	Lie algebra of infinitesimal characters of a Hopf algebra H, 18
$\mathrm{Ch}(H)$	character group of a Hopf algebra H, 18
$\chi(D,\alpha)$	Murnaghan–Nakayama coefficient, 80
$\mathrm{col}(D)$	number of columns of a skew shape D, 75
Comp	set of all integer compositions, 7
Δ_C	coproduct of a coalgebra C (subscript often omitted)
$\mathrm{dia}(D)$	number of diagonals occupied by a skew shape D, 75
$e(P)$	number of linear extensions of a poset P
ε_C	counit of a coalgebra C (subscript often omitted)
FdB	Faà di Bruno Hopf algebra, 28
$\Gamma(H)$	grouplike elements of bialgebra H, 12
$\mathcal{H}$	undecorated Connes–Kreimer Hopf algebra of rooted trees, 23
$\mathcal{H}_I$	I-decorated Connes–Kreimer Hopf algebra of rooted trees, 35
$\widetilde{\mathcal{H}}_r$	edge-decorated Connes–Kreimer Hopf algebra of rooted trees, 47
$\mathrm{H}^1(H,M)$	first cohomology of the comodule M over the bialgebra H, 31
$\mathrm{ht}\,A$	height of a ribbon A, equal to the number of rows minus 1, 80
id_V	identity operator on a vector space V (subscript often omitted)
$\mathbb{K}$	the base field
$K_\downarrow(E)$	bottom key ribbon of edge shape E, 96

$K_{\uparrow}(E)$	top key ribbon of edge shape E, 96
$\ell(\alpha)$	length of a composition α, 7
lin	operator sending a univariate polynomial to the coefficient of its linear term, 20
$\mathrm{mel}(\tau)$	Mellin monomial of a binary tubing τ, 41 and 50
$\mathbb{N}$	natural numbers (including 0)
$\mathbb{N}_+$	natural numbers excluding 0
$\mathrm{NW}(D)$	northwest border of a skew shape D, 80
$\mathbf{NW}(D)$	northwest decomposition of a skew shape D, 90
$\mathrm{od}(v)$	outdegree of a vertex v, 22
$\mathbf{od}(v)$	outdegree vector of a vertex v of a decorated tree, 62
$\mathrm{pack}(\alpha)$	composition obtained by deleting zeroes from the exponent vector α, 27
Par	set of all integer partitions, 7
$\mathrm{Prim}\, H$	Lie algebra of primitive elements of a bialgebra H, 12
$\mathbb{Q}$	rational numbers
QSym	quasisymmetric functions, 26
$\mathbb{R}$	real numbers
Rio	Riordan Hopf algebra, 65
$\mathrm{rk}(\tau, v)$	rank of a vertex v in a binary tubing τ, 37
$\mathbf{rk}(\tau, v)$	rank vector of a vertex v in a binary tubing τ of an edge-decorated tree, 49
$\mathrm{rt}\, t$	root of a rooted tree t, 22
$\mathrm{RT}(D, \alpha)$	set of ribbon tableaux of shape D and weight α, 80
$\mathrm{row}(D)$	number of rows of a skew shape D, 75
$\mathrm{SE}(D)$	southeast border of a skew shape D, 80
$\mathbf{SE}(D)$	southeast decomposition of a skew shape D, 90
$\mathcal{S}$	Hopf algebra of skew shapes, 76
$\mathrm{sort}(\alpha)$	partition obtained by sorting an exponent vector, 7
$\mathrm{Span}\, S$	subspace spanned by a subset S of a vector space
$\mathrm{SSYT}(D)$	set of semistandard tableaux of shape D, 77
Sym	symmetric functions, 23
$\mathcal{T}(S)$	set of unlabelled trees with vertices decorated by elements of S, 35
$\mathrm{Tub}(t)$	set of binary tubings of a rooted tree t, 37
$\xi(D)$	number of northbound or southbound boxes of a skew shape D, 110

$\mathbb{Y}$	Young's lattice, 74
$\mathbb{Z}$	integers
$\mathrm{Z}^1(H, M)$	vector space of 1-cocycles on the comodule M over the bialgebra H, 31
$a^{\underline{k}}$	falling factorial $a(a-1)\cdots(a-k+1)$, 6
$\alpha \vDash n$	α is a composition of n, 7
$\lambda \vdash n$	λ is a partition of n, 7
$V^\vee$	graded dual of graded vector space V, 15
$\alpha \rightharpoonup m$	left action of an element $\alpha \in C^*$ on an element m of a right C-comodule, 17
$m \leftharpoonup \alpha$	right action of an element $\alpha \in C^*$ on an element m of a left C-comodule, 17
$\lessdot$	covering relation of a poset, 21
$f^\perp$	operator on Sym adjoint to multiplication by f, 26
$E \sqcup_W E$	amalgamation of WOW shape E with itself, 90
$E \cdot_W E$	near-amalgamation of WOW shape E with itself, 90
$D \circ_W E$	WOW composition of D with E, 90
$D \square_W E$	modified WOW composition of D with E, 102 and 112

Chapter 1

Introduction

Hopf algebras are a well-known [10] class of algebraic structures. They first appeared in topology, taking their name from the work of Hopf [32] on the homology and cohomology of Lie groups. The cohomology of any manifold has a product (the *cup product*) making it into a ring, but Hopf observed that the group structure allows one to also define a so-called *coproduct*, and that the existence of this structure could be used to prove results about the Betti numbers of Lie groups and generalizations thereof. A structure of this nature, with both a product and a coproduct, is a *bialgebra*, and with a little bit of extra structure it is what we now call a *Hopf algebra*. (These terms will be defined precisely in Chapter 2.) Since Hopf's work these structures have extended their tendrils into geometry, representation theory, and many other branches of mathematics.

In a Hopf algebra H, the product is a map $H \otimes H \to H$ and the coproduct a map $H \to H \otimes H$. Joni and Rota [35] observed that these structures appear naturally in combinatorics. From a combinatorial perspective we think of multiplication as some way of *joining objects together* and comultiplication as *splitting objects apart*. Operations of this nature are the essence of combinatorics. It transpires that many examples can be found that fit together in the right way to form a bialgebra. Most of these bialgebras are in fact Hopf algebras, as was perhaps first clearly observed by Schmitt [52, 54]. In the decades since this pioneering work, combinatorial Hopf algebras[1] have become a significant area of algebraic combinatorics, with connections to representation theory [1, 65], category theory [4, 48, 53], polyhedral combinatorics [2], and many other areas. Perhaps the crown jewel of the field is the Aguiar–Bergeron–Sottile theorem [3] which shows how many deep and important invariants of combinatorial objects can be made to appear as if by magic from seemingly trivial Hopf-algebraic structures.

Two strands within the theory of combinatorial Hopf algebras are of particular interest to us in this thesis. The older of the two is the theory of symmetric functions. Of course, symmetric functions are a classical subject in mathematics, and have been a central theme of algebraic combinatorics for as long as such a subject has existed. Geissinger [27] observed that symmetric functions form a Hopf algebra, and that many of the most important opera-

[1] We do not use the term "combinatorial Hopf algebra" with any precise technical meaning, although multiple attempts have been made to give it one (e.g. [3, 39]). For us, both here and in the title of this thesis, we really mean nothing more or less than simply a Hopf algebra that somehow captures interesting combinatorial information.

tions and identities in symmetric function theory are naturally captured by this Hopf algebra structure. Despite this, the application of Hopf algebra techniques to symmetric function problems has remained a somewhat minor part of symmetric function theory as a whole, but we will see that it is capable of pulling its weight on certain problems. In particular, building on ideas of Yeats [63], we use it to attack a conjecture of McNamara and van Willigenburg [43] on identities between the *skew Schur functions*, a family of symmetric functions of deep importance in combinatorics and representation theory. Skew Schur functions are indexed by *skew Ferrers shapes* (these terms are defined in Section 5.1), and the particular problem we will be interested in is the question of when two two distinct shapes give rise to the *same* skew Schur function. McNamara and van Willigenburg conjectured necessary and sufficient combinatorial conditions for this; we will prove sufficiency (Theorem 5.3.9). While this ought to be the "easy" direction of the problem, it has remained open with little progress for 15 years.

The other strand of interest is the application of Hopf algebras to renormalization in quantum field theory, which began with the work of Connes and Kreimer [15, 38]. While at first blush this may seem rather far afield from combinatorics, it turns out that the Hopf algebras which appear in this theory are quite combinatorial indeed, being built from objects like graphs and rooted trees. The idea, roughly, is that the the coproduct on these objects—that is, the appropriate notion of *splitting objects apart*—encodes the recursive structures of the loop integrals that appear in Feynman's approach [20, 21] to quantum field theory. In perturbative quantum field theory one needs to *renormalize* such integrals, i.e. tweak their values in some hopefully well-defined way, in order to obtain answers that are correct (or even finite). The central theme of the work of Connes and Kreimer is that this procedure is naturally expressed as a computation in a Hopf algebra of Feynman diagrams.

A slightly different approach (and the one we will more or less take) is to instead begin with the equations of motion of the theory, the *Dyson–Schwinger equations*. These can be expressed in terms of certain integro-differential operators which, regarded as operating on polynomials, turn out to behave nicely with respect to the natural Hopf algebra structure: operators of this form are precisely the *1-cocycles* with respect to a certain cohomology theory for bialgebras. Abstracting away the details of these operators and remembering *only* this 1-cocycle property leads to considering Hopf algebras of rooted trees (the so-called *Connes–Kreimer Hopf algebras*) with the Dyson–Schwinger equations becoming functional equations for counting classes of trees in the spirit of classical enumerative combinatorics. This is the idea behind *combinatorial Dyson–Schwinger equations*, introduced by Bergbauer and Kreimer [9]. These have been studied extensively from a Hopf algebra perspective, notably by Foissy [22, 23, 24, 25]. Our goal is slightly different: we wish to use the Hopf algebra structure to find solutions to the *genuine* Dyson–Schwinger equations as combinatorial sums over trees. Our most general result of this nature (Theorem 4.3.2) encompasses some cases that are were already known to have combinatorial solutions, some cases that weren't, and some that had not been seriously studied before at all from this perspective.

1.1 Summary of this thesis

In Chapter 2 we will cover the relevant background material on Hopf algebras and other subjects that we need for the rest of the thesis.

1.1.1 Binary tubings and 1-cocycles

In Chapter 3 we will study the 1-cocycle operators we mentioned in passing above. These originate in a cohomology theory for bicomodules over coalgebras originally introduced by Doi [16] but we will specialize to the case of a left comodule M over a bialgebra H. (All of these terms are defined in Chapter 2.) We will define exactly what 1-cocycles are in Section 3.1, but for now it is enough to say that they are certain linear maps $M \to H$, and that the primary example is the case $M = H$. Despite the abstract nonsense, 1-cocycles have a natural combinatorial interpretation in terms of trees. Connes and Kreimer [15] showed that a certain Hopf algebra $\mathcal{H}$ of rooted forests (which we will meet in Section 2.3.1) is the *universal* commutative bialgebra possessing a 1-cocycle, which is simply the operator B_+ that joins together a forest with a new root to form a tree. For any other bialgebra H with a 1-cocycle Λ there is an induced map $\varphi\colon \mathcal{H} \to H$ satisfying $\varphi B_+ = \Lambda\varphi$.

In the case $H = \mathbb{K}[z]$, a polynomial algebra in one variable with its natural Hopf algebra structure, 1-cocycles are certain integro-differential operators. In this case, we give a non-recursive combinatorial formula (Theorem 3.3.1) for the induced map $\mathcal{H} \to \mathbb{K}[z]$. This result is due to the author and previously appeared in a joint paper [8] with Balduf, Cantwell, Ebrahimi-Fard, Nabergall, and Yeats in the context of solving Dyson–Schwinger equations. However, while this is the motivation, it has nothing intrinsically to do with Dyson–Schwinger equations and can be simply viewed a result in pure Hopf algebra. The formula is a sum over certain decompositions of a tree known as *binary tubings*. These are a special case of the notion of tubings (also called *pipings*) of posets which appear in the theory of Galashin's poset associahedron [26].

In fact, Theorem 3.3.1 applies more generally to maps that are universal with respect to a *family* of 1-cocycles, in which case we must replace $\mathcal{H}$ with a similarly defined Hopf algebra of forests with decorated vertices. We can further generalize to consider 1-cocycles defined not on the comodule H but on a tensor power $H^{\otimes r}$. In this case it turns out that the universal object is a Hopf algebra of forests which have decorated *edges* (and also vertices, if we want to consider families). This Hopf algebra and its universal property have not to the author's knowledge been considered before. After proving the basic properties, we derive a tubing formula (Theorem 3.5.1) for this context as well.

1.1.2 Dyson–Schwinger equations

In Chapter 4 we study several properties of combinatorial Dyson–Schwinger equations. First, we give some background (Section 4.1) on the physical significance of the equations, then we get to work applying the results of Chapter 3 to solve these equations combinatorially. In the past, solutions to certain Dyson–Schwinger equations have been found as certain generating functions for *rooted connected chord diagrams* by Marie and Yeats [41] in a special case and then Hihn and Yeats [31] in a more general one. The virtue of these expansions is that while the number of chord diagrams of size n (and hence the number of degree n terms in the expansion) grows superexponentially fast, the contribution of each diagram is relatively "small" and easy to reason about or compute.

We give an expansion (Theorem 4.2.6) with similar properties to the chord diagram expansion but using binary tubings of trees instead. The expansion itself offers slightly more generality than the chord diagram expansion (though it is likely that the chord diagram expansion could be extended to these cases with some work). More important, however, is that while conceptual explanations have been lacking for the chord diagram expansion, the tubing expansion is ultimately derived directly from the universal property of Connes–Kreimer Hopf algebras. Moreover, the Hopf-algebraic framework lends itself nicely to generalization: with no extra work, we also get an expansion (Theorem 4.2.10) for *systems* of Dyson–Schwinger equations, which did not have a known chord diagram expansion. Both of these results appear in [8] and are easy corollaries of Theorem 3.3.1, using the fact that the operators which appear in the Dyson–Schwinger equations are in fact 1-cocycles.

We then move to studying *Dyson–Schwinger equations with distinguished insertion places*, a physically natural generalization of the previously considered Dyson–Schwinger equations. While these have been studied before to some degree, little is known, and we have to develop the appropriate analogue of the theory of combinatorial Dyson–Schwinger equations for this case. These turn out to be formulated in terms of 1-cocycles on tensor products, and with the combinatorial theory in hand we can apply Theorem 3.5.1 to get tubing expansions for these types of equations as well (Theorem 4.3.2).

One property that ordinary Dyson–Schwinger equations are expected to have, at least in nicer cases, is that their solutions also satisfy the *renormalization group equation*. In Section 4.4 we give a (somewhat) novel interpretation of this equation in terms of the *Riordan group*. We then use this to prove that solutions of Dyson–Schwinger equations do indeed satisfy the equation; for ordinary DSEs this is a known fact (and more or less a dressed up version of the known proof) but for the equations with distinguished insertion places it is a new result which was previously conjectured by Nabergall [45].

1.1.3 Skew equivalence

Finally, in Chapter 5 we make a sharp turn away from physics and towards symmetric functions. As previously mentioned, we are interested in the problem of determining when two skew shapes define the same skew Schur function; such shapes are said to be *skew-equivalent*. Much work has been put into this problem [11, 34, 43, 51, 61, 63] but the general case is quite difficult. However, McNamara and van Willigenburg [43] gave a conjectural answer in terms of an operation called *WOW composition*. While defining this operation and stating the conjecture are beyond the scope of an introduction, we summarize as follows: if D and D' are skew-equivalent, we should be able to build new shapes $D \circ_W E$ and $D' \circ_W E$ in some yet-to-be-specified way such that these shapes too are equivalent, and moreover all equivalences should arise this way starting from a certain trivial case.

Perhaps surprisingly, even the sufficiency direction of the conjecture is difficult; in fact, McNamara and van Willigenburg were unable to prove that $D \circ_W E$ and $D' \circ_W E$ are actually equivalent in the full generality that they conjectured they should be (although they obtained substantial partial results). Our main result (Theorem 5.3.9) is that this is indeed the case.

Our methods are quite distinct from those used by McNamara and van Willigenburg: while they mainly use certain determinantal identities, we use Hopf-algebraic methods. The idea, pioneered by Yeats [63], is to consider a certain Hopf algebra defined directly on skew shapes and consider the map to the Hopf algebra of symmetric functions that sends a shape to its skew Schur function. Notably, the Hopf algebra of symmetric functions is *cocommutative* while the Hopf algebra of shapes is not, and this can be exploited to prove certain relations between skew Schur functions. In Section 5.2 we develop a framework based on this idea which allows Yeats's combinatorial arguments to be drastically simplified.

Having built our framework, we will spend Section 5.3 fully explaining the definitions and stating the conjectures. The problem naturally falls into two cases which we call the *edge case* and the *corner case*. McNamara and van Willigenburg already proved the result in the corner case, so our most important work is done in Section 5.4 where we prove the edge case. However, in Section 5.5 we also give a novel proof of the corner case using the same Hopf-algebraic ideas. Finally, in Section 5.6 we prove another of McNamara and van Willigenburg's conjectures and in doing so build a bridge between our techniques and theirs.

Chapter 2

Background

2.1 Polynomials, power series, and exponent vectors

2.1.1 Notation and conventions

We will work throughout over a field $\mathbb{K}$ of characteristic 0.[1] All vector spaces are over $\mathbb{K}$, and the symbol $\otimes$ will always denote the tensor product over $\mathbb{K}$.

> **Remark 2.1.1.** It is not strictly necessary to assume that $\mathbb{K}$ is a field: in almost all cases a commutative $\mathbb{Q}$-algebra will do so long as the term "vector space" is interpreted as "free $\mathbb{K}$-module". More generality is possible in many cases but is not required for the intended combinatorial and physical applications.

As is standard, for an algebra A we write $A[[x]]$ for the algebra of formal power series in the indeterminate x with coefficients in A, and $A[x]$ for the subalgebra of polynomials. For a power series $A(x) \in A[[x]]$, we write $[x^n]A(x)$ for the coefficient of the monomial x^n. In some cases it is more convenient to expand not in terms of monomials but divided powers $x^n/n!$; we will then use the notation $[x^n/n!]A(x) = n![x^n]A(x)$.

Let I be any set, finite or infinite. For a vector $\alpha \in \mathbb{N}^I$, we write

$$|\alpha| = \sum_{i \in I} \alpha_i$$

and say that α is an *exponent vector* if $|\alpha| < \infty$. Given an exponent vector α and an element $\mathbf{a} \in A^I$ for some algebra A, we write

$$\mathbf{a}^\alpha = \prod_{i \in I} a_i^{\alpha_i}.$$

In particular, if $\mathbf{x} = (x_i)_{i \in I}$ and A is any algebra, we can consider the multivariate polynomial algebra $A[\mathbf{x}]$ and power series algebra $A[[\mathbf{x}]]$, which consist respectively of finite and infinite A-linear combinations of monomials $\mathbf{x}^\alpha$.

We will use the underline notation for falling factorial powers

$$a^{\underline{k}} = \prod_{j=0}^{k-1} (a - j)$$

[1] Determining which, if any, of our results hold in characteristic p is left as an exercise to the reader.

and we will generalize this to exponent vectors in the same way as we do with ordinary powers:

$$\mathbf{a}^{\underline{\alpha}} = \prod_{i \in I} a_i^{\underline{\alpha_i}}.$$

2.1.2 Binomial coefficients

As usual, the binomial coefficients are defined by

$$\binom{a}{k} = \frac{a^{\underline{k}}}{k!}$$

where a is an element of some algebra. They satisfy the *Pascal recurrence*

$$\binom{a}{k} = \binom{a-1}{k} + \binom{a-1}{k-1}. \tag{2.1}$$

Most important is the case where $a \in \mathbb{N}$, in which case the Pascal recurrence can be recursively applied to the first term on the right side of (2.1) repeatedly to get the *hockey-stick identity*

$$\binom{n}{k} = \sum_{j=1}^{n} \binom{j}{k-1}. \tag{2.2}$$

We will also generalize binomial coefficients to vectors; for $\mathbf{a} \in A^I$ and α an exponent vector write

$$\binom{\mathbf{a}}{\alpha} = \prod_{i \in I} \binom{a_i}{\alpha_i}.$$

We partially order $\mathbb{N}^I$ by treating it as a product of copies of $\mathbb{N}$; i.e. $\alpha \geq \beta$ if and only if $\alpha_i \geq \beta_i$ for all $i \in I$. Using this and the vector binomial coefficient notation, we get a kind of multivariate binomial theorem:

$$(\mathbf{x} + \mathbf{y})^\alpha = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \mathbf{x}^\beta \mathbf{y}^{\alpha - \beta}. \tag{2.3}$$

2.1.3 Compositions and partitions

An *(integer) composition* of size n and length k is a finite list $\alpha = (\alpha_1, \ldots, \alpha_k)$ of positive integers such that $\alpha_1 + \cdots + \alpha_k = n$. While strictly speaking these are a special case of the integer vectors considered in the previous subsections, we treat them somewhat differently: when working with integer vectors we fix some indexing set I, whereas we want to think of compositions of different lengths as being "the same kind of object". We denote the length of a composition α by $\ell(\alpha)$ and the set of all compositions by Comp. We will also write $\alpha \vDash n$ to mean α is a composition of size n.

A composition is called an *(integer) partition* if its entries are weakly decreasing. We write $\lambda \vdash n$ to mean that λ is a partition of size n, and denote the set of all partitions by Par. Partitions are closely related to exponent vectors: if α is an exponent vector (over any indexing set) there is a unique partition $\mathrm{sort}(\alpha)$ consisting of the nonzero entries of α arranged in decreasing order. Clearly two exponent vectors α and β are equal up to permuting entries if and only if $\mathrm{sort}(\alpha) = \mathrm{sort}(\beta)$. On the other hand, we also often find it useful to identify a partition $\lambda = (\lambda_1, \dots, \lambda_k)$ with the $\mathbb{N}_+$-indexed exponent vector $(\lambda_1, \dots, \lambda_k, 0, 0, \dots)$ obtained by concatenating an infinite sequence of zeroes to it, and we will silently make this identification wherever it is convenient. In particular, λ_i should always be interpreted as zero when $i > \ell(\lambda)$. With this in mind, we partially order partitions by saying $\lambda \geq \mu$ if $\lambda_i \geq \mu_i$ for all $i \in \mathbb{N}_+$; thinking of partitions as exponent vectors this agrees with the ordering defined in the previous subsection.

2.1.4 Lagrange inversion

The *Lagrange inversion formula* allows solving for the compositional inverse of a power series. It will be of mild interest to us.

Theorem 2.1.2 (Lagrange Inversion Theorem [56, Theorem 5.4.2])**.** *Suppose $G(x) \in \mathbb{K}[[x]]$ has zero constant term and nonzero linear term. Then there is a unique series $R(x)$ satisfying $G(R(x)) = R(G(x)) = x$. Moreover, for any series $F(x)$, we have*

$$[x^n]F(R(x)) = \frac{1}{n}[x^{n-1}]F'(x)\left(\frac{x}{G(x)}\right)^n.$$

The *Lagrange implicit function theorem* is an equivalent formulation that is useful in enumeration of tree-like structures.

Theorem 2.1.3 (Lagrange Implicit Function Theorem)**.** *Let $A(x) \in \mathbb{K}[[x]]$ be a power series with nonzero constant term. There exists a unique series $R(x) \in \mathbb{K}[[x]]$ such that $R(x) = xA(R(x))$. For any series $F(x)$,*

$$[x^n]F(R(x)) = \frac{1}{n}[x^{n-1}]F'(x)A(x)^n$$

for $n > 0$.

Proof. Note that $R(x) = xA(R(x))$ can be rewritten as

$$\frac{R(x)}{A(R(x))} = x$$

i.e. $R(x)$ is the compositional inverse of $x/A(x)$. The result follows from Theorem 2.1.2. □

2.2 Hopf algebras

We give a thorough, albeit somewhat terse, overview of the of the relevant parts of the theory of Hopf algebras. Nearly everything covered in this section is contained in the union of [14, 28, 44], but unfortunately there seems to be no single existing reference that includes all we need.

2.2.1 Algebras, coalgebras, and bialgebras

Let A be an algebra. Writing $m_A\colon A\otimes A\to A$ for the multiplication map and $u_A\colon \mathbb{K}\to A$ for the unique linear map sending 1 to 1, the associativity and unitality conditions can be rephrased as the statement that the diagrams

$$\begin{tikzcd}
A\otimes A\otimes A \arrow[r, "m_A\otimes \mathrm{id}_A"] \arrow[d, "\mathrm{id}_A\otimes m_A"] & A\otimes A \arrow[d, "m_A"] \\
A\otimes A \arrow[r, "m_A"] & A
\end{tikzcd}
\qquad
\begin{tikzcd}
A \arrow[r, "u_A\otimes \mathrm{id}_A"] \arrow[d, "\mathrm{id}_A\otimes u_A"'] \arrow[dr, "\mathrm{id}_A"] & A\otimes A \arrow[d, "m_A"] \\
A\otimes A \arrow[r, "m_A"] & A
\end{tikzcd}$$

commute.

A (counital coassociative) *coalgebra* is a vector space C equipped with linear maps $\Delta_C\colon C\to C\otimes C$ (the *coproduct*) and $\varepsilon_C\colon C\to\mathbb{K}$ (the *counit*) making the dual diagrams

$$\begin{tikzcd}
C \arrow[r, "\Delta_C"] \arrow[d, "\Delta_C"] & C\otimes C \arrow[d, "\mathrm{id}_C\otimes\Delta_C"] \\
C\otimes C \arrow[r, "\Delta_C\otimes \mathrm{id}_C"] & C\otimes C\otimes C
\end{tikzcd}
\qquad
\begin{tikzcd}
C \arrow[r, "\Delta_C"] \arrow[d, "\Delta_C"'] \arrow[dr, "\mathrm{id}_C"] & C\otimes C \arrow[d, "\varepsilon_C\otimes \mathrm{id}_C"] \\
C\otimes C \arrow[r, "\mathrm{id}_C\otimes\varepsilon_C"] & C
\end{tikzcd}$$

commute. We will often drop the subscripts and simply write Δ and ε when the coalgebra in question is clear from context.

We can think of coassociativity as saying that for any k there is only one map $C\to C^{\otimes k}$ that can be built from tensor products of Δ_C and id_C. Let us denote this by Δ_C^k, again omitting the subscript when clear.[2] Explicitly, it can be recursively defined by

$$\Delta_C^{k+1} = (\Delta_C^k\otimes \mathrm{id}_C)\Delta_C$$

with the base case $\Delta_C^0=\varepsilon_C$. We may check that by coassociativity we also have $\Delta_C^{k+1} = (\mathrm{id}_C\otimes\Delta_C^k)\Delta_C$.

A coalgebra C is *cocommutative* if $T\Delta_C=\Delta_C$, where $T\colon C\otimes C\to C\otimes C$ is the *twist map* given by swapping tensor factors. (Note that for an algebra, the identity $m_AT=m_A$ is easily seen to be equivalent to commutativity.)

Another use of twist maps is in defining the tensor product of algebras or coalgebras. Suppose A and B are algebras. Recall that the tensor product $A\otimes B$ is naturally an algebra with multiplication

$$(a\otimes b)(a'\otimes b') = aa'\otimes bb'$$

with the identity clearly given by $1\otimes 1$. In the linear-map formulation, this says

$$m_{A\otimes B} = (m_A\otimes m_B)(\mathrm{id}_A\otimes T\otimes \mathrm{id}_B)$$

and

$$u_{A\otimes B} = u_A\otimes u_B.$$

[2] Other notations for this exist: what we call Δ^k is written as Δ_k in [14], Δ_{k-1} in [44], and $\Delta^{(k-1)}$ in [28]. Our convention has the flaw that we somewhat unnervingly have $\Delta^1=\mathrm{id}$ and $\Delta^2=\Delta$ but we consider this preferable to having an off-by-one between the notations for the map and its codomain. In any case, while the iterated coproduct notation will occasionally be convenient, we will not be making heavy use of it.

Dually, if C and D are coalgebras then $C \otimes D$ is naturally a coalgebra with coproduct

$$\Delta_{C\otimes D} = (\mathrm{id}_C \otimes T \otimes \mathrm{id}_D)(\Delta_C \otimes \Delta_D)$$

and counit

$$\varepsilon_{C\otimes D} = \varepsilon_C \otimes \varepsilon_D.$$

If A and B are algebras, a linear map $\varphi\colon A \to B$ is an algebra morphism if $\varphi(ab) = \varphi(a)\varphi(b)$ and $\varphi(1) = 1$. This is equivalent to requiring that the diagrams

$$\begin{tikzcd} A \otimes A \arrow[r, "m_A"] \arrow[d, "\varphi\otimes\varphi"] & A \arrow[d, "\varphi"] \\ B \otimes B \arrow[r, "m_B"] & B \end{tikzcd} \qquad \begin{tikzcd} \mathbb{K} \arrow[r, "u_A"] \arrow[dr, "u_B"] & A \arrow[d, "\varphi"] \\ & B \end{tikzcd}$$

commute. Dually, if C and D are coalgebras, a linear map $\varphi\colon C \to D$ is a *coalgebra morphism* if

$$\begin{tikzcd} C \arrow[r, "\Delta_C"] \arrow[d, "\varphi"] & C \otimes C \arrow[d, "\varphi\otimes\varphi"] \\ D \arrow[r, "\Delta_D"] & D \otimes D \end{tikzcd} \qquad \begin{tikzcd} C \arrow[r, "\varphi"] \arrow[dr, "\varepsilon_C"] & D \arrow[d, "\varepsilon_D"] \\ & \mathbb{K} \end{tikzcd}$$

commute.

A *bialgebra* is a vector space H equipped with maps $m_H, u_H, \Delta_H, \varepsilon_H$ making it into an algebra and a coalgebra such that the diagrams

$$\begin{tikzcd} H \otimes H \arrow[r, "m_H"] \arrow[d, "\Delta_H\otimes\Delta_H"] & H \arrow[dd, "\Delta_H"] \\ H \otimes H \otimes H \otimes H \arrow[d, "\mathrm{id}_H\otimes T\otimes \mathrm{id}_H"] & \\ H \otimes H \otimes H \otimes H \arrow[r, "m_H\otimes m_H"] & H \otimes H \end{tikzcd} \qquad \begin{tikzcd} \mathbb{K} \arrow[r, "u_H"] \arrow[dr, "u_H\otimes u_H"'] & H \arrow[d, "\Delta_H"] \\ & H \otimes H \end{tikzcd}$$

$$\begin{tikzcd} H \otimes H \arrow[r, "m_H"] \arrow[dr, "\varepsilon_H\otimes\varepsilon_H"'] & H \arrow[d, "\varepsilon_H"] \\ & \mathbb{K} \end{tikzcd} \qquad \begin{tikzcd} \mathbb{K} \arrow[r, "u_H"] \arrow[dr, "\mathrm{id}_{\mathbb{K}}"] & H \arrow[d, "\varepsilon_H"] \\ & \mathbb{K} \end{tikzcd}$$

commute. Comparing these diagrams to those we used to define morphisms, the first row says exactly that Δ_H is an algebra morphism and the second that ε_H is an algebra morphism. On the other hand, the first column says that m_H is a coalgebra morphism and the second that u_H is a coalgebra morphism. Thus these conditions are equivalent, and are both equivalent to being a bialgebra.

Suppose C is a coalgebra and A is an algebra. If $\varphi, \psi\colon C \to A$ are linear maps, their *convolution* is $\varphi * \psi = m_A(f \otimes g)\Delta_C$. Convolution is associative, and has an identity given by $u_A\varepsilon_C$. Thus the space of linear maps $C \to A$ is itself an algebra. If C is cocommutative and A is commutative the convolution product is commutative. In particular, taking $A = \mathbb{K}$, the dual space C^* is an algebra with unit ε_C which is commutative if and only if C is cocommutative. (We will discuss the duality between algebras and coalgebras further in the next subsection.)

If H is a bialgebra, it makes sense to convolve two linear maps $H \to H$. We say H is a *Hopf algebra* if id_H has a (two-sided) convolution inverse. This inverse is unique if it exists, and is known as the *antipode* and denoted S_H (as usual, without subscript when clear). The condition that S_H be the convolution inverse of the identity can also be expressed as the assertion that the diagram

$$\begin{tikzcd}
& H \otimes H \arrow[r, "S_H \otimes \mathrm{id}_H"] & H \otimes H \arrow[dr, "m_H"] & \\
H \arrow[ur, "\Delta_H"] \arrow[r, "\varepsilon_H"] \arrow[dr, "\Delta_H"'] & \mathbb{K} \arrow[r, "u_H"] & & H \\
& H \otimes H \arrow[r, "\mathrm{id}_H \otimes S_H"'] & H \otimes H \arrow[ur, "m_H"'] &
\end{tikzcd}$$

commutes. Let us now establish some basic properties of the antipode.

Proposition 2.2.1 ([14, Proposition 3.1.1] or [28, Proposition 1.4.10] or [44, Proposition 1.5.10])**.** *Let H be a Hopf algebra.*

(i) The antipode reverses multiplication: $S_H(ab) = S_H(b)S_H(a)$.

(ii) The antipode reverses comultiplication: $\Delta_H S_H = (S_H \otimes S_H)T\Delta_H$ where T is the twist map.

In general, if A and A' are algebras, a map $A \to A'$ which reverses multiplication is an *algebra anti-morphism*. Similarly if C and C' are coalgebras a map $C \to C'$ which reverses comultiplication is a *coalgebra anti-morphism*. If H and H' are bialgebras, we will a map $H \to H'$ which reverses both multiplication and comultiplication is a *bialgebra anti-morphism*; thus the antipode is the primary example.[3] Another important property is the following.

Proposition 2.2.2 ([14, Lemma 3.1.2] or [28, Corollary 1.4.12] or [44, Corollary 1.5.12])**.** *If H is either commutative or cocommutative, $S_H^2 = \mathrm{id}_H$.*

Let us now, finally, see some actual examples of Hopf algebras.

Example 2.2.3. Let G be a group. The *group algebra* $\mathbb{K}G$ is the free vector space on G with multiplication given by linearly extending the group operation. We make $\mathbb{K}G$ into a bialgebra by setting $\Delta g = g \otimes g$ and $\varepsilon(g) = 1$ for all $g \in G$. For maps $\varphi, \psi \colon \mathbb{K}G \to A$ for any algebra A, the convolution is given by

$$(\varphi * \psi) = \varphi(g)\psi(g) \tag{2.4}$$

and from this it easily follows that $\mathbb{K}G$ is a Hopf algebra with antipode $S(g) = g^{-1}$.

[3] In general one might also be interested in maps which preserve multiplication and reverse comultiplication or vice versa, giving four different variations in total, but since the bialgebras of interest to us will be commutative or cocommutative there will only be morphisms and anti-morphisms.

Example 2.2.4. We make the polynomial algebra $\mathbb{K}[z]$ into a (graded) bialgebra by taking the coproduct and counit as the unique algebra morphisms extending $\Delta z = 1 \otimes z + z \otimes 1$ and $\varepsilon(z) = 0$. Explicitly, by the binomial theorem, the coproduct is given on the basis of monomials by

$$\Delta(z^n) = (\Delta z)^n = \sum_{k=0}^{n} \binom{n}{k} z^k \otimes z^{n-k}. \tag{2.5}$$

and the counit is given (on all polynomials) by $\varepsilon(f(z)) = f(0)$. This too is a Hopf algebra, with antipode $S(f(z)) = f(-z)$. (Indeed, from (2.5) we easily see $(\mathrm{id} * S)(z^n) = (z - z)^n$, which agrees with the counit.)

It is often profitable to think of this coproduct in a different way. We can identify $\mathbb{K}[x] \otimes \mathbb{K}[z]$ with $\mathbb{K}[z_1, z_2]$ (where $z^j \otimes z^k$ corresponds to $z_1^j z_2^k$). The coproduct then simply corresponds to the map $f(z) \mapsto f(z_1 + z_2)$. We will make extensive use of this perspective in various parts of this thesis.

Example 2.2.5. One can also construct Hopf algebras of a more combinatorial flavour. A representative example is the bialgebra $\mathcal{G}$ of graphs (sometimes known as the *chromatic Hopf algebra* for reasons that will be explained in Example 2.2.25). As a vector space, $\mathcal{G}$ is freely spanned by unlabelled simple graphs. The multiplication is disjoint union. Equivalently, $\mathcal{G}$ is therefore a free commutative algebra generated by the *connected* graphs. The coproduct of a graph G is

$$\Delta G = \sum_{S \sqcup T = V(G)} G|_S \otimes G|_T$$

where $G|_S$ denotes the subgraph induced by S. That $\mathcal{G}$ is actually a *Hopf* algebra is not as clear as the previous examples (though it will follow from Lemma 2.2.11) but Humpert and Martin [33] gave an explicit and reasonably nice formula for the antipode as a sum over acyclic orientations.

Examples 2.2.3 and 2.2.4 illustrate two very important special types of elements of bialgebras. Given a bialgebra H, a *grouplike element* is an element $g \in H$ such that $\Delta g = g \otimes g$. We denote the set of these by $\Gamma(H)$. A *primitive element* is an element p such that $\Delta p = 1 \otimes p + p \otimes 1$; we denote the vector space of these by $\operatorname{Prim} H$. Some of the properties from Examples 2.2.3 and 2.2.4 generalize to such elements in other Hopf algebras.

Proposition 2.2.6 ([14, Lemma 3.4.1])**.** *Let H be a bialgebra. The product of any two grouplike elements is grouplike, hence $\Gamma(H)$ forms a monoid under multiplication. Moreover, if H is Hopf, every grouplike element g is invertible with $g^{-1} = S(g)$, and thus $\Gamma(H)$ forms a group.*

Proposition 2.2.7 ([14, Lemma 3.1.1])**.** *The commutator of any two primitive elements is primitive, and thus* $\operatorname{Prim} H$ *forms a Lie algebra.*

Thus from any Hopf algebra we can get a group and a Lie algebra. We have already seen that, conversely, we can get a Hopf algebra from a group. We can also get a Hopf algebra from a Lie algebra.

Example 2.2.8. Let $\mathfrak{g}$ be any Lie algebra. Recall that the universal enveloping algebra $\mathcal{U}(\mathfrak{g})$ satisfies the following universal property: for any algebra A, any linear map $\varphi\colon \mathfrak{g} \to A$ satisfying $\varphi([x,y]) = \varphi(x)\varphi(y) - \varphi(y)\varphi(x)$ extends uniquely to an algebra morphism $\mathcal{U}(\mathfrak{g}) \to A$. With this in mind, we can make $\mathcal{U}(\mathfrak{g})$ into a bialgebra by defining the coproduct and the counit to be the unique algebra morphisms extending $\Delta x = 1 \otimes x + x \otimes 1$ and $\varepsilon(x) = 0$ for $x \in \mathfrak{g}$. This too is a Hopf algebra, with antipode given by $S(x) = -x$ for $x \in \mathfrak{g}$. It can be shown[4] that $\operatorname{Prim} \mathcal{U}(\mathfrak{g}) = \mathfrak{g}$.

Example 2.2.8 is in fact more general than it may seem: a large class of Hopf algebras, which we will now define, are actually of this form. Let us say that a bialgebra H is *unipotent* if for every $h \in H$ there exists some $k > 0$ such that $(\mathrm{id} - \varepsilon)^{*k}(h) = 0$. Note that if h is product of n primitive elements, we can always take $k = n + 1$, so any bialgebra generated by primitive elements (such as a universal enveloping algebra) is unipotent.

Theorem 2.2.9 (Takeuchi)**.** *Any unipotent bialgebra is a Hopf algebra, with antipode given by*

$$S = \sum_{k \geq 0} (-1)^k (\mathrm{id} - \varepsilon)^{*k}.$$

Proof. The unipotency implies that the infinite sum is well-defined, as for any given element only finitely many terms are nonzero. But this is just a geometric series for $(\varepsilon - (\mathrm{id} - \varepsilon))^{*-1} = \mathrm{id}^{*-1}$ so it is indeed an antipode. $\square$

A celebrated result of Cartier[5] gives a complete structure theorem for cocommutative unipotent Hopf algebras.

Theorem 2.2.10 ([14, Theorem 4.3.1] or [44, Theorem 5.6.5])**.** *Suppose H is a cocommutative unipotent Hopf algebra. Then $H \cong \mathcal{U}(\operatorname{Prim} H)$.*

A *graded vector space* is a vector space V equipped with a direct sum decomposition $V = \bigoplus_{n \in \mathbb{N}} V_n$. The tensor product of two graded vector spaces V and W is also graded, with the grading given by

$$(V \otimes W)_n = \bigoplus_{0 \leq k \leq n} V_k \otimes W_{n-k}.$$

A *graded algebra* is an algebra A with a grading on its underlying vector space such that $1 \in A_0$ and $A_m A_n \subseteq A_{m+n}$; that is, the unit and product maps preserve the grading (thinking of $\mathbb{K}$ as concentrated in degree 0). Dually, a *graded coalgebra* is a coalgebra C with a grading such that $\Delta_C C_n \subseteq (C \otimes C)_n$ and ε_C vanishes on the positive-degree pieces. A *graded bialgebra* is a bialgebra with a grading respected by both the algebra and coalgebra structures. A graded vector space V is *connected* if $\dim V_0 = 0$ and is of *finite type* (also called *locally finite-dimensional*) if all graded pieces are finite-dimensional. Many bialgebras of interest to us satisfy both of these conditions. Connectedness is an especially important condition due to the following key lemma.

[4] For instance using the Poincaré–Birkhoff–Witt theorem.
[5] Commonly known as the Milnor–Moore theorem.

Lemma 2.2.11. *Every connected graded bialgebra is unipotent, and in particular a Hopf algebra.*

Proof. Suppose H is a connected graded bialgebra. Then H_0 is spanned by 1, so $\mathrm{id} - \varepsilon$ vanishes on H_0. But then if h is any element of degree less than k, every term in an expansion of $\Delta^k h$ must have an H_0 tensor factor, so $(\mathrm{id} - \varepsilon)^{*k}(h) = 0$. $\square$

Remark 2.2.12. Throughout this thesis we will deal with various combinatorially graded bialgebras. Most of these are connected: if the bialgebra has a basis indexed by some class of combinatorial objects and is graded by some natural notion of size, connectedness corresponds to the very reasonable condition of having a unique object of size 0. Thus by Lemma 2.2.11 these bialgebras are Hopf algebras. However, the role played by the antipode will in most cases be somewhat small.

2.2.2 Duality

We have already seen that if C is a coalgebra, its dual C^* forms an algebra with the convolution product. A different way to think about this is in terms of the map $\Delta_C^* \colon (C \otimes C)^* \to C^*$ adjoint to the coproduct on C. By definition, this sends $\rho \in (C \otimes C)^*$ to the composition $\rho\Delta_C$. Restricted to the subspace $C^* \otimes C^*$, this says

$$\Delta_C^*(\varphi \otimes \psi) = (\varphi \otimes \psi)\Delta_C = \varphi * \psi.$$

Thus m_{C^*} is exactly the restriction of Δ_C^*. Similarly, using the natural identification of $\mathbb{K}$ with $\mathbb{K}^*$, we also have $u_{C^*} = \varepsilon_C^*$. One virtue of thinking of the dual in this way is that it makes it immediately clear that if $f \colon C \to D$ is a coalgebra morphism, then f^* is an algebra morphism. As such, duality defines a contravariant functor from the category of coalgebras to the category of algebras.[6]

We would like to also be able to dualize algebras to get coalgebras. Indeed, if A is a *finite-dimensional* algebra, then $(A \otimes A)^* = A^* \otimes A^*$ and we can simply make A^* into a coalgebra using m_A^* as the coproduct and u_A^* as the counit. Moreover, we clearly have that $A \cong A^{**}$ as coalgebras. When A is infinite-dimensional, however, m_A^* need not take values in the proper subspace $A^* \otimes A^*$, so this construction does not work. To resolve it we introduce the *restricted dual* or *finite dual* A°, which consists of those elements of A^* which vanish on some ideal of finite codimension.[7]

Theorem 2.2.13 ([14, Theorem 2.12.1] or [44, Proposition 9.1.2])**.** *Let A be an algebra. Then $m_A^* A^\circ \subseteq A^\circ \otimes A^\circ$, and A° forms a coalgebra with coproduct m_A^* and counit u_A^*.*

Suppose $f \colon A \to B$ is an algebra morphism. If $\varphi \in B^\circ$ then there is some ideal I of finite codimension contained in the kernel of φ. Then $f^{-1}(I)$ is an ideal, necessarily of finite codimension, which is contained in the kernel of $f^*(\varphi)$. Thus $f^*(B^\circ) \subseteq A^\circ$, and restricted duality defines a contravariant functor from algebras to coalgebras. In particular, if H is a

[6] The reader may be either relieved or disappointed to learn that the word "functor" does not appear outside of this section.

[7] Warning: the notation A° is used in [28] for the *graded* dual, which we denote $A^\vee$.

bialgebra, then Δ_H^* gives a coalgebra morphism $(H \otimes H)^\circ \to H^\circ$. This observation gets us most of the way to the following result.

Theorem 2.2.14 ([44, Theorem 9.1.3])**.** *Let H be a bialgebra. Then H° is a subalgebra of H^*, and forms a bialgebra with the inherited algebra structure and the coalgebra structure from Theorem 2.2.13. Moreover, if H is a Hopf algebra then $S_H^* H^\circ \subseteq H^\circ$ and H° is a Hopf algebra with antipode S_H^*.*

Example 2.2.15. By (2.4), the dual $(\mathbb{K}G)^*$ can be identified with the algebra of functions $G \to \mathbb{K}$ with pointwise multiplication. The restricted dual $(\mathbb{K}G)^\circ$ is the subalgebra spanned by matrix coefficients of finite-dimensional representations of G, with a coproduct dual to matrix multiplication. See [14, Section 2.7].

The *graded dual* $V^\vee$ of a graded vector space V is the subspace of V^* consisting of linear forms that vanish on V_n for all but finitely many n. Equivalently,

$$V^\vee = \bigoplus_{n \in \mathbb{N}} V_n^*$$

and this decomposition gives a grading on $V^\vee$ as well. Using this grading, we clearly have $V^{\vee\vee} \cong V$ when V is of finite type. If C is a graded coalgebra, then the fact that the coproduct respects the grading implies that $C^\vee$ is a subalgebra of C^*, and that the product respects the grading on $C^\vee$. Hence $C^\vee$ is a graded algebra. (Note that C^* is not naturally graded in general.) On the other hand, if A is a graded algebra of finite type then $\bigoplus_{n \geq N} A_n$ is an ideal of finite codimension for any N, and we similarly get that $A^\vee$ is a graded coalgebra. It follows that if H is a graded bialgebra (resp. Hopf algebra) of finite type then $H^\vee$ is a graded bialgebra (resp. Hopf algebra): the compatibility between the algebra and coalgebra structures is immediate as they are both inherited from H°.

Example 2.2.16. An element $\varphi \in \mathbb{K}[x]^*$ is uniquely determined by the sequence $(\varphi(1), \varphi(x), \varphi(x^2), \dots)$. From (2.5) we see that convolution corresponds to multiplying the exponential generating functions of these sequences. Thus $\mathbb{K}[x]^*$ can be identified with $\mathbb{K}[[x]]$. The corresponding pairing $\langle -, - \rangle \colon \mathbb{K}[x] \otimes \mathbb{K}[[x]] \to \mathbb{K}$ is given by

$$\langle x^n, A(x) \rangle = [x^n/n!]A(x)$$

or equivalently

$$\langle f(x), A(x) \rangle = f(\partial/\partial x)A(x)|_{x=0}. \tag{2.6}$$

From this and the fact that ideals in $\mathbb{K}[x]$ are principal, one can deduce that the restricted dual $\mathbb{K}[x]^\circ$ is isomorphic to the subalgebra of $\mathbb{K}[[x]]$ consisting of power series that satisfy a linear differential equation with constant coefficients. (See [44, Example 9.1.7] for a proof of this in different language.)

By definition, $\varphi \in \mathbb{K}[x]^\vee$ if there are only finitely many n such that $\varphi(x^n) \neq 0$; in other words, if its generating function is a polynomial. Thus we have $\mathbb{K}[x]^\vee \cong \mathbb{K}[x]$ as algebras, and indeed the coproducts can be seen to match as well, so $\mathbb{K}[x]$ is (graded) self-dual.

2.2.3 Modules and comodules

Let A be an algebra. A *left module* over A is a vector space M with a linear map $\alpha\colon A\otimes M \to M$ (the *action*) such that the diagrams

$$\begin{tikzcd} A\otimes A\otimes M \arrow[r, "\mathrm{id}_A\otimes\alpha"] \arrow[d, "m_A\otimes\mathrm{id}_M"] & A\otimes M \arrow[d, "\alpha"] \\ A\otimes M \arrow[r, "\alpha"] & M \end{tikzcd} \qquad \begin{tikzcd} M \arrow[r, "u_A\otimes\mathrm{id}_M"] \arrow[dr, "\mathrm{id}_M"'] & A\otimes M \arrow[d, "\alpha"] \\ & M \end{tikzcd}$$

commute. Writing $am = \alpha(a\otimes m)$, these say $(ab)m = a(bm)$ and $1m = m$, matching the more familiar definition of modules. We can similarly define a *right module* with a map $M\otimes A \to A$ making the obvious analogous diagrams commute. Basic examples of modules include A itself, ideals and quotients of A, and direct sums of these. If $\varphi\colon A \to B$ is a morphism of algebras and M we can "pull back" B-modules to A-modules: that is, if M is a left B-module we can make the same vector space M into a left A-module by defining $am = \varphi(a)m$, and similarly for right modules.

If A and A' are algebras, an (A, A')-*bimodule* is a vector space M equipped with action maps $\alpha\colon A\otimes M \to M$ and $\alpha'\colon M\otimes A' \to M$ making it into a left A-module and right A'-module respectively, such that the diagram

$$\begin{tikzcd} A\otimes M\otimes A' \arrow[r, "\alpha\otimes\mathrm{id}_{A'}"] \arrow[d, "\mathrm{id}_A\otimes\alpha'"] & M\otimes A' \arrow[d, "\alpha'"] \\ A\otimes M \arrow[r, "\alpha"] & M \end{tikzcd}$$

also commutes. In particular, A is an (A, A)-bimodule (in which case this compatibility is just associativity) and bimodules pull back to bimodules, so an algebra morphism $A \to B$ makes B into an (A, A)-bimodule.

Let H be a bialgebra. The coalgebra structure on H translates to certain structures on the category of (left, right) H-modules. Since ε_H is an algebra morphism, we can pull back $\mathbb{K}$-modules (i.e. vector spaces) to H-modules using ε_H. In particular, we can make $\mathbb{K}$ itself into an H-module this way, which we call the *trivial module*. On the other hand, since Δ_H is also an algebra morphism, we can pull back $(H\otimes H)$-modules to H-modules as well. In particular, if M and N are H-modules then $M\otimes N$ is naturally an $(H\otimes H)$-module so becomes an H-module as well by pulling back using the coproduct. The coassociativity and counitality properties imply that this tensor product of H-modules is associative up to natural isomorphism with unit object the trivial module $\mathbb{K}$. These operations also make sense for (H, H')-bimodules when H and H' are both bialgebras.

Straight from the definition, we see that grouplike and primitive elements act particularly nicely on tensor products: grouplikes satisfy

$$g(a\otimes b) = (ga)\otimes(gb) \tag{2.7}$$

and primitives satisfy

$$p(a\otimes b) = a\otimes(pb) + (pa)\otimes b. \tag{2.8}$$

Remark 2.2.17. We see that, (2.7) matches the usual definition from group representation theory of the tensor product of representations, and similarly (2.8) matches the usual definition from Lie algebra theory. The notion of trivial module for bialgebras also generalizes the notion of trivial representation for groups and Lie algebras.

The concept of modules can be dualized. Let C be a coalgebra. A *left C-comodule* is a vector space M equipped with a map $\delta\colon M \to C \otimes M$ (the *coaction*) such that the diagrams

$$\begin{tikzcd} M \arrow[r, "\delta"] \arrow[d, "\delta"] & C \otimes M \arrow[d, "\mathrm{id}_C \otimes \delta"] \\ C \otimes M \arrow[r, "\Delta_C \otimes \mathrm{id}_M"] & C \otimes C \otimes M \end{tikzcd} \qquad \begin{tikzcd} M \arrow[r, "\delta"] \arrow[dr, "\mathrm{id}_M"'] & C \otimes M \arrow[d, "\varepsilon_C \otimes \mathrm{id}_M"] \\ & M \end{tikzcd}$$

commute. Analogously, a *right C-comodule* is a vector space M equipped with a linear map $\delta\colon M \to M \otimes C$ such that the analogous diagrams commute. Comodules over coalgebras are the dual notion to modules over algebras in the same way that coalgebras are the dual notion to algebras. Any coalgebra C is both a left and right comodule over itself, in both cases with $\delta = \Delta$. Rather than pulling back, comodules can be pushed forward: if $\varphi\colon C \to C'$ is a coalgebra morphism, and M is a left C-comodule with coaction δ, we can make M into a left C'-comodule the coaction $(\varphi \otimes \mathrm{id}_M)\delta$, and similarly with right comodules.

Inevitably, a *bicomodule* is the dual notion to a bimodule. That is, if C and C' are coalgebras, a (C, C')-bicomodule is a vector space M with maps $\delta\colon M \to C \otimes M$ and $\delta'\colon M \to M \otimes C'$ making it into a left C-comodule and right C' comodule, such that the diagram

$$\begin{tikzcd} M \arrow[r, "\delta"] \arrow[d, "\delta'"] & C \otimes M \arrow[d, "\mathrm{id}_C \otimes \delta'"] \\ M \otimes C' \arrow[r, "\delta \otimes \mathrm{id}_{C'}"] & C \otimes M \otimes C' \end{tikzcd}$$

commutes. In particular, C itself is a (C, C)-bicomodule, and any coalgebra map $C \to D$ makes C into a (D, D)-bicomodule by pushing forward.

In the case of a bialgebra H rather than a mere coalgebra, the unit and product give operations on comodules similar to the operations on modules we previously mentioned. We will describe these for left comodules, though the same ideas work for right comodules. First, if V is any vector space, the *trivial coaction* of H on V is given by $\delta(v) = 1 \otimes v$. Second, if M and N are left H-comodules (with coactions δ_M and δ_N respectively) then $M \otimes N$ becomes a left H-comodule as well, with its coaction given as the composite map

$$M \otimes N \xrightarrow{\delta_M \otimes \delta_N} H \otimes M \otimes H \otimes N \to H \otimes H \otimes M \otimes N \xrightarrow{m_H \otimes \mathrm{id}_M \otimes \mathrm{id}_N} H \otimes M \otimes N$$

where the middle arrow switches the two middle tensor factors. This construction is associative up to natural isomorphism and has a unit object given by $\mathbb{K}$ with the trivial coaction.

If M is a left C-comodule with coaction δ and $\alpha \in C^*$, define $m \leftharpoonup \alpha = (\alpha \otimes \mathrm{id}_M)\delta(m)$. We can compute

$$\begin{aligned} (m \leftharpoonup \alpha) \leftharpoonup \beta &= (\beta \otimes \mathrm{id}_M)\delta((\alpha \otimes \mathrm{id}_M)\delta(m)) \\ &= (\beta \otimes \mathrm{id}_M)(\alpha \otimes \delta)\delta(m) \end{aligned}$$

$$
\begin{aligned}
&= (\alpha \otimes \beta \otimes \mathrm{id}_M)(\mathrm{id}_C \otimes \delta)\delta(m) \\
&= (\alpha \otimes \beta \otimes \mathrm{id}_M)(\Delta_C \otimes \mathrm{id}_M)\delta(m) \\
&= ((\alpha * \beta) \otimes \mathrm{id}_M)\delta(m) \\
&= m \leftharpoonup (\alpha * \beta)
\end{aligned}
$$

so $\leftharpoonup$ makes M into a *right* module over C^*. Analogously, if M is a right C-comodule then we define $\alpha \rightharpoonup m = (\mathrm{id}_M \otimes \alpha)\delta(m)$ and this makes M into a left C^*-module. If M is a (C, D)-bicomodule then these operations are compatible and make M into a (D^*, C^*)-bimodule. As a particular case, of this, any coalgebra C is naturally a (C^*, C^*)-bimodule.

> **Remark 2.2.18.** If A is a finite-dimensional algebra, one can easily dualize this construction to turn modules over A into comodules over A^*. When A is infinite-dimensional, it turns out that not all A-modules can be naturally made into A°-comodules, but some of them can. A necessary and sufficient condition is that every element of M is contained in a finite-dimensional submodule; see [44, Lemma 1.6.4].

Consider a bialgebra H. Then H becomes an (H^*, H^*)-bimodule by the above construction, and we can restrict these actions to the subalgebra H°. It turns out that the actions also play nicely with the coalgebra structure of H°, more specifically with the tensor product of H°-bimodules. The following is an exercise in unravelling definitions.

Proposition 2.2.19. *Let H be a bialgebra. Then the multiplication map $H \otimes H \to H$ is an (H°, H°)-bimodule morphism.*

> **Example 2.2.20.** Continuing from Example 2.2.16 we identify $\mathbb{K}[z]^*$ with $\mathbb{K}[[z]]$. The action is given by taking the coproduct, which we think of as substituting in a sum of two variables as in Example 2.2.4, and then applying the pairing (2.6) in one tensor factor i.e. in one of the two variables:
>
> $$A(z) \rightharpoonup f(z) = A(\partial/\partial y) f(y+z)|_{y=0} = A(d/dz) f(z).$$

2.2.4 Characters

A *character* of a bialgebra H is an algebra morphism $H \to \mathbb{K}$. This can equivalently be expressed in terms of the operator m_H^* adjoint to multiplication: $\zeta \in H^*$ is a character if and only if

$$m_H^* \zeta = \zeta m_H = \zeta \otimes \zeta.$$

In other words, characters of H are precisely grouplike elements of H°. (Note that the kernel of a character is an ideal, necessarily of codimension 1, so they do lie in the restricted dual.) It follows from Proposition 2.2.6 that characters of H form a monoid $\mathrm{Ch}(H) = \Gamma(H^\circ)$ under convolution, with identity element the counit, and that if H is Hopf then $\mathrm{Ch}(H)$ is a group with inverse given by $\zeta^{*-1} = \zeta S$.

An *infinitesimal character* of a bialgebra H is a map $\sigma \colon H \to \mathbb{K}$ satisfying

$$\sigma(ab) = \sigma(a)\varepsilon(b) + \varepsilon(a)\sigma(b),$$

or in other words a derivation of H into the trivial H-module $\mathbb{K}$. Infinitesimal characters are the same as primitive elements of H°. They form a Lie algebra $\mathfrak{ch}(H) = \operatorname{Prim} H^\circ$.

Characters and infinitesimal characters behave nicely with respect to the actions of H^* on H discussed in the previous section.

Proposition 2.2.21. *Suppose H is a bialgebra.*

(i) For $\zeta \in \mathrm{Ch}(H)$ and $a, b \in H$ we have

$$\zeta \rightharpoonup ab = (\zeta \rightharpoonup a)(\zeta \rightharpoonup b)$$

and

$$ab \leftharpoonup \zeta = (a \leftharpoonup \zeta)(b \leftharpoonup \zeta).$$

In other words, ζ acts on both left and right as an automorphism of H.

(ii) For $\sigma \in \mathfrak{ch}(H)$ and $a, b \in H$ we have

$$\sigma \rightharpoonup ab = a(\sigma \rightharpoonup b) + (\sigma \rightharpoonup a)b$$

and

$$ab \leftharpoonup \sigma = a(b \leftharpoonup \sigma) + (a \leftharpoonup \sigma)b.$$

In other words, σ acts on both left and right as a derivation of H.

Proof. Immediate from Proposition 2.2.19. □

As the notation suggests, $\mathfrak{ch}(H)$ and $\mathrm{Ch}(H)$ are, in nice cases, related by an exponential map. In particular, if H is a unipotent (e.g. connected graded) Hopf algebra and σ is an infinitesimal character, the infinite sum

$$\exp_*(\sigma) = \sum_{n \geq 0} \frac{\sigma^{*n}}{n!}$$

is well-defined. Moreover, we can compute

$$\begin{aligned} m_H^* \exp_*(\sigma) &= \exp_*(\sigma \otimes \varepsilon + \varepsilon \otimes \sigma) \\ &= \exp_*(\sigma \otimes \varepsilon) \exp_*(\varepsilon \otimes \sigma) \\ &= \exp_*(\sigma) \otimes \exp_*(\sigma) \end{aligned}$$

so $\exp_*(\sigma)$ is a character. In fact, it turns out that all characters are obtained this way.

Theorem 2.2.22 ([14, Corollary 3.4.1])**.** *Let H be a unipotent Hopf algebra. The map $\exp_* \colon \mathfrak{ch}(H) \to \mathrm{Ch}(H)$ is a bijection, with inverse the convolution logarithm*

$$\log_*(\zeta) = \sum_{k \geq 1} \frac{(-1)^k}{k} (\zeta - \varepsilon)^{*k}.$$

Characters and infinitesimal characters are intimately related to the polynomial Hopf algebra $\mathbb{K}[z]$. If $\sigma \in \mathfrak{ch}(H)$, then the same calculation with exponentials shows that $\exp_*(z\sigma) \colon H \to \mathbb{K}[z]$ is an algebra morphism. But we also have

$$\exp_*((z_1 + z_2)\sigma) = \exp_*(z_1\sigma) * \exp_*(z_2\sigma)$$

which, by the discussion in Example 2.2.4, is equivalent to

$$\Delta_{\mathbb{K}[z]} \exp_*(z\sigma) = (\exp_*(z\sigma) \otimes \exp_*(z\sigma))\Delta_H$$

i.e. $\exp_*(z\sigma)$ is also a coalgebra morphism. Again, we have a converse to this. Let $\operatorname{lin} \colon \mathbb{K}[z] \to \mathbb{K}$ denote the map which extracts the linear coefficient of a polynomial. Note that lin is itself an infinitesimal character.

Lemma 2.2.23. *Let H be a unipotent Hopf algebra and $\varphi \colon H \to \mathbb{K}[z]$ be a bialgebra morphism. Then $\operatorname{lin}\varphi$ is an infinitesimal character, and $\varphi = \exp_*(z \operatorname{lin}\varphi)$.*

Proof. Since φ is a coalgebra morphism, it induces an algebra morphism $\varphi^* \colon \mathbb{K}[[z]] \to H^*$ where we have identified $\mathbb{K}[z]^*$ with $\mathbb{K}[[z]]$ as in Example 2.2.16. By definition this map satisfies

$$\varphi^*(F(z))(a) = F(d/dz)\varphi(a)|_{z=0}.$$

Thus in particular we have

$$\varphi^*(z^n) = n![z^n]\varphi(a)$$

but since φ^* is an algebra morphism, it also satisfies

$$\varphi^*(z^n) = \varphi^*(z)^{*n} = (\operatorname{lin}\varphi)^{*n}.$$

The result follows. □

We can package things up nicely as follows.

Theorem 2.2.24. *Let H be a unipotent Hopf algebra and $\varphi \colon H \to \mathbb{K}[z]$ be an algebra morphism. Then $\operatorname{lin}\varphi$ is an infinitesimal character if and only if $\varphi|_{z=0} = \varepsilon$. Moreover, the following are equivalent:*

(i) φ *is a bialgebra morphism,*

(ii) $\varphi = \exp_*(z \operatorname{lin}\varphi)$,

(iii) $\varphi|_{z=0} = \varepsilon$ *and* $\frac{d}{dz}\varphi = (\operatorname{lin}\varphi) * \varphi$,

(iv) $\varphi|_{z=0} = \varepsilon$ *and* $\frac{d}{dz}\varphi = \varphi * (\operatorname{lin}\varphi)$.

Proof. Observe that

$$\operatorname{lin}\varphi(ab) = \operatorname{lin}(\varphi(a))\varphi(b)|_{z=0} + \varphi(a)|_{z=0}\operatorname{lin}(\varphi(b))$$

so it is immediate that $\operatorname{lin}\varphi$ is an infinitesimal character if and only if $\varphi|_{z=0} = \varepsilon$.

For the equivalence, note that (i) implies (ii) by Lemma 2.2.23. Conversely, (ii) implies (i) because if $\varphi = \exp_*(z \operatorname{lin}\varphi)$ then $\varphi|_{z=0} = \varepsilon$, so $\operatorname{lin}\varphi$ is an infinitesimal character and hence φ is a bialgebra morphism. Finally, that both (iii) and (iv) are equivalent to (ii) is a routine calculation. □

Combing Theorem 2.2.22 and Theorem 2.2.24, we get a diagram

$$\mathrm{Ch}(H) \underset{\exp_*}{\overset{\log_*}{\rightleftarrows}} \mathfrak{ch}(H) \underset{\mathrm{lin}}{\overset{\sigma\mapsto\exp_*(z\sigma)}{\rightleftarrows}} \mathbf{Bialg}(H, \mathbb{K}[z]).$$

of bijections and their inverses, where $\mathbf{Bialg}(H, \mathbb{K}[z])$ denotes the set of bialgebra morphisms $H \to \mathbb{K}[z]$. By composition we also have bijections $\mathrm{Ch}(H) \rightleftarrows \mathbf{Bialg}(H, \mathbb{K}[z])$ which are of interest as well. The bijection $\mathbf{Bialg}(H, \mathbb{K}[z]) \to \mathrm{Ch}(H)$ is simply evaluation at $z = 1$. Its inverse sends ζ to $\exp_*(z \log \zeta) = \zeta^{*z}$ which we can expand as a binomial series:

$$\zeta^{*z} = \sum_{n\geq 0} \binom{z}{n} (\zeta - \varepsilon)^{*n}.$$

Example 2.2.25. The bijection between $\mathrm{Ch}(H)$ and $\mathbf{Bialg}(H, \mathbb{K}[z])$ is especially important in the applications of Hopf algebras to combinatorics. As an illustrative example, consider the graph Hopf algebra $\mathcal{G}$ (Example 2.2.5) and let ζ be the incredibly dull character that sends a graph to 0 if it has any edges and 1 if it has none. Then from the definition of the coproduct we see that $\zeta^{*k}(G)$ counts ways to partition the vertices of G into k subsets each of which induces a subgraph with no edges, i.e. proper k-colourings of G. The corresponding morphism $\mathcal{G} \to \mathbb{K}[z]$ thus sends a graph to its *chromatic polynomial*, perhaps the most famous polynomial invariant of graphs that there is.

2.3 Partially ordered sets

For a more in-depth overview of posets see Stanley [57, Chapter 3]. We here review some basic terminology which we will make use of. A subset S of a poset is called a *downset* if $p \in X$ and $q \leq p$ implies $q \in X$ or an *upset* if $p \in X$ and $q \geq p$ implies $q \in X$. Note that S is a downset if and only if $P \setminus S$ is an upset. The set of all downsets of P is denoted $J(P)$ and is itself a poset (more specifically a distributive lattice) with the ordering given by inclusion.

A subset S is *convex* if $p, p' \in X$ and $p \leq q \leq p'$ implies $q \in X$; equivalently a convex subset is the intersection of an upset and a downset. Given $p \leq q$ the *(closed) interval* from p to q is the set

$$[p, q] = \{x \in P \colon p \leq x \leq q\}.$$

Clearly $[p, q]$ is a convex subset.

Suppose $p, q \in P$. We write $p \lessdot q$ and say that q *covers* p if $p < q$ and there is no $x \in P$ such that $p < x < q$.

Two elements $p, q \in P$ are *comparable* if $p \leq q$ or $q \leq p$. The *comparability graph* of P has vertex set P with edges between all pairs of distinct comparable elements. A poset is *connected* if its comparability graph is connected.

Let $\mathcal{P}$ denote the vector space freely generated by isomorphism classes of finite posets. We make $\mathcal{P}$ into a bialgebra with multiplication given by disjoint union. (Equivalently, $\mathcal{P}$ is the commutative algebra freely generated by isomorphism classes of *connected* posets.) The coproduct is given by

$$\Delta P = \sum_{X \in J(P)} X \otimes (P \setminus X). \tag{2.9}$$

We equip $\mathcal{P}$ with a grading such that the degree of a poset is the number of elements. This makes $\mathcal{P}$ into a connected graded bialgebra and hence a Hopf algebra. One can construct many similar Hopf algebras by considering special classes of posets and/or posets with additional structure; most of the combinatorial Hopf algebras considered throughout this thesis are of this type.

Example 2.3.1. As another nice example of the idea considered in Example 2.2.25, consider the character ζ which sends all posets to 1. Then $\zeta^{*k}(P)$ counts ways to partition P into disjoint subsets $P_1, \ldots, P_k$ where P_1 is a downset in P, then P_2 is a downset in $P \setminus P_1$, and so on. These are clearly in bijection with weakly order-preserving maps $P \to \{1, \ldots, k\}$. Thus the bialgebra map $\mathcal{P} \to \mathbb{K}[z]$ sends a poset to a polynomial $\Omega(P; z)$ with the property that $\Omega(P; k)$ counts such maps; this is known as the *order polynomial*.

Remark 2.3.2. Schmitt [54] defined a quite general way to construct a Hopf algebra from an appropriate family $\mathcal{F}$ of posets, the so-called *incidence Hopf algebra*. These work somewhat differently from the Hopf algebra $\mathcal{P}$ we have defined here. The posets in the family are required to be intervals. The product in the Hopf algebra is given by the product of posets, while the coproduct is

$$\Delta[a, b] = \sum_{a \leq p \leq b} [a, p] \otimes [p, b]. \tag{2.10}$$

The intervals are not necessarily considered up to isomorphism; rather one may choose any equivalence relation satisfying appropriate properties (a so-called *Hopf relation*).

Despite these differences, our Hopf algebra $\mathcal{P}$ is isomorphic to a certain an incidence Hopf algebra. Namely, consider the family of all finite distributive lattices with the Hopf relation just being isomorphism. The map $P \mapsto J(P)$ is a bialgebra isomorphism from $\mathcal{P}$ to the incidence Hopf algebra of this family; see [54, Section 16].

2.3.1 Rooted trees and Connes–Kreimer Hopf algebras

A *(rooted) forest* is a finite poset such that each element is covered by at most one element. A connected forest is a *(rooted) tree*.[8] A rooted tree can be decomposed as a unique maximal element, the *root*, together with a forest. For a rooted tree t we denote by root by $\operatorname{rt} t$. On the other hand a forest uniquely decomposes as a disjoint union of trees. Thus this notion of rooted tree is combinatorially equivalent to more common graph-theoretic definitions.

Although we prefer to think of trees as posets, we will sometimes use graph or tree-specific terminology. In particular, we often refer to elements of trees as vertices and covering relations as edges. The unique vertex covering a non-root vertex is its *parent*, vertices covered by a vertex are its *children*. We will think of a tree as oriented downwards (opposite the order) so that the number of children of a vertex is the *outdegree* and denoted $\operatorname{od}(v)$.

Two special classes of trees will be of some significance. The first is the *ladder* ℓ_n, which as a poset is simply a chain (and as a graph is simply a path rooted at one of its ends) but is conventionally called a ladder in this context. The second is the *corolla* s_n consisting of

[8] We will not have any need for unrooted trees or forests.

just a root with $n-1$ children.

It is clear that a disjoint union of forests is a forest and that any upset or downset in a forest is a forest. Thus the subspace of the poset Hopf algebra $\mathcal{P}$ spanned by forests is a Hopf subalgebra, the (undecorated) *Connes–Kreimer Hopf algebra*, which we denote $\mathcal{H}$. We can also characterize $\mathcal{H}$ in a more algebraic way. Consider the linear operator B_+ on $\mathcal{P}$ which sends each poset P to the poset obtained by adjoining a new element larger than all elements of P. Then $\mathcal{H}$ is the unique minimal subalgebra of $\mathcal{P}$ which is mapped to itself by B_+. From this perspective it is not immediately obvious that $\mathcal{H}$ should be a Hopf subalgebra. We can understand this by considering the relationship between B_+ and the coproduct. Note that the only downset of B_+P which contains the new element is the entirety of B_+P. The other downsets coincide with the downsets of P, and if D is such a downset we have $(B_+P) \setminus D \cong B_+(P \setminus D)$. It follows that

$$\Delta B_+P = B_+P \otimes 1 + \sum_{D \in J(P)} D \otimes B_+(P \setminus D)$$

or in other words

$$\Delta B_+ = B_+ \otimes 1 + (\mathrm{id} \otimes B_+)\Delta. \tag{2.11}$$

An operator satisfying (2.11) is a *1-cocycle*. (These form part of a cohomology theory which will be defined in Section 3.1.) The key significance of the Connes–Kreimer Hopf algebra is that it possesses a universal property with respect to 1-cocycles.

Theorem 2.3.3 (Connes–Kreimer [15, Theorem 2])**.** *Let A be a commutative algebra and Λ be a linear operator on A. There exists a unique algebra morphism $\varphi\colon \mathcal{H} \to A$ such that $\varphi B_+ = \Lambda\varphi$. Moreover, if A is a bialgebra and Λ is a 1-cocycle then φ is a bialgebra morphism.*

Note that there is nothing mysterious about the map φ guaranteed by Theorem 2.3.3. Since any tree t can be written *uniquely* (up to reordering) in the form $t = B_+(t_1 \cdots t_k)$ for some $t_1, \ldots, t_k$ we can and must define φ recursively by

$$\varphi(t) = \Lambda(\varphi(t_1) \cdots \varphi(t_k)). \tag{2.12}$$

A natural question is whether we can find an explicit, non-recursive formula for φ. Without knowing anything about A and Λ there is clearly nothing we can do, but in Chapter 3 we will solve certain cases of this problem.

2.4 Hopf algebras of note

In this section we set up some of the main examples of Hopf algebras which will be of interest to us in this thesis and collect together some their properties.

2.4.1 Symmetric functions

Surely the most important Hopf algebra in algebraic combinatorics is Sym, the Hopf algebra of symmetric functions. We can only scratch the surface of this subject; see [56, Chapter

7] for a full reference or [28, Chapter 2] for a more Hopf-algebraic perspective. Consider a sequence $\mathbf{x} = (x_1, x_2, \dots)$ of indeterminates indexed by $\mathbb{N}_+$. As an algebra, Sym is the subalgebra of $\mathbb{K}[[\mathbf{x}]]$ consisting of those power series which are of finite degree and which are invariant under all permutations of the variables. Since we take only the bounded-degree series, Sym is graded by degree, unlike the full power series algebra.

Recalling the notation from Section 2.1.3, for any integer partition λ we define the *monomial symmetric function*

$$m_\lambda = \sum_{\mathrm{sort}(\alpha)=\lambda} \mathbf{x}^\alpha.$$

summing over $\mathbb{N}_+$-indexed exponent vectors α. Since two exponent vectors are equal up to permuting entries if and only if the partitions obtained by sorting them agree, it follows that the monomial symmetric functions are a basis for Sym. In particular, the dimension of the graded piece Sym_n is the number of partitions of size n.

For $n \geq 1$, we define the *elementary symmetric function*

$$e_n = \sum_{i_1 < \cdots < i_n} x_{i_1} \cdots x_{i_n},$$

the *complete symmetric function*

$$h_n = \sum_{i_1 \leq \cdots \leq i_n} x_{i_1} \cdots x_{i_n},$$

and the *power sum*

$$p_n = \sum_i x_i^n.$$

For a partition $\lambda = (\lambda_1, \dots, \lambda_k)$, we define $e_\lambda = e_{\lambda_1} \cdots e_{\lambda_k}$, and analogously h_λ and p_λ.

Theorem 2.4.1 ([56, Theorem 7.4.4, Corollary 7.6.2, Corollary 7.7.2]). *All of $\{e_n\}_{n\in\mathbb{N}_+}$, $\{h_n\}_{n\in\mathbb{N}_+}$, and $\{p_n\}_{n\in\mathbb{N}_+}$ are algebraically independent generating sets for Sym. Consequently, $\{e_\lambda\}_{\lambda\in\mathrm{Par}}$, $\{h_\lambda\}_{\lambda\in\mathrm{Par}}$, and $\{p_\lambda\}_{\lambda\in\mathrm{Par}}$ are bases.*

The coproduct on Sym can most naturally be described as follows. We identify $\mathsf{Sym}\otimes\mathsf{Sym}$ with symmetric functions in two separate sets of variables $\mathbf{x}$ and $\mathbf{y}$. The coproduct then corresponds to the map $f(\mathbf{x}) \mapsto f(\mathbf{x}, \mathbf{y})$.

$$\Delta e_n = \sum_{k=0}^{n} e_k \otimes e_{n-k}, \tag{2.13}$$

$$\Delta h_n = \sum_{k=0}^{n} h_k \otimes h_{n-k}, \tag{2.14}$$

and

$$\Delta p_n = 1 \otimes p_n + p_n \otimes 1. \tag{2.15}$$

(Here for convenience we take $e_0 = h_0 = 1$.)

Remark 2.4.2. One might object that our definition of the coproduct doesn't actually make any sense. Certainly, for an arbitrary series $f(\mathbf{x}) \in \mathbb{K}[[\mathbf{x}]]$ it is completely unclear what $f(\mathbf{x}, \mathbf{y})$ should mean; we seem to be trying to make a substitution for the variables where the values we are substituting are not indexed by the same set as the variables themselves! Thus it is essential that our attention is restricted to symmetric functions. In this case there are (at least) two ways we can sense of what is going on. The first is to simply interleave the x-variables and y-variables arbitrarily to make a single sequence of variables; the symmetry ensures that the choice of how this is done does not affect the answer. The other is to declare that what we really mean is

$$f(\mathbf{x}, \mathbf{y}) = \lim_{k \to \infty} f(x_1, \dots, x_k, y_1, \dots, y_k, 0, 0, \dots) \tag{2.16}$$

where (as usual for power series) the notion of convergence here is that the coefficient of each monomial $\mathbf{x}^\alpha$ on the RHS eventually stabilizes to the coefficient on the LHS. We leave it as a straightforward exercise that this limit does in fact exist when f is symmetric.

An alternative, less satisfying, approach is to appeal to Theorem 2.4.1 and simply take one of (2.13) to (2.15) as the definition of the coproduct. One significant downside of this approach is that it is far from immediately obvious that the three formulas are actually equivalent without appealing to an interpretation like ours.

Remark 2.4.3. Since the power sums are primitive and are algebraically independent generators, we see that Sym is simply isomorphic to a polynomial bialgebra $\mathbb{K}[p_1, p_2, \dots]$. (Indeed, Theorem 2.2.10 implies that any commutative and cocommutative unipotent bialgebra is isomorphic to a polynomial bialgebra, since the Lie algebra structure is trivial.) Note however that this only works because we assumed $\mathbb{K}$ to be a field of characteristic zero. In Chapter 5 we will be interested in certain identities between symmetric functions with integer coefficients; even though we use characteristic-zero methods these identities must hold over *any* field because they ultimately boil down to some equalities between integers. This is one reason why we do not want to simply identify symmetric functions with polynomials like this.

Since Sym is a connected graded bialgebra, it is a Hopf algebra. Indeed, we have already seen how to generate Sym by primitives, so we can easily write down a formula for the antipode: $S(p_n) = -p_n$, and hence (by Proposition 2.2.1(i)) $S(p_\lambda) = (-1)^{\ell(\lambda)} p_\lambda$. More interesting, however, is that the antipode relates the elementary and complete bases.

Theorem 2.4.4 ([28, Theorem 2.4.1(ii, iii)]). *The antipode of* Sym *satisfies* $S(e_n) = (-1)^n h_n$ *and* $S(h_n) = (-1)^n e_n$.

In symmetric function theory it is traditional to work not with the antipode but with the *fundamental involution* ω, defined by $\omega(e_n) = h_n$. This is essentially equivalent to the antipode: if f is homogeneous of degree n then $\omega(f) = (-1)^n S(f)$.

By Remark 2.4.3 it follows (as in Example 2.2.16) that Sym is graded self-dual. To obtain an isomorphism $\mathsf{Sym} \to \mathsf{Sym}^\vee$ we may choose a primitive element $\psi_n \in \mathsf{Sym}_n^\vee$ for each n;

then the map defined by $p_n \mapsto \psi_n$ will do the job. The standard choice is to use

$$\psi_n(p_\lambda) = \begin{cases} n, & \lambda = (n) \\ 0, & \text{otherwise} \end{cases}$$

which has the nice property that it preserves integrality of coefficients. The *Hall inner product*[9] $\langle -, - \rangle$ is the bilinear form on Sym corresponding to this map. By definition, we can see $\langle p_n, p_n \rangle = n$ and more generally

$$\langle p_n^k, p_n^k \rangle = \psi_n^{*k}(p_n^k) = k!\psi_n(p_n)^k = k!n^k.$$

To write down a general formula in the power sum basis we introduce notation. For a partition λ we write $m_j(\lambda)$ for the number of i such that $\lambda_i = j$. Then define

$$z_\lambda = \prod_j m_j(\lambda)! j^{m_j(\lambda)}. \tag{2.17}$$

Then we have

$$\langle p_\lambda, p_\mu \rangle = \begin{cases} z_\lambda, & \lambda = \mu \\ 0, & \text{otherwise} \end{cases} \tag{2.18}$$

which can be verified by a similar convolution argument.

We will henceforth identify Sym with $\mathsf{Sym}^\vee$ using the map we have described. In particular, this gives left and right actions of Sym on itself as in Section 2.2.3. Since Sym is cocommutative these actions are the same. Most often these actions are described in terms of *perp operators*,

$$f^\perp g = f \rightharpoonup g = g \leftharpoonup f,$$

where an alternative definition is that $f^\perp$ is the operator adjoint to multiplication by f with respect to the Hall inner product.

> **Remark 2.4.5.** Those who are familiar with symmetric functions will surely notice that there is one more commonly studied basis which we have not yet mentioned, the *Schur functions*. Schur functions and their cousins the skew Schur functions will be the main subject of Chapter 5, but we defer their definition until then.

2.4.2 Quasisymmetric functions

While our main interest will be in symmetric functions, we would be remiss not to make a brief mention of their cousins the *quasisymmetric functions*. Let us return our attention to the map $f(\mathbf{x}) \mapsto f(\mathbf{x}, \mathbf{y})$ that we used to describe the coproduct of Sym. We may heuristically think of quasisymmetric functions as being precisely the class of (bounded-degree) series on which this defines a coproduct.

[9]Strictly speaking, this is only an "inner product" in the usual sense when $\mathbb{K}$ is a subfield of $\mathbb{R}$; otherwise it is merely a nondegenerate symmetric bilinear form. This (ab)use of terminology is nonetheless standard.

To give a proper definition, let us again consider indeterminates $\mathbf{x} = (x_1, x_2, \dots)$. A series $f(\mathbf{x}) \in \mathbb{K}[[\mathbf{x}]]$ is a *quasisymmetric function* if it has bounded degree and for all indices i,

$$f(\mathbf{x}) = f(x_1, \dots, x_{i-1}, 0, x_i, x_{i+1}, \dots).$$

Quasisymmetric functions form an algebra QSym. Clearly, by repeating this rule we may insert any finite number of zeroes in whichever positions we like. This ensures that the limit on the right side of (2.16) exists, and we define a coproduct Δ as the map $\mathsf{QSym} \to \mathsf{QSym} \otimes \mathsf{QSym}$ corresponding to $f(\mathbf{x}) \mapsto f(\mathbf{x}, \mathbf{y})$ as we did with symmetric functions.

Quasisymmetric functions are related to compositions similarly to how symmetric functions are related to partitions. For an exponent vector $\alpha \in \mathbb{N}^{\mathbb{N}_+}$, define $\operatorname{pack}(\alpha)$ to be the composition consisting of the nonzero entries of α in their original order. An alternative definition of quasisymmetry is that f is quasisymmetric if $[\mathbf{x}^\alpha]f = [\mathbf{x}^\beta]f$ whenever $\operatorname{pack}(\alpha) = \operatorname{pack}(\beta)$. With this in mind, for a composition α we define the *monomial quasisymmetric function*

$$M_\alpha = \sum_{\operatorname{pack}(\alpha')=\alpha} \mathbf{x}^{\alpha'}.$$

These form a basis for QSym. Note that the monomial *symmetric* functions can be expanded in terms of these as

$$m_\lambda = \sum_{\substack{\alpha \in \mathrm{Comp} \\ \operatorname{sort}(\alpha)=\lambda}} M_\alpha.$$

In the monomial basis the coproduct is given by

$$\Delta M_\alpha = \sum_{\beta \cdot \gamma = \alpha} M_\beta \otimes M_\gamma$$

where $\cdot$ denotes concatenation of compositions. Note that this formula makes it clear that Δ is in fact coassociative, which is not entirely obvious from the way we defined it. This formula also makes it clear that QSym is *not* cocommutative as Sym is. Indeed, this is in a sense the main difference between them, as we will see shortly.

From the perspective of combinatorial Hopf algebras, the most important result about quasisymmetric functions is the Aguiar–Bergeron–Sottile theorem, which states that QSym has a certain universal property. We need some notation. For a character ζ of some *graded* Hopf algebra H, write ζ_n for the linear map that equals ζ on H_n and is zero on all other graded pieces. For a composition $\alpha = (\alpha_1, \dots, \alpha_k)$ write

$$\zeta_\alpha = \zeta_{\alpha_1} * \cdots * \zeta_{\alpha_k}.$$

Let η be the character of QSym given by mapping x_1 to 1 and x_i to 0 for all other i. Explicitly, on the monomial basis this is

$$\eta(M_\alpha) = \begin{cases} 1 & \ell(\alpha) \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 2.4.6 (Aguiar–Bergeron–Sottile [3, Theorems 4.1 and 4.3])**.** *Let H be a connected graded Hopf algebra and $\zeta \in \mathrm{Ch}(H)$. There exists a unique graded bialgebra morphism $\varphi\colon H \to \mathsf{QSym}$ such that $\eta\varphi = \zeta$. This map has the explicit formula*

$$\varphi(h) = \sum_{\alpha \in \mathrm{Comp}} \zeta_\alpha(h) M_\alpha.$$

Moreover, if H is cocommutative then φ takes values in Sym.

Of course, one can conversely define a character $\zeta = \eta\varphi$ given a graded bialgebra morphism φ, so this gives a *bijection* between $\mathrm{Ch}(H)$ and the set of graded bialgebra morphisms $H \to \mathsf{QSym}$. Recall from Section 2.2.4 that $\mathrm{Ch}(H)$ is also in bijection with the set of (not necessarily graded) bialgebra morphisms $H \to \mathbb{K}[z]$ under the more general hypothesis that H is unipotent. Thus we should see quasisymmetric functions as playing a role analogous to polynomials in a world where we insist on strictly preserving degree.

Example 2.4.7. In the same spirit as Example 2.2.25, we may use the Aguiar–Bergeron–Sottile theorem to obtain interesting symmetric and quasisymmetric invariants of combinatorial objects from seemingly uninteresting characters of associated Hopf algebras. For instance, letting ζ be the same character of $\mathcal{G}$ as in that example, the induced map $\mathcal{G} \to \mathsf{Sym}$ sends a graph to its *chromatic symmetric function*, an invariant introduced by Stanley [58] that has become a significant topic of study in algebraic combinatorics.

2.4.3 Faà di Bruno

Let $\widetilde{\mathfrak{D}} \subset \mathbb{K}[[x]]$ be the set of all series with zero constant term and nonzero linear term. These are also known as *formal diffeomorphisms* and form a group with respect to composition of series. Let $\mathfrak{D}$ be the subgroup consisting of series with linear term x; these are sometimes known as *δ-series*. It turns out that $\mathfrak{D}$ is essentially isomorphic to the character group of a graded Hopf algebra, the *Faà di Bruno Hopf algebra* FdB.

As an algebra, FdB should be thought of as the algebra of polynomial functions on $\mathfrak{D}$. Explicitly, it is the polynomial algebra $\mathbb{K}[\pi_1, \pi_2, \dots]$ in an $\mathbb{N}_+$-indexed set of variables. We organize these variables into a power series

$$\Pi(x) = x + \sum_{n \geq 1} \pi_n x^{n+1}.$$

Then the map $\mathrm{Ch}(\mathsf{FdB}) \to \mathfrak{D}$ given by $\zeta \mapsto \zeta(\Pi(x))$ is clearly a bijection. (Note that here and throughout the thesis, notation like $\zeta(\Pi(x))$ implicitly means applying ζ coefficientwise.) We define a coproduct

$$\Delta\pi_n = \sum_{k=0}^{n} [x^{n+1}]\Pi(x)^{k+1} \otimes \pi_k. \tag{2.19}$$

(where $\pi_0 = 1$). Observe that this makes FdB into a connected graded bialgebra if we define π_n to have degree n; this is the reason for the off-by-one in the definition. The following proposition is essentially immediate from (2.19).

Proposition 2.4.8. *Let A be a commutative algebra and $\varphi, \psi \colon \mathsf{FdB} \to A$ be algebra morphisms. Let $\Phi(x) = \varphi(\Pi(x))$ and $\Psi(x) = \psi(\Pi(x))$. Then $(\varphi * \psi)(\Pi(x)) = \Psi(\Phi(x))$.*

In particular, this actually implies that the map $\mathrm{Ch}(\mathsf{FdB}) \to \mathfrak{D}$ described above is an *anti*-isomorphism of groups. Clearly we could have defined the coproduct with the tensor factors flipped in order to make it an isomorphism, but the way we have defined it is both traditional and will turn out to be convenient for our purposes.

Remark 2.4.9. The Faà di Bruno Hopf algebra is isomorphic to the incidence Hopf algebra (see Remark 2.3.2) of finite partition lattices. Indeed, this is essentially the way it was originally defined by Doubilet [17] though this precedes the general theory of incidence Hopf algebras. The isomorphism between the two comes by expressing the power series coefficients that appear in (2.19) as sums over set partitions. This is a combinatorial formulation of *Faà di Bruno's formula*, whence $\mathcal{F}$ is named.

We can also use Proposition 2.4.8 to get a formula for the antipode of FdB. Indeed, Proposition 2.4.8 implies that $S(\Pi(x))$ is the compositional inverse of $\Pi(x)$, so by Lagrange inversion,

$$S(\pi_n) = [x^{n+1}]S(\Pi(x)) = \frac{1}{n+1}[x^n]\left(\frac{x}{\Pi(x)}\right)^{n+1}. \tag{2.20}$$

Remark 2.4.10. Haiman and Schmitt [29] were able to derive (2.20) directly from Takeuchi's formula (Theorem 2.2.9) using a combinatorial argument, thus giving a Hopf-algebraic proof of Lagrange inversion.

We can also work out what the actions of FdB^* described in Section 2.2.3 look like. By Proposition 2.2.19 it is sufficient to understand how an element of FdB^* acts on the generators, or equivalently on the series $\Pi(x)$ itself. The following result follows more or less immediately from (2.19).

Proposition 2.4.11. *Suppose $\varphi \in \mathsf{FdB}^*$ and let $\Phi(x) = \varphi(\Pi(x))$. Then:*

(i) $\varphi \rightharpoonup \Pi(x) = \Phi(\Pi(x))$.

(ii) If $\varphi \in \mathrm{Ch}(\mathsf{FdB})$ *then* $\Pi(x) \leftharpoonup \varphi = \Pi(\Phi(x))$.

(iii) If $\varphi \in \mathfrak{ch}(\mathsf{FdB})$ *then* $\Pi(x) \leftharpoonup \varphi = \Phi(x)\Pi'(x)$.

As a particular consequence of Proposition 2.4.11(iii), we get a nice description of the Lie algebra $\mathfrak{ch}(\mathsf{FdB})$: the map $\varphi \mapsto \varphi(\Pi(x))\frac{d}{dx}$ gives a faithful representation by differential operators on $\mathbb{K}[[x]]$. We can also combine this with Theorem 2.2.24 to characterize bialgebra morphisms $\mathsf{FdB} \to \mathbb{K}[z]$.

Theorem 2.4.12. *Let $\varphi \colon \mathsf{FdB} \to \mathbb{K}[z]$ be an algebra morphism and let $\Phi(x, z) = \varphi(\Pi(x))$. Let $\beta(x)$ be the linear term in z of $\Phi(x, z)$. Then φ is a bialgebra morphism if and only if $\Phi(x, 0) = x$ and*

$$\frac{\partial \Phi(x, z)}{\partial z} = \beta(x)\frac{\partial \Phi(x, z)}{\partial x}. \tag{2.21}$$

Proof. By Theorem 2.2.24, φ is a bialgebra morphism if and only if $\varphi|_{z=0} = \varepsilon$ and $\frac{d}{dz}\varphi = (\operatorname{lin}\varphi) * \varphi$. Since φ is an algebra morphism, its behaviour is determined by what it does to the generators, so these are respectively equivalent to $\Phi(x, 0) = \varepsilon(\Pi(x)) = x$ and

$$\frac{\partial \Phi(x, z)}{\partial z} = ((\operatorname{lin}\varphi) * \varphi)(\Pi(x)) = \varphi(\Pi(x) \leftharpoonup \operatorname{lin}(\varphi)).$$

Since $\operatorname{lin}\varphi$ is an infinitesimal character, by Proposition 2.4.11(iii) the right-hand side is $\beta(x)\frac{\partial \Phi(x,z)}{\partial x}$ as wanted. $\square$

In Chapter 4 we will return to these ideas and explore their applications in quantum field theory.

Chapter 3

Binary Tubings and 1-Cocycles

In this chapter we return to the Connes–Kreimer Hopf algebra and its universal property (Theorem 2.3.3). The main problem we will be interested in will be to find explicit formulas for the bialgebra morphisms which come from this property as well as some generalizations thereof.

3.1 Cohomology of comodules

In Section 2.3.1 we briefly discussed a class of operators we mysteriously referred to as "1-cocycles". Let us now put this into its correct context. Let H be a bialgebra and M a left comodule over H, with coaction δ. For $k \geq 0$, a *k-cochain on M* is a linear map $M \to H^{\otimes k}$. Denote the vector space of k-cochains by $\mathrm{C}^k(H, M)$. The *coboundary map* $d_k \colon \mathrm{C}^k(H, M) \to \mathrm{C}^{k+1}(H, M)$ is defined by

$$d_k \Lambda = (\mathrm{id}_H \otimes \Lambda)\delta + \sum_{j=1}^{k} (-1)^j (\mathrm{id}_H^{\otimes(j-1)} \otimes \Delta \otimes \mathrm{id}_H^{\otimes(k-j)})\Lambda + (-1)^k \Lambda \otimes 1. \tag{3.1}$$

The kernel and image of this map are the spaces of k-cocycles and $(k+1)$-coboundaries respectively. The space of k-cocycles is denoted $\mathrm{Z}^k(H, M)$. A tedious but routine calculation shows that $d_{k+1} d_k = 0$, so every coboundary is a cocycle. The quotient $\mathrm{H}^k(H, M) = \mathrm{Z}^k(H, M)/d_{k-1}\mathrm{C}^{k-1}(H, M)$ is the kth cohomology of the comodule M. Most often considered is the case $M = H$, in which case we simply write $\mathrm{Z}^k(H)$ and $\mathrm{H}^k(H)$.

Remark 3.1.1. For the reader familiar with homological algebra, we mention that $\mathrm{H}^k(H, M)$ is nothing more than $\mathrm{Ext}_H^k(M, H)$ in the category of comodules over H, with the above definition amounting to a computation of this Ext-group with an explicit injective resolution of H as a comodule over itself. In particular, this implies that the cohomology groups depend only on the coalgebra structure of H and not on its multiplication, even though the definition makes use of the multiplicative unit. However, this fact will be of no real use to us as we are more interested in the cocycles than the cohomology.

Remark 3.1.2. More generally, we could consider a *bicomodule* M with left coaction δ

and right coaction δ', and define the coboundary to be

$$d_k\Lambda = (\mathrm{id}_H \otimes \Lambda)\delta + \sum_{j=1}^{k}(-1)^j(\mathrm{id}_H^{\otimes(j-1)} \otimes \Delta \otimes \mathrm{id}_H^{\otimes(k-j)})\Lambda + (-1)^k(\Lambda \otimes \mathrm{id}_H)\delta'.$$

This is the original notion of cohomology for coalgebras introduced by Doi [16]. The version we are considering is simply the special case where the right coaction is trivial.

We will only be interested in the case $k = 1$. Moreover, for the remainder of this section we will focus on the case $M = H$. (We will consider some other comodules in Section 3.4.) In this case the cocycle condition $d_1\Lambda = 0$ can be written

$$\Delta\Lambda = \Lambda \otimes 1 + (\mathrm{id}_H \otimes \Lambda)\Delta \tag{3.2}$$

which is the form we saw in Section 2.3.1 with a combinatorial interpretation. There are also some very natural examples of a more algebraic flavour.

Example 3.1.3. Let $\mathcal{I}$ be the integration operator on $\mathbb{K}[z]$:

$$\mathcal{I}f(z) = \int_0^z f(u)\,du.$$

Recall (from Example 2.2.4) that the coproduct on $\mathbb{K}[z]$ can be interpreted as substituting a sum $z_1 + z_2$ in place of the variable z. That $\mathcal{I}$ is a 1-cocycle then boils down to some familiar properties of integrals:

$$\begin{aligned}\int_0^{z_1+z_2} f(u)\,du &= \int_0^{z_1} f(u)\,du + \int_{z_1}^{z_1+z_2} f(u)\,du \\ &= \int_0^{z_1} f(u)\,du + \int_0^{z_2} f(z_1+u)\,du.\end{aligned}$$

By Theorem 2.3.3, $\mathcal{I}$ defines a morphism $\varphi\colon \mathcal{H} \to \mathbb{K}[z]$. An easy induction with the recurrence (2.12) gives

$$\varphi(t) = \frac{z^{|t|}}{\prod_{v\in t}|t_v|}$$

where t_v denotes the subtree (principal downset) rooted at v. The denominator is also known as the *tree factorial*. A formula of Knuth [36, Section 5.1, Exercise 20] gives the number of linear extensions of a tree as

$$e(t) = \frac{|t|!}{\prod_{v\in t}|t_v|}$$

and hence we can alternatively write

$$\varphi(t) = \frac{e(t)z^{|t|}}{|t|!}.$$

This latter formula is the simplest special case of the formula we will derive for arbitrary 1-cocycles on $\mathbb{K}[z]$ as Theorem 3.3.1.

Example 3.1.4. Another natural example of a 1-cocycle on $\mathbb{K}[z]$ is the "discrete integral" operator $\mathcal{S}$ (not to be confused with the antipode S) which is uniquely defined by the property that for $n \in \mathbb{N}$,

$$\mathcal{S}f(z)\Big|_{z=n} = \sum_{k=1}^{n} f(k). \tag{3.3}$$

We can explicitly define such an $\mathcal{S}$ on the basis of binomial coefficients by

$$\mathcal{S}\binom{z}{r} = \binom{z+1}{r+1}$$

which satisfies (3.3) by the hockey-stick identity (2.2). Uniqueness follows from the fact that two polynomials which agree at infinitely many elements of $\mathbb{K}$ are equal in $\mathbb{K}[z]$. Using this fact again, we can show that $\mathcal{S}$ is a cocycle by exactly the same calculation as in Example 3.1.3 but with a sum instead of an integral:

$$\begin{aligned}\sum_{k=1}^{n_1+n_2} f(k) &= \sum_{k=1}^{n_1} f(k) + \sum_{k=n_1+1}^{n_1+n_2} f(k) \\ &= \sum_{k=1}^{n_1} f(k) + \sum_{k=1}^{n_2} f(n_1+k).\end{aligned}$$

The map $\mathcal{H} \to \mathbb{K}[z]$ induced by this cocycle is the restriction to $\mathcal{H}$ of the map from Example 2.3.1 sending a poset to its order polynomial. The recurrence

$$\Omega(B_+f;n) = \sum_{k=1}^{n} \Omega(f;k)$$

comes from summing over the different values that such a map may send the root to. It is known (see e.g. [57, Section 3.12]) that for any poset P, the leading term of $\Omega(P;z)$ is $e(P)z^{|P|}/|P|!$, so the leading term of this map is the same as the map from Example 3.1.3. As we will soon see, this is no coincidence.

We begin with some basic properties of 1-cocycles. For stating these it is convenient to generalize the notion of convolution to maps defined on comodules: for an algebra A and maps $\alpha\colon H \to A$ and $\beta\colon M \to A$, write

$$\alpha *_\delta \beta = m_A(\alpha \otimes \beta)\delta$$

where m_A is the multiplication map on A.

Lemma 3.1.5. *Let M be a comodule over H and $\Lambda \in \mathrm{Z}^1(H,M)$. Then:*

(i) If $\alpha, \beta\colon H \to A$ for some algebra A then $(\alpha * \beta)\Lambda = \beta(1)\alpha\Lambda + \alpha *_\delta \beta\Lambda$.

(ii) $\varepsilon\Lambda = 0$.

(iii) If $\varphi\colon N \to M$ is a homomorphism of comodules then $\Lambda\varphi \in \mathrm{Z}^1(H,N)$.

Proof. (i) Immediate from the definition of 1-cocycles and convolution.

(ii) Follows from (i) since ε is the identity for convolution.

(iii) Write δ and δ' for the coactions on M and N respectively. We have

$$\Delta\Lambda\varphi = (\Lambda \otimes 1)\varphi + (\mathrm{id} \otimes \Lambda)\delta\varphi = \Lambda\varphi \otimes 1 + (\mathrm{id} \otimes \Lambda\varphi)\delta'. \qquad \square$$

Note in particular that if $\beta(1) = 0$ (e.g. if β is an infinitesimal character) then (i) just says $(\alpha * \beta)\Lambda = \alpha * \beta\Lambda$.

Suppose $\Lambda \in \mathrm{Z}^1(H)$. We can use Λ to build new cocycles on various comodules. Given a left comodule M with coaction δ and a linear map $\psi\colon M \to \mathbb{K}$, we define $\Lambda \circledast \psi = (\Lambda \otimes \psi)\delta$.

Lemma 3.1.6. *Subject to the above assumptions, $\Lambda \circledast \psi \in \mathrm{Z}^1(H, M)$.*

Proof. We compute

$$\begin{aligned}
\Delta(\Lambda \circledast \psi) &= (\Delta\Lambda \otimes \psi)\delta \\
&= (\Lambda \otimes \psi)\delta \otimes 1 + ((\mathrm{id} \otimes \Lambda)\Delta \otimes \psi)\delta \\
&= (\Lambda \circledast \psi) \otimes 1 + (\mathrm{id} \otimes \Lambda \otimes \psi)(\Delta \otimes \mathrm{id})\delta \\
&= (\Lambda \circledast \psi) \otimes 1 + (\mathrm{id} \otimes \Lambda \otimes \psi)(\mathrm{id} \otimes \delta)\delta \\
&= (\Lambda \circledast \psi) \otimes 1 + (\mathrm{id} \otimes (\Lambda \circledast \psi))\delta.
\end{aligned} \qquad \square$$

As a special case of this, note that $d\varepsilon \circledast \psi = d\psi$. When $M = H$ we can write $\Lambda \circledast \psi$ using the left action of H^* on H described in Section 2.2.3:

$$(\Lambda \circledast \psi)h = \Lambda(\psi \rightharpoonup h).$$

Using this operation and the integral cocycle from Example 3.1.3, we can describe all 1-cocycles on $\mathbb{K}[z]$.

Theorem 3.1.7 (Panzer [47, Theorem 2.6.4])**.** *For any series $A(z) \in \mathbb{K}[[z]]$, the operator*

$$f(z) \mapsto \int_0^z A(d/du) f(u)\, du \tag{3.4}$$

is a 1-cocycle on $\mathbb{K}[z]$. Moreover, all 1-cocycles on $\mathbb{K}[z]$ are of this form.

Proof sketch. That this is a cocycle follows from Lemma 3.1.6, identifying $\mathbb{K}[z]^*$ with $\mathbb{K}[[z]]$ as discussed in Example 2.2.16. The other direction is a calculation; see the proof in [47]. (We will also prove a more general result later as Theorem 3.4.1.) $\square$

Corollary 3.1.8. *The cohomology $\mathrm{H}^1(\mathbb{K}[z])$ is 1-dimensional and generated by the class of the integral cocycle $\mathcal{I}$.*

Proof. Note $\mathcal{I}(1) = z$, so $\mathcal{I}$ is not a coboundary. Now suppose Λ is a 1-cocycle given by (3.4). Write $A(z) = a_0 + zB(z)$ for some series $B(z)$. Then we have

$$\Lambda f(z) = \int_0^z A(d/du) f(u)\, du$$

$$
\begin{aligned}
&= a_0 \int_0^z f(u)\,du + \int_0^z \frac{d}{du} B(d/du) f(u)\,du \\
&= a_0 \mathcal{I} f(z) + B(d/dz) f(z) - B(d/dz) f(z)\big|_{z=0}
\end{aligned}
$$

hence $\Lambda = a_0 \mathcal{I} + d\beta$ where β is the linear form $z^n \mapsto [z^n]B(z)$. ☐

Example 3.1.9. To put the discrete integral cocycle (Example 3.1.4) in the form (3.4), we take $A(z) = ze^z/(e^z - 1)$, the exponential generating function for the Bernoulli numbers[1]. The identity

$$
f(1) + \cdots + f(n) = \int_0^n A(d/du) f(u)\,du
$$

is the *Euler–Maclaurin formula*; see for instance [50, Section 7].[2]

Remark 3.1.10. We can also write 1-cocycles on $\mathbb{K}[z]$ in a different form, namely

$$
f(z) \mapsto f(\partial/\partial\rho) \frac{e^{z\rho} - 1}{\rho} A(\rho)\big|_{\rho=0}.
$$

Checking on the basis of monomials quickly reveals that this is equivalent to the 1-cocycle that appears in the statement of Theorem 3.1.7. Operators of this form appear in the theory of Dyson–Schwinger equations, as we will discuss further in Section 4.1.

We will take Theorem 3.1.7 as the starting point for the first part of our quest. Given a 1-cocycle Λ expressed in the form (3.4) we will be interested in the problem of determining an explicit formula for the map $\mathcal{H} \to \mathbb{K}[z]$ induced by Λ, in terms of the coefficients of the series $A(z)$. As a first step, we can easily compute the leading term, generalizing our observation from Example 3.1.4.

Proposition 3.1.11. *Let Λ be given by (3.4) and let $\varphi\colon \mathcal{H} \to \mathbb{K}[z]$ satisfy $\varphi B_+ = \Lambda\varphi$. Let a_0 be the constant term of $A(z)$. Then for any forest t,*

$$
\varphi(t) = \frac{e(t)(a_0 z)^{|t|}}{|t|!} + (\textit{lower-order terms}).
$$

Proof. Write $t = B_+ f$. Then from the form of (3.4) it is clear that the leading term of $\varphi(t)$ is simply a_0 times the integral of the leading term of $\varphi(f)$, which in turn is the product of the leading terms of the components of f. Thus inductively assuming the result holds for smaller trees, the leading term is $e(f)(a_0 z)^{|t|}/|t|!$ where clearly $e(t) = e(f)$. ☐

In the next section, we will introduce a generalization of linear extensions that will allow us to give an analogous interpretation to the lower-order terms. In fact, we will be able to do this in a more general setting where we have an arbitrary *family* of 1-cocycles indexed

[1] There are two conventions for Bernoulli numbers. The "other" Bernoulli numbers, with exponential generating function $z/(e^z - 1)$, would appear if we defined our discrete integral as a sum from 0 to $n - 1$ rather than 1 to n.

[2] Most often the Euler–Maclaurin formula is stated as an asymptotic expansion for the sum of a smooth function f of a real variable, but it holds exactly – and makes sense over any field – when f is a polynomial.

by some set I. The universal object here is the *decorated Connes–Kreimer Hopf algebra* $\mathcal{H}_I$, which we now define. By an *I-tree* (resp. *I-forest*) we mean a tree (resp. forest) with each vertex decorated by an element of I. We will write $\mathcal{T}(I)$ for the set of I-trees and $\mathcal{F}(I)$ for the set of I-forests. Then $\mathcal{H}_I$ is the free vector space on $\mathcal{F}(I)$, made into a bialgebra with disjoint union as multiplication and the same coproduct as in $\mathcal{H}$ but preserving the decorations on all vertices. As usual, we can grade by the number of vertices and find that $\mathcal{H}_I$ is a connected graded bialgebra and hence a Hopf algebra.

Remark 3.1.12. We could instead choose some weight function $w\colon I \to \mathbb{N}$ and grade $\mathcal{H}_I$ by total weight. If w takes only *positive* values then this grading will also make $\mathcal{H}_I$ connected. In the application to Dyson–Schwinger equations we will have such a weight function already and so it is natural to grade the algebra this way, but it won't really matter for anything we do.

For each $i \in I$, we have an operator $B_+^{(i)}$ on $\mathcal{H}_I$ that sends an I-forest to the I-tree obtained by adding a root with decoration i. For the same combinatorial reasons as the usual B_+ operator on $\mathcal{H}$, all of these are 1-cocycles. Theorem 2.3.3 generalizes easily to this setting.

Theorem 3.1.13 ([25, Proposition 2 and 3])**.** *Let A be a commutative algebra and $\{\Lambda_i\}_{i\in I}$ be a family of linear operators on A. There exists a unique algebra morphism $\varphi\colon \mathcal{H}_I \to A$ such that $\varphi B_+^{(i)} = \Lambda_i \varphi$. Moreover, if A is a bialgebra and Λ_i is a 1-cocycle for each i then φ is a bialgebra morphism.*

3.2 Binary tubings

We now introduce the combinatorial objects that we will use to understand 1-cocycles. We will only be interested in trees, but the basic definitions can be given in the context of an arbitrary finite poset P. A *tube* is a connected convex subset of P.[3] For tubes X, Y write $X \to Y$ if $X \cap Y = \emptyset$ and there exist $x \in X$ and $y \in Y$ such that $x < y$. A collection τ of tubes is called a *tubing* if it satisfies the following conditions:

- (Laminarity) If $X, Y \in \tau$ then either $X \cap Y = \emptyset$, $X \subseteq Y$, or $X \supseteq Y$.
- (Acyclicity) There do not exist tubes $X_1, \ldots, X_k \in \tau$ with $X_1 \to X_2 \to \cdots \to X_k \to X_1$.

Tubings of posets (also called *pipings*) were introduced by Galashin [26] to index the vertices of a certain polytope associated to P, the *P-associahedron*. They were rediscovered (in the case of trees) by the authors of [8] in the present context. Note that for trees the acyclicity condition is trivial.

Remark 3.2.1. Galashin defines a *proper tube* to be one which is neither a singleton nor the entirety of P, and a *proper tubing* to be one consisting only of proper tubes. Only the proper tubes and tubings play a role in the definition of the poset associahedron, but for

[3] Note that being both connected and convex is equivalent to inducing a connected subgraph of the Hasse diagram of P.

us it will be sensible to include the improper ones. Note that if one restricts attention to *maximal* tubings (which we largely will do) then this makes no combinatorial difference, as a maximal tubing contains *all* of the improper tubes and removing them maps the set of maximal tubings bijectively to the set of maximal proper tubings.

Remark 3.2.2. Tubings of posets are only loosely related to the better-known notion of tubings of graphs introduced by Carr and Devadoss [13]. For graphs, a *tube* is defined to be a set of vertices which induces a connected subgraph and a *tubing* is a set of tubes satisfying the laminarity condition along with a certain *non-adjacency* condition which is entirely different from the acyclicity condition for poset tubings. Thus the notions should not be confused. However, in the case of trees there is a close connection: tubings of a rooted tree (as a poset) are in bijection with tubings of the *line graph* of the tree. (This is essentially a special case of a result of Ward [60, Lemma 3.17] which is stated in terms of related objects called *nestings*. See [8, Section 6] for a discussion of this in our language.)

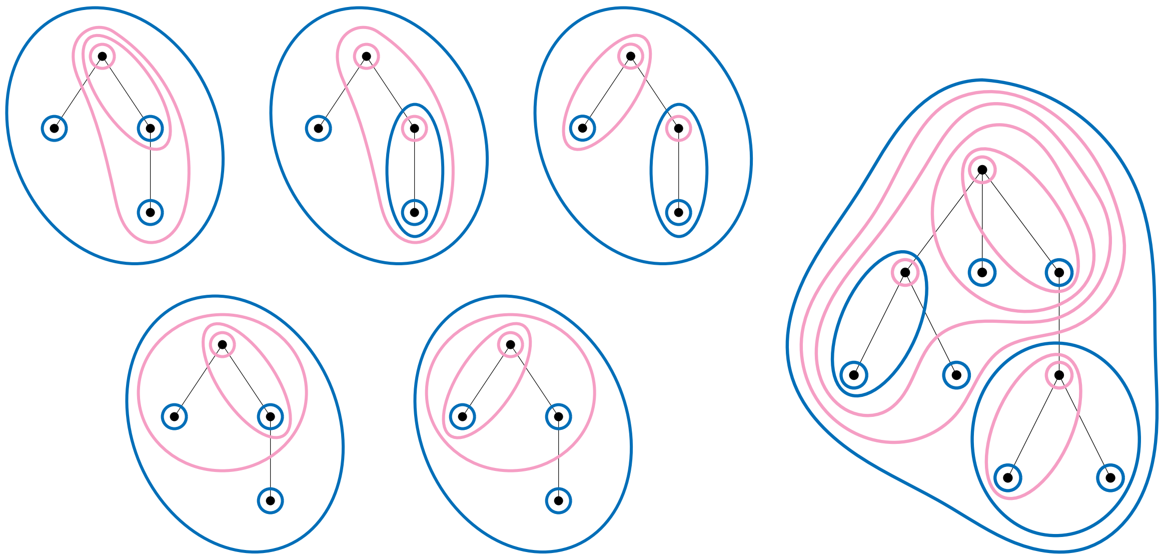
Remark 3.2.3. A subset of a rooted tree is convex and connected (in the order-theoretic sense) if and only if it is connected in the graph-theoretic sense. Thus the set of tubings of a rooted tree is really an invariant of the underlying unrooted tree. However, the statistics on tubings in which we will be interested do depend on the root and are best thought of in terms of the partial ordering.

The laminarity condition implies that if τ is a tubing of P, the poset of tubes ordered by inclusion is a forest. In the case of a maximal tubing of a connected poset, there is a unique maximal tube (namely P itself) and each non-singleton tube has exactly two maximal tubes properly contained within it. Relative to X, one of these tubes is a downset and one is an upset. Taking the downset as the left child and the upset as the right child, the tubes of a maximal tubing thus have the structure of a binary plane tree. For this reason (and to avoid confusion with graph tubings), maximal tubings of rooted trees were referred to as *binary tubings* in [8] and we will also use this language.

Henceforth we shall restrict our attention exclusively to binary tubings of rooted trees. We will write $\mathrm{Tub}(t)$ for the set of binary tubings of t. We will call a tube a *lower tube* (resp. *upper tube*) if it is a downset (resp. upset) in its parent in the tree of tubes. We will also consider t itself to be a lower tube; this ensures that each vertex is the root of exactly one lower tube. Given a vertex v, define the *rank* $\mathrm{rk}(\tau, v)$ of v in τ to be the number of upper tubes rooted at v.[4] We will write $b(\tau)$ for the number of tubes of τ containing the root of t; note that we clearly have $b(\tau) = \mathrm{rk}(\tau, \mathrm{rt}\, t) + 1$.

Remark 3.2.4. Our definition of the rank refers only to the upper tubes, but it can be equivalently defined in terms of *lower* tubes only: for each upper tube rooted at v there is a corresponding lower tube, with the property that there is no lower tube of τ lying strictly between it and the unique lower tube rooted at v. Conversely any such lower tube corresponds to an upper tube rooted at v. In other words, considering the lower tubes of τ as a poset ordered by containment, $\mathrm{rk}(\tau, v)$ is the number of lower tubes that are covered by the unique lower tube rooted at v.

[4]In [8] the equivalent statistic $b(\tau, v) = \mathrm{rk}(\tau, v) + 1$ counting the total number of tubes rooted at v was used instead. It has since become clear that the rank is really the fundamental quantity.



(a) All binary tubings of a small tree.

(b) A binary tubing of a larger tree.

Figure 3.1. Examples of binary tubings. Upper and lower tubes highlighted in different colours.

Binary tubings of rooted trees have a recursive structure which we will make essential use of.

Proposition 3.2.5. *Let t be a rooted tree with $|t| > 1$. There is a bijection between binary tubings of t and triples (t', τ', τ'') where t' is a proper subtree (principal downset) of t, $\tau' \in \mathrm{Tub}(t')$, and $\tau'' \in \mathrm{Tub}(t \setminus t')$. Moreover this bijection satisfies*

$$\mathrm{rk}(\tau, v) = \begin{cases} \mathrm{rk}(\tau', v) & v \in t' \\ \mathrm{rk}(\tau'', v) + 1 & v = \mathrm{rt}\, t \\ \mathrm{rk}(\tau'', v) & \textit{otherwise} \end{cases} \tag{3.5}$$

and $b(\tau) = b(\tau'') + 1$.

Proof. By the discussion above there are two maximal tubes t', t'' properly contained in the largest tube t, where t' is a downset and $t'' = t \setminus t'$ is an upset and both are connected. A connected downset in a rooted tree is a subtree; since the complement is nonempty it must be that t' is a proper subtree. Let

$$\tau' = \{u \in \tau : u \subseteq t'\}$$

and

$$\tau'' = \{u \in \tau : u \cap t' = \emptyset\}.$$

By the laminarity condition, we have $\tau = \{t\} \cup \tau' \cup \tau''$. Since each tube in τ' and τ'' still contains two maximal tubes within it, these are still maximal tubings of t' and t'' respectively, i.e. $\tau' \in \mathrm{Tub}(t')$ and $\tau'' \in \mathrm{Tub}(t \setminus t')$. Note that the upper tubes of τ' and τ'' are also upper tubes of τ, whereas t'' is a lower tube in τ'' and an upper tube in τ. Thus the statement about ranks follows. Since $b(\tau) = \mathrm{rk}(\tau, \mathrm{rt}\, t) + 1$ and $\mathrm{rt}\, t \in t''$ the statement about the b-statistic also follows. □

Remark 3.2.6. Observe that in a binary tubing we split the tree into an upper and lower part just as in the definition of the coproduct, but with the key difference that they are both required to be trees. To make this more algebraic, let $P_{\mathrm{lin}} \colon \mathcal{H} \to \mathcal{H}$ be the projection onto the subspace spanned by trees. Then the *linearized coproduct* is $\Delta_{\mathrm{lin}} = (P_{\mathrm{lin}} \otimes P_{\mathrm{lin}})\Delta$. In effect, Δ_{lin} looks the same as the usual coproduct but only includes terms where both tensor factors are trees rather than arbitrary forests. Unlike the coproduct, the linearized coproduct fails to be coassociative: there are multiple different maps $\mathcal{H} \to \mathcal{H}^{\otimes k}$ that can be built by iterating it. For instance in the case $k = 3$ there are distinct maps $(\Delta_{\mathrm{lin}} \otimes \mathrm{id})\Delta_{\mathrm{lin}}$ and $(\mathrm{id} \otimes \Delta_{\mathrm{lin}})\Delta_{\mathrm{lin}}$. It is not hard to see that if we iterate all the way to $k = |t|$, the terms that we can get from all of these maps taken collectively correspond exactly to the binary tubings.

3.2.1 Enumeration of binary tubings

While not strictly relevant to our application, it is also interesting to consider the purely combinatorial question of counting binary tubings of a tree. Write $N(t)$ for the number of binary tubings of t. As an immediate consequence of Proposition 3.2.5 we get a recursive formula for $N(t)$.

Corollary 3.2.7 ([8, Lemma 2.8]). *Let t be a rooted tree with $|t| > 1$. Then*

$$N(t) = \sum_{t'} N(t')N(t \setminus t') \tag{3.6}$$

summing over proper subtrees t'.

Using Corollary 3.2.7 we are able to easily enumerate binary tubings for corollas and ladders. (See Section 2.3.1 for definitions and notation.)

Proposition 3.2.8. *We have $N(s_n) = (n-1)!$ and $N(\ell_n) = \mathrm{cat}(n-1)$, the $(n-1)$st Catalan number.*

Proof. The recurrence gives

$$N(s_n) = (n-1)N(s_{n-1})$$

and

$$N(\ell_n) = \sum_{k=1}^{n-1} N(\ell_k)N(\ell_{n-k})$$

with the base case $N(s_1) = N(\ell_1) = 1$. The result follows. □

Corollas and ladders turn out to be the extreme cases. We can easily observe that the number of tubings of a tree is bounded above by $(n-1)!$ simply because for a subtree t' of size $k < n$, we have

$$N(t')N(t \setminus t') \leq (k-1)!(n-k-1)! \leq (n-2)!$$

and hence each term on the RHS of (3.6) is at most $(n-2)!$ so the sum is at most $(n-1)!$ as desired. More generally we have the following result.[5]

Theorem 3.2.9 ([8, Lemma 3.7])**.** *For any tree t with n vertices,* $\mathrm{cat}(n-1) \leq N(t) \leq (n-1)!$.

The following common generalization of the two formulas of Proposition 3.2.8 was suggested to the author by Alejandro Morales. It has been independently proved by Nguyen and Sack [46].

Theorem 3.2.10. *Let $t_{r,s} = B_+(\ell_1^r \ell_s)$. Then*

$$N(t_{r,s}) = \frac{(r+1)!}{r+s+1}\binom{r+2s}{s}.$$

Proof. We have $N(t_{r,0}) = r!$ and $N(t_{0,s}) = \mathrm{cat}(s)$, which match the formula.

From Corollary 3.2.7, for $r > 0$ and $s > 0$, we have

$$N(t_{r,s}) = rN(t_{r-1,s}) + \sum_{k=1}^{s} \mathrm{cat}(k-1)N(t_{r,s-k}). \tag{3.7}$$

Consider the generating function

$$A(x,y) = \sum_{r\geq 0}\sum_{s\geq 0} \frac{N(t_{r,s})x^r y^s}{r!}.$$

In terms of this series, (3.7) says that for $r > 0$ and $s > 0$, we have

$$[x^r y^s]A(x,y) = [x^r y^s](x + yC(y))A(x,y) \tag{3.8}$$

where $C(y)$ is the generating function for the Catalan numbers, which satisfies

$$C(y) = \frac{1}{1 - yC(y)}.$$

Explicitly solving (3.8) using the known base cases we get

$$A(x,y) = \frac{1}{1 - x - yC(y)}$$

to which we apply the Lagrange implicit function theorem (Theorem 2.1.3) with $F(u) = \frac{1}{1-x-u}$ and $G(u) = \frac{1}{1-u}$:

$$N(t_{r,s}) = r![x^r y^s]A(x,y)$$

[5] The proof for the lower bound is far less pleasant.

$$
\begin{aligned}
&= \frac{r!}{s}[x^r u^{s-1}]\frac{1}{(1-x-u)^2(1-u)^s} \\
&= \frac{(r+1)!}{s}[u^{s-1}]\frac{1}{(1-u)^{r+s+2}} \\
&= \frac{(r+1)!}{s}\binom{r+2s}{s-1} \\
&= \frac{(r+1)!}{r+s+1}\binom{r+2s}{s}.
\end{aligned}
$$

□

3.3 The tubing expansion (part 1)

We are now ready to give the promised formula for maps $\mathcal{H}_I \to \mathbb{K}[z]$ induced by the universal property Theorem 3.1.13. Let us fix a set I of decorations and a family of 1-cocycles

$$
\Lambda_i f(z) = \int_0^z A_i(d/du) f(u)\, du
$$

where

$$
A_i(z) = \sum_{n\geq 0} a_{i,n} z^n.
$$

Given an I-tree t, let us write $d(t)$ for the decoration of the root vertex, and $d(v)$ for the decoration of a vertex v. For a tubing τ of t, we define the *Mellin monomial*[6]

$$
\mathrm{mel}(\tau) = \prod_{\substack{v\in t \\ v\neq \mathrm{rt}\, t}} a_{d(v),\mathrm{rk}(\tau,v)}. \tag{3.9}
$$

With these definitions in hand, we can state the formula.

Theorem 3.3.1 ([8, Theorem 4.2])**.** *With the above setup, the unique map $\varphi\colon \mathcal{H}_I \to \mathbb{K}[z]$ satisfying $\varphi B_+^{(i)} = \Lambda_i \varphi$ is given on trees by the formula*

$$
\varphi(t) = \sum_{\tau\in \mathrm{Tub}(t)} \mathrm{mel}(\tau) \sum_{k=1}^{b(\tau)} a_{d(t),b(\tau)-k} \frac{z^k}{k!}. \tag{3.10}
$$

The proof of this result will make up the remainder of this section.

Example 3.3.2. Let t be the tree that appears in Figure 3.1a. Computing the contributions of the five tubings and summing them up, we get

$$
\begin{aligned}
\varphi(t) = c_0^3\left(c_3 z + c_2\frac{z^2}{2!} + c_1\frac{z^3}{3!} + c_0\frac{z^4}{4!}\right) + 2c_0^2 c_1\left(c_2 z + c_1\frac{z^2}{2!} + c_0\frac{z^3}{3!}\right) \\
+ 2c_0^3\left(c_2 z + c_1\frac{z^2}{2!} + c_0\frac{z^3}{3!}\right)
\end{aligned}
$$

[6] The name is inspired by the physics application, where the series $A(z)$ is, modulo some minor details, the Mellin transform of a primitive Feynman diagram. See Section 4.1.

where the second and third tubing in the figure give the same contribution, as do the fourth and fifth. These coincidences can be explained combinatorially by the fact that in both cases the offending pair of tubings differ in the upper tubes but have the exact same set of lower tubes, which in light of Remark 3.2.4 is sufficient to determine the Mellin monomial and b-statistic.

Remark 3.3.3. The leading term of (3.10) counts binary tubings τ with $b(\tau) = |t|$, i.e. where every upper tube contains the root. Since a lower tube of size greater than 1 must contain an upper tube which cannot contain the root, this implies that every lower tube except the outermost tube is a singleton. Such tubings were called *leaf tubings* in [8]. If we think of tubings as iteratively pulling off subtrees, leaf tubings are the case where we just pull off single leaves, and are completely determined by the order in which we do so. Since we must pull off each vertex before its parent, this must be a linear extension of the tree, and this gives a bijection between leaf tubings and linear extensions. Each non-root vertex in a leaf tubing has rank 0, so the leading coefficient of $\varphi(t)$ is

$$\frac{e(t)}{|t|!}\prod_{v\in t} a_{d(v),0}$$

which in the case of a single 1-cocycle matches what we computed in Proposition 3.1.11.

Remark 3.3.4. By Example 3.1.4 and Example 3.1.9, we can use Theorem 3.3.1 we can get a formula for the order polynomial of a tree as a sum over binary tubings of some terms involving products of Bernoulli numbers.

For the moment, let ψ denote the algebra map defined by the right side of (3.10). Let σ be the linear term of ψ. Explicitly we have

$$\sigma(t) = \sum_{\tau\in\mathrm{Tub}(t)} a_{d(t),b(\tau)-1}\operatorname{mel}(\tau) \tag{3.11}$$

and σ vanishes on disconnected forests (including the empty forest). Since the constant term of ψ is zero on all trees (agreeing with the counit), by Theorem 2.2.24, σ is an infinitesimal character.

Lemma 3.3.5 ([8, Lemma 4.3])**.** *For any tree t and $k \geq 1$,*

$$\sigma^{*k}(t) = \sum_{\substack{\tau\in\mathrm{Tub}(t)\\ b(\tau)\geq k}} a_{d(t),b(\tau)-k}\operatorname{mel}(\tau).$$

Note that for $k > 1$, σ^{*k} is *not* an infinitesimal character and does not vanish on all disconnected forests, though it trivially must vanish on the empty forest. However, it will be sufficient for our purposes to understand its value on trees.

Proof. By induction on k. Since $b(\tau) \geq 1$ for all tubings τ, the base case is exactly (3.11). Then by (2.9) we have

$$\sigma^{*k+1}(t) = (\sigma * \sigma^{*k})(t) = \sum_f \sigma(f)\sigma^{*k}(t\setminus f)$$

as f ranges over subforests (i.e. downsets). However, since σ vanishes on disconnected forests only terms where f is a tree contribute, and since σ^{*k} vanishes on the empty forest we may further restrict to *proper* subtrees $t' \subset t$. Thus, using the induction hypothesis we have

$$\begin{aligned}\sigma^{*k+1}(t) &= \sum_{t'} \sigma(t')\sigma^{*k}(t \setminus t') \\ &= \sum_{t'} \sum_{\tau' \in \mathrm{Tub}(t')} \sum_{\substack{\tau'' \in \mathrm{Tub}(t \setminus t') \\ b(\tau'') \geq k}} a_{d(t'),b(\tau')-1} \operatorname{mel}(\tau') a_{d(t),b(\tau'')-k} \operatorname{mel}(\tau'').\end{aligned}$$

Now by the recursive construction of tubings (Proposition 3.2.5), the pair (τ', τ'') uniquely determines a tubing τ of t, and

$$b(\tau) = b(\tau'') + 1.$$

It follows from (3.5) that the Mellin monomials satisfy

$$\operatorname{mel}(\tau) = a_{d(t'),b(\tau')-1} \operatorname{mel}(\tau') \operatorname{mel}(\tau'')$$

as we get a factor for each non-root vertex of t' and t'' as well as for the root of t' which does not contribute to $\operatorname{mel}(\tau')$. Hence we can rewrite the above triple sum as

$$\sigma^{*k+1}(t) = \sum_{\substack{\tau \in \mathrm{Tub}(t) \\ b(\tau) \geq k+1}} a_{d(t),b(\tau)-k-1} \operatorname{mel}(\tau)$$

as wanted. □

Remark 3.3.6. With the exception of the unique tubing of the one-vertex tree, every binary tubing τ satisfies $b(\tau) \geq 2$. Thus for trees on more than one vertex, σ and σ^{*2} count the same set with slightly different weights. In particular, consider the case that all of the coefficients equal 1, so $\sigma(t)$ is just the number of binary tubings of t. By Lemma 3.3.5, this σ agrees with σ^{*2} on such trees. This recovers Corollary 3.2.7.

Lemma 3.3.7 ([8, Lemma 4.4])**.** *$\psi = \exp_*(z\sigma)$.*

(By Theorem 2.2.24, this is equivalent to ψ being a Hopf algebra morphism.)

Proof. Both sides are algebra morphisms, so we only need to show they agree on trees. We compute

$$\begin{aligned}\psi(t) &= \sum_{\tau \in \mathrm{Tub}(t)} \operatorname{mel}(\tau) \sum_{k=1}^{b(\tau)} a_{d(t),b(\tau)-k} \frac{z^k}{k!} \\ &= \sum_{k \geq 1} \frac{z^k}{k!} \sum_{\substack{\tau \in \mathrm{Tub}(t) \\ b(\tau) \geq k}} a_{d(t),b(\tau)-k} \operatorname{mel}(\tau) \\ &= \sum_{k \geq 1} \frac{z^k}{k!} \sigma^{*k}(t) \\ &= \exp_*(z\sigma)(t).\end{aligned}$$

□

We can now examine how σ interacts with the B_+ operators.

Lemma 3.3.8. *For each $i \in I$,*

$$\sigma B_+^{(i)} = \sum_{k \geq 0} a_{i,k} \sigma^{*k}.$$

Proof. For $k \geq 1$ and $i \in I$ define an infinitesimal character $\sigma_{i,k}$ by

$$\sigma_{i,k}(t) = \sum_{\substack{\tau \in \mathrm{Tub}(t) \\ b(\tau)=k}} c(\tau)$$

when t is a tree with root decorated by i and otherwise zero. Clearly, we then have

$$\sigma = \sum_{i \in I} \sum_{k \geq 1} a_{i,k-1} \sigma_{i,k}.$$

By construction we have $\sigma_{i,k} B_+^{(j)} = 0$ for $j \neq i$. Thus the result follows if we can show that $\sigma_{i,k} B_+^{(i)} = \sigma^{*k-1}$. We do this by induction on i. For the base case, note that the only tubing τ satisfying $b(\tau) = 1$ is the unique tubing of the one-vertex tree, which also has $\mathrm{mel}(\tau) = 1$. Thus $\sigma_{i,1}$ sends the one-vertex tree with decoration i to 1 and all other trees to 0, so $\sigma_{i,1} B_+^{(i)}$ sends the empty forest to 1 and all other forests to 0, i.e. $\sigma_{i,1} B_+^{(i)} = \varepsilon = \sigma^{*0}$.

Now suppose $k \geq 1$ and consider $\sigma_{i,k+1}$. This vanishes on one-vertex trees. For t a tree with more than one vertex and $d(t) = i$ we have (where the sums over t' are summing over proper subtrees)

$$\begin{aligned}
\sigma_{i,k+1}(t) &= \sum_{\substack{\tau \in \mathrm{Tub}(t) \\ b(\tau)=k+1}} \mathrm{mel}(\tau) \\
&= \sum_{t'} \left(\sum_{\tau' \in \mathrm{Tub}(t')} a_{i,b(\tau')-1} \, \mathrm{mel}(\tau') \right) \left(\sum_{\substack{\tau'' \in \mathrm{Tub}(t \setminus t') \\ b(\tau'')=k}} \mathrm{mel}(\tau'') \right) \\
&= \sum_{t'} \sigma(t') \sigma_{i,k}(t \setminus t') \\
&= (\sigma * \sigma_{i,k})(t).
\end{aligned}$$

and thus while $\sigma_{i,k+1} \neq \sigma * \sigma_{i,k}$ the two do agree on the image of $B_+^{(i)}$.

Thus

$$\begin{aligned}
\sigma_{i,k+1} B_+^{(i)} &= (\sigma * \sigma_{i,k}) B_+^{(i)} \\
&= \sigma * \sigma_{i,k} B_+^{(i)} && \text{by Lemma 3.1.5(i)} \\
&= \sigma * \sigma^{*k-1} && \text{by the induction hypothesis} \\
&= \sigma^{*k}
\end{aligned}$$

as wanted. □

Finally, we are ready to prove our formula.

Proof of Theorem 3.3.1. To show $\varphi = \psi$, by uniqueness we need only show that ψ satisfies the required formula $\psi B_+^{(i)} = \Lambda^{(i)}\psi$. Note that both of these have zero constant term, so it is sufficient to show that they agree after differentiating with respect to z. By construction, we have

$$\frac{d}{dz}\Lambda^{(i)} = A_i(d/dz).$$

On the other hand, since $\psi = \exp_*(z\sigma)$, we have

$$\frac{d}{dz}\psi = \psi * \sigma.$$

Now using our lemmas we can compute

$$\begin{aligned}
\frac{d}{dz}\psi B_+^{(i)} &= (\psi * \sigma)B_+^{(i)} \\
&= \psi * \sigma B_+^{(i)} && \text{by Lemma 3.1.5(i)} \\
&= \sum_{k\geq 0} a_{i,k}\psi * \sigma^{*k} && \text{by Lemma 3.3.8} \\
&= \sum_{k\geq 0} a_{i,k}\frac{d^k}{dz^k}\psi && \text{by Lemma 3.3.8} \\
&= \frac{d}{dz}\Lambda^{(i)}\psi.
\end{aligned}$$

The result follows. □

3.4 1-cocycles on tensor powers

We now return to 1-cocycles. In Section 3.1 we defined 1-cocycles for arbitrary comodules over a bialgebra H but focused mainly on the case where the comodule is H itself. In this section, we will consider the case where the comodule is a tensor power of H. (Recall the tensor product of comodules from Section 2.2.3.) This theory is not as well developed, and the results of this section have not previously appeared.

Note that there is a canonical comodule homomorphism $\mu\colon H^{\otimes r} \to H$, namely the multiplication map $\mu(h_1 \cdots h_r) = h_1 \cdots h_r$. By Lemma 3.1.5(iii) we can build various 1-cocycles $\Lambda\mu \in \mathrm{Z}^1(H, H^{\otimes r})$ for various $\Lambda \in \mathrm{Z}^1(H)$. We will call such cocycles *boring*; as we will see, they are generally the trivial case for our results.

As in Section 3.1 we focus on the case of $H = \mathbb{K}[z]$. Similarly to Example 2.2.4 we identify $\mathbb{K}[z]^{\otimes r}$ with $\mathbb{K}[z_1, \ldots, z_r]$ made into a comodule with coaction

$$\delta \mathbf{z}^\alpha = \sum_{\beta\leq\alpha} \binom{\alpha}{\beta} z^{|\alpha|-|\beta|} \otimes \mathbf{z}^\beta. \tag{3.12}$$

We now give a generalization of Theorem 3.1.7 to tensor powers.

Theorem 3.4.1. *For any series $A(\mathbf{z}) \in \mathbb{K}[[z_1, \ldots, z_r]]$, the map $\mathbb{K}[z_1, \ldots, z_r] \to \mathbb{K}[z]$ given by*

$$f(\mathbf{z}) \mapsto \int_0^z A(\partial/\partial u_1, \ldots, \partial/\partial u_r) f(u_1, \ldots, u_r)\big|_{u_1=\cdots=u_r=u}\, du \tag{3.13}$$

is a 1-cocycle. Moreover, all 1-cocycles $\mathbb{K}[z_1, \ldots, z_r] \to \mathbb{K}[z]$ are of this form.

Proof. Let $\psi\colon \mathbb{K}[z_1, \ldots, z_r] \to \mathbb{K}$ be given by

$$\psi(\mathbf{z}^\alpha) = [\mathbf{z}^\alpha]A(\mathbf{z}).$$

Then the operator defined by (3.13) is simply $\mathcal{I} \circledast \psi$ where $\mathcal{I}$ is the usual integral cocycle on $\mathbb{K}[z]$ (see Example 3.1.3) and the $\circledast$ notation was defined in Section 3.1. Thus by Lemma 3.1.5(iii) and Lemma 3.1.6, this operator is indeed a 1-cocycle.

Conversely, suppose Λ is a 1-cocycle. We wish to find a series $A(\mathbf{z})$ such that Λ has the form (3.13). For $\alpha \in \mathbb{N}^r$ take $a_\alpha = \operatorname{lin}\Lambda(\mathbf{z}^\alpha)$ and let $A(\mathbf{z})$ be the exponential generating function for these:

$$A(\mathbf{z}) = \sum_{\alpha \in \mathbb{N}^r} \frac{a_\alpha \mathbf{z}^\alpha}{\alpha_1! \cdots \alpha_r!}.$$

Now observe that for a polynomial $f(z)$ we have $\frac{df(z)}{dz} = \operatorname{lin} \rightharpoonup f(z)$. (This is a special case of Example 2.2.16.) With this in mind,

$$\begin{aligned}
\frac{d\Lambda(\mathbf{z}^\alpha)}{dz} &= \operatorname{lin} \rightharpoonup \Lambda(\mathbf{z}^\alpha) \\
&= (\mathrm{id} \otimes \operatorname{lin})(\Lambda \otimes 1 + (\mathrm{id} \otimes \Lambda)\delta)\mathbf{z}^\alpha \\
&= (\mathrm{id} \otimes \operatorname{lin}\Lambda)\delta\mathbf{z}^\alpha \\
&= \sum_{\beta \le \alpha} \binom{\alpha}{\beta} a_\beta z^{|\alpha|-|\beta|} \\
&= \sum_{\beta \le \alpha} \frac{a_\beta}{\beta_1! \cdots \beta_r!} \prod_{i=1}^r \frac{d^{\beta_i} z^{\alpha_i}}{dz^{\beta_i}} \\
&= \sum_{\beta \le \alpha} \frac{a_\beta}{\beta_1! \cdots \beta_r!} \frac{\partial^{|\beta|}\mathbf{u}^\alpha}{\partial u_1^{\beta_1} \cdots \partial u_r^{\beta_r}}\bigg|_{u_1=\cdots=u_r=z} \\
&= A\left(\frac{\partial}{\partial u_1}, \ldots, \frac{\partial}{\partial u_r}\right)\mathbf{u}^\alpha\bigg|_{u_1=\cdots=u_r=z}
\end{aligned}$$

and hence by linearity, for any polynomial $f(\mathbf{z})$ we have

$$\frac{d\Lambda f(\mathbf{z})}{dz} = A\left(\frac{\partial}{\partial u_1}, \ldots, \frac{\partial}{\partial u_r}\right) f(\mathbf{u})\bigg|_{u_1=\cdots=u_r=z}$$

which is also the derivative of the right-hand side of (3.13). But since $\Lambda f(\mathbf{z})$ must have zero constant term by Lemma 3.1.5(ii), it is exactly given by (3.13), as wanted. □

Remark 3.4.2. The multiplication map $\mathbb{K}[z]^{\otimes r} \to \mathbb{K}[z]$ corresponds to the map $\mathbb{K}[\mathbf{z}] \to \mathbb{K}[z]$ that substitutes z for all of the variables. The adjoint map $\mathbb{K}[[z]] \to \mathbb{K}[[\mathbf{z}]]$ is the substitution $z \mapsto z_1 + \cdots + z_r$. Thus the boring 1-cocycles correspond to series that expand in powers of the sum $z_1 + \cdots + z_r$.

Next we construct the analogue in this setting of the Connes–Kreimer Hopf algebra. Let $\widetilde{\mathcal{T}}_r$ be the set of unlabelled rooted trees with edges decorated by elements of $\{1, \ldots, r\}$ and $\widetilde{\mathcal{F}}_r$ the corresponding set of forests. Define $\widetilde{\mathcal{H}}_r$ to be the free vector space on $\widetilde{\mathcal{F}}_r$, made into an algebra with disjoint union as multiplication and an upset-downset coproduct exactly as in $\mathcal{H}$ but preserving the decorations on the (remaining) edges. Now define $\tilde{B}_+ \colon \widetilde{\mathcal{H}}_r^{\otimes r} \to \widetilde{\mathcal{H}}_r$ as follows: for $f_1, \ldots, f_r \in \widetilde{\mathcal{F}}_r$, $\tilde{B}_+(f_1 \otimes \cdots \otimes f_r)$ is the tree obtained from the forest $f_1 \cdots f_r$ by adding a new root with an edge to the root of each component, where the edges to f_i have decoration i. Note that clearly $\widetilde{\mathcal{H}}_1 \cong \mathcal{H}$ and in this case $\tilde{B}_+$ is just the usual B_+.

Proposition 3.4.3. $\tilde{B}_+ \in \mathrm{Z}^1(\widetilde{\mathcal{H}}_r, \widetilde{\mathcal{H}}_r^{\otimes r})$.

Proof. Let $t = \tilde{B}_+(f_1 \otimes \cdots \otimes f_r)$. The only downset in t that contains the root is all of t, and each other downset is the union of a downset in f_i for each i. In the complementary upset, all edges on the root have the same decoration as they do in t, so

$$\begin{aligned}
\Delta t &= \sum_{f \in J(t)} f \otimes (t \setminus f) \\
&= 1 \otimes t + \sum_{f_1' \in J(f_1)} \cdots \sum_{f_r' \in J(f_r)} f_1' \cdots f_r' \otimes \tilde{B}_+((f_1 \setminus f_1') \otimes \cdots \otimes (f_r \setminus f_r')).
\end{aligned}$$

On the other hand, the left coaction δ of $\widetilde{\mathcal{H}}_r$ on $\widetilde{\mathcal{H}}_r^{\otimes r}$ is, by definition,

$$\delta(f_1 \otimes \cdots \otimes f_r) = \sum_{f_r' \in J(f_r)} f_1' \cdots f_r' \otimes (f_1 \setminus f_1') \otimes \cdots \otimes (f_r \setminus f_r').$$

Comparing these, we see that indeed

$$\Delta \tilde{B}_+ = 1 \otimes \tilde{B}_+ + (\mathrm{id} \otimes \tilde{B}_+)\delta. \qquad \square$$

The universal property of $\mathcal{H}$ naturally extends to $\widetilde{\mathcal{H}}_r$. The proof is essentially the same as the original Connes–Kreimer result (Theorem 2.3.3).

Theorem 3.4.4. *Let A be a commutative algebra and $\Lambda \colon A^{\otimes r} \to A$ a linear map. There exists a unique map $\varphi \colon \widetilde{\mathcal{H}}_r \to A$ such that $\varphi \tilde{B}_+ = \Lambda \varphi^{\otimes r}$. Moreover, if A is a bialgebra and Λ is a 1-cocycle then φ is a bialgebra morphism.*

Proof. Suppose $t \in \widetilde{\mathcal{T}}_r$. We can uniquely write t in the form $t = \tilde{B}_+(f_1 \otimes \cdots \otimes f_r)$. Then we can recursively set

$$\varphi(t) = \Lambda(\varphi(f_1) \otimes \cdots \otimes \varphi(f_r))$$

where for a forest, $\varphi(f)$ is the product over the components. Clearly this is a well-defined algebra map and the unique one satisfying the desired identity.

If A is a bialgebra and Λ is a 1-cocycle, we compute

$$\begin{aligned}\Delta\varphi(t) &= \Delta\Lambda(\varphi(f_1)\otimes\cdots\otimes\varphi(f_r))\\ &= \varphi(t)\otimes 1 + (\mathrm{id}\otimes\Lambda)\delta(\varphi(f_1)\otimes\cdots\otimes\varphi(f_r))\end{aligned}$$

where δ is the coaction of A on $A^{\otimes r}$. Now suppose that φ preserves coproducts for each of the forests, i.e.

$$\Delta\varphi(f_i) = \sum_{f_i'\in J(f_i)} \varphi(f_i')\otimes\varphi(f_i\setminus f_i').$$

Then

$$\delta(\varphi(f_1)\otimes\cdots\otimes\varphi(f_r)) = \sum_{f_1'\in J(f_1)}\cdots\sum_{f_r'\in J(f_r)} \varphi(f_1'\cdots f_r')\otimes\varphi(f_1\setminus f_1')\otimes\cdots\otimes\varphi(f_r\setminus f_r')$$

and hence

$$\begin{aligned}\Delta\varphi(t) &= \varphi(t)\otimes 1 + \sum_{f_1'\in J(f_1)}\cdots\sum_{f_r'\in J(f_r)} \varphi(f_1'\cdots f_r')\otimes\Lambda(\varphi(f_1\setminus f_1')\otimes\cdots\otimes\varphi(f_r\setminus f_r'))\\ &= \varphi(t)\otimes 1 + \sum_{f_1'\in J(f_1)}\cdots\sum_{f_r'\in J(f_r)} \varphi(f_1'\cdots f_r')\otimes\varphi(\tilde{B}_+((f_1\setminus f_1')\otimes\cdots\otimes(f_r\setminus f_r'))\\ &= \varphi(t)\otimes 1 + \sum_{f\in J(t)}\varphi(f)\otimes\varphi(t\setminus f)\\ &= (\varphi\otimes\varphi)(\Delta t)\end{aligned}$$

as desired. □

Example 3.4.5. Consider the (boring) 1-cocycle $B_+\mu \in \mathrm{Z}^1(\mathcal{H},\mathcal{H}^{\otimes r})$. By Theorem 3.4.4, there is a unique $\beta\colon \widetilde{\mathcal{H}}_r \to \mathcal{H}$ such that $\beta\tilde{B}_+ = B_+\mu\beta^{\otimes r} = B_+\beta\mu$. It is easily seen that such a map is given by simply forgetting the edge decorations. More generally, consider any boring 1-cocycle $\Lambda\mu$ on any bialgebra H. Let $\psi\colon \mathcal{H}\to H$ be the unique map such that $\psi B_+ = \Lambda\psi$ and let $\beta\colon\widetilde{\mathcal{H}}_r\to\mathcal{H}$ continue to denote the edge-undecorating map. Now we observe that

$$\psi\beta\tilde{B}_+ = \psi B_+\mu\beta^{\otimes r} = \Lambda\psi^{\otimes r}\mu\beta^{\otimes r} = \Lambda\mu(\psi\beta)^{\otimes r}$$

so the universal map is simply $\varphi = \psi\beta$ and the boring case is indeed boring.

In the next section we will generalize the tubing expansion to the map from Theorem 3.4.4. As before, we will want a more general version including several 1-cocycles at once. For this it is convenient to allow tensor products indexed by arbitrary finite sets rather than just ordinals. Note that everything we did with $\mathbb{K}[z]$ still works here: we can identify $\mathbb{K}[z]^{\otimes E}$ with $\mathbb{K}[z_e\colon e\in E]$, and all 1-cocycles are integro-differential operators as in Theorem 3.4.1 but slightly more notationally challenging.

Let I be a set and $\mathcal{E} = \{E_i\}_{i\in I}$ be a family of finite sets. Define a $(I,\mathcal{E})$-*tree* to be a rooted tree where each vertex is decorated by an element of I and each edge from a parent of type i is decorated by an element of E_i. Denote the set of $(I,\mathcal{E})$-trees by $\widetilde{\mathcal{T}}(I,\mathcal{E})$. Analogously we have $(I,\mathcal{E})$-forests and we denote the set of these by $\widetilde{\mathcal{F}}(I,\mathcal{E})$. Let $\widetilde{\mathcal{H}}_{I,\mathcal{E}}$ denote the free vector space on $\widetilde{\mathcal{F}}(I,\mathcal{E})$. As usual, we make this into a bialgebra with disjoint union as the product and an upset-downset coproduct preserving the decorations on the vertices and (remaining) edges. For $i \in I$, let $\tilde{B}_+^{(i)} : \widetilde{\mathcal{H}}_{I,\mathcal{E}}^{\otimes E_i} \to \widetilde{\mathcal{H}}_{I,\mathcal{E}}$ be the operator that adds a new root joined to each component with the appropriate decoration. These are 1-cocycles by the same argument as in Proposition 3.4.3.

Theorem 3.4.6. *Let A be a commutative algebra and $\{\Lambda_i\}_{i\in I}$ a family of linear maps, $\Lambda_i : A^{\otimes E_i} \to A$. There exists a unique algebra morphism $\varphi : \widetilde{\mathcal{H}}_{I,\mathcal{E}} \to A$ such that $\varphi B_+^{(i)} = \Lambda_i \varphi^{\otimes E_i}$. Moreover, if Λ_i is a 1-cocycle for each i then φ is a bialgebra morphism.*

Proof. Analogous to Theorem 3.4.4. □

Remark 3.4.7. Analogously to Example 3.4.5, if *all* of the 1-cocycles are boring there will be a factorization of φ through the map $\widetilde{\mathcal{H}}_{I,\mathcal{E}} \to \mathcal{H}_I$ that forgets edge decorations. However, we can make a more refined statement: if Λ_i is boring, then φ is independent of the edge decorations for vertices of type i. Proving this is left as an exercise (though in the case that the target algebra is $\mathbb{K}[z]$ it will follow from the results of the next section).

3.5 The tubing expansion (part 2)

Fix a family $\mathcal{E} = \{E_i\}_{i\in I}$ of finite sets. For each i, introduce indeterminates $\mathbf{z}_i = (z_e : e \in E_i)$ and choose a 1-cocycle $\Lambda_i \in \mathrm{Z}^1(\mathbb{K}[z], \mathbb{K}[\mathbf{z}_i])$. Let $A_i(\mathbf{z}_i)$ be the corresponding power series given by Theorem 3.4.1. We will choose to expand the series using multinomial coefficients:

$$A_i(\mathbf{z}_i) = \sum_{\alpha\in\mathbb{N}^{E_i}} a_{i,\alpha} \binom{|\alpha|}{\alpha} \mathbf{z}_i^\alpha.$$

This curious-looking convention can be justified by the observation that if Λ_i is boring then, by Remark 3.4.2, we can write

$$A_i(\mathbf{z}_i) = B\left(\sum_{e\in E_i} z_e\right)$$

for some series $B(z)$. Our convention is such that in this case we have $a_{i,\alpha} = [z^{|\alpha|}]B(z)$. In particular, this will make our expansion manifestly identical to Theorem 3.3.1 in the case that all of the cocycles are boring, and more generally make it obviously independent of the edge decorations for those vertex types with boring 1-cocycles as suggested by Remark 3.4.7.

To generalize the tubing expansion to this case, we need the appropriate generalizations of the statistics that appear in Theorem 3.3.1. Let t be an $(I,\mathcal{E})$-tree and τ be a binary tubing of t. Suppose $t' \in \tau$ is an upper tube and t'' the corresponding lower tube (i.e. the unique lower tube such that $t' \cup t'' \in \tau$). We define the *type* of t' to be the decoration of the first edge on the unique path from $\mathrm{rt}\, t'$ to $\mathrm{rt}\, t''$. By construction, the type is an element of E_i where $i = d(t')$ is the decoration of the root vertex of t'. Note we only assign types to upper tubes, not lower tubes. For each vertex $v \in t$ and edge type $e \in E_{d(v)}$ define the *e-rank* $\mathrm{rk}_e(\tau, v)$ to be the number of upper tubes of type e rooted at v. Collect these together to get

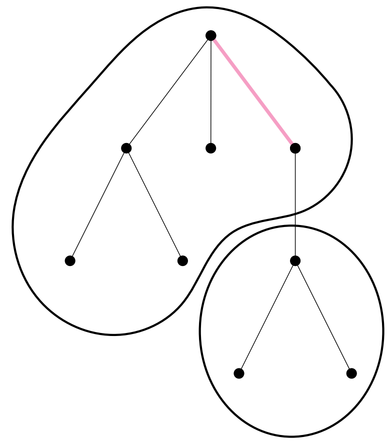


Figure 3.2. An upper tube and its corresponding lower tube. The type of the upper tube is the decoration of the highlighted edge.

the *rank vector* $\mathbf{rk}(\tau, v) \in \mathbb{N}^{E_{d(v)}}$. Clearly $|\,\mathbf{rk}(\tau, v)| = \mathrm{rk}(\tau, v)$. Then we define the *Mellin monomial* of τ to be

$$\mathrm{mel}(\tau) = \prod_{\substack{v \in t \\ v \neq \mathrm{rt}\, t}} a_{d(v), \mathbf{rk}(\tau, v)}.$$

The analogue of the b-statistic is slightly more complicated. For $1 \le k \le b(\tau)$ write $\beta_i^k(\tau)$ the number of upper tubes of type i containing (hence rooted at) $\mathrm{rt}\, t$, excluding the outermost $k-1$ upper tubes. Collect these into a vector $\beta^k(\tau)$ (which for lack of a better name we simply term the k*th* β*-vector* of τ). Thus $\beta^1(\tau) = \mathbf{rk}(\tau, \mathrm{rt}\, t)$ and $|\beta^k(\tau)| = b(\tau) - k$. We are now ready to state our expansion.

Theorem 3.5.1. *With the above setup, the unique map* $\varphi \colon \widetilde{\mathcal{H}}_{I,\mathcal{E}} \to \mathbb{K}[z]$ *satisfying* $\varphi \tilde{B}_+^{(i)} = \Lambda_i \varphi$ *is given on trees by the formula*

$$\varphi(t) = \sum_{\tau \in \mathrm{Tub}(t)} \mathrm{mel}(\tau) \sum_{k=1}^{b(\tau)} a_{d(t), \beta^k(\tau)} \frac{z^k}{k!}. \tag{3.14}$$

To prove this, we essentially follow the same argument as in Section 3.3, with some added complexity. As we did then, we will temporarily write ψ for the right side of (3.14). Let σ be the linear term of ψ, i.e.

$$\sigma(t) = \sum_{\tau \in \mathrm{Tub}(t)} a_{d(t), \beta^1(\tau)} \,\mathrm{mel}(\tau).$$

and σ vanishes on disconnected forests.

Lemma 3.5.2. *For any tree* t *and* $k \ge 1$,

$$\sigma^{*k}(t) = \sum_{\substack{\tau \in \mathrm{Tub}(t) \\ b(\tau) \ge k}} a_{d(t), \beta^k(\tau)} \,\mathrm{mel}(\tau).$$

(As in Lemma 3.3.5, σ^{*k} can be nonzero on disconnected forests, but we are claiming this equality only for trees.)

Proof. By induction on k. The base case is true by definition. Then

$$\sigma^{*k+1}(t) = \sum_{f \in J(t)} \sigma(f)\sigma^{*k}(t \setminus f)$$

but σ is an infinitesimal character so $\sigma(f) \neq 0$ only if f is actually a tree. Moreover, clearly $\sigma^{*k}(1) = 0$, so we can restrict the sum to proper subtrees t'. Then inductively we have

$$\begin{aligned}\sigma^{*k+1}(t) &= \sum_{t'} \sigma(t')\sigma^{*k}(t \setminus t') \\ &= \sum_{t'} \sum_{\tau' \in \mathrm{Tub}(t')} \sum_{\substack{\tau'' \in \mathrm{Tub}(t \setminus t') \\ b(\tau'') \geq k}} a_{d(t'),\beta^1(\tau')} \operatorname{mel}(\tau') a_{d(t),\beta^k(\tau'')} \operatorname{mel}(\tau'').\end{aligned}$$

By the recursive construction of tubings (Proposition 3.2.5), τ' and τ'' uniquely determine a tubing $\tau \in \mathrm{Tub}(t)$. Note that for any vertex $v \in t'$, the upper tubes of τ rooted at v are the same as those of τ', so $\mathbf{rk}(\tau, v) = \mathbf{rk}(\tau', v)$. The same is true for vertices of τ'' other than $\mathrm{rt}\, t$. Thus we have have

$$\operatorname{mel}(\tau) = \operatorname{mel}(\tau') a_{d(t'),\mathbf{rk}(\tau,\mathrm{rt}\, t')} \operatorname{mel}(\tau'') = \operatorname{mel}(\tau') a_{d(t'),\beta^1(\tau')} \operatorname{mel}(\tau'').$$

Moreover, since β^k ignores outermost the $k-1$ upper tubes containing $\mathrm{rt}\, t$ by definition, we have $\beta^{k+1}(\tau) = \beta^k(\tau'')$. Finally, there is one additional tube in τ containing the root, so $b(\tau) = b(\tau'') + 1$. Hence the triple sum simplifies to

$$\sigma^{*k+1}(t) = \sum_{\substack{\tau \in \mathrm{Tub}(t) \\ b(\tau) \geq k+1}} a_{d(t),\beta^{k+1}(\tau)} \operatorname{mel}(\tau)$$

as wanted. □

For notational convenience, for $\alpha \in \mathbb{N}^{E_i}$ let us write $\sigma^{[\alpha]}$ for the linear form on $\mathcal{H}_{D,\mathcal{I}}^{\otimes I_d}$ given by

$$\sigma^{[\alpha]} = \bigotimes_{e \in E_d} \sigma^{*\alpha_e}.$$

Note that the coalgebra structure on $\mathcal{H}_{I,\mathcal{E}}^{\otimes E_i}$ gives a convolution product on linear forms; one easily checks that $\sigma^{[\alpha+\beta]} = \sigma^{[\alpha]} * \sigma^{[\beta]}$.

Lemma 3.5.3. *For $i \in I$,*

$$\sigma \tilde{B}_+^{(i)} = \sum_{\alpha \in \mathbb{N}^{E_i}} \binom{|\alpha|}{\alpha} a_{i,\alpha} \sigma^{[\alpha]}.$$

Proof. For each $i \in I$ and $\alpha \in \mathbb{N}^{E_i}$ let $\sigma_{i,\alpha}$ be the infinitesimal character defined by

$$\sigma_{i,\alpha}(t) = \begin{cases} \sum_{\tau \in \mathrm{Tub}(t), \beta^1(\tau)=\alpha} \mathrm{mel}(\tau) & d(t) = i \\ 0 & \text{otherwise} \end{cases}$$

so that we have

$$\sigma = \sum_{i \in I} \sum_{\alpha \in \mathbb{N}^{E_i}} a_{i,\alpha} \sigma_{i,\alpha}.$$

Note that we have $\sigma_{i,\alpha} \tilde{B}_+^{(j)} = 0$ when $i \neq j$, so to get the desired formula it suffices to show that $\sigma_{i,\alpha} \tilde{B}_+^{(i)} = \binom{|\alpha|}{\alpha} \sigma^{[\alpha]}$ for $i \in I$ and $\alpha \in \mathbb{N}^{E_i}$. We do this by induction on $m = |\alpha|$.

Now, for $\alpha = 0$ we have that $\sigma_{i,\alpha}(t) = 1$ if t is the one-vertex tree with decoration i and 0 otherwise. Of course $\sigma^{[0]}$ is 1 when each forest is empty and 0 otherwise, so we do indeed see that $\sigma_{i,0} \tilde{B}_+^{(i)} = \sigma^{[0]}$ as wanted.

Suppose now that $m > 0$ and that the desired identity holds for smaller values. Then $\sigma_{i,\alpha}$ vanishes on one-vertex trees, so all tubings of interest have at least one upper tube. For $e \in E_i$, let $\sigma_{i,\alpha}^e(t)$ be the sum only for those tubings where the outermost upper tube has type e. Thus

$$\sigma_{i,\alpha} = \sum_{e \in E_i} \sigma_{i,\alpha}^e.$$

Note $\sigma_{i,\alpha}^e = 0$ when $\alpha_e = 0$, as in this case there must be no tube of type i containing the root. Suppose now that $\alpha_e \neq 0$ and $t = \tilde{B}_+^{(i)}\left(\bigotimes_{e' \in E_i} f_{e'}\right)$. Let τ be a binary tubing of t which recursively corresponds to (τ', τ''). Then the outermost upper tube of τ is type e precisely when the lower subtree t' is contained in f_e. Moreover in this case we have $\beta^1(\tau) = \beta^1(\tau'') + 1_e$, where 1_e is the indicator vector for e. Then we have

$$\sigma_{i,\alpha}^e(t) = \sum_{\substack{t' \subseteq f_i \\ \text{subtree}}} \sigma(t') \sigma_{i,\alpha - 1_i}(t \setminus t').$$

Now $t \setminus t' = \tilde{B}_+^{(i)}(\bigotimes_{e'} f'_{e'})$ where $f'_e = f_e \setminus t'$ and $f'_{e'} = f_{e'}$ for $e \neq e'$. It follows that

$$\begin{aligned} \sigma_{i,\alpha}^e B_+^{(i)} &= \sigma^{[1_e]} * \sigma_{i,\alpha-1_e} \tilde{B}_+^{(i)} \\ &= \sigma^{[1_e]} * \binom{m-1}{\alpha - 1_e} \sigma^{[\alpha - 1_e]} \\ &= \binom{m-1}{\alpha - 1_e} \sigma^{[\alpha]}. \end{aligned}$$

Finally, by summing over the values of e we get

$$\sigma_{i,\alpha} \tilde{B}_+^{(i)} = \sum_{e \in E_i} \binom{m-1}{\alpha - 1_e} \sigma^{[\alpha]} = \binom{m}{\alpha} \sigma^{[\alpha]}$$

by the multinomial analogue of the Pascal recurrence. □

Proof of Theorem 3.5.1. Let $\Psi_i\colon \widetilde{\mathcal{H}}_{I,\mathcal{E}}^{\otimes E_i} \to \mathbb{K}[\mathbf{z}_i]$ be given by

$$\Psi_i\left(\bigotimes_{e\in E_i} f_e\right) = \prod_{e\in E_i}\Big(\psi(f_e)\big|_{z=z_e}\Big).$$

This is simply the map $\psi^{\otimes E_i}$ carried through the identification of $\mathbb{K}[z]^{\otimes E_i}$ with $\mathbb{K}[\mathbf{z}_i]$. Thus, our goal is to show $\psi\tilde{B}_+^{(i)} = \Lambda_i\Psi_i$; by uniqueness this shows $\varphi = \psi$. Note that since ψ is an algebra morphism, we have

$$\Psi_i\left(\bigotimes_{e\in E_i} f_e\right)\Bigg|_{z_e=z\,\forall e\in E_i} = \prod_{e\in E_i}\psi(f_e) = \psi\left(\prod_{e\in E_i} f_e\right).$$

By Lemma 3.5.2 we have $\psi = \exp_*(z\sigma)$. Thus

$$\frac{d}{dz}\psi\tilde{B}_+^{(i)} = (\psi * \sigma)\tilde{B}_+^{(i)} = \psi *_\delta \sigma\tilde{B}_+^{(i)}$$

where δ is the coaction and the second equality is by Lemma 3.1.5(i). But observe that

$$\begin{aligned}
\left(\psi *_\delta \sigma\tilde{B}_+^{(i)}\right)\left(\bigotimes_{e\in E_i} f_e\right) &= \sum_{\substack{f'_e\in J(f_e)\\ \forall e\in E_i}} \psi\left(\prod_{e\in E_i} f'_e\right)\sigma\left(\tilde{B}_+^{(i)}\left(\bigotimes_{e\in e_i}(f_e\setminus f'_e)\right)\right)\\
&= \sum_{\substack{f'_e\in J(f_e)\\ \forall e\in E_i}} \Psi_i\left(\bigotimes_{e\in E_i} f'_e\right)\sigma\left(\tilde{B}_+^{(i)}\left(\bigotimes_{e\in e_i}(f_e\setminus f'_e)\right)\right)\Bigg|_{z_e=z\,\forall e\in E_i}\\
&= \left(\Psi_i * \sigma\tilde{B}_+^{(i)}\right)\left(\bigotimes_{e\in E_i} f_e\right)\Bigg|_{z_e=z\,\forall e\in E_i}.
\end{aligned}$$

Thus we have

$$\begin{aligned}
\frac{d}{dz}\psi\tilde{B}_+^{(i)} &= \Psi_i * \sigma\tilde{B}_+^{(i)}\big|_{z_e=z\,\forall e\in E_i}\\
&= \left(\Psi_i * \sum_{\alpha\in\mathbb{N}^{E_i}}\binom{|\alpha|}{\alpha}a_{i,\alpha}\sigma^{[\alpha]}\right)\Bigg|_{z_e=z\,\forall e\in E_i} && \text{by Lemma 3.5.3}\\
&= A_i(\partial/\partial z_e\colon e\in E_i)\Psi\big|_{z_e=z\,\forall e\in E_i}\\
&= \frac{d}{dz}\Lambda_i\Psi.
\end{aligned}$$

Since we also have

$$\psi(\tilde{B}_+^{(i)}1) = a_{i,0} = \Lambda_i(1)$$

this implies $\psi\tilde{B}_+^{(i)} = \Lambda_i\Psi_i$, as wanted. ∎

Chapter 4

Dyson–Schwinger Equations

4.1 Dyson–Schwinger equations in physics

In the introduction, we claimed that the Dyson–Schwinger equations are "quantum equations of motion". In this section we will attempt to elaborate on that remark and to explain what these equations actually are. This section serves primarily as motivation for the remainder of the chapter and can be safely skipped by the physics-averse.

Let us begin by wondering aloud what a *quantum equation of motion* should be. Suppose our universe consists of a single lonely particle drifting endlessly through a featureless vacuum. Classical equations of motion can tell us, for instance, where the particle will be tomorrow, assuming we know where it is today. In a quantum universe, however, things are rarely so certain. An electron, for instance, may spontaneously emit a photon and lose energy, leaving the universe slightly less lonely but making a deterministic answer to the question of where the electron may be found in the future impossible. Moreover, electrons are indistinguishable, so even the question is somewhat suspicious: asking whether the electron we observed today is the same one we saw yesterday is meaningless.

We are thus reduced to asking questions like: given that we detected a particle at a certain point in spacetime, what is the probability (or rather *amplitude*) that we will also find a particle of the same type at some other point in spacetime? This type of question is answered in quantum field theory by a (two-point) *correlation function*. Equations of motion in this context should then be equations satisfied by the correlation functions that are sufficient to determine them completely. The Dyson–Schwinger equations are one formulation of such equations.

Before we think about what these equations might look like, we should answer a more basic question: what exactly is the correlation function a function of? Naïvely, the answer should be (in our example) a pair of spacetime positions. We will assume that the laws of physics in our lonely particle's universe are translation invariant, so that the amplitude really depends only on the (spacetime) *displacement* rather than the specific positions. For reasons we refuse to delve into, we will take a Fourier transform and write our correlation function in *momentum space*, as a function of the (four-)momentum.

In Feynman's approach to quantum theory [20, 21], the amplitude we are interested in is a sum or integral over the possible paths that the particle may take. These may be represented graphically using Feynman diagrams. These are essentially just graphs, with each edge representing a particle travelling at constant momentum and each vertex an interaction between particles. If there are several types of particles in our theory, the edges will be decorated with these types. The vertices will be constrained by what types of interactions between particles exist in the theory we are considering. The diagrams that contribute to the two-point function have two "external edges", both of the type corresponding to the particle we are looking at. The correlation function $\overline{G}$ is then a combinatorial sum over these diagrams:

$$\overline{G}(\mathbf{p}) = \sum_{\gamma} \varphi(\gamma).$$

Here φ, the *Feynman rules*, assigns each graph its contribution, which is an integral over the space of all possible processes of that "shape". These are given by assigning a momentum to each edge, satisfying momentum conservation; in graph-theoretic language they are $\mathbb{R}^4$-valued flows on the graph.

We can simplify this picture by doing some combinatorics. First, observe that if the graph has a bridge (cut edge) then the momentum through that edge must equal the overall momentum $\mathbf{p}$. Thus the graph can be thought of as a concatenation of two (or more) bridgeless[1] pieces, and we get a geometric series expansion

$$\overline{G}(\mathbf{p}) = \frac{1}{1 - \sum_{\gamma \text{ bridgeless}} \varphi(\gamma)}. \tag{4.1}$$

We denote the denominator by $G(\mathbf{p})$; it will be more convenient to formulate the DSEs in terms of this function.

Bridgeless diagrams still have recursive structure. Given any edge of a Feynman diagram, we may make a more complicated diagram by replacing that edge with another Feynman diagram with two external edges; this is called *inserting* into the edge. Bridgeless diagrams which cannot be built by insertions are called *primitive*.[2] Thus we can decompose

$$G(\mathbf{p}) = 1 - \sum_{\gamma \text{ primitive}} G_\gamma(\mathbf{p}) \tag{4.2}$$

where $G_\gamma(\mathbf{p})$ is the sum over diagrams obtained by (perhaps repeatedly) inserting arbitrary diagrams into γ. In the case of only one type of edge, such diagrams are in bijection with families of diagrams indexed by the internal edges of γ. Such families are more directly enumerated by a power $\overline{G}(\mathbf{p})^{|E(\gamma)|}$. If our Feynman rules behave nicely, we can rearrange the sum making up $G_\gamma(\mathbf{p})$ to actually express it in terms of this power, leaving us with

$$G_\gamma(\mathbf{p}) = \mathcal{L}_\gamma(\overline{G}(\mathbf{p})^{|E(\gamma)|}). \tag{4.3}$$

for some integral operator $\mathcal{L}_\gamma$. Combining (4.1), (4.2), and (4.3) gives us the *Dyson–Schwinger equation* for the two-point function in a theory with only one type of particle:

$$G(\mathbf{p}) = 1 - \sum_{\gamma \text{ primitive}} \mathcal{L}_\gamma(G(\mathbf{p})^{-|E(\gamma)|}). \tag{4.4}$$

[1] Physicists would call these *one-particle irreducible* or *1PI* for short.

[2] Primitive diagrams are in fact primitive elements in a certain Hopf algebra of Feynman diagrams; see [62, Chapter 5].

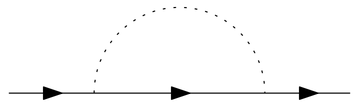


Figure 4.1. A primitive diagram in Yukawa theory.

More generally, we may have several types of particles each with its own two-point function, in which case we will have a system of equations.

We now give an explicit example of what these equations look like. This example comes from a quantum field theory called *(massless) Yukawa theory*. This theory has two types of particles and many primitive diagrams, but can be approximated by considering only those diagrams built by inserting the single primitive diagram shown in Section 4.1 into itself repeatedly. This will give us a single equation with a single integral operator:

$$G(\mathbf{p}) = 1 - x \int_{\mathbb{R}^4} \frac{\langle \mathbf{p}, \mathbf{k} \rangle \, d^4\mathbf{k}}{\|\mathbf{p}\|^2 \|\mathbf{k}\|^2 \|\mathbf{p}+\mathbf{k}\|^2 G(\mathbf{k})}. \tag{4.5}$$

Unfortunately, as is often the case in quantum field theory, we are now forced to admit that what we have written is nonsense: the integrand has obvious singularities at $\mathbf{k} = 0$ and $\mathbf{k} = -p$ and there is no possible way that this integral can yield a finite answer. To make sense of it requires *renormalization*. For fear of awakening shadow and flame we choose not to delve too deep into this subject and only note that there are many different ways to renormalize but the simplest for our purposes is *kinematic renormalization* which can be done at the level of the Dyson–Schwinger equation itself.[3] Here we choose a reference scale μ and replace our equation with

$$G_{\text{ren}}(\mathbf{p}) = 1 - x \int_{\mathbb{R}^4} \left(\frac{\langle \mathbf{p}, \mathbf{k} \rangle \, d^4\mathbf{k}}{\|\mathbf{p}\|^2 \|\mathbf{k}\|^2 \|\mathbf{p}+\mathbf{k}\|^2 G_{\text{ren}}(\mathbf{k})} - \frac{\langle \mathbf{p}_0, \mathbf{k}_0 \rangle \, d^4\mathbf{k}_0}{\|\mathbf{p}_0\|^2 \|\mathbf{k}_0\|^2 \|\mathbf{p}_0+\mathbf{k}_0\|^2 G_{\text{ren}}(\mathbf{k}_0)} \right). \tag{4.6}$$

where $\mathbf{p}_0$ is a vector of norm μ parallel to $\mathbf{p}$ (i.e. $\mathbf{p}_0 = \mu\mathbf{p}/\|\mathbf{p}\|$), and $\mathbf{k}_0 = \mu\mathbf{k}/\|\mathbf{p}\|$. Observe that this cancels out the singularities in the integrand. There are still some analytic difficulties here but this equation can at least be solved for a formal power series in x. The coefficients in the series are functions of $\mathbf{p}$, but note that due to the rotation invariance of the integrand they actually depend only on the norm $\|\mathbf{p}\|$. More specifically we will expand them in the variable $L = \log(\|\mathbf{p}\|^2/\mu^2)$ and rename $G_{\text{ren}}(\mathbf{p})$ to $G(x, L)$.

Substituting $\mathbf{p} = e^{L/2}\mathbf{p}_0$ and $\mathbf{k} = e^{L/2}\mathbf{k}_0$ into (4.6) gives us

$$G(x, L) = 1 - \frac{x}{\mu^2} \int_{\mathbb{R}^4} \left(\frac{1}{G(x, L + \log(\|\mathbf{k}_0\|^2/\mu^2))} - \frac{1}{G(x, \log(\|\mathbf{k}_0\|^2/\mu^2))} \right) \frac{\langle \mathbf{p}_0, \mathbf{k}_0 \rangle \, d^4\mathbf{k}_0}{\|\mathbf{k}_0\|^2 \|\mathbf{p}_0 + \mathbf{k}_0\|^2}.$$

We can put this in a more convenient form by the following trick: for a polynomial or convergent power series f and a positive real number ξ, we have $f(\log \xi) = f(\partial/\partial\rho)\xi^\rho|_{\rho=0}$. Thus

$$\frac{1}{G(x, L + \log(\|\mathbf{k}_0\|^2/\mu^2))} - \frac{1}{G(x, \log(\|\mathbf{k}_0\|^2/\mu^2))} = G(x, \partial/\partial\rho)^{-1}(e^{L\rho} - 1)(\|\mathbf{k}_0\|^2/\mu^2)^\rho \Big|_{\rho=0}$$

[3]For more details and a comparison of different renormalization schemes in the context of DSEs, see [7].

and we can then pull all of the G's and L's out of the integral:

$$G(x, L) = 1 - xG(x, \partial/\partial\rho)^{-1}(e^{L\rho} - 1)F(\rho)\big|_{\rho=0} \tag{4.7}$$

where

$$F(\rho) = \frac{1}{\mu^{2+2\rho}} \int_{\mathbb{R}^4} \frac{\langle \mathbf{p}_0, \mathbf{k}_0 \rangle \, d^4\mathbf{k}_0}{\|\mathbf{k}_0\|^{2-2\rho} \|\mathbf{p}_0 + \mathbf{k}_0\|^2}$$

is the *Mellin transform* of the primitive diagram, which is meromorphic with a simple pole at 0. (Note that this implies that the right-hand side is well-defined as a nonzero formal power series, since $(e^{L\rho} - 1)F(\rho)$ is analytic at zero.)

Clearly, (4.7) can be solved uniquely as it is easily seen that the coefficient of x^n is recursively determined by the earlier coefficients. These coefficients will be polynomials in L, so the solution lies in $\mathbb{K}[L][[x]]$. Moreover, the operator on polynomials that implicitly appears in the equation is one we have seen before in Remark 3.1.10, taking $A(z) = -zF(z)$. In particular it is a 1-cocycle, so we can use our work in Chapter 3 to study this equation.

4.2 Solving Dyson–Schwinger equations

4.2.1 Setup

We begin by laying out an algebraic framework for Dyson–Schwinger equations. Let P be a set (finite or infinite) and assign each $p \in P$ a weight $w_p \in \mathbb{N}_+$, such that there are only finitely many elements of each weight, and an *insertion exponent* $\mu_p \in \mathbb{K}$. In the physical application P is the set of primitive Feynman diagrams, w_p is the dimension of the cycle space of p, and μ_p is – at least in examples like the one in the previous section – the negative of the number of edges of p.[4] To each $p \in P$ we also associate a 1-cocycle Λ_p on the polynomial Hopf algebra $\mathbb{K}[L]$. The *Dyson–Schwinger equation* (DSE) defined by these data is

$$G(x, L) = 1 + \sum_{p \in P} x^{w_p} \Lambda_p(G(x, L)^{\mu_p}). \tag{4.8}$$

(Note that here and throughout the chapter, expressions such as $\Lambda_p(G(x, L)^{\mu_p})$ are to be interpreted as meaning that we expand the argument as a series in x and apply the operator coefficientwise.) By Theorem 3.1.7 we can write

$$\Lambda_p f(L) = \int_0^L A_p(d/du) f(u) \, du \tag{4.9}$$

for some series $A_p(z) \in \mathbb{K}[[z]]$ which physically is more or less the Mellin transform of p.

A particularly nice case, which covers most of the physically reasonable examples, is when there is a linear relationship between the weights and the insertion exponents: $\mu_p = 1 + sw_p$

[4] Note however that we are not even requiring it to be an integer! Non-integer insertion exponents do not seem to fit into the story we told in Section 4.1 and may or may not have any physical relevance, but this generality will come for free from our approach, in which we essentially treat the insertion exponents as indeterminates.

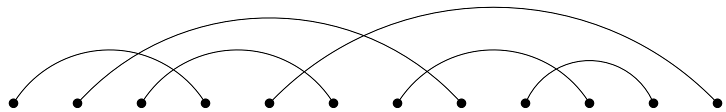


Figure 4.2. A connected chord diagram represented linearly.

for some $s \in \mathbb{K}$. In this case we can combine terms to get

$$G(x, L) = 1 + \sum_{k\geq 1} x^k \Lambda_k(G(x, L)^{1+sk}). \tag{4.10}$$

Previous work on combinatorial solutions to Dyson–Schwinger equations has focused on this case, and indeed equations of this form have some special properties which we discuss in Section 4.5. However, our main results apply in the more general form (4.8).

Remark 4.2.1. Marie and Yeats [41] gave a combinatorial solution to (4.10) in the case $s = -1$ and only one cocycle Λ_1; this was extended by Hihn and Yeats [31] to solve (4.10) in general for s a negative integer. We give a brief summary of the former result, which is representative of the general flavour.

First we introduce the combinatorial objects involved. (These will not appear in our results.) A *(rooted) chord diagram* of size n, for our purposes, is a partition of $\{1, \dots, 2n\}$ into n sets of size 2 which we call *chords*. The terminology suggests that the ground set should be thought of as points on a circle, but we will prefer to think of them as linearly ordered, as in Figure 4.2. The *root chord* is the one that contains 1. Two chords $\{i < j\}$ and $\{i' < j'\}$ *cross* if $i < i' < j < j'$ or $i' < i < j' < j$. The *intersection graph* of a chord diagram is the graph with the chords as vertices and edges between pairs of chords that cross. We may orient the graph so that the edge goes from $\{i, j\}$ to $\{i', j'\}$ when $i < i'$; this defines the *directed intersection graph*. A chord diagram is *connected* if its intersection graph is connected.

The Marie–Yeats expansion for the solution of the Dyson–Schwinger equation is a sum over connected chord diagrams involving some slightly unusual statistics. The *intersection order* for the chords of a connected chord diagram C is defined as follows: the root comes first. Delete the root to obtain a (possibly disconnected) chord diagram of size 1 less, order its connected components according to their starting positions from left to right, and then within each component order the chords recursively using the insertion order. Say that a chord is *terminal* if it has outdegree 0 in the directed intersection graph; explicitly this means that it does not cross any chord that starts to the right. (Note that the rightmost chord is always terminal, so there is at least one.)

Let C be a connected chord diagram with k terminal chords and let $t_1 < \cdots < t_k$ be their positions in the intersection order. Then define the *Mellin monomial* of C as

$$\mathrm{mel}(C) = a_0^{|C|-k} \prod_{i=1}^{k-1} a_{t_{k+1}-t_k}$$

in terms of the coefficients of the power series corresponding to the 1-cocycle by (4.9). Also write $b(C) = t_1$. Then the solution to the Dyson–Schwinger equation

$$G(x, L) = 1 + x\Lambda(G(x, L)^{-1})$$

is given by

$$G(x, L) = 1 + \sum_C \operatorname{mel}(C) \sum_{k=1}^{b(C)} a_{b(C)-k} \frac{x^{|C|} L^k}{k!}.$$

One may immediately see a resemblance to the formulas we saw in Chapter 3; indeed, the original motivation for the work on tubings in [8] was precisely to try and connect this chord diagram expansion to Connes–Kreimer tree combinatorics. In the next section we will prove expansions for various Dyson–Schwinger equations in terms of tubings using the results of Chapter 3. These expansions resemble the chord diagram expansions but have fewer terms, apply in greater generality, and have simpler and more conceptual proofs. In [8, Section 5] a bijection is given between tubings of plane trees and connected chord diagrams which preserves the relevant statistics, allowing the chord diagram expansions to be derived from the tubing expansion (or vice versa, in the cases where the former exists).

As alluded to in the previous section, we are interested not only in single equations but also systems. The setup here is the same, but we partition our indexing set P into $\{P_i\}_{i \in I}$ for some finite set I which will index the equations in the system. Each $p \in P$ is still assigned a simple weight $w_p \in \mathbb{N}_+$ but the insertion exponent is now an insertion exponent vector $\mu_p \in \mathbb{K}^I$. We are now solving for a vector of series $\mathbf{G}(x, L) = (G_i(x, L))_{i \in I}$. The system of equations we consider is

$$G_i(x, L) = 1 + \sum_{p \in P_i} x^{w_p} \Lambda_p(\mathbf{G}(x, L)^{\mu_p}). \tag{4.11}$$

The analogue of the special case (4.10) is the existence of a so-called *invariant charge* for the system. This will be further discussed in Section 4.5.

Remark 4.2.2. There is no known generalization of the chord diagram expansion to systems, though one could possibly be derived from our results. The tubing expansions we will prove are the first combinatorial solutions of this type for systems.

4.2.2 Single equations

We will take a two-step approach to solving Dyson–Schwinger equations. First, we will solve the so-called *combinatorial Dyson–Schwinger equation*, in which each 1-cocycle Λ_p is replaced by $B_+^{(p)}$. This gives a series in $T(x) \in \mathcal{H}_P[[x]]$ which encodes the recursive structure of the DSE. We then apply the unique map $\varphi \colon \mathcal{H}_P \to \mathbb{K}[L]$ satisfying $\varphi B_+^{(p)} = \Lambda_p \varphi$ which exists by Theorem 3.1.13 to get $G(x, L) = \varphi(T(x))$ which will solve the Dyson–Schwinger equation. Since we already have a combinatorial expansion for φ (Theorem 3.3.1) this will give us a combinatorial expansion for the solution of the DSE.

The combinatorial version of (4.8) is the equation

$$T(x) = 1 + \sum_{p \in P} x^{w_p} B_+^{(p)}(T(x)^{\mu_p}). \tag{4.12}$$

The solution to the equation is essentially due to Bergbauer and Kreimer [9] although our statement is somewhat more general. To state it we need some more notation. For a vertex v with decoration p, we will write $w(v) = w_p$ and $\mu(v) = \mu_p$. We will write

$$w(t) = \sum_{v \in t} w(v).$$

Finally, by an *automorphism* of a decorated tree we mean an automorphism of the underlying tree (as a poset) which preserves the decorations, and as one would expect we denote the automorphism group of t by $\mathrm{Aut}(t)$.

Proposition 4.2.3 ([9, Lemma 4])**.** *The unique solution to (4.12) is*

$$T(x) = 1 + \sum_{t \in \mathcal{T}(P)} \left(\prod_{v \in t} \mu(v)^{\underline{\mathrm{od}(v)}} \right) \frac{t x^{w(t)}}{|\mathrm{Aut}(t)|}. \tag{4.13}$$

(Recall from Section 2.1 that the underline notation denotes a falling factorial.)

Proof. Let $T(x)$ be given by (4.13); we will show that $T(x)$ satisfies (4.12).

We introduce some notation. For a forest $f \in \mathcal{F}(P)$ write $\kappa(f)$ for the number of connected components. Write $\mathcal{F}_k(P)$ for the set of forests $f \in \mathcal{F}(P)$ with $\kappa(f) = k$. Write $\tilde{T}(x) = T(x) - 1$, so $\tilde{T}(x)$ is a kind of exponential generating function for P-trees. By the compositional formula (see e.g. [56, Theorem 5.5.4]), divided powers count forests:

$$\frac{\tilde{T}(x)^k}{k!} = \sum_{f \in \mathcal{F}_k(P)} \left(\prod_{v \in f} \mu(v)^{\underline{\mathrm{od}(v)}} \right) \frac{f x^{w(f)}}{|\mathrm{Aut}(f)|}.$$

Then by the binomial series expansion, for any $u \in \mathbb{K}$ we have

$$\begin{aligned} T(x)^u &= (1 + \tilde{T}(x))^u \\ &= \sum_{k \geq 0} \binom{u}{k} \tilde{T}(x)^k \\ &= \sum_{f \in \mathcal{F}(P)} u^{\underline{\kappa(f)}} \left(\prod_{v \in f} \mu(v)^{\underline{\mathrm{od}(v)}} \right) \frac{f x^{w(f)}}{|\mathrm{Aut}(f)|}. \end{aligned}$$

Now, any tree $t \in \mathcal{T}(P)$ can be uniquely written as $t = B_+^{(p)} f$ for some $p \in P$ and $f \in \mathcal{F}(P)$. In this case, we have $w(t) = w(f) + w_p$. We have $\mathrm{Aut}(t) \cong \mathrm{Aut}(f)$ and the outdegrees of all non-root vertices are the same in t as in f. The outdegree of the root is $\kappa(f)$. Using this bijection we get

$$\begin{aligned} T(x) &= 1 + \sum_{p \in P} \sum_{f \in \mathcal{F}(P)} \mu_p^{\underline{\kappa(f)}} \left(\prod_{v \in f} \mu(v)^{\underline{\mathrm{od}(v)}} \right) \frac{(B_+^{(p)} f) x^{w(f) + w_p}}{|\mathrm{Aut}(f)|} \\ &= 1 + \sum_{p \in P} x^{w_p} B_+^{(p)} (T(x)^{\mu_p}) \end{aligned}$$

as desired. $\square$

Example 4.2.4. Suppose we have just a single cocycle (so we are essentially in the undecorated Connes–Kreimer Hopf algebra $\mathcal{H}$) and a nonnegative integer insertion exponent k, so (4.13) becomes

$$T(x) = 1 + xB_+(T(x)^k).$$

If we ignore the B_+, this would simply give the ordinary generating function for k-ary trees (in the computer scientists' sense, where the children of each vertex are totally ordered including the "missing" ones). With the B_+ included, it should still be generating function for k-ary trees but now each tree contributes its underlying ordinary rooted tree. It is not too hard to show that $\left(\prod_{v\in t} k^{\underline{\mathrm{od}(v)}}\right)/|\mathrm{Aut}(t)|$ counts the number of ways to make a tree t into a k-ary tree, so this agrees with (4.13).

Of particular note is the case $k = 1$, the *linear* Dyson–Schwinger equation, which produces 1-ary trees i.e. ladders.

Example 4.2.5. Again consider a single cocycle, but now with insertion exponent -1. The equation

$$T(x) = 1 + xB_+(T(x)^{-1})$$

can be rewritten in terms of $\tilde{T}(x) = T(x) - 1$ as

$$\tilde{T}(x) = xB_+\left(\frac{1}{1+\tilde{T}(x)}\right).$$

If the plus sign were replaced by a minus this would give a generating function for plane trees. With the plus, it gives plane trees with a sign corresponding to the number of edges. On the other hand, noting that $(-1)^{\underline{d}} = (-1)^d d!$, we have that the contribution of t to the right side of (4.13) is $(-1)^{|t|-1}\left(\prod_{v\in t}\mathrm{od}(v)!\right)/|\mathrm{Aut}(t)|$ which up to the sign is the number of ways to make t into a plane tree, so this also matches the formula.

Combining Proposition 4.2.3 with our work on binary tubings gives the desired combinatorial expansion for the solution of the Dyson–Schwinger equation (4.8). Note that in [8] only equations of the form (4.10) are explicitly addressed, but on inspection the same proof works for the more general setup. To state the expansion, we use the notation of Chapter 3; in particular, we have

$$a_{p,n} = [z^n]A_p(z)$$

and for a binary tubing τ of a tree $t \in \mathcal{T}(P)$ the Mellin monomial $\mathrm{mel}(\tau)$ is defined by (3.9).

Theorem 4.2.6 ([8, Theorem 2.12])**.** *The unique solution to (4.8) is*

$$G(x,L) = 1 + \sum_{t\in\mathcal{T}(P)}\left(\prod_{v\in t}\mu(v)^{\underline{\mathrm{od}(v)}}\right)\sum_{\tau\in\mathrm{Tub}(t)}\mathrm{mel}(\tau)\sum_{k=1}^{b(\tau)} a_{d(t),b(\tau)-k}\frac{x^{w(t)}L^k}{|\mathrm{Aut}(t)|k!}.$$

In particular, the solution to the special case (4.10) is

$$G(x,L) = 1 + \sum_{t\in\mathcal{T}(\mathbb{N}_+)}\left(\prod_{v\in t}(1+sw(v))^{\underline{\mathrm{od}(v)}}\right)\sum_{\tau\in\mathrm{Tub}(t)}\mathrm{mel}(\tau)\sum_{k=1}^{b(\tau)} a_{w(\mathrm{rt}\,t),b(\tau)-k}\frac{x^{w(t)}L^k}{|\mathrm{Aut}(t)|k!}.$$

Proof. Immediate from Proposition 4.2.3 and Theorem 3.3.1. □

Remark 4.2.7. In Section 4.1 our derivation of the Dyson–Schwinger using Feynman diagrams was already essentially combinatorial in nature, but here we seem to have discarded the Feynman diagram combinatorics and replaced it with trees. The two interpretations can be reconciled by thinking of the trees that appear in these expansions as *insertion trees* which encode the way a Feynman diagram is built from primitive diagrams. One can in principle recover the contribution of an individual Feynman diagram by an appropriately weighted sum over (tubings of) insertion trees for that diagram. (See for instance [37].)

Remark 4.2.8. Recall from Section 3.2 that corollas have factorially many binary tubings. Thus, for generic choices of Mellin coefficients where we do not have cancellation, the coefficients of $G(x, L)$ grow factorially fast regardless of L and the series will have zero radius of convergence. Indeed, a well-known heuristic argument of Dyson [19] suggests that this is exactly what we should expect: if the radius of convergence were nonzero it would imply that the correlation function could be analytically continued to (sufficiently small) negative values of the coupling parameter x, but on physical grounds there simply should not be a well-defined correlation function in that regime. Nonetheless, despite being divergent, these series can (at least in some cases) carry nontrivial information about the actual correlation function which can be extracted via techniques such as Borel summation. (See [40] for this approach in a setup close to ours.)

4.2.3 Systems

The combinatorial version of the system (4.11) is

$$T_i(x) = 1 + \sum_{p \in P_i} x^{w_p} B_+^{(p)}(\mathbf{T}(x)^{\mu_p}). \tag{4.14}$$

As with the single-equation case, we first need to solve this combinatorial system. Unlike the single-equation case, this is due to the author and first appears in [8], though it is a straightforward generalization of Proposition 4.2.3. Again, [8] only considers a special case but the proof works in our more general setup.[5] We need some additional notation: let us write $\mathcal{T}_i(P)$ for the subset of $\mathcal{T}(P)$ for which the root has a decoration in P_i; clearly these are the trees that can contribute to $T_i(x)$. For a vertex v let $\mathrm{od}_i(v)$ be the number of children which have their decoration in P_i; we collect all of these together to form the *outdegree vector* $\mathbf{od}(v) \in \mathbb{N}^I$.

Theorem 4.2.9 ([8, Theorem 4.7])**.** *The unique solution to the system (4.14) is*

$$T_i(x) = 1 + \sum_{t \in \mathcal{T}_i(P)} \left(\prod_{v \in t} \mu(v)^{\underline{\mathbf{od}(v)}} \right) \frac{t x^{w(t)}}{|\mathrm{Aut}(t)|}.$$

Proof. Analogous to Proposition 4.2.3. □

[5]Indeed, allowing the insertion exponents to be arbitrary as we do allows the formula to be much cleaner.

Again, applying Theorem 3.3.1 immediately gives us a solution to the original system.

Theorem 4.2.10 ([8, Theorem 4.8]). *The unique solution to the system (4.11) is*

$$G_i(x, L) = 1 + \sum_{t \in \mathcal{T}_i(P)} \left(\prod_{v \in t} \mu(v)^{\underline{\mathbf{od}(v)}} \right) \sum_{\tau \in \mathrm{Tub}(t)} \mathrm{mel}(\tau) \sum_{k=1}^{b(\tau)} a_{d(t), b(\tau)-k} \frac{x^{w(t)} L^k}{|\mathrm{Aut}(t)|}.$$

4.3 Distinguishing insertion places

In this section we will introduce a generalization of the Dyson–Schwinger framework of the previous section. To motivate this, let us briefly recall one of the more poorly motivated parts of Section 4.1, namely the equation (4.3). This was justified by the idea that a Feynman diagram can be recursively built starting with a primitive diagram and repeatedly inserting into edges. However, there is an implicit assumption that the Feynman rules do not depend in a serious way on *which* edge one inserts into. This is not completely unreasonable (see [64, Section 2.3.3]) but is still a deficiency in the framework. To remedy this, we would like to allow equations like

$$G(x, L) = 1 + xG(x, \partial/\partial\rho_1)^{-1} \cdots G(x, \partial/\partial\rho_m)^{-1} (e^{L(\rho_1 + \cdots + \rho_m)} - 1) F(\rho_1, \ldots, \rho_m)\big|_{\rho=0} \quad (4.15)$$

where each edge has its own variable in the Mellin transform. Such equations have been considered by Yeats [64] and Nabergall [45, Section 4.2] but significantly less is known about them than the equations we have looked at thus far. None of the results of this section have previously appeared.

The approach we will take to studying this generalization should be no surprise: this is the purpose for which our results in Chapter 3 about 1-cocycles on tensor products were brought into the world.

4.3.1 Setup

Our setup is similar to the one in Section 4.2.1 with some extra details similar to what we saw in Section 3.4. We again have a set P which will index our cocycles, but to each $p \in P$ we associate a finite set E_p of *insertion places*. Each insertion place e has its own insertion exponent μ_e; sometimes it will still be convenient to refer to the *overall insertion exponent*

$$\mu_p = \sum_{e \in E_p} \mu_e.$$

Finally, to each p we associate a vector of indeterminates $\mathbf{L}_p = (L_e)_{e \in E_p}$ and a 1-cocycle $\Lambda_p \in \mathrm{Z}^1(\mathbb{K}[\mathbf{L}_p], \mathbb{K}[L])$. The Dyson–Schwinger associated to these data is

$$G(x, L) = 1 + \sum_{p \in P} x^{w_p} \Lambda_p \left(\prod_{e \in E_p} G(x, L_e)^{\mu_e} \right). \quad (4.16)$$

We will also consider systems. As in Section 4.2.1 we partition our index set P into $\{P_i\}_{i\in I}$ and replace the insertion exponents with insertion exponent vectors. Our system is then

$$G_i(x,L) = 1 + \sum_{p\in P_i} x^{w_p}\Lambda_p\left(\prod_{e\in E_p}\mathbf{G}(x,L_e)^{\mu_e}\right). \tag{4.17}$$

4.3.2 Combinatorial version and tubing expansion

We follow the same strategy as in Section 4.2, first lifting (4.16) and (4.17) to combinatorial versions on the Hopf algebra $\widetilde{\mathcal{H}}_{P,\mathcal{E}}$ we introduced in Section 3.4, and then applying Theorem 3.5.1 to get a solution to the original equations. This time we will work in the full generality of systems from the start. The combinatorial version of the system (4.17) is

$$T_i(x) = 1 + \sum_{p\in P_i} x^{w_p}\tilde{B}_+^{(p)}\left(\bigotimes_{e\in E_p}\mathbf{T}(x)^{\mu_e}\right). \tag{4.18}$$

We are slightly abusing notation here by neglecting to notate the obvious (but non-injective!) map $\widetilde{\mathcal{H}}_{P,\mathcal{E}}[[x]]^{\otimes E_p} \to \widetilde{\mathcal{H}}_{P,\mathcal{E}}^{\otimes E_p}[[x]]$. In effect we want to treat x as though it were a scalar, in line with our policy of always applying operators coefficientwise. With this pedantry out of the way we move on to solving the system. The formula generalizes that of Theorem 4.2.9. Each vertex will now have several outdegree vectors, one for each edge type: we write $\mathrm{od}_i(v,e)$ for the number of children of v which have decorations lying in P_i and such that the edge connecting them has decoration e. These are collected together into the outdegree vector $\mathbf{od}(v,e)\in\mathbb{N}^I$.

Theorem 4.3.1. *The unique solution to (4.18) is*

$$T_i(x) = 1 + \sum_{t\in\mathcal{T}(P_i)}\left(\prod_{v\in t}\prod_{e\in E_{d(v)}}\mu_e^{\mathbf{od}(v,e)}\right)\frac{tx^{w(t)}}{|\mathrm{Aut}(t)|}.$$

Proof. Analogous to Proposition 4.2.3. □

As before, we can get a combinatorial expansion for the solution of the Dyson–Schwinger equation by applying our results on 1-cocycles.

Theorem 4.3.2. *The unique solution to (4.17) is*

$$G_i(x,L) = 1 + \sum_{t\in\mathcal{T}(P_i)}\left(\prod_{v\in t}\prod_{e\in E_{d(v)}}\mu_e^{\mathbf{od}(v,e)}\right)\sum_{\tau\in\mathrm{Tub}(t)}\mathrm{mel}(\tau)\sum_{k=1}^{b(\tau)} a_{d(t),\beta^k(\tau)}\frac{x^{w(t)}L^k}{|\mathrm{Aut}(t)|k!}.$$

Proof. Immediate from Theorem 4.3.1 and Theorem 3.5.1. □

While Theorem 4.3.2 is really the main result of this chapter, we are also interested in other properties of Dyson–Schwinger equations with distinguished insertion places. For ordinary Dyson–Schwinger equations it is known that in nice cases the solutions also satisfy a *renormalization group equation*. We will generalize this to Dyson–Schwinger equations with distinguished insertion places, but first we will explain what these equations are and their significance from a Hopf-algebraic perspective.

4.4 Interlude: The renormalization group equation and the Riordan group

Let $\beta(x)$ and $\gamma(x)$ be formal power series, with $\beta(0) = 0$. The *renormalization group equation* (or *Callan–Symanzik equation*) is

$$\left(\frac{\partial}{\partial L} - \beta(x)\frac{\partial}{\partial x} - \gamma(x)\right)G(x, L) = 0. \tag{4.19}$$

As suggested by the notation, we will ultimately want to think of this $G(x, L)$ as the same one which appears in the Dyson–Schwinger equation, but for the purposes of this section we can consider it to be simply notation for the (potential) solution to this PDE. The goal of this section is to explain how (4.19) is intimately related to a certain Hopf algebra. As a starting point, notice that if $\gamma(x) = 0$ we have already seen this equation: by Theorem 2.4.12 it describes a bialgebra morphism $\mathsf{FdB} \to \mathbb{K}[L]$. We will show that something similar is true for (4.19). Let us make clear at the outset that more or less everything in this section is already known in one form or another, but perhaps not as well-known as it should be.

Recall from Section 2.4.3 that $\widetilde{\mathfrak{D}}$ denotes the group of formal power series with zero constant term and nonzero linear term under composition and $\mathfrak{D}$ the subgroup of δ-series, and that $\mathfrak{D}^{\mathrm{op}}$ is isomorphic to the character group of FdB. Now observe that for $\Phi(x) \in \widetilde{\mathfrak{D}}$ the map $F(x) \mapsto F(\Phi(x))$ is a ring automorphism of $\mathbb{K}[[x]]$. Moreover, composing these automorphisms corresponds to composing the series in reverse, so $\widetilde{\mathfrak{D}}^{\mathrm{op}}$ (and hence also $\mathfrak{D}^{\mathrm{op}}$) acts by automorphisms on $\mathbb{K}[[x]]$. Consequently they also act on $\mathbb{K}[[x]]^\times$, the multiplicative group of power series with nonzero constant term. Let $\mathbb{K}[[x]]^\times_1$ be the subgroup of $\mathbb{K}[[x]]^\times$ consisting of those series with constant term 1. The *Riordan group* is the semidirect product $\mathfrak{R} = \mathbb{K}[[x]]^\times_1 \rtimes \mathfrak{D}$. Explicitly, the elements consist of pairs $(F(x), \Phi(x))$ of series with $F(x) \in \mathbb{K}[[x]]^\times_1$ and $\Phi(x) \in \mathfrak{D}$, with the operation

$$(F(x), \Phi(x)) * (G(x), \Psi(x)) = (F(x)G(\Phi(x)), \Psi(\Phi(x))).$$

Remark 4.4.1. The Riordan group was first introduced—at least under that name—by Shapiro, Getu, Woan, and Woodson [55]. It is usually thought of as a group of infinite matrices, via the correspondence

$$(F(x), \Phi(x)) \mapsto \Big[[x^i]F(x)G(x)^j\Big]_{i,j\in\mathbb{N}}$$

sending a pair of series to their *Riordan matrix*. (This is simply a matrix representation of the natural action of $\mathfrak{R}$ on $\mathbb{K}[[x]]$.) Conventions vary on whether or not to include the restrictions on coefficients; our choice matches the original definition in [55] as well as being convenient for relating $\mathfrak{R}$ to a Hopf algebra.

We now wish to define a Hopf algebra with $\mathfrak{R}$ as its character group, similar to the Faà di Bruno Hopf algebra. We will call it the *Riordan Hopf algebra* and denote it by Rio. As an algebra, Rio is a free commutative algebra in two sets of generators $\{\pi_1, \pi_2, \dots\}$ and $\{y_1, y_2, \dots\}$. The π's will generate a copy of FdB; in particular, their coproduct is still given by (2.19). (This inclusion $\mathsf{FdB} \to \mathsf{Rio}$ is dual to the quotient map $\mathfrak{R} \to \mathfrak{D}^{\mathrm{op}}$ coming from the semidirect product.) We assemble the y's into a power series as well, this time in the more obvious way:

$$Y(x) = 1 + \sum_{n\geq 1} y_n x^n.$$

Then the coproduct is given by

$$\Delta y_n = \sum_{j=0}^{n} [x^n] Y(x) \Pi(x)^j \otimes y_j. \tag{4.20}$$

Analogously to Proposition 2.4.8, we easily get the following result.

Proposition 4.4.2. *Let A be a commutative algebra and $\varphi, \psi \colon \mathsf{Rio} \to A$ be algebra morphisms. Let $F(x) = \varphi(Y(x))$, $\Phi(x) = \varphi(\Pi(x))$, $G(x) = \psi(Y(x))$, and $\Psi(x) = \psi(\Pi(x))$. Then*

$$(\varphi * \psi)(Y(x)) = F(x)G(\Phi(x))$$

and

$$(\varphi * \psi)(\Pi(x)) = \Psi(\Phi(x)).$$

Consequently, $\mathrm{Ch}(\mathsf{Rio}) \cong \mathfrak{R}$.

We also have an analogue of Proposition 2.4.11. Note that since the π's generate a copy of FdB we can simply apply Proposition 2.4.11 itself to see how elements of the dual act on them. Thus we only need the actions on $Y(x)$.

Proposition 4.4.3. *Suppose $\varphi \in \mathsf{Rio}^*$ and let $F(x) = \varphi(Y(x))$ and $\Phi(x) = \varphi(\Pi(x))$. Then*

(i) $\varphi \rightharpoonup Y(x) = F(\Pi(x))Y(x)$.

(ii) If $\varphi \in \mathrm{Ch}(\mathsf{FdB})$ *then* $Y(x) \leftharpoonup \varphi = F(x)Y(\Phi(x))$.

(iii) If $\varphi \in \mathfrak{ch}(\mathsf{FdB})$ *then* $\Pi(x) \leftharpoonup \varphi = \Phi(x)Y'(x) + F(x)Y(x)$.

Finally we reach the main result of this section.

Theorem 4.4.4. *Let $\varphi \colon \mathsf{Rio} \to \mathbb{K}[z]$ be an algebra morphism and let $F(x,z) = \varphi(Y(x))$ and $\Phi(x,z) = \varphi(\Pi(x))$. Let $\beta(x)$ be the linear term in z of $\Phi(x,z)$ and $\gamma(x)$ the linear term in z of $F(x,z)$. Suppose φ is a bialgebra morphism when restricted to the subalgebra FdB. Then φ is a bialgebra morphism on all of Rio if and only if $F(x,z)$ satisfies the renormalization group equation*

$$\left(\frac{\partial}{\partial z} - \beta(x)\frac{\partial}{\partial x} - \gamma(x)\right) F(x,z) = 0. \tag{4.21}$$

Proof. By the same argument as Theorem 2.4.12 it is necessary and sufficient to have $\frac{d}{dz}\varphi = (\operatorname{lin}\varphi) * \varphi$. Applying Proposition 4.4.3(iii) gives the result. $\square$

Obviously, if we assume that φ is merely an algebra map $\mathsf{Rio} \to \mathbb{K}[z]$ then it is a bialgebra morphism if and only if it satisfies the conditions of both Theorem 2.4.12 and Theorem 4.4.4.

Remark 4.4.5. A result equivalent to Theorem 4.4.4 was proved by Bacher [6, Proposition 7.1]. He does not take a Hopf algebra perspective but instead essentially works with the Lie algebra $\mathfrak{ch}(\mathsf{Rio})$ in a matrix representation and for an element $\sigma \in \mathfrak{ch}(\mathsf{Rio})$ corresponding to the pair $(\gamma(x), \beta(x))$ characterizes $\exp_*(z\sigma)$ as (the Riordan matrix of) the solution to (4.21) and (2.21), which is equivalent to our result by Theorem 2.2.24. That the PDE in question is in fact the renormalization group equation seems not to have been noticed.

4.5 The invariant charge

Now we relate the previous section to Dyson–Schwinger equations. As alluded to, it is known that the solution to a genuine physical Dyson–Schwinger equation or system should satisfy a renormalization group equation. This is not true in the complete generality of the framework we have set up but is known to be true for ordinary Dyson–Schwinger equations under the additional hypothesis of the existence of a so-called *invariant charge*. In this section we will define what this means and give a proof using the ideas of the previous section, then generalize to the case of distinguished insertion places.

We will begin with the simplest case (4.8). By Theorem 4.4.4, we see that $G(x, L)$ satisfies a renormalization group equation if there exists a bialgebra morphism $\mathsf{Rio} \to \mathbb{K}[[L]]$ that sends $Y(x)$ to $G(x, L)$. It is natural to lift to the combinatorial equation (4.12) and ask instead for a bialgebra morphism $\mathsf{Rio} \to \mathcal{H}_P$ that sends $Y(x)$ to $T(x)$. The question then is where $\Pi(x)$ should be mapped. We wish to construct from $T(x)$ an auxiliary series $Q(x) \in \mathcal{H}_P[[x]]$—the invariant charge—such that the map sending $\Pi(x)$ to $Q(x)$ and $Y(x)$ to $T(x)$ is a bialgebra morphism. Note that this is unique if it exists since the coproduct formula (4.20) allows us to recover it from the coproducts of coefficients of $T(x)$. It turns out that the case in which we can ensure this exists is exactly the special case (4.10).[6]

Proposition 4.5.1. *Let $T(x) \in \mathcal{H}_{\mathbb{N}_+}[[x]]$ be the solution of the combinatorial Dyson–Schwinger equation*

$$T(x) = 1 + \sum_{k \geq 1} x^k B_+^{(k)}(T(x)^{1+sk}).$$

Then the algebra morphism $\varphi\colon \mathsf{Rio} \to \mathcal{H}_{\mathbb{N}_+}$ defined by $\varphi(Y(x)) = T(x)$ and $\varphi(\Pi(x)) = xT(x)^s$ is a bialgebra morphism. As a consequence, the solution $G(x, L)$ to the corresponding Dyson–Schwinger equation (4.10) satisfies the renormalization group equation

$$\left(\frac{\partial}{\partial L} - sx\gamma(x)\frac{\partial}{\partial x} - \gamma(x)\right) G(x, L) = 0$$

where $\gamma(x)$ is the linear term in L of $G(x, L)$.

Remark 4.5.2. While the *phrasing* of Proposition 4.5.1 seems to be new, its content is not: the coproduct formula for $T(x)$ implied by combining this result with (4.20) is well-known. (See [64, Lemma 4.6] for exactly this formula and for instance [12, Theorem 1],

[6] We will only show sufficiency; for necessity see [23, Proposition 10] although note that the setup there is somewhat different from ours.

[49, Proposition 4.2], and [59, Proposition 7] for essentially equivalent formulas appearing in slightly different contexts.) Our proof will also be isomorphic to the proof in [64], but presented in what we hope is a more conceptually clear way. The genuinely new result of this section is the generalization to distinguished insertion places.

We now work towards proving Proposition 4.5.1. As in Section 4.3.2 we will find it convenient to abuse notation by writing tensor products of power series when we really mean power series with tensor coefficients. With this in mind, we can rewrite (2.19) and (4.20) simply as

$$\Delta\Pi(x) = \sum_{j\geq 0} \Pi(x)^{j+1} \otimes \pi_j$$

and

$$\Delta Y(x) = \sum_{j\geq 0} Y(x)\Pi(x)^j \otimes y_j.$$

Our first lemma is a common generalization of both formulas.

Lemma 4.5.3. *For any $s \in \mathbb{K}$ and $k \in \mathbb{N}$,*

$$\Delta\big(Y(x)^s\Pi(x)^k\big) = \sum_{j\geq 0} Y(x)^s\Pi(x)^j \otimes [x^j]Y(x)^s\Pi(x)^k.$$

(Note that since $\Pi(x)$ has zero constant term, we can only raise it to natural powers if we want to stay in the realm of power series.)

Proof. Both sides are power series with coefficients that are polynomials in s, so it is sufficient to prove the case $s \in \mathbb{N}$. Then by the coproduct formulas we can write

$$\begin{aligned}\Delta\big(Y(x)^s\Pi(x)^k\big) &= \sum_{i_1,\dots,i_s,j_1,\dots,j_k} Y(x)^s\Pi(x)^{i_1+\cdots+i_s+j_1+\cdots+j_k+k} \otimes y_{i_1}\cdots y_{i_s}\pi_{j_1}\cdots\pi_{j_k} \\ &= \sum_{j\geq 0} Y(x)^s\Pi(x)^j \otimes \sum_{i_1+\cdots+i_s+j_1+\cdots+j_k+k=j} y_{i_1}\cdots y_{i_s}\pi_{j_1}\cdots\pi_{j_k} \\ &= \sum_{j\geq 0} Y(x)^s\Pi(x)^j \otimes [x^j]Y(x)^s\Pi(x)^k\end{aligned}$$

as desired. □

For $n \geq 0$, let $\mathsf{FdB}^{(n)}$ denote the subalgebra of FdB generated by $\pi_1, \dots, \pi_{n-1}$ (this should not be confused with the graded piece FdB_n) and let $\mathsf{Rio}^{(n)}$ denote the subalgebra of Rio generated by $\pi_1, \dots, \pi_{n-1}$ and $y_1, \dots, y_n$. From the coproduct formulas it is clear that are in fact sub-bialgebras. The following result is new as stated but encapsulates the main calculation used in standard proofs of Proposition 4.5.1.

Lemma 4.5.4. *Suppose H is a bialgebra, $\varphi\colon \mathsf{Rio} \to H$ is an algebra morphism, and $\{\Lambda_k\}_{k\in\mathbb{N}_+}$ is a family of 1-cocycles on H. Let $\Phi(x) = \varphi(\Pi(x))$ and suppose $\varphi(Y(x)) = F(x)$ where $F(x)$ is the unique solution to*

$$F(x) = 1 + \sum_{k\geq 1} \Lambda_k(F(x)\Phi(x)^k). \tag{4.22}$$

Then for $n \geq 0$, if φ is a bialgebra morphism when restricted to $\mathsf{FdB}^{(n)}$, it is also a bialgebra morphism when restricted to $\mathsf{Rio}^{(n)}$.

Recall that by definition $\Pi(x)$ and hence also $\Phi(x)$ has zero constant term, so only the terms with $k \leq n$ on the right side of (4.22) can contribute to the coefficient of x^n. Thus the equation really does have a unique solution.

Proof. Since we are given that φ is an algebra morphism we must only prove it preserves the coproducts of the generators. We prove this by induction on n. In the base case, $\mathsf{FdB}^{(0)} = \mathsf{Rio}^{(0)} = \mathbb{K}$ so there is nothing to prove. Now suppose that $n > 0$ and that φ is a bialgebra morphism when restricted to $\mathsf{Rio}^{(n-1)}$ and also preserves the coproduct of π_{n-1}. Then we must show it preserves the coproduct of y_n. Note that when $k > 1$, the coefficient $[x^n]Y(x)\Pi(x)^k$ does not contain y_n, so its coproduct agrees with the formula from Lemma 4.5.3. Thus

$$
\begin{aligned}
\Delta\varphi(y_n) &= \Delta([x^n]F(x)) \\
&= \Delta\left(\sum_{k\geq 1}\Lambda_k\big([x^n]F(x)\Phi(x)^k\big)\right) \\
&= (\Lambda_k \otimes 1 + (\mathrm{id}\otimes\Lambda_k)\Delta)\left(\sum_{k\geq 1}[x^n]F(x)\Phi(x)^k\right) \\
&= [x^n]F(x)\otimes 1 + \sum_{k\geq 1}\sum_{j=0}^{n-k}[x^n]F(x)\Phi(x)^j \otimes [x^j]\Lambda_k(F(x)\Phi(x)^k) \\
&= [x^n]F(x)\otimes 1 + \sum_{j=0}^{n-1}[x^n]F(x)\Phi(x)^j \otimes [x^j]\left(\sum_{k\geq 1}\Lambda_k(F(x)\Phi(x)^k)\right) \\
&= \sum_{j=0}^{n}[x^n]F(x)\Phi(x)^j \otimes [x^j]F(x) \\
&= (\varphi\otimes\varphi)(\Delta y_n).
\end{aligned}
$$

□

Remark 4.5.5. An obvious consequence of Lemma 4.5.4 is that if φ is already known to be a bialgebra morphism when restricted to FdB then it is a bialgebra morphism on all of Rio. This is not quite the right version of the statement for the application to Dyson–Schwinger equations, but it does give some interesting examples of series satisfying renormalization group equations. For instance, consider the map $\mathsf{FdB} \to \mathcal{H}$ given by $\pi_n \mapsto \ell_1^n$. (Recall that ℓ_n is the n-vertex ladder so in particular ℓ_1 is the unique one-vertex tree.) It is a straightforward exercise to show that this is in fact a bialgebra morphism. Thus we can extend this map to Rio by sending $Y(x)$ to the series $T(x)$ defined by

$$
T(x) = 1 + xB_+\left(\frac{T(x)}{1-\ell_1 x}\right),
$$

an example due to Dugan [18] of a series not coming from a Dyson–Schwinger equation which nonetheless satisfies a renormalization group equation after applying a bialgebra morphism $\mathcal{H} \to \mathbb{K}[L]$. In the spirit of Examples 4.2.4 and 4.2.5, we can think of $T(x)$ as a generating function for plane trees with the property that one obtains a ladder after deleting all leaves.

We can now prove Proposition 4.5.1.

Proof of Proposition 4.5.1. We prove by induction on n that φ is a bialgebra morphism on $\mathsf{Rio}^{(n)}$. For $n = 0$ this is trivial. Now suppose $n > 0$ and that φ is a bialgebra morphism on $\mathsf{Rio}^{(n-1)}$. In particular, φ is a bialgebra morphism on $\mathsf{FdB}^{(n-1)}$, and we observe that since $[x^{n-1}]T(x)^s \in \varphi(\mathsf{Rio}^{(n-1)})$, by Lemma 4.5.3 we have

$$\begin{aligned}
\Delta\varphi(\pi_{n-1}) &= \Delta([x^{n-1}]T(x)^s) \\
&= \sum_{j\geq 0} [x^{n-1}]T(x)^s(xT(x)^s)^j \otimes [x^j]T(x)^s \\
&= \sum_{j\geq 0} [x^{n-1}](xT(x)^s)^{j+1} \otimes [x^j]T(x)^s \\
&= \sum_{j\geq 0} [x^{n-1}]\varphi(\Pi(x))^{j+1} \otimes \varphi(\pi_j) \\
&= (\varphi\otimes\varphi)(\Delta\pi_{n-1})
\end{aligned}$$

so φ is a bialgebra morphism on $\mathsf{FdB}^{(n)}$. By Lemma 4.5.4, φ is thus a bialgebra morphism on $\mathsf{Rio}^{(n)}$ as wanted. The renormalization group equation then follows from Theorem 4.4.4. $\square$

Now we consider systems. The idea is the same, that we would like to write each equation of the system in a form that looks like (4.22). In general this will only work if we have the same invariant charge for each equation. In terms of the setup in Section 4.2.1, for $p \in P_i$ we want a linear relation

$$\mu_p = 1_i + w_p\mathbf{s}$$

for some $\mathbf{s} = (s_i)_{i\in I} \in \mathbb{K}^I$. As in the single-equation case, we may as well combine terms together to write the system in the form

$$G_i(x,L) = 1 + \sum_{k\geq 1} x^k \Lambda_{i,k}\left(G_i(x,L)\prod_j G_j(x,L)^{s_j k}\right). \tag{4.23}$$

The corresponding combinatorial system then looks like

$$T_i(x) = 1 + \sum_{k\geq 1} B_+^{(i,k)}(T_i(x)Q(x)^k) \tag{4.24}$$

where

$$Q(x) = x\prod_{i\in I} T_i(x)^{s_i}. \tag{4.25}$$

We then have the following generalization of Proposition 4.5.1. (Note that most of the papers referenced in Remark 4.5.2 are actually for this version already.)

Theorem 4.5.6. *Let $\mathbf{T}(x) \in \mathcal{H}_{I\times\mathbb{N}_+}[[x]]^I$ be the solution to the combinatorial Dyson–Schwinger system (4.24). Then for any $i \in I$, the map $\varphi_i \colon \mathsf{Rio} \to \mathcal{H}_{I\times\mathbb{N}^+}$ defined by $\varphi_i(Y(x)) = T_i(x)$ and $\varphi_i(\Pi(x)) = Q(x)$ is a bialgebra morphism. As a consequence, the solution $\mathbf{G}(x, L)$ to the corresponding Dyson–Schwinger system (4.23) satisfies the renormalization group equations*

$$\left(\frac{\partial}{\partial L} - \beta(x)\frac{\partial}{\partial x} - \gamma_i(x)\right)G_i(x, L) = 0$$

where $\gamma_i(x)$ is the linear term in L of $G_i(x, L)$ and

$$\beta(x) = \sum_{i\in I} s_i x\gamma(x).$$

Proof. We prove by induction on n that φ_i is a bialgebra morphism on $\mathsf{Rio}^{(n)}$ for all i. Supposing they are all bialgebra morphisms on $\mathsf{Rio}^{(n-1)}$. Then, as in the proof of Proposition 4.5.1, we have

$$\begin{aligned}
\Delta\varphi_i(\pi_{n-1}) &= \Delta([x^{n-1}]Q(x)) \\
&= [x^{n-1}]\prod_{i\in I}\Delta(T_i(x)^{s_i}) \\
&= [x^{n-1}]\sum_{\alpha\in\mathbb{N}^I}\prod_{i\in I}(T_i(x)^{s_i}Q(x)^{\alpha_i}\otimes[x^{\alpha_i}]T_i(x)^{s_i}) \\
&= \sum_{\alpha\in\mathbb{N}^I}[x^n]Q(x)^{|\alpha|+1}\otimes\prod_{i\in I}[x^{\alpha_i}]T_i(x)^{s_i} \\
&= \sum_{j\geq 0}[x^n]Q(x)^{j+1}\otimes[x^{j+1}]Q(x) \\
&= (\varphi\otimes\varphi)(\Delta\pi_{n-1})
\end{aligned}$$

and thus φ_i is a bialgebra morphism on $\mathsf{FdB}^{(n)}$ and hence on $\mathsf{Rio}^{(n)}$ by Lemma 4.5.4. The renormalization group equation then follows from Theorem 4.4.4. $\square$

We can similarly generalize to the case of distinguished insertion places, i.e. systems of the form (4.18). Naturally, this will come from a generalization of Lemma 4.5.4 to allow cocycles on tensor powers of the target bialgebra H. In this case the statement gets more complicated but the proof is much the same. First we need a generalization of Lemma 4.5.3.

Lemma 4.5.7. *Let δ be the left coaction of Rio on $\mathsf{Rio}^{\otimes E}$ for E a finite set. Then for any $\mathbf{u} \in \mathbb{K}^r$ and any exponent vector $\alpha \in \mathbb{N}^E$,*

$$\delta\left(\bigotimes_{e\in E} Y(x)^{u_e}\Pi(x)^{\alpha_e}\right) = \sum_{j\geq 0} Y(x)^{|\mathbf{u}|}\Pi(x)^j \otimes [x^j]\bigotimes_{e\in E} Y(x)^{u_e}\Pi(x)^{\alpha_e}.$$

Proof. Immediate from Lemma 4.5.3 by the definition of the coaction. $\square$

With this we can prove the following.

Lemma 4.5.8. *Let H be a bialgebra P be a set, and $\{E_p\}_{p\in P}$ a family of finite sets. For $p \in P$ let $\Lambda_p\colon H^{\otimes E_p} \to H$ be a 1-cocycle, let $\mathbf{u}_p = (u_e)_{e\in E_p}$ be a vector with $|\mathbf{u}| = 1$, and let $\mathbf{w}_p = (w_e)_{e\in E_p}$ be a nonzero exponent vector. Suppose $\varphi\colon \mathsf{Rio} \to H$ is an algebra morphism. Let $\Phi(x) = \varphi(\Pi(x))$ and suppose $\varphi(Y(x)) = F(x)$ where $F(x)$ is the unique solution to*

$$F(x) = 1 + \sum_{p\in P} \Lambda_p\left(\bigotimes_{e\in E_p} F(x)^{u_e}\Phi(x)^{\alpha_e}\right).$$

Then for $n \geq 0$, if φ is a bialgebra morphism when restricted to $\mathsf{FdB}^{(n)}$, it is also a bialgebra morphism when restricted to $\mathsf{Rio}^{(n)}$.

Proof. The setup for our induction argument is identical to that in the proof of Lemma 4.5.4. Thus, supposing that φ is a bialgebra morphism on $\mathsf{Rio}^{(n-1)}$ for some $n \geq 1$, we set out to show that φ preserves the coproduct of y_n. Note that we have no y_n in $[x^n]\bigotimes_{e\in E_p} Y(x)^{u_e}\Pi(x)^{w_e}$ since $|\mathbf{w}_p| > 0$. Thus by Lemma 4.5.7 we have

$$\begin{aligned}
\Delta\varphi(y_n) &= \Delta([x^n]F(x)) \\
&= [x^n]F(x) \otimes 1 + \sum_{p\in P}(\mathrm{id}\otimes\Lambda_p)[x^n]\delta\left(\bigotimes_{e\in E_p} F(x)^{u_e}\Phi(x)^{w_e}\right) \\
&= [x^n]F(x) \otimes 1 + \sum_{p\in P}\sum_{j\geq 0}[x^n]F(x)\Phi(x)^j \otimes \Lambda_p\left([x^j]\bigotimes_{e\in E_p} F(x)^{u_e}\Phi(x)^{w_e}\right) \\
&= \sum_{j\geq 0}[x^n]F(x)\Phi(x)^j \otimes [x^j]F(x)
\end{aligned}$$

as desired. □

With Lemma 4.5.8 in mind we can formulate an appropriate notion of invariant charge for Dyson–Schwinger equations with distinguished insertion places. We want a series $Q(x)$ to play the role of $\Phi(x)$ in the statement of the lemma. In the case of a single equation (4.16) we see that the condition we want is that for some $s \in \mathbb{K}$, the insertion exponents can be written in the form

$$\mu_e = u_e + s w_e \tag{4.26}$$

where $u_e \in \mathbb{K}$ and $\alpha_e \in \mathbb{N}$ are such that

$$\sum_{e\in E_p} u_e = 1 \tag{4.27}$$

and

$$\sum_{e\in E_p} w_e = w_p \tag{4.28}$$

for each p. In this case we take $Q(x) = xT(x)^s$ as in the case of a single ordinary DSE. We can then write our (combinatorial) equation as

$$T(x) = \sum_{p\in P} \tilde{B}_+\left(\bigotimes_{e\in E_p} T(x)^{u_e}Q(x)^{\alpha_e}\right). \tag{4.29}$$

Remark 4.5.9. The choice of exactly how to write the insertion exponents in the form (4.26) is not unique in general. However the value of s and hence of $Q(x)$ does not depend on this choice, because the *overall* insertion exponents satisfy $\mu_p = 1 + sw_p$ regardless.

Recall our original motivating example was equation (4.15) which has m edge types and all insertion exponents equal to -1. Thus we here have $s = -(1+m)$. We can write it in the form (4.26) by choosing one edge type e_0 to have $u_{e_0} = m$ and $w_{e_0} = 1$, and all others to have $u_e = -1$ and $w_e = 0$. Thus while writing this equations in the form (4.29) is convenient for our purposes in this section, it does involve arbitrarily breaking the symmetry of the original equation.

For systems the situation is similar. For $e \in E_p$ where $p \in P_i$, we want the insertion exponent vector to satisfy a relation

$$\mu_e = u_e 1_i + \alpha_e \mathbf{s} \tag{4.30}$$

where u_e and α_e still satisfy (4.27) and (4.28). Thus the form of our system is

$$T_i(x) = \sum_{p \in P_i} \tilde{B}_+^{(p)} \left(\bigotimes_{e \in E_p} T_i(x)^{u_e} Q(x)^{\alpha_e} \right). \tag{4.31}$$

where as before

$$Q(x) = x \prod_{i \in I} T_i(x)^{s_i}.$$

With the setup done, we can state the main result of this section. This proves a conjecture of Nabergall [45, Conjecture 4.2.3] which corresponds to the case all insertion exponents equal -1.

Theorem 4.5.10. *Let $\mathbf{T}(x) \in \widetilde{\mathcal{H}}_{P,\mathcal{E}}[[x]]^I$ be the solution to the combinatorial Dyson–Schwinger system (4.31). Then for any $i \in I$, the map $\varphi_i \colon \mathsf{Rio} \to \widetilde{\mathcal{H}}_{P,\mathcal{E}}$ defined by $\varphi_i(Y(x)) = T_i(x)$ and $\varphi_i(\Pi(x)) = Q(x)$ is a bialgebra morphism. As a consequence, the solution $\mathbf{G}(x, L)$ to the corresponding Dyson–Schwinger system satisfies the renormalization group equations*

$$\left(\frac{\partial}{\partial L} - \beta(x) \frac{\partial}{\partial x} - \gamma_i(x) \right) G_i(x, L) = 0$$

where $\gamma_i(x)$ is the linear term in L of $G_i(x, L)$ and

$$\beta(x) = \sum_{i \in I} s_i x \gamma(x).$$

Proof. This follows by an identical proof to Theorem 4.5.6 but using Lemma 4.5.8 in place of Lemma 4.5.4. ∎

This completes for now our exploration of Dyson–Schwinger equations. In the next chapter we will shift gears entirely and enter the world of symmetric functions.

Chapter 5

Skew Equivalence

5.1 Shapes

In this section we review the necessary facts about skew shapes and skew Schur functions. While there are no new results, the presentation is somewhat nonstandard and not all of the material is well-known, so the author humbly suggests that even those readers who are experts on symmetric functions may not wish to skip it entirely.

5.1.1 Skew shapes and the shape Hopf algebra

For our purposes, a *Ferrers shape* is a finite downset in $\mathbb{N}_+ \times \mathbb{N}_+$. Ordered by inclusion, these form a distributive lattice $\mathbb{Y}$, known as *Young's lattice*. Ferrers shapes are in bijection with integer partitions, with the correspondence given by

$$\lambda \leftrightarrow \{(i,j) \in \mathbb{N}_+ \times \mathbb{N}_+ \colon j \leq \lambda_i\}.$$

Moreover this bijection is easily seen to be order-preserving with respect to the componentwise ordering on partitions (that is, the ordering induced by the inclusion $\mathrm{Par} \to \mathbb{N}^{\mathbb{N}_+}$; see Section 2.4.1). As is conventional, we will notationally identify partitions with their corresponding shapes. We will draw shapes as Young diagrams in English notation, with elements represented by boxes placed in the plane with the first coordinate increasing from north to south and the second coordinate increasing from west to east. (See Figure 5.1 for examples.) In light of this we will typically call the elements of a Ferrers shapes *boxes*. (Others may prefer the term *cells*.)

A *skew shape*, or simply *shape*, is a finite convex[1] subset of $\mathbb{N}_+ \times \mathbb{N}_+$. Given partitions λ and μ with $\lambda \geq \mu$, the skew shape λ/μ is defined to be the difference of their Ferrers shapes; explicitly

$$\lambda/\mu = \{(i,j) \in \mathbb{N}_+ \times \mathbb{N}_+ \colon \mu_i < j \leq \lambda_i\}.$$

It is clear that every skew shape arises this way. Note however that the shape λ/μ does not uniquely determine the partitions λ and μ; for instance, $\lambda/\lambda = \emptyset$ regardless of λ. A *subshape* of a skew shape is a subset which is itself a skew shape, i.e. is convex. A skew

[1] In the order-theoretic sense, as in Section 2.3.

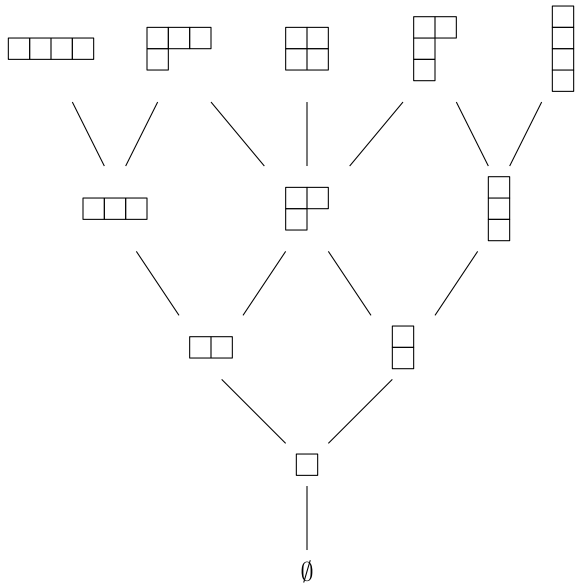


Figure 5.1. Shapes of size at most 4 in $\mathbb{Y}$.

(a) (b) (c) (d)

Figure 5.2. Some skew shapes.

shape is *connected* if it is connected as a poset. We can rephrase this as follows: two boxes are *adjacent* if they share an edge, and the shape is connected if the graph defined by this adjacency relation is connected. Note in particular that "adjacency" here is horizontal or vertical, never diagonal. For instance, in Figure 5.2, shape (a) is connected while the rest are disconnected.

Remark 5.1.1. Note that if we are identifying partitions with their shapes, λ/μ could also be written as $\lambda \setminus \mu$ using the usual notation for set difference. Since the notation λ/μ is traditional, we retain it for this special case of a pair of partitions with $\lambda \geq \mu$ but will use the ordinary set-difference notation for other cases of deleting a subshape (which need not be a downset) from a skew shape.

By a *row* of D we mean a *nonempty* subshape of D which is the intersection with a horizontal line $\{i\} \times \mathbb{N}_+$, and similarly a *column* is a nonempty intersection with a vertical line $\mathbb{N}_+ \times \{j\}$. We denote the number of rows and columns of D by $\operatorname{row}(D)$ and $\operatorname{col}(D)$ respectively. In the case $D = \lambda$ is a partition the rows simply correspond to the parts, so $\operatorname{row}(\lambda) = \ell(\lambda)$.

The *content* of a box $(i, j) \in \mathbb{N}_+ \times \mathbb{N}_+$ is $c(i, j) = j - i$. The set of all boxes with content k is a northwest-southeast diagonal in the plane; we will always mean one of these when we say *diagonal*. Write $\operatorname{dia}(D)$ for the number of diagonals which have nonempty intersection with a shape D; we will also say that such diagonals are *occupied* by D. Equivalently, $\mathrm{dia}(D)$ is the number of distinct values which appear as contents of boxes in D. In the case of a connected shape, $\mathrm{dia}(D) = \mathrm{row}(D) + \mathrm{col}(D) - 1$. Since all three statistics clearly add over connected components we obtain the following.

Lemma 5.1.2. *For any shape D, the number of connected components of D is* $\mathrm{dia}(D) - \mathrm{row}(D) - \mathrm{col}(D)$.

For skew shapes D and E we will write $D \leq E$ if D is a downset of E; note this is stronger than being a subshape. Equivalently, $D \leq \lambda/\mu$ if and only if $D = \rho/\mu$ for some partition ρ with $\mu \leq \rho \leq \lambda$. In particular, for ordinary Ferrers shapes this agrees with the ordering on $\mathbb{Y}$. If $D \leq E$ then $E \setminus D$ is also a skew shape; we say that D is *removable on the left* and $E \setminus D$ is *removable on the right* from E.

If D and D' are connected skew shapes, we write $D \approx D'$ and say that D and D' are *congruent* if D' is a translation of D. We extend this to disconnected shapes by allowing each connected component to be translated separately. (For instance, in Figure 5.2, shapes (b), (c), and (d) are all congruent.) In general everything we do with skew shapes will respect congruence, and we will often identify congruent shapes when it is not confusing to do so. In particular, many of the operations on skew shapes that we consider will not come with a canonical way to view the constructed shapes as subsets of $\mathbb{N}_+ \times \mathbb{N}_+$ or as intervals in $\mathbb{Y}$, so they are better thought of as operations on congruence classes of shapes. The first and simplest of these is the *edge-disjoint union* $D \sqcup E$, consisting of copies of D and E placed so that no box of D is adjacent to any box of E.

Let $\mathcal{S}$ be the free vector space on congruence classes of shapes. We will write $[D]$ for the congruence class of D thought of as an element of $\mathcal{S}$. We make $\mathcal{S}$ into an algebra with product given by $[D][E] = [D \sqcup E]$ and coproduct

$$\Delta[D] = \sum_{X \leq D} [X] \otimes [D \setminus X]. \tag{5.1}$$

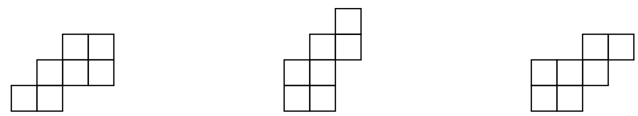
These are easily seen to make $\mathcal{S}$ into a connected graded bialgebra and hence a Hopf algebra, which we call the *shape Hopf algebra*. This Hopf algebra seems to have first been explicitly considered by Yeats [63]. Note that the definition of congruence for disconnected shapes is exactly what it needs to be in order to ensure that, as an algebra, $\mathcal{S}$ is a free commutative algebra on congruence classes of connected shapes.

Remark 5.1.3. Note that the coproduct in $\mathcal{S}$ is an upset-downset coproduct just like that in the poset Hopf algebra $\mathcal{P}$ considered in Section 2.3. Thus we have a natural bialgebra morphism $\mathcal{S} \to \mathcal{P}$ sending a skew shape to its underlying poset. However, this map is not injective, as non-congruent shapes can be isomorphic as posets.

Remark 5.1.4. We can lift the notion of congruence from skew shapes to intervals in $\mathbb{Y}$ by declaring that $[\mu, \lambda] \approx [\mu', \lambda']$ if and only if $\lambda/\mu \approx \lambda'/\mu'$. This is an example of a Hopf relation and $\mathcal{S}$ is the associated incidence Hopf algebra. (See Remark 2.3.2.)

We mention two other operations on skew shapes which play nicely with the Hopf algebra structure of $\mathcal{S}$. Let D be a skew shape. The *transpose* of D is

$$D^T = \{(j, i) \colon (i, j) \in D\}$$



(a) D **(b)** D^T **(c)** D^*

Figure 5.3. Transpose and antipodal rotation of a shape D.

which is clearly also a skew shape. (In the case of partitions, the transpose is commonly called the *conjugate*.) The *antipodal rotation* D^* of D is obtained by rotating D by 180°. Note that the antipodal rotation is only defined up to translation as we have not indicated what point in the plane we are rotating about, nor *can* we choose a single point that ensures we stay in the positive quadrant for all skew shapes. As usual, we are really interested in operations on congruence classes so this is not an issue. To be explicit, however, we may choose any m and n such that $D \subseteq [m] \times [n]$; then

$$D^* \approx \{(m-j, n-i) \colon (i,j) \in D\}.$$

See Figure 5.3 for examples of these operations. It is easy to verify the following result. (Recall that an *anti-automorphism* of a bialgebra is an invertible linear map that reverses the order of multiplication and comultiplication. Of course, since we are in the commutative setting, only the latter is relevant.)

Proposition 5.1.5. *The maps $D \mapsto D^T$ and $D \mapsto D^*$ respectively extend to an automorphism and an anti-automorphism of $\mathcal{S}$.*

5.1.2 Skew Schur functions

Let D be a skew shape. A *weak tableau* (or *reverse plane partition*) of shape D is a weakly increasing map $T \colon D \to \mathbb{N}_+$. These are typically illustrated by filling each box $b \in D$ with the value $T(b)$. For $i \in \mathbb{N}_+$ we will write T_i for the subshape consisting of boxes filled with i. The condition that T is weakly increasing equivalently says that

$$T_1 \leq T_1 \cup T_2 \leq T_1 \cup T_2 \cup T_3 \leq \cdots$$

and indeed a weak tableau is really just a convenient way to encode a chain of skew shapes. With (5.1) in mind, iterated coproducts in $\mathcal{S}$ can be written as sums over weak tableaux:

$$\Delta^k D = \sum_T T_1 \otimes \cdots \otimes T_k \tag{5.2}$$

where the sum is over weak tableaux of shape D with values in $\{1, \ldots, k\}$.

The *weight* of a weak tableau T is the vector $\operatorname{wt} T = (|T_1|, |T_2|, \ldots)$. A *semistandard Young tableau* is a weak tableau which is *strictly* increasing down columns. The set of semistandard Young tableaux of shape D is denoted $\mathrm{SSYT}(D)$. The *skew Schur function* associated to D is

$$s_D = \sum_{T \in \mathrm{SSYT}(D)} \mathbf{x}^{\operatorname{wt} T} \tag{5.3}$$

(see Section 2.1 for notation). While not completely obvious, it well-known these are symmetric functions. When $D = \lambda$ is a partition, the skew Schur function s_λ is a *Schur function*.[2] (Since we will generally be interested in skew Schur functions, we will sometimes say *ordinary Schur function* for emphasis when dealing with the non-skew case.)

Theorem 5.1.6 ([56, Corollary 7.12.2])**.** *The Schur functions $\{s_\lambda\}_{\lambda\in\mathrm{Par}}$ form a basis for* Sym*. Moreover, they are orthonormal with respect to the Hall inner product.*

The map $\mathsf{schur}\colon \mathcal{S} \to \mathsf{Sym}$ sending D to s_D is in fact a bialgebra morphism. This fact can be seen by a direct calculation, but it is perhaps more pleasant to derive it from the Aguiar–Bergeron–Sottile theorem (Theorem 2.4.6). Define a *vertical strip* to a be a skew shape with no more than one box in each column. The condition that a weak tableau T be semistandard is equivalent to requiring that the subshapes T_i are vertical strips. Let $\zeta \in \mathrm{Ch}(\mathcal{S})$ be given by

$$\zeta([D]) = \begin{cases} 1, & D \text{ is a horizontal strip} \\ 0, & \text{otherwise.} \end{cases}$$

Then, in terms of the notation defined before Theorem 2.4.6, for any composition $\alpha = (\alpha_1, \dots, \alpha_k)$ we have

$$\zeta_\alpha([D]) = \sum_T \zeta_{\alpha_1}(T_1)\cdots\zeta_{\alpha_k}(T_k)$$

by (5.2). In other words, $\zeta_\alpha([D])$ counts semistandard tableaux of weight α. Comparing (5.3) to the explicit formula of Theorem 2.4.6 we see that the Aguiar–Bergeron–Sottile map induced by ζ is none other than schur.

Remark 5.1.7. The Aguiar–Bergeron–Sottile theorem guarantees that the induced map takes values in Sym when the source Hopf algebra is cocommutative. However, in this case we see that schur takes values in Sym even though $\mathcal{S}$ is *not* cocommutative. This slightly mysterious fact is key to everything we do in this chapter.

Since schur is a bialgebra morphism, skew Schur functions satisfy the coproduct formula

$$\Delta s_D = \sum_{X \le D} s_X \otimes s_{D\setminus X}. \tag{5.4}$$

Note that if $D = \lambda$ is a partition, then a shape removable on the left from D is also a partition. Thus by orthogonality, we have

$$s_\mu^\perp s_\lambda = \sum_{\nu\le\lambda} \langle s_\mu, s_\nu\rangle s_{\lambda/\nu} = \begin{cases} s_{\lambda/\mu}, & \lambda \ge \mu \\ 0, & \text{otherwise.} \end{cases} \tag{5.5}$$

Consider expanding a product of two Schur functions in the Schur basis. We can write

$$s_\mu s_\nu = \sum_{\lambda\vdash|\mu|+|\nu|} c^\lambda_{\mu,\nu} s_\lambda \tag{5.6}$$

[2] Restricted to n variables, these are *Schur polynomials*, the irreducible characters of $\mathrm{GL}_n(\mathbb{C})$.

for some coefficients $c^\lambda_{\mu,\nu}$.[3] Note that by commutativity, $c^\lambda_{\mu,\nu} = c^\lambda_{\nu,\mu}$. Now, since $s_\mu^\perp$ is by definition the operator adjoint to multiplication by s_μ, the matrix representing it is just the transpose. This means that these same coefficients appear in the expansion of a skew Schur function in the Schur basis:

$$s_{\lambda/\mu} = s_\mu^\perp s_\lambda = \sum_{\nu \vdash |\lambda|-|\mu|} c^\lambda_{\mu,\nu} s_\nu. \tag{5.7}$$

The transpose and antipodal rotation operations also behave nicely on Schur functions. (Recall the fundamental involution ω from Section 2.4.1.)

Proposition 5.1.8 ([56, Theorem 7.15.6])**.** *For any skew shape D, $\omega(s_D) = s_{D^T}$.*

Proposition 5.1.9 ([56, Exercise 7.56(a)])**.** *For any skew shape D, $s_D = s_{D^*}$.*

Proposition 5.1.9 is our first example of a situation in which two distinct skew shapes have equal skew Schur functions. Our primary goal in this chapter will be to understand when this can happen. It is useful to introduce some terminology and notation: two skew shapes D, D' are said to be *skew-equivalent*, written $D \sim D'$, if $s_D = s_{D'}$. Thus by Proposition 5.1.9, we have $D \sim D^*$ for all shapes D. Van Willigenburg proved a partial converse.

Theorem 5.1.10 (Van Willigenburg [61, Theorem 2.2])**.** *Suppose λ is a partition. If D is a skew shape such that $D \sim \lambda$ then $D \approx \lambda$ or $D \approx \lambda^*$.*

However, the situation becomes far more complicated when we leave behind the comfortable world of partitions and venture into the wilderness of arbitrary skew shapes. The study of skew equivalence for more general shapes was initiated by Billera, Thomas, and van Willigenburg [11] who gave a complete characterization in the case of so-called *ribbon* shapes; we will explain this in the next subsection. Building on this work, Reiner, Shaw, and van Willigenburg [51] gave some necessary and some sufficient conditions for various classes of shapes. Finally, McNamara and van Willigenburg [43] gave conditions that they conjectured to be both necessary and sufficient for the general case but were unable to prove either direction in complete generality. Our main result is that these conditions are indeed sufficient. Merely stating the conjecture requires a considerable amount of setup; we will devote much of Section 5.3 to this. For now, let us simply mention some basic facts about skew equivalence.

Lemma 5.1.11. *Suppose $D \sim D'$. Then D and D' have the same number of boxes, connected components, rows, columns, and diagonals.*

Proof. Each of these can be read off of the skew Schur function $s_D = s_{D'}$:

- The number of boxes is simply the degree of s_D.
- A result of Reiner, Shaw, and van Willigenburg [51, Corollary 6.3] says that skew Schur functions corresponding to connected shapes are irreducible elements of the ring Sym. Since Sym is isomorphic to a polynomial algebra over a field it is a unique factorization domain. Thus the number of connected components of D is the number of factors in an irreducible factorization of s_D

[3] These are the *Littlewood–Richardson coefficients*, for which various combinatorial rules are known. (See [56, Section A1.3].) We will never need to make explicit use of these.

- Only the northernmost entry of each column can be filled with a 1 in a semistandard Young tableau; conversely it is clearly possible to find such a tableau in which each column has its northernmost entry filled with 1. Thus the number of columns is the maximum power of x_1 that appears in any monomial in s_D.
- The number of rows of D is the number of columns of D^T, hence is the maximum power of x_1 that appears in any monomial in $\omega(s_D)$.
- By Lemma 5.1.2 the number of diagonals can be derived from the numbers of connected components, rows, and columns. □

Remark 5.1.12. Reiner, Shaw, and van Willigenburg introduced more refined invariants, the *row overlap partitions*, from which all of the data mentioned in Lemma 5.1.11 can be derived. They show [51, Corollary 8.11] that these too are preserved by skew equivalence. McNamara [42] showed that these agree under the weaker hypothesis that the two skew Schur functions have the same support when expanded in terms of fundamental quasisymmetric functions, and conjectured that the converse also holds.

5.1.3 Ribbons

A *ribbon* (also called a *border strip*, *rim hook*, or *skew hook*) is a connected skew shape containing no 2×2 square. Equivalently, a connected shape D is a ribbon if and only if $|D| = \mathrm{dia}(D)$; that is, there are no two boxes with the same content. In particular, the boxes are totally ordered by content and we will always think of boxes of ribbons in terms of this ordering. (Note that while content is not translation-invariant, the *difference* between the contents of two cells is. Thus it makes sense to talk about this ordering even when we are considering shapes up to congruence.)

For any connected skew shape D, the set of boxes of D with no box to the northwest is a ribbon, the *northwest border* which we denote $\mathrm{NW}(D)$. Similarly, the set of boxes with no box to the southeast is the *southeast border* $\mathrm{SE}(D)$. These notions also make sense for D disconnected, in which case $\mathrm{NW}(D)$ and $\mathrm{SE}(D)$ are not ribbons but are disjoint unions of ribbons. Regardless of whether or not D is connected, both borders contain exactly one box from each diagonal occupied by D and so we have $|\mathrm{NW}(D)| = |\mathrm{SE}(D)| = \mathrm{dia}(D)$.

The *height* of a ribbon is its number of rows minus 1. Equivalently, it is the vertical distance between the first and last cell viewed as elements of $\mathbb{N}_+ \times \mathbb{N}_+$. This statistic plays a key role in perhaps the most important result relating to ribbons, the *Murnaghan–Nakayama rule*, which we will summarize now. For a shape D and a composition $\alpha \vDash |D|$, a *ribbon tableau* of shape D and weight α is a weak tableau of weight α such that the shapes $T_1, \ldots, T_{\ell(\alpha)}$ are ribbons. Denote the set of these by $\mathrm{RT}(D, \alpha)$. Define the *height* $\mathrm{ht}\, T$ of a ribbon tableau T to be the sum of the heights of the ribbons comprising it. The *Murnaghan–Nakayama coefficients* are defined by

$$\chi(D, \alpha) = \sum_{T \in \mathrm{RT}(D,\alpha)} (-1)^{\mathrm{ht}\, T}.$$

The best known version of the Murnaghan–Nakayama rule concerns the multiplication of a Schur function with power sums. (This can be found in standard references on symmetric functions, e.g. Stanley's book [56, Section 7.17].)

Theorem 5.1.13 (Murnaghan–Nakayama Rule I)**.** *For $\mu \vdash m$ and $\alpha \vDash n$,*

$$s_\mu p_\alpha = \sum_{\lambda \vdash m+n} \chi(\lambda/\mu, \alpha) s_\lambda. \tag{5.8}$$

We can make some observations from this. First, since $p_\alpha = p_\beta$ when α is a rearrangement of β, it is immediate from (5.8) that $\chi(D, \alpha) = \chi(D, \beta)$ in this case as well. Thus we can safely restrict attention to the case of α a partition. Secondly, taking $\mu = \emptyset$ we get the expansion of power sums in terms of Schur functions: for $\nu \vdash n$,

$$p_\nu = \sum_{\lambda \vdash n} \chi(\lambda, \nu) s_\lambda. \tag{5.9}$$

The version of the Murnaghan–Nakayama rule we will find most useful is not Theorem 5.1.13 but an equivalent statement in terms of inner products.

Theorem 5.1.14 (Murnaghan–Nakayama Rule II)**.** *For any skew shape D and $\nu \vdash |D|$, $\langle p_\nu, s_D \rangle = \chi(D, \nu)$.*

Proof. Writing $D = \lambda/\mu$ we have $\langle p_\nu, s_{\lambda/\mu} \rangle = \langle s_\mu p_\nu, s_\lambda \rangle$ and the result follows immediately from Theorem 5.1.13 by orthonormality of Schur functions. □

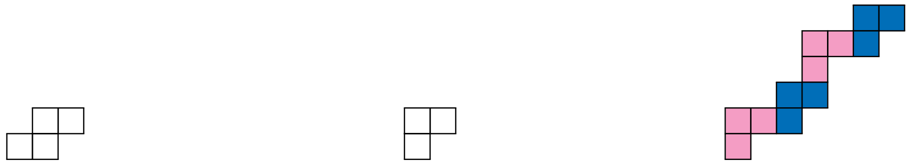
Corollary 5.1.15. *For any skew shape D and natural number m,*

$$\langle p_m, s_D \rangle = \begin{cases} (-1)^{\operatorname{ht} D} & D \text{ is a ribbon of size } m \\ 0 & \text{otherwise.} \end{cases}$$

Note that Theorem 5.1.14 implies that Murnaghan–Nakayama coefficients respect skew equivalence: $\chi(D, \nu) = \chi(D', \nu)$ whenever $D \sim D'$. Using Theorem 5.1.14 and the orthogonality of power sums, we can also derive the expansion of skew Schur functions in the power sum basis:

$$s_D = \sum_{\nu \vdash |D|} \frac{\chi(D, \nu)}{z_\nu} p_\nu. \tag{5.10}$$

We remarked in the previous section that skew equivalences involving ribbons are completely characterized; we will now discuss this. The key idea is the operation of *composition* of ribbons introduced by Billera, Thomas, and van Willigenburg [11]. This is defined in terms of two simpler operations, horizontal and vertical attachment, which make sense for non-ribbon shapes as well. Given two skew shapes D and E, the *horizontal attachment* $D \odot E$ is the shape obtained by placing E in the plane with its southwesternmost box immediately east of the northeasternmost box of D. The *vertical attachment* $D \cdot E$ is similar, but instead the southwesternmost box of E is placed immediately north of the northeasternmost box of D.



(a) A **(b)** B **(c)** $A \circ B$

Figure 5.4. Ribbon composition. In (c) the copies of B are highlighted in alternating colours.

For A and B, the *composition* $A \circ B$ is defined to be a ribbon consisting of a copy of B for each box of A. Copies corresponding to adjacent vertices are attached using $\odot$ when the boxes are horizontally adjacent and $\cdot$ when vertically adjacent. Composition of ribbons is algebraically nice: it is associative [11, Proposition 3.3] and satisfies a certain unique factorization property [11, Theorem 3.6]. Combining this operation with antipodal rotation produces all skew-equivalences between ribbon shapes.

Theorem 5.1.16 (Billera–Thomas–van Willigenburg [11, Theorem 4.1])**.** *Suppose A and A' are ribbons. Then $A \sim A'$ if and only if there exist factorizations*

$$A = B_1 \circ \cdots \circ B_k$$

and

$$A' = B'_1 \circ \cdots \circ B'_k$$

where for each i, either $B'_i = B_i$ or $B'_i = B^*_i$.

(Note also that Corollary 5.1.15 immediately implies that a non-ribbon cannot be equivalent to a ribbon.)

Implicit in the statement of Theorem 5.1.16 is the fact that ribbon composition respects skew equivalence: if $A \sim A'$ and $B \sim B'$ then $A \circ B \sim A' \circ B'$. To generalize Theorem 5.1.17, one can define $D \circ B$ for an arbitrary skew shape D in exactly the same way as we did in the case D is a ribbon.[4] This was considered by Reiner, Shaw, and van Willigenburg [51] who showed that it still respects skew equivalence.

Theorem 5.1.17 (Reiner–Shaw–van Willigenburg [51, Theorem 7.6])**.** *Suppose D and D' are skew shapes and B and B' are ribbons such that $D \sim D'$ and $B \sim B'$. Then $D \circ B \sim D' \circ B'$.*

Unfortunately, Theorem 5.1.17 is far from a complete characterization of skew equivalence among general shapes. Indeed, [51] introduces several other constructions that produce equivalent shapes. Moreover, [51, Section 9] gives examples of skew equivalences that do not follow from any of these results. All of these are unified in the theory of WOW composition, the subject of Section 5.3. Before doing so, in the next section we will introduce our Hopf-algebraic framework for skew equivalence problems and use it to give a new proof of Theorem 5.1.17.

[4] One must check that this is well-defined; in [51] this was left as an exercise to the reader, and we choose to continue the tradition.

5.2 Actions of symmetric functions on shapes

We return to the story of the shape Hopf algebra $\mathcal{S}$ and the map $\mathsf{schur}\colon \mathcal{S} \to \mathsf{Sym}$ discussed in Section 5.1.2. Since schur is a coalgebra morphism, it induces a $(\mathsf{Sym}, \mathsf{Sym})$-bicomodule structure on $\mathcal{S}$. Then by self-duality, this in turn induces commuting left and right actions of Sym on $\mathcal{S}$ as in Section 2.2.3. Explicitly these are given by

$$f \rightharpoonup [D] = \sum_{X \le D} \langle f, s_{D \setminus X} \rangle X \tag{5.11}$$

and

$$[D] \leftharpoonup f = \sum_{X \le D} \langle f, s_X \rangle (D \setminus X). \tag{5.12}$$

We will make use of these actions repeatedly, so let us establish now some basic facts about them.

Proposition 5.2.1. *The left and right actions of* Sym *on* $\mathcal{S}$ *satisfy the following properties:*

(i) For any $f \in \mathsf{Sym}$ *and* $h \in \mathcal{S}$*, we have* $\mathsf{schur}(f \rightharpoonup h) = \mathsf{schur}(h \leftharpoonup f) = f^{\perp}\mathsf{schur}(h)$*.*

(ii) For any $f, g \in \mathsf{Sym}$ *and* $h \in \mathcal{S}$*, we have* $(f \rightharpoonup h \leftharpoonup g)^T = \omega(f) \rightharpoonup h^T \leftharpoonup \omega(g)$*.*

(iii) For any $f, g \in \mathsf{Sym}$ *and* $h \in \mathcal{S}$*, we have* $(f \rightharpoonup h \leftharpoonup g)^* = g \rightharpoonup h^* \leftharpoonup f$*.*

(iv) For partitions $\lambda \ge \mu$*, we have* $[\lambda/\mu] = [\lambda] \leftharpoonup s_\mu$ *and* $[\lambda/\mu]^* = s_\mu \rightharpoonup [\lambda]^*$*.*

(v) For partitions $\rho \ge \lambda \ge \mu$ *with* ρ *a rectangle, we have* $[\lambda/\mu] = s_{\rho/\lambda} \rightharpoonup [\rho] \leftharpoonup s_\mu$*.*

Proof. (i) It is immediate from the definition that $\mathsf{schur}(f \rightharpoonup h) = f \rightharpoonup \mathsf{schur}(h)$ and $\mathsf{schur}(h \leftharpoonup f) = \mathsf{schur}(h) \leftharpoonup f$. By cocommutativity these agree, and $f^{\perp}\mathsf{schur}(h)$ is simply alternative notation.

(ii) For this one and the next it is convenient to express the actions in terms of the map $\mathsf{schur}^\vee\colon \mathsf{Sym} \to \mathcal{S}^\vee$ adjoint to schur. Note that $f \rightharpoonup h \leftharpoonup g$ is just the action of the operator $\mathsf{schur}^\vee(g) * \mathrm{id} * \mathsf{schur}^\vee(f)$ on the element h. Let us temporarily denote $h \mapsto h^T$ by φ. Then Proposition 5.1.8 gives $\omega \circ \mathsf{schur} = \mathsf{schur} \circ \varphi$. Taking adjoints gives

$$\mathsf{schur}^\vee(\omega(f)) = \varphi^\vee \mathsf{schur}^\vee(f) = \mathsf{schur}^\vee(f)\varphi$$

and so since φ is an involution and an automorphism

$$\begin{aligned}
\varphi(\mathsf{schur}^\vee(g) * \mathrm{id} * \mathsf{schur}^\vee(f)) &= (\mathsf{schur}^\vee(g) \otimes \varphi \otimes \mathsf{schur}^\vee(f))\Delta_3 \\
&= (\mathsf{schur}^\vee(g)\varphi \otimes \mathrm{id} \otimes \mathsf{schur}^\vee(f)\varphi)(\varphi \otimes \varphi \otimes \varphi)\Delta_3 \\
&= (\mathsf{schur}^\vee(\omega(g)) \otimes \mathrm{id} \otimes \mathsf{schur}^\vee(\omega(f)))\Delta_3\varphi \\
&= (\mathsf{schur}^\vee(\omega(g)) * \mathrm{id} * \mathsf{schur}^\vee(\omega(f)))\varphi
\end{aligned}$$

as desired.

(iii) Let us temporarily denote $h \mapsto h^D$ by ψ. Then Proposition 5.1.9 gives $\mathsf{schur} = \mathsf{schur} \circ \psi$. Taking adjoints gives

$$\mathsf{schur}^\vee(f) = \psi^\vee \mathsf{schur}^\vee(f) = \mathsf{schur}^\vee(f)\psi$$

and so since ψ is an involution and an anti-automorphism

$$\begin{aligned}
\psi(\mathsf{schur}^\vee(g) * \mathrm{id} * \mathsf{schur}^\vee(f)) &= (\mathsf{schur}^\vee(g) \otimes \psi \otimes \mathsf{schur}^\vee(f))\Delta_3 \\
&= (\mathsf{schur}^\vee(g)\psi \otimes \mathrm{id} \otimes \mathsf{schur}^\vee(f)\psi)(\psi \otimes \psi \otimes \psi)\Delta_3 \\
&= (\mathsf{schur}^\vee(f) \otimes \mathrm{id} \otimes \mathsf{schur}^\vee(g))\Delta_3\psi \\
&= (\mathsf{schur}^\vee(f) * \mathrm{id} * \mathsf{schur}^\vee(g))\psi
\end{aligned}$$

as desired.

(iv) By orthonormality of Schur functions,

$$[\lambda] \leftharpoonup s_\mu = \sum_{\nu \leq \lambda} \langle s_\mu, s_\nu \rangle [\lambda/\nu] = [\lambda/\mu].$$

By (ii), applying antipodal rotation to both sides produces the other formula.

(v) By (iv) we have

$$[\lambda] = [\rho/\nu]^* = s_{\rho/\lambda} \rightharpoonup [\rho]$$

Applying (iv) a second time gives the desired formula. □

A consequence of Proposition 5.2.1(iv) is that $\mathcal{S}$ is generated as a right Sym-module by partitions and as a left Sym-module by antipodal rotations of partitions. Since every partition fits into some rectangle, Proposition 5.2.1(v) gives that the rectangles generate $\mathcal{S}$ as a $(\mathsf{Sym}, \mathsf{Sym})$-bimodule.

It is natural to extend the notion of skew equivalence from shapes to arbitrary elements of the shape Hopf algebra: that is, for $h, h' \in \mathcal{S}$ we write $h \sim h'$ if $\mathsf{schur}(h) = \mathsf{schur}(h')$. This is, of course, simply the congruence induced by the kernel of schur. Thus Proposition 5.2.1(i) says $f \rightharpoonup h \sim h \leftharpoonup f$. Our first new result of the chapter shows that in fact *this is the root of all skew equivalence.*

Theorem 5.2.2. *The kernel of* schur *is spanned by elements of the form* $f \rightharpoonup h - h \leftharpoonup f$ *for* $f \in \mathsf{Sym}$ *and* $h \in \mathcal{S}$.

Proof. The images of partitions, the ordinary Schur functions, are a basis. Thus by (5.7), the kernel is spanned by elements of the form

$$[\lambda/\mu] - \sum_\nu c^\lambda_{\mu,\nu}[\nu]$$

so it is sufficient to show that these elements are of the desired form. We have already seen that $[\lambda/\mu] = [\lambda] \leftharpoonup s_\mu$. On the other hand,

$$s_\mu \rightharpoonup [\lambda] = \sum_\nu \langle s_\mu, s_{\lambda/\nu} \rangle [\nu]$$

$$= \sum_{\nu} c^{\lambda}_{\mu,\nu}[\nu].$$

The result follows. □

Remark 5.2.3. The Hopf-algebraic approach to skew equivalence was pioneered by Yeats [63], who used the cocommutativity of Sym to derive a partial result towards the McNamara–van Willigenburg conjecture. She did not phrase her results in terms of the actions of Sym on $\mathcal{S}$, but one can re-frame her argument in this way as we explain in Section 5.3.3. The author's work on the subject began with a modest attempt to generalize Yeats's methods. We may view Theorem 5.2.2 as the ultimate vindication of this approach, showing that *every* skew equivalence can be derived this way. (Indeed, by the end of this chapter, we will have done so for all known constructions of equivalent shapes as well as some previously conjectural ones.)

5.2.1 Ribbon composition revisited

We now use Theorem 5.2.2, along with the Murnaghan–Nakayama rule, to give a new proof of Theorem 5.1.17, and along the way prove some lemmas which we will be of use later in proving new results. Let us linearly extend the map $[D] \mapsto [D \circ B]$ to a map (indeed an algebra morphism) $\mathcal{S} \to \mathcal{S}$, which we also denote by $h \mapsto h \circ B$. What we will actually prove is the following strengthening of Theorem 5.1.17.

Proposition 5.2.4. *Let $h, h' \in \mathcal{S}$ be arbitrary and B be a ribbon. If $h \sim h'$ then $h \circ B \sim h' \circ B$.*

Remark 5.2.5. While Proposition 5.2.4 is stronger than Theorem 5.1.17 as stated, it can still be proved by the methods of [51]. In particular, [51, Proposition 7.5] states that there exists a well-defined algebra morphism $\mathsf{Sym} \to \mathsf{Sym}$ that sends s_D to $s_{D\circ B}$, from which Proposition 5.2.4 easily follows. However, our proof is somewhat different and acts as a pleasant warmup to the proofs of our main results.

Of course, one cannot prove a result such as this by general abstract nonsense. Having constructed the required algebraic abstractions, we have finally arrived at the point where we need to do some combinatorics. The key idea is to understand which ribbons can be removed on the left and/or right from shapes of the form $D \circ B$; this will be a running theme through all of our main results in this chapter.

Lemma 5.2.6. *Let B be a ribbon of size k. Then:*

- *The only ribbon of size k that is removable on the left from $B \odot B$ is the first copy of B.*
- *The only ribbon of size k that is removable on the right from $B \odot B$ is the second copy of B.*
- *The only ribbon of size k that is removable on the left from $B \cdot B$ is the second copy of B.*

- *The only ribbon of size k that is removable on the right from $B \cdot B$ is the first copy of B.*

Proof. It is clear that the claimed ribbons are removable; we will show no other ribbon of size k is removable on either side. Let B_1 and B_2 denote the two copies of B in either $B \odot B$ or $B \cdot B$. Suppose A is a ribbon of size k which is removable on the left and which is not equal to B_1 or B_2. Let a and a' be the boxes of A of minimum and maximum content respectively. Clearly A is not contained in B_1 or B_2, so $a \in B_1$ and $a' \in B_2$. Since A is removable on the left, a does not have a box to its west; thus it has a box a'' to its south. But then a'' is at a distance of k from a', so it is the box corresponding to a' in B_1. Since a'' has a box to its north, so does a', contradicting the assumption that A is removable on the left. A similar argument shows that A cannot be removable on the right either. □

Lemma 5.2.7. *Let D be any skew shape and B be a ribbon of size k. The only ribbons of size k which are removable on the left (resp. right) from $D \circ B$ are the copies of B coming from boxes of D which are removable on the left (resp. right).*

Proof. A ribbon removable on the left from $D \circ B$ is contained in $\mathrm{NW}(D \circ B) = \mathrm{NW}(D) \circ B$. If the ribbon has size k, it overlaps at most two copies of B, which are attached using $\cdot$ or $\odot$. By Lemma 5.2.6, if it does overlap two copies, it is not removable. So any removable ribbon of size k is a copy of B. Moreover, it is removable only if the copy before it (if present) is attached using $\cdot$ and the one after (if present) is attached using $\odot$. Thus the corresponding box in $\mathrm{NW}(D)$ does not have a box to its right or above it; this is precisely the criterion for a box to be removable. The argument for ribbons removable on the right is analogous. □

Note that for any subshape X of D, there is a corresponding subshape $X \circ B$ of $D \circ B$. Moreover, if X is removable (on the left or right) then so is $X \circ B$, and $(D \circ B) \setminus (X \circ B) = (D \setminus X) \circ B$. Using these, we can extend Lemma 5.2.7 from ribbons of size k to ribbons of size divisible by k.

Lemma 5.2.8. *Let D be any skew shape and B be a ribbon of size k. The only ribbons of size divisible by k which are removable on the left (resp. right) from $D \circ B$ are shapes of the form $A \circ B$ where A is a ribbon removable on the left (resp. right) from D.*

Proof. Suppose C is a ribbon of size jk removable on the left. We can write $C = C' \odot C''$ or $C = C' \cdot C''$ where $|C'| = k$ and $|C''| = (j-1)k$. In the former case, C' is removable on the left as well, so by Lemma 5.2.7 is a copy of B coming from a removable box b of D. Then C'' is removable from $(D \setminus b) \circ B$ so inductively is of the form $A' \circ B$ for some A' removable on the left from $D \setminus b$. Since C' is attached horizontally to C'', the box b is to the west of the first box of A' and $C = A \circ B$ where $A = b \odot A'$. In the latter case, C'' is removable on the left so similarly we have $C'' = A' \circ B$ where now A' is removable from D and the box b is to the south of the first box of A' (as it must be removable from $D \setminus A$). Thus $C' = A \circ B$ where $A = b \cdot A'$. The argument for ribbons removable on the right is again analogous. □

The key idea of our techniques is to use the Murnaghan–Nakayama rule to relate these combinatorial results about ribbons to the actions of Sym on $\mathcal{S}$. First we need the following straightforward result.

Lemma 5.2.9. *Let A and B be ribbons. Then:*

(i) $\operatorname{ht}(A \cdot B) = \operatorname{ht} A + \operatorname{ht} B + 1$

(ii) $\operatorname{ht}(A \odot B) = \operatorname{ht} A + \operatorname{ht} B$

(iii) $\operatorname{ht}(A \circ B) = |A| \operatorname{ht} B + \operatorname{ht} A$.

Proof. Immediately from the definitions we have

$$\operatorname{ht}(A \cdot B) = \operatorname{row}(A \cdot B) - 1 = \operatorname{row}(A) + \operatorname{row}(B) - 1 = \operatorname{ht} A + \operatorname{ht} B + 1$$

and

$$\operatorname{ht}(A \odot B) = \operatorname{row}(A \odot B) - 1 = \operatorname{row}(A) + \operatorname{row}(B) - 2 = \operatorname{ht} A + \operatorname{ht} B.$$

Then by definition we have

$$A \circ B = B \vartriangle_1 B \vartriangle_2 \cdots \vartriangle_{|A|-1} B$$

where $\vartriangle_i$ is $\odot$ if the ith and $(i+1)$st boxes of A are horizontally adjacent and $\cdot$ if they are vertically adjacent. The latter happens precisely when i is the last box of its row, so the number of such i is the number of rows excluding the last one, i.e. $\operatorname{ht} A$. Then the third formula follows from the first two. $\square$

With this we can apply the Murnaghan–Nakayama rule to ribbon composition.

Lemma 5.2.10. *Let D be any skew shape D and B be a ribbon of size k. For any m we have*

$$p_{km} \rightharpoonup [D \circ B] = (-1)^{m \operatorname{ht} B} (p_m \rightharpoonup [D]) \circ B$$

and

$$[D \circ B] \leftharpoonup p_{km} = (-1)^{m \operatorname{ht} B} ([D] \leftharpoonup p_m) \circ B.$$

Proof. By definition, we have

$$[D \circ B] \leftharpoonup p_{km} = \sum_{Y \leq D \circ B} \langle p_{km}, s_Y \rangle [(D \circ B) \setminus Y].$$

By Corollary 5.1.15, the only terms that contribute are those for which Y is a ribbon of size km. By Lemma 5.2.8, these are of the form $A \circ B$ for A a ribbon of size m removable on the left from D. By Lemma 5.2.9 we have $\operatorname{ht}(A \circ B) = (-1)^{\operatorname{ht} A + |A| \operatorname{ht} B} = (-1)^{|A| \operatorname{ht} B} \langle p_m, s_A \rangle$ and thus

$$\begin{aligned}
[D \circ B] \leftharpoonup p_{km} &= (-1)^{m \operatorname{ht} B} \sum_{A \leq D} \langle p_m, s_A \rangle [(D \circ B) \setminus (A \circ B)] \\
&= (-1)^{m \operatorname{ht} B} \sum_{A \leq D} \langle p_m, s_A \rangle [D \setminus A] \circ B \\
&= (-1)^{m \operatorname{ht} B} ([D] \leftharpoonup p_m) \circ B.
\end{aligned}$$

The argument for the right action is analogous. $\square$

Since the power sums are an algebraically independent generating set for Sym (Theorem 2.4.1), there is a unique morphism $\mathsf{Sym} \to \mathsf{Sym}$ that sends p_m to p_{km} for all m. Denote this by $f \mapsto f[p_k]$.[5] Then Lemma 5.2.10 can be extended to a more general result.

Proposition 5.2.11. *Let f be a symmetric function which is homogeneous of degree m. For any $h \in \mathcal{S}$,*

$$f[p_k] \rightharpoonup (h \circ B) = (-1)^{m \operatorname{ht} B} (f \rightharpoonup h) \circ B$$

and

$$(h \circ B) \leftharpoonup f[p_k] = (-1)^{m \operatorname{ht} B} (h \leftharpoonup f) \circ B.$$

Proof. By linearity it is sufficient to prove this when $f = p_\nu$ for some $\nu \vdash m$ and $h = [D]$ for some shape D. This then follows by an easy induction from Lemma 5.2.10. $\square$

We now have all we need to prove the desired result.

Proof of Proposition 5.2.4. By Theorem 5.2.2, it is sufficient to prove $(f \rightharpoonup h) \circ B \sim (h \leftharpoonup f) \circ B$ for any $f \in \mathsf{Sym}$ and $h \in \mathcal{S}$. But this is immediate from Proposition 5.2.11. $\square$

5.3 WOW shapes

The notions of WOW shape and WOW composition were introduced by McNamara and van Willigenburg [43] in order to generalize Theorem 5.1.16 to all shapes. In this section we review their work as well as the contributions of Yeats [63]. Let W and O be connected skew shapes. A WOW *shape*[6] is a connected skew shape E with the following properties:

1. E contains two copies of W, one of which contains the southwesternmost box and one of which contains the northeasternmost box. (These are clearly unique once they exist; we will refer to them as W_{sw} and W_{ne}.)

2. W is maximal among shapes for which the first property holds and which occupy the same diagonals as W.

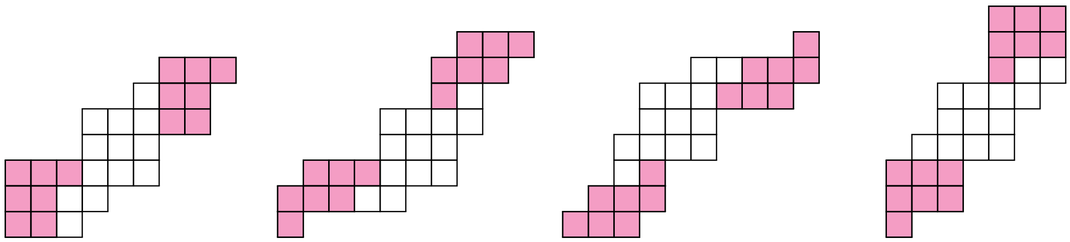
3. There exists at least one diagonal which lies strictly between W_{sw} and W_{ne}. (This in particular means that the two copies are disjoint.)

4. The complement in E of the two copies of W is a copy of O.

The third and fourth properties together have some important consequences for the structure of E. Let b be the northeasternmost box of W_{sw}. Since E is connected, there is some box of E adjacent to b. More specifically, by properties 1 and 3, b cannot be the

[5] It is a special case of the more general operation of *plethysm*, but we do not need this.

[6] Here we are following the terminology used by Yeats. McNamara and van Willigenburg do not use the term "WOW shape" but rather write $E = WOW$ to mean that E satisfies the first of our four properties, with O denoting the complement of the copies of W whether or not it is connected or even a valid skew shape. They refer to our second and third property as Hypotheses I and II, and the assumption that O is actually a connected skew shape as Hypothesis III. Since all of their results, and ours, assume (at least) these hypotheses we prefer to absorb them into the definition.



(a) E_1 (b) E_2 (c) E_3 (d) E_4

Figure 5.5. Some WOW shapes (for varying W and O) with the W subshapes highlighted.

northeasternmost box of E, so it has a box to its east and/or north. Moreover, by property 3 this cannot be a box of W_{ne} so by property 4 it is in O. But since O is itself a skew shape, it is not possible that there are boxes of O to *both* the east and the north of b. Thus exactly one of the two possibilities hold: we say that W_{sw} is *attached horizontally* to O if b has a box of O to its east, and *attached vertically* if it has a box of O to its north. Analogously, the southwesternmost box of W_{ne} has either a box to its west or its south and we respectively say that W_{ne} is attached vertically or horizontally. We say that E is a $W \to O \to W$, $W \to O \uparrow W$, $W \uparrow O \to W$, or $W \uparrow O \uparrow W$ shape where the first arrow is horizontal or vertical depending on how W_{sw} is attached and the second for W_{ne}. Figure 5.5 shows one shape of each type.

It is easy to see that if E is a WOW shape then E^T is a $W^TO^TW^T$ shape and E^* is a $W^*O^*W^*$ shape. In both cases, the two copies of W are swapped. In the case of the transpose, horizontal adjacency in E becomes vertical adjacency in E^T and vice versa, whereas in the case of the antipodal rotation these are preserved. As such, we see that transpose exchanges the $\to\to$ and $\uparrow\uparrow$ cases while preserving the other two, whereas antipodal rotation exchanges the $\to\uparrow$ and $\uparrow\to$ cases while preserving the other two.[7] (For instance, in Figure 5.5 we see that E_1 and E_4 are transposes of one another while E_2 and E_3 are antipodal rotations of one another.) It will transpire that everything we do transforms sensibly under these involutions, so this often allows us to reduce from four cases to just two. Let us call the shapes where both sides are attached the same way ($W \to O \to W$ and $W \uparrow O \uparrow W$) *edge shapes* and those where they are attached differently ($W \to O \uparrow W$ and $W \uparrow O \to W$) *corner shapes*.[8] Each of these classes of shapes is closed under both involutions. Thus by using the properties of the two involutions one can simplify the work of dealing with shapes of one class or another, but the two classes will generally have to be dealt with separately.

While we do not consider the empty shape to be connected, it turns out to be convenient to also allow the case $W = \emptyset$. We will simply say that any connected shape O is an $\emptyset O \emptyset$ shape (the only one), and following [43] we will consider it more specifically to be $\emptyset \to O \to \emptyset$.[9]

[7] This observation is used in [43] but is not explicitly stated as a result there.

[8] This terminology does not appear in [43] or [63]; only the present author can be blamed for it.

[9] Note that this is not *completely* arbitrary: this case does behave like an edge shape. However, the choice to think of it as $W \to O \to W$ rather than $W \uparrow O \uparrow W$ is purely a matter of convention. The definitions we make in the next section are chosen such that it *does* behave like a $W \to O \to W$ shape with regards to composition but they could have been swapped.

5.3.1 Composition of shapes

We now begin the process of building up the McNamara–van Willigenburg composition operation on shapes. Like composition with ribbons, this will be built up from two binary operations. The simpler of these is the *amalgamation* $E \sqcup_W E'$ where E is WOW and E' is $WO'W$. This is obtained from E and E' by identifying W_{ne} in E with W_{sw} in E'. (Note that the second property in the definition of WOW shapes ensures that the overlap between the two copies consists only of the identified copies of W.) If $W = \emptyset$, since we are thinking of E as the horizontal type of edge shape, this should be interpreted as meaning that we place the northeasternmost box of E one unit west of the southwesternmost box of E'.

The *near-amalgamation* $E \cdot_W E$ of a shape E with itself is defined differently in the four cases. Like the amalgamation, it will contain two copies of E which we will call E_1 and E_2 to state the definition. In all cases, the overlap of the two will occupy the same diagonals as W_{ne} in E_1 and W_{sw} in E_2:

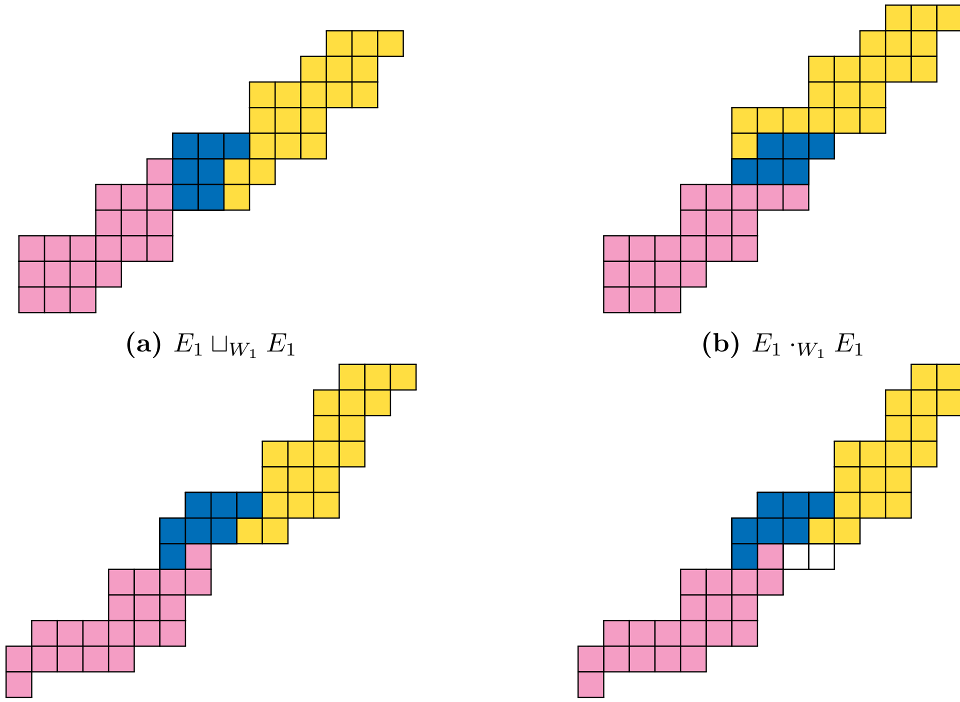
- If $W = \emptyset$, place the northeasternmost box of E_1 one unit south of the southwesternmost box of E_2.
- Otherwise, if E is $W \to O \to W$, W_{ne} in E_1 is one unit southeast of W_{sw} in E_2.
- If E is $W \uparrow O \uparrow W$, W_{ne} in E_1 is one unit northwest of W_{sw} in E_2.
- If E is $W \to O \uparrow W$, overlap the two W's as in the amalgamation but add an additional copy of W offset by one unit southeast.
- If E is $W \uparrow O \to W$, overlap the two W's as in the amalgamation but add an additional copy of W offset by one unit northwest.

Note that no permutation of these possibilities will produce skew shapes in general.

Remark 5.3.1. Notice that the definition is such that $(E \cdot_W E)^T = E^T \cdot_{W^T} E^T$. Thus, despite the notation, this is not analogous to the usual vertical and horizontal attachment operations which instead have $(E \cdot E)^T = E^T \odot E^T$ and $(E \odot E)^T = E^T \cdot E^T$. We will see that the operation $\cdot_W$ on $W \to O \to W$ shapes is closely related to the operation $\cdot$ on ribbons, as the notation suggests. However, on $W \uparrow O \uparrow W$ shapes it is instead related to the $\odot$ operation and on corner shapes both WOW operations will relate to both ribbon operations.

If D is a ribbon and E is a WOW shape, the *composition* $D \circ_W E$ has one copy of E for each box of D, joined using $\sqcup_W$ when the boxes are horizontally adjacent and $\cdot_W$ when they are vertically adjacent. In the edge case, this still makes sense for any skew shape D and is the definition of the composition in general.[10] In the corner case, however, this fails to be a sensible definition: attempting to apply it to a box with neighbours both the north and east, the copies of E corresponding to those boxes should by the definitions of $\sqcup_W$ and $\cdot_W$ overlap *completely*, and so one of the copies of E is redundant. To define it correctly for corner shapes, we make use of certain decompositions of the shape D into ribbons which we will now define.

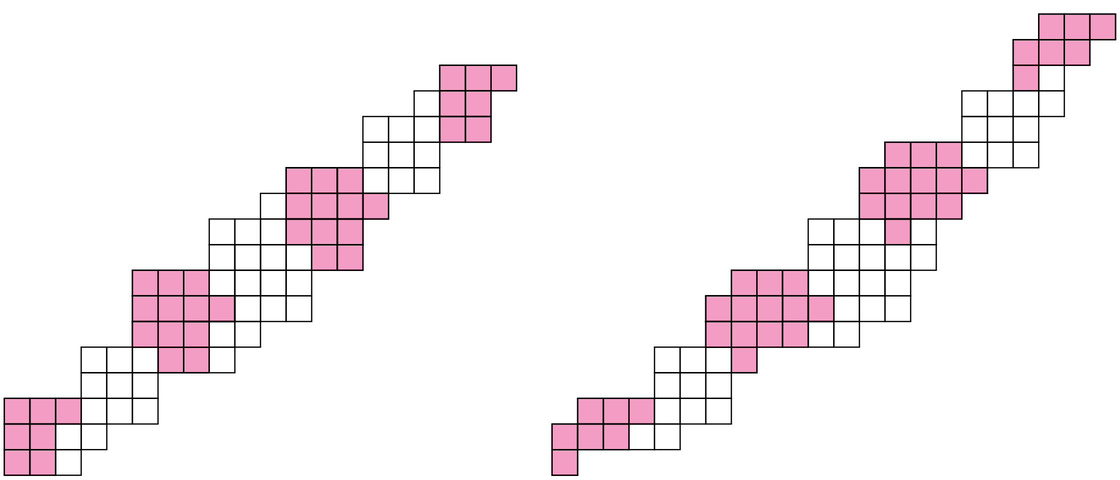
[10]Once again we defer to tradition and omit a proof of well-definedness.



(a) $E_1 \sqcup_{W_1} E_1$ (b) $E_1 \cdot_{W_1} E_1$

(c) $E_2 \sqcup_{W_2} E_2$ (d) $E_2 \cdot_{W_2} E_2$

Figure 5.6. Amalgamations and near-amalgamations of some of the shapes from Figure 5.5. The two copies of E and the overlap between them are highlighted in different colours, while boxes which come from neither copy are in white.



(a) $D \circ_{W_1} E_1$ (b) $D \circ_{W_2} E_2$

Figure 5.7. Compositions of a 2×2 square D with some of the shapes from Figure 5.5. Copies of W are highlighted.

Recall that $\mathrm{NW}(D)$ denotes the northwest border of D. For $k \geq 1$, we recursively define the shapes $\mathrm{NW}_k(D)$ by

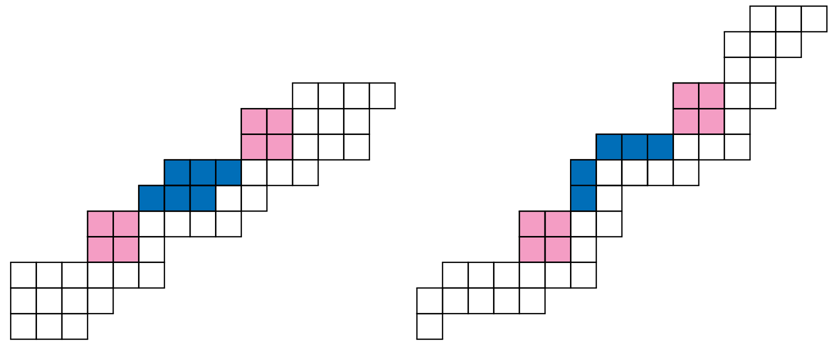
$$\mathrm{NW}_k(D) = \mathrm{NW}(D \setminus (\mathrm{NW}_1(D) \cup \cdots \cup \mathrm{NW}_{k-1}(D))).$$

Clearly these shapes are disjoint and partition the boxes of D. Each connected component of $\mathrm{NW}_k(D)$ is a ribbon; the multiset of ribbons which occur as connected components of these shapes is the *northwest decomposition* of D, which we denote $\mathbf{NW}(D)$. Analogously, we define the *southeast decomposition* $\mathbf{SE}(D)$.

If E is a $W \to O \uparrow W$ shape, we now define $D \circ_W E$ as follows: for each ribbon $A \in \mathbf{NW}(D)$, form the shape $A \circ_W E$. Since every box of D appears in exactly one of these ribbons we have within these shapes a copy E_b of E for each $b \in D$. We construct $D \circ_W E$ by placing the shapes such that if b' is immediately to the southeast of b, then $E_{b'}$ appears offset by one unit to the southeast of E_b, overlapping in all but $\mathrm{NW}(E_b) \cup \mathrm{SE}(E_{b'})$. (We note that the connected components of $\mathrm{NW}(D)$ are in bijection with the connected components of D. Thus we can place the shapes coming from these components first, disjointly, and for the remaining copies of E we always have a box to the northwest telling where us to place it.) If E is a $W \uparrow O \to W$ shape, we define $D \circ_W E$ in exactly the same way but with $\mathbf{SE}(D)$ in place of $\mathbf{NW}(D)$.

Remark 5.3.2. We could have defined the composition for edge shapes similarly to how we did for corner shapes. Indeed, suppose b' is one unit southeast of b in D, and let c be the box south of b and hence west of b'. Then by definition, within $D \circ_W E$, the relevant copies of E appear as $E_c \cdot_W E_b$ and $E_c \sqcup_W E_{b'}$. If E is $W \to O \to W$, this means that $E_{b'}$ is one unit southwest of E_b, so this construction applied to *any* ribbon decomposition would produce $D \circ_W E$. For $W \uparrow O \uparrow W$ shapes we would need a slightly modified construction that places $E_{b'}$ one unit northwest of E_b when b' is one unit southeast of b but could again use any ribbon decomposition. On the other hand, McNamara and van Willigenburg show [43, Lemma 3.18] that this modified construction also works for corner shapes if one swaps the roles of $\mathbf{NW}(D)$ and $\mathbf{SE}(D)$.

Remark 5.3.3. Our definition of composition differs from that of McNamara and van Willigenburg in the case $D = \emptyset$. We do *not* treat this case specially: following directions as written, we find $\emptyset \circ_W E = \emptyset$ regardless of E. McNamara and van Willigenburg instead define $\emptyset \circ_W E = W$. This convention is required to make one of their key lemmas [43, Lemma 3.25], relating composition of shapes to a certain operation on symmetric functions they introduce, true as stated. We nonetheless argue that our convention is superior for three reasons. Firstly, it makes $(D \sqcup D') \circ_W E \approx (D \circ_W E) \sqcup (D' \circ_W E)$ true in all cases rather than requiring an exception in the case that D or D' is empty. This property means that $D \mapsto D \circ_W E$ extends to an algebra morphism $\mathcal{S} \to \mathcal{S}$ which is clearly desirable from the perspective we are taking. Secondly, we will show in Section 5.6 that the definition of the symmetric function operation can be slightly modified in such a way that the resulting operation is better behaved in general *and* does not need this strange convention regarding the empty shape. Finally, one of McNamara and van Willigenburg's main results is a formula [43, Theorem 3.28] which appears to the author to be false as stated according to their convention for the empty shape (reading $s_W = 1$ in this case)



(a) **(b)**

Figure 5.8. The two copies of $\overline{O}$ and the subshape $\overline{W}$ within $E_1 \sqcup_{W_1} E_1$ and $E_2 \sqcup_{W_2} E_2$.

but correct with ours.

The basic properties of WOW composition used by McNamara and van Willigenburg are summarized in the following lemma.

Lemma 5.3.4 (McNamara–van Willigenburg [43, Lemma 3.19])**.** *Let D be any shape and E be a WOW shape. Then:*

(i) $D \circ_W E$ is a well-defined skew shape.

(ii) $(D \circ_W E)^* \approx D^* \circ_{W^*} E^*$

(iii) If $W \neq \emptyset$, $(D \circ_W E)^T \approx D^* \circ_{W^T} E^T$.

Remark 5.3.5. Most of Lemma 5.3.4 follows more or less immediately from the definitions. The exception is Lemma 5.3.4(iii) in the corner case, which requires the alternative construction of $D \circ_W E$ using the southeast rather than northwest decomposition previously mentioned in Remark 5.3.2. We will not make any explicit use of this construction but in this way we do implicitly depend upon it.

This composition operation is the desired generalization of the ribbon composition, but the naïve extension of Theorem 5.1.17 turns out to be false in general: one more technical assumption will be needed in order to make the composition well-behaved. From a WOW shape E, we consider two additional shapes $\overline{W}$ and $\overline{O}$. The latter is more straightforward: $\overline{O}$ consists of all boxes of O with another box of O to the southeast; in other words this is the shape obtained from O by deleting the southeast border. (Note that we get the same shape by deleting the northwest border instead; we will call this copy $\underline{O}$.)

Consider the amalgamation $E \sqcup_W E$, and let W_{mid} be the "middle" copy of W, where the two copies of E overlap. Then $\overline{W}$ is the shape consisting of all those boxes of W_{mid} which have a box of $E \sqcup_W E$ to the southeast, together with those boxes of $E \sqcup_W E$ which have a box of W_{mid} to the southeast. (Again, we get another copy $\underline{W}$ of the same shape by replacing the word "southeast" with "northwest".) We say that E is *proper* if $\overline{W}$ is not adjacent to either copy of $\overline{O}$ in $E \sqcup E$, or equivalently, $\underline{W}$ is not adjacent to either copy of $\underline{O}$.

Remark 5.3.6. The properness condition is what McNamara and van Willigenburg call *Hypothesis IV*, but their definition is slightly different. Rather than looking at $E \sqcup_W E$ they consider an *infinite* shape $E^{\sqcup_W \infty}$, and define both $\overline{W}$ and the properness condition relative to this shape rather than $E \sqcup_W E$. However, it is clear from the third condition in the definition of WOW shapes that a given copy of W can only share diagonals with the two adjacent copies of O and no other part of $E^{\sqcup_W \infty}$ and that the corresponding copy of $\overline{W}$ can only possibly be adjacent to the copies of $\overline{O}$ in those adjacent copies of O. Thus our definitions are equivalent.

Remark 5.3.7. It is clear from the definitions that $\overline{O}$ and $\underline{O}$ are subshapes of O, whereas $\overline{W}$ and $\underline{W}$ need not be subshapes of W. However, as mentioned in the previous remark, the third condition in the definition of WOW shapes shows that $\overline{W}$ is contained within the union of W_{mid} and the two copies of O in $E \sqcup_W E$. Let us write the first and second copy of E as E_1 and E_2 respectively, and similarly their copies of O as O_1 and O_2. Observe that if W_{sw} is attached horizontally in E, then no box of O can have a box of W_{sw} to its southeast, and so no box of O_2 can be in $\overline{W}$. Thus $\overline{W}$ is contained in E_1 in this case. On the other hand, if W_{ne} is attached vertically, then $\overline{W}$ is contained in E_2 by a similar argument. Similarly, if W_{sw} is attached vertically then $\underline{W}$ is contained in E_1 and if W_{ne} is attached horizontally then $\underline{W}$ is contained in E_2. Putting this together, we see that:

- If E is $W \to O \uparrow W$ then $\overline{W}$ is contained in $E_1 \cap E_2 = W_{\mathrm{mid}}$ and $\underline{W}$ may intersect both copies of O.
- If E is $W \uparrow O \to W$ then $\underline{W}$ is contained in W_{mid} and $\overline{W}$ may intersect both copies of O.
- If E is $W \to O \to W$ then $\overline{W}$ is contained in E_1 and $\underline{W}$ is contained in E_2.
- If E is $W \uparrow O \uparrow W$ then $\underline{W}$ is contained in E_1 and $\overline{W}$ is contained in E_2.

5.3.2 The McNamara–van Willigenburg conjecture

We are finally ready to state the main conjecture of McNamara and van Willigenburg generalizing Theorem 5.1.16 to all shapes. To avoid drowning the quagmire of parentheses found in the statement of the conjecture in [43] we will take the convention that WOW composition always associates to the left, i.e. $D \circ_{W_1} E_1 \circ_{W_2} E_2$ will be interpreted as $(D \circ_{W_1} E_1) \circ_{W_2} E_2$.

Conjecture 5.3.8 (McNamara–van Willigenburg [43, Conjecture 5.7])**.** *Suppose D and D' are skew shapes. Then $D \sim D'$ if and only if there exist factorizations*

$$D = E_0 \circ_{W_1} E_1 \circ_{W_2} \cdots \circ_{W_k} E_k$$

and

$$D' = E'_0 \circ_{W'_1} E'_1 \circ_{W'_2} \cdots \circ_{W'_k} E'_k$$

where E_0 is any shape and E_i is a proper $W_iO_iW_i$ shape for $1 \le i \le k$, and for each i either $E'_i = E_i$ or $E'_i = E^*_i$.

Showing that such factorizations are necessary seems extremely difficult, but even showing that they are sufficient in all cases has remained open since the publication of [43]. Our main result is the following, which implies inductively that the sufficiency direction is true.

Theorem 5.3.9. *If $D \sim D'$ and E is a proper WOW shape, then $D \circ_W E \sim D' \circ_W E$.*

We will spend much of the remainder of the chapter proving this, but let us first show that it does indeed imply sufficiency. (Note that the observation that Theorem 5.3.9 would imply this direction does essentially appear in [43] but not as an explicit result.)

Theorem 5.3.10. *Suppose E_0 is any shape and E_i is a proper $W_iO_iW_i$ shape for $1 \leq i \leq k$, and for each i either $E_i' = E_i$ or $E_i' = E_i^*$. Then*

$$E_0 \circ_{W_1} E_1 \circ_{W_2} \cdots \circ_{W_k} E_k \sim E_0' \circ_{W_1'} E_1' \circ_{W_2'} \cdots \circ_{W_k'} E_k'$$

Proof. Inductively suppose the result has already been proven for shorter factorizations, so we have

$$E_0 \circ_{W_1} E_1 \circ_{W_2} \cdots \circ_{W_{k-1}} E_{k-1} \sim E_0' \circ_{W_1'} E_1' \circ_{W_2'} \cdots \circ_{W_{k-1}'} E_{k-1}'$$

If $E_k' = E_k$ we are immediately done by applying Theorem 5.3.9. Suppose $E_k' = E_k^*$. Then

$$\begin{aligned} E_0 \circ_{W_1} E_1 \circ_{W_2} \cdots \circ_{W_k} E_k &\sim (E_0 \circ_{W_1} E_1 \circ_{W_2} \cdots \circ_{W_k} E_k)^* \\ &= (E_0 \circ_{W_1} E_1 \circ_{W_2} \cdots \circ_{W_{k-1}} E_{k-1})^* \circ_{W_k^*} E_k^* \\ &\sim E_0 \circ_{W_1} E_1 \circ_{W_2} \cdots \circ_{W_{k-1}} E_{k-1} \circ_{W_k^*} E_k^* \\ &\sim E_0' \circ_{W_1'} E_1' \circ_{W_2'} \cdots \circ_{W_{k-1}'} E_{k-1}' \circ_{W_{k^*}} E_k^* \end{aligned}$$

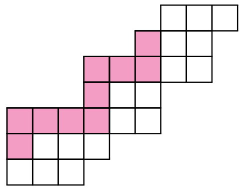
where the steps are respectively by Proposition 5.1.9, Lemma 5.3.4(ii), and then two applications of Theorem 5.3.9 (using the equivalences from Proposition 5.1.9 and the induction hypothesis respectively). □

While McNamara and van Willigenburg were unable to prove their conjecture in general, they did have strong partial results towards it. In particular they say that a WOW shape E satisfies *Hypothesis V* if either E is a corner shape or E is an edge shape such that at least one of W_{ne} and W_{sw} is only adjacent to O in a single box.

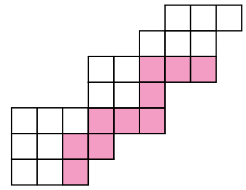
Theorem 5.3.11 (McNamara–van Willigenburg [43, Theorem 3.31])**.** *If $D \sim D'$ and E is a proper WOW shape satisfying Hypothesis V, then $D \circ_W E \sim D' \circ_W E$.*

Thus our contribution is only removing Hypothesis V. However, our techniques are quite different. McNamara and van Willigenburg derive Theorem 5.3.11 from another result [43, Theorem 3.28] which relates the WOW composition to a certain algebraic operation on symmetric functions. They prove the relation using the *Hamel–Goulden identity*, a certain determinantal identity related to ribbon decompositions of skew shapes. Our approach instead uses the ideas from Section 5.2 as well as the combinatorics of Yeats's *key ribbons*

We will discuss key ribbons and Yeats's results in the next subsection, and then we will get down to the business of proving Theorem 5.3.9. In Section 5.4 we will prove it for the edge case; since McNamara and van Willigenburg already proved the corner case this is all that is strictly required to complete the proof. However, our techniques can also be used to give a novel proof of the corner case, which we do in Section 5.5.



(a) top key ribbon



(b) bottom key ribbon

Figure 5.9. The key ribbons of the shape E_1 from Figure 5.5.

5.3.3 Key ribbons

We are not the first to show that Hypothesis V can be removed in some cases. Yeats [63] was able to prove some very special cases of Theorem 5.3.9 in the edge case without it. As previously mentioned in Remark 5.2.3 her results were Hopf-algebraic in nature and were the inspiration for our approach. The essential combinatorial idea used in her results, as well as ours (at least in the edge case) is that of the *key ribbons* of an edge shape.

Let E be a $W \to O \to W$ shape. Thinking of E as the middle copy in $E \sqcup_W E \sqcup_W E$, each of the two copies of W in E has a corresponding $\overline{W}$ and $\underline{W}$. By Remark 5.3.7, $\underline{W}_{\mathrm{sw}}$ is contained in E (as this would be W_{mid} when considering just the first and second copies of E), and similarly (considering the second and third copies) $\overline{W}_{\mathrm{ne}}$ is also contained in E, whereas the other two copies need not be. The subshapes W_{sw}, $\overline{O}$, and $\overline{W}_{\mathrm{ne}}$ are disjoint and cover most of the boxes of E. Consider the subshape consisting of those boxes which are not in any of these. By the definitions of $\overline{O}$ and $\overline{W}$ this shape consists of precisely those boxes of O and W_{ne} that lie in the southeast border of $E \sqcup_W E$. In particular, this shape is a segment of the southeast border of E itself, starting at the first box of $\mathrm{SE}(O)$ and ending at the box in W_{ne} corresponding to the box to the west of where it started. This is the *bottom key ribbon* introduced by Yeats [63]. We denote it by $K_\downarrow(E)$.

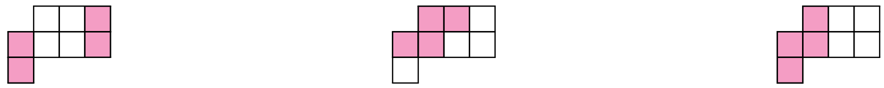
Analogously, the *top key ribbon* $K_\uparrow(E)$ consists of those boxes not contained in any of the disjoint subshapes $\underline{W}_{\mathrm{sw}}$, $\underline{O}$, or W_{ne}. This is the segment of $\mathrm{NW}(E)$ which ends at the last box of $\mathrm{NW}(O)$ and starts at the box of W_{sw} corresponding to the box of W_{ne} to the east of where it ends. We also define the key ribbons for $W \uparrow O \uparrow W$ shapes by the relations $K_\downarrow(E^T) = K_\downarrow(E)^T$ and $K_\uparrow(E^T) = K_\uparrow(E)^T$. (Clearly, these could also be defined explicitly in a similar way to the $W \to O \to W$ case.)

Remark 5.3.12. Yeats did not define the key ribbons in terms of $\overline{W}$ and $\overline{O}$ but only by the explicit definitions as segments of $\mathrm{NW}(E)$ and $\mathrm{SE}(E)$. Our reformulation of the definition, while straightforward, is our first step in synthesizing Yeats's approach with that of McNamara and van Willigenburg.

The two key ribbons are distinct but share many of their properties.

Lemma 5.3.13 (Yeats [63, Lemma 3])**.** *The top and bottom key ribbon have the same size, number of rows, and number of columns.*

Let us write $k(E)$ for the common size of the top and bottom key ribbons and $h(E)$ for their height.



(a) W's highlighted **(b)** top key ribbon **(c)** loose end

Figure 5.10. A $W \to O \to W$ shape with a loose end ribbon.

Remark 5.3.14. From the construction of either key ribbon we see that $k(E)$ is simply the "Manhattan distance" (the sum of the vertical and horizontal distances) from a certain box in W_{sw} to the corresponding box in W_{ne}. Since everything is connected, it is therefore also equal to the distance from *any* box in W_{sw} to the corresponding box in W_{ne}.

It is clear that the top key ribbon is removable on the left and the bottom key ribbon is removable on the right. The parameter $k(E)$ will play the role in the edge case that the size of the ribbon B played in Section 5.2.1. Thus we will be interested in which ribbons of size $k(E)$ are removable from various shapes $D \circ_W B$. Of course the simplest case is taking D to be a single box, i.e. looking at which ribbons are removable from E itself. In general there may be others besides the key ribbons. For $W \to O \to W$ shapes, Yeats defined a *loose end ribbon* to be a ribbon of size $k(E)$ which is either removable on the left and starts before the top key ribbon, or removable on the right and starts after the top key ribbon. (For $W \uparrow O \uparrow W$ shapes we must swap "before" and "after".) Yeats showed that the presence of such ribbons is essentially the only thing that can go wrong.

Lemma 5.3.15 (Yeats [63, Lemma 8])**.** *Suppose λ is a partition and E is a $W \to O \to W$ or $W \uparrow O \uparrow W$ shape with no loose end ribbons. Then:*

(i) The only ribbon of size $k(E)$ which is removable on the left from $\lambda \circ_W E$ is the copy of $K_\uparrow(E)$ in the copy of E corresponding to the northwesternmost box of λ.

(ii) The only ribbon of size $k(E)$ which is removable on the right from $\lambda^* \circ_W E$ is the copy of $K_\downarrow(E)$ in the copy of E corresponding to the southeasternmost box of λ^*.

Remark 5.3.16. Of course, taking λ to be a single box, this also implies that there can be no ribbons of size $k(E)$ removable from E itself other than key ribbons and loose ends. Indeed, this follows from Remark 5.3.14: a segment of $\mathrm{NW}(E)$ that starts in W_{sw} *after* the start of $K_\uparrow(E)$ will end in W_{ne} at a box adjacent to the one corresponding to the one where it started, and thus cannot be removable by a similar argument to the one used to prove Lemma 5.2.6. On the other hand, there can *be* no segment of length $k(E)$ starting in O; it would simply run off the edge. Clearly the same ideas apply to ribbons removable on the right.

Using Lemma 5.3.15 and cocommutativity, Yeats showed the following result.

Theorem 5.3.17 (Yeats [63, Theorem 13])**.** *Let D be the shape obtained by deleting the northwest or southeast corner of a rectangle and let E be a $W \to O \to W$ or $W \uparrow O \uparrow W$ shape with no loose end ribbons. Then $D \circ_W E \sim D^* \circ_W E$.*

Remark 5.3.18. The proof of Theorem 5.3.17 in [63] makes use of some rather complicated combinatorial contortions, but the framework we developed in Section 5.2 allows it to be summarized surprisingly succinctly: let ρ be a rectangle and b its northwesternmost box. Observe that deleting the copy of $K_\uparrow(E)$ in $\rho \circ_W E$ corresponding to b actually leaves behind $(\rho \setminus b) \circ_W E$ and similarly deleting the $K_\downarrow(E)$ from the other corner gives $(\rho \setminus b)^* \circ_W E$. Thus, by Lemma 5.3.15 and the Murnaghan–Nakayama rule,

$$p_{k(E)} \rightharpoonup [\rho \circ_W E] = (-1)^{h(E)}[(\rho \setminus b)^* \circ_W E]$$

and

$$[\rho \circ_W E] \leftharpoonup p_{k(E)} = (-1)^{h(E)}[(\rho \setminus b) \circ_W E]$$

and the result follows by Proposition 5.2.1(i).

While the hypothesis on D is clearly extremely restrictive, Theorem 5.3.17 nonetheless implies some equivalences that don't follow from Theorem 5.3.11. (An example is given in [63, Section 5].) On the other hand, it is not immediately obvious how to relate Theorem 5.3.17 to Theorem 5.3.9. Note that we have not required properness, but in exchange we have the no-loose-ends condition. Addressing this discrepancy will be our first task of the next section.

5.4 The edge case

At last, our preparations are complete and we may begin proving things for the edge case. Note that using the identity $(D \circ_W E)^T = D^* \circ_{W^T} E^T$, we can swap between the two different types of edge shapes easily, so for the combinatorial part of this section we will restrict attention to $W \to O \to W$ shapes. As promised, we start by tying up loose ends and resolving the discrepancy between the hypotheses of Theorem 5.3.17 and Theorem 5.3.9. We are able to do this in the best way possible: it turns out there is no discrepancy at all.

Theorem 5.4.1. *An edge shape is proper if and only if it has no loose end ribbons.*

Proof. Suppose E is a $W \to O \to W$ shape with a loose end ribbon. We will show that E is not proper. Since properness is preserved by antipodal rotation, we may assume that E has a loose end ribbon A which is removable on the left. This ribbon starts at some box $a \in W_{\mathrm{sw}}$. Let a' be the box in W_{ne} corresponding to a. By Remark 5.3.14, A ends on the diagonal before the one containing a', at some box $b \in O$. Note that b cannot be the northeasternmost box of E (since E is $W \to O \to W$) and cannot have a box to its north (since A is removable) so it has a box to its west. Since A is removable it must be that a has no box to its west, and hence also no box to its northwest. Thus a' does not have a box of W_{ne} to its west or northwest. But $a \notin K_\uparrow(E)$, so $a \in \underline{W}_{\mathrm{sw}}$; thus it must be that a' has a box of O to its northwest and hence also one to its west. The latter is on the same diagonal as b, and cannot be equal to b as it has a box to its north. Thus $b \in \overline{O}$, but the box to the west of b is on the same diagonal as a' and hence is in $\overline{W}_{\mathrm{ne}}$, contradicting properness.

For the converse, suppose E is improper. Again, rotating if necessary, we assume $\overline{O}$ and $\overline{W}_{\mathrm{ne}}$ contain adjacent boxes. Clearly no box of $\overline{O}$ can be adjacent to a box of W_{ne} in a

$W \to O \to W$ shape, so any box of $\overline{W}_{\mathrm{ne}}$ which is adjacent to a box of $\overline{O}$ is in $\mathrm{SE}(O)$. Let b be the northeasternmost box of $\overline{O}$ and c be the box to its east, which is in $\mathrm{NW}(O) \cap \mathrm{SE}(O)$. Now, since W is connected, if any box that comes before c in $\mathrm{SE}(O)$ has a box of W on the same diagonal then so does c. On the other hand, any box that comes after c is not adjacent to anything on $\overline{O}$. Since we have assumed there is some adjacency between $\overline{W}_{\mathrm{ne}}$ and $\overline{O}$, it must be the case that $c \in \overline{W}_{\mathrm{ne}}$. Thus c has a box $a' \in W_{\mathrm{ne}}$ to its southeast. Note that to the west of a' is the box southeast of b, which is in O since $b \in \overline{O}$. Let a be the box of W_{sw} corresponding to a' and let A the segment of $\mathrm{NW}(E)$ from a to b. Since a' is on the diagonal after the one containing b, A has size $k(E)$. But a' has a box of O to its west so a has no box to its west, and by construction b has no box to its north, so A is removable. □

We now begin our quest to prove Theorem 5.3.9. We will adapt the strategy we used to prove Proposition 5.2.4. To do this we will need to prove analogues for edge shapes of the various lemmas from Section 5.2.1. We start with an analogue of Lemma 5.2.7. We have already seen a partial result of this nature in Lemma 5.3.15; with Theorem 5.4.1 in hand we can even apply it with the desired hypothesis of properness. However, we need a slightly stronger version.

Lemma 5.4.2. *Let D be a skew shape and E be a proper $W \to O \to W$ shape. The only ribbons of size $k(E)$ removable on the left (resp. right) from $D \circ_W E$ are the copies of $K_\uparrow(E)$ (resp. $K_\downarrow(E)$) coming from boxes of D which are removable on the left (resp. right).*

Proof. Any ribbon removable on the left is contained in $\mathrm{NW}(D \circ_W E)$, so consider what $\mathrm{NW}(D \circ_W E)$ looks like: it consists of $\mathrm{NW}(D) \circ K_\uparrow(E)$ with extra bits at the beginning and end. Any ribbon of size $k(E)$ that starts within the extra bit at the start must end before the corresponding key ribbon and hence is contained within one copy of E, and analogously for ribbons that end within the extra bit at the end. Thus by Theorem 5.4.1 no such ribbon can be removable. Therefore any ribbon of size $k(E)$ removable from $\mathrm{NW}(D \circ_W E)$ is contained in $\mathrm{NW}(D) \circ K_\uparrow(E)$. But by Lemma 5.2.7 these are exactly the copies of $K_\uparrow(E)$ coming from removable boxes as wanted. The argument for removing ribbons on the right is analogous. □

In the ribbon case, if b is a removable box of D and B_b is the corresponding copy of B in $D \circ B$ then $(D \circ B) \setminus B_b = (D \setminus b) \circ B$. In the WOW case the behaviour is more subtle, but what we find is that deleting the copy of $K_\uparrow(E)$ or $K_\downarrow(E)$ corresponding to B will produce $(D \setminus b) \circ_W E$ plus some manageable extra junk. Indeed we see this already when D is just a single box: deleting either key ribbon from E produces $\overline{W} \sqcup \overline{O} \sqcup W$ as already established.

Lemma 5.4.3. *Let D be a shape and b be a box which is removable on the left (resp. right). Let E be a proper $W \to O \to W$ shape, and let S be the shape obtained from $D \circ_W E$ by deleting the copy of $K_\uparrow(E)$ (resp. $K_\downarrow(E)$) corresponding to b. Then S is the edge-disjoint union of:*

- *A copy of $(D \setminus b) \circ_W E$.*
- *A copy of $\overline{O}$, if b does not have a box to its southeast (resp. northwest).*
- *A copy of $\overline{W}$, if b does not have a box to its south (resp. north).*

- *A copy of W, if b does not have a box to its east (resp. west).*

Proof. It is clearly sufficient to prove the version for b removable on the left, as the other version follows by applying the result to D^* and E^*.

First suppose b has a box to its southeast i.e. b is the top left corner of a 2×2 square. Then, in the decomposition $E = \underline{W}_{\mathrm{sw}} \cup \underline{O} \cup W_{\mathrm{ne}} \cup K_{\uparrow}(E)$, we see that W_{ne} overlaps with the copy of E corresponding to the box east of b while $\underline{W}_{\mathrm{sw}}$ and $\underline{O}$ overlap with the copy corresponding to the box southeast of b. Thus $K_{\uparrow}(E)$ is the only part that does not overlap any other copy of E, and deleting it gives $(D \setminus b) \circ_W E$ as wanted.

On the other hand, if b has no box to its southeast, then the corresponding copy of E can overlap only with the copies of E coming from the boxes immediately south and east of b (if these exist). Thus $\underline{O}$ overlaps with no other copy of E, whereas $\underline{W}_{\mathrm{sw}}$ and W_{ne} overlap with the copies coming from the boxes south and east respectively if these exist, and otherwise with nothing. Thus S is the union of the specified pieces. Since E is proper, deleting $K_{\uparrow}(E)$ disconnects $\underline{W}_{\mathrm{sw}}$, $\underline{O}$, and W_{ne}, so the union is edge-disjoint as wanted. □

To apply Lemma 5.4.2 inductively we also need to know that there are no "unexpected" ribbons of size $k(E)$ created when we delete a key ribbon. The next lemma shows this is the case.

Lemma 5.4.4. *Let E be a proper $W \to O \to W$ shape. Then W, $\overline{W}$, and $\overline{O}$ have no removable ribbons of size divisible by $k(E)$.*

Proof. We show there are no removable ribbons of size $jk(E)$ by induction on j. For $j = 1$, since $K_{\downarrow}(E)$ is removable on the right from E, its complement $\overline{W}_{\mathrm{sw}} \sqcup \overline{O} \sqcup W_{\mathrm{ne}}$ is removable on the left. Thus for any ribbon removable on the left from W, $\overline{W}$, or $\overline{O}$ there is a copy of that ribbon removable on the left from E. Since $K_{\uparrow}(E)$ cannot be contained in any of these pieces, any such ribbon of length $k(E)$ would be a loose end.

For $j > 1$, suppose there is a ribbon C of size $jk(E)$ which is removable on the left. Then we can write $C = C' \odot C''$ or $C = C' \cdot C''$ where $|C'| = k(E)$ and $|C''| = (j-1)k(E)$. In the former case C' is removable and in the latter case C'' is removable; either way this is a contradiction. □

We can now prove the analogue of Lemma 5.2.8.

Lemma 5.4.5. *Let D be a skew shape and E be a proper $W \to O \to W$ shape. Then:*

(1) The only ribbons of size divisible by $k(E)$ removable on the left from $D \circ_W E$ are the ribbons $A \circ K_{\uparrow}(E)$ where A is a ribbon removable on the left from D.

(2) The only ribbons of size divisible by $k(E)$ removable on the right from $D \circ_W E$ are the ribbons $A \circ K_{\downarrow}(E)$ where A is a ribbon removable on the right from D.

Proof. We prove (1); the proof of (2) is analogous. As in the proof of Lemma 5.2.8, suppose C is a ribbon of size $jk(E)$ removable on the left. Write $C = C' \odot C''$ or $C = C' \cdot C''$ where $|C'| = k(E)$ and $|C''| = (j-1)k(E)$. First suppose $C = C' \odot C''$. Then C' is removable on the left from $D \circ_W E$ and C'' is removable from $(D \circ_W E) \setminus C'$. By Lemma 5.4.2, C' is a copy of $K_{\uparrow}(E)$ coming from a removable box b of D and by Lemma 5.4.3, $(D \circ_W E) \setminus C'$

is an edge-disjoint union of $(D \setminus b) \circ_W E$ with possibly a copy of W, $\overline{W}$, and/or $\overline{O}$. But by Lemma 5.4.4, the latter shapes do not have ribbons of size divisible by $k(E)$ that are removable on the left, so C'' must be removable on the left from $(D \setminus b) \circ_W E$. Inductively, $C'' = A' \circ K_\uparrow(E)$ for some A' removable on the left from $D \setminus b$ and hence $C = A \circ K_\uparrow(E)$ where $A = b \odot C''$.

Similarly, if $C = C' \cdot C''$ then C'' is removable, so inductively $C'' = A' \circ K_\uparrow(E)$ where A' is removable from D. Then C' is removable from $(D \circ_W E) \setminus C''$ which consists of $(D \setminus A') \circ_W E$ together with some copies of W, $\overline{W}$, and $\overline{O}$. Again by Lemma 5.4.4 none of the latter can have ribbons of size $k(E)$, so C' is removable from $(D \setminus A') \circ_W E$. Thus C' is a copy of $K_\uparrow(E)$ corresponding to a removable box b from $D \setminus A'$, and $C = A \circ K_\uparrow(E)$ where $A = b \cdot A'$. $\square$

We can use Lemma 5.4.3 and Lemma 5.4.5 to compute the action of certain power sums on $[D \circ_W B]$, similarly to Lemma 5.2.10.

Lemma 5.4.6. *Let D be a shape and E be a proper $W \to O \to W$ shape. For any m,*

$$[D \circ_W E] \leftharpoonup p_{mk(E)} = (-1)^{mh(E)} \sum_A (-1)^{\operatorname{ht} A} [W]^{\operatorname{row}(D)-\operatorname{row}(D \setminus A)} [\overline{W}]^{\operatorname{col}(D)-\operatorname{col}(D \setminus A)} \times [\overline{O}]^{\operatorname{dia}(D)-\operatorname{dia}(D \setminus A)} [(D \setminus A) \circ_W E]$$

where A ranges over ribbons removable from D on the left, and $p_{mk(E)} \rightharpoonup [D \circ_W E]$ is given by an identical formula with A ranging over ribbons removable on the right instead.

Proof. By (5.12) and the Murnaghan–Nakayama rule (in the form of Theorem 5.1.14),

$$[D \circ_W E] \leftharpoonup p_{mk(E)} = \sum_{Y \le D \circ_W E} \chi(Y, mk(E)) [(D \circ_W E) \setminus Y]$$

By Corollary 5.1.15, $\chi(Y, mk(E)) = 0$ unless Y is a ribbon of size $mk(E)$. By Lemma 5.4.5, such a ribbon is of the form $X \circ K_\uparrow(E)$ for some ribbon X of size m removable on the left from D. In this case we have

$$\chi(Y, mk(E)) = (-1)^{\operatorname{ht}(X \circ K_\uparrow(E))} = (-1)^{mh(E)+\operatorname{ht} X} = (-1)^{mh(E)} \chi(X, m)$$

and again by the Murnaghan–Nakayama rule, $\chi(X, m)$ is zero for all other $X \le D$. Thus we have (using Lemma 5.2.9)

$$[D \circ_W E] \leftharpoonup p_{mk(E)} = (-1)^{mh(E)} \sum_{X \le D \circ_W E} \chi(X, k(E)) [(D \circ_W E) \setminus (X \circ K_\uparrow(E))].$$

Now, note that $\operatorname{row}(D) - \operatorname{row}(D \setminus X)$ is the number of rows of D completely contained in X, i.e. the number of boxes of X with no box of D to the east. Similarly, $\operatorname{col}(D) - \operatorname{col}(D \setminus X)$ is the number of boxes of X with no box of D to the south and $\operatorname{dia}(D) - \operatorname{dia}(D \setminus X)$ the number with no box of D to the southeast. Thus by Lemma 5.4.3, we have

$$[(D \circ_W E) \setminus (X \circ K_\uparrow(E))] = [W]^{\operatorname{row}(D)-\operatorname{row}(D \setminus X)} [\overline{W}]^{\operatorname{col}(D)-\operatorname{col}(D \setminus X)} [\overline{O}]^{\operatorname{dia}(D)-\operatorname{dia}(D \setminus X)} [(D \setminus X) \circ_W E]$$

as desired. The proof for the left action is analogous. $\square$

While this is in some sense an analogue of Lemma 5.2.10, we must acknowledge that the statement of Lemma 5.4.6 is quite hideous. Beyond aesthetic concerns, there is the more serious problem that we cannot trivially generalize the statement from shapes to arbitrary elements of $\mathcal{S}$, as the right side does not appear to be linear in $[D]$. To address these problems, we introduce an operation we call *modified composition*:

$$[D] \mathbin{\square_W} E = [W]^{|D|-\mathrm{row}(D)}[\overline{W}]^{|D|-\mathrm{col}(D)}[\overline{O}]^{|D|-\mathrm{dia}(D)}[D \circ_W E] \tag{5.13}$$

which we extend linearly to $\mathcal{S}$. This operation will absorb the extra fudge factors that appear on the right-hand side of the formulas in Lemma 5.4.6. Note that (5.13) will be our definition of modified composition for both $W \to O \to W$ and $W \uparrow O \uparrow W$ shapes E. (In the next section we will introduce a slightly different variation for the corner case.) We can now restate Lemma 5.4.6 in a far more pleasing way.

Lemma 5.4.7. *Let D be a shape and E be a proper $W \to O \to W$ shape. For any m we have*

$$p_{mk(E)} \rightharpoonup ([D] \mathbin{\square_W} E) = ((-1)^{h(E)}[W][\overline{W}][\overline{O}])^m((p_m \rightharpoonup [D]) \mathbin{\square_W} E)$$

and

$$([D] \mathbin{\square_W} E) \leftharpoonup p_{mk(E)} = ((-1)^{h(E)}[W][\overline{W}][\overline{O}])^m(([D] \leftharpoonup p_m) \mathbin{\square_W} E).$$

Proof. As usual, the proofs of the two are analogous so we only prove the right action. First suppose E is $W \to O \to W$. By Lemma 5.4.4, $p_{mk} \rightharpoonup [W] = 0 = [W] \leftharpoonup p_{mk}$, and similarly for $\overline{W}$ and $\overline{O}$. Since $p_{mk(E)}$ is primitive, by it acts as a derivation by Proposition 2.2.21. Thus

$$([D] \mathbin{\square_W} E) \leftharpoonup p_{mk(E)} = [W]^{|D|-\mathrm{row}(D)}[\overline{W}]^{|D|-\mathrm{col}(D)}[\overline{O}]^{|D|-\mathrm{dia}(D)}([D \circ_W E] \leftharpoonup p_{mk(E)}).$$

If we expand this using Lemma 5.4.6, the right-hand side becomes

$$\begin{aligned}
&(-1)^{mh(E)}\sum_A (-1)^{\mathrm{ht}\,A}[W]^{|D|-\mathrm{row}(D\setminus A)}[\overline{W}]^{|D|-\mathrm{col}(D\setminus A)}[\overline{O}]^{|D|-\mathrm{dia}(D\setminus A)}[(D \setminus A) \circ_W E] \\
&= (-1)^{mh(E)}\sum_A (-1)^{\mathrm{ht}\,A}[W]^{|D|-|D\setminus A|}[\overline{W}]^{|D|-|D\setminus A|}[\overline{O}]^{|D|-|D\setminus A|}([D \setminus A] \mathbin{\square_W} E) \\
&= ((-1)^{h(E)}[W][\overline{W}][\overline{O}])^m \sum_A (-1)^{\mathrm{ht}\,A}([D \setminus A] \mathbin{\square_W} E) \\
&= ((-1)^{h(E)}[W][\overline{W}][\overline{O}])^m(([D] \leftharpoonup p_m) \mathbin{\square_W} E)
\end{aligned}$$

as wanted. ∎

With this formulation in hand we can prove the analogue of Proposition 5.2.11.

Theorem 5.4.8. *Let f be a symmetric function which is homogeneous of degree m and E be a proper $W \to O \to W$ shape. For any $h \in \mathcal{S}$,*

$$f[p_{k(E)}] \rightharpoonup (h \mathbin{\square_W} E) = ((-1)^{h(E)}[W][\overline{W}][\overline{O}])^m((f \rightharpoonup h) \mathbin{\square_W} E)$$

and

$$(h \mathbin{\square_W} E) \leftharpoonup f[p_{k(E)}] = ((-1)^{h(E)}[W][\overline{W}][\overline{O}])^m((h \leftharpoonup f) \mathbin{\square_W} E).$$

Proof. By linearity it is sufficient to prove it for $f = p_\nu$ for $\nu \vdash m$ and $h = [D]$ for some shape D. This follows by an easy induction from Lemma 5.4.7, again using the fact that $p_{rk(E)}$ acts as a derivation which annihilates $[W]$, $[\overline{W}]$, and $[\overline{O}]$. $\square$

We will also need a $W \uparrow O \uparrow W$ version. It turns out that in this case the two actions are swapped and twisted by the fundamental involution.

Theorem 5.4.9. *Let f be a symmetric function which is homogeneous of degree m and E be a proper $W \uparrow O \uparrow W$ shape. For any $h \in \mathcal{S}$,*

$$f[p_{k(E)}] \rightharpoonup (h \,\Box_W\, E) = ((-1)^{h(E)}[W][\overline{W}][\overline{O}])^m((h \leftharpoonup \omega(f)) \,\Box_W\, E)$$

and

$$(h \,\Box_W\, E) \leftharpoonup f[p_{k(E)}] = ((-1)^{h(E)}[W][\overline{W}][\overline{O}])^m((\omega(f) \rightharpoonup h) \,\Box_W\, E).$$

Proof. Note that E^T is a proper $W^T \to O^T \to W^T$ shape with $k(E^T) = k(E)$ and $h(E^T) = k(E)-h(E)-1$. Moreover, by Lemma 5.3.4(iii) and the fact that antipodal rotation preserves the number of rows, columns, and diagonals of a shape, we have $([D]\Box_W E)^T = ([D^*]\Box_{W^T} E^T)$ and by linearity this also holds for an arbitrary element of $\mathcal{S}$.

We make use of the identity $\omega(f[p_k(E)]) = (-1)^{m(k(E)-1)}\omega(f)[p_{k(E)}]$. (This is due to Alexandersson and Uhlin [5, Lemma 2.13] but can easily be seen by checking it on the power sum basis.) Thus by Proposition 5.2.1(ii) we have

$$\begin{aligned}
(h \,\Box_W\, E) \leftharpoonup f[p_{k(E)}] &= (h^* \,\Box_{W^T}\, E^T)^T \leftharpoonup f[p_{k(E)}] \\
&= \big((h^* \,\Box_{W^T}\, E^T) \leftharpoonup \omega(f[p_{k(E)}])\big)^T \\
&= (-1)^{m(k(E)-1)}\big((h^* \,\Box_{W^T}\, E^T) \leftharpoonup \omega(f)[p_{k(E)}]\big)^T \\
&= \Big((-1)^{k(E)-h(E^T)-1}\Big)^m\Big(([W^T][\overline{W}^T][\overline{O}^T])^m((h^* \leftharpoonup \omega(f)) \,\Box_{W^T}\, E^T)\Big)^T \\
&= ((-1)^{h(E)}[W][\overline{W}][\overline{O}])^m((h^* \leftharpoonup \omega(f))^* \,\Box_W\, E) \\
&= ((-1)^{h(E)}[W][\overline{W}][\overline{O}])^m((\omega(f) \rightharpoonup h) \,\Box_W\, E).
\end{aligned}$$

The proof for the left action is analogous. $\square$

We can now prove our main result, the analogue of Proposition 5.2.4.

Theorem 5.4.10. *Let $h, h' \in \mathcal{S}$ be arbitrary and E be a proper $W \to O \to W$ or $W \uparrow O \uparrow W$ shape. If $h \sim h'$ then $h \,\Box_W\, E \sim h' \,\Box_W\, E$.*

Proof. By Theorem 5.2.2, it is sufficient to prove $(f \rightharpoonup h) \,\Box_W\, E \sim (h \leftharpoonup f) \,\Box_W\, E$ for any $f \in \mathsf{Sym}$ and $h \in \mathcal{S}$. But this is immediate from Theorem 5.4.9 and Theorem 5.4.8. $\square$

Corollary 5.4.11. *Let D and D' be skew shapes and E be a proper $W \to O \to W$ or $W \uparrow O \uparrow W$ shape. If $D \sim D'$ then $D \circ_W E \sim D' \circ_W E$.*

Proof. By Theorem 5.4.10, if $D \sim D'$ then $[D] \,\Box_W\, E \sim [D'] \,\Box_W\, E$. Since equivalent shapes have the same number of rows, columns, and diagonals it follows that $s_{D \circ_W E}$ and $s_{D' \circ_W E}$ are equal up to (the same) powers of s_W, $s_{\overline{W}}$, and $s_{\overline{O}}$. Since multiplication in Sym is cancellative, this implies that $s_{D \circ_W E} = s_{D' \circ_W E}$. $\square$

Combining this with Theorem 5.3.11 completes the proof of Theorem 5.3.9.

Remark 5.4.12. Note that Theorem 5.4.10 does not hold if $\square_W$ is replaced by $\circ_W$, or in other words Corollary 5.4.11 does not hold for arbitrary elements of $\mathcal{S}$. This is because while equivalent *shapes* must have the same number of rows, columns, and diagonals, it is not the case these must agree for all terms in equivalent *sums* of shapes. For instance, there is an equivalence

$$[(2,1)] \sim [(2)][(1)] - [(3)].$$

Theorem 5.4.10 thus gives

$$[W][\overline{W}][(2,1) \circ_W E] \sim [W][E \sqcup_W E][E] - [W]^2[E \sqcup_W E \sqcup_W E]$$

and hence

$$[\overline{W}][(2,1) \circ_W E] \sim [E \sqcup_W E][E] - [W][E \sqcup_W E \sqcup_W E]$$

but we cannot eliminate these remaining factors of $[W]$ and $[\overline{W}]$.

5.5 The corner case

As previously remarked, while proving the edge case is sufficient to complete the proof of Theorem 5.3.9, we are also able to give a novel proof for the corner case using the same method. To do this we need to generalize the ideas of the previous section. Our first step, as in the edge case, is to identify some special ribbons within E. Let E be a $W \to O \uparrow W$ shape; we will be able to use Lemma 5.3.4(ii) to reduce to this case. Recall from Remark 5.3.7 that we have $\overline{W} \leq W$ in this case. Let $K_W = W \setminus \overline{W}$. There are two copies $(K_W)_{\text{sw}}$ and $(K_W)_{\text{ne}}$ within E. Analogously, let $K_O = O \setminus \overline{O} = \text{SE}(O)$. Then within E, the special ribbons appear in the form $(K_W)_{\text{sw}} \odot K_O \cdot (K_W)_{\text{ne}}$. The two ribbons $K_W \odot K_O$ and $K_O \cdot K_W$ will be our analogues of the key ribbons; however, we will also find that the ribbons $K_W \cdot K_O$ and $K_O \odot K_W$ make appearances. We will call both $K_W \cdot K_O$ and $K_W \odot K_O$ *top key ribbons*, and both $K_O \cdot K_W$ and $K_O \odot K_W$ *bottom key ribbons*. All of these have the same size $k(E) = |K_W| + |K_O|$. (Note that, as in the edge case, $k(E)$ is the Manhattan distance from any box of W_{sw} to the corresponding box in W_{ne}.)

Unlike in the edge case, the two key ribbons that appear in E in height by 1 and both are removable on the right rather than one on each side. This will clearly complicate any attempt to find an analogue of Lemma 5.4.7, but we press on nonetheless. We first show that, like in the edge case, proper corner shapes have "no loose ends".

Remark 5.5.1. Consider the shape $E \sqcup_W E$ and the middle copy W_{mid} of W. Then, by its very definition, $\overline{W}_{\text{mid}}$ consists of those boxes of W_{mid} with noting to the southeast; in other words, $(K_W)_{\text{mid}}$ is exactly the portion of $\text{SE}(E \sqcup_W E)$ that comes from W_{mid}. Moreover, if we add another copy of W translated one unit southeast to form $E \cdot_W E$, the "extra" boxes that are added form another copy of K_W one unit southeast of $(K_W)_{\text{mid}}$.

Lemma 5.5.2. *Let E be any $W \to O \uparrow W$ shape. There are no ribbons of size $k(E)$ removable on the left from E.*

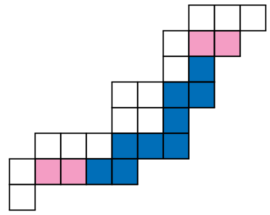


Figure 5.11. The special ribbons in the $W \to O \uparrow W$ shape E_2 from Figure 5.5.

Proof. Since the shape is $W \to O \uparrow W$, on any diagonal containing a box of either copy of W, the northwesternmost box is in that copy of W. Now, suppose A is a ribbon of size $k(E)$ starting at a box $a \in W_{\mathrm{sw}}$ and continuing along the northwest border. Let a' be the box corresponding to a in W_{ne}. Then A ends on the diagonal immediately before the diagonal containing a'; since A runs along the northwest border it ends at a box b to the south or west of a'. In order for the ribbon to be removable, b must be to the west of a', but there must be no box to the west of a. Thus $b \in O$, but this contradicts the shape being $W \to O \uparrow W$. ☐

Lemma 5.5.3. *Let E be a proper $W \to O \uparrow W$ shape. The only ribbons of size $k(E)$ removable on the right from E are $K_{W_{\mathrm{sw}}} \odot K_O$ and $K_O \cdot K_{W_{\mathrm{ne}}}$.*

Proof. Suppose A is such a ribbon. Since $k(E) > |K_O| = \mathrm{dia}(O)$, A cannot be contained entirely in O, so it intersects (at least) one copy of W. Replacing E with E^T if necessary, we may assume without loss of generality that A starts in W_{sw}, say at some box a. Now the two suggested ribbons are the only ones of this size removable from $K_W \odot K_O \cdot K_W$, so we must have $a \in \overline{W}_{\mathrm{sw}}$. Let a' be the box corresponding to a in $\overline{W}_{\mathrm{ne}}$. Then A must end on the diagonal before the one containing a', at some box $b \in O$. Now, a is the start of a ribbon removable on the right, so a cannot have a box to its south, and hence also cannot have one to the southeast. Thus a' does not have a box of W_{ne} to its south or southeast. But since $a' \in \overline{W}_{\mathrm{ne}}$ and a does not have a box of O to its southeast, it must be that a' has a box of O to its southeast and hence also one to its south. The latter is on the same diagonal as b, so is either in $\overline{O}$ or is equal to b. By properness, it must be the latter, so b is immediately south of a', but since a' has a box to its southeast this contradicts the removability of A. ☐

Next, as in the edge case, we wish to look at ribbons of size divisible by $k(E)$ that appear in compositions $D \circ_W E$, as well as what remains when we remove them. First, we will look at the structure of the northwest and southeast border ribbons of these shapes. In the edge case, we essentially had $\mathrm{NW}(D) \circ K_{\uparrow}(E)$ and $\mathrm{SE}(D) \circ K_{\downarrow}(E)$, plus some extra bits at the beginning and the end. In the corner case this is subtler: the two ribbons are quite different and neither one is exactly built by ribbon composition.

Lemma 5.5.4. *Suppose D is a connected shape and E is a $W \to O \uparrow W$ shape. Then*

$$\mathrm{NW}(D \circ_W E) \approx \underbrace{\mathrm{NW}(E) \sqcup_{\mathrm{NW}(W)} \cdots \sqcup_{\mathrm{NW}(W)} \mathrm{NW}(E)}_{\mathrm{dia}(D)\ \textit{copies}}.$$

Proof. Immediate from the definition. ☐

The southeast ribbon is far more complicated.

Lemma 5.5.5. *Suppose D is a connected shape and E is a $W \to O \uparrow W$ shape. Let* $\mathrm{SE}(E) = A \odot K_{W_{\mathrm{sw}}} \odot K_O \cdot K_{W_{\mathrm{ne}}} \cdot B$. *Then*

$$\mathrm{SE}(D \circ_W E) \approx A \odot K_W \vartriangle_0 K_O \blacktriangle_1 K_W \vartriangle_1 K_O \blacktriangle_2 \cdots \vartriangle_{d-1} K_O \blacktriangle_d K_W \cdot B$$

where $d = \mathrm{dia}(D)$ and

- *$\vartriangle_i$ is $\odot$ if $i = 0$ or if the ith and $(i+1)$st boxes of $\mathrm{SE}(D)$ are horizontally adjacent, $\cdot$ if they are vertically adjacent*
- *$\blacktriangle_i$ is $\cdot$ if $i = d$ or if the ith and $(i+1)$st boxes of $\mathrm{NW}(D)$ are horizontally adjacent, $\odot$ if they are vertically adjacent.*

Proof. Let $b_1, \ldots, b_d$ be the boxes of $\mathrm{SE}(D)$ in order. As usual write E_b for the copy of E in $D \circ_W E$ corresponding to a box $b \in D$, and similarly for any subshape of E. The ribbons $(K_O)_{b_i}$ clearly each form a segment of $\mathrm{SE}(D)$. The segment preceding $(K_O)_{b_1}$ comes entirely from E_{b_1} so looks like $A \odot K_W$ and is attached horizontally as in E. Similarly, the segment following $(K_O)_{b_d}$ comes entirely from E_{b_d} so looks like $K_W \cdot B$ and is attached vertically. Thus what remains to show is that, for $1 \le i \le d-1$, the segment in between $(K_O)_{b_i}$ and $(K_O)_{b_{i+1}}$ is always a copy of K_W and that it is attached to the two copies of K_O in the claimed manner. There are four cases to consider:

- If the ith steps of the northwest and southeast border walks are both horizontal, then b_i and b_{i+1} are horizontally adjacent and are in the same ribbon of the northwest decomposition of D. Thus E_{b_i} and $E_{b_{i+1}}$ appear in the form $E_{b_i} \sqcup_W E_{b_{i+1}}$, and hence by Remark 5.5.1 they are joined by a copy of K_W, namely the one from the overlapping copy of W. This corresponds to $(K_W)_{\mathrm{ne}}$ in E_{b_i} and to $(K_W)_{\mathrm{sw}}$ in $E_{b_{i+1}}$, so it is attached vertically to $(K_O)_{b_i}$ and horizontally to $(K_O)_{b_{i+1}}$. Thus this segment looks like $K_O \cdot K_W \odot K_O$.
- If the ith steps of the northwest and southeast border walks are both vertical, then b_i and b_{i+1} are vertically adjacent and are in the same ribbon of the northwest decomposition of D. Thus E_i and E_{i+1} appear in the form $E_{b_i} \cdot_W E_{b_{i+1}}$. Again by Remark 5.5.1, they are joined by a copy of K_W, this time the one consisting of the extra boxes. Since this copy is one unit southeast of the copy in the overlapping W, it is instead attached horizontally to $(K_O)_{b_i}$ and vertically to $(K_O)_{b_{i+1}}$. Thus this segment looks like $K_O \odot K_W \cdot K_O$.
- If the ith step of the northwest border walk is horizontal but the ith step of the southeast border walk is vertical, then b_i and b_{i+1} are vertically adjacent but are not in the same ribbon of the northwest decomposition. Clearly this implies that $b_i \notin \mathrm{NW}(D)$, so let c be the box immediately northwest of b_i. Then c is on the same diagonal as the ith box of $\mathrm{NW}(D)$ and b_{i+1} is immediately to its east, so c precedes b_{i+1} in a ribbon of the northwest decomposition. Thus E_c and $E_{b_{i+1}}$ appear in the form $E_c \sqcup_W E_{b_{i+1}}$ and E_{b_i} appears one unit southwest of E_c. Since there is no box to the east of b_i, the copy of K_W in $(W_{\mathrm{ne}})_{b_i}$ is part of the southeast border and is attached to $(K_O)_{b_i}$ the same way as in E, i.e. vertically. On the other hand, the copy of K_W in $(W_{\mathrm{ne}})_c$

is attached horizontally to $(K_O)_{b_{i+1}}$. Since E_{b_i} is one unit southeast of E_c, the copy of K_W in $(W_{\mathrm{ne}})_{b_i}$ is instead attached vertically to $(K_O)_{b_{i+1}}$, so this segment looks like $K_O \cdot K_W \cdot K_O$.

- If the ith step of the northwest border walk is horizontal but the ith step of the southeast border walk is vertical, then b_i and b_{i+1} are horizontally adjacent but are not in the same ribbon of the northwest decomposition. Clearly this implies that $b_{i+1} \notin \mathrm{NW}(D)$, so let c be the box immediately northwest of b_{i+1}. Then c is on the same diagonal as the $(i+1)$st box of $\mathrm{NW}(D)$ and b_i is immediately to its south, so b_i precedes c in the a ribbon of the northwest decomposition. Thus E_{b_i} and E_c appear in the form $E_{b_i} \cdot_W E_c$ and $E_{b_{i+1}}$ appears one unit southwest of E_c. Since there is no box to the south of b_{i+1}, the copy of K_W in $(W_{\mathrm{sw}})_{b_{i+1}}$ is part of the southeast border and is attached to $(K_O)_{b_{i+1}}$ the same way as in E, i.e. horizontally. On the other hand, the copy of K_W in $(W_{\mathrm{ne}})_c$ is attached vertically to $(K_O)_{b_i}$. Since $E_{b_{i+1}}$ is one unit southeast of E_c, the copy of K_W in $(W_{\mathrm{ne}})_{b_i}$ is instead attached horizontally to $(K_O)_{b_{i+1}}$, so this segment looks like $K_O \odot K_W \odot K_O$.

We see that all four cases match the formula. □

We are now nearly ready to establish the analogue of Lemma 5.4.2. We first introduce some terminology that will make life easier when applying Lemma 5.5.5. Let us say that a box $b \in D$ is *eastbound* (resp. *northbound*) if it has a box to its east (resp. north) in the same ribbon of the northwest decomposition. Let us also say that the final box of each ribbon of the northwest decomposition is eastbound. Note that if a box is northbound or eastbound then so are all boxes on the same diagonal. (For boxes that are not the final box of their ribbon this basically follows immediately from the definition of the northwest decomposition, whereas for those that are it follows from the fact that every ribbon of the northwest decomposition ends on the eastern border.)

Dually, we will say that a box is *westbound* (resp. *southbound*) if it has a box to its west (resp. south) in same ribbon of the southeast decomposition, and we will also consider the first box in each ribbon of the southeast decomposition to be westbound. This terminology implies that a box is westbound (resp. southbound) in D if and only if it is eastbound (resp. northbound) in D^*.

Remark 5.5.6. The concepts of northbound, eastbound, southbound, and westbound boxes originate in the theory of *outside decompositions* introduced by Hamel and Goulden [30]. The Hamel–Goulden results can be stated in terms of one of the two pairs of notions or the other: Hamel and Golden originally used the terms "approached from below" and "approached from the left" for what we call southbound and westbound boxes, whereas when McNamara and van Willigenburg apply their work in [43] they instead use the other pair, saying that boxes "go north" or "go east" matching our northbound/eastbound terminology. For us it is useful to have names for all four concepts despite the slight terminological awkwardness it leads to.

Lemma 5.5.7. *Let D be any shape and E be a proper $W \to O \uparrow W$ shape. Then:*

(i) For each box b removable from D on the right, there is a bottom key ribbon removable on the right from $D \circ_W E$. This key ribbon is congruent to $K_O \cdot K_W$ if b is eastbound and $K_O \odot K_W$ if b is northbound, where the copy of K_O is the one in the copy of E corresponding to b.

(ii) For each box b removable from D on the left, there is a top key ribbon removable on the right from $D \circ_W E$. This key ribbon is congruent to $K_W \cdot K_O$ if b is southbound and $K_W \odot K_O$ if b is westbound, where the copy of K_O is the one in the copy of E corresponding to the box of $\mathrm{SE}(D)$ *on the same diagonal as b.*

(iii) There are no other ribbons of size $k(E)$ removable from $D \circ_W E$. (In particular, there are none at all removable on the left.)

Proof. Clearly it is sufficient to prove these in the case that D is connected; the result follows for disconnected shapes by applying it to each component separately. Thus we may use Lemma 5.5.4 and Lemma 5.5.5.

(i) Suppose b is the ith box of $\mathrm{SE}(D)$. Then since b is removable on the right, the $(i-1)$st box of $\mathrm{SE}(D)$ is to its west (or $i = 1$) and the $(i+1)$st is to its north (or $i = \mathrm{dia}(D)$). Thus, by Lemma 5.5.5, $K_O \blacktriangle_i K_W$ is removable on the right, where $\blacktriangle_i$ is $\cdot$ if $i = \mathrm{dia}(D)$ or if the ith and $(i+1)$st boxes of $\mathrm{NW}(D)$ are horizontally adjacent—in other words if b is eastbound—and $\odot$ otherwise.

(ii) Suppose b is the ith box of $\mathrm{NW}(D)$. Then since b is removable on the right, the $(i-1)$st box of $\mathrm{SE}(D)$ is to its south (or $i = 1$) and the $(i+1)$st is to its east (or $i = \mathrm{dia}(D)$). Thus, by Lemma 5.5.5, $K_W \vartriangle_{i-1} K_O$ is removable on the right, where $\vartriangle_{i-1}$ is $\cdot$ if $i = 1$ or if the ith and $(i-1)$st boxes of $\mathrm{SE}(D)$ are horizontally adjacent—in other words if b is westbound—and $\odot$ otherwise.

(iii) By Lemma 5.5.4, a segment of $\mathrm{NW}(D \circ_W E)$ of size $k(E)$ is contained within a single copy of E, so cannot be removable by Lemma 5.5.2. Thus there are no ribbons of size $k(E)$ removable on the left.

By Lemma 5.5.5, the ribbons $K_W \vartriangle_i K_O$ or $K_O \blacktriangle_i K_W$ are not removable when the ith box of the relevant border ribbon of D is not removable. Any segment of $\mathrm{SE}(D \circ_W E)$ of size $k(E)$ that starts strictly within a copy of K_O cannot be removable, as it would end at the box in the next copy of K_O corresponding to the box preceding it, and similarly for a ribbon that starts strictly within one of the copies of K_W. This only leaves ribbons that are entirely contained within the first or last copy of E, but these are not removable by Lemma 5.5.3. □

Next we establish the analogue of Lemma 5.4.3.

Lemma 5.5.8. *Let D be a shape and b be a box which is removable on the left (resp. right) from D. Let E be a proper $W \to O \uparrow W$ shape, and let S be the shape obtained from $D \circ_W E$ by deleting the key ribbon corresponding to b via Lemma 5.5.7. Then S is the edge-disjoint union of:*

- *A copy of $(D \setminus b) \circ_W E$.*

- *A copy of $\overline{O}$, if b does not have a box to its southeast (resp. northwest).*
- *A copy of $\overline{W}$, if b does not have a box to its south (resp. north).*
- *A copy of W, if b does not have a box to its east (resp. west).*

Proof. First suppose b is removable on the right. Deleting b does not change the northwest decomposition of D other than by removing b from its ribbon (which may split it into two ribbons) so $(D \setminus b) \circ_W E$ just looks like $D \circ_W E$ but without including E_b in the construction. Now, the key ribbon consists of the copy of K_O in E_b followed by a copy of K_W. Since b has no box to its southeast, this K_O does not overlap with anything. As for the K_W there are two possibilities:

- If b is northbound, it must have a box b_{n} to the north. In this case the relevant copy of K_W the extra K_W that comes from the fact that the corresponding copies of E appear as $E_b \cdot_W E_{b_{\mathrm{n}}}$. This clearly does not overlap with any other part of the shape.
- If b is eastbound, the relevant K_W is the copy of $(K_W)_{\mathrm{ne}}$ in E_b. This cannot overlap with the copy of E coming from a box to the north on a different ribbon, and b cannot have a box to the west, so again this copy of K_W does not overlap with any other part of the shape.

Thus deleting the key ribbon leaves all the boxes of the subshape $(D \setminus b) \circ_W E$ intact. It remains to understand what else is left behind. Note that since b is removable on the right, there cannot be a box to the south of b, so there cannot be any extra boxes coming from the $\cdot_W$ operation in the southwestern part of the shape. On the other hand, any extra boxes coming from the $\cdot_W$ operation in the northeastern part of the shape are part of the key ribbon as we have already touched on. Thus any other remaining boxes must actually come from E_b. Note K_O is always part of the key ribbon so we must only divine the fates of $\overline{O}$ and the two copies of W. We may observe that:

- If b has a box b_{nw} to its northwest, then $\overline{O}$ in E_b completely overlaps with $E_{b_{\mathrm{nw}}}$ (and hence with $(D \setminus b) \circ_W E$). If b does not have a box to the northwest then $\overline{O}$ in E_b cannot overlap with anything so appears in S disjointly from $(D \setminus b) \circ_W E$.
- If b is northbound, then it has a box b_{n} to its north and W_{ne} in E_b completely overlaps with $E_{b_{\mathrm{n}}}$.
- If b is eastbound then $(K_W)_{\mathrm{ne}}$ is part of the key ribbon so does not appear in S. If b has a box b_{n} to its north then $\overline{W}_{\mathrm{ne}}$ in E_b completely overlaps with $E_{b_{\mathrm{n}}}$ while if b does not have a box to its north then $\overline{W}_{\mathrm{ne}}$ does not overlap with anything and appears in S disjointly from $(D \setminus b) \circ_W E$.
- If b has an eastbound box b_{w} to its west then W_{sw} in E_b overlaps completely with $E_{b_{\mathrm{w}}}$.
- If b has a northbound box b_{w} to its west then b also has a box b_{nw} to its northwest. Then $\overline{W}_{\mathrm{sw}}$ in E_b overlaps completely with E_{nw} while $(K_W)_{\mathrm{sw}}$ overlaps with the extra K_W coming from $E_{\mathrm{w}} \cdot_W E_{\mathrm{nw}}$.

- If b has no box to its west then W_{sw} in E_b does not overlap with anything and so appears in S disjointly from $(D \setminus b) \circ_W E$.

Thus the shape does indeed consist of the disjoint pieces listed in the statement. The union is edge-disjoint by properness. The result for boxes removable on the left follows from Lemma 5.3.4(iii), by applying the result for boxes removable on the right to $D^* \circ_{W^T} E^T$ and taking the transpose. (Note that here we are really harnessing the nontriviality of this result for corner shapes to save work!) □

We also need an analogue of Lemma 5.4.4. In fact, we can show something far stronger.

Lemma 5.5.9. *If E is a $W \to O \uparrow W$ shape then $k(E) > \max(\mathrm{dia}(W), \mathrm{dia}(O))$. In particular, none of W, $\overline{W}$, O, or $\overline{O}$ can contain any ribbon of size $k(E)$ or greater, removable or otherwise.*

Proof. By construction, $k(E) > |K_O| = \mathrm{dia}(O)$. For W, consider any ribbon A in E of size $k(E)$ starting at some box $a \in W_{\mathrm{sw}}$. Then A ends one diagonal before the one containing the box of W_{ne} corresponding to a. Since there must be a diagonal strictly between the two copies of W, it follows that A cannot end in W_{sw}. Since $\overline{W} \subseteq W$ and $\overline{O} \subseteq O$ the result follows for them as well. □

Before we can state the analogue of Lemma 5.4.5, we will need a small result relating the left and right aspects of Lemma 5.5.7. Denote by $\xi(D)$ the number of northbound boxes of D. Clearly $\xi(D^*)$ is the number of southbound boxes.

Lemma 5.5.10. *For any shape D, we have $\xi(D) = \xi(D^*)$.*

Proof. Within a given ribbon of the northwest decomposition, the northbound boxes are precisely the last box of each row other than the top row. Thus

$$\xi(D) = \sum_{A \in \mathbf{NW}(D)} \mathrm{ht}\, A$$

and

$$\xi(D^*) = \sum_{A \in \mathbf{SE}(D)} \mathrm{ht}\, A.$$

Now we show by induction that these are equal. It is sufficient to show this in the case of connected shapes (since both clearly sum over components) so $\mathrm{NW}(D)$ and $\mathrm{SE}(D)$ are ribbons. Thus we can write

$$\xi(D) = \mathrm{ht}\,\mathrm{NW}(D) + \xi(D \setminus \mathrm{NW}(D))$$

and

$$\xi(D^*) = \mathrm{ht}\,\mathrm{SE}(D) + \xi((D \setminus \mathrm{SE}(D))^*).$$

But note that $\mathrm{NW}(D)$ and $\mathrm{SE}(D)$ are ribbons which start and end at the same boxes, so they have the same height. On the other hand $D \setminus \mathrm{NW}(D) \approx D \setminus \mathrm{SE}(D)$, so inductively we may assume $\xi(D \setminus \mathrm{NW}(D)) = \xi((D \setminus \mathrm{SE}(D))^*)$ and the result follows. □

With this new statistic in hand we can state the result. As expected, the ribbons of size divisible by $k(E)$ removable from $D \circ_W E$ are determined by ribbons removable from D. However, the correspondence is once again not exactly ribbon composition. As a consequence, the heights of the corresponding ribbons—which we must understand in order to relate this to the action of power sums—are more subtle in this case but can be expressed using the ξ statistic. For convenience of notation let us write $h(E) = \operatorname{ht} K_W + \operatorname{ht} K_O + 1$; this is of course the height of the key ribbons $K_W \odot K_O$ and $K_O \odot K_W$, but the other two types have height $h(E) - 1$ instead.

Lemma 5.5.11. *Let D be a shape and E a proper $W \to O \uparrow W$ shape. Then:*

(i) For each ribbon A removable from D on the right, there is a ribbon $\widetilde{A}$ of size $|A|k(E)$ removable on the right from $D \circ_W E$, consisting of the union of the key ribbons corresponding to the boxes of A by Lemma 5.5.7(i). Moreover,

$$\operatorname{ht} \widetilde{A} = |A|h(E) + \operatorname{ht} A - \xi(D) + \xi(D \setminus A).$$

(ii) For each ribbon A removable from D on the left, there is a ribbon $\widetilde{A}$ of size $|A|k(E)$ removable on the right from $D \circ_W E$, consisting of the union of the key ribbons corresponding to the boxes of A by Lemma 5.5.7(ii). Moreover,

$$\operatorname{ht} \widetilde{A} = |A|h(E) - \operatorname{ht} A + \xi(D) - \xi(D \setminus A) - 1.$$

(iii) There are no other ribbons of size divisible by $k(E)$ removable from $D \circ_W E$. (In particular, there are none at all removable on the left.)

Proof. (i) It is clear from Lemma 5.5.7 that the described ribbon is removable. Suppose A consists of boxes $b_p, \dots, b_q$ where $b_1, \dots, b_{\operatorname{dia}(D)}$ are the boxes of $\operatorname{SE}(D)$. Then in the notation of Lemma 5.5.5,

$$\widetilde{A} \approx K_O \blacktriangle_p K_W \vartriangle_p \cdots \vartriangle_{q-1} K_O \blacktriangle_q K_W.$$

Here $\blacktriangle_i$ is $\cdot$ if b_i is eastbound and $\odot$ if b_i is northbound (in D), while $\vartriangle_i$ is $\odot$ if boxes b_i and b_{i+1} are horizontally adjacent and $\cdot$ if b_i and b_{i+1} are vertically adjacent. By Lemma 5.2.9, the height of $\widetilde{A}$ is $|A|(\operatorname{ht} K_O + \operatorname{ht} K_W)$ plus the number of $\cdot$ operations that appear. The number of i such that b_i and b_{i+1} are vertically adjacent is $\operatorname{ht} A$ (see the proof of Lemma 5.2.9). The number of boxes of A which are northbound in D is $\xi(D) - \xi(D \setminus A)$ as deleting boxes on the right does not change whether other boxes are northbound or eastbound. Thus the number of boxes which are eastbound is $|A| - \xi(D) + \xi(D \setminus A)$. Thus

$$\begin{aligned}\operatorname{ht} \widetilde{A} &= |A|(\operatorname{ht} K_O + \operatorname{ht} K_W) + \operatorname{ht} A + |A| - \xi(D) + \xi(D \setminus A) \\ &= |A|(\operatorname{ht} K_O + \operatorname{ht} K_W + 1) + \operatorname{ht} A - \xi(D) + \xi(D \setminus A) \\ &= |A|h(E) + \operatorname{ht} A - \xi(D) + \xi(D \setminus A).\end{aligned}$$

(ii) Again it is clear that the ribbon is removable. Suppose A consists of $b'_p, \dots, b'_q$ where $b'_1, \dots, b'_{\mathrm{dia}(D)}$ are the boxes of $\mathrm{NW}(D)$ in order. In this case we have

$$\widetilde{A} \approx K_W \, \triangle_{p-1} \, K_O \, \blacktriangle_p \cdots \blacktriangle_{q-1} \, K_W \, \triangle_{q-1} \, K_O$$

where $\triangle_{i-1}$ is $\odot$ if b'_i is westbound and $\cdot$ if b'_i is southbound (in D), while $\blacktriangle_i$ is $\cdot$ if b'_i and b'_{i+1} are horizontally adjacent and $\odot$ if they are vertically adjacent. The number of i for which b'_i and b'_{i+1} are vertically adjacent is $\mathrm{ht}\, A$, so the number for which they are horizontally adjacent is $|A| - 1 - \mathrm{ht}\, A$. The number of boxes of A which are northbound in D is $\xi(D) - \xi(D \setminus A)$ since deleting boxes on the left does not change whether other boxes are westbound or southbound. Thus

$$\begin{aligned} \mathrm{ht}\, \widetilde{A} &= |A|(\mathrm{ht}\, K_W + \mathrm{ht}\, K_O) + |A| - 1 - \mathrm{ht}\, A + \xi(D) - \xi(D \setminus A) \\ &= |A|(\mathrm{ht}\, K_W + \mathrm{ht}\, K_O + 1) - \mathrm{ht}\, A + \xi(D) - \xi(D \setminus A) - 1 \\ &= |A| h(E) - \mathrm{ht}\, A + \xi(D) - \xi(D \setminus A) - 1. \end{aligned}$$

(iii) Follows inductively from Lemma 5.5.7(iii) and Lemma 5.5.9 analogously to the proof of Lemma 5.4.5. □

It is now time to relate the combinatorics we have worked out to the actions of symmetric functions on shapes. Rather than first prove an ugly result as we did in the edge case with Lemma 5.4.6, we will skip ahead to defining the appropriate variation of modified composition. Up to a sign, this will be the same as in the edge case:

$$[D] \,\square_W\, E = (-1)^{\xi(D)} [W]^{|D| - \mathrm{row}(D)} [\overline{W}]^{|D| - \mathrm{col}(D)} [\overline{O}]^{|D| - \mathrm{dia}(D)} [D \circ_W E] \tag{5.14}$$

and we again extend this linearly to $\mathcal{S}$. (Note that unlike the edge case, the presence of this sign factor makes it really matter that we are defining this as an operation on the shape Hopf algebra rather than simply on shapes!) We are now ready to state the analogue of Lemma 5.4.7.

Lemma 5.5.12. *Let D be a shape and E be a proper $W \to O \uparrow W$ shape. For any m we have*

$$p_{mk(E)} \rightharpoonup ([D] \,\square_W\, E) = ((-1)^{h(E)} [W][\overline{W}][\overline{O}])^m ((p_m \rightharpoonup [D] - [D] \leftharpoonup p_m) \,\square_W\, E)$$

and

$$([D] \,\square_W\, E) \leftharpoonup p_{mk(E)} = 0.$$

Proof. Since $p_{mk(E)}$ acts as a derivation (by Proposition 2.2.21) and annihilates $[W]$, $[\overline{W}]$, and $[\overline{O}]$ (by Lemma 5.5.9) we have

$$p_{mk(E)} \rightharpoonup ([D] \,\square_W\, E) = (-1)^{\xi(D)} [W]^{|D| - \mathrm{row}(D)} [\overline{W}]^{|D| - \mathrm{col}(D)} [\overline{O}]^{|D| - \mathrm{dia}(D)} (p_{mk(E)} \rightharpoonup [D \circ_W E])$$

and similarly for the right action. In the case of the right action this immediately gives that $([D] \,\square_W\, E) \leftharpoonup p_{mk(E)} = 0$ since $D \circ_W E$ has no ribbons of size divisible by $k(E)$ removable on the left. On the other hand,

$$p_{mk(E)} \rightharpoonup ([D] \circ_W E) = \sum_{\widetilde{A}} (-1)^{\mathrm{ht}\, \widetilde{A}} [(D \circ_W E) \setminus \widetilde{A}]$$

summing over ribbons $\widetilde{A}$ of size $mk(E)$ removable on the right from $D \circ_W E$. By Lemma 5.5.11 these are in bijection to ribbons A of size m removable on the left or right from D. By Lemma 5.5.8, we have

$$[(D \circ_W E) \setminus \widetilde{A}] = [W]^{\mathrm{row}(D)-\mathrm{row}(D\setminus A)}[\overline{W}]^{\mathrm{col}(D)-\mathrm{col}(D\setminus A)}[\overline{O}]^{\mathrm{dia}(D)-\mathrm{dia}(D\setminus A)}[(D \setminus A) \circ_W E]$$

(where the exponents come from counting boxes as in the proof of Lemma 5.4.6). Thus the overall contribution of A to $p_{mk(E)} \rightharpoonup ([D] \square_W E)$ is

$$\begin{aligned} &(-1)^{\xi(D)+\mathrm{ht}\,\widetilde{A}}[W]^{|D|-\mathrm{row}(D)}[\overline{W}]^{|D|-\mathrm{col}(D)}[\overline{O}]^{|D|-\mathrm{dia}(D)}[(D \circ_W E) \setminus \widetilde{A}] \\ &= (-1)^{\xi(D)+\mathrm{ht}\,\widetilde{A}}[W]^{|D|-\mathrm{row}(D\setminus A)}[\overline{W}]^{|D|-\mathrm{col}(D\setminus A)}[\overline{O}]^{|D|-\mathrm{dia}(D\setminus A)}[(D \setminus A) \circ_W E] \\ &= (-1)^{\xi(D)-\xi(D\setminus A)+\mathrm{ht}\,\widetilde{A}}([W][\overline{W}][\overline{O}])^m([D \setminus A] \square_W E). \end{aligned}$$

Using the formulas for $\mathrm{ht}\,\widetilde{A}$ from Lemma 5.5.11, this becomes

$$\pm(-1)^{mh(E)+\mathrm{ht}\,A}([W][\overline{W}][\overline{O}])^m([D \setminus A] \square_W E)$$

where $\pm$ is a plus for A removable on the right and a minus for A removable on the left. Thus the sum over A removable on the right gives $((-1)^{h(E)}[W][\overline{W}][\overline{O}])^m(p_m \rightharpoonup [D]) \square_W E$ and the sum over A removable on the left gives $-((-1)^{h(E)}[W][\overline{W}][\overline{O}])^m([D] \leftharpoonup p_m) \square_W E$. The result follows. $\square$

Unlike the edge case, we cannot extrapolate to a version of Theorem 5.4.8 here. Nonetheless, this will be enough to prove the main result. In particular, note that since the left sides of the two formulas are equivalent, looking at the right sides gives $(p_m \rightharpoonup [D] - [D] \leftharpoonup p_m) \square_W E \sim 0$, i.e. $(p_m \rightharpoonup [D]) \square_W E \sim ([D] \leftharpoonup p_m) \square_W E$. We will also need the $W \uparrow O \to W$ version. When E is $W \uparrow O \to W$ let us define $k(E) = k(E^*)$ and $h(E) = h(E^*)$.

Lemma 5.5.13. *Let D be a shape and E be a proper $W \uparrow O \to W$ shape. For any m we have*

$$p_{mk(E)} \rightharpoonup ([D] \square_W E) = 0$$

and

$$([D] \square_W E) \leftharpoonup p_{mk(E)} = ((-1)^{h(E)}[W][\overline{W}][\overline{O}])^m(([D] \leftharpoonup p_m - p_m \rightharpoonup [D]) \square_W E).$$

Proof. Note that E^* is a proper $W^* \to O^* \uparrow W^*$ shape. By Lemma 5.3.4(ii) and Lemma 5.5.10, along with the fact that antipodal rotation preserves the numbers of rows, columns, and diagonals we have $([D] \square_W E)^* = [D^*] \square_{W^*} E^*$. Thus

$$\begin{aligned} p_{mk(E)} \rightharpoonup ([D] \square_W E) &= p_{mk(E)} \rightharpoonup ([D^*] \square_{W^*} E^*)^* && \text{by Proposition 5.2.1(iii)} \\ &= (([D^*] \square_{W^*} E^*) \leftharpoonup p_{mk(E)})^* \\ &= 0 && \text{by Lemma 5.5.12} \end{aligned}$$

and similarly

$$([D] \square_W E) \leftharpoonup p_{mk(E)} = ([D^*] \square_{W^*} E^*)^* \leftharpoonup p_{mk(E)}$$

$$
\begin{aligned}
&= (p_{mk(E)} \rightharpoonup ([D^*] \,\square_{W^*}\, E^*))^* \\
&= \Big(((-1)^{h(E)}[W^*][\overline{W}^*][\overline{O}^*])^m ((p_m \rightharpoonup [D^*] - [D^*] \leftharpoonup p_m) \,\square_{W^*}\, E^*) \Big)^* \\
&= ((-1)^{h(E)}[W][\overline{W}][\overline{O}])^m ((p_m \rightharpoonup [D^*] - [D^*] \leftharpoonup p_m)^* \,\square_W\, E) \\
&= ((-1)^{h(E)}[W][\overline{W}][\overline{O}])^m (([D] \leftharpoonup p_m - p_m \rightharpoonup [D]) \,\square_W\, E). \qquad \square
\end{aligned}
$$

We now arrive at the main result, the corner analogue of Theorem 5.4.10.

Theorem 5.5.14. *Let $h, h' \in \mathcal{S}$ be arbitrary and E be a proper $W \to O \uparrow W$ or $W \uparrow O \to W$ shape. If $h \sim h'$ then $h \,\square_W\, E \sim h' \,\square_W\, E$.*

Proof. By Theorem 5.2.2 it is sufficient to prove $(f \rightharpoonup h) \,\square_W\, E \sim (h \leftharpoonup f) \,\square_W\, E$ for all $f \in \mathsf{Sym}$ and $h \in \mathcal{S}$. By linearity it is sufficient to prove this when $f = p_\nu$ for some partition ν. It follows from Lemma 5.5.12 and Lemma 5.5.13 that $(p_m \rightharpoonup h) \,\square_W\, E \sim (h \leftharpoonup p_m) \,\square_W\, E$. We prove it for general partitions ν by induction on the length of ν. Let $\nu' = (\nu_2, \ldots, \nu_{\ell(\nu)})$ and suppose we have already established that $(p_{\nu'} \rightharpoonup h) \,\square_W\, E \sim (h \leftharpoonup p_{\nu'}) \,\square_W\, E$ for all $h \in \mathcal{S}$. Then for any $h \in \mathcal{S}$ we have

$$
\begin{aligned}
p_\nu \rightharpoonup h &= p_{\nu_1} \rightharpoonup (p_{\nu'} \rightharpoonup h) \\
&\sim p_{\nu'} \rightharpoonup h \leftharpoonup p_{\nu_1} \\
&\sim (h \leftharpoonup p_{\nu_1}) \leftharpoonup p_{\nu'} \\
&= h \leftharpoonup p_\nu
\end{aligned}
$$

as desired. $\square$

This implies the corner case of Theorem 5.3.9, thus completing our proof.

5.6 Algebraic composition

Our final task is to relate our work to McNamara and van Willigenburg's original approach to their conjecture. Firstly, let us observe that the following result is an immediate consequence of of Theorem 5.4.10 and Theorem 5.5.14.

Theorem 5.6.1. *Let E be a proper WOW shape. There exists a unique algebra morphism* $\mathsf{Sym} \to \mathsf{Sym}$ *such that $s_D \mapsto \mathsf{schur}([D] \,\square_W\, E)$ for all skew shapes D.*

We will denote this map by $f \mapsto f \,\square_W\, s_E$. On the other hand, in [43], McNamara and van Willigenburg consider a certain *nonlinear* map $\mathsf{Sym} \to \mathsf{Sym}$ which they denote $f \mapsto f \circ_W s_E$ and which plays a central role in their approach to the skew equivalence problem. The purpose of this section is to show that these operations are essentially equivalent.

First we give the definition of their operation. Given a symmetric function f, we write it as a polynomial in the complete symmetric functions. We then minimally homogenize this polynomial by adding "h_0" factors to each term to match the maximum number of factors that appear in any term. (This is the nonlinear part.) Finally, we map $h_n \mapsto s_{(n) \circ_W E}$ for $n > 0$ and $h_0 \mapsto s_W$.[11] Their main result (rewritten in our notation and terminology) is the following.

[11] We note that while the latter is consistent with the convention $\emptyset \circ_W E = W$, this construction still sends $s_\emptyset = 1$ to 1, which is not. This is the source of the error mentioned in Remark 5.3.3.

Theorem 5.6.2 (McNamara–van Willigenburg [43, Theorem 3.28])**.** *Let D be a shape and E a be a proper WOW shape satisfying Hypothesis V. Then*

$$s_D \circ_W s_E = \pm s_{\overline{W}}^{|D|-\mathrm{col}(D)} s_{\overline{O}}^{|D|-\mathrm{dia}(D)} s_{D \circ_W E}$$

where the sign is a plus in the edge case and depends only on D in the corner case.

It was using Theorem 5.6.2 that McNamara and van Willigenburg proved Theorem 5.3.11, and they conjectured [43, Conjecture 3.26] that Hypothesis V can be removed, implying Theorem 5.3.9. Though we have already proved the latter by other means, it turns out that we can also prove this conjecture. Indeed, it follows easily from Theorem 5.6.1. First, we show how the two operations are related in general.

Proposition 5.6.3. *Let f be a symmetric function of degree n and let r be the maximum number of factors that appears in any term of the expansion of f in complete symmetric functions. Then for any proper WOW shape E,*

$$f \,\square_W\, s_E = s_{\overline{W}}^{n-r}(f \circ_W s_E).$$

Proof. Note that

$$h_n \,\square_W\, s_E = \mathsf{schur}((n) \,\square_W\, s_E) = s_{\overline{W}}^{n-1} s_{(n)\circ_W E} = s_{\overline{W}}^{n-1}(h_n \circ_W s_E).$$

Thus, since both operations are multiplicative on complete symmetric functions, if we write

$$f = \sum_\lambda a_\lambda h_\lambda$$

then

$$f \,\square_W\, s_E = \sum_\lambda a_\lambda s_{\overline{W}}^{n-\ell(\lambda)}(h_\lambda \circ_W s_E)$$

whereas by definition

$$f \circ_W s_E = \sum_\lambda a_\lambda s_{\overline{W}}^{r-\ell(\lambda)}(h_\lambda \circ_W s_E).$$

The result follows. ∎

Note that it follows from the Jacobi–Trudi formula [56, Theorem 7.16.1] that the maximum number of factors that appear in a term of the h-expansion of s_D is $\mathrm{row}(D)$. Thus comparing Proposition 5.6.3 with (5.13) and (5.14) gives the following result.

Theorem 5.6.4. *Let D be a shape and E be a proper WOW shape. Then*

$$s_D \circ_W s_E = \pm s_{\overline{W}}^{|D|-\mathrm{col}(D)} s_{\overline{O}}^{|D|-\mathrm{dia}(D)} s_{D \circ_W E}$$

where the sign is a plus in the edge case and equals $(-1)^{\xi(D)}$ in the corner case.

Note that in addition to removing Hypothesis V, our result also comes with the slight improvement over Theorem 5.6.2 of giving a reasonably natural interpretation of the sign that appears in the corner case.

References

[1] Marcelo Aguiar, Carlos André, Carolina Benedetti, Nantel Bergeron, Zhi Chen, Persi Diaconis, Anders Hendrickson, Samuel Hsiao, I. Martin Isaacs, Andrea Jedwab, Kenneth Johnson, Gizem Karaali, Aaron Lauve, Tung Le, Stephen Lewis, Huilan Li, Kay Magaard, Eric Marberg, Jean-Christophe Novelli, Amy Pang, Franco Saliola, Lenny Tevlin, Jean-Yves Thibon, Nathaniel Thiem, Vidya Venkateswaran, C. Ryan Vinroot, Ning Yan, and Mike Zabrocki. "Supercharacters, symmetric functions in noncommuting variables, and related Hopf algebras". *Advances in Mathematics* 229, 2012, pp. 2310–2337. arXiv: 1009.4134 [math.CO].

[2] Marcelo Aguiar and Federico Ardila. "Hopf monoids and generalized permutahedra". *Memoirs of the American Mathematical Society* 142, 2023. arXiv: 1709.07504 [math.CO].

[3] Marcelo Aguiar, Nantel Bergeron, and Frank Sottile. "Combinatorial Hopf algebras and generalized Dehn–Sommerville relations". *Compositio Mathematica* 142, 2006, pp. 1–30. arXiv: math/0310016 [math.CO].

[4] Marcelo Aguiar and Swapneel Mahajan. *Monoidal Functors, Species and Hopf Algebras*. CRM Monograph Series 29. American Mathematical Society, 2010.

[5] Per Alexandersson and Joakim Uhlin. "Cyclic sieving, skew Macdonald polynomials and Schur positivity". *Algebraic Combinatorics* 3, 2020, pp. 913–939. URL: https://alco.centre-mersenne.org/item/ALCO_2020__3_4_913_0/.

[6] Roland Bacher. "Sur le groupe d'interpolation". Preprint. 2006. arXiv: math/0609736 [math.CO].

[7] Paul-Hermann Balduf. "Dyson–Schwinger equations in minimal subtraction". *Annales de l'Institut Henri Poincaré. Combinatorics, Physics and their Interactions*, 2023. Published online first. arXiv: 2109.13684 [hep-th].

[8] Paul-Hermann Balduf, Amelia Cantwell, Kurusch Ebrahimi-Fard, Lukas Nabergall, Nicholas Olson-Harris, and Karen Yeats. "Tubings, chord diagrams, and Dyson–Schwinger equations". Preprint. 2023. arXiv: 2302.02019 [math.CO].

[9] Christoph Bergbauer and Dirk Kreimer. "Hopf algebras in renormalization theory: locality and Dyson–Schwinger equations from Hochschild cohomology". In: *Physics and Number Theory*. Ed. by Louise Nyssen. IRMA Lectures in Mathematics and Theoretical Physics 10. EMS Press, 2006, pp. 133–164. arXiv: hep-th/0506190.

[10] George M. Bergman. "Everybody knows what a Hopf algebra is". In: *Group Actions on Rings*. Ed. by Susan Montgomery. Contemporary Mathematics 43. American Mathematical Society, 1985, pp. 25–48.

[11] Louis J. Billera, Hugh Thomas, and Stephanie van Willigenburg. "Decomposable compositions, symmetric quasisymmetric functions and equality of ribbon Schur functions". *Advances in Mathematics* 204, 2006, pp. 204–206. arXiv: math/0405434 [math.CO].

[12] Michael Borinsky. "Feynman graph generation and calculations in the Hopf algebra of Feynman graphs". *Computer Physics Communications* 185, 2014, pp. 3317–3330. arXiv: 1402.2613 [hep-th].

[13] Michael Carr and Satyan L. Devadoss. "Coxeter complexes and graph-associahedra". *Topology and its Applications* 153, 2006, pp. 2155–2168. arXiv: math/0407229 [math.QA].

[14] Pierre Cartier and Frédéric Patras. *Classical Hopf Algebras and Their Applications*. Algebra and Applications 29. Springer, 2021.

[15] Alain Connes and Dirk Kreimer. "Hopf algebras, renormalization and noncommutative geometry". *Communications in Mathematical Physics* 199, 1998, pp. 203–242. arXiv: hep-th/9808042.

[16] Yukio Doi. "Homological coalgebra". *Journal of the Mathematical Society of Japan* 33, 1981, pp. 31–50.

[17] Peter Doubilet. "A Hopf algebra arising from the lattice of partitions of a set". *Journal of Algebra* 28, 1974, pp. 127–132.

[18] William T. Dugan. "Sequences of trees and higher-order renormalization group equations". Master's thesis. University of Waterloo, 2019. URL: https://uwspace.uwaterloo.ca/handle/10012/14957.

[19] F. J. Dyson. "Divergence of perturbation theory in quantum electrodynamics". *Physical Review* 84, 1952, pp. 631–632.

[20] R. P. Feynman. "Space-time approach to non-relativistic quantum mechanics". *Review of Modern Physics* 20, 1948, pp. 367–387.

[21] R. P. Feynman. "Space-time approach to quantum electrodynamics". *Physical Review* 76, 1949, pp. 769–789.

[22] Loïc Foissy. "Faà di Bruno subalgebras of the Hopf algebra of planar trees from combinatorial Dyson–Schwinger equations". *Advances in Mathematics* 218, 2008, pp. 136–162. arXiv: 0707.1204 [math.RA].

[23] Loïc Foissy. "General Dyson–Schwinger equations and systems". *Communications in Mathematical Physics* 327, 2014, pp. 151–179. arXiv: 1112.2606 [math.RA].

[24] Loïc Foissy. "Lie algebra associated to systems of Dyson-Schwinger equations". *Advances in Mathematics* 226, 2011, pp. 4702–4730. arXiv: 0909.0358 [math.RA].

[25] Loïc Foissy. "Multigraded Dyson-Schwinger systems". *Journal of Mathematical Physics* 61, 2020, p. 051703. arXiv: 1511.06859 [math.RA].

[26] Pavel Galashin. "P-associahedra". *Selecta Mathematica*. New Series 30, 6, 2023. arXiv: 2110.07257 [math.CO].

[27] Ladnor Geissinger. "Hopf algebras of symmetric functions and class functions". In: *Combinatoire et Representation du Groupe Symetrique*. Ed. by Dominique Foata. Lecture Notes in Mathematics 579. Springer-Verlag, 1977, pp. 168–181.

[28] Darij Grinberg and Victor Reiner. *Hopf Algebras in Combinatorics*. URL: http://www.cip.ifi.lmu.de/~grinberg/algebra/HopfComb.pdf.

[29] Mark Haiman and William Schmitt. "Incidence algebra antipodes and Lagrange inversion in one and several variables". *Journal of Combinatorial Theory*. Series A 50, 1987, pp. 172–185.

[30] A. M. Hamel and I. P. Goulden. "Planar decompositions of tableaux and Schur function determinants". *European Journal of Combinatorics* 16, 1995, pp. 461–477.

[31] Markus Hihn and Karen Yeats. "Generalized chord diagram expansions of Dyson–Schwinger equations". *Annales de l'Institut Henri Poincaré. Combinatorics, Physics and their Interactions* 6, 2019, pp. 573–605. arXiv: 1602.02550 [math-ph].

[32] Heinz Hopf. "Uber die Topologie der Gruppen-Mannigfaltigkeiten und ihre Verallgemeinerungen". *Annals of Mathematics*. Second Series 42, 1941, pp. 22–52.

[33] Brandon Humpert and Jeremy L. Martin. "The incidence Hopf algebra of graphs". *SIAM Journal on Discrete Mathematics* 26, 2012, pp. 555–570. arXiv: 1012.4786 [math.CO].

[34] Emma Yu Jin and Shu Xiao Li. "Towards equivalent thickened ribbon Schur functions". Preprint. 2024. arXiv: 2403.01843 [math.CO].

[35] S. A. Joni and G.-C. Rota. "Coalgebras and bialgebras in combinatorics". *Studies in Applied Mathematics* 61, 1979, pp. 93–139.

[36] Donald Knuth. *The Art of Computer Programming*. Vol. 3: *Searching and Sorting*. 2nd ed. Addison-Wesley, 1998.

[37] Dirk Kreimer. "On overlapping divergences". *Communications in Mathematical Phyiscs* 204, 1999, pp. 669–689. arXiv: q-alg/9707029.

[38] Dirk Kreimer. "On the Hopf algebra structure of perturbative quantum field theories". *Advances in Theoretical and Mathematical Physics* 2, 1998, pp. 303–334. arXiv: q-alg/9707029.

[39] Jean-Louis Loday and María Ronco. "Combinatorial Hopf algebras". In: *Quanta of Maths*. Ed. by Etienne Blanchard, David Ellwood, Masoud Khalkhali, Matilde Marcolli, Henri Moscovici, and Sorin Popa. Clay Mathematics Proceedings 11. American Mathematical Society, 2007, pp. 347–383.

[40] Nicolas Marie. "On Laplace-Borel resummation of Dyson-Schwinger equations". PhD thesis. Simon Fraser University, 2014. URL: https://summit.sfu.ca/item/14938.

[41] Nicolas Marie and Karen Yeats. "A chord diagram expansion coming from some Dyson-Schwinger equations". *Communications in Number Theory and Physics* 7, 2014, pp. 251–291. arXiv: 1210.5457 [math.CO].

[42] Peter R. W. McNamara. "Comparing skew Schur functions: a quasisymmetric perspective". *Journal of Combinatorics* 5, 2014, pp. 51–58. arXiv: 1307.6233 [math.CO].

[43] Peter R. W. McNamara and Stephanie van Willigenburg. "Towards a combinatorial classification of skew Schur functions". *Transactions of the American Mathematical Society* 361, 2009, pp. 4437–4470. arXiv: math/0608446 [math.CO].

[44] Susan Montgomery. *Hopf Algebras and Their Actions on Rings*. CBMS Regional Conference Series in Mathematics 82. American Mathematical Society, 1993.

[45] Lukas Nabergall. "Enumerative perspectives on chord diagrams". PhD thesis. University of Waterloo, 2022. URL: https://uwspace.uwaterloo.ca/handle/10012/18873.

[46] Son Nguyen and Andrew Sack. "Poset associahedra and stack-sorting". Preprint. 2023. arXiv: 2310.02512 [math.CO].

[47] Erik Panzer. "Hopf-algebraic renormalization of Kreimer's toy model". Master's thesis. Humboldt University of Berlin, 2011. arXiv: 1202.3552 [math.QA].

[48] F. Patras and C. Reutenauer. "On descent algebras and twisted bialgebras". *Moscow Mathematical Journal* 4, 2004, pp. 199–216.

[49] David Prinz. "Gauge symmetries and renormalization". *Mathematical Physics, Analysis and Geometry* 25, 20, 2022. arXiv: 2001.00104 [hep-th].

[50] Hans Rademacher. *Topics in Analytic Number Theory*. Springer-Verlag, 1973.

[51] Victor Reiner, Kristin M. Shaw, and Stephanie van Willigenburg. "Coincidences among skew Schur functions". *Advances in Mathematics* 216, 2007, pp. 118–152. arXiv: math/0602634 [math.CO].

[52] William R. Schmitt. "Antipodes and incidence coalgebras". *Journal of Combinatorial Theory*. Series A 46, 1987, pp. 264–290.

[53] William R. Schmitt. "Hopf algebras of combinatorial structures". *Canadian Journal of Mathematics* 45, 1993, pp. 412–428.

[54] William R. Schmitt. "Incidence Hopf algebras". *Journal of Pure and Applied Algebra* 96, 1994, pp. 299–330.

[55] Louis W. Shapiro, Seyoum Getu, Wen-Jin Woan, and Leon C. Woodson. "The Riordan group". *Discrete Applied Mathematics* 34, 1991, pp. 229–239.

[56] Richard P. Stanley. *Enumerative Combinatorics*. Vol. 2. Cambridge Studies in Advanced Mathematics 62. Cambridge University Press, 1999.

[57] Richard P. Stanley. *Enumerative Combinatorics*. Vol. 1. 2nd ed. Cambridge Studies in Advanced Mathematics 49. Cambridge University Press, 2012.

[58] Richard P. Stanley. "A symmetric function generalization of the chromatic polynomial of a graph". *Advances in Mathematics* 111, 1995, pp. 166–194.

[59] Walter D. van Suijlekom. "Renormalization of gauge fields using Hopf algebras". In: *Quantum Field Theory. Competitive Models*. Ed. by Bertfried Fauser, Jürgen Tolksdorf, and Eberhard Zeidler. Springer, 2009, pp. 135–154. arXiv: 0801.3170 [math-ph].

[60] Benjamin C. Ward. "Massey products for graph homology". *International Mathematics Research Notices* 2022, 2021, pp. 8086–8161. URL: https://academic.oup.com/imrn/article/2022/11/8086/6071864.

[61] Stephanie van Willigenburg. "Equality of Schur and skew Schur functions". *Annals of Combinatorics* 9, 2005, pp. 355–562. arXiv: math/0410044 [math.CO].

[62] Karen Yeats. *A Combinatorial Perspective on Quantum Field Theory*. Springer Briefs in Mathematical Physics 15. Springer, 2017.

[63] Karen Yeats. "A Hopf algebraic approach to Schur function identities". *Electronic Journal of Combinatorics* 24, P3.10, 2017. URL: https://www.combinatorics.org/ojs/index.php/eljc/article/view/v24i3p10.

[64] Karen Yeats. "Growth estimates for Dyson-Schwinger equations". PhD thesis. Boston University, 2008.

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