Combinatorial Methods in Multilinear Algebra

by

Jöns Richard Gustaf Ehrenborg

B.Sc., University of Stockholm (1988)

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1993

Signature of Author ....
Department of Mathematics
5th May, 1993

Certified by ....
Gian-Carlo Rota
Professor of Mathematics
Thesis Supervisor

Accepted by ....
Sigurdur Helgason
Chairman, Departmental Graduate Committee
Department of Mathematics

Combinatorial Methods in Multilinear Algebra

by

Jöns Richard Gustaf Ehrenborg

Submitted to the Department of Mathematics
on 5th May, 1993, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This dissertation consists of two parts. The first chapter is a study canonical forms. The main theorem of canonical forms and apolarity reduces the question of whether a form in an algebra is canonical to a question about apolarity. We give applications of this theorem to tensors and skew-symmetric tensors. The chapter ends with a complete study of the invariants and covariants of two by two by two matrices, and how these covariants relate to known covariants of symmetric and skew-symmetric tensors.

In the second chapter we study inversion formulas for formal power series. We use the theory of colored species to prove the plethystic Lagrange inversion formula and the infinite variated Good's inversion formula. These inversion formulas are shown to be equivalent to transfer formulas in the infinite variated umbral calculus. Lastly, we give two enumerative proofs of the plethystic Lagrange inversion formula. Chapter 2 is joint work with Miguel Méndez.

Thesis Supervisor: Gian-Carlo Rota
Title: Professor of Mathematics

Dedicated to Laura and Ana Méndez.

Acknowledgements

First I would like to thank Anders Björner, who advised me to apply to MIT and thus started my venture in America.

Here at MIT I would like to thank Robert C. Pace and his numerous friends, the social club of room 2-342, and the SCC 24-Hour Coffeehouse for support and friendships through the years. Among all these people, I would like to mention by name those with whom I talked about mathematics. They are Sara Billey, Carlos Celentano, Wendy Chan, William Chen, Lenore Cowen, Frank D'Ippolito, Lars Ernström, Wayne Goddard, Mike Hawrylycz, Jennifer Huang, Dan Klain, Marcos Kiwi, Yuan Ma, Hsin-Fei Meng, Paul Milewski, Noël Nachtigal, Joe Oliveira, Satomi Okazaki, Dan Port, Satish Reddy, Neil Ross, Peter Schmid, Leonard Schulman, Ray Sidney, Dan Spielman, Esteban Tabak, Brian Taylor, Glenn Tesler, Dmitriy Vaynblat, Shinya Watanabe, Ethan Wolf, Rodney Worthing, Jianke Yang, Julia Yang, Yiqun Yin, Francis Zane, and Peter Zhuang. In our many conversations I have learned much about mathematics and related subjects. Especially close to my heart are Bonnie Berger, Chris Skinner, and Heather and Jihad Touma. Thanks also to Gábor Hetyei, who has been a great help to me in understanding mathematics and clearing my thoughts on many issues.

The administrative and technical assistance of Maureen Lynch and Robert Becker has been invaluable.

It was a pleasure to collaborate with Miguel Méndez. We have spent many hours, many days discussing ideas, and I am indebted to him. I am also grateful to Dan Kleitman and Richard Stanley for sharing their knowledge of combinatorics with me, both through classes, discussions, and also by letting me be their teaching assistant.

Most of all, I would like express my deepest gratitude to my advisor Gian-Carlo Rota, who has shown me beautiful mathematics and has taught me the *attitudes* of combinatorics and of life.

At last but not at least I would like to thank my parents, my family and my ancestors, without whose help and support this adventure wouldn't have been possible.

Contents

Introduction

Among all branches of mathematics that have developed in modern times, multilinear algebra is the one that has had the most halting and irregular development. Already in the nineteenth century, the total neglect of the work of Grassmann, following in the wake of an almost equally shameful oblivion of Möbius's barycentric calculus and even of von Staudt's algebra of throws (which was to be embarrassingly rediscovered as late as the 1930's by von Neumann), was coupled with an equally inexplicable duplication of work under disparate headings. Thus, towards the end of the past century and in the beginning of the present, the newly invented tensors, triumphantly unveiled by Ricci and Levi-Civita as the saving concept of differential geometry, were inspiringly worked upon by elite cadres of vanguardists, blissfully ignorant of the already largely developed theory of invariants, where they would have found their tensors ready and waiting with powerfully developed techniques, though in an altogether different language. At a more shocking level, the first two installments of the Rev. Alfred Young's "Quantitative Substitutional Analysis" and Issai Schur's celebrated thesis (written under Frobenius's guidance), published exactly the same year, 1900, were not only not seen as related to one another for some twenty years hence, but were not even noticed by differential geometers for at least another 50 years, after vector bundles and global techniques had made representation theory a sine qua non. Perhaps the most ironic episode betokening this mutual ignorance of schools is a comparison between Schouten's ponderous treatise "Riccikalkul" and Weitzenbock's equally ponderous "Invarianthentheorie", both books published in roughly the same period. True, the first author being a Dutch electrical engineer by upbringing and the second an Austrian Nazi general might have prevented their developing a convivial friendship. Yet, they both taught in the Netherlands, and the subjects of their books can in retrospect be viewed as the same; nevertheless, little if any mention is made of either in the other's writings. The principal victim of this state of affairs was classical invariant theory. We shall leave to another occasion the precise tracing of the historical events that led to the temporary eclipse of this field, which after the combinatorial pyrotechnics of Major McMahon, after the clamorous failure of Emmy Noether in her thesis (written under the guidance of Paul Gordan) to determine a set of generators for the ring of invariants of the ternary quartic (which led to a prohibition of even uttering the name "invariant" among some of Emmy Noether's disciples), and after the profound but slightly epigonic outpour of the Scottish school, led both continental and American mathematicians to the mistaken impression that the field of invariant theory was an amusing hobby to be left in the care of British gentlemen and reverends. The recent rebirth of the field, heralded by the finer points of representation theory as well as by the necessities of physics (without which Young's name might well still be a dead letter), is presently risking yet another embarrassing duplication. The requirements of the fine-tuned theory of irreducible representations, while acknowledging and amply making use of the newly developed powerful techniques of combinatorics, are on the one hand skirting with a gingerly ignorance of the classical heritage, on the other, they are reluctant to adopt the recent sweeping notational reforms of bijective combinatorics. The present thesis situates itself at the intersection of these two relatively new trends. While on the one hand my background is decidedly invariant- theoretic rather than representation-theoretic, and while my language is uncompromisingly bijective, I should like to stress the underlying unity of style as well as motivation in the differently named sections that follow. The first section decidedly harks back to what is perhaps the most beautiful and the least known idea of classical invariant theory, namely, apolarity. In another work **[E-R1]** (to be published separately and not included in the present thesis) I have given an up-to-date exposition of this concept, together with varied applications to canonical forms for ordinary polynomials in several variables. In the course of rethinking the classical theory along contemporary lines, it occurred to me that the concept of apolarity has a much wider scope than the one the classical theory limits it to. In fact, I was lucky enough to be able to develop the theory of apolarity in the context of associative but non-commutative algebras, more specifically, in the free ring. I was thereby led to a very general result on canonical forms of polynomials in non- commuting variables (Theorem 1 below). Luckily, the techniques I have used in the proof of this theorem ultimately rely on calculations with Jacobians, as in the classical theory. There are several applications of this main theorem, which I have left out of the present thesis for reasons of time, but which I intend to include in a later published version. I only briefly hint at some applications to commutative algebras, which go beyond the range of the classical theory (Section 1.5). I should like to stress that the algebra of polarizations has been extended in the present thesis in a direction which is different from the ones previously used even with non-commutative algebras, for example, the one used by P.M. Cohn's thesis **[Cohn]** (written under the guidance of Philip Hall), and I surmise that, much as the algebra of polarization has been effectively used in recent work in invariant theory and Hopf algebras, the algebra of polarizations presented in the present thesis may lend itself to further application beyond the confines of the theory of apolarity. The second and third sections have taken up much of my efforts for a long time, and I would like to believe that they contain permanent results. Briefly, the problem that is tackled is the representation of tensors as sums of minimal numbers of decomposable tensors (sometimes called the rank of a tensor, especially for skew-symmetric tensors). My main contribution consists in relating the problem to various considerations of the theory of block designs (or the simpler rook coverings defined in Section 1.3). Of several results obtained, I should like to call attention to Proposition 1.2.7, perhaps the least trivial of the lot. I should like to add that, unlike other problems in the invariant theory of arbitrary tensors, the problems treated here do not seem to be amenable to reduction to the commutative case, as happens for other invariant-theoretic problems relative to general tensors. In Section 1.5 I develop the invariant theory of binary tensors of degree three. The results I obtain could be compared with those obtained in Clebsch and Gordan's well-known computation of a generating set of concomitants for a ternary cubic, although their results do not imply mine, nor (probably) mine theirs. Little has been written on this subject, and even less on its relation to classical invariant theory of symmetric tensors. In fact, we find that our determination of the invariants of our tensors, as well as their syzygies, implies some classical results on invariants not only of commutative (for the binary cubic, for instance) but also of skew-symmetric tensors of step three as well. The remaining portion of the thesis is concerned with an altogether different problem, which, to be sure, also arose in invariant theory. It is the problem of finding and explicitly computing an analog of the Lagrange inversion formula for the plethystic composition of formal power series in infinitely many variables. To this end, we have extended and adapted the language of Joyal's theory of species. Our starting point is the notion of a c-monoid, introduced by Mendez and Yang, together with some notions that originated in the theory of Witt vectors, namely, Verschiebung and Frobenius operators. We also use W.Y.C. Chen's theory of plethystic trees. We give two versions of our "Lagrange inversion formula" for plethystic composition: the first is decidedly bijective, and couched in the language of plethystic species already used by the above-mentioned authors (Theorem 3); the second (Theorem 4) gives the formula in the ordinary language of formal power series. In the following section, we provide a bijective interpretation to an inversion formula due to I.J. Good, which may be viewed as yet another generalization of the Lagrange inversion formula. Our proof is entirely bijective. This is followed with a development of an umbral calculus for plethystic composition, culminating in a plethystic analog of the so-called "transfer formula" of Roman and Rota, along lines initiated by W.Y.C. Chen; in the present treatment, the proof of the transfer formula is made to depend on our previous work on plethystic inversion. Finally, in the last section, we obtain several results, some new and some overlapping with the work of Chen, on plethystic trees. In the last sections, we have provided two enumerative proofs of our plethystic Lagrange inversion formula, for mathematicians who prefer the classical language of formal power series to the language of species. I very much hope that the present thesis will contribute to the symbiosis of invariant theory and combinatorics, as envisaged by the early workers, such as Sylvester and McMahon.

Cambridge, May 5, 1993.

Chapter 1

Apolarity and Canonical Forms

1.1 Main theorem on apolarity and canonical forms

1.1.1 Polynomials

Define $\mathbf{C}\langle x_1, \ldots, x_n\rangle$ be the algebra of polynomials over $\mathbf{C}$ with non-commuting variables $x_1, \ldots, x_n$.

Definition 1.1.1 *Let A be an algebra over a the field of complex numbers $\mathbf{C}$. Define*

$$A\{x_1, \ldots, x_n\} = A * \mathbf{C}\langle x_1, \ldots, x_n\rangle,$$

where the product $*$ is the free product of algebras. The elements of $A\{x_1, \ldots, x_n\}$ are called polynomials *in n variables.*

We can also view $A\{x_1, \ldots, x_n\}$ as the smallest algebra containing the algebra A and the variables $x_1, \ldots, x_n$. In fact an element of $A\{x_1, \ldots, x_n\}$ can be constructed in a finite number of steps by the following rules.

- $x_i \in A\{x_1, \ldots, x_n\}$ for $1 \leq i \leq n$,
- $a \in A\{x_1, \ldots, x_n\}$ for $a \in A$,

- $p + q \in A\{x_1, \ldots, x_n\}$ for $p, q \in A\{x_1, \ldots, x_n\}$,
- $p \cdot q \in A\{x_1, \ldots, x_n\}$ for $p, q \in A\{x_1, \ldots, x_n\}$.

Hence we can prove statements about elements in $A\{x_1, \ldots, x_n\}$ by induction.

We will view the elements of $A\{x_1, \ldots, x_n\}$ as polynomials over the algebra A. Observe that a general monomial are in the form $a_1 x_{i_1} a_2 x_{i_2} a_3 \cdots a_m x_{i_m} a_{m+1}$. In fact we can evaluate the values of such a polynomial.

Lemma 1.1.1 *Let B an an algebra that contains the algebra A. Then there is an unique evaluation map, denoted by* eval, *from $A\{x_1, \ldots, x_n\} \times B^n$ to B such that for $(b_1, \ldots, b_n) \in B^n$*

- $\mathrm{eval}\,(x_i\,;\ b_1, \ldots, b_n) = b_i$,
- $\mathrm{eval}\,(a\,;\ b_1, \ldots, b_n) = a$,
- $\mathrm{eval}\,(\cdot\,;\ b_1, \ldots, b_n)$ *is an algebra homomorphism from $A\{x_1, \ldots, x_n\}$ to B.*

This lemma follows from the definition of the free product. To avoid confusion we will sometimes write $\mathrm{eval}\,(p;\ x_1 \leftarrow b_1, \ldots, x_n \leftarrow b_n)$ instead of $\mathrm{eval}\,(p;\ b_1, \ldots, b_n)$.

1.1.2 Polarizations and Apolarity

Definition 1.1.2 *A polarization D_{t,x_i} is a linear map from the algebra $A\{x_1, \ldots, x_n\}$ to the algebra $A\{t, x_1, \ldots, x_n\}$, that satisfies*

- $D_{t,x_i}(a) = 0$ *for* $a \in A$,
- $D_{t,x_i}(x_i) = t$,
- $D_{t,x_i}(x_j) = 0$ *for* $j \neq i$,
- $D_{t,x_i}(p \cdot q) = D_{t,x_i}(p) \cdot q + p \cdot D_{t,x_i}(q)$.

Since the algebra $A\{y\}$ has the algebra A as subalgebra, an expression in the form eval $(p;\ q)$ makes sense and is in $A\{y\}$, where $p \in A\{x\}$. and $q \in A\{y\}$. Thus we have the following chain rule.

Proposition 1.1.2 *Let $p \in A\{x\}$ and $q \in A\{y\}$*

$$D_{u,y}\text{eval}\,(p;\ x \leftarrow q) = \text{eval}\,(D_{t,x}p;\ t \leftarrow D_{u,y}q, x \leftarrow q)\,.$$

Proof: The proof is by induction on p.

- $p = a \in A$. Both sides vanish

$$D_{u,y}\text{eval}\,(a\,;\ x \leftarrow q) = D_{u,y}a = 0,$$

$$\text{eval}\,(D_{t,x}a\,;\ t \leftarrow D_{u,y}q, x \leftarrow q) = \text{eval}\,(0\,;\ x \leftarrow D_{u,y}q, t \leftarrow q) = 0.$$

- $p = x$.

$$\begin{aligned} D_{u,y}\text{eval}\,(x\,;\ x \leftarrow q) &= D_{u,y}q \\ &= \text{eval}\,(t\,;\ t \leftarrow D_{u,y}q, x \leftarrow q) \\ &= \text{eval}\,(D_{t,x}x\,;\ t \leftarrow D_{u,y}q, x \leftarrow q)\,, \end{aligned}$$

  and thus they are equal.

- $p = r + s$, where $r, s \in A\{x\}$.

$$\begin{aligned} D_{u,y}\text{eval}\,(r+s\,;\ x \leftarrow q) &= D_{u,y}\text{eval}\,(r\,;\ x \leftarrow q) + D_{u,y}\text{eval}\,(s\,;\ x \leftarrow q) \\ &= \text{eval}\,(D_{t,x}r\,;\ t \leftarrow D_{u,y}q, x \leftarrow q) + \\ &\quad +\text{eval}\,(D_{t,x}s\,;\ t \leftarrow D_{u,y}q, x \leftarrow q) \\ &= \text{eval}\,(D_{t,x}(r+s)\,;\ t \leftarrow D_{u,y}q, x \leftarrow q)\,. \end{aligned}$$

- $p = r \cdot s$, where $r, s \in A\{x\}$.

$$\begin{aligned}
D_{u,y}\text{eval}\,(r \cdot s\,;\, x \leftarrow q) &= D_{u,y}\,(\text{eval}\,(r\,;\, x \leftarrow q) \cdot \text{eval}\,(s\,;\, x \leftarrow q)) \\
&= D_{u,y}\,(\text{eval}\,(r\,;\, x \leftarrow q)) \cdot \text{eval}\,(s\,;\, x \leftarrow q) + \\
&\quad + \text{eval}\,(r\,;\, x \leftarrow q) \cdot D_{u,y}\,(\text{eval}\,(s\,;\, x \leftarrow q)) \\
&= \text{eval}\,(D_{t,x}r\,;\, t \leftarrow D_{u,y}q, x \leftarrow q) \cdot \text{eval}\,(s\,;\, x \leftarrow q) + \\
&\quad + \text{eval}\,(r\,;\, x \leftarrow q) \cdot \text{eval}\,(D_{t,x}s\,;\, t \leftarrow D_{u,y}q, x \leftarrow q) \\
&= \text{eval}\,(D_{t,x}(r) \cdot s\,;\, t \leftarrow D_{u,y}q, x \leftarrow q) + \\
&\quad + \text{eval}\,(r \cdot D_{t,x}(s)\,;\, t \leftarrow D_{u,y}q, x \leftarrow q) \\
&= \text{eval}\,(D_{t,x}(r) \cdot s + r \cdot D_{t,x}(s)\,;\, t \leftarrow D_{u,y}q, x \leftarrow q) \\
&= \text{eval}\,(D_{t,x}(r \cdot s)\,;\, t \leftarrow D_{u,y}q, x \leftarrow q)\,.
\end{aligned}$$

□

We have also have the following chain rule for polarizations.

Proposition 1.1.3 *Let $p \in A\{x\}$ and let ϕ be a differentiable function from $\mathbf{C}$ to B, where B is an algebra that contains A as a subalgebra.*

$$\frac{\partial}{\partial\alpha}\text{eval}\,(p;\, \phi(\alpha)) = \text{eval}\,(D_{t,x}p;\, \phi'(\alpha), \phi(\alpha))\,.$$

Proof: The proof is by induction on p.

- $p = a \in A$. Both sides vanish

$$\frac{\partial}{\partial\alpha}\text{eval}\,(a\,;\, \phi(\alpha)) = \frac{\partial}{\partial\alpha}a = 0,$$

$$\text{eval}\,(D_{t,x}a\,;\, \phi'(\alpha), \phi(\alpha)) = \text{eval}\,(0\,;\, \phi'(\alpha), \phi(\alpha)) = 0.$$

- $p = x$.

$$\frac{\partial}{\partial\alpha}\text{eval}\,(x\,;\ \phi(\alpha)) = \frac{\partial}{\partial\alpha}\phi(\alpha) = \phi'(\alpha),$$

$$\text{eval}\,(D_{t,x}x\,;\ \phi'(\alpha),\phi(\alpha)) = \text{eval}\,(t\,;\ \phi'(\alpha),\phi(\alpha)) = \phi'(\alpha),$$

and thus they are equal.

- $p = q + r$, where $q, r \in A\{x\}$.

$$\begin{aligned}
\frac{\partial}{\partial\alpha}\text{eval}\,(p+q\,;\ \phi(\alpha)) &= \frac{\partial}{\partial\alpha}\left(\text{eval}\,(p\,;\ \phi(\alpha)) + \text{eval}\,(q\,;\ \phi(\alpha))\right)\\
&= \frac{\partial}{\partial\alpha}\text{eval}\,(p\,;\ \phi(\alpha)) + \frac{\partial}{\partial\alpha}\text{eval}\,(q\,;\ \phi(\alpha))\\
&= \text{eval}\,(D_{t,x}p\,;\ \phi'(\alpha),\phi(\alpha)) + \text{eval}\,(D_{t,x}q\,;\ \phi'(\alpha),\phi(\alpha))\\
&= \text{eval}\,(D_{t,x}p + D_{t,x}q\,;\ \phi'(\alpha),\phi(\alpha))\\
&= \text{eval}\,(D_{t,x}(p+q)\,;\ \phi'(\alpha),\phi(\alpha))\,.
\end{aligned}$$

- $p = q \cdot r$, where $q, r \in A\{x\}$.

$$\begin{aligned}
\frac{\partial}{\partial\alpha}\text{eval}\,(p\cdot q\,;\ \phi(\alpha)) &= \frac{\partial}{\partial\alpha}\left(\text{eval}\,(p\,;\ \phi(\alpha))\cdot\text{eval}\,(q\,;\ \phi(\alpha))\right)\\
&= \frac{\partial}{\partial\alpha}\left(\text{eval}\,(p\,;\ \phi(\alpha))\right)\cdot\text{eval}\,(q\,;\ \phi(\alpha))\\
&\quad+\text{eval}\,(p\,;\ \phi(\alpha))\cdot\frac{\partial}{\partial\alpha}\left(\text{eval}\,(q\,;\ \phi(\alpha))\right)\\
&= \text{eval}\,(D_{t,x}(p)\,;\ \phi'(\alpha),\phi(\alpha))\cdot\text{eval}\,(q\,;\ \phi(\alpha))\\
&\quad+\text{eval}\,(p\,;\ \phi(\alpha))\cdot\text{eval}\,(D_{t,x}(q)\,;\ \phi'(\alpha),\phi(\alpha))\\
&= \text{eval}\,(D_{t,x}(p)\cdot q\,;\ \phi'(\alpha),\phi(\alpha)) +\\
&\quad+\text{eval}\,(p\cdot D_{t,x}(q)\,;\ \phi'(\alpha),\phi(\alpha))\\
&= \text{eval}\,(D_{t,x}(p)\cdot q + p\cdot D_{t,x}(q)\,;\ \phi'(\alpha),\phi(\alpha))\\
&= \text{eval}\,(D_{t,x}(p\cdot q)\,;\ \phi'(\alpha),\phi(\alpha))\,.
\end{aligned}$$

□

Similarly we can prove the more general chain rule.

Proposition 1.1.4 *Let $p \in A\{x_1, \ldots, x_n\}$ and let ϕ_i be a differentiable function from* C *to B for $i = 1, \ldots, n$, where B is an algebra that has A as a subalgebra.*

$$\frac{\partial}{\partial \alpha}\text{eval}\,(p;\ \phi_1(\alpha), \ldots, \phi_n(\alpha)) = \sum_{i=1}^{n} \text{eval}\,(D_{t,x_i}p;\ \phi_i'(\alpha), \phi_1(\alpha), \ldots, \phi_n(\alpha)).$$

Lemma 1.1.5 *Let B be an algebra that has A as a subalgebra. Let $p \in A\{x_1, \ldots, x_n\}$ and let $b_1, \ldots, b_n \in B$. Then the following map from B to B is linear*

$$y \longmapsto \text{eval}\,(D_{t,x_i}p;\ y, b_1, \ldots, b_n).$$

Proof: The proof is by induction on p.

- $p = a \in A$. Then $D_{t,x_i}p = 0$, so the map is the zero map, which is linear.
- $p = x_j$. Then $D_{t,x_i}p = \delta_{i,j}t$. So $y \longrightarrow \delta_{i,j}y$, which is linear.
- $p = q + r$, where $q, r \in A\{x_1, \ldots, x_n\}$.

$$\begin{aligned} \text{eval}\,(D_{t,x_i}(q+r);\ y, b_1, \ldots, b_n) &= \text{eval}\,(D_{t,x_i}(q) + D_{t,x_i}(r);\ y, b_1, \ldots, b_n) \\ &= \text{eval}\,(D_{t,x_i}(q);\ y, b_1, \ldots, b_n) + \\ &\quad +\text{eval}\,(D_{t,x_i}(r);\ y, b_1, \ldots, b_n), \end{aligned}$$

  and the sum of to linear maps is linear.

- $p = q \cdot r$, where $q, r \in A\{x_1, \ldots, x_n\}$.

$$\text{eval}\,(D_{t,x_i}(q \cdot r);\ y, b_1, \ldots, b_n)$$

$$
\begin{aligned}
&= \text{eval}\,(D_{t,x_i}(q)\cdot r + q\cdot D_{t,x_i}(r)\,;\ y, b_1,\ldots,b_n)\\
&= \text{eval}\,(D_{t,x_i}(q)\cdot r\,;\ y, b_1,\ldots,b_n) +\\
&\quad + \text{eval}\,(q\cdot D_{t,x_i}(r)\,;\ y, b_1,\ldots,b_n)\\
&= \text{eval}\,(D_{t,x_i}(q)\,;\ y, b_1,\ldots,b_n)\cdot \text{eval}\,(r\,;\ b_1,\ldots,b_n) +\\
&\quad + \text{eval}\,(q\,;\ b_1,\ldots,b_n)\cdot \text{eval}\,(D_{t,x_i}(r)\,;\ y, b_1,\ldots,b_n)\,.
\end{aligned}
$$

The above expression is a linear combination of two linear maps, thus it is a linear map.

□

Definition 1.1.3 *Let A and B be algebras over the field* $\mathbf{C}$ *such that A is a subalgebra of B. An element $p \in A\{x_1,\ldots,x_s\}$ is* homogeneous *with respect to the linear spaces $V, W_1,\ldots,W_s$, where these linear spaces are finite dimensional subspaces of B, if for all $w_1 \in W_1,\ldots,w_s \in W_s$ we have that* $\text{eval}\,(p;\ w_1,\ldots,w_s) \in V$.

Lemma 1.1.6 *Assume that the polynomial $p \in A\{x_1,\ldots,x_s\}$ is homogeneous with respect to the linear spaces $V, W_1,\ldots,W_s$, then the polynomial $D_{t,x_i}(p)$ is homogeneous with respect to the linear spaces $V, W_i, W_1,\ldots,W_s$.*

Proof: Since p is homogeneous with respect to the linear spaces $V, W_1,\ldots,W_s$, we know that the following element lies in V:

$$\text{eval}\,(p;\ w_1,\ldots,w_i+\alpha\cdot w,\ldots,w_s)\,,$$

where $\alpha \in \mathbf{C}$, $w \in W_i$, and $w_j \in W_j$ for $j = 1,\ldots,s$. Consider the above expression as a function of α. The derivative of this function in variable α will also take values in V. By Proposition 1.1.3 the derivative is equal to

$$\frac{\partial}{\partial\alpha}\text{eval}\,(p;\ w_1,\ldots,w_i+\alpha\cdot w,\ldots,w_s)$$

$$= \operatorname{eval}\left(D_{t,x_i}p;\ \frac{\partial}{\partial\alpha}(w_i + \alpha\cdot w), w_1, \ldots, w_i + \alpha\cdot w, \ldots, w_s\right)$$
$$= \operatorname{eval}\left(D_{t,x_i}p;\ w, w_1, \ldots, w_i + \alpha\cdot w, \ldots, w_s\right).$$

Now by letting $\alpha = 0$ the result will follow. □

Definition 1.1.4 *Let V and W be finite dimensional linear spaces. Let $f : W \longrightarrow V$, and let $L \in V^*$. We say that f is apolar to L relative to W if for all $w \in W$*

$$\langle L \mid f(w)\rangle = 0.$$

1.1.3 Main Theorem

Definition 1.1.5 *Let V be a finite dimensional linear space. We say that a generic element $v \in V$ has a property P, if the set of all elements in V that has this property form a dense set in V, where V has the Euclidean topology.*

Theorem 1 *Let $V, W_1, \ldots, W_s$ be finite dimensional linear subspaces of the algebra A. Let p in $A\{x_1, \ldots, x_s\}$ be homogeneous with respect to the linear spaces $V, W_1, \ldots, W_s$. A generic element $v \in V$ can be written in the form*

$$v = \operatorname{eval}(p;\ w_1, \ldots, w_s)$$

for some $w_1, \ldots, w_s$ if and only if there exist $w'_1, \ldots, w'_s$ so that there is no nonzero dual element in V^* which is apolar to the linear map

$$y_j \longmapsto \operatorname{eval}\left(D_{t,x_j}p;\ y_j, w'_1, \ldots, w'_s\right)$$

relative to W_j, for all $1 \le j \le s$.

Observe that Lemma 1.1.5 shows that the map $y_j \mapsto \operatorname{eval}\left(D_{t,x_j}p;\ y_j, w'_1, \ldots, w'_s\right)$ is linear, and that Lemma 1.1.6 guarantees that it maps the linear space W_j into V.

To be able to prove the theorem, we need the two following propositions, which we state without proof. Let $\overline{\mathbf{C}(x_1,\ldots,x_q)}$ be the field of all algebraically functions in the variables $x_1,\ldots,x_q$.

Proposition 1.1.7 *Let $p_1(x_1,\ldots,x_q),\ldots,p_r(x_1,\ldots,x_q) \in \overline{\mathbf{C}(x_1,\ldots,x_q)}$, where $r \leq q$. Then the algebraic functions $p_1,\ldots,p_r$ are algebraically independent if and only if the matrix*

$$\left(\frac{\partial p_i}{\partial x_j}\right)_{1\leq i\leq r, 1\leq j\leq q}$$

has full rank.

Proposition 1.1.8 *Let $p_1(x_1,\ldots,x_q),\ldots,p_r(x_1,\ldots,x_q) \in \overline{\mathbf{C}(x_1,\ldots,x_q)}$, where $r \leq q$. Let $P : \mathbf{C}^q \to \mathbf{C}^r$ be defined by*

$$P(x_1,\ldots,x_q) = (p_1(x_1,\ldots,x_q),\ldots,p_r(x_1,\ldots,x_q)).$$

Then the algebraic functions $p_1,\ldots,p_r$ are algebraically independent if and only if the range of the map P is dense in $\mathbf{C}^r$.

Proof of Theorem 1: Let $d = \dim(V)$ and $d_j = \dim(W_j)$ for $j = 1,\ldots,s$. Choose a bases for V, say $u_1,\ldots,u_d$ and a bases for W_j, say $z_{j,1},\ldots,z_{j,d_j}$. Thus an element $w_j \in W_j$ can be written in the form

$$w_j = \sum_{i=1}^{d_j} \alpha_{j,i} z_{j,i},$$

where $\alpha_{j,i} \in \mathbf{C}$. We will call the coefficients $\alpha_{j,i}$ *parameters*. Let

$$\mathrm{Par} = \{(j,i) \ : \ 1 \leq j \leq s,\ 1 \leq i \leq d_j\}.$$

Thus a parameter is on the form α_k where $k \in \mathrm{Par}$. Observe that the number of parameters is $|\mathrm{Par}| = d_1 + \cdots + d_s$.

Assume that a generic element $v = \sum_{i=1}^{d} \beta_i u_i$ of V can be written in the form $p(w_1, \ldots, w_s)$. By counting coefficients on the left hand side and parameters on the right hand side, we obtain the inequality $d \leq |\mathrm{Par}| = d_1 + \cdots + d_s$.

Expand

$$\mathrm{eval}\,(p;\ w_1, \ldots, w_s) = \mathrm{eval}\left(p;\ \sum_{i=1}^{d_1} \alpha_{1,i} z_{1,i}, \ldots, \sum_{i=1}^{d_s} \alpha_{s,i} z_{s,i}\right)$$

into the bases $u_1, \ldots, u_d$. That is, we write

$$v = \sum_{i=1}^{s} \beta_i u_i = \mathrm{eval}\left(p;\ \sum_{i=1}^{d_1} \alpha_{1,i} z_{1,i}, \ldots, \sum_{i=1}^{d_s} \alpha_{s,i} z_{s,i}\right) = \sum_{i=1}^{d} \phi_i \left(\alpha_k\right)_{k \in \mathrm{Par}} u_i.$$

We obtain d identities

$$\beta_i = \phi_i \left(\alpha_k\right)_{k \in \mathrm{Par}}.$$

We view the coefficients β_i as polynomials in parameters.

Consider the map $\Phi : \mathbf{C}^{\mathrm{Par}} \longrightarrow \mathbf{C}^d$ defined by

$$\Phi\left(\left(\alpha_k\right)_{k \in \mathrm{Par}}\right) = \left(\phi_i \left(\alpha_k\right)_{k \in \mathrm{Par}}\right)_{1 \leq i \leq d},$$

where the coordinates of $\mathbf{C}^{\mathrm{Par}}$ are indexed by the set Par.

The assumption is that the range of the map Φ is dense in $\mathbf{C}^d$. By Proposition 1.1.8 we infer that the d polynomials ϕ_i are algebraically independent. Hence, by Proposition 1.1.7, the matrix

$$\left(\frac{\partial \phi_i}{\partial \alpha_k}\right)_{1 \leq i \leq d, k \in \mathrm{Par}} \tag{1.1}$$

has full rank, where the rows are indexed by i and the columns by the set Par.

Since the matrix (1.1) has full rank, we can choose values for the parameters such that the matrix (1.1) still has full rank. Denote these values we choose for the parameters by γ_k for $k \in \mathrm{Par}$. Let

$$w'_j = \sum_{i=1}^{d_j} \gamma_{j,i} z_{j,i}.$$

Thus $w'_j \in W_j$. Thus the matrix

$$\left(\left[\frac{\partial \phi_i}{\partial \alpha_k} \right]_{\alpha_m = \gamma_m} \right)_{1 \leq i \leq d, k \in \mathrm{Par}}$$

has full rank.

Because the matrix (1.1) has full rank, the columns of the matrix span the linear space $\mathbf{C}^d$. But $\mathbf{C}^d$ is canonically isomorphic to V. In particular, via this isomorphism we get

$$\left(\frac{\partial \phi_i}{\partial \alpha_k} \right)_{1 \leq i \leq d} \longmapsto \sum_{i=1}^{d} \frac{\partial \phi_i}{\partial \alpha_k} u_i = \frac{\partial v}{\partial \alpha_k}.$$

Thus, the elements

$$\left[\frac{\partial v}{\partial \alpha_k} \right]_{\alpha_m = \gamma_m}$$

span the linear space V.

Hence there is no nonzero functional $L \in V^*$ such that

$$\left\langle L \,\middle|\, \left[\frac{\partial v}{\partial \alpha_k} \right]_{\alpha_m = \gamma_m} \right\rangle = 0$$

for all $k \in \mathrm{Par}$.

Each of the parameters will only occur in one of the vectors $w_1, \ldots, w_s$. The parameter $\alpha_{j,i}$ occurs only in w_j. In particular, we have

$$\frac{\partial w_j}{\partial \alpha_{j,i}} = z_{j,i}.$$

Hence by the chain rule, Proposition 1.1.4, we conclude that

$$\frac{\partial v}{\partial \alpha_{j,i}} = \frac{\partial}{\partial \alpha_{j,i}} \mathrm{eval}\,(p;\ w_1, \ldots, w_s)$$

$$= \text{eval}\left(D_{t,x_j}p;\ \frac{\partial w_j}{\partial \alpha_{j,i}}, w_1, \ldots, w_s\right)$$
$$= \text{eval}\left(D_{t,x_j}p;\ z_{j,i}, w_1, \ldots, w_s\right).$$

Observe that

$$\left[\text{eval}\left(D_{t,x_j}p;\ z_{j,i}, w_1, \ldots, w_s\right)\right]_{\alpha_m=\gamma_m} = \text{eval}\left(D_{t,x_j}p;\ z_{j,i}, w'_1, \ldots, w'_s\right).$$

Thus we can write our condition as follows: there is no nonzero functional $L \in V^*$ such that

$$\left\langle L \,\middle|\, \text{eval}\left(D_{t,x_j}p;\ z_{j,i}, w'_1, \ldots, w'_s\right)\right\rangle = 0$$

for all $1 \leq j \leq s$ and $1 \leq i \leq d_j$. Observe that the above expression is linear in $z_{j,i}$ and recall that the elements $z_{j,1}, \ldots, z_{j,d_j}$ form a bases for W_j. Hence the statement above is equivalent to that there is no nonzero functional $L \in V^*$ such that

$$\left\langle L \,\middle|\, \text{eval}\left(D_{t,x_j}p;\ y_j, w'_1, \ldots, w'_s\right)\right\rangle = 0$$

for all $1 \leq j \leq s$ and for all $y_j \in W_j$.

Thus we have proven that there is no nonzero element in V^* apolar to all the maps

$$y_j \longmapsto \text{eval}\left(D_{t,x_j}p;\ y_j, w'_1, \ldots, w'_s\right)$$

relative to W_j for $j = 1, \ldots, s$. This provides us with half the proof.

To prove the second part, all we need to do is trace the equivalences above in opposite direction. Hence assume that there is no nonzero element in V^* apolar to all the maps

$$y_j \longmapsto \text{eval}\left(D_{t,x_j}p;\ y_j, w'_1, \ldots, w'_s\right)$$

relative to W_j for $j = 1, \ldots, s$. This implies that the elements

$$\text{eval}\left(D_{t,x_j}p;\ z_{j,i}, w_1', \ldots, w_s'\right),$$

where $1 \leq j \leq s$ and $1 \leq i \leq d_j$, span the linear space V. Hence $d_1 + \cdots + d_s$ elements span a linear space of dimension d. Thus

$$d \leq d_1 + \cdots + d_s. \tag{1.2}$$

By identities above, we can rewrite the above elements, and by using the canonical isomorphism between V and $\mathbf{C}^d$ we get that the matrix

$$\left(\left[\frac{\partial \phi_i}{\partial \alpha_k}\right]_{\alpha_m = \gamma_m}\right)_{1 \leq i \leq d, k \in \text{Par}}$$

has rank d. The inequality (1.2) implies that the above matrix has full rank. Thus the matrix

$$\left(\frac{\partial \phi_i}{\partial \alpha_k}\right)_{1 \leq i \leq d, k \in \text{Par}},$$

where we remove the values of the β's cannot have lower rank. But the rank can not increase so the last matrix has also full rank.

By Proposition 1.1.7 we know that the polynomials $\phi_1, \ldots, \phi_d$ are algebraically independent. Proposition 1.1.8 implies that the range of the map $\Phi : \mathbf{C}^{\text{Par}} \longrightarrow \mathbf{C}^d$ is dense. But this is equivalent to the second implication of the theorem. □

1.2 Multi dimensional matrices

1.2.1 Rook coverings

Let $A_1, \ldots, A_n$ be a finite sets. Let $A_1 \times \cdots \times A_n$ be the Cartesian product of these sets. That is,

$$A_1 \times \cdots \times A_n = \{(a_1, \ldots, a_n) \;:\; a_1 \in A_1, \ldots, a_n \in A_n\}.$$

Definition 1.2.1 *A* rook covering *of $A_1 \times \cdots \times A_n$ is a subset R of $A_1 \times \cdots \times A_n$ such that for all $(a_1, \ldots, a_n)$ in $A_1 \times \cdots \times A_n$ there exist $(r_1, \ldots, r_n) \in R$ such that $(a_1, \ldots, a_n)$ and $(r_1, \ldots, r_n)$ differ in at most one place. That is,*

$$|\{i \;:\; a_i \neq r_i\}| \leq 1.$$

We call a rook covering R exact if for all $(a_1, \ldots, a_n) \in A_1 \times \cdots \times A_n$ there exists a unique $(r_1, \ldots, r_n) \in R$ with the above conditions.

Let $d_1, \ldots, d_n$ be integers. We will say that there is a rook covering of $d_1 \times \cdots \times d_n$ of cardinality N, if there a is rook covering of $A_1 \times \cdots \times A_n$ of cardinality N, where A_i has size d_i for $i = 1, \ldots, n$.

We are interested in finding rook coverings R, where the cardinality of R is as small as possible. Since each rook covers $d_1 + \cdots + d_n - n + 1$ elements, the cardinality of a rook covering is at least

$$|R| \geq \frac{d_1 \cdots d_n}{d_1 + \cdots + d_n - n + 1}.$$

This bound is an equality if the rook covering is exact.

Trivially we have that there is a rook covering of d of size 1, and this is best possible. Just place the rook arbitrarily.

Lemma 1.2.1 *The smallest rook covering of $d_1 \times d_2$ has size* $\min(d_1, d_2)$.

Proof: Let $A_i = \{1, \ldots, d_i\}$. Then the following is a rook covering of $A_1 \times A_2$.

$$\{(k,k) \; : \; 1 \leq k \leq \min(d_1, d_2)\} .$$

To show that it is minimal, assume that there is a rook covering with size less than $\min(d_1, d_2)$. Then there is a row that do not contain a rook, and there is a column that do not contain a rook. The intersection of this row and this column is an element not covered by R, thus leading to contradiction. □

Lemma 1.2.2 *If there is a rook covering of $d_1 \times \cdots \times d_n$ using N rooks, there is a rook covering of $d \times d_1 \times \cdots \times d_n$ using $d \cdot N$ rooks.*

Proof: Assume that R is a rook covering of $A_1 \times \cdots \times A_n$. Then $A \times R$ is a rook covering of $A \times A_1 \times \cdots \times A_n$. □

Proposition 1.2.3 *There is a rook covering of $d \times d \times d$ using N rooks, where*

$$N = \begin{cases} \frac{1}{2} \cdot d^2 & \text{if } d \text{ is even,} \\ \frac{1}{2} \cdot (d^2 + 1) & \text{if } d \text{ is even.} \end{cases}$$

Moreover, this is the minimal cardinality of such rook coverings.

Proof: Let $b = \lfloor \frac{d}{2} \rfloor$, and $c = \lceil \frac{d}{2} \rceil$. Thus $b + c = d$, and $c - b \leq 1$. Observe that $N = b^2 + c^2$. Let $B = \{1, \ldots, b\}$, $C = \{b+1, \ldots, d\}$, and $A = \{1, \ldots, d\}$. We claim that the following set is a rook covering of A^3.

$$\begin{aligned} R \;&=\; \left\{(x,y,z) \in B^3 \; : \; x + y + z \equiv 0 \pmod{b}\right\} \cup \\ &\qquad \cup \left\{(x,y,z) \in C^3 \; : \; x + y + z \equiv 0 \pmod{c}\right\} . \end{aligned}$$

It is easy to see this claim, since given three numbers belonging to the set A, two of them belong to the same set B or C.

To see that N is the minimal cardinality of a rook covering, consider a rook covering R. Let S be a maximal subset of A^2 subject to the following two conditions.

- $(S \times A) \cap R = \emptyset$. That is, for all rooks $(x, y, z) \in R$ we have that $(x, y) \notin S$.

- S is a partial permutation. That is, let $(x, y), (x', y') \in S$. Then the following two implications hold $x = x' \Longrightarrow y = y'$, and $y = y' \Longrightarrow x = x'$.

Let $|S| = m$. Consider the two projections of S to first and second coordinate. That is, $X = \{x \in A : \text{there is a } y \in A \text{ such that } (x, y) \in S\}$, and $Y = \{y \in A : \text{there is a } x \in A \text{ such that } (x, y) \in S\}$. Consider the following partition of the set A^3 into three disjoints parts

$$\begin{aligned} E &= X \times Y \times A, \\ F &= X \times (A - Y) \times A \cup (A - X) \times Y \times A, \\ G &= (A - X) \times (A - Y) \times A. \end{aligned}$$

Let e, f, g be the number of rooks in each of these subsets.

Since S is maximal, we know that for all $x \in A - X$ and for all $y \in A - Y$ there exists $z \in A$ such that $(x, y, z) \in R$, since otherwise we could extend the set S with the pair (x, y). Hence the set G contains at least $(d - m)^2$ rooks. Thus $g \geq (d - m)^2$.

Let $(x, y) \in S$. Assume there exist $u \in A - X$ such that there is no $z \in A$ satisfying $(u, y, z) \in R$. Assume similarly that there exist $v \in A - Y$ such that there is no rook at (x, v, z). This leads to contradiction since

$$S - \{(x, y)\} \cup \{(x, v), (u, y)\}$$

is larger than S, and thus S is not maximal. Hence we conclude that there are at least $d-m$ rooks in the set

$$\{x\} \times (A-Y) \times A \cup (A-X) \times \{y\} \times A.$$

By summing over all $(x,y) \in S$ we conclude that $f \geq m \cdot (d-m)$.

A rook in E can cover at most two elements in the set $S \times A$. Similarly a rook in F can cover only one element of $S \times A$. Since every element of $S \times A \subseteq A^3$ is covered by some rook, we conclude that $md \leq 2e + f$.

Now we have that

$$\begin{aligned} 2(e+f) &= (2e+f)+f \\ &\geq md + m \cdot (d-m) = 2md - m^2. \end{aligned}$$

Thus the number of rooks in A^3 is

$$\begin{aligned} e+f+g &\geq md - \frac{1}{2} \cdot m^2 + (d-m)^2 \\ &= d^2 - md + \frac{1}{2} \cdot m^2 \\ &= \frac{1}{2} d^2 + \frac{1}{2} \cdot (d-m)^2 \geq \frac{1}{2} \cdot d^2. \end{aligned}$$

Since $\left\lceil \frac{1}{2} \cdot d^2 \right\rceil = N$, the proof is complete. □

From **[Ro]** we have the following proposition

Proposition 1.2.4 *For $n \geq 2$ a rook covering of $d^n = d \times \cdots \times d$ has at least $\frac{d^{n-1}}{n-1}$ elements. This lower bound cannot be attained unless d is divisible by $n-1$.*

Proposition 1.2.5 *If there is a rook covering of d^n using N rooks, then there is rook covering of $(c \cdot d)^n$ using $c^{n-1} \cdot N$ rooks.*

Proof: Let A be a set of size d and let $C = \{1, \ldots, c\}$. Let R be rook covering of A^n. Consider the set $(A \times C)^n$ and consider the subset R' of $(A \times C)^n$ defined as follows.

$$\{((a_1, x_1), \ldots, (a_n, x_n)) \in (A \times C)^n \ : \ (a_1, \ldots, a_n) \in R, \ x_1 + \cdots + x_n \equiv 0 \pmod{c}\}.$$

It is a direct verification that R' is rook covering of cardinality $c^{n-1} \cdot N$. □

Proposition 1.2.6 *Let q be a prime power, and let k be a positive integer. Let*

$$n = \frac{q^k - 1}{q - 1}.$$

Then there is a perfect rook covering of

$$q^n = \underbrace{q \times \cdots \times q}_{n}$$

of size q^{n-k}.

Proof: Observe that a perfect rook covering of q^d can be viewed as a error correcting code that corrects one error. Example of such error correcting codes are Hamming codes, and they exists under the above conditions **[Lint]**. □

This result and its proof was already pointed out in **[Lo]**.

The two smallest examples of this proposition are:

- 2 rooks cover $2^3 = 2 \times 2 \times 2$,
- 9 rooks cover $3^4 = 3 \times 3 \times 3 \times 3$.

In general to find a minimal rook covering is hard. In the case of of $3^5 = 3 \times 3 \times 3 \times 3 \times 3$ one can easily find a rook covering that has 27 rooks. In fact, such a covering is minimal **[K-L]**. Other references on this problem are **[Bl-Lam]**, **[F-R]**, **[We]** and **[Wi]**.

1.2.2 Multi dimensional matrices

Let U be a linear space over the complex numbers. We define the tensor algebra $\mathrm{Tens}[U]$ to be

$$\mathrm{Tens}[U] = \bigoplus_{n\geq 0} U^{\otimes n},$$

where $U^{\otimes 0} = \mathbf{C}$, and $U^{\otimes n+1} = U \otimes U^{\otimes n}$. The multiplication on $\mathrm{Tens}[U]$ is defined by

$$(u_1 \otimes \cdots \otimes u_i) \cdot (u_1' \otimes \cdots \otimes u_j') = u_1 \otimes \cdots \otimes u_i \otimes u_1' \otimes \cdots \otimes u_j',$$

and extended by linearity.

Let $W_1, \ldots, W_s$ be finite dimensional linear subspaces of U, where W_j has dimension d_j. Let $z_{j,1}, \ldots, z_{j,d_j}$ be a basis for W_j. Observe now that the linear space of matrices of size $d_1 \times \cdots \times d_s$ is isomorphic to to the space $W_1 \otimes \cdots \otimes W_s$ by the linear map

$$M = (m_{i_1,\ldots,i_s})_{1\leq i_1\leq d_1,\ldots,1\leq i_s\leq d_s} \longmapsto \sum_{1\leq i_1\leq d_1,\ldots,1\leq i_s\leq d_s} m_{i_1,\ldots,i_s} z_{1,i_1} \otimes \cdots \otimes z_{s,i_s}.$$

Definition 1.2.2 *An element v in the linear space $W_1 \otimes \cdots \otimes W_s$ has* rank k *if the k is the smallest integer m such that there exists $y_{j,i} \in W_j$ for $1 \leq j \leq s$ and $1 \leq i \leq m$ such that*

$$v = \sum_{i=1}^{m} y_{1,i} \otimes \cdots \otimes y_{s,i}.$$

This definition of rank agrees with the ordinary rank of two dimensional matrices.

Definition 1.2.3 *The linear space $W_1 \otimes \cdots \otimes W_s$ has* essential rank k *if the k is the smallest integer m such that the set of elements v in $W_1 \otimes \cdots \otimes W_s$ of rank at most m form a dense set in the Euclidean topology.*

Observe that the maximal rank of $W_1 \otimes \cdots \otimes W_s$ might be different from the essential rank of $W_1 \otimes \cdots \otimes W_s$. That it can be so, we will see in Section 1.5 about $2 \times 2 \times 2$ matrices.

1.2.3 Essential rank and rook coverings

Theorem 2 *Let W_i be linear spaces of dimension d_i. If there is rook covering of $d_1 \times \cdots \times d_s$ of cardinality N, then the essential rank of $V = W_1 \odot \cdots \odot W_s$ is less than or equal to N.*

Proof: Let $z_{j,1}, \ldots, z_{j,d_j}$ be a basis for W_j. Similarly let $z^*_{j,1}, \ldots, z^*_{j,d_j}$ be a the dual basis for W^*_j. That is $\left\langle z^*_{j,i} \,\middle|\, z_{j,k} \right\rangle = \delta_{i,k}$. Let U be a linear space which contains $W_1, \ldots, W_s$ as subspaces. Then Tens$[U]$ is an algebra, where $W_1, \ldots, W_s$ and V are linear subspaces. Clearly the polynomial

$$p = \sum_{i=1}^{N} x_{1,i} \cdots x_{s,i},$$

in $C\{x_{1,1}, \ldots, x_{1,N}, \ldots, x_{s,1}, \ldots, x_{s,N}\}$ is homogeneous with respect to V, $\underbrace{W_1, \ldots, W_1}_{N}, \ldots, \underbrace{W_s, \ldots, W_s}_{N}$.

In the language of Theorem 1 we would like to show that any generic element of V can be written in the form

$$\text{eval}\,(p;\ w_{1,1}, \ldots, w_{1,N}, \ldots, w_{s,1}, \ldots, w_{s,N}), \tag{1.3}$$

where $w_{j,i} \in W_j$. By Theorem 1 we would like to show that there exist $w'_{j,i} \in W_j$ for $j = 1, \ldots, s$ and $i = 1, \ldots, N$ such that there is no nonzero dual element $L \in V^*$ apolar to all the linear maps

$$y_{j,i} \longmapsto \text{eval}\left(D_{t,x_{j,i}}p;\ y_{j,i}, w'_{1,1}, \ldots, w'_{s,N}\right), \tag{1.4}$$

where $y_{j,i} \in W_j$. But the polarization computes to

$$\begin{aligned} y_{j,i} \quad &\longmapsto \quad \text{eval}\left(x_{1,i} \cdots x_{j-1,i} \cdot t \cdot x_{j+1,i} \cdots x_{s,i};\ y_{j,i}, w'_{1,1}, \ldots, w'_{s,N}\right) \\ &= \quad w'_{1,i} \cdots w'_{j-1,i} \cdot y_{j,i} \cdot w'_{j+1,i} \cdots w'_{s,i}. \end{aligned}$$

Let A_j be the set $\{1, \ldots, d_j\}$. Let R be a rook covering of $A_1 \times \cdots \times A_s$ of cardinality

R. Thus let

$$R = \{(r_{1,i}, \ldots, r_{s,i}) \ : \ 1 \leq i \leq N\}.$$

Choose $w'_{j,i} = z_{j,r_{j,i}}$

Since we only have finite dimensional linear spaces, we know that the dual of a tensor product is the tensor product of the duals. That is,

$$V^* = W_1^* \otimes \cdots \otimes W_s^*.$$

Assume that L in V^* is apolar to all the linear maps in (1.4). We can write the dual element L in terms of the dual basis.

$$L = \sum_{(i_1,\ldots,i_s)\in A_1\times\cdots\times A_s} \beta_{i_1,\ldots,i_s} z_{1,i_1}^* \otimes \cdots \otimes z_{s,i_s}^*.$$

Consider an element $(i_1, \ldots, i_s) \in A_1 \times \cdots \times A_s$. Since R is a rook covering, there is an element $(r_{1,k}, \ldots, r_{s,k}) \in R$ that only differ in at most one coordinate from $(i_1, \ldots, i_s)$. Let the coordinate where $(i_1, \ldots, i_s)$ and $(r_{1,k}, \ldots, r_{s,k})$ differ be j. If they are the same, choose j arbitrary. Thus we know that L is apolar to the linear map $y_{j,k} \mapsto z_{1,r_{1,k}} \cdots z_{j-1,r_{j-1,k}} \cdot y_{j,k} \cdot z_{j+1,r_{j+1,k}} \cdots z_{s,r_{s,k}}$. Let $y_{j,k}$ take the value of z_{j,i_j}. Hence

$$\begin{aligned} 0 &= \left\langle L \,\middle|\, z_{1,r_{1,k}} \cdots z_{j-1,r_{j-1,k}} \cdot z_{j,i_j} \cdot z_{j+1,r_{j+1,k}} \cdots z_{s,r_{s,k}} \right\rangle \\ &= \left\langle L \,\middle|\, z_{1,i_1} \cdots z_{j-1,i_{j-1}} \cdot z_{j,i_j} \cdot z_{j+1,i_{j+1}} \cdots z_{s,i_s} \right\rangle \\ &= \beta_{i_1,\ldots,i_s}. \end{aligned}$$

We conclude that all coefficients of L vanish. Thus we know that $L = 0$. Theorem 1 implies that any generic element of V can be written in the form (1.3). Thus the linear space $V = W_1 \otimes \cdots \otimes W_s$ has essential rank less than or equal to N. □

Corollary 1.2.1 *Let W_1 and W_2 be two linear space of dimensions d_1 and d_2. Then the* *linear space* $W_1 \otimes W_2$ *has essential rank* $\min(d_1, d_2)$.

Proof: By Lemma 1.2.1 we know that the essential rank of $W_1 \otimes W_2$ is at most $\min(d_1, d_2)$. Let us assume that that the essential rank is less than $\min(d_1, d_2)$ in order to reach a contradiction. That implies that a generic element of $W_1 \otimes W_2$ can be written in the form

$$\sum_{i=1}^{k} w_{1,i} \otimes w_{2,i},$$

where $w_{j,i} \in W_j$ and $k < \min(d_1, d_2)$.

By Theorem 1 we know that we can find elements $w'_{j,i} \in W_j$ for $j = 1, 2$ and $i = 1, \ldots, k$ such that there is no nonzero $L \in W_1^* \otimes W_2^*$ apolar to all the maps

$$t \longmapsto t \otimes w_{2,i} \quad \text{and } t \longmapsto w_{1,i} \otimes t.$$

Since W_j has dimension greater than k, we can find a nonzero element $y_j \in W_j^*$ such that $\langle y_j \mid w'_{j,i} \rangle = 0$ for all $i = 1, \ldots, k$. Thus $y_1 \otimes y_2$ is a nonzero element in $(W_1 \otimes W_2)^*$. It is easy to see that $y_1 \otimes y_2$ is apolar to all the maps above, which leads us to the desired contradiction. □

Corollary 1.2.2 *Let* W *be a linear space of dimension* d. *Then* $W^{\otimes 3}$ *has essential rank less than or equal than* $\left\lceil \frac{d^2}{2} \right\rceil$.

Corollary 1.2.3 *Let* W *be a linear space of dimension* q, *where* q *is a prime power. Let* k *be a positive integer, and*

$$n = \frac{q^k - 1}{q - 1}.$$

Then $W^{\otimes n}$ *has essential rank less than or equal than* q^{n-k}.

Analogously to Proposition 1.2.5 we have that

Proposition 1.2.7 *Let* W *and* U *be linear spaces. If the linear space* $W^{\otimes s}$ *has essential rank* N, *then the linear space* $(W \otimes U)^{\otimes s}$ *has essential rank at most* $d^{s-1} \cdot N$, *where* d *is the dimension of* U.

Proof: We know that a generic element of $W^{\otimes s}$ can be written in the form

$$\sum_{i=1}^{N} w_{1,i} \otimes \cdots \otimes w_{s,i}.$$

By Theorem 1 this canonical form implies that there exist $w'_{j,i}$, $1 \le j \le s$ and $1 \le i \le N$, such that there is no nonzero element in $(W^{\otimes s})^*$ apolar to all the maps

$$t_{j,i} \longmapsto w'_{1,i} \otimes \cdots \otimes t_{j,i} \otimes \cdots \otimes w'_{s,i},$$

where $t_{j,i} \in W$.

Let $z_1, \ldots, z_n$ be a basis of W, and let $z_1^*, \ldots, z_n^*$ be the dual basis of W^*. Given $L \in (W^{\otimes s})^*$, which we can expand as

$$L = \sum_{(i_1,\ldots,i_s)\in\{1,\ldots,n\}^s} \beta_{i_1,\ldots,i_s} z_{i_1}^* \otimes \cdots \otimes z_{i_s}^*.$$

Thus we know that for all $(i_1, \ldots, i_s) \in \{1, \ldots, n\}^s$ there exist $y_{j,i} \in W$ for $1 \le j \le s$ and $1 \le i \le N$, such that

$$\left\langle L \,\middle|\, \sum_{j=1}^{s}\sum_{i=1}^{N} w'_{1,i} \otimes \cdots \otimes y_{j,i} \otimes \cdots \otimes w'_{s,i} \right\rangle = \beta_{i_1,\ldots,i_s}.$$

This can also be written as

$$\sum_{j=1}^{s}\sum_{i=1}^{N} w'_{1,i} \otimes \cdots \otimes y_{j,i} \otimes \cdots \otimes w'_{s,i} = z_{i_1} \otimes \cdots \otimes z_{i_s}. \tag{1.5}$$

Let $u_1, \ldots, u_d$ be a basis of U. Consider the following subset of $\{1, \ldots, d\}^s$.

$$P = \{(p_1, \ldots, p_s) \in \{1, \ldots, d\}^s \ : \ p_1 + \cdots + p_s \equiv 0 \pmod{d}\}.$$

The cardinality of P is d^{s-1}. Define the elements $u'_{j,i,\mathbf{p}} \in W \otimes U$, where $j = 1, \ldots, s$,

$i = 1, \ldots, N$, and $\mathbf{p} \in P$, by

$$u'_{j,i,\mathbf{p}} = w'_{j,i} \otimes u_{p_j}.$$

Thus there are $s \cdot N \cdot d^{s-1}$ such elements. Consider now the linear maps

$$t_{j,i,\mathbf{p}} \longmapsto u'_{1,i,\mathbf{p}} \otimes \cdots \otimes t_{j,i,\mathbf{p}} \otimes \cdots \otimes u'_{s,i,\mathbf{p}}, \tag{1.6}$$

where $t_{j,i,\mathbf{p}} \in W \otimes U$. We would like to show that there is no nonzero element in $((W \otimes U)^{\otimes s})^*$ apolar to all the maps above. It is sufficient to prove the equivalent statement that the images of the maps above span the linear space $(W \otimes U)^{\otimes s}$. Choose $(i_1, \ldots, i_s) \in \{1, \ldots, n\}^s$ and $(q_1, \ldots, q_s) \in \{1, \ldots, d\}^s$. They correspond to the basis element

$$(z_{i_1} \otimes u_{q_1}) \otimes \cdots \otimes (z_{i_s} \otimes u_{q_s})$$

of $(W \otimes U)^{\otimes s}$. As observed before, we can find $y_{j,i} \in W$ for $1 \leq j \leq s$ and $1 \leq i \leq N$, such that equation (1.5) is satisfied. Let $\mathbf{p}^j$ be the element of P such that

$$p^j_k = \begin{cases} q_k & \text{if } k \neq j, \\ q_k - (q_1 + \cdots + q_s) \pmod{d} & \text{if } k = j. \end{cases}$$

Thus $\mathbf{q}$ and $\mathbf{p}^j$ only differ in the jth coordinate. Let

$$\hat{y}_{j,i} = y_{j,i} \otimes u_{q_j}.$$

The element $u'_{1,i,\mathbf{p}^j} \otimes \cdots \otimes y_{j,i,\mathbf{q}^i} \otimes \cdots \otimes u'_{s,i,\mathbf{p}^j}$ lies in the image of one maps in (1.6). Consider now the sum of such elements.

$$\begin{aligned} & \sum_{i=1}^{N} \sum_{j=1}^{s} u'_{1,i,\mathbf{p}^j} \otimes \cdots \otimes y_{j,i,\mathbf{q}^i} \otimes \cdots \otimes u'_{s,i,\mathbf{p}^j} \\ = & \sum_{i=1}^{N} \sum_{j=1}^{s} \left(w'_{1,i} \otimes u_{p^j_1}\right) \otimes \cdots \otimes \left(y_{j,i} \otimes u_{q_j}\right) \otimes \cdots \otimes \left(w'_{s,i} \otimes u_{p^j_s}\right) \end{aligned}$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{s}\left(w'_{1,i}\otimes u_{q_1}\right)\otimes\cdots\otimes\left(y_{j,i}\otimes u_{q_j}\right)\otimes\cdots\otimes\left(w'_{s,i}\otimes u_{q_s}\right).$$

Observe that the two linear spaces $(W\otimes U)^{\otimes s}$ and $W^{\otimes s}\otimes U^{\otimes s}$ are naturally isomorphic. Let Φ be this isomorphism, that is, $\Phi : W^{\otimes s}\otimes U^{\otimes s}\longrightarrow (W\otimes U)^{\otimes s}$. We can view Φ as a reordering of the terms. Now, the above element in $(W\otimes U)^{\otimes s}$ can be written as

$$\begin{aligned}
& \Phi\left(\sum_{i=1}^{N}\sum_{j=1}^{s} w'_{1,i}\otimes\cdots\otimes y_{j,i}\otimes\cdots\otimes w'_{s,i}\otimes u_{q_1}\otimes\cdots\otimes u_{q_j}\otimes\cdots\otimes u_{q_s}\right)\\
= & \Phi\left(\left(\sum_{i=1}^{N}\sum_{j=1}^{s} w'_{1,i}\otimes\cdots\otimes y_{j,i}\otimes\cdots\otimes w'_{s,i}\right)\otimes u_{q_1}\otimes\cdots\otimes u_{q_s}\right)\\
= & \Phi\left(z_{i_1}\otimes\cdots\otimes z_{i_s}\otimes u_{q_1}\otimes\cdots\otimes u_{q_s}\right)\\
= & \left(z_{i_1}\otimes u_{q_1}\right)\otimes\cdots\otimes\left(z_{i_s}\otimes u_{q_s}\right).
\end{aligned}$$

But this is the basis element we chose. Hence we conclude that the images of the linear maps (1.6) span $(W\otimes U)^{\otimes s}$. □

1.3 Skew-symmetric tensors

1.3.1 Skew-symmetric tensors and rank

Let W be a linear space and let $\mathrm{Ext}(W)$ be the exterior algebra of W. Recall that the exterior algebra is graded, thus we can write

$$\mathrm{Ext}(W) = \bigoplus_{s\geq 0}\mathrm{Ext}_s(W),$$

where $\mathrm{Ext}_s(W)$ are the skew-symmetric tensor of step s.

Definition 1.3.1 *An element v in the linear space* $\mathrm{Ext}_s(W)$ *has* rank k *if the k is the smallest integer m such that there exists $y_{j,i} \in W$ for $1 \leq j \leq s$ and $1 \leq i \leq m$ such that*

$$v = \sum_{i=1}^{m} y_{1,i} \wedge \cdots \wedge y_{s,i}.$$

Definition 1.3.2 *The linear space* $\mathrm{Ext}_s(W)$ *has* essential rank k *if the k is the smallest integer m such that the set of elements v in* $\mathrm{Ext}_s(W)$ *that has rank less or equal to m form a dense set in the Euclidean topology.*

Observe that the maximal rank of $\mathrm{Ext}_s(W)$ might be different from the essential rank of $\mathrm{Ext}_s(W)$.

Proposition 1.3.1 *Let* $\dim(W) = n$. *Then the essential rank of* $\mathrm{Ext}_2(W)$ *is equal to* $N = \lfloor \frac{n}{2} \rfloor$.

Proof: Let $z_1, \ldots, z_n$ be a basis for W. Similarly let $z_1^*, \ldots, z_n^*$ be a the dual basis for W^*. That is $\langle z_i^* \mid z_k \rangle = \delta_{i,k}$. The polynomial

$$p = \sum_{i=1}^{N} x_{1,i} x_{2,i},$$

in $C\{x_{1,1}, \ldots, x_{2,N}\}$ is homogeneous with respect to $V, \underbrace{W, \ldots, W}_{2 \cdot N}$ in the exterior algebra.

We would like to show that any generic element of V can be written in the form

$$\mathrm{eval}\,(p;\ w_{1,1}, \ldots, w_{1,N}, w_{2,1}, \ldots, w_{2,N}), \tag{1.7}$$

where $w_{j,i} \in W$. Thus by Theorem 1 we would like to show that there exist $w'_{j,i} \in W$ for $j = 1, 2$ and $i = 1, \ldots, N$ such that there is no nonzero dual element $L \in V^*$ apolar to all the linear maps

$$y_{j,i} \longmapsto \mathrm{eval}\left(D_{t,x_{j,i}} p;\ y_{j,i}, w'_{1,1}, \ldots, w'_{3,N}\right), \tag{1.8}$$

where $y_{j,i} \in W$. But the polarization computes to

$$\begin{aligned} y_{1,i} &\longmapsto \text{eval}\left(t \cdot x_{2,i}\,;\, y_{j,i}, w'_{1,1}, \ldots, w'_{2,N}\right) \\ &= y_{1,i} \wedge w'_{2,i}, \end{aligned}$$

for $j = 1$. Similarly for $j = 2$ we get $y_{2,i} \longmapsto w'_{1,i} \wedge y_{2,i}$.

Let $w'_{j,i} = z_{j+2\cdot(i-1)}$, for $j = 1, 2$, and $i = 1, \ldots, N$.

Since W is a finite dimensional linear space, we know that

$$V^* = \text{Ext}_2(W^*).$$

Assume that L in V^* is apolar to all the linear maps in (1.8). We can write the dual element L in terms of the dual basis.

$$L = \sum_{1 \leq p < q \leq n} \beta_{p,q} z_p^* \wedge z_q^*.$$

Consider a pair $\{p, q\}$, such that $p < q$. Since $p < q \leq n$, we have that $1 \leq p \leq 2 \cdot N$. We can write $p = k + 2 \cdot (i - 1)$, where $1 \leq k \leq 2$ and $1 \leq i \leq N$. Now consider the map for $j = 3 - k$ and for this value of i, and with $y_{j,i} = z_q$. If $j = 1$ the computation looks like

$$\begin{aligned} \left\langle L \,\middle|\, y_{1,i} \wedge w'_{2,i} \right\rangle &= \langle L \mid z_q \wedge z_p \rangle \\ &= -\beta_{p,q}. \end{aligned}$$

Similarly for $j = 2$ we get $\left\langle L \,\middle|\, w'_{1,i} \wedge y_{2,i} \right\rangle = \beta_{p,q}$. The dual element L is apolar to this map, we get that $\beta_{p,q} = 0$. Since p and q is arbitrarily, we have proven that $L = 0$. Theorem 1 implies that any generic element of V can be written in the form (1.7). Thus the linear space $V = \text{Ext}_2(W)$ has essential rank less than or equal to N. □

In fact, one can also show that the maximal rank of $\text{Ext}_2(W)$ is equal to $\lfloor \frac{n}{2} \rfloor$.

1.3.2 Essential rank of skew-symmetric tensors of step 3

For a set A define the set $\binom{A}{k}$ to be the set of all subsets of A of cardinality k. This notation is analog to the notation of the binomial coefficient, since $\left|\binom{A}{k}\right| = \binom{|A|}{k}$.

Definition 1.3.3 *A* Steiner triple system *on a non-empty set A is a subset S of $\binom{A}{3}$ such that for all pairs $P \in \binom{A}{2}$ there is a unique triple $Q \in S$ such that $P \subseteq Q$.*

Proposition 1.3.2 *A necessary and sufficient condition for a Steiner triple system to exist on a set A of cardinality n is that $n \equiv 1, 3 \bmod 6$.*

It is easy to see that the condition that $n \equiv 1, 3 \bmod 6$ is a necessary condition for a Steiner triple system to exist on a set of cardinality n. Observe also that the size of a Steiner triple system is $\frac{1}{3} \cdot \binom{n}{2} = \frac{n \cdot (n-1)}{6}$.

Proposition 1.3.3 *If there is Steiner triple systems of order n and m, then there is Steiner triple systems of order $n \cdot m$.*

This proposition could be directly proven by Proposition 1.3.2. But one can construct a triple system on a $n \cdot m$ set, by using the triple systems on the n set and the m set, see **[Ry]** Chapter 8, Theorem 1.2.

Proposition 1.3.4 *Let W be a linear spaces of dimension $n + m$. If there is Steiner triple system on a n-set and on a m-set, then the essential rank of $V = \mathrm{Ext}_3(W)$ is less than or equal to $N = \frac{n \cdot (n-1) + m \cdot (m-1)}{6}$.*

Proof: Let $z_1, \ldots, z_{n+m}$ be a basis for W. Similarly let $z_1^*, \ldots, z_{n+m}^*$ be a the dual basis for W^*. That is $\langle z_i^* \mid z_k \rangle = \delta_{i,k}$. Clearly the polynomial

$$p = \sum_{i=1}^{N} x_{1,i} x_{2,i} x_{3,i},$$

in $C\{x_{1,1}, \ldots, x_{3,N}\}$ is homogeneous with respect to V, $\underbrace{W, \ldots, W}_{3 \cdot N}$ in the exterior algebra.

We would like to show that any generic element of V can be written in the form

$$\operatorname{eval}(p;\ w_{1,1},\ldots,w_{1,N},\ldots,w_{3,1},\ldots,w_{3,N}), \tag{1.9}$$

where $w_{j,i} \in W$. Thus by Theorem 1 we would like to show that there exist $w'_{j,i} \in W$ for $j = 1,2,3$ and $i = 1,\ldots,N$ such that there is no nonzero dual element $L \in V^*$ apolar to all the linear maps

$$y_{j,i} \longmapsto \operatorname{eval}\left(D_{t,x_{j,i}}p;\ y_{j,i}, w'_{1,1},\ldots,w'_{3,N}\right), \tag{1.10}$$

where $y_{j,i} \in W$. But the polarization computes to

$$\begin{aligned} y_{1,i} &\longmapsto \operatorname{eval}\left(t\cdot x_{2,i}\cdot x_{3,i};\ y_{j,i}, w'_{1,1},\ldots,w'_{3,N}\right) \\ &= y_{1,i}\wedge w'_{2,i}\wedge w'_{3,i}, \end{aligned}$$

for $j = 1$. Similarly for $j = 2$ and $j = 3$ we get

$$\begin{aligned} y_{2,i} &\longmapsto w'_{1,i}\wedge y_{2,i}\wedge w'_{3,i}, \\ y_{3,i} &\longmapsto w'_{1,i}\wedge w'_{2,i}\wedge y_{3,i}. \end{aligned}$$

Let S_1 be a Steiner triple system on the set $\{1,\ldots,n\}$ and let S_2 be a Steiner triple system on the set $\{n+1,\ldots,n+m\}$. For $i = 1,2$, let $N_i = |S_i|$. Thus we have that $N_1 + N_2 = N$. Assume that we can write

$$S_1 = \{\{a_{1,i}, a_{2,i}, a_{3,i}\}\}_{i=1,\ldots,N_1}, \quad S_2 = \{\{a_{1,i}, a_{2,i}, a_{3,i}\}\}_{i=N_1+1,\ldots,N_1+N_2},$$

where $a_{1,i} < a_{2,i} < a_{3,i}$. Let

$$w'_{j,i} = z_{a_{j,i}},$$

for $j = 1,2,3$, and $i = 1,\ldots,N$.

Since W is a finite dimensional linear space, we know that

$$V^* = \mathrm{Ext}_3(W^*).$$

Assume that L in V^* is apolar to all the linear maps in (1.10). We can write the dual element L in terms of the dual basis.

$$L = \sum_{1 \leq p < q < r \leq n+m} \beta_{p,q,r} z_p^* \wedge z_q^* \wedge z_r^*.$$

Consider a triplet $\{p, q, r\}$, such that $p < q < r$. Either $q \leq n$ or $q \geq n+1$. The two case are similar, thus with out loss of generality we can assume that $q \leq n$. Thus $\{p, q\} \in \binom{\{1,\ldots,n\}}{2}$. Since S_1 is a Steiner triple system there exist $s \in \{1, \ldots, n\}$ such that $\{p, q, s\} \in S_1$. Say that $\{p, q, s\} = \{a_{1,i}, a_{2,i}, a_{3,i}\}$ for some $i = 1, \ldots, N_1$. So $s = a_{j,i}$ for some $j = 1, 2, 3$. Since these three cases are similar, we can assume that $j = 3$. Thus we know that L is apolar to the linear map $y_{3,i} \mapsto z_{a_{1,i}} \wedge z_{a_{2,i}} \wedge y_{3,i}$. Hence, by letting $y_{3,i} = z_r$, we get

$$\begin{aligned} 0 &= \left\langle L \,\middle|\, z_{a_{1,i}} \wedge z_{a_{2,i}} \wedge z_r \right\rangle \\ &= \langle L \mid z_p \wedge z_q \wedge z_r \rangle \\ &= \beta_{p,q,r}. \end{aligned}$$

We conclude that all coefficients of L vanish. Thus we know that $L = 0$. Theorem 1 implies that any generic element of V can be written in the form (1.9). Thus the linear space $V = \mathrm{Ext}_3(W)$ has essential rank less than or equal to N. □

Corollary 1.3.1 *Let* $\dim(W) = n$. *If* $n \equiv 2, 6 \bmod 12$ *then the essential rank of* $\mathrm{Ext}_3(W)$ *is less than or equal to* $\frac{n \cdot (n-2)}{12}$.

Proof: Since $\frac{n}{2} \equiv 1, 3 \bmod 6$, and by previous proposition, the essential rank of $\mathrm{Ext}_3(W)$ is less than or equal to

$$\frac{2 \cdot \frac{n}{2} \cdot \left(\frac{n}{2} - 1\right)}{6} = \frac{n \cdot (n-2)}{12}.$$

□

Corollary 1.3.2 *Let* $\dim(W) = n$. *If* $n \equiv 4 \bmod 12$ *then the essential rank of* $\mathrm{Ext}_3(W)$ *is less than or equal to* $\frac{n \cdot (n-2)+4}{12}$.

Proof: Since $\frac{n}{2} \equiv 2 \bmod 6$. Observe that $\frac{n}{2} - 1 \equiv 1 \bmod 6$, and $\frac{n}{2} + 1 \equiv 3 \bmod 6$. Thus by previous proposition, the essential rank of $\mathrm{Ext}_3(W)$ is less than or equal to

$$\frac{\left(\frac{n}{2} - 1\right) \cdot \left(\frac{n}{2} - 2\right) + \left(\frac{n}{2} + 1\right) \cdot \frac{n}{2}}{6} = \frac{n \cdot (n-2) + 4}{12}.$$

□

Corollary 1.3.3 *Let* $\dim(W) = n$. *If* $n \equiv 0, 8 \bmod 12$ *then the essential rank of* $\mathrm{Ext}_3(W)$ *is less than or equal to* $\frac{n \cdot (n-2)+36}{12}$.

Proof: Since $\frac{n}{2} \equiv 0, 4 \bmod 6$. Observe that $\frac{n}{2} - 3 \equiv 1, 3 \bmod 6$, and $\frac{n}{2} + 3 \equiv 1, 3 \bmod 6$. Thus by previous proposition, the essential rank of $\mathrm{Ext}_3(W)$ is less than or equal to

$$\frac{\left(\frac{n}{2} - 3\right) \cdot \left(\frac{n}{2} - 4\right) + \left(\frac{n}{2} + 3\right) \cdot \left(\frac{n}{2} + 2\right)}{6} = \frac{n \cdot (n-2) + 36}{12}.$$

□

Corollary 1.3.4 *Let* $\dim(W) = n$. *If* $n \equiv 10 \bmod 12$ *then the essential rank of* $\mathrm{Ext}_3(W)$ *is less than or equal to* $\frac{n \cdot (n-2)+16}{12}$.

Proof: Since $\frac{n}{2} \equiv 5 \bmod 6$. Observe that $\frac{n}{2} - 2 \equiv 3 \bmod 6$, and $\frac{n}{2} + 2 \equiv 1 \bmod 6$. Thus by previous proposition, the essential rank of $\mathrm{Ext}_3(W)$ is less than or equal to

$$\frac{\left(\frac{n}{2}-2\right)\cdot\left(\frac{n}{2}-3\right)+\left(\frac{n}{2}+2\right)\cdot\left(\frac{n}{2}+1\right)}{6} = \frac{n\cdot(n-2)+16}{12}.$$

□

In the case when the dimension of the linear space W is 8, Corollary 1.3.3 says that the essential rank of $\mathrm{Ext}_3(W)$ is less than or equal to 7. But in fact, we do even better as we will see in the next lemma.

Lemma 1.3.5 *Let W be a linear space of dimension* 8. *Then the linear space $V = \mathrm{Ext}_3(W)$ has essential rank less than or equal to* 4.

Proof: Let $z_1, \ldots, z_8$ be a basis for W. Let $z_1^*, \ldots, z_8^*$ be the dual basis for W^*. Then $z_i^* \wedge z_j^* \wedge z_k^*$, where $1 \le i < j < k \le 8$, form a basis for V^*.

Choose $w'_{1,i} = z_{2i-1}$, $w'_{2,i} = z_{2i}$, and $w'_{3,i} = z_{2i-2} - z_{2i+1}$, for $i = 1, 2, 3, 4$ and where indices are counted modulo 8. Assume that $L \in V^*$ is apolar to the following twelve linear maps

$$\begin{aligned} y_{1,i} &\longmapsto y_{1,i} \wedge w'_{2,i} \wedge w'_{3,i} = y_{1,i} \wedge z_{2i} \wedge (z_{2i-2} - z_{2i+1}), \\ y_{2,i} &\longmapsto w'_{1,i} \wedge y_{2,i} \wedge w'_{3,i} = z_{2i-1} \wedge y_{2,i} \wedge (z_{2i-2} - z_{2i+1}), \\ y_{3,i} &\longmapsto w'_{1,i} \wedge w'_{2,i} \wedge y_{3,i} = z_{2i-1} \wedge z_{2i} \wedge y_{3,i}. \end{aligned}$$

We can write

$$L = \sum_{1 \le i < j < k \le 8} \alpha_{\{i,j,k\}} \cdot z_i^* \wedge z_j^* \wedge z_k^*.$$

By using that L is apolar third map with $y_{3,i} = z_k$, we get that $\alpha_{\{2i-1,2i,k\}} = 0$. Since L is apolar to first map with $y_{1,i} = z_{2i-2}$, we have $\alpha_{\{2i-2,2i,2i+1\}} = 0$. Similarly, L is apolar to second map with $y_{2,i} = z_{2i-2}$, we have $\alpha_{\{2i-2,2i-1,2i+1\}} = 0$.

Since L is apolar to the first map with $y_{1,i} = z_{2i+2}$, we get

$$\begin{aligned} 0 &= \langle L \mid z_{2i+2} \wedge z_{2i} \wedge (z_{2i-2} - z_{2i+1}) \rangle \\ &= \text{sign}(2i+2, 2i, 2i-2) \cdot \alpha_{\{2i+2,2i,2i-2\}} - \text{sign}(2i+2, 2i, 2i+1) \cdot \alpha_{\{2i+2,2i,2i+1\}} \\ &= \text{sign}(2i+2, 2i, 2i-2) \cdot \alpha_{\{2i+2,2i,2i-2\}}. \end{aligned}$$

Again, use the first map with $y_{1,i} = z_{2i+4}$.

$$\begin{aligned} 0 &= \langle L \mid z_{2i+4} \wedge z_{2i} \wedge (z_{2i-2} - z_{2i+1}) \rangle \\ &= \text{sign}(2i+4, 2i, 2i-2) \cdot \alpha_{\{2i+4,2i,2i-2\}} - \text{sign}(2i+4, 2i, 2i+1) \cdot \alpha_{\{2i+4,2i,2i+1\}} \\ &= -\text{sign}(2i+4, 2i, 2i+1) \cdot \alpha_{\{2i+4,2i,2i+1\}}. \end{aligned}$$

Use the first linear map with $y_{1,i} = z_{2i+3}$.

$$\begin{aligned} 0 &= \langle L \mid z_{2i+3} \wedge z_{2i} \wedge (z_{2i-2} - z_{2i+1}) \rangle \\ &= \text{sign}(2i+3, 2i, 2i-2) \cdot \alpha_{\{2i+3,2i,2i-2\}} - \text{sign}(2i+3, 2i, 2i+1) \cdot \alpha_{\{2i+3,2i,2i+1\}} \\ &= \text{sign}(2i+3, 2i, 2i-2) \cdot \alpha_{\{2i+3,2i,2i-2\}}. \end{aligned}$$

By symmetry, use the second map with $y_{2,i} = z_{2i-3}$, $y_{2,i} = z_{2i+3}$, and $y_{2,i} = z_{2i+4}$, to conclude that $\alpha_{2i-1,2i-3,2i+1} = 0$, $\alpha_{2i-1,2i+3,2i-2} = 0$, and $\alpha_{2i-1,2i+4,2i+1} = 0$. It is easy to see that we have showned that all coefficients of L vanish, and thus L is equal to 0. By Theorem 1 $\sum_{i=1}^{4} w_{i,1} \wedge w_{i,2} \wedge w_{i,3}$ is a generic canonical form for the linear space $\text{Ext}_3(W)$, and thus the space has essential rank 4. □

Lemma 1.3.6 *Let W be a linear space of dimension* 5. *Then a generic element of the linear space $V = \text{Ext}_3(W)$ can be written in the form*

$$x \wedge y \wedge w + x \wedge u \wedge v,$$

where $x, y, w, u, v \in W$.

Proof: Thus we will consider the following 5 linear maps $t \mapsto t \wedge y' \wedge w' + t \wedge u' \wedge v'$, $t \mapsto x' \wedge t \wedge w'$, $t \mapsto x' \wedge y' \wedge t$, $t \mapsto x' \wedge t \wedge v'$, and $t \mapsto x' \wedge u' \wedge t$.

Let $x' = z_1$, $y' = z_2$, $w' = z_3$, $u' = z_4$, and $v' = z_5$, where $z_1, \ldots, z_5$ is a basis for W.

Assume that L is an element in $\mathrm{Ext}_3(W)^*$ apolar to all the five linear maps above. We can write

$$L = \sum_{1 \leq i < j < k \leq 5} \alpha_{\{i,j,k\}} \cdot z_i^* \wedge z_j^* \wedge z_k^*,$$

where $z_1^*, \ldots, z_5^*$ is the dual basis for W^*.

Since L is apolar to the first map, apply this apolarity condition with $t = z_2$. This implies that $\alpha_{\{2,4,5\}} = 0$. Similarly, the first map with $t = z_3$, $t = z_4$, and $t = z_5$ implies that $\alpha_{\{3,4,5\}} = 0$, $\alpha_{\{2,3,4\}} = 0$, and $\alpha_{\{2,3,5\}} = 0$.

Since L is apolar to second map, apply this condition with $t = z_i$. We see that $\alpha_{\{1,i,3\}} = 0$. Similary, by using the three other linear maps, we conclude that $\alpha_{\{1,i,j\}} = 0$. Hence L is equal to 0, and by Theorem 1 the proof of the lemma is done. □

1.4 Commutative algebras

When the underlying algebras are commutative there is an alternative way to rewrite theorem 1.

1.4.1 Polarizations and derivations

Define an algebra homomorphism ϕ that maps $A\{x_1, \ldots, x_n\}$ into the algebra of polynomials $A[x_1, \ldots, x_n]$, where the variables commute, by $\phi(x_i) = x_i$, and $\phi(a) = a$ for $a \in A$. Hence ϕ maps the monomial $a_1 x_{i_1} a_2 x_{i_2} a_3 \cdots a_m x_{i_m} a_{m+1}$ into $a_1 a_2 a_3 \cdots a_m a_{m+1} \cdot x_{i_1} x_{i_2} \cdots x_{i_m}$. Observe that ϕ is a surjective homomorphism.

Lemma 1.4.1 *Let B be a commutative algebra that has the algebra A as a subalgebra. Then for any element $p \in A\{x_1, \ldots, x_n\}$ and for $b_1, \ldots, b_n \in B$ we have that*

$$\text{eval}\,(p;\ b_1, \ldots, b_n) = \phi(p)(b_1, \ldots, b_n).$$

For $f \in A[x_1, \ldots, x_n]$ denote the derivative of f in the variable x_i by $D_{x_i}(f)$.

Lemma 1.4.2 *Let B be a commutative algebra that has the algebra A as a subalgebra. Then for any element $p \in A\{x_1, \ldots, x_n\}$ and for $b, b_1, \ldots, b_n \in B$ we have that*

$$\text{eval}\,(D_{t,x_i}p;\ b, b_1, \ldots, b_n) = b \cdot (D_{x_i}\phi(p))\,(b_1, \ldots, b_n).$$

Proof: The proof is by induction on $p \in A\{x_1, \ldots, x_n\}$.

- $p = a \in A$. Both sides vanish

$$\text{eval}\,(D_{t,x_i}a\,;\ b, b_1, \ldots, b_n) = \text{eval}\,(0\,;\ b, b_1, \ldots, b_n) = 0,$$

$$b \cdot (D_{x_i}\phi(a))\,(b_1, \ldots, b_n) = b \cdot (D_{x_i}a)\,(b_1, \ldots, b_n) = b \cdot (0)\,(b_1, \ldots, b_n) = 0.$$

- $p = x_j$, where $j \neq i$. Both sides vanish as above.

$$\begin{aligned}
\text{eval}\,(D_{t,x_i}x_i\,;\ b, b_1, \ldots, b_n) &= \text{eval}\,(t\,;\ b, b_1, \ldots, b_n) \\
&= b \\
&= b \cdot (1)\,(b_1, \ldots, b_n) \\
&= b \cdot (D_{x_i}x_i)\,(b_1, \ldots, b_n) \\
&= b \cdot (D_{x_i}\phi(x_i))\,(b_1, \ldots, b_n).
\end{aligned}$$

- $p = r + s$, where $r, s \in A\{x_1, \ldots, x_n\}$.

$$\begin{aligned}
\text{eval}\,(D_{t,x_i}(r+s)\,;\, b, b_1, \ldots, b_n) &= \text{eval}\,(D_{t,x_i}r + D_{t,x_i}s\,;\, b, b_1, \ldots, b_n) \\
&= \text{eval}\,(D_{t,x_i}r\,;\, b, b_1, \ldots, b_n) + \\
&\quad + \text{eval}\,(D_{t,x_i}s\,;\, b, b_1, \ldots, b_n) \\
&= b \cdot (D_{x_i}\phi(r))\,(b_1, \ldots, b_n) + \\
&\quad + b \cdot (D_{x_i}\phi(s))\,(b_1, \ldots, b_n) \\
&= b \cdot (D_{x_i}\phi(r) + D_{x_i}\phi(s))\,(b_1, \ldots, b_n) \\
&= b \cdot (D_{x_i}\phi(r+s))\,(b_1, \ldots, b_n).
\end{aligned}$$

- $p = r \cdot s$, where $r, s \in A\{x_1, \ldots, x_n\}$.

$$\begin{aligned}
&\text{eval}\,(D_{t,x_i}(r \cdot s)\,;\, b, b_1, \ldots, b_n) \\
= \;&\text{eval}\,(D_{t,x_i}(r) \cdot s + r \cdot D_{t,x_i}(s)\,;\, b, b_1, \ldots, b_n) \\
= \;&\text{eval}\,(D_{t,x_i}(r)\,;\, b, b_1, \ldots, b_n) \cdot \text{eval}\,(s\,;\, b_1, \ldots, b_n) + \\
&+ \text{eval}\,(r\,;\, b_1, \ldots, b_n) \cdot \text{eval}\,(D_{t,x_i}(s)\,;\, b, b_1, \ldots, b_n) \\
= \;&b \cdot (D_{x_i}(\phi(r)) \cdot \phi(s) + \phi(r) \cdot D_{x_i}(\phi(s)))\,(b_1, \ldots, b_n) \\
= \;&b \cdot (D_{x_i}\,(\phi(r \cdot s)))\,(b_1, \ldots, b_n).
\end{aligned}$$

□

Definition 1.4.1 *Let A and B be commutative algebras over the field $\mathbf{C}$ such that A is a subalgebra of B. An element $f \in A[x_1, \ldots, x_s]$ is* homogeneous *with respect to the linear spaces $V, W_1, \ldots, W_s$, where these linear spaces are finite dimensional subspaces of B, if for all $w_1 \in W_1, \ldots, w_s \in W_s$ we have that $f(w_1, \ldots, w_s) \in V$.*

Lemma 1.4.3 *Assume that the polynomial $f \in A[x_1, \ldots, x_s]$ is homogeneous with respect to the linear spaces $V, W_1, \ldots, W_s$, then the polynomial $t \cdot D_{x_i}(f)$ is homogeneous with* *respect to the linear spaces* $V, W_i, W_1, \ldots, W_s$.

Proof: We can find $p \in A\{x_1, \ldots, x_s\}$ such that $\phi(p) = f$. For $w_1 \in W_1, \ldots, w_s \in W_s$ we have that

$$\mathrm{eval}\,(p;\ w_1, \ldots, w_s) = f(w_1, \ldots, w_s) \in V.$$

Hence p is homogeneous with respect to the linear spaces $V, W_1, \ldots, W_s$. By Lemma 1.1.6 we know that $D_{t,x_i}p$ is homogeneous with respect to the linear spaces $V, W_i, W_1, \ldots, W_s$. Hence for $w \in W, w_1 \in W_1, \ldots, w_s \in W_s$ we have that

$$\begin{aligned}(t \cdot f)(w, w_1, \ldots, w_s) &= w \cdot f(w_1, \ldots, w_s) \\ &= \mathrm{eval}\,(D_{t,x_i}p;\ w, w_1, \ldots, w_s) \in V,\end{aligned}$$

and the conclusion of the lemma follows. □

Definition 1.4.2 *Let V and W be finite dimensional linear spaces of the commutative algebra B. Let c be an element of B such that $c \cdot W \subseteq V$. We say that f is apolar to $L \in V^*$ relative to W if for all $w \in W$*

$$\langle L \mid c \cdot w \rangle = 0.$$

1.4.2 The main theorem for commutaive algebras

Proposition 1.4.4 *Let $V, W_1, \ldots, W_s$ be finite dimensional linear subspaces of the commutative algebra A. Let $f \in A[x_1, \ldots, x_s]$ be homogeneous with respect to the linear spaces $V, W_1, \ldots, W_s$. A generic element $v \in V$ can be written in the form*

$$v = f(w_1, \ldots, w_s)$$

for some $w_1, \dots, w_s$ if and only if there exist $w'_1, \dots, w'_s$ so that there is no nonzero dual element in V^* which is apolar to the elements

$$\left(D_{x_j} f\right)(w'_1, \dots, w'_s)$$

relative to W_j, for all $1 \le j \le s$.

Observe that lemma 1.4.3 guarantees that apolarity condition is well defined.

Proof: We can find p in $A\{x_1, \dots, x_s\}$ such that $\phi(p) = f$. That is,

$$\mathrm{eval}\,(p;\, a_1, \dots, a_s) = f(a_1, \dots, a_s).$$

Hence if a generic element $v \in V$ can be written in the form $v = f(w_1, \dots, w_s)$, then it can be written in the form $v = \mathrm{eval}\,(p;\, w_1, \dots, w_s)$.

Observe that the statement that the element $\left(D_{x_j} f\right)(w'_1, \dots, w'_s)$ is apolar to $L \in V^*$ relative to W_j is equivalent to that the linear map

$$y_j \longmapsto y_j \cdot \left(D_{x_j} f\right)(w'_1, \dots, w'_s)$$

is apolar to $L \in V^*$ relative to W_j. But by Lemma 1.4.2 we know that

$$y_j \cdot \left(D_{x_j} f\right)(w'_1, \dots, w'_s) = \mathrm{eval}\left(D_{t,x_j} p;\, y_j, w'_1, \dots, w'_s\right).$$

Thus we can use Theorem 1 and the proof of the proposition follows. □

The classical case of this proposition is when the algebra A is graded,

$$A = \bigoplus_{n \ge 0} A_n,$$

and we consider the linear subspaces $V, W_1, \dots, W_s$ to be homogeneous subspaces A_n. This is the way how the proposition is presented in [**E-R1**].

1.5 Two by two by two matrices

In the following we will make a thoroughly study of the canonical forms of $2 \times 2 \times 2$ matrices, and of invariants and covariants of these matrices.

1.5.1 Invariants and covariants of polynomials

We start by recalling some facts about two-dimensional vector spaces. For $\mathbf{a}, \mathbf{b} \in \mathbf{C}^2$ define

$$
\begin{aligned}
(\mathbf{a}|\mathbf{b}) &= a_1 b_1 + a_2 b_2, \\
[\mathbf{a}, \mathbf{b}] &= a_1 b_2 - a_2 b_1,
\end{aligned}
$$

where $\mathbf{a} = (a_1, a_2)$ and $\mathbf{b} = (b_1, b_2)$. Let ϕ be a linear map from $\mathbf{C}^2$ to itself. Define the determinant of ϕ by

$$
\det(\phi) = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix},
$$

where $\phi(\mathbf{e}_1) = (a_{1,1}, a_{1,2})$, and $\phi(\mathbf{e}_2) = (a_{2,1}, a_{2,2})$. Observe that the determinant of the linear map ϕ is independent of the basis of the linear space. Define ϕ^* to be the linear map defined by $\phi^*(\mathbf{e}_1) = (a_{1,1}, a_{2,1})$, and $\phi^*(\mathbf{e}_2) = (a_{1,2}, a_{2,2})$.

Lemma 1.5.1 *Let $\mathbf{a}, \mathbf{b} \in \mathbf{C}^2$, and let $\phi, \psi : \mathbf{C}^2 \longrightarrow \mathbf{C}^2$ be two linear maps. Then the following is true*

$$
\begin{aligned}
(\phi(\mathbf{a})|\mathbf{b}) &= (\mathbf{a}|\phi^*(\mathbf{b})), \\
[\phi(\mathbf{a}), \phi(\mathbf{b})] &= \det(\phi) \cdot (\mathbf{a}|\mathbf{b}), \\
\det(\phi^*) &= \det(\phi), \\
\det(\phi \circ \psi) &= \det(\phi) \cdot \det(\psi), \\
(\phi \circ \psi)^* &= \psi^* \circ \phi^*.
\end{aligned}
$$

Let y_1, y_2 be variables and let $\mathrm{span}(y_1, y_2)$ be the two dimensional linear space spanned by y_1 and y_2. Let A be an algebra over the field of complex numbers $\mathbf{C}$, and let $A[y_1, y_2]$ be the polynomial algebra with variables y_1, y_2, and the coefficients in A. A linear map $\phi : \mathrm{span}(y_1, y_2) \longrightarrow \mathrm{span}(y_1, y_2)$ extends to a map $\hat{\phi} : A[y_1, y_2] \longrightarrow A[y_1, y_2]$ by substitution. That is, for $p(\mathbf{y}) \in A[y_1, y_2]$,

$$\hat{\phi}p(y_1, y_2) = p(\phi(y_1), \phi(y_2)).$$

Thus the $\hat{\phi}$ correspond to change of variables. Moreover notice that $\widehat{\phi \circ \psi} = \hat{\phi} \circ \hat{\psi}$.

Let W_p be the linear subspace of $\mathbf{C}[\mathbf{y}] = \mathbf{C}[y_1, y_2]$ consisting of all homogeneous elements of degree p.

Definition 1.5.1 *A polynomial map f from a finite dimensional linear space U to a finite dimensional linear space V is a function from U to V that can be written in the form*

$$f\left(\sum_{i=1}^{n} a_i u_i\right) = \sum_{j=1}^{m} p_j(a_1, \ldots, a_n) v_j,$$

where $p_1, \ldots, p_m$ are polynomials, $u_1, \ldots, u_n$ are a basis for U, and $v_1, \ldots, v_n$ are a basis for V.

Definition 1.5.2 *A covariant C of W_p is a polynomial map from W_p to $\mathbf{C}[\mathbf{y}]$, for all linear maps $\phi : \mathrm{span}(y_1, y_2) \longrightarrow \mathrm{span}(y_1, y_2)$, we have that*

$$C\left(\hat{\phi} w\right) = \det(\phi)^g \cdot \hat{\phi} C(w),$$

where the non-negative integer g is called the index of C.

Definition 1.5.3 *An invariant I of W_p is a covariant of W_p, which maps W_p into $\mathbf{C}$.*

Since $\hat{\phi} c = c$ for all $c \in \mathbf{C}$ we have that the condition for I being an invariant is

$$I\left(\hat{\phi} w\right) = \det(\phi)^g \cdot I(w),$$

for the index g.

As an example of these concepts we will give four covariants of the linear space W_2 of binary quadratics.

1.

$$\Delta\left(a_0y_1^2+a_1y_1y_2+a_2y_2^2\right)=a_1^2-4a_0a_2.$$

This invariant is called the discriminant and has index 2. It can also be described by

$$\Delta(w)=-\begin{vmatrix}\frac{\partial^2 w}{\partial y_1^2} & \frac{\partial^2 w}{\partial y_1\partial y_2}\\ \frac{\partial^2 w}{\partial y_2\partial y_1} & \frac{\partial^2 w}{\partial y_2^2}\end{vmatrix}.$$

2. This is the identity covariant. That is, $\mathrm{Id}(w)=w$.

3. 1. This is the constant covariant.

4. 0. This is the zero covariant.

It is well known that a binary quadratic can be written in one of the following three forms pq, p^2, 0, where p and q are independent vectors in $\mathrm{span}(y_1,y_2)$. Moreover, define the bracket on $\mathrm{span}(y_1,y_2)$ by $[a_1y_1+a_2y_2,b_1y_1+b_2y_2]=a_1b_2-a_2b_1$. We can show the correspondence between covariants and canonical forms of the binary cubics, by introducing the following table.

	pq	p^2	0
0	0	0	0
Δ	$[p,q]^2$	0	0
Id	pq	p^2	0
1	1	1	1

Observe that we can also describe the discriminant by

$$\Delta\left(p^2+q^2\right)=-4[p,q]^2.$$

1.5.2 Invariants and covariants of $2 \times 2 \times 2$ matrices

Let $\mathbf{x}$ denote the sextet $\left(x_j^{(i)}\right)_{i=1,2,3,\, j=1,2}$, where each of these six elements are considered to be variables. For $i = 1,2,3$ let $\mathbf{x}^{(i)} = (x_1^{(i)}, x_2^{(i)})$. Let $\mathbf{C}[\mathbf{x}]$ denote the algebra of polynomials in these six variables. Let W be the linear subspace of $\mathbf{C}[\mathbf{x}]$, which consists of polynomials that are homogeneous and have degree 1 in $\mathbf{x}^{(i)}$ for all $i = 1,2,3$. An element w of W can be written in the form

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} a_{i,j,k} x_i^{(1)} x_j^{(2)} x_k^{(3)}.$$

Observe W is naturally isomorphic to the linear space of $2 \times 2 \times 2$ matrices by sending w to the matrix

$$\left(\begin{pmatrix} a_{1,1,1} & a_{1,1,2} \\ a_{1,2,1} & a_{1,2,2} \end{pmatrix}, \begin{pmatrix} a_{2,1,1} & a_{2,1,2} \\ a_{2,2,1} & a_{2,2,2} \end{pmatrix}\right).$$

Let V_i be the linear space spanned by $x_1^{(i)}$ and $x_2^{(i)}$ for $i = 1,2,3$. Let $\phi_i : V_i \longrightarrow V_i$ be a linear map. Thus $\widehat{\phi}_i$ is a change of variables in the $\mathbf{x}^{(i)}$ variables.

Definition 1.5.4 *A covariant C of W is a polynomial map from W to $\mathbf{C}[\mathbf{x}]$, for all linear maps ϕ_1, ϕ_2, ϕ_3, where $\phi_i : V_i \longrightarrow V_i$ for $i = 1,2,3$, we have that*

$$C\left(\widehat{\phi}_1\widehat{\phi}_2\widehat{\phi}_3 w\right) = \det(\phi_1)^{g_1} \cdot \det(\phi_2)^{g_2} \cdot \det(\phi_3)^{g_3} \cdot \widehat{\phi}_1\widehat{\phi}_2\widehat{\phi}_3 C(w),$$

for some non-negative integers g_1, g_2, and g_3. Call the triple (g_1, g_2, g_3) the index of C.

There are two trivial covariants. The first one is $C \equiv c$ where c is a constant. The second covariant is the identity map, that is, $C(w) = w$ for all $w \in W$. Both these covariants has the index $(g_1, g_2, g_3) = (0,0,0)$.

Observe that if C' and C'' are two covariants with corresponding indices (g_1', g_2', g_3') and (g_1'', g_2'', g_3'') then $C = C' \cdot C''$ is also a covariant with the index $(g_1' + g_1'', g_2' + g_2'', g_3' + g_3'')$.

The statement is a straightforward calculation.

$$\begin{aligned} C\left(\hat\phi_1\hat\phi_2\hat\phi_3 w\right) &= C'\left(\hat\phi_1\hat\phi_2\hat\phi_3 w\right)\cdot C''\left(\hat\phi_1\hat\phi_2\hat\phi_3 w\right) \\ &= \prod_{i=1}^{3}\det(\phi_i)^{g_i'}\cdot\hat\phi_1\hat\phi_2\hat\phi_3 C'(w)\cdot\prod_{i=1}^{3}\det(\phi_i)^{g_i''}\cdot\hat\phi_1\hat\phi_2\hat\phi_3 C''(w) \\ &= \prod_{i=1}^{3}\det(\phi_i)^{g_i'+g_i''}\cdot\hat\phi_1\hat\phi_2\hat\phi_3 C(w). \end{aligned}$$

Moreover, if C' and C'' are two covariants with the same index (g_1, g_2, g_3) then $C'+C''$ is also a covariant with index (g_1, g_2, g_3).

Let C be a covariant with index (g_1, g_2, g_3), and let $\psi_i : V_i \longrightarrow V_i$ be invertible linear maps for $i = 1, 2, 3$. Then

$$D(w) = \hat\psi_1^{-1}\hat\psi_2^{-1}\hat\psi_3^{-1} C\left(\hat\psi_1\hat\psi_2\hat\psi_3 w\right)$$

is also a covariant with index (g_1, g_2, g_3). This follows from the following calculation

$$\begin{aligned} D\left(\hat\phi_1\hat\phi_2\hat\phi_3 w\right) &= \hat\psi_1^{-1}\hat\psi_2^{-1}\hat\psi_3^{-1} C\left(\hat\psi_1\hat\psi_2\hat\psi_3\hat\phi_1\hat\phi_2\hat\phi_3 w\right) \\ &= \prod_{i=1}^{3}\det(\psi\phi)^{g_i}\cdot\hat\psi_1^{-1}\hat\psi_2^{-1}\hat\psi_3^{-1}\hat\psi_1\hat\psi_2\hat\psi_3\hat\phi_1\hat\phi_2\hat\phi_3 C(w) \\ &= \prod_{i=1}^{3}\det(\phi)^{g_i}\cdot\prod_{i=1}^{3}\det(\psi)^{g_i}\cdot\hat\phi_1\hat\phi_2\hat\phi_3 C(w) \\ &= \prod_{i=1}^{3}\det(\phi)^{g_i}\cdot\prod_{i=1}^{3}\det(\psi)^{g_i}\cdot\hat\phi_1\hat\phi_2\hat\phi_3\hat\psi_1^{-1}\hat\psi_2^{-1}\hat\psi_3^{-1}\hat\psi_1\hat\psi_2\hat\psi_3 C(w) \\ &= \prod_{i=1}^{3}\det(\phi)^{g_i}\cdot\hat\phi_1\hat\phi_2\hat\phi_3\hat\psi_1^{-1}\hat\psi_2^{-1}\hat\psi_3^{-1} C\left(\hat\psi_1\hat\psi_2\hat\psi_3 w\right) \\ &= \prod_{i=1}^{3}\det(\phi)^{g_i}\cdot\hat\phi_1\hat\phi_2\hat\phi_3 D(w). \end{aligned}$$

Definition 1.5.5 *An invariant I of W is a covariant of W, which maps W into $\mathbf{C}$.*

Since $\widehat{\phi}_1\widehat{\phi}_2\widehat{\phi}_3 c = c$ for all $c \in \mathbf{C}$ we have that the condition for I being an invariant is

$$I\left(\widehat{\phi}_1\widehat{\phi}_2\widehat{\phi}_3 w\right) = \prod_{i=1}^{3} \det(\phi_i)^{g_i} \cdot I(w),$$

for the index (g_1, g_2, g_3).

Proposition 1.5.2 *All invariants of $2 \times 2 \times 2$ matrices are in the form $s \cdot \Gamma^k$, where s is a complex constant, k a non-negative integer and Γ is defined by*

$$\begin{aligned}\Gamma(w) &= (a_{1,1,1}a_{2,2,2} + a_{1,2,1}a_{2,1,2} - a_{1,1,2}a_{2,2,1} - a_{1,2,2}a_{2,1,1})^2 - \\ &\quad -4(a_{1,1,1}a_{2,1,2} - a_{1,1,2}a_{2,1,1})(a_{1,2,1}a_{2,2,2} - a_{1,2,2}a_{2,2,1})\end{aligned}$$

where

$$w = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} a_{i,j,k} x_i^{(1)} x_j^{(2)} x_k^{(3)}.$$

The invariant $s \cdot \Gamma^k$ has the index $(2k, 2k, 2k)$.

Proof: It is a straightforward calculation that Γ is an invariant. By the following rewriting of the above definition of Γ, it is easy to check that $\Gamma\left(\widehat{\phi}_3 w\right) = \det(\widehat{\phi}_3)^2 \cdot \Gamma(w)$.

$$\Gamma(w) = \left(\begin{vmatrix} a_{1,1,1} & a_{2,2,1} \\ a_{1,1,2} & a_{2,2,2} \end{vmatrix} + \begin{vmatrix} a_{1,2,1} & a_{2,1,1} \\ a_{1,2,2} & a_{2,1,2} \end{vmatrix}\right)^2 - 4 \cdot \begin{vmatrix} a_{1,1,1} & a_{2,1,1} \\ a_{1,1,2} & a_{2,1,2} \end{vmatrix} \cdot \begin{vmatrix} a_{1,2,1} & a_{2,2,1} \\ a_{1,2,2} & a_{2,2,2} \end{vmatrix}.$$

The two other conditions $\Gamma\left(\widehat{\phi}_i w\right) = \det(\widehat{\phi}_i)^2 \cdot \Gamma(w)$, $i = 1, 2$, follows by similar reformulations of $\Gamma(w)$. Then it directly follows that $s \cdot \Gamma^k$ is an invariant. This argument proves one implication of the proposition.

We will start with a generic element of W, and try to compute $I(w)$ by using the fact that I is an invariant. This means that we will apply different changes of variables, so that we will reduce as many coefficients as possible to 0. To make this process easy to follow, we will write the element w as a $2 \times 2 \times 2$ matrix. Moreover, we will only apply maps $\widehat{\phi}_i$ that have determinant equal to 1. This corresponds to multiplying one plane in three dimensional matrix by a scalar and adding it to the parallel plane. We begin with the matrix

$$\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} e & f \\ g & h \end{pmatrix}\right).$$

$$\left(\begin{pmatrix} a & b \\ c+pa & d+pb \end{pmatrix}, \begin{pmatrix} e & f \\ g+pe & h+pf \end{pmatrix}\right).$$

$$\left(\begin{pmatrix} a+qb & b \\ c+pa+qd+qpb & d+pb \end{pmatrix}, \begin{pmatrix} e+qf & f \\ g+pe+qh+qpf & h+pf \end{pmatrix}\right).$$

Now find complex numbers p and q such that

$$\begin{aligned} c+pa+qd+qpb &= 0, \\ g+pe+qh+qpf &= 0. \end{aligned}$$

This is an equation system of second order, and the solution is

$$p = \frac{-(ah+cf-bg-de) \pm \sqrt{D}}{2(af-be)},$$

$$q = \frac{+(ah-cf+bg-de) \mp \sqrt{D}}{2(df-bh)},$$

where

$$D = (ah+cf-bg-de)^2 - 4(af-be)(ch-dg).$$

That is, D is the discriminant of the equation system in p and q. Since we started with a generic element, we may assume that $af-be \neq 0$ and $df-bh \neq 0$ so p and q are well defined. Thus the above $2 \times 2 \times 2$ matrix looks like

$$\left(\begin{pmatrix} a+qb & b \\ 0 & d+pb \end{pmatrix}, \begin{pmatrix} e+qf & f \\ 0 & h+pf \end{pmatrix}\right).$$

If the two vectors $(a+qb, e+qf)$ and $(d+pb, h+pf)$ spans the linear space $\mathbf{C}^2$, then the vector (b,f) is in the linear span of the two first vectors. Since we started with a generic element w, we may assume that these two vectors are linearly independent, and hence they span $\mathbf{C}^2$. Using this information in two operations we get

$$\left(\begin{pmatrix} a+qb & 0 \\ 0 & d+pb \end{pmatrix}, \begin{pmatrix} e+qf & 0 \\ 0 & h+pf \end{pmatrix}\right).$$

Continue with the following sequences of operations.

$$\left(\begin{pmatrix} a+qb & 0 \\ 0 & d+pb \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & h+pf-(d+pb)\frac{e+qf}{a+qb} \end{pmatrix}\right).$$

$$\left(\begin{pmatrix} a+qb & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & h+pf-(d+pb)\frac{e+qf}{a+qb} \end{pmatrix}\right).$$

By dividing the top plane of the matrix with $(a+qb)$, and multiplying the bottom plane by the same value we get

$$\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & (a+qb)(h+pf)-(d+pb)(e+qf) \end{pmatrix}\right).$$

Observe that this last operation correspond to a linear map with determinant equal to 1. But an easy calculation shows that

$$(a+qb)(h+pf)-(d+pb)(e+qf) = \pm\sqrt{D}.$$

Hence we know that

$$I\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} e & f \\ g & h \end{pmatrix}\right) = I\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & \pm\sqrt{D} \end{pmatrix}\right)$$

$$= \left(\pm\sqrt{D}\right)^{g_1} \cdot I\left(\left(\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right), \left(\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right)\right)$$
$$= \left(\pm\sqrt{D}\right)^{g_1} \cdot s,$$

where s is a constant and g_1 is the first coordinate of the index. If s is equal to zero then the invariant I vanishes. Assume then that s is non-zero. Since I is a polynomial map, g_1 is an even nonnegative integer. Hence $D = \Gamma(w)$. By symmetry it follows that $g_1 = g_2 = g_3$, which concludes the proof. □

1.5.3 Umbral notation

We will now introduce umbral notation. The umbral notation is very helpful in describing covariants. We will use greek letters, α, β, …, δ to denote umbrae. For each of these umbrae there will be six variables. Namely for the umbra α we have the variables $\alpha_1^{(1)}$, $\alpha_2^{(1)}$, $\alpha_1^{(2)}$, $\alpha_2^{(2)}$, $\alpha_1^{(3)}$, and $\alpha_2^{(3)}$. As before we will denote the pair $(\alpha_1^{(i)}, \alpha_2^{(i)})$ by $\alpha^{(i)}$.

Consider the algebra $\mathbf{C}[\mathbf{x}, \alpha, \ldots, \delta]$ which consists of polynomials in the variables $x_1^{(1)}$, …, $x_2^{(3)}$, $\alpha_1^{(1)}$, …, $\delta_2^{(3)}$. Define the umbral map U from this algebra to the algebra $\mathbf{C}[\mathbf{x}]$ as follows

1. The map U is linear,
2. $U(p(\mathbf{x}) \cdot q(\alpha) \cdots r(\delta)) = U(p(\mathbf{x})) \cdot U(q(\alpha)) \cdots U(r(\delta))$, where $p(\mathbf{x}) \in \mathbf{C}[\mathbf{x}]$, $q(\alpha) \in \mathbf{C}[\alpha]$, …, $r(\delta) \in \mathbf{C}[\delta]$,
3. $U(p(\mathbf{x})) = p(\mathbf{x})$, for all $p(\mathbf{x}) \in \mathbf{C}[\mathbf{x}]$.
4. for each umbra α, we have that $U\left(\alpha_i^{(1)}\alpha_j^{(2)}\alpha_k^{(3)}\right) = a_{i,j,k}$, and the umbral map U vanishes on any other monomial in the variables belonging to the umbra α.

By using the umbral map, we get a short notation to write with. We begin by noticing

Lemma 1.5.3

$$U\left(\left(\alpha^{(1)}\middle|\mathbf{x}^{(1)}\right)\left(\alpha^{(2)}\middle|\mathbf{x}^{(2)}\right)\left(\alpha^{(3)}\middle|\mathbf{x}^{(3)}\right)\right)=\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2}a_{i,j,k}x_i^{(1)}x_j^{(2)}x_k^{(3)}.$$

Proof: The proof is a straight forward verification.

$$\begin{aligned}U\left(\left(\alpha^{(1)}\middle|\mathbf{x}^{(1)}\right)\left(\alpha^{(2)}\middle|\mathbf{x}^{(2)}\right)\left(\alpha^{(3)}\middle|\mathbf{x}^{(3)}\right)\right) &= U\left(\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2}\alpha_i^{(1)}\alpha_j^{(2)}\alpha_k^{(3)}\cdot x_i^{(1)}x_j^{(2)}x_k^{(3)}\right)\\ &= \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2}a_{i,j,k}\cdot x_i^{(1)}x_j^{(2)}x_k^{(3)}.\end{aligned}$$

□

Definition 1.5.6 *A bracket monomial is a product of terms in the forms* $\left[\alpha^{(i)},\beta^{(i)}\right]$ *and* $\left(\alpha^{(i)}\middle|\mathbf{x}^{(i)}\right)$.

Observe that a bracket monomial M would vanish when applying the umbral operator U, if for some umbra α in the monomial M, the umbral variable $\alpha^{(i)}$ does not occur exactly once.

Proposition 1.5.4 *The umbral operator of a bracket monomial is a covariant.*

Proof: Let M be a bracket monomial. Say that

$$M=\prod_{j=1}^{n}\left[\alpha(j)^{(i)},\beta(j)^{(i)}\right]\cdot\prod_{j=1}^{m}\left(\gamma(j)^{(i)}\middle|\mathbf{x}^{(i)}\right),$$

where $\alpha(j)$, $\beta(j)$, and $\gamma(j)$ are umbrae. The proposition claims that

$$C(w)=U(M)$$

is a covariant, where $w=\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2}a_{i,j,k}x_i^{(1)}x_j^{(2)}x_k^{(3)}$.

We know that an element $w \in W$ can be represented umbrally by

$$\left(\alpha^{(1)}\middle|\mathbf{x}^{(1)}\right)\left(\alpha^{(2)}\middle|\mathbf{x}^{(2)}\right)\left(\alpha^{(3)}\middle|\mathbf{x}^{(3)}\right).$$

Thus $\hat{\phi}_1\hat{\phi}_2\hat{\phi}_3 w$ can be represented umbrally by

$$\begin{aligned} &\left(\alpha^{(1)}\middle|\phi_1(\mathbf{x}^{(1)})\right)\left(\alpha^{(2)}\middle|\phi_2(\mathbf{x}^{(2)})\right)\left(\alpha^{(3)}\middle|\phi_3(\mathbf{x}^{(3)})\right) \\ = &\left(\phi_1^*(\alpha^{(1)})\middle|\mathbf{x}^{(1)}\right)\left(\phi_2^*(\alpha^{(2)})\middle|\mathbf{x}^{(2)}\right)\left(\phi_3^*(\alpha^{(3)})\middle|\mathbf{x}^{(3)}\right). \end{aligned}$$

To compute the umbral operator of the bracket polynomial corresponding to $\hat{\phi}_1\hat{\phi}_2\hat{\phi}_3 w$, we replace $\alpha^{(i)}$ with $\phi_i^*(\alpha^{(i)})$ for each umbrae α and $i = 1, 2, 3$. Hence

$$C(\hat{\phi}_1\hat{\phi}_2\hat{\phi}_3 w) = U\left(\prod_{j=1}^{n}\left[\phi^*(\alpha(j)^{(i)}), \phi^*(\beta(j)^{(i)})\right] \cdot \prod_{j=1}^{m}\left(\phi^*(\gamma(j)^{(i)})\middle|\mathbf{x}^{(i)}\right)\right).$$

But observe that

$$\left[\phi_i^*(\alpha(j)^{(i)}), \phi_i^*(\beta(j)^{(i)})\right] = \det(\phi_i) \cdot \left[\alpha(j)^{(i)}, \beta(j)^{(i)}\right],$$

and

$$\left(\phi_i^*(\gamma(j)^{(i)})\middle|\mathbf{x}^{(i)}\right) = \left(\gamma(j)^{(i)}\middle|\phi_i(\mathbf{x}^{(i)})\right).$$

Let g_i be the number of terms in M on the form $\left[\alpha^{(i)}, \beta^{(i)}\right]$. Then the above expression is equal to

$$\prod_{i=1}^{3}\det(\phi_i)^{g_i} \cdot U\left(\prod_{j=1}^{n}\left[\alpha(j)^{(i)}, \beta(j)^{(i)}\right] \cdot \prod_{j=1}^{m}\left(\gamma(j)^{(i)}\middle|\phi(\mathbf{x}^{(i)})\right)\right).$$

By evaluating this umbral expression we get

$$\prod_{i=1}^{3}\det(\phi_i)^{g_i} \cdot \hat{\phi}_1\hat{\phi}_2\hat{\phi}_3 C(w).$$

Hence we conclude that $C(w) = U(M)$ is a covariant. $\square$

We now have a tool for expressing covariants of $2 \times 2 \times 2$ matrices. We will give a list of eight covariants, which are of interest to us.

1.
$$-\frac{1}{2} \cdot \left[\alpha^{(1)}, \beta^{(1)}\right] \left[\alpha^{(2)}, \beta^{(2)}\right] \left[\alpha^{(3)}, \gamma^{(3)}\right] \left[\beta^{(3)}, \delta^{(3)}\right] \left[\gamma^{(1)}, \delta^{(1)}\right] \left[\gamma^{(2)}, \delta^{(2)}\right].$$

The umbral operator of the above bracket monomial evaluates to Γ, where Γ is the invariant described in Proposition 1.5.2. We will prove this in the end of this section.

2.
$$-\left[\alpha^{(1)}, \beta^{(1)}\right] \left[\alpha^{(2)}, \beta^{(2)}\right] \left[\alpha^{(3)}, \gamma^{(3)}\right] \left(\beta^{(3)} \middle| \mathbf{x}^{(3)}\right) \left(\gamma^{(1)} \middle| \mathbf{x}^{(1)}\right) \left(\gamma^{(2)} \middle| \mathbf{x}^{(2)}\right).$$

This covariant is denoted by S. Observe that S maps W into itself.

3.
$$\frac{1}{2} \left[\alpha^{(2)}, \beta^{(2)}\right] \left[\alpha^{(3)}, \beta^{(3)}\right] \left(\alpha^{(1)} \middle| \mathbf{x}^{(1)}\right) \left(\beta^{(1)} \middle| \mathbf{x}^{(1)}\right).$$

Denote the covariant that the above bracket monomial evaluates to by $H^{(1)}$. This covariant is called the Hessian, and can also be described by

$$H^{(1)}(w) = \begin{vmatrix} \frac{\partial^2 w}{\partial x_1^{(2)} \partial x_1^{(3)}} & \frac{\partial^2 w}{\partial x_1^{(2)} \partial x_2^{(3)}} \\ \frac{\partial^2 w}{\partial x_2^{(2)} \partial x_1^{(3)}} & \frac{\partial^2 w}{\partial x_2^{(2)} \partial x_2^{(3)}} \end{vmatrix}.$$

4.
$$\frac{1}{2} \left[\alpha^{(1)}, \beta^{(1)}\right] \left[\alpha^{(3)}, \beta^{(3)}\right] \left(\alpha^{(2)} \middle| \mathbf{x}^{(2)}\right) \left(\beta^{(2)} \middle| \mathbf{x}^{(2)}\right).$$

Similarly, this is also a Hessian, but in the second variable. It is denoted by $H^{(2)}$.

5.
$$\frac{1}{2} \left[\alpha^{(1)}, \beta^{(1)}\right] \left[\alpha^{(2)}, \beta^{(2)}\right] \left(\alpha^{(3)} \middle| \mathbf{x}^{(3)}\right) \left(\beta^{(3)} \middle| \mathbf{x}^{(3)}\right).$$

This Hessian is denoted by $H^{(3)}$.

6.

$$\left(\alpha^{(1)}\middle|\mathbf{x}^{(1)}\right)\left(\alpha^{(2)}\middle|\mathbf{x}^{(2)}\right)\left(\alpha^{(3)}\middle|\mathbf{x}^{(3)}\right).$$

This is the identity covariant. That is, $\mathrm{Id}(w) = w$.

7. 1. This is the constant covariant.

8. 0. This is the zero covariant.

Let us now show that the umbral expression given for Γ actually evaluates to the expression given in Proposition 1.5.2. We know that invariant given by the umbral expression has index $(2,2,2)$. Thus by the result of Proposition 1.5.2 we only need to check that the constant is equal to 1. To do this, evaluate the umbral expression for the element $x_1^{(1)}x_1^{(2)}x_1^{(3)} + x_2^{(1)}x_2^{(2)}x_2^{(3)}$.

$$\begin{aligned} & U\left(-\frac{1}{2}\cdot\left[\alpha^{(1)},\beta^{(1)}\right]\left[\alpha^{(2)},\beta^{(2)}\right]\left[\alpha^{(3)},\gamma^{(3)}\right]\left[\beta^{(3)},\delta^{(3)}\right]\left[\gamma^{(1)},\delta^{(1)}\right]\left[\gamma^{(2)},\delta^{(2)}\right]\right) \\ = & -\frac{1}{2}\cdot U\left(\left(\alpha_1^{(1)}\beta_2^{(1)}-\alpha_2^{(1)}\beta_1^{(1)}\right)\left(\alpha_1^{(2)}\beta_2^{(2)}-\alpha_2^{(2)}\beta_1^{(2)}\right)\left(\alpha_1^{(3)}\gamma_2^{(3)}-\alpha_2^{(3)}\gamma_1^{(3)}\right)\cdot\right. \\ & \left.\cdot\left[\beta^{(3)},\delta^{(3)}\right]\left[\gamma^{(1)},\delta^{(1)}\right]\left[\gamma^{(2)},\delta^{(2)}\right]\right). \end{aligned}$$

But the only monomials of the umbra α that will not vanish are the monomial $\alpha_1^{(1)}\alpha_1^{(2)}\alpha_1^{(3)}$ and the monomial $\alpha_2^{(1)}\alpha_2^{(2)}\alpha_2^{(3)}$. Instead these two monomials will evaluate to 1. Thus the above expression is equal to

$$-\frac{1}{2}\cdot U\left(\left(\beta_2^{(1)}\beta_2^{(2)}\gamma_2^{(3)}-\beta_1^{(1)}\beta_1^{(2)}\gamma_1^{(3)}\right)\left[\beta^{(3)},\delta^{(3)}\right]\left[\gamma^{(1)},\delta^{(1)}\right]\left[\gamma^{(2)},\delta^{(2)}\right]\right).$$

By a similar argument on the umbra δ we have that it is equal to

$$-\frac{1}{2}\cdot U\left(\left(\beta_2^{(1)}\beta_2^{(2)}\gamma_2^{(3)}-\beta_1^{(1)}\beta_1^{(2)}\gamma_1^{(3)}\right)\left(\beta_1^{(3)}\gamma_1^{(1)}\gamma_1^{(2)}-\beta_2^{(3)}\gamma_2^{(1)}\gamma_2^{(2)}\right)\right),$$

which easily computes to 1 and the equality between the two expressions of Γ follows.

For the other covariants above we have chosen the constants such that there will be no rational factor appearing when we compute the covariant. That is, the expression $C\left(x_1^{(1)}x_1^{(2)}x_1^{(3)} + x_2^{(1)}x_2^{(2)}x_2^{(3)}\right)$ is a monomial without a constant.

1.5.4 Canonical forms

Proposition 1.5.5 *An element $w \in W$ can be written in exactly one of the following seven forms*

$$pqr + stu,$$
$$sqr + ptr + pqu,$$
$$pqr + ptu,$$
$$pqr + squ,$$
$$pqr + str,$$
$$pqr,$$
$$0$$

where $p, s \in V_1$, $q, t \in V_2$, $r, u \in V_3$, p and s are linearly independent, q and t are linearly independent, and r and u are linearly independent.

Recall that if two vectors are linearly independent, then they are both nonzero.

Proof: We begin by showing that any $w \in W$ can be written in one of these two following forms

$$pqr + stu, \qquad sqr + ptr + pqu,$$

where $p, s \in V_1$, $q, t \in V_2$, and $r, u \in V_3$. Observe that both forms above are invariant under changes of variables. To prove that such changes of variables is possible is quite similar to the proof of Proposition 1.5.2. But we need to be more careful, since we don't start with a generic element of W, but any element of W.

We claim that by a change of variables, we can transform any $2 \times 2 \times 2$ matrix to a $2 \times 2 \times 2$ matrix where two adjacent entries are equal to zero. Following the proof of Proposition 1.5.2 we can do this if $af - be \neq 0$ and $df - bh \neq 0$. Without loss of generality we can assume that $af - be = 0$, since the case $df - bh = 0$ is symmetric. Since $af - be = 0$, the two vectors (a, e) and (b, f) are linearly dependent. Now by change of variables, we can make one of the vectors vanish, and thus there is two adjacent entries in the matrix that are equal to zero. This finish the claim and we can assume with that the matrix looks like

$$\left(\begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix}, \begin{pmatrix} e' & f' \\ 0 & h' \end{pmatrix}\right).$$

If the vectors (a', e') and (d', h') are linearly independent, then by two more changes of variables we can eliminate b' and f'. Thus we have the matrix

$$\left(\begin{pmatrix} a' & 0 \\ 0 & d' \end{pmatrix}, \begin{pmatrix} e' & 0 \\ 0 & h' \end{pmatrix}\right).$$

This matrix corresponds to the element

$$a'x_1^{(1)}x_1^{(2)}x_1^{(3)} + d'x_1^{(1)}x_2^{(2)}x_2^{(3)} + e'x_2^{(1)}x_1^{(2)}x_1^{(3)} + h'x_2^{(1)}x_2^{(2)}x_2^{(3)},$$

which can be written as

$$\left(a'x_1^{(1)} + e'x_2^{(1)}\right) x_1^{(2)}x_1^{(3)} + \left(d'x_1^{(1)} + h'x_2^{(1)}\right) x_2^{(2)}x_2^{(3)}.$$

This is on the first desired form.

It remains to handle the case when the vectors (a', e') and (d', h') are linearly dependent. If they are both equal to zero, then the matrix is trivially of one of the above forms. Hence assume that $(a', e') \neq \mathbf{0}$, so we can write $(d', h') = j \cdot (a', e')$. Now the matrix corresponds to the element

$$a'x_1^{(1)}x_1^{(2)}x_1^{(3)} + b'x_1^{(1)}x_1^{(2)}x_2^{(3)} + j \cdot a'x_1^{(1)}x_2^{(2)}x_2^{(3)} + e'x_2^{(1)}x_1^{(2)}x_1^{(3)} + f'x_2^{(1)}x_1^{(2)}x_2^{(3)} + j \cdot e'x_2^{(1)}x_2^{(2)}x_2^{(3)},$$

which we can write as

$$\left(b'x_1^{(1)} + f'x_2^{(1)}\right) x_1^{(2)} x_2^{(3)} + \left(a'x_1^{(1)} + e'x_2^{(1)}\right) \left(jx_2^{(2)}\right) x_2^{(3)} + \left(a'x_1^{(1)} + e'x_2^{(1)}\right) x_1^{(2)} x_1^{(3)},$$

which is in the second desired form. Thus we have proved that any $w \in W$ can either be written in the form $pqr + stu$, or in the form $sqr + ptr + pqu$. If p and s are dependent, q and t are dependent, or r and u are dependent then it easy to see that we can reduce the expression further to one of the following five forms $pqr + ptu$, $pqr + squ$, $pqr + str$, pqr, 0.

So far we have proven that $w \in W$ can be written in one of the seven forms. What remains to prove is that w can be written in exactly one of the seven forms.

To prove this claim, evaluate the seven covariants on the seven canonical forms. To do this in a quick way, evaluate the covariant C on the element $x_1^{(1)} x_1^{(2)} x_1^{(3)} + x_2^{(1)} x_2^{(2)} x_2^{(3)}$. Then by a suitable change of variable, one computes $C(pqr+stu)$. Now since the covariant is continuous, use the limit

$$\lim_{R \to \infty} (-Rp)(-Rq)(-Rr) + (Rp + R^{-2}s)(Rq + R^{-2}t)(Rr + R^{-2}u) = sqr + ptr + pqu,$$

to find $C(sqr + ptr + pqu)$. Similarly use the four limits

$$\begin{aligned} \lim_{\epsilon \to 0} pqr + (p + \epsilon s)tu &= pqr + ptu, \\ \lim_{\epsilon \to 0} pqr + s(q + \epsilon t)u &= pqr + squ, \\ \lim_{\epsilon \to 0} pqr + st(r + \epsilon u) &= pqr + str \\ \lim_{\epsilon \to 0} pqr + (\epsilon s)tu &= pqr, \end{aligned}$$

to find $C(pqr + ptu)$, $C(pqr + squ)$, $C(pqr + str)$, and $C(pqr)$. Moreover, define the bracket on V_i by

$$\left[a_1 x_1^{(i)} + a_2 x_2^{(i)}, b_1 x_1^{(i)} + b_2 x_2^{(i)}\right] = a_1 b_2 - a_2 b_1.$$

The computations are summarized in the following table.

	$pqr+stu$	$sqr+ptr+pqu$	$pqr+ptu$	$pqr+squ$	$pqr+str$	pqr	0
0	0	0	0	0	0	0	0
Γ	$[p,s]^2[q,t]^2[r,u]^2$	0	0	0	0	0	0
S	$[p,s][q,t][r,u](pqr-stu)$	$2[p,s][q,t][r,u]pqr$	0	0	0	0	0
$H^{(1)}$	$[q,t][r,u]ps$	$-[q,t][r,u]p^2$	$[q,t][r,u]p^2$	0	0	0	0
$H^{(2)}$	$[p,s][r,u]qt$	$-[p,s][r,u]q^2$	0	$[p,s][r,u]q^2$	0	0	0
$H^{(3)}$	$[p,s][q,t]ru$	$-[p,s][q,t]r^2$	0	0	$[p,s][q,t]r^2$	0	0
Id	$pqr+stu$	$sqr+ptr+pqu$	$pqr+ptu$	$pqr+squ$	$pqr+str$	pqr	0
1	1	1	1	1	1	1	1

Observe that for every pair of forms in the table above there is at least one covariant that sends one of the forms to zero, and the other form not to zero. Thus no element of W can be written on two of the above forms. This argument concludes the proof of the proposition. □

We will conclude our discussion by showing two partial ordered sets. One poset on the eight covariants, and one poset on the canonical forms.

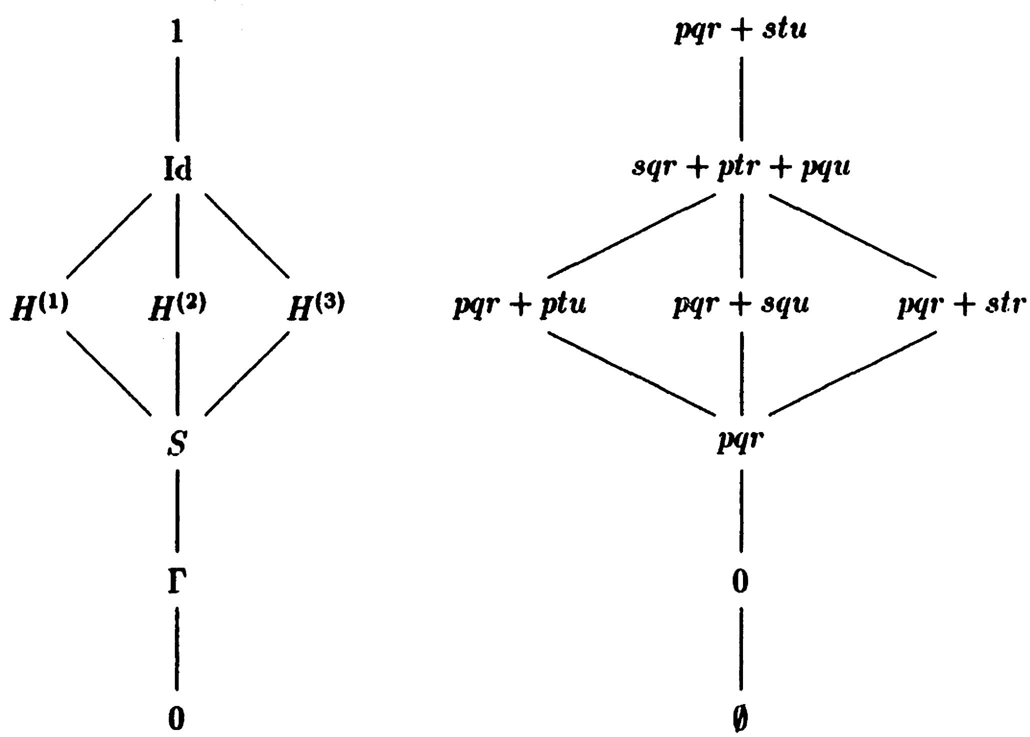
Definition 1.5.7 *For two covariants C and C' we say that $C \leq C'$ if $C'(w) = 0$ implies $C(w) = 0$ for all $w \in W$.*

Definition 1.5.8 *For two forms f and f' we say that $f \leq f'$ if*

$$\overline{\{w \in W \ : \ w \text{ is of the form } f\}} \subseteq \overline{\{w \in W \ : \ w \text{ is of the form } f'\}},$$

where the closure is the topological closure on W, where W has the Euclidean topology.

Lemma 1.5.6 *Consider the following two partial orders. The first partial order is on the eight covariants Γ, S, $H^{(1)}$, $H^{(2)}$, $H^{(3)}$, Id, 1, and 0. The second partial order is on the eight canonical forms $pqr+stu$, $sqr+ptr+pqu$, $pqr+ptu$, $pqr+squ$, $pqr+str$, pqr, 0, and $\emptyset$, where $\emptyset$ is the form no element can be written in.*



Proof: The order relations between the covariant in the first poset follows by studying the table in the proof above. By Corollary 1.2.2 we know that $pqr + stu$ is a generic canonical form. Hence it is the maximal element. The other order relations of the second poset follow easily by letting one of the variables go to zero in limit. □

Corollary 1.5.1 *The following algebraic relation holds between the six covariants* Γ, S, $H^{(1)}$, $H^{(2)}$, $H^{(3)}$, Id.

$$S^2 = \Gamma \cdot \mathrm{Id}^2 - 4 \cdot H^{(1)} \cdot H^{(2)} \cdot H^{(3)}.$$

Proof: By continuity it is enough to consider $w = pqr + stu$.

$$\begin{aligned} S(pqr + stu)^2 &= ([p,s][q,t][r,u](pqr - stu))^2 \\ &= ([p,s][q,t][r,u](pqr + stu))^2 - 4[p,s]^2[q,t]^2[r,u]^2 psqtru \\ &= \Gamma(pqr + stu) \cdot \mathrm{Id}(pqr + stu)^2 - \end{aligned}$$

$$-4 \cdot H^{(1)}(pqr + stu) \cdot H^{(2)}(pqr + stu) \cdot H^{(3)}(pqr + stu).$$

□

Corollary 1.5.2

$$\Gamma = \Delta^{(i)} \circ H^{(i)},$$

where $\Delta^{(i)}$ be the discriminant on quadratic polynomials in the two variables $x_1^{(i)}$ and $x_2^{(i)}$.

Proof: By continuity and symmetry it is enough to consider $w = pqr + stu$ and $i = 1$.

$$\begin{aligned} \Delta^{(1)}\left(H^{(1)}(pqr + stu)\right) &= \Delta^{(1)}([q,t][r,u]ps) \\ &= [[q,t][r,u]p,s]^2 \\ &= [p,s]^2[q,t]^2[r,u]^2 \\ &= \Gamma(pqr + stu). \end{aligned}$$

□

Recall that covariant S maps W into W. Thus we can consider the covariant $C \circ S$, where C is a covariant.

Corollary 1.5.3 *The following three identities hold.*

$$\begin{aligned} \Gamma \circ S &= \Gamma^3, \\ S \circ S &= -\Gamma^2 \cdot \mathrm{Id}, \\ H^{(i)} \circ S &= -\Gamma \cdot H^{(i)}. \end{aligned}$$

Proof: The proof is straightforward. We will only give the proof of the first identity. By continuity it is enough to consider $w = pqr + stu$.

$$\begin{aligned}\Gamma(S(pqr+stu)) &= \Gamma([p,s][q,t][r,u]pqr - [p,s][q,t][r,u]stu) \\ &= [[p,s][q,t][r,u]p, -[p,s][q,t][r,u]s]^2 [q,t]^2[r,u]^2 \\ &= [p,s]^6[q,t]^6[r,u]^6 \\ &= \Gamma(pqr+stu)^3.\end{aligned}$$

□

1.5.5 The binary cubic

We will show in this section how binary cubics relate to $2\times2\times2$ matrices. That will be a study in their covariants abd their canonical forms. Readers are also referred to **[K-R]**.

Recall that W_3 is the linear subspace of $\mathbf{C}[\mathbf{y}] = \mathbf{C}[y_1, y_2]$ consisting of all homogeneous elements of degree 3. An element p of $\mathbf{C}[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}]$ is called symmetric if

$$p(\mathbf{x}^{(\sigma(1))}, \mathbf{x}^{(\sigma(2))}, \mathbf{x}^{(\sigma(3))}) = p(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}),$$

for all permutations σ of $1,2,3$. The symmetric elements of the linear space W of $2\times2\times2$ matrices are of the form

$$\left(\begin{pmatrix} a_0 & a_1 \\ a_1 & a_2 \end{pmatrix}, \begin{pmatrix} a_1 & a_2 \\ a_2 & a_3 \end{pmatrix}\right).$$

Define the algebra homomorphism $\Omega : \mathbf{C}[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}] \longrightarrow \mathbf{C}[\mathbf{y}]$ which is defined by $\Omega\left(x_j^{(i)}\right) = y_j$. Moreover define a linear map $\omega : W_3 \longrightarrow W$ by

$$\omega\left(a_0y_1^3 + 3a_1y_1^2y_2 + 3a_2y_1y_2^2 + a_3y_2^3\right) = \left(\begin{pmatrix} a_0 & a_1 \\ a_1 & a_2 \end{pmatrix}, \begin{pmatrix} a_1 & a_2 \\ a_2 & a_3 \end{pmatrix}\right).$$

Note that for all $w \in W_3$ we have that $\Omega(\omega(w)) = w$.

Lemma 1.5.7 *If C is a covariant of W with index (g_1, g_2, g_3), then the map $\Omega \circ C \circ \omega$ is a covariant of W_3 with index $g_1 + g_2 + g_3$.*

Proof: For $\phi(y_j) = a_{j,1}y_1 + a_{j,2}y_2$, define $\phi_i : V_i \longrightarrow V_i$ by $\phi_i\left(x_j^{(i)}\right) = a_{j,1}x_1^{(i)} + a_{j,2}x_2^{(i)}$. Directly we have that $\det(\phi_i) = \det(\phi)$. More important is the two following identities.

$$\begin{aligned}\Omega(\widehat{\phi}_1\widehat{\phi}_2\widehat{\phi}_3 p(\mathbf{x})) &= \widehat{\phi}\Omega(p(\mathbf{x})),\\ \omega(\widehat{\phi}w) &= \widehat{\phi}_1\widehat{\phi}_2\widehat{\phi}_3\omega(w).\end{aligned}$$

The proof of both of them are straightforward calculations.

Now, we can study the map $\Omega \circ C \circ \omega$.

$$\begin{aligned}\left(\Omega \circ C \circ \omega \circ \widehat{\phi}\right)(w) &= \left(\Omega \circ C \circ (\widehat{\phi}_1\widehat{\phi}_2\widehat{\phi}_3)\right)(\omega(w))\\ &= \Omega\left(\prod_{i=1}^{3}\det(\phi_i)^{g_i}\cdot\widehat{\phi}_1\widehat{\phi}_2\widehat{\phi}_3 C(\omega(w))\right)\\ &= \prod_{i=1}^{3}\det(\phi_i)^{g_i}\cdot\Omega\left(\widehat{\phi}_1\widehat{\phi}_2\widehat{\phi}_3 C(\omega(w))\right)\\ &= \prod_{i=1}^{3}\det(\phi)^{g_i}\cdot\widehat{\phi}\Omega\left(C(\omega(w))\right)\\ &= \det(\phi)^{g_1+g_2+g_3}\cdot\widehat{\phi}(\Omega \circ C \circ \omega)(w).\end{aligned}$$

Thus $\Omega \circ C \circ \omega$ is a covariant of W_3. □

Let us now see what the covariants Γ, S, $H^{(1)}$, $H^{(2)}$, $H^{(3)}$, Id, 1, and 0 of $2 \times 2 \times 2$ matrices correspond to for binary cubics. Let $w = a_0y_1^3 + 3a_1y_1^2y_2 + 3a_2y_1y_2^2 + a_3y_2^3$.

1. The covariant given by $\Omega \circ \Gamma \circ \omega$ is called the discriminant. It is computed by

$$\Delta(w) = (a_0a_3 - a_1a_2)^2 - 4(a_0a_2 - a_1^2)(a_1a_3 - a_2^2).$$

2. The covariant $\Omega \circ S \circ \omega$ of binary cubics is called the Jacobian, and is denoted by T.

3. For the three Hessians we have that

$$\Omega \circ H^{(1)} \circ \omega = \Omega \circ H^{(2)} \circ \omega = \Omega \circ H^{(3)} \circ \omega.$$

Denote this covariant by H. It is also described by

$$H(w) = \frac{1}{36} \cdot \begin{vmatrix} \frac{\partial^2 w}{\partial x_1{}^2} & \frac{\partial^2 w}{\partial x_1 \partial x_2} \\ \frac{\partial^2 w}{\partial x_2 \partial x_1} & \frac{\partial^2 w}{\partial x_2{}^2} \end{vmatrix}.$$

4. Clearly $\Omega \circ \mathrm{Id} \circ \omega = \mathrm{Id}$, the identity map on W_3.

5. $\Omega \circ 1 \circ \omega = 1$, the constant covariant.

6. $\Omega \circ 0 \circ \omega = 0$, the zero covariant.

A binary cubic can be written in one of the following forms $p^3 + q^3$, $3p^2q$, p^3, and 0, where $p, q \in \mathrm{span}(y_1, y_2)$ and p and q are linearly independent.

We conclude the the discussion about covariants and canonical forms of the binary cubics, by presenting the following table.

	$p^3 + q^3$	$3p^2q$	p^3	0
0	0	0	0	0
Δ	$[p,q]^6$	0	0	0
T	$[p,q]^3(p^3 - q^3)$	$2[p,q]^3p^3$	0	0
H	$[p,q]^2pq$	$-[p,q]^2p^2$	0	0
Id	$p^3 + q^3$	$3p^2q$	p^3	0
1	1	1	1	1

Observe that a binary cubic w can be written in the form $p^3 + q^3$ where p and q are independent, if and only if w does not have any multiple roots. A binary cubic without multiple roots can be written in the form stu, where s, t and u are pairwise linearly independent. Thus one can ask how to express covariants of the binary cubic stu.

$$\Delta(stu) = -\frac{1}{27} \cdot [s,t]^2[t,u]^2[u,s]^2.$$

This is the classical way to describe the discriminant. Similarly, we can express the Hessian of a binary cubic as

$$H(stu) = \frac{1}{9} \cdot e_2([t,u]s, [u,s]t, [s,t]u),$$

where e_2 is the second elementary symmetric function, that is, $e_2(x,y,z) = xy + xz + yz$. Finally, the Jacobian T can be written as

$$T(stu) = \frac{1}{27} \cdot ([t,u]s - [u,s]t) \cdot ([u,s]t - [s,t]u) \cdot ([s,t]u - [t,u]s).$$

This can also be written as

$$T(stu) = \frac{1}{27} \cdot \det \begin{pmatrix} 1 & 1 & 1 \\ [t,u]s & [u,s]t & [s,t]u \\ [t,u]^2s^2 & [u,s]^2t^2 & [s,t]^2u^2 \end{pmatrix}.$$

Recall that we have the identity $[t,u]s + [u,s]t + [s,t]u = 0$.

1.5.6 Some skew-tensors

Let U be a 6 dimensional linear space and consider the linear space $Ext_3(U)$. By Corollary 1.3.1 we know that this linear space has essential rank 2. That is, a generic skew-tensor of step 3 can be written in the form $p \wedge q \wedge r + s \wedge t \wedge u$, where $p, q, r, s, t, u \in U$.

Proposition 1.5.8 *An element $w \in \mathrm{Ext}_3(U)$ can be written in exactly one of the following five forms*

$$\begin{array}{l} p \wedge q \wedge r + s \wedge t \wedge u, \\ s \wedge q \wedge r + p \wedge t \wedge r + p \wedge q \wedge u, \\ p \wedge q \wedge r + p \wedge t \wedge u, \\ p \wedge q \wedge r, \\ 0 \end{array}$$

where $p, q, r, s, t, u \in U$, and the elements p, q, r, s, t, and u are linearly independent.

For a proof see **[G-R-S]** pages 69-72. They also present three covariants C_1, C_2, and C_3, and show how they relate to the above canonical forms. Their relations are that C_1 vanishes on the two last forms, C_2 vanishes on the three last forms, and C_3 is only nonzero on the first form.

Observe that

$$\begin{array}{rcl} C_1: \ \mathrm{Ext}_3(U) & \longrightarrow & \mathrm{Ext}_5(U) \otimes U, \\ C_2: \ \mathrm{Ext}_3(U) & \longrightarrow & \mathrm{Ext}_6(U) \otimes \mathrm{Ext}_3(U) \cong \mathbf{C} \otimes \mathrm{Ext}_3(U) \cong \mathrm{Ext}_3(U), \\ C_3: \ \mathrm{Ext}_3(U) & \longrightarrow & \mathrm{Ext}_6(U) \otimes \mathrm{Ext}_6(U) \cong \mathbf{C} \otimes \mathbf{C} \cong \mathbf{C}. \end{array}$$

The covariant C_1 has index 0, and the covariant C_2 has index 1. Lastly C_3 is an invariant with index 2.

We present here a complete description of the relation between these covariants and the canonical forms of the linear space $\mathrm{Ext}_3(U)$. To be able to do so in a nice way, we will multiply the covariants C_1, C_2, and C_3 by appropriate constants. Thus we will consider the following covariants.

$$\begin{array}{rcl} D_1 & = & -C_1, \\ D_2 & = & -\frac{1}{3} \cdot C_2, \\ D_3 & = & -\frac{1}{6} \cdot C_3. \end{array}$$

In the following table we have surpressed the writing out the wedge product between the vectors.

	$pqr+stu$	$sqr+ptr+pqu$	$pqr+ptu$	pqr	0
0	0	0	0	0	0
D_3	$[p,q,r,s,t,u]^2$	0	0	0	0
D_2	$[p,q,r,s,t,u](pqr-stu)$	$2[p,q,r,s,t,u]pqr$	0	0	0
D_1	$(psqtr)\otimes u$ $+(psqtu)\otimes r$ $+(psruq)\otimes t$ $+(psrut)\otimes q$ $+(qtrup)\otimes s$ $+(qtrus)\otimes p$	$-2\cdot(psqtr)\otimes r$ $-2\cdot(psruq)\otimes q$ $-2\cdot(qtrup)\otimes p$	$2\cdot(qtrup)\otimes p$	0	0
Id	$pqr+stu$	$sqr+ptr+pqu$	$pqr+ptu$	pqr	0
1	1	1	1	1	1

The table is constructed in the same manner as the corresponding table for $2\times 2\times 2$ matrices. The only work is to compute the covariants on the element $e_1\wedge e_2\wedge e_3+e_4\wedge e_5\wedge e_6$, where $e_1,\ldots,e_6$ is a basis for U That completes the first column. The other columns follows by continuity arguments.

Similarly to Corollary 1.5.3 we have the next corollary.

Corollary 1.5.4 *The following two identities hold.*

$$\begin{aligned} D_3\circ D_2 &= D_3^3, \\ D_2\circ D_2 &= -D_3^2\cdot \mathrm{Id}. \end{aligned}$$

Proof: The proof is straightforward. By continuity it is enough to consider $v = p \wedge q \wedge r + s \wedge t \wedge u$.

$$\begin{aligned} D_3\left(D_2(p \wedge q \wedge r + s \wedge t \wedge u)\right) &= D_3\left([p,q,r,s,t,u]p \wedge q \wedge r - [p,q,r,s,t,u]s \wedge t \wedge u\right) \\ &= [[p,q,r,s,t,u]p,q,r,-[p,q,r,s,t,u]s,t,u]^2 \\ &= [p,q,r,s,t,u]^6 \\ &= D_3(pqr+stu)^3. \end{aligned}$$

$$\begin{aligned} D_2\left(D_2(p \wedge q \wedge r + s \wedge t \wedge u)\right) &= D_2\left([p,q,r,s,t,u]p \wedge q \wedge r - [p,q,r,s,t,u]s \wedge t \wedge u\right) \\ &= [[p,q,r,s,t,u]p,q,r,-[p,q,r,s,t,u]s,t,u] \cdot \\ &\quad \cdot \left([p,q,r,s,t,u]p \wedge q \wedge r + [p,q,r,s,t,u]s \wedge t \wedge u\right) \\ &= -[p,q,r,s,t,u]^4 \cdot (p \wedge q \wedge r + s \wedge t \wedge u) \\ &= -D_3(pqr+stu)^2 \cdot (p \wedge q \wedge r + s \wedge t \wedge u). \end{aligned}$$

□

Let $e_1, \ldots, e_6$ be a basis for U, such that $[e_1, \ldots, e_6] = 1$. Let $\mathcal{I}$ be the set of all three elements subsets of the set $\{1, \ldots, 6\}$. Define for $I \in \mathcal{I}$ the element e_I in $\mathrm{Ext}_3(U)$ by

$$e_I = e_i \wedge e_j \wedge e_k,$$

where $I = \{i,j,k\}$ and $i < j < k$. Observe that $\{e_I\}_{I \in \mathcal{I}}$ form a basis for $\mathrm{Ext}_3(U)$.

Proposition 1.5.9 *The invariant D_3 on the element $v = \sum_{I \in \mathcal{I}} a_I e_I$ is given by*

$$\begin{aligned} D_3(v) &= \sum_{\mathbf{i} \in L_1} a^2_{\{i_1,i_2,i_3\}} \cdot a^2_{\{i_4,i_5,i_6\}} - \\ &\quad -2 \cdot \sum_{\mathbf{i} \in L_2} a_{\{i_1,i_3,i_4\}} \cdot a_{\{i_1,i_5,i_6\}} \cdot a_{\{i_2,i_3,i_4\}} \cdot a_{\{i_2,i_5,i_6\}} + \end{aligned}$$

$$+4 \cdot \sum_{\mathbf{i} \in L_3} \Big(a_{\{i_1,i_3,i_5\}} \cdot a_{\{i_1,i_4,i_6\}} \cdot a_{\{i_2,i_3,i_6\}} \cdot a_{\{i_2,i_4,i_5\}} +$$

$$+ a_{\{i_2,i_4,i_6\}} \cdot a_{\{i_2,i_3,i_5\}} \cdot a_{\{i_1,i_4,i_5\}} \cdot a_{\{i_1,i_3,i_6\}} \Big),$$

where

$$\begin{aligned} L_1 &= \{\mathbf{i} \;:\; i_1,\ldots,i_6 \text{ distinct}, i_1 < i_2 < i_3, i_4 < i_5 < i_6, i_1 < i_4\}, \\ L_2 &= \{\mathbf{i} \;:\; i_1,\ldots,i_6 \text{ distinct}, i_1 < i_2, i_3 < i_4, i_5 < i_6, i_3 < i_5\}, \\ L_3 &= \{\mathbf{i} \;:\; i_1,\ldots,i_6 \text{ distinct}, i_1 < i_3 < i_5, i_1 < i_2, i_3 < i_4, i_5 < i_6\}. \end{aligned}$$

Proof: Since D_3 is an invariant, $D_3(v)$ is a polynomial of the a_I's. Thus we can write $D_3(v)$ as

$$D_3(v) = \sum_{f:\mathcal{I}\to\mathbf{N}} b_f \cdot \prod_{I\in\mathcal{I}} a_I^{f(I)}.$$

Consider the map $\phi(e_i) = \alpha_i \cdot e_i$, where $\alpha_i \in \mathbf{C}$. Observe that $\det(\phi) = \alpha_1 \cdots \alpha_6$. Thus we have that

$$\begin{aligned} (\alpha_1 \cdots \alpha_6)^2 \cdot \sum_{f:\mathcal{I}\to\mathbf{N}} b_f \cdot \prod_{I\in\mathcal{I}} a_I^{f(I)} &= (\alpha_1 \cdots \alpha_6)^2 \cdot D_3(v) \\ &= \det(\phi)^2 \cdot D_3(v) \\ &= D_3(\widehat{\phi} v) \\ &= \sum_{f:\mathcal{I}\to\mathbf{N}} b_f \cdot \prod_{I\in\mathcal{I}} \alpha_I^{f(I)} a_I^{f(I)}. \end{aligned}$$

Since this is true for all α_I, we see that when b_f is nonzero, we have that for all $i \in \{1,\ldots,6\}$ that $\sum_{i\in I\in\mathcal{I}} f(I) = 2$. So $3\cdot\sum_{I\in\mathcal{I}} f(I) = \sum_{i\in\{1,\ldots,6\}} \sum_{i\in I\in\mathcal{I}} f(I) = \sum_{i\in\{1,\ldots,6\}} 2 = 12$. Thus each monomial in $D_3(v)$ correspond to a four elements of the set $\mathcal{I}$, such that each $i \in \{1,\ldots,6\}$ lies in exactly two of the four sets.

There are then three possible types of monomials.

(i) The four sets are J, J, K, K, where $J, K \in \mathcal{I}$ and K is the complement of J. Let g be the corresponding function, that is, $g(J) = g(K) = 2$, and the other values of g

vanish.

$$\begin{aligned}
b_g &= b_g \cdot \prod_{I\in\mathcal{I}} (1)^{g(I)} \\
&= D_3(e_J + e_K) \\
&= [e_{j_1}, e_{j_2}, e_{j_3}, e_{k_1}, e_{k_2}, e_{k_3}]^2 \\
&= (\pm 1)^2 = 1.
\end{aligned}$$

Hence the coefficient of all these terms are 1. There are 10 monomials in this form.

(ii) The four set are $\{j\} \cup A, \{k\} \cup A, \{j\} \cup B, \{k\} \cup B$, where $\{j\} \cup \{k\} \cup A \cup B = \{1, \ldots, 6\}$, and $|A| = |B| = 2$. Let g be the corresponding function, that is, $g(\{j\} \cup A) = g(\{k\} \cup A) = g(\{j\} \cup B) = g(\{k\} \cup B) = 1$, and otherwise g vanish. Also define $g_1(\{j\} \cup A) = g_1(\{k\} \cup B) = g_2(\{k\} \cup A) = g_2(\{j\} \cup B) = 2$.

$$\begin{aligned}
b_g + b_{g_1} + b_{g_2} &= D_3(e_{\{j\}\cup A} + e_{\{k\}\cup A} + e_{\{j\}\cup B} + e_{\{k\}\cup B}) \\
&= D_3((\epsilon_j + \epsilon_k)\epsilon_{a_1}\epsilon_{a_2} + (\epsilon_j + \epsilon_k)\epsilon_{b_1}\epsilon_{b_2}) \\
&= 0.
\end{aligned}$$

Thus we conclude that $b_g = -b_{g_1} - b_{g_2} = -2$. There are 45 monomials in this form.

(iii) The four set are $\{i_1, i_3, i_5\}, \{i_1, i_4, i_6\}, \{i_2, i_3, i_6\}, \{i_2, i_4, i_5\}$, where $i_1, \ldots, i_6$ are distinct. Let g be the corresponding function to this monomial.

$$\begin{aligned}
b_g &= D_3(e_{i_1}e_{i_3}e_{i_5} + e_{i_1}e_{i_4}e_{i_6} + e_{i_2}e_{i_3}e_{i_6} + e_{i_2}e_{i_4}e_{i_5}) \\
&= D_3\left(\frac{1}{2}(\epsilon_1 + \epsilon_2)(\epsilon_3 + \epsilon_4)(\epsilon_5 + \epsilon_6) + \frac{1}{2}(\epsilon_1 - \epsilon_2)(\epsilon_3 - \epsilon_4)(\epsilon_5 - \epsilon_6)\right) \\
&= \left[\frac{1}{2}(\epsilon_1 + \epsilon_2), \epsilon_3 + \epsilon_4, \epsilon_5 + \epsilon_6, \frac{1}{2}(\epsilon_1 - \epsilon_2), \epsilon_3 - \epsilon_4, \epsilon_5 - \epsilon_6\right]^2 \\
&= 4.
\end{aligned}$$

Thus we conclude that $b_g = 4$. There are 30 monomials in this form.

Hence the proof is complete. □

Recall that V_i is the linear space spanned by $x_1^{(i)}$ and $x_2^{(i)}$ for $i = 1, 2, 3$. We can assume that

$$U = \bigoplus_{i=1}^{3} V_i,$$

since the dimension of U is 6. We can naturally think of V_i as linear subspace of U, such that V_1, V_2, and V_3 pairwise only intersect in 0. Also recall that $V_1 \odot V_2 \odot V_3 \cong W$, which is the linear space of $2 \times 2 \times 2$ matrices.

Define the linear map $\xi : V_1 \odot V_2 \odot V_3 \longrightarrow \mathrm{Ext}_3(U)$ which is defined by

$$\xi\,(p \odot q \odot r) = p \wedge q \wedge r,$$

where $p \in V_1$, $q \in V_2$, and $r \in V_3$. Observe that this map is injective. Hence we can define the inverse map Ξ from the image of ξ to $V_1 \odot V_2 \odot V_3$, by

$$\Xi\,(p \wedge q \wedge r) = p \odot q \odot r,$$

where $p \in V_1$, $q \in V_2$, and $r \in V_3$. Note that Ξ is linear. Also we have that $\Xi(\xi(w)) = w$ for all $w \in W \cong V_1 \odot V_2 \odot V_3$.

Moreover, on the linear space U we can define a bracket of step 6. That is, a skew-symmetric multilinear form from $U^{\otimes 6}$ to $\mathbf{C}$. Recall that we view U as the sum of V_1, V_2, and V_3. Since on each of these three linear spaces we have brackets of step 2, we can lift these brackets to U. That is, we can define a bracket on U such that for $p_1, p_2 \in V_1$, $q_1, q_2 \in V_2$, and $r_1, r_2 \in V_3$ we have that

$$[p_1, p_2, q_1, q_2, r_1, r_2] = [p_1, p_2] \cdot [q_1, q_2] \cdot [r_1, r_2].$$

Lemma 1.5.10 *The invariant* Γ *on* $2 \times 2 \times 2$ *matrices is related to the invariant* D_3 *by the following identity*

$$D_3 \circ \xi = \Gamma.$$

Similarly the covariant S *on* $2 \times 2 \times 2$ *matrices is related to* D_2 *by*

$$\Xi \circ D_2 \circ \xi = S.$$

Proof: It is enough to prove these two identities for a dense subset of $W \cong V_1 \otimes V_2 \otimes V_3$, since invariants and covariants are continuous. We know that a generic element of W is in the form $p \otimes q \otimes r + s \otimes t \otimes u$, where $p, s \in V_1$, $q, t \in V_2$, and $r, u \in V_3$.

$$\begin{aligned} D_3(\xi(p \otimes q \otimes r + s \otimes t \otimes u)) &= D_3(p \wedge q \wedge r + s \wedge t \wedge u) \\ &= [p, q, r, s, t, u]^2 \\ &= [p, q]^2 \cdot [r, s]^2 \cdot [t, u]^2 \\ &= \Gamma(p \otimes q \otimes r + s \otimes t \otimes u). \end{aligned}$$

$$\begin{aligned} \Xi(D_3(\xi(p \otimes q \otimes r + s \otimes t \otimes u))) &= \Xi(D_3(p \wedge q \wedge r + s \wedge t \wedge u)) \\ &= \Xi([p, q, r, s, t, u] \cdot (p \wedge q \wedge r - s \wedge t \wedge u)) \\ &= [p, q, r, s, t, u] \cdot (p \otimes q \otimes r - s \otimes t \otimes u) \\ &= [p, q] \cdot [r, s] \cdot [t, u] \cdot (p \otimes q \otimes r - s \otimes t \otimes u) \\ &= S(p \otimes q \otimes r + s \otimes t \otimes u). \end{aligned}$$

□

Chapter 2

A Bijective Proof of Infinite Variated Good's Inversion

2.1 Formal power series and colored sets

Let $\mathfrak{I}$ be a set, possible infinite.

Definition 2.1.1 *A* multi index $\mathbf{n}$ *is a vector whose entries are nonnegative integers whose sum is finite and the entries are indexed by* $\mathfrak{I}$. *That is* $\mathbf{n} = (n_i)_{i \in \mathfrak{I}}$, *where* $n_i \in \mathbb{N}$. *Define the base multi indices by* $\mathbf{e}_j = (\delta_{j,i})_{i \in \mathfrak{I}}$ *for* $j \in \mathfrak{I}$.

Notice that any multi index $\mathbf{n}$ can be written as a linear combination of the base multi indices

$$\mathbf{n} = \sum_{i \in \mathfrak{I}} n_i \cdot \mathbf{e}_i,$$

where there are only a finite number of nonzero terms in the sum. For a finite subset $I \subseteq \mathfrak{I}$, define

$$\mathbf{e}_I = \sum_{i \in I} \mathbf{e}_i.$$

Let $\mathbf{K}$ be a field of characteristic 0. We will be considering formal power series in the variables $(x_i)_{i \in \mathfrak{I}}$. A power series is denoted with $f(\mathbf{x})$. To be able write readable formulas we introduce the following notations.

$$\begin{aligned}
\mathbf{n} &= (n_i)_{i\in\mathfrak{I}} \\
\mathbf{n}! &= \prod_{i\in\mathfrak{I}} n_i! \\
\mathbf{x} &= (x_i)_{i\in\mathfrak{I}} \\
\mathbf{x}^{\mathbf{n}} &= \prod_{i\in\mathfrak{I}} x_i^{n_i} \\
\binom{\mathbf{n}}{\mathbf{k}} &= \prod_{i\in\mathfrak{I}} \binom{n_i}{k_i} \\
(\mathbf{n})_{\mathbf{k}} &= \prod_{i\in\mathfrak{I}} (n_i)_{k_i}
\end{aligned}$$

Thus a formal power series can be written as an exponential power series

$$f(\mathbf{x}) = \sum_{\mathbf{n}} a_{\mathbf{n}} \frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!}.$$

The sum of two power series $f(\mathbf{x}) = \sum_{\mathbf{n}} a_{\mathbf{n}} \frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!}$ and $g(\mathbf{x}) = \sum_{\mathbf{n}} b_{\mathbf{n}} \frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!}$, is defined componentwise. That is

$$(f+g)(\mathbf{x}) = \sum_{\mathbf{n}} (a_{\mathbf{n}} + b_{\mathbf{n}}) \frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!}.$$

The product of two power series $f(\mathbf{x})$ and $g(\mathbf{x})$ is

$$(f \cdot g)(\mathbf{x}) = \sum_{\mathbf{n}} c_{\mathbf{n}} \frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!},$$

where

$$c_{\mathbf{n}} = \sum_{0 \le \mathbf{k} \le \mathbf{n}} \binom{\mathbf{n}}{\mathbf{k}} a_{\mathbf{k}} b_{\mathbf{n}-\mathbf{k}}.$$

Definition 2.1.2 *A* collection *power series* $\mathbf{g}(\mathbf{x})$ *is a set of formal power series indexed by the set* $\mathfrak{I}$. *That is*

$$\mathbf{g}(\mathbf{x}) = (g_i(\mathbf{x}))_{i\in\mathfrak{I}}.$$

A summable *collection* $\mathbf{g}(\mathbf{x})$ *is a collection such that for every multi index* $\mathbf{n}$ *the coefficient* $\left[\frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!}\right] g_i(\mathbf{x})$ *is nonzero for only a finite number of* $i \in \mathfrak{I}$.

Definition 2.1.3 *Let* $f(\mathbf{x})$ *be a formal power series and* $\mathbf{g}(\mathbf{x})$ *be a summable collection formal power series, such that* $g_i(\mathbf{x})$ *has no constant coefficient. Define the* composition $f \circ \mathbf{g}$ *as*

$$(f \circ \mathbf{g})(\mathbf{x}) = f\left((g_i(\mathbf{x}))_{i \in \mathfrak{I}}\right).$$

Observe that $\mathbf{x}^{\mathbf{n}} \circ \mathbf{g} = \prod_{i \in \mathfrak{I}} g_i(\mathbf{x})^{n_i}$. We will also write this expression as $\mathbf{g}^{\mathbf{n}}$.

Definition 2.1.4 *A* colored set (E, f) *is a set* E *with a function* $f : E \longrightarrow \mathfrak{I}$. *The* color *of an element* $a \in E$ *is the value* $f(a)$. *The colored set* (E, f) *is finite if* E *is a finite set. The cardinality of a finite colored set* (E, f) *is a multi index card*$(E, f) = \mathbf{n}$ *such that*

$$n_i = \left|\{a \in E \;:\; f(a) = i\}\right|.$$

If $\text{card}(E, f) = \mathbf{n}$ we say that (E, f) is a $\mathbf{n}$ set. When we need to speak about a generic colored set of cardinality $\mathbf{n}$, we are going to write $\mathbf{n}$ for this generic set.

2.2 Colored species

We introduce now the theory of colored species. This theory was developed in **[M-N]**. We will only give a short sketch of definitions and main results. The reader interested in this subject is referred to **[M-N]**.

Let $\mathbf{B}$ be the category of finite sets and bijections. Recall that a *species* is a functor from $\mathbf{B}$ to $\mathbf{B}$. Similarly we can define the category of colored sets.

Definition 2.2.1 *Let* $\mathbf{B}_{\mathfrak{I}}$ *be the category of finite colored sets and bijections, which preserves color. A* colored species M *is functor from* $\mathbf{B}_{\mathfrak{I}}$ *to* $\mathbf{B}$.

Define the generating function of colored species M to be

$$\text{card}(M; \mathbf{x}) = \sum_{\mathbf{n}} |M[\mathbf{n}]| \cdot \frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!},$$

where $M[\mathbf{n}]$ is the species applied to a generic $\mathbf{n}$ set.

For $i \in \mathfrak{I}$ define the colored species X_i by

$$X_i[(E,f)] = \begin{cases} \{E\} & \text{if } \mathrm{card}(E,f) = \mathbf{e}_i \\ \emptyset & \text{otherwise} \end{cases}.$$

Observe that $\mathrm{card}(X_i;\mathbf{x}) = x_i$.

Define sum and product of two colored species by

$$\begin{aligned} (M+N)[(E,f)] &= M[(E,f)] \,\dot\cup\, N[(E,f)] \\ (M \cdot N)[(E,f)] &= \dot{\bigcup_{E_1+E_2=E}} M[(E_1, f|_{E_1})] \times N[(E_2, f|_{E_2})] \end{aligned}$$

For a colored species M and for $i \in \mathfrak{I}$ define the colored species $M'^{(i)}$ by

$$M'^{(i)}[(E,f)] = M[(E \,\dot\cup\, \{*\}, f)],$$

where $*$ is a ghost element of color i. That is we extend f such that $f(*) = i$. Moreover let

$$M^{\bullet(i)} = X_i \cdot M'^{(i)}.$$

That means that we mark an element of color i in the underlying set.

A *colored partition* of a colored set (E,f) is a partition π of the set E, with a function $g : \pi \longrightarrow \mathfrak{I}$. Let $\Pi[(E,f)]$ be the set of all colored partitions of (E,f).

Definition 2.2.2 *A* collection *of colored species* $\vec{\mathbf{M}}$ *is a set of colored species indexed by the set* $\mathfrak{I}$. *That is*

$$\vec{\mathbf{M}} = (M_i)_{i\in\mathfrak{I}}.$$

A summable *collection* $\vec{\mathbf{M}}$ *is a collection such that for every colored set* (E,f) *the set* $M_i[(E,f)]$ *is nonempty for only a finite number of* $i \in \mathfrak{I}$.

Observe that is $\vec{\mathbf{M}}$ is a summable collection of colored species, then $(\mathrm{card}(M_i;\mathbf{x}))_{i\in\mathfrak{I}}$ is a summable collection of power series.

Let M be a colored species and let $\vec{\mathbf{N}}$ be a summable collection of colored species, such that $N_i[\emptyset] = \emptyset$ for all $i \in \mathfrak{I}$. Define the divided power $\Gamma_{\mathbf{k}}(\vec{\mathbf{N}})$ as

$$\left(\Gamma_{\mathbf{k}}(\vec{\mathbf{N}})\right)[(E,f)] = \dot{\bigcup_{(\pi,g)\in\Pi[(E,f)],\ \mathrm{card}(\pi,g)=\mathbf{k}}} \prod_{B\in\pi} N_{g(B)}[(B, f|_B)].$$

Define the composition $M \circ \vec{\mathbf{N}}$ by

$$(M \circ \vec{\mathbf{N}})[(E,f)] = \dot{\bigcup_{(\pi,g)\in\Pi[(E,f)]}} M[(\pi,g)] \times \prod_{B\in\pi} N_{g(B)}[(B, f|_B)].$$

Observe that the summability condition of the collection $\vec{\mathbf{N}}$ implies that the two sets $\left(\Gamma_{\mathbf{k}}(\vec{\mathbf{N}})\right)[(E,f)]$ and $(M \circ \vec{\mathbf{N}})[(E,f)]$ are finite.

Note the following identity

$$\prod_{i\in\mathfrak{I}} X_i^{n_i} \circ \vec{\mathbf{N}} = \prod_{i\in\mathfrak{I}} N_i^{n_i}.$$

We will be writing the above expression as $\vec{\mathbf{N}}^{\mathbf{n}}$.

Now we can show the relationship between operations on colored species and operations on the respective generating functions.

Proposition 2.2.1

$$\begin{aligned}
card(M+N;\mathbf{x}) &= card(M;\mathbf{x}) + card(N;\mathbf{x}) \\
card(M\cdot N;\mathbf{x}) &= card(M;\mathbf{x}) \cdot card(N;\mathbf{x}) \\
card(M'^{(i)};\mathbf{x}) &= \frac{\partial}{\partial x_i} card(M;\mathbf{x}) \\
card(M^{\bullet(i)};\mathbf{x}) &= x_i \frac{\partial}{\partial x_i} card(M;\mathbf{x}) \\
card((M\circ\vec{\mathbf{N}});\mathbf{x}) &= (card(M;\mathbf{x}) \circ ((card(N_i;\mathbf{x}))_{i\in\mathfrak{I}}))(\mathbf{x}) \\
card(\Gamma_{\mathbf{k}}(\vec{\mathbf{N}});\mathbf{x}) &= \left(\frac{\mathbf{x}^{\mathbf{k}}}{\mathbf{k}!} \circ ((card(N_i;\mathbf{x}))_{i\in\mathfrak{I}})\right)(\mathbf{x})
\end{aligned}$$

$$= \frac{1}{\mathbf{k}!} \prod_{i \in \Im} card(N_i; \mathbf{x})^{k_i}$$

2.3 Colored functions and colored trees

Let $\vec{\mathbf{M}} = (M_i)_{i \in \Im}$ be a collection of colored species, not necessarily a summable collection.

Definition 2.3.1 *An* $\vec{\mathbf{M}}$-enriched function ϕ *from* (E, f) *to* (F, g) *is a function* ϕ *from* E *to* F *such that for all* $b \in F$ *the colored set*

$$(\{a \in E \mid \phi(a) = b\}, f) = \left(\phi^{-1}(b), f\right)$$

is enriched with a $M_{g(b)}$ *structure.*

Lemma 2.3.1 *Let* (F, g) *be a fixed colored set. The set of structures defined on the colored set* (E, f) *by*

$$\left(\prod_{b \in F} M_{g(b)}\right) [(E, f)],$$

is isomorphic to the set of $\vec{\mathbf{M}}$*-enriched functions from* (E, f) *to* (F, g).

Definition 2.3.2 *A* $\vec{\mathbf{M}}$-enriched tree (forest) *on a* $\mathbf{n}$ *set* (E, f) *is a tree (forest) on* E *such that for every node* $a \in E$ *we put a* $M_{f(a)}$ *structure on its set of sons.*

Let $A^{(i)}_{\vec{\mathbf{M}}}$ be the colored species of $\vec{\mathbf{M}}$-enriched colored trees with the root of color i. Let $\vec{\mathbf{A}}_{\vec{\mathbf{M}}}$ be the collection $\left(A^{(i)}_{\vec{\mathbf{M}}}\right)_{i \in \Im}$.

Proposition 2.3.2 *The collection* $\vec{\mathbf{A}}_{\vec{\mathbf{M}}}$ *is a summable collection of colored species and the colored species* $A^{(i)}_{M}$ *fulfills the following functional equation*

$$A^{(i)}_{\vec{\mathbf{M}}} = X_i \cdot (M_i \circ \vec{\mathbf{A}}_{\vec{\mathbf{M}}}).$$

Proof: Look at the trees $A^{(i)}_{\vec{\mathbf{M}}}[(E,f)]$. The root of the trees has color i and the root lies in the colored set (E,f). But the colored set (E,f) has only a finite number of colors. Hence $A^{(i)}_{\vec{\mathbf{M}}}[(E,f)]$ will only be nonempty for a finite number of $i \in \mathfrak{I}$. Thus the collection $\vec{\mathbf{A}}_{\vec{\mathbf{M}}}$ is summable.

The colored species $A^{(i)}_{\vec{\mathbf{M}}}$ puts a enriched tree with the root of color i on a colored set. Thus $(M_i \circ (\vec{\mathbf{A}}_{\vec{\mathbf{M}}})_{i\in\mathfrak{I}})$ puts a colored partition on a set, where each block of color i the structure $A^{(i)}_{\vec{\mathbf{M}}}$, and the set of blocks receives a M_i structure. But note that $A^{(i)}_{\vec{\mathbf{M}}}$ is a colored tree. Moreover this colored tree has the root of color i. Since every block contains a unique root, we can view it as putting the M_i structure on the roots of all the trees. Thus $(M_i \circ \vec{\mathbf{A}}_{\vec{\mathbf{M}}})$ is a colored forest with an M_i structure on the roots.

Now by multiplying with X_i, we will start by selecting an element of color i. Thus in the set of structures the colored species $X_i \cdot (M \circ \vec{\mathbf{A}}_{\vec{\mathbf{M}}})$ describes, join this selected element to the roots of the forest. Thus we have a $\vec{\mathbf{M}}$-enriched colored tree, and the equation follows. □

The implicit species theorem **[M-N]** implies that the equation system $\vec{\mathbf{Y}} = \vec{\mathbf{X}} \cdot (\vec{\mathbf{M}} \circ \vec{\mathbf{Y}})$ has a unique solution $\vec{\mathbf{Y}}$, that is summable.

Define the colored species $End_{\vec{\mathbf{M}}}$ by letting $End_{\vec{\mathbf{M}}}[(E,f)]$ be all $\vec{\mathbf{M}}$-enriched colored functions from the colored set (E,f) to itself. Such a colored function is called a $\vec{\mathbf{M}}$-*enriched colored endofunction.*

An $\vec{\mathbf{M}}$-*enriched contraction* on a colored set (E,f) is an $\vec{\mathbf{M}}$-enriched function ϕ such that there exist a node $a \in E$, such that for all $b \in E$ there exists a positive integer k such that for all $n \geq k$ we have $\phi^n(b) = a$. Observe that $\phi(a) = a$. We call the vertex a the attracting point. The contraction has depth m if $\phi^m(b) = a$ for all $b \in E$.

Lemma 2.3.3 *The colored species of* $\vec{\mathbf{M}}$*-enriched contractions with the attracting point of color* i *is described by*

$$X_i \cdot (M_i'^{(i)} \circ \vec{\mathbf{A}}_{\vec{\mathbf{M}}}).$$

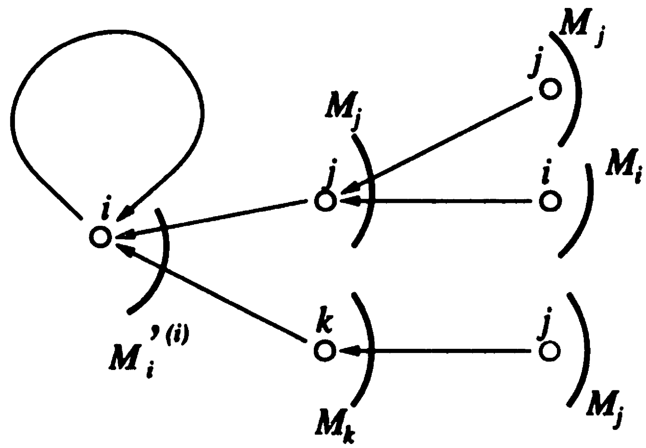


Figure 2-1: An example of an $\vec{\mathbf{M}}$-enriched contraction.

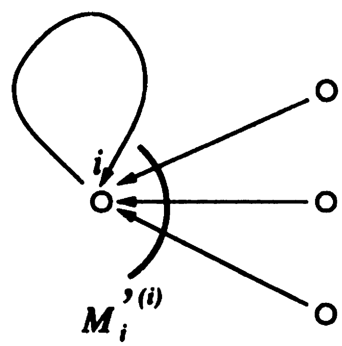


Figure 2-2: An example of an $\vec{\mathbf{M}}$-enriched contraction of depth 1 without enriched leaves.

Proof: The colored species X_i chooses the attracting point a of color i. Then the set $(E - \{a\}, f|_{E-\{a\}})$ has the structure of an $\vec{\mathbf{M}}$-enriched colored forest. But the roots and the attracting vertex a has a M_i structure on them. This is equivalent to put an $M_i'^{(i)}$ structure on the roots. □

Lemma 2.3.4 *The colored species $M_i^{\bullet(i)}$ is naturally isomorphic to $\vec{\mathbf{M}}$-enriched contractions ϕ, which has depth 1, attracting vertex of color i and no structure put on $\phi^{-1}(b)$ if b is not the attracting vertex.*

Let ϕ be a colored function from the colored set $(E, f|_E)$ to the colored set $(R \dot\cup E, f)$. An element a in the set E is called periodic if there exists m, a positive integer such that $\phi^m(a) = a$.

Definition 2.3.3 *Define the species $\mathcal{N}^{\mathbf{k}}_{\vec{\mathbf{M}}}$ of non periodic $\vec{\mathbf{M}}$-enriched colored functions as follows, let $\mathcal{N}^{\mathbf{k}}_{\vec{\mathbf{M}}}[(E, f)]$ be the set of all $\vec{\mathbf{M}}$-enriched colored functions ϕ from the colored set (E, f) to the colored set $(E, f) \dot\cup \mathbf{k}$, such that there is no periodic element in the colored set (E, f).*

Lemma 2.3.5 *Let $(R \dot\cup E, f)$ be an $\mathbf{n}$ set and assume that $(R, f|_R)$ is a $\mathbf{k}$ set. The set of all $\vec{\mathbf{M}}$-enriched colored functions from $(E, f|_E)$ to $(R \dot\cup E, f)$ is described by*

$$\left(\mathcal{N}^{\mathbf{k}}_{\vec{\mathbf{M}}} \cdot End_{\vec{\mathbf{M}}}\right)[(E, f|_E)].$$

Lemma 2.3.6 *For a collection of colored species $\vec{\mathbf{M}}$ we have that*

$$\left|\left(\Gamma_{\mathbf{k}}(\vec{\mathbf{A}}_{\vec{\mathbf{M}}})\right)[\mathbf{n}]\right| = \binom{\mathbf{n}}{\mathbf{k}} \cdot \left|\mathcal{N}^{\mathbf{k}}_{\vec{\mathbf{M}}}[\mathbf{n} - \mathbf{k}]\right|.$$

Proof: $\left(\Gamma_{\mathbf{k}}(\vec{\mathbf{A}}_{\vec{\mathbf{M}}})\right)[\mathbf{n}]$ describes the set of forest on a $\mathbf{n}$ set such that there are k_i roots of color i. The set of roots R can be chosen in $\binom{\mathbf{n}}{\mathbf{k}}$ possible ways. Let E be the complemented set. Now the forest can be consider as non periodic $\vec{\mathbf{M}}$-enriched function from $(E, f|_E)$ to $(E, f|_E) \dot\cup [\mathbf{k}]$. But the number of such functions are $\left|\mathcal{N}^{\mathbf{k}}_{\vec{\mathbf{M}}}[\mathbf{n} - \mathbf{k}]\right|$, and the lemma follows. □

2.4 C-monoids

In the following sections we will study how Good's inversion formula specializes to Lagrange's inversion formula. To do so we need to define the plethystic composition of two colored species and of two formal power series. We begin defining c-monoids, and from this concept we can define the general plethysm.

Definition 2.4.1 *A* c-monoid $(\mathfrak{I}, \cdot, 1)$ *is a set* $\mathfrak{I}$ *with an associative binary operation* $\cdot$ *with identity element* $1 \in \mathfrak{I}$ *(that is a monoid) which satisfies the additional properties:*

(1) For all $i, j \in \mathfrak{I}$, *if* $i \cdot j = 1$ *then* $i = j = 1$. *(Indivisibility of the identity.)*

(2) For all $i, j, j' \in \mathfrak{I}$, *if* $i \cdot j = i \cdot j'$ *then* $j = j'$. *(Left cancellation law.)*

Example 2.4.1 *The c-monoid of natural integers under addition. It is denoted by* $(\mathfrak{I}, \cdot, 1) = (\mathbf{N}, +, 0)$. *Clearly this is a c-monoid.*

Example 2.4.2 *The c-monoid of positive integers under multiplication. Let* $(\mathfrak{I}, \cdot, 1) = (\mathbf{P}, \cdot, 1)$. *Observe that this c-monoid is isomorphic to an infinite countable direct product of the c-monoid in the previous example.*

Example 2.4.3 *The c-monoid of words over an alphabet. Let* A *be an alphabet. Denote* A^* *to be the set of all words with letters in* A. *Then* $(A^*, \cdot, \epsilon)$ *is a c-monoid, where* $\cdot$ *is concatenation and* ϵ *is the empty word.*

Definition 2.4.2 *If* $(\mathfrak{I}, \cdot, 1)$ *is a c-monoid we define the* divisibility *relation on* $\mathfrak{I}$ *as the following for* $i, j \in \mathfrak{I}$ *we have that* $i \leq j$ *if and only if there is* $k \in \mathfrak{I}$ *such that* $j = i \cdot k$.

Notice that $i \leq j$ implies that $k \cdot i \leq k \cdot j$.

Example 2.4.4 *In the first example above, the divisibility relation is the ordinary linear order on nonnegative integers. In the second example, the positive integer* i *is less than or equal to the positive integer* j, *if* i *divide* j. *In the last example the word* i *is less than or equal to the word* j, *if* i *is a prefix of* j.

Lemma 2.4.1 *Let* $(\mathfrak{I}, \cdot, 1)$ *be a c-monoid, and let* $\leq$ *be the induced divisibility relation* $\mathfrak{I}$. *Then:*

(1) $(\mathfrak{I}, \leq)$ *is a partial order on* $\mathfrak{I}$ *with minimal element* 1.

(2) For each $i \in \mathfrak{I}$, *the dual order ideal* $\mathfrak{I}_i = \{j \in \mathfrak{I} : i \leq j\}$ *is isomorphic to* $\mathfrak{I}$ *via the function* $\phi_i : \mathfrak{I} \longrightarrow \mathfrak{I}_i$ *given by* $\phi_i(j) = i \cdot j$. *In particular, if* $\mathfrak{I} \neq \{1\}$ *then* $\mathfrak{I}$ *is infinite.*

Definition 2.4.3 *A* finite factorization *monoid (FF c-monoid)* $(\mathfrak{I}, \cdot, 1)$ *is a monoid where every element has only a finite number of factorizations into different elements.*

Lemma 2.4.2 *Let* $(\mathfrak{I}, \cdot, 1)$ *be a c-monoid, and let* $\leq$ *be the induced divisibility relation* $\mathfrak{I}$. *Then the condition that* $(\mathfrak{I}, \cdot, 1)$ *is a finite factorization monoid is equivalent to that the partial order* $(\mathfrak{I}, \leq)$ *is locally finite.*

2.5 Plethystic composition of formal power series

Before we defined composition between a power series and a summable collection of power series. We also made a similar definition for colored species. If the index set $\mathfrak{I}$ has the structure of a c-monoid, we are now able to define the plethystic composition between two power series. Likewise we will defined the plethystic composition between two colored species.

Definition 2.5.1 *Define the* Verschiebung operator $\mathcal{V}_i$ *on a multi index by* $\mathcal{V}_i(\mathbf{e}_j) = \mathbf{e}_{i \cdot j}$ *and extend by linearity.*

Definition 2.5.2 *Define the* Frobenius operator $\mathcal{F}_i$ *on formal power series by*

$$\mathcal{F}_i\left(g\left((x_j)_{j \in \mathfrak{I}}\right)\right) = g\left((x_{i \cdot j})_{j \in \mathfrak{I}}\right),$$

where $i \in \mathfrak{I}$.

Notice that $\mathcal{F}_i$ is an injective algebra homomorphism on formal power series. Moreover we have the following fact

$$\mathcal{F}_i\left(\mathbf{x}^{\mathbf{n}}\right) = \mathbf{x}^{\mathcal{V}_i(\mathbf{n})}.$$

Definition 2.5.3 *Let $f(\mathbf{x})$ and $g(\mathbf{x})$ be formal power series, such that $g(\mathbf{x})$ has no constant coefficient. Define the* plethystic composition *$f * g$ as*

$$(f * g)(\mathbf{x}) = f\left((\mathcal{F}_i(g(\mathbf{x})))_{i\in\mathfrak{I}}\right).$$

Let $\mathbf{g}(\mathbf{x})$ be the collection of power series defined by $\mathbf{g}(\mathbf{x}) = (\mathcal{F}_i(g(\mathbf{x})))_{i\in\mathfrak{I}}$. Since $g(\mathbf{x})$ do not have a constant term, the collection $\mathbf{g}(\mathbf{x})$ is summable. Observe now that

$$(f * g)(\mathbf{x}) = (f \circ \mathbf{g})(\mathbf{x}).$$

This identity connects the two different compositions.

Example 2.5.1 *For the c-monoid of positive integers and multiplication, $(\mathbf{P}, \cdot, 1)$, the plethystic composition is the classical plethysm defined by Polya*

$$(f * g)(x_1, x_2, x_3, \ldots) = f\left(g(x_1, x_2, x_3, \ldots), g(x_2, x_4, x_6, \ldots), g(x_3, x_6, x_9, \ldots), \ldots\right).$$

Example 2.5.2 *For the c-monoid of natural integers and addition, $(\mathbf{N}, +, 0)$, the plethystic composition we obtain is the shift-plethysm*

$$(f * g)(x_0, x_1, x_2, \ldots) = f\left(g(x_0, x_1, x_2, \ldots), g(x_1, x_2, x_3, \ldots), g(x_2, x_3, x_4, \ldots), \ldots\right).$$

Example 2.5.3 *The c-monoid $(A^*, \cdot, \epsilon)$ leads to an infinite family of plethysm, one for each cardinality of the alphabet A.*

$$(f * g)\left((x_w)_{w\in A^*}\right) = f\left(\left(g\left((x_{w'w})_{w\in A^*}\right)\right)_{w'\in A^*}\right).$$

2.6 Plethystic composition of colored species

Define the Frobenius operator $\mathcal{F}_i$ on colored species by the following identity

$$(\mathcal{F}_i(M))\,[(E,f)] = \begin{cases} M[(E,g)] & \text{if } i\cdot g(a) = f(a) \text{ for all } a \in E \\ \emptyset & \text{otherwise} \end{cases}.$$

The Frobenius operator on colored species has the following combinatorial interpretation. Each structure in $\mathcal{F}_i(M)[(E,f)]$ is obtained from a unique structure in $M[(E,f)]$ by multiplying the colors of the underlying set (E,f) on the left by i.

We can rewrite this as

$$(\mathcal{F}_i(M))\,[(E,i\cdot f)] = M[(E,f)].$$

Directly we see that

$$\begin{aligned} \mathcal{F}_i(M+N) &= \mathcal{F}_i(M)+\mathcal{F}_i(N), \\ \mathcal{F}_i(M\cdot N) &= \mathcal{F}_i(M)\cdot\mathcal{F}_i(N). \end{aligned}$$

Moreover

A *plethystic partition* of a colored set (E,f) is a partition π of the set E, with a function $g:\pi\longrightarrow\mathfrak{I}$ such that for all $B\in\pi$ and for all $e\in B$

$$g(B)\leq f(e).$$

Let $\Pi_p[(E,f)]$ be the set of all plethystic partitions of (E,f).

Let M and N be two colored species, such that $N[\emptyset]=\emptyset$. Define the divided power $\gamma_{\mathbf{k}}(N)$ as

$$(\gamma_{\mathbf{k}}(N))\,[(E,f)] = \dot{\bigcup_{(\pi,g)\in\Pi_p[(E,f)],\ \mathrm{card}(\pi,g)=\mathbf{k}}}\ \prod_{B\in\pi}\mathcal{F}_{g(B)}(N)[(B,f|_B)].$$

Define the plethystic composition $M * N$ by

$$(M * N)[(E,f)] = \dot{\bigcup_{(\pi,g)\in\Pi_p[(E,f)]}} M[(\pi,g)] \times \prod_{B\in\pi} \mathcal{F}_{g(B)}(N)[(B,f|_B)].$$

Let $\vec{\mathbf{N}}$ be the collection of colored species defined by $\vec{\mathbf{N}} = (\mathcal{F}_i(N))_{i\in\mathfrak{I}}$. Since $N[\emptyset] = \emptyset$, the collection $\vec{\mathbf{N}}$ will be summable. It is now true that

$$M * N = M \circ \vec{\mathbf{N}}.$$

Moreover, it also true that

$$\gamma_{\mathbf{k}}(N) = \Gamma_{\mathbf{k}}(\vec{\mathbf{N}}).$$

Proposition 2.6.1

$$\begin{aligned}
card(\mathcal{F}_i(M);\mathbf{x}) &= \mathcal{F}_i(card(M;\mathbf{x})) \\
card((M * N);\mathbf{x}) &= (card(M;\mathbf{x}) * card(N;\mathbf{x}))(\mathbf{x}) \\
card(\gamma_{\mathbf{k}}(N);\mathbf{x}) &= \left(\frac{\mathbf{x}^{\mathbf{k}}}{\mathbf{k}!} * card(N;\mathbf{x})\right)(\mathbf{x}) \\
&= \frac{1}{\mathbf{k}!}\prod_{i\in\mathfrak{I}} \mathcal{F}_i(card(N;\mathbf{x}))^{k_i}
\end{aligned}$$

2.7 Colored species of permutations

Let S be the species of permutations on elements of color 1. Let S_0 be the species of nonempty permutations on elements of color 1. Similarly define L, L_0 as linear orders on elements of color 1. Their generating functions are

$$\begin{aligned}
\text{card}(S;\mathbf{x}) &= \text{card}(L;\mathbf{x}) = \tfrac{1}{1-x_1} \\
\text{card}(S_0;\mathbf{x}) &= \text{card}(L_0;\mathbf{x}) = \tfrac{1}{1-x_1} - 1 = \tfrac{x_1}{1-x_1}
\end{aligned}$$

From [J] comes the following equipotent identity

Lemma 2.7.1

$$S_0 \equiv X_1 \cdot S.$$

Proof: We have that

$$S_0 \equiv L_0 = X_1 \cdot L \equiv X_1 \cdot S.$$

□

Proof: It is easy to construct a natural bijection to see that following is true.

$$S_0^\bullet = X_1 \cdot S + X_1 \cdot S^\bullet.$$

This can be written as

$$\begin{aligned} S_0^\bullet &= X_1 \cdot S + X_1 \cdot S^\bullet \\ &= X_1^\bullet \cdot S + X_1 \cdot S^\bullet \\ &= (X_1 \cdot S)^\bullet \end{aligned}$$

from which the result follows. □

Definition 2.7.1 *Let I be a subset of $\mathfrak{I}$. Define the colored species S^I by*

$$S^I = \prod_{i \in I} \mathcal{F}_i(S).$$

Similarly, for a finite subset I of $\mathfrak{I}$, define the colored species S_0^I and X_I by

$$\begin{aligned} S_0^I &= \prod_{i \in I} \mathcal{F}_i(S_0), \\ X_I &= \prod_{i \in I} X_i. \end{aligned}$$

The set of structures $S^I[(E, f)]$ consists permutations on on each fiber $f^{-1}(i)$ where $i \in I$.

Recall that a permutation might be empty. Similarly the set of structures $S_0^I[(E,f)]$ consists *nonempty* permutations on on each fiber $f^{-1}(i)$ where $i \in I$. Observe then that the colored species $S_0^I \cdot S^{\mathfrak{I}-I}$ puts a permutation on each fiber $f^{-1}(i)$ for $i \in \mathfrak{I}$, and demands that there are elements with the colors of the set I.

Lemma 2.7.2 *The following equipotent identity is true*

$$S_0^I \cdot S^{\mathfrak{I}-I} \equiv X_I \cdot S^{\mathfrak{I}}.$$

Proof: Directly we have that

$$\begin{aligned} S_0^I \cdot S^{\mathfrak{I}-I} &= \prod_{i \in I} \mathcal{F}_i(S_0) \cdot \prod_{i \in \mathfrak{I}-I} \mathcal{F}_i(S) \\ &\equiv \prod_{i \in I} \mathcal{F}_i(X_1 \cdot S) \cdot \prod_{i \in \mathfrak{I}-I} \mathcal{F}_i(S) \\ &= \prod_{i \in I} X_i \cdot \prod_{i \in \mathfrak{I}} \mathcal{F}_i(S) \\ &= X_I \cdot S^{\mathfrak{I}}. \end{aligned}$$

□

2.8 Plethystic functions and plethystic trees

Let M be a colored species.

Definition 2.8.1 *An* M-enriched plethystic function/contraction/tree/forest *is a* $\vec{\mathbf{M}}$-*enriched function/contraction/tree/forest, where the collection* $\vec{\mathbf{M}}$ *is defined by*

$$\vec{\mathbf{M}} = (\mathcal{F}_i(M))_{i \in \mathfrak{I}}.$$

Observe that an M-enriched plethystic function ϕ from the colored set (E,f) to the colored set (F,g) fulfill $f(a) \geq g(\phi(a))$ for all $a \in E$.

Notice that if b is a node in a plethystic tree on a colored set (E,f), and if a is the father of the node b, then $f(a) \leq f(b)$. This is the same definition of plethystic trees as in **[M-N]**.

We have three important colored species to define.

1. Let A_M be the colored species of M-enriched plethystic trees with the root of color 1. That is $A_M = A^{(1)}_{\vec{\mathbf{M}}}$. Observe that $(\mathcal{F}_i(A_M))_{i\in\mathbf{J}} = \vec{\mathbf{A}}_{\vec{\mathbf{M}}}$.

2. Let End_M be the colored species of M-enriched plethystic endofunctions. That is $End_M = End_{\vec{\mathbf{M}}}$.

3. Let $\mathcal{N}^{\mathbf{k}}_M$ be the colored species of non periodic M-enriched plethystic functions. That is $\mathcal{N}^{\mathbf{k}}_M = \mathcal{N}^{\mathbf{k}}_{\vec{\mathbf{M}}}$.

Thus we can rewrite Lemma 2.3.1 to the following.

Lemma 2.8.1 *Let (E,f) and (F,g) be two colored sets. The set of all M-enriched plethystic functions from (E,f) to (F,g) is described by*

$$\left(\prod_{b\in F} \mathcal{F}_{g(b)}(M)\right)[(E,f)].$$

Example 2.8.1 *Let the underlying c-monoid be positive integers and multiplication. A M-enriched plethystic tree, where M is a colored species, is an enriched plethystic tree as defined in [C1].*

Directly from Proposition 2.3.2 we get the following lemma.

Lemma 2.8.2 *The colored species A_M fulfills the following functional equation*

$$A_M = X_1 \cdot (M * A_M).$$

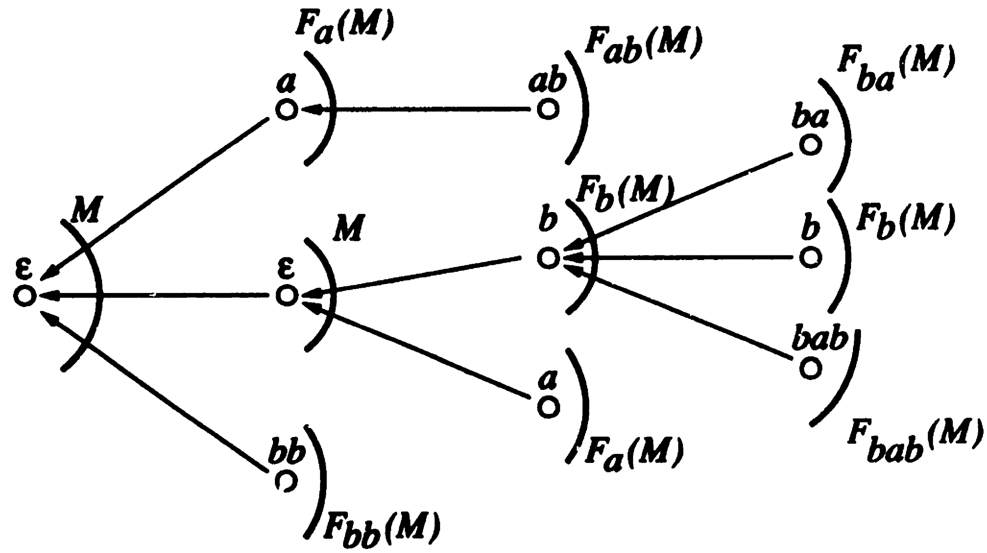


Figure 2-3: An example of a M-enriched plethystic tree over the c-monoid $(\{a, b\}^*, \cdot, \epsilon)$.

Lemma 2.3.3 and Lemma 2.3.4 translates into

Lemma 2.8.3 *The colored species of M-enriched contractions with the attracting point of color 1 is described by*

$$X_1 \cdot (M'^{(1)} * A_M).$$

Lemma 2.8.4 *The colored species $M^{\bullet(1)} = X_1 \cdot M'^{(1)}$is naturally isomorphic to M-enriched contractions ϕ, which has depth 1, attracting vertex of color 1 and no structure put on $\phi^{-1}(b)$ if b is not the attracting vertex.*

Let ϕ be a plethystic function from the colored set $(E, f|_E)$ to to the colored set $(R \dot{\cup} E, f)$. Let a be a periodic element of color i. Since $f(a) \geq f(\phi(a)) \geq f(\phi^2(a)) \geq \cdots \geq f(\phi^m(a))$ we know that all nodes in the same cycle as a has the color i.

Lemma 2.8.5 *The colored species End_M is naturally isomorphic to*

$$S^{\circ} * \left(X_1 \cdot (M'^{(1)} * A_M)\right).$$

Lemma 2.8.6 *Let $(R \dot\cup E, f)$ be a* **n** *set and assume that $(R, f|_R)$ is a* **k** *set. The set of all M-enriched plethystic functions from $(E, f|_E)$ to $(R \dot\cup E, f)$ is described by*

$$\left(\mathcal{N}_M^{\mathbf{k}} \cdot End_M\right)[(E, f|_E)].$$

2.9 Lagrange inversion formula

Let $R \dot\cup E$ be a colored **n** set and R a colored **k** set. Let $J = \{i : n_i \neq 0\}$, which is a finite set. Moreover, in this section, let $M' = M'^{(1)}$ and $M^\bullet = M^{\bullet(1)}$.

Proposition 2.9.1 *Let I be a subset of J. Then there is a natural bijection between the set of M-enriched plethystic functions from E to $R \dot\cup E$ such that there exists a cyclic point of color i for all $i \in I$, and the set*

$$\left(\prod_{i\in I}\mathcal{F}_i(M^\bullet M^{n_i-1}) \cdot \prod_{i\in J-I}\mathcal{F}_i(M^{n_i})\right)[(E, f)].$$

Proof: Observe that

$$\prod_{i\in I}\mathcal{F}_i(M^\bullet M^{n_i-1}) \cdot \prod_{i\in J-I}\mathcal{F}_i(M^{n_i}) = \prod_{i\in I}\mathcal{F}_i(M^\bullet) \cdot \prod_{i\in I}\mathcal{F}_i(M^{n_i-1}) \cdot \prod_{i\in J-I}\mathcal{F}_i(M^{n_i}).$$

By Lemma 2.8.4 $\mathcal{F}_i(M^\bullet)$ chooses an M-enriched contraction of depth 1 with the attracting vertex of color i. Thus

$$\prod_{i\in I}\mathcal{F}_i(M^\bullet)$$

chooses for each $i \in I$ an attracting vertex of color i, and to each of them a contraction of depth 1. Let C be the set of attracting vertices, and let E_1 be the underlying set on which these contractions are built. Hence we have chosen an M-enriched plethystic function from the colored set $(E_1, f|_{E_1})$ to the colored set $(C, f|_C)$.

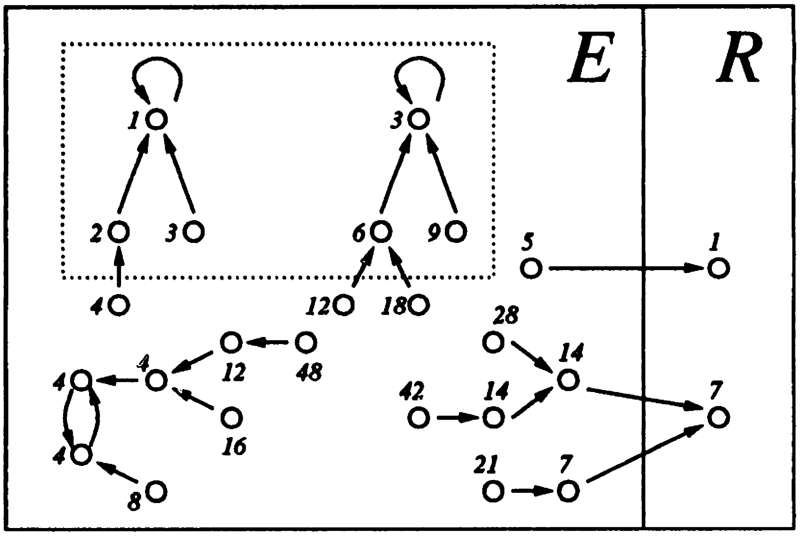


Figure 2-4: An example of a plethystic function with marked contractions of depth 1.

The cardinality of $(R \dot{\cup} E) - C$ is $\mathbf{n} - \mathbf{e}_I$. Hence the colored species

$$\prod_{i\in I} \mathcal{F}_i(M^{n_i-1}) \cdot \prod_{i\in \mathfrak{I}-I} \mathcal{F}_i(M^{n_i})$$

chooses an M-enriched plethystic function from the colored set $(E_2, f|_{E_2})$ to the colored set $((R \dot{\cup} E) - C, f|_{(R\dot{\cup}E)-C})$, where $E_2 = E - E_1$.

Recall that $E = E_1 + E_2$, thus by joining these two plethystic functions we get an M-enriched plethystic function from the colored set (E, f) to the colored set $(R \dot{\cup} E, f)$. Moreover we know that this function has an attracting vertex of color i for each $i \in I$.

The above set of structures can be written as

$$\left(\mathcal{N}_M^{\mathbf{k}} \cdot \prod_{i\in I} \mathcal{F}_i(X_1 \cdot (M' * A_M))) \cdot End_M\right)[(E, f|_E)].$$

The first factor describes the structure of elements that image of repeated applications of the function will be in R. The second part is all the contractions which have the attracting vertices of the given colors. The third part is the structure on those elements that will be in a cycle after repeated applications of the function.

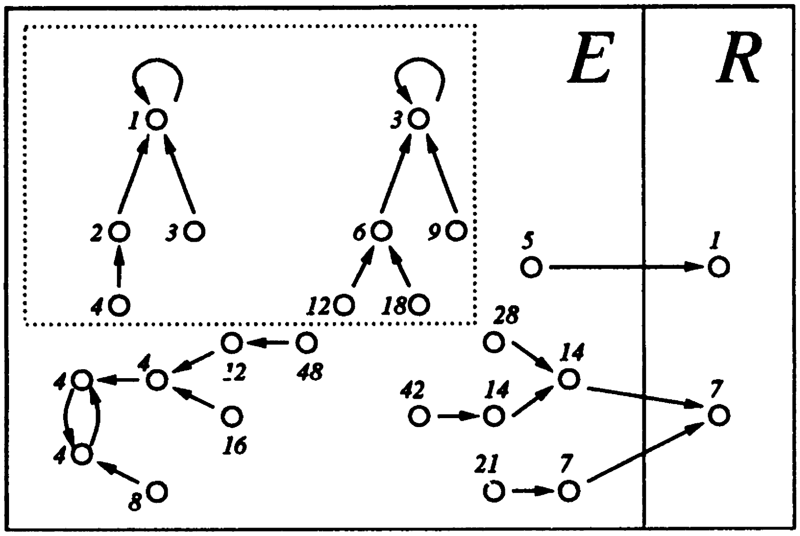


Figure 2-5: An example of a plethystic function with marked contractions.

The above colored species can be written as

$$
\begin{aligned}
& \mathcal{N}_M^{\mathbf{k}} \cdot \prod_{i \in I} \mathcal{F}_i(X_1 \cdot (M' * A_M)) \cdot End_M \\
= \; & \mathcal{N}_M^{\mathbf{k}} \cdot (X_I * (X_1 \cdot (M' * A_M))) \cdot \left(S^J * (X_1 \cdot (M' * A_M))\right) \\
= \; & \mathcal{N}_M^{\mathbf{k}} \cdot \left((X_I \cdot S^J) * (X_1 \cdot (M' * A_M))\right).
\end{aligned}
$$

Since

$$X_I \cdot S^J \equiv S_0^I \cdot S^{J-I},$$

we conclude that the set of structures above is equipotent to

$$\left(\mathcal{N}_M^{\mathbf{k}} \cdot \left((S_0^I \cdot S^{J-I}) * (X_1 \cdot (M' * A_M))\right)\right)[(E, f|_E)].$$

But this is the structure of enriched plethystic functions such that there will be at least one periodic element of color i for each $i \in I$. This concludes the proof of the proposition. □

Theorem 3 (Lagrange inversion formula, species version) *Let M be a colored species and A_M be the colored species of M-enriched plethystic trees. Assume that* $\mathbf{n} \geq \mathbf{k}$, *and let $J = \{i \in \mathfrak{I} : n_i \neq 0\}$. Then*

$$|(\gamma_{\mathbf{k}}(A_M))[\mathbf{n}]| \equiv \binom{\mathbf{n}}{\mathbf{k}} \cdot \left| \left(\prod_{i \in J} \mathcal{F}_i \left(M^{n_i} - M^{\bullet} \cdot M^{n_i - 1} \right) \right) [\mathbf{n} - \mathbf{k}] \right|,$$

*written by help of abuse of notation. (The notation could be made strict by using Möbius species [***M-Y***].)*

Proof: By Proposition 2.9.1 we know that

$$\left| \left(\prod_{i \in I} \mathcal{F}_i(M^{\bullet} \cdot M^{n_i - 1}) \cdot \prod_{i \in J - I} \mathcal{F}_i(M^{n_i}) \right) [(E, f|_E)] \right|,$$

counts the plethystic functions which has periodic elements of color i for each $i \in I$. But we would like to count plethystic functions that do not have any periodic elements at all. By inclusion and exclusion the number of such plethystic functions is

$$\sum_{I \subseteq J} (-1)^{|I|} \left| \left(\prod_{i \in I} \mathcal{F}_i(M^{\bullet} \cdot M^{n_i - 1}) \cdot \prod_{i \in J - I} \mathcal{F}_i(M^{n_i}) \right) [(E, f|_E)] \right|.$$

By abuse of notation we can write the above

$$\begin{aligned} & \sum_{I \subseteq J} (-1)^{|I|} \left| \left(\prod_{i \in I} \mathcal{F}_i(M^{\bullet} M^{n_i - 1}) \prod_{i \in J - I} \mathcal{F}_i(M^{n_i}) \right) [(E, f|_E)] \right| \\ = & \left| \sum_{I \subseteq J} (-1)^{|I|} \left(\prod_{i \in I} \mathcal{F}_i(M^{\bullet} M^{n_i - 1}) \prod_{i \in J - I} \mathcal{F}_i(M^{n_i}) \right) [(E, f|_E)] \right| \\ = & \left| \left(\prod_{i \in J} \mathcal{F}_i \left(M^{n_i} - M^{\bullet} M^{n_i - 1} \right) \right) [(E, f|_E)] \right|. \end{aligned}$$

This is the number of non periodic M-enriched plethystic functions from the set $(E, f|_E)$ to the set $(E, f|_E) \dot{\cup} \mathbf{k}$. Thus we have the identity

$$\left|\mathcal{N}_M^{\mathbf{k}}[(E, f|_E)]\right| = \left|\left(\prod_{i \in J} \mathcal{F}_i \left(M^{n_i} - M^{\bullet} M^{n_i - 1}\right)\right)[(E, f|_E)]\right|.$$

Now by Lemma 2.3.6 and the above identity Lagrange inversion formula follows. □

By equating the coefficients of the equation $f(\mathbf{x}) = x_1 \cdot (G * f)(\mathbf{x})$ a system of recurrences occur for the coefficients of $f(\mathbf{x})$, and this system is easily seen to have a unique solution. Hence the equation $f(\mathbf{x}) = x_1 \cdot (G * f)(\mathbf{x})$ uniquely determines the power series $f(\mathbf{x})$.

Theorem 4 (Lagrange inversion formula) *Let $f(\mathbf{x})$ and $G(\mathbf{x})$ be power series in the variables $(x_i)_{i \in \mathfrak{I}}$ such that*

$$f(\mathbf{x}) = x_1 \cdot (G * f)(\mathbf{x}).$$

Assume that $\mathbf{n} \geq \mathbf{k}$, and let $J = \{i \in \mathfrak{I} \, : \, n_i \neq 0\}$. Then

$$[\mathbf{x}^{\mathbf{n}}]\left(\mathbf{x}^{\mathbf{k}} * f\right)(\mathbf{x}) = \left[\mathbf{x}^{\mathbf{n}-\mathbf{k}}\right] \prod_{i \in J} H_i(\mathbf{x}).$$

where

$$H_i(\mathbf{x}) = \mathcal{F}_i \left(G(\mathbf{x})^{n_i} - x_1 \cdot \frac{\partial G(\mathbf{x})}{\partial x_1} \cdot G(\mathbf{x})^{n_i - 1}\right).$$

Proof: Let M be a colored species and $G(\mathbf{x})$ its generating function. That is $G(\mathbf{x}) = \mathrm{card}(M; \mathbf{x})$. Let $\overline{f}(\mathbf{x}) = \mathrm{card}(A_M; \mathbf{x})$ where A_M is the colored species of M-enriched plethystic trees. Since $A_M = X_1 \cdot (M * A_M)$, we get

$$\begin{aligned} \overline{f}(\mathbf{x}) &= \mathrm{card}(A_M; \mathbf{x}) \\ &= \mathrm{card}\left(X_1 \cdot (M * A_M); \mathbf{x}\right) \end{aligned}$$

$$= x_1 \cdot (\text{card}(M;\mathbf{x}) * \text{card}(A_M;\mathbf{x}))$$
$$= x_1 \cdot \left(G(\mathbf{x}) * \overline{f}(\mathbf{x})\right)$$

But $f(\mathbf{x})$ is uniquely determined by the above equation. Hence we have that $f(\mathbf{x}) = \overline{f}(\mathbf{x}) = \text{card}(A_M;\mathbf{x})$.

Now the left hand side of the species version of Lagrange inversion formula is equal to

$$\begin{aligned} |(\gamma_{\mathbf{k}}(A_M))[\mathbf{n}]| &= \left[\frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!}\right] \text{card}\,(\gamma_{\mathbf{k}}(A_M);\mathbf{x}) \\ &= \left[\frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!}\right] \left(\frac{\mathbf{x}^{\mathbf{k}}}{\mathbf{k}!} * f\right)(\mathbf{x}) \\ &= \frac{\mathbf{n}!}{\mathbf{k}!} [\mathbf{x}^{\mathbf{n}}] \left(\mathbf{x}^{\mathbf{k}} * f\right)(\mathbf{x}). \end{aligned}$$

And the right hand side

$$\begin{aligned} & \binom{\mathbf{n}}{\mathbf{k}} \cdot \left| \left(\prod_{i \in J} \mathcal{F}_i \left(M^{n_i} - M^{\bullet} \cdot M^{n_i - 1} \right) \right) [\mathbf{n} - \mathbf{k}] \right| \\ = & \binom{\mathbf{n}}{\mathbf{k}} \cdot \left[\frac{\mathbf{x}^{\mathbf{n}-\mathbf{k}}}{(\mathbf{n}-\mathbf{k})!}\right] \text{card}\left(\prod_{i \in J} \mathcal{F}_i \left(M^{n_i} - M^{\bullet} \cdot M^{n_i - 1} \right); \mathbf{x} \right) \\ = & \frac{\mathbf{n}!}{\mathbf{k}!} \cdot \left[\mathbf{x}^{\mathbf{n}-\mathbf{k}}\right] \prod_{i \in J} \mathcal{F}_i \left(G(\mathbf{x})^{n_i} - x_1 G'(\mathbf{x}) \cdot G(\mathbf{x})^{n_i - 1} \right) \\ = & \frac{\mathbf{n}!}{\mathbf{k}!} \cdot \left[\mathbf{x}^{\mathbf{n}-\mathbf{k}}\right] \prod_{i \in J} H_i(\mathbf{x}) \end{aligned}$$

Thus we have proven the theorem for formal power series $G(\mathbf{x})$ such that $\left[\frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!}\right] G(\mathbf{x})$ is a nonnegative integer for all multi indices $\mathbf{n}$. By the principle of extension of algebraic identities the theorem follows for all $G(\mathbf{x})$. □

We would like to point out that one can also prove the above theorem by directly enumerating plethystic trees. This method of proving Theorem 4 will be presented in Section 2.15 and Section 2.16.

2.10 Good's inversion formula

Let $R \dot{\cup} E$ be a colored **n** set and R a colored **k** set. Let $J = \{i : n_i \neq 0\}$.

Definition 2.10.1 *Let I be a finite subset of $\mathfrak{I}$, and let π be a permutation on I. Define the colored species $P^{\pi}_{\vec{\mathbf{M}}}$ by $P^{\pi}_{\vec{\mathbf{M}}}[(E,f)]$ is all $\vec{\mathbf{M}}$-enriched colored functions ϕ, such that there is a subset $\{e_i\}_{i\in I}$ of E such that $f(e_i) = i$, $\phi(e_i) = e_{\pi(i)}$ and for all $b \in E$ there exists a positive integer k such that $\phi^k(b) \in \{e_i\}_{i\in I}$. The elements e_i are called the attracting vertices.*

If I only consist of one element, that is $I = \{i\}$, then $P^{\pi}_{\vec{\mathbf{M}}}$ is just a colored contraction, with the attracting vertex of color i. If I is the empty set, then $P^{\emptyset}_{\vec{\mathbf{M}}} = 1$.

Lemma 2.10.1 *The colored species*

$$\prod_{i\in I} M_i^{\bullet(\pi^{-1}(i))}$$

is naturally isomorphic to $\vec{\mathbf{M}}$-enriched endofunction ψ, such that there are elements e_i such that $f(e_i) = i$ and $\psi(e_i) = e_{\pi(i)}$ for all $i \in I$, and for all elements b we have that $\psi(b) \in \{e_i\}_{i\in I}$. The last condition makes the function ψ to be of depth 1. Moreover, we do not put any M structure on the fiber $\psi^{-1}(b)$, if $b \notin \{e_i\}_{i\in I}$.

Proof: By Lemma 2.3.4 $M_i^{\bullet(\pi^{-1}(i))}$ chooses a M_i-enriched contraction of depth 1 and the attracting vertex of color $\pi^{-1}(i)$. Thus

$$\prod_{i\in I} M_i^{\bullet(\pi^{-1}(i))}$$

chooses a function ϕ such for each $i \in I$ a fixed point of color $\pi^{-1}(i)$, and to each of them a contraction of depth 1. Now define $\omega(e_{\pi^{-1}(i)}) = e_i$. Define a new function ψ by $\psi(b) = \omega(\phi(b))$. Notice that ψ is a colored function. Enrich the colored set $\psi^{-1}(b)$ with the structure $\phi^{-1}(\omega^{-1}(b))$. This completes the bijection. □

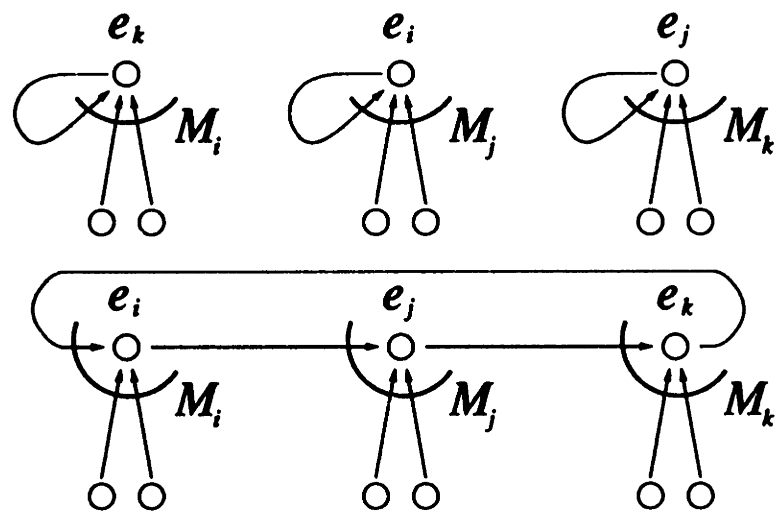


Figure 2-6: The construction for the permutation $\pi(i) = j$, $\pi(j) = k$ and $\pi(k) = i$.

Proposition 2.10.2 *Let I be a subset of J. Let π be a permutation on the set I. Then there is a natural bijection between the set*

$$\left(\prod_{i\in I} M_i^{\bullet(\pi^{-1}(i))} M_i^{n_i - 1} \cdot \prod_{i \in \mathfrak{I} - I} M_i^{n_i}\right)[(E, f|_E)].$$

and the set

$$\left(\mathcal{N}_{\vec{\mathbf{M}}}^{\mathbf{k}} \cdot P_{\vec{\mathbf{M}}}^{\pi} \cdot End_{\vec{\mathbf{M}}}\right)[(E, f|_E)].$$

Proof: We can write

$$\prod_{i\in I} M_i^{\bullet(\pi^{-1}(i))} M_i^{n_i - 1} \cdot \prod_{i \in \mathfrak{I} - I} M_i^{n_i} = \prod_{i\in I} M_i^{\bullet(\pi^{-1}(i))} \quad \cdot \quad \vec{\mathbf{M}}^{\mathbf{n} - \mathbf{e}_I}.$$

By Lemma 2.10.1 the first term in the above product chooses a $\vec{\mathbf{M}}$-enriched endofunction such that there are elements e_i so that $f(e_i) = i$ and $\psi(e_i) = e_{\pi(i)}$ for all $i \in I$, and for all elements b we have that $\phi(b) \in \{e_i\}_{i\in I}$. Let $C = \{e_i : i \in I\}$, and we call these elements the attracting vertices. Let E_1 be the underlying set of elements that this structure is built on.

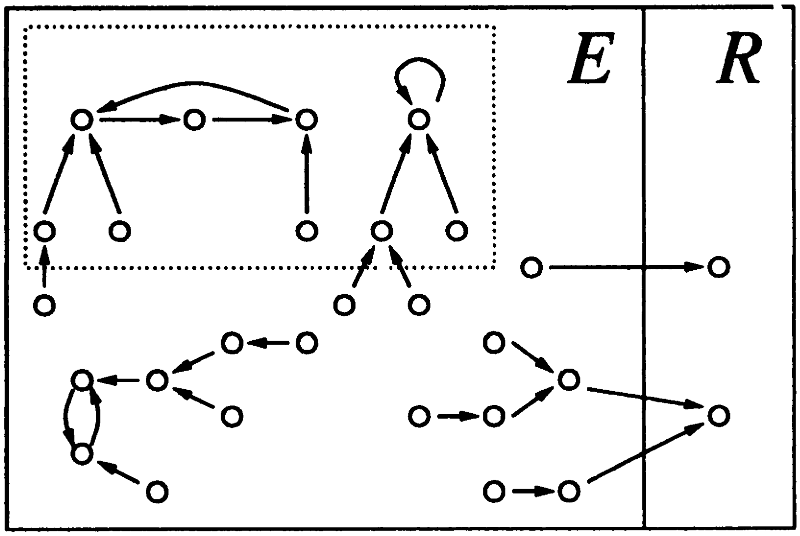


Figure 2-7: An example with a colored function with the product $\prod_{i \in I} M_i^{\bullet(\pi^{-1}(i))}$ marked.

The set $(R \dot{\cup} E) - C$ has the cardinality $\mathbf{n} - \mathbf{e}_I$. Hence the colored species $\vec{\mathbf{M}}^{\mathbf{n}-\mathbf{e}_I}$ chooses an $\vec{\mathbf{M}}$-enriched function from the colored set $(E_2, f|_{E_2})$ to the colored set $((R \dot{\cup} E) - C, f|_{(R\dot{\cup}E)-C})$.

Thus by taking the union between these two functions we get an $\vec{\mathbf{M}}$-enriched colored function ψ from the colored set $(E, f|_E)$ to the colored set $(E \dot{\cup} R, f)$ with attracting vertex e_i of color i for each $i \in I$, such that $\psi(e_i) = e_{\pi(i)}$ for all $i \in I$.

Now consider the right hand side of the proposition.

$$\mathcal{N}^{\mathbf{k}}_{\vec{\mathbf{M}}} \cdot P^{\pi}_{\vec{\mathbf{M}}} \cdot End_{\vec{\mathbf{M}}}.$$

The first term $\mathcal{N}^{\mathbf{k}}_{\vec{\mathbf{M}}}$ describes the part of the structure of the function, whose underlying elements will reach the colored set R after repeatedly application of the function. The second term $P^{\pi}_{\vec{\mathbf{M}}}$ describes the part of the structure, whose underlying elements will reach the attracting vertices. Finally, the third term $End_{\vec{\mathbf{M}}}$ describes the part of the function, whose elements will reach cycles. Hence this is the same set of structures as above and thus the result follows. □

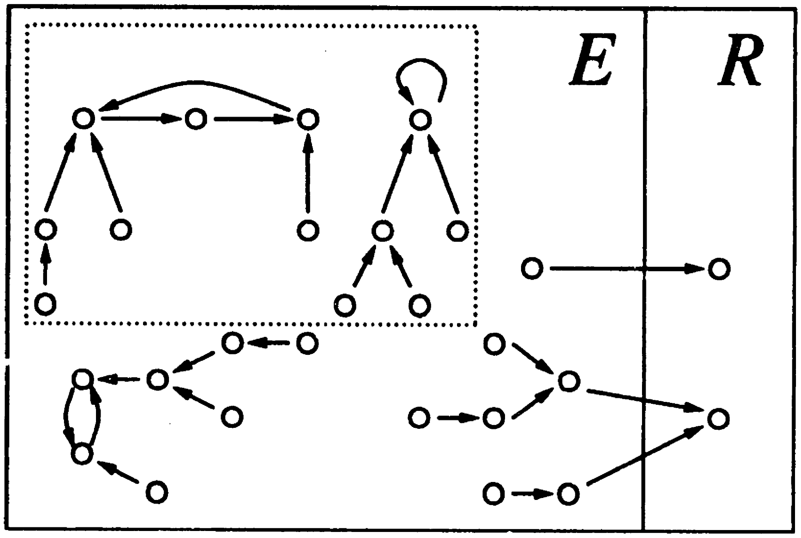


Figure 2-8: An example with a colored function with the colored species $P^{\pi}_{\vec{\mathbf{M}}}$ marked.

Define $c(\pi)$, where π is a permutation, as the number of cycles in π.

Proposition 2.10.3 *Let $j \in \mathfrak{J}$. Then we have that*

$$\sum_{I \subseteq J} \sum_{\pi \in S[I]} (-1)^{c(\pi)} \left(P^{\pi}_{\vec{\mathbf{M}}} \cdot End_{\vec{\mathbf{M}}}\right)^{\bullet(j)} = 0.$$

Proof: Let T be the set

$$T = \dot{\bigcup_{I \subseteq J}} \dot{\bigcup_{\pi \in S[I]}} \left(P^{\pi}_{\vec{\mathbf{M}}} \cdot End_{\vec{\mathbf{M}}}\right)^{\bullet(j)} [(E, f)].$$

An element of T is written as (ϕ, π, K, c), where ϕ is the $\vec{\mathbf{M}}$-enriched function, π is the marked permutation of colors, K is elements which have the colors the permutation π acts upon, and c is the marked element of color I. Define the sign of (ϕ, π, K, c) by $\text{sign}((\phi, \pi, K, c)) = (-1)^{c(\pi)}$.

We will construct an involution χ on the set T, that is sign reversing. That is, $\chi^2((\phi, \pi, K, c)) = (\phi, \pi, K, c)$, and $\text{sign}(\chi((\phi, \pi, K, c))) = -\text{sign}((\phi, \pi, K, c))$. From this will the proposition follow since χ defines a bijection between the two sets

$$\{t \in T : \text{sign}(t) = 1\} \quad \text{and} \quad \{t \in T : \text{sign}(t) = -1\}.$$

Given $(\phi, \pi, K, c) \in T$. We will now start constructing the involution $\chi((\phi, \pi, K, c)) = (\psi, \sigma, L, c)$.

Define the sequence $x_0, x_1, x_2, \ldots$ by the following rules. Let m be the smallest nonnegative integer such that $\phi^m(c)$ is a periodic element. Let $x_0 = \phi^m(c)$. Define $x_k = \phi^k(x_0)$ for $k \geq 1$.

Let n be the smallest nonnegative integer such that $f(x_n) \in \{f(x_0), \ldots, f(x_{n-1})\} \dot\cup f(K)$. Observe that such a integer exists since x_0 is periodic element. Four cases can occur

(i) $n = 0$ and $x_0 \in K$. Then remove the cycle $\{x_0, x_1, \ldots\}$ from the marked permutation. That is $L = K - \{x_0, x_1, \ldots\}$. Restrict also π to the set $f(L) = f(K) - \{f(x_0), f(x_1), \ldots\}$, to obtain σ. But let $\psi = \phi$, and let them have the same marked element c.

(ii) $x_n = x_0$. Then add the cycle $\{x_0, x_1, \ldots, x_{n-1}\}$ to the marked permutation. That is $L = K \dot\cup \{x_0, x_1, \ldots, x_{n-1}\}$. Extend also π to the set $f(L) = f(K) \dot\cup \{f(x_0), f(x_1), \ldots, f(x_{n-1})\}$, to obtain σ. But let $\psi = \phi$, and let them have the same marked element c.

(iii) $f(x_n) \in f(K)$. Assume that $f(x_n) = f(z_0)$ for some element z_0 such that z_0 is in the permutation $\hat{\pi}$. Assume that the length of the cycle the element z_0 is in is k. Let $z_i = \hat{\pi}^i(z_0)$. Thus $z_k = z_0$ and $\{z_1, \ldots, z_k\}$ are the elements of the cycle. Define ψ by

$$\psi(b) = \begin{cases} x_{n+1} & \text{if } b = z_k \\ z_1 & \text{if } b = x_n \\ \phi(b) & \text{if } b \neq z_k, b \neq x_n \end{cases}$$

As in the previous case, the colored sets $\phi^{-1}(b)$ and $\psi^{-1}(b)$ have the same cardi-

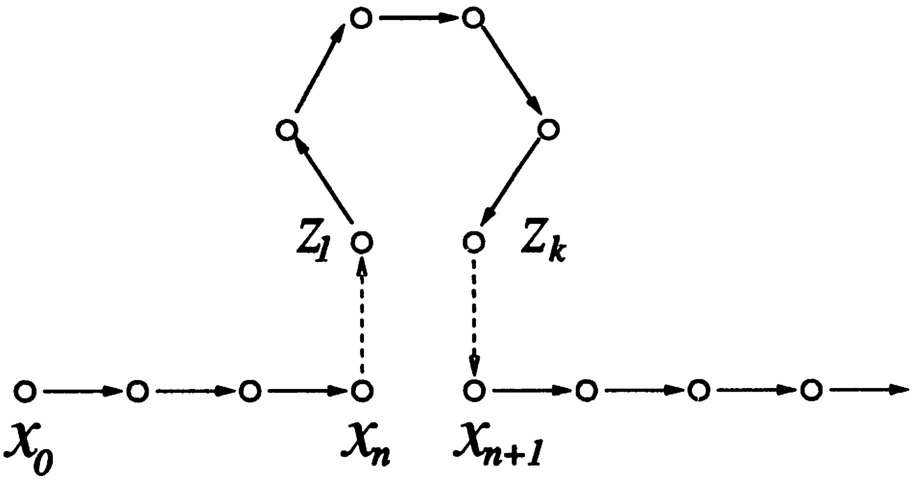


Figure 2-9: The construction of ψ in case (iii).

nality for all $b \in E$ and thus the functions ϕ and ψ have the same enrichment.

But remove from the marked permutation the cycle $\{z_1, \ldots, z_k\}$. That is $L = K - \{z_1, \ldots, z_k\}$. Restrict π to the set $f(L) = f(K) - \{f(z_1), \ldots, f(z_k)\}$, to obtain the permutation σ. Let still c be the marked periodic element of ϕ.

(iv) $f(x_n) \in \{f(x_0), \ldots, f(x_{n-1})\}$. Assume that $f(x_n) = f(x_m)$ for some integer m such that $0 \leq m \leq n-1$. Define ψ by

$$\psi(b) = \begin{cases} x_{m+1} & \text{if } b = x_n \\ x_{n+1} & \text{if } b = x_m \\ \phi(b) & \text{if } b \neq x_m, b \neq x_n \end{cases}$$

Observe that the colored sets $\phi^{-1}(b)$ and $\psi^{-1}(b)$ have the same cardinality for all $b \in E$. Thus the functions ϕ and ψ have the same enrichment of the colored species $\vec{\mathbf{M}}$.

But extend the marked permutation with the cycle $\{x_{m+1}, \ldots, x_n\}$. Hence we can write is $L = K \dot{\cup} \{x_{m+1}, \ldots, x_n\}$. Extend also π to the set $f(L) = f(K) \dot{\cup} \{f(x_{m+1}), \ldots, f(x_n)\}$, to obtain σ. Let still c be the marked periodic element of ϕ.

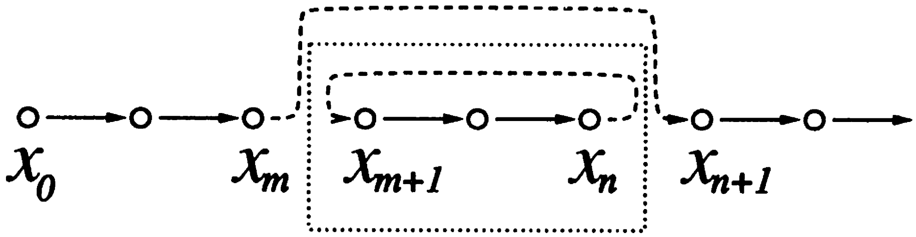


Figure 2-10: The construction of ψ in case (iv).

Clearly $\mathrm{sign}((\phi,\pi,K,c)) = -\mathrm{sign}((\psi,\sigma,L,c))$, since the difference in the number of cycles of π and σ is one.

Left to show is that χ is an involution. Now if we apply χ twice observe that we are going to add and remove the same cycle from the partial permutation π. This fact checks in all four cases above. Observe that cases (i) and (ii) are dual, and that the cases (iii) and (iv) are dual. Thus χ is an involution, and the proposition follows. □

Lemma 2.10.4

$$\sum_{I\subseteq J}\sum_{\pi\in S[I]}(-1)^{c(\pi)}P^{\pi}_{\tilde{\mathbf{M}}}\cdot End_{\tilde{\mathbf{M}}} = 1.$$

Proof: Let

$$A(\mathbf{x}) = \mathrm{card}\left(\sum_{I\subseteq J}\sum_{\pi\in S[I]}(-1)^{c(\pi)}P^{\pi}_{\tilde{\mathbf{M}}}\cdot End_{\tilde{\mathbf{M}}};\mathbf{x}\right).$$

Observe for all $j\in\mathfrak{J}$ that

$$x_j\cdot\frac{\partial}{\partial x_j}A(\mathbf{x}) = \mathrm{card}\left(\sum_{I\subseteq J}\sum_{\pi\in S[I]}(-1)^{c(\pi)}\left(P^{\pi}_{\tilde{\mathbf{M}}}\cdot End_{\tilde{\mathbf{M}}}\right)^{\bullet(j)};\mathbf{x}\right) = 0.$$

Hence for all $j\in\mathfrak{J}$ we have that $\frac{\partial}{\partial x_j}A(\mathbf{x}) = 0$. Thus solving for $A(\mathbf{x})$ by integration, we observe that $A(\mathbf{x})$ will be a constant. To find this constant observe that the colored species $P_{\emptyset}$ will contain the empty function. Thus the constant is equal 1 and the result follows. □

Lemma 2.10.5 *Let $(a_{i,j})_{i,j\in J}$ be a matrix. Then*

$$\det(\delta_{i,j}\cdot b_i - a_{i,j})_{i,j\in J} \;=\; \sum_{I\subseteq J}\sum_{\pi\in S[I]}(-1)^{c(\pi)}\prod_{i\notin I} b_i \prod_{i\in I} a_{i,\pi^{-1}(i)}.$$

Proof:

$$\begin{aligned}
\det(\delta_{i,j}\cdot b_i - a_{i,j})_{i,j\in J} &= \sum_{I\subseteq J}(-1)^{|I|}\prod_{i\notin I} b_i \det(a_{i,j})_{i,j\in I}\\
&= \sum_{I\subseteq J}(-1)^{|I|}\sum_{\pi\in S[I]}\operatorname{sign}(\pi)\prod_{i\notin I} b_i \prod_{i\in I} a_{i,\pi^{-1}(i)}\\
&= \sum_{I\subseteq J}\sum_{\pi\in S[I]}(-1)^{c(\pi)}\prod_{i\notin I} b_i \prod_{i\in I} a_{i,\pi^{-1}(i)}.
\end{aligned}$$

□

Theorem 5 (Good's inversion formula, the species version) *Let $\vec{\mathbf{M}}$ be a collection of colored species. Let $A^{(i)}_{\vec{\mathbf{M}}}$ be the colored species of $\vec{\mathbf{M}}$-enriched trees, with the root of color i. Let $J = \{i \in \mathfrak{I} : n_i \neq 0\}$ and assume that $\mathbf{n} \geq \mathbf{k}$. Then we have that*

$$\left|\left(\Gamma_{\mathbf{k}}(\vec{\mathbf{A}}_{\vec{\mathbf{M}}})\right)[\mathbf{n}]\right| \equiv \binom{\mathbf{n}}{\mathbf{k}}\cdot\left|\left(\det\left(\delta_{i,j}M_i^{n_i} - M_i^{\bullet(j)}\cdot M_i^{n_i-1}\right)_{i,j\in J}\right)[\mathbf{n}-\mathbf{k}]\right|,$$

*written by help of abuse of notation. (The notation could be made strict by using Möbius species [**M-Y**].)*

Proof: By Lemma 2.10.5, Proposition 2.10.2, and Lemma 2.10.4 it follows that

$$\begin{aligned}
&\left|\left(\det\left(\delta_{i,j}M_i^{n_i} - M_i^{\bullet(j)}\cdot M_i^{n_i-1}\right)_{i,j\in J}\right)[\mathbf{n}-\mathbf{k}]\right|\\
=\;&\left|\left(\sum_{I\subseteq J}\sum_{\pi\in S[I]}(-1)^{c(\pi)}\prod_{i\notin I} M_i^{n_i}\prod_{i\in I} M_i^{\bullet(\pi^{-1}(i))}M_i^{n_i-1}\right)[\mathbf{n}-\mathbf{k}]\right|\\
=\;&\left|\left(\sum_{I\subseteq J}\sum_{\pi\in S[I]}(-1)^{c(\pi)}\mathcal{N}^{\mathbf{k}}_{\vec{\mathbf{M}}}\cdot P_\pi\cdot End_{\vec{\mathbf{M}}}\right)[\mathbf{n}-\mathbf{k}]\right|
\end{aligned}$$

$$= \left| \left(\mathcal{N}_{\vec{\mathbf{M}}}^{\mathbf{k}} \cdot \sum_{I \subseteq J} \sum_{\pi \in S[I]} (-1)^{c(\pi)} P_\pi \cdot End_{\vec{\mathbf{M}}} \right) [\mathbf{n} - \mathbf{k}] \right|$$
$$= \left| \mathcal{N}_{\vec{\mathbf{M}}}^{\mathbf{k}} [\mathbf{n} - \mathbf{k}] \right|$$

Now by Lemma 2.3.6 the result will follow. □

Lagrange inversion formula follows easily from Good's inversion formula. To see this implication, use the collection $\vec{\mathbf{M}} = (\mathcal{F}_i(M))_{i \in \mathfrak{I}}$. Observe that $i \not\leq j$ implies that $(\mathcal{F}_i(M))^{\bullet(j)} = 0$. Thus the determinant in Good's inversion formula is upper triangular, and it follows that its value is the product of the elements on the main diagonal. Thus Lagrange inversion formula is proved.

By equating the coefficients of the equation system $\mathbf{f}(\mathbf{x}) = \mathbf{x} \cdot (\mathbf{G} \circ \mathbf{f})(\mathbf{x})$ a set of recurrences occur for the coefficients of $\mathbf{f}(\mathbf{x})$, and this set of recurrences is easily seen to have a unique solution. Hence the equations $f_i(\mathbf{x}) = x_i \cdot (G_i \circ \mathbf{f})(\mathbf{x})$ for $i \in \mathfrak{I}$ uniquely determines the collection $\mathbf{f}(\mathbf{x})$. Moreover, it is easy to see that $\mathbf{f}(\mathbf{x})$ is a summable collection.

Theorem 6 (Good's inversion formula) *Let $\mathbf{f}(\mathbf{x})$ be a summable collection of formal power series and let $\mathbf{G}(\mathbf{x})$ be a collection of formal power series, such that for $i \in \mathfrak{I}$*

$$f_i(\mathbf{x}) = x_i \cdot (G_i \circ \mathbf{f}(\mathbf{x})).$$

Let $J = \{i \in \mathfrak{I} : n_i \neq 0\}$ and assume that $\mathbf{n} \geq \mathbf{k}$. Then we have that

$$[\mathbf{x}^{\mathbf{n}}] \prod_{i \in \mathfrak{I}} f_i(\mathbf{x})^{k_i} = \left[\mathbf{x}^{\mathbf{n}-\mathbf{k}}\right] \det \left(\delta_{i,j} \cdot G_i(\mathbf{x})^{n_i} - x_j \frac{\partial G_i(\mathbf{x})}{\partial x_j} G_i(\mathbf{x})^{n_i - 1} \right)_{i,j \in J}.$$

Proof: Let M_i be a colored species and $G_i(\mathbf{x})$ its generating function. That is $G_i(\mathbf{x}) = \mathrm{card}(M_i; \mathbf{x})$. Let $\overline{f}_i(\mathbf{x}) = \mathrm{card}(A_{\vec{\mathbf{M}}}^{(i)}; \mathbf{x})$ where $A_{\vec{\mathbf{M}}}^{(i)}$ is the colored species of $\vec{\mathbf{M}}$-enriched colored trees with root of color i. Since $A_{\vec{\mathbf{M}}}^{(i)} = X_i \cdot (M_i \circ \vec{\mathbf{A}}_{\vec{\mathbf{M}}})$, we get

$$\begin{aligned}
\overline{f}_i(\mathbf{x}) &= \operatorname{card}(A_{\vec{\mathbf{M}}}^{(i)}; \mathbf{x}) \\
&= \operatorname{card}\left(X_i \cdot (M_i \circ \vec{\mathbf{A}}_{\vec{\mathbf{M}}}); \mathbf{x}\right) \\
&= x_i \cdot \left(\operatorname{card}(M_i; \mathbf{x}) \circ \operatorname{card}(\vec{\mathbf{A}}_{\vec{\mathbf{M}}}; \mathbf{x})\right) \\
&= x_i \cdot \left(G_i \circ \overline{\mathbf{f}}\right)(\mathbf{x})
\end{aligned}$$

But $f_i(\mathbf{x})$ is uniquely determined by the above equation. Hence $f_i(\mathbf{x}) = \overline{f}_i(\mathbf{x}) = \operatorname{card}(A_{\vec{\mathbf{M}}}^{(i)}; \mathbf{x})$.

Now the left hand side of the species version of Good's inversion formula is equal to

$$\begin{aligned}
\left|\left(\Gamma_{\mathbf{k}}(\vec{\mathbf{A}}_{\vec{\mathbf{M}}})\right)[\mathbf{n}]\right| &= \left[\frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!}\right] \operatorname{card}\left(\Gamma_{\mathbf{k}}(\vec{\mathbf{A}}_{\vec{\mathbf{M}}}); \mathbf{x}\right) \\
&= \left[\frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!}\right] \left(\frac{\mathbf{x}^{\mathbf{k}}}{\mathbf{k}!} \circ \mathbf{f}\right)(\mathbf{x}) \\
&= \frac{\mathbf{n}!}{\mathbf{k}!} [\mathbf{x}^{\mathbf{n}}] \left(\mathbf{x}^{\mathbf{k}} \circ \mathbf{f}\right)(\mathbf{x}).
\end{aligned}$$

And the right hand side

$$\begin{aligned}
&\binom{\mathbf{n}}{\mathbf{k}} \cdot \left|\left(\det\left(\delta_{i,j} M_i^{n_i} - M_i^{\bullet(j)} \cdot M_i^{n_i-1}\right)_{i,j\in J}\right)[\mathbf{n}-\mathbf{k}]\right| \\
= &\binom{\mathbf{n}}{\mathbf{k}} \cdot \left[\frac{\mathbf{x}^{\mathbf{n}-\mathbf{k}}}{(\mathbf{n}-\mathbf{k})!}\right] \operatorname{card}\left(\det\left(\delta_{i,j} M_i^{n_i} - M_i^{\bullet(j)} \cdot M_i^{n_i-1}\right)_{i,j\in J}; \mathbf{x}\right) \\
= &\frac{\mathbf{n}!}{\mathbf{k}!} \cdot \left[\mathbf{x}^{\mathbf{n}-\mathbf{k}}\right] \det\left(\delta_{i,j} G_i(\mathbf{x})^{n_i} - x_j \cdot \frac{\partial G_i(\mathbf{x})}{\partial x_j} \cdot G_i(\mathbf{x})^{n_i-1}\right)_{i,j\in J}
\end{aligned}$$

Thus we have proven the theorem for formal power series $G_i(\mathbf{x})$ such that $\left[\frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!}\right] G_i(\mathbf{x})$ is a nonnegative integer for all multi indices $\mathbf{n}$. By the principle of extension of algebraic identities the theorem follows for all $\mathbf{G}(\mathbf{x})$ and $\mathbf{f}(\mathbf{x})$. □

2.11 Umbral calculus

We give here a short review of some result in the infinite variated multi-variable umbral calculus. The theory was developed in **[C1]** by W. Chen. We will only present those result that we need for this presentation.

Let $\mathcal{K}[\mathbf{x}]$ be the set of all polynomials in the variables $(x_i)_{i\in\mathcal{I}}$.

An operator M on $\mathcal{K}[\mathbf{x}]$ is a linear map from $\mathcal{K}[\mathbf{x}]$ to itself. Three classical operators are

(i) The partial differentiation with respect to x_i. That is the map

$$D_i p(\mathbf{x}) = \frac{\partial p(\mathbf{x})}{\partial x_i}.$$

(ii) The multiplication with respect to x_i,

$$\underline{\mathbf{x}}_i p(\mathbf{x}) = x_i p(\mathbf{x}).$$

(iii) The shift operator. Let $\mathbf{a}$ be a vector, then the shift is defined to be

$$E^{\mathbf{a}} p(\mathbf{x}) = p(\mathbf{x} + \mathbf{a}).$$

We say that an operator T is invertible if there is another operator S such that $TS = 1$, where 1 is the identity operator.

Definition 2.11.1 *An operator T is called* shift invariant *if it commutes with all shift operators, that is for every vector* $\mathbf{a}$,

$$TE^{\mathbf{a}} = E^{\mathbf{a}}T.$$

From **[C1]** we have the following classification of shift invariant operators.

Proposition 2.11.1 *An operator T on $\mathcal{K}[\mathbf{x}]$ is shift invariant if and only if it is a formal power series in the differential operators $(D_i)_{i\in\mathfrak{I}}$. Thus we can write*

$$T = \sum_{\mathbf{n}} a_{\mathbf{n}} \frac{\mathbf{D}^{\mathbf{n}}}{\mathbf{n}!}.$$

The formal power series $\sum_{\mathbf{n}} a_{\mathbf{n}} \frac{\mathbf{t}^{\mathbf{n}}}{\mathbf{n}!}$ is called the indicator series of T.

Definition 2.11.2 *A* delta operator *is a shift invariant operator Q such that $Q1 = 0$.*

Definition 2.11.3 *A* summable collection *of delta operators $\mathbf{Q} = (Q_i)_{i\in\mathfrak{I}}$ is a set of delta operators Q_i indexed by the set $\mathfrak{I}$, such that their indicator sequences are summable.*

We say that a summable collection of delta operators $\mathbf{Q}$ is *admissible* if there exists a summable collection of formal power series $\mathbf{g}(\mathbf{t}) = (g_i(\mathbf{t}))_{i\in\mathfrak{I}}$ such that

$$(\mathbf{q} \circ \mathbf{g})(\mathbf{t}) = \mathbf{t},$$

or

$$(q_i \circ \mathbf{g})(\mathbf{t}) = t_i,$$

where $q_i(\mathbf{t})$ is the indicator sequence of Q_i. If $\mathbf{Q}$ is admissible, we will denote $g_i(\mathbf{t})$ by $q_i^{\langle -1\rangle}(\mathbf{t})$ and $\mathbf{g}(\mathbf{t})$ by $\mathbf{q}^{\langle -1\rangle}(\mathbf{t})$.

Proposition 2.11.2 *Let $\mathbf{Q} = (Q_i)_{i\in\mathfrak{I}}$ be a summable collection of delta operators. Then there exist a unique polynomial sequence $p_{\mathbf{n}}(\mathbf{x})$, indexed by multi indices $\mathbf{n}$, such that*

$$Q_i p_{\mathbf{n}}(\mathbf{x}) = n_i p_{\mathbf{n}-\mathbf{e}_i}(\mathbf{x}),$$

and

$$p_{\mathbf{n}}(\mathbf{0}) = \delta_{\mathbf{n},\mathbf{0}}.$$

A such sequence is called a basic sequence *of the collection of delta operators $\mathbf{Q}$.*

Proposition 2.11.3 *Let* $\mathbf{Q}$ *be a summable collection of delta operators, which is admissible. Assume that* Q_i *has indicator series* $q_i(\mathbf{t})$*. Let* $(p_{\mathbf{n}}(\mathbf{x}))$ *be the basic sequence of* $\mathbf{Q}$*. Then we have*

$$\sum_{\mathbf{n}} p_{\mathbf{n}}(\mathbf{x})\frac{\mathbf{t}^{\mathbf{n}}}{\mathbf{n}!} = \exp\left(\sum_{i\in\mathfrak{I}} x_i q_i^{\langle -1\rangle}(\mathbf{t})\right),$$

where $q_i^{\langle -1\rangle}(\mathbf{t})$ *is the inverse defined above.*

2.12 The general transfer formula and Good's inversion formula

Definition 2.12.1 *The* Pincherle derivate *of an operator* T *is defined by*

$$\partial_i(T) = T\mathbf{X}_i - \mathbf{X}_i T.$$

Observe that the indicator sequence of $\partial_i(T)$ is the derivative of the indicator sequence of T with respect to t_i.

Lemma 2.12.1 *For a formal power series* $M(\mathbf{x})$ *we have that*

$$[\mathbf{x}^{\mathbf{k}}]M(\mathbf{D})\mathbf{x}^{\mathbf{n}} = [\mathbf{x}^{\mathbf{n}-\mathbf{k}}]\frac{\mathbf{n}!}{\mathbf{k}!}M(\mathbf{x}).$$

Proof: Since the identity is linear in $M(\mathbf{x})$, it is enough to consider the case $M(\mathbf{x}) = \mathbf{x}^{\mathbf{h}}$.

$$\begin{aligned}
[\mathbf{x}^{\mathbf{k}}]\,\mathbf{D}^{\mathbf{h}}\mathbf{x}^{\mathbf{n}} &= [\mathbf{x}^{\mathbf{k}}]\,(\mathbf{n})_{\mathbf{h}}\mathbf{x}^{\mathbf{n}-\mathbf{h}} \\
&= \delta_{\mathbf{k},\mathbf{n}-\mathbf{h}}\cdot(\mathbf{n})_{\mathbf{h}} \\
&= \delta_{\mathbf{n}-\mathbf{k},\mathbf{h}}\cdot(\mathbf{n})_{\mathbf{n}-\mathbf{k}} \\
&= [\mathbf{x}^{\mathbf{n}-\mathbf{k}}]\,\frac{\mathbf{n}!}{\mathbf{k}!}\mathbf{x}^{\mathbf{h}}.
\end{aligned}$$

□

Lemma 2.12.2 *Let* $\mathbf{Q}$ *be a summable collection of delta operators, where* Q_i *has indicator series* $f_i^{\langle -1\rangle}(\mathbf{t})$*. Assume that* $\mathbf{Q}$ *is admissible. Then we have*

$$p_{\mathbf{n}}(\mathbf{x}) = \sum_{\mathbf{k}} \mathbf{x}^{\mathbf{k}} \frac{\mathbf{n}!}{\mathbf{k}!} [\mathbf{t}^{\mathbf{n}}] \left(\mathbf{t}^{\mathbf{k}} \circ \mathbf{f}\right)(\mathbf{t}),$$

where $(p_{\mathbf{n}}(\mathbf{x}))$ *is the basic sequence associated with* $\mathbf{Q}$*.*

Proof: Then we have by Proposition 2.11.3 that

$$\sum_{\mathbf{n}} p_{\mathbf{n}}(\mathbf{x}) \cdot \frac{\mathbf{t}^{\mathbf{n}}}{\mathbf{n}!} = \exp\left(\sum_{i\in\mathfrak{I}} x_i \cdot f_i(\mathbf{t})\right).$$

Thus by looking at the coefficient of $\frac{\mathbf{t}^{\mathbf{n}}}{\mathbf{n}!}$ we get

$$\begin{aligned} p_{\mathbf{n}}(\mathbf{x}) &= \left[\frac{\mathbf{t}^{\mathbf{n}}}{\mathbf{n}!}\right] \exp\left(\sum_{i\in\mathfrak{I}} x_i \cdot f_i(\mathbf{t})\right) \\ &= \mathbf{n}!\,[\mathbf{t}^{\mathbf{n}}] \sum_{\mathbf{k}} \prod_{i\in\mathfrak{I}} \frac{(x_i \cdot f_i(\mathbf{t}))^{k_i}}{k_i!} \\ &= \mathbf{n}!\,[\mathbf{t}^{\mathbf{n}}] \sum_{\mathbf{k}} \frac{\mathbf{x}^{\mathbf{k}}}{\mathbf{k}!} \prod_{i\in\mathfrak{I}} f_i(\mathbf{t})^{k_i} \\ &= \sum_{\mathbf{k}} \mathbf{x}^{\mathbf{k}}\, [\mathbf{t}^{\mathbf{n}}]\, \frac{\mathbf{n}!}{\mathbf{k}!} \left(\mathbf{t}^{\mathbf{k}} \circ \mathbf{f}\right)(\mathbf{t}). \end{aligned}$$

□

Theorem 7 (The general transfer formula) *Let* $\mathbf{Q} = (Q_i)_{i\in\mathfrak{I}}$ *be a summable collection of delta operators, such that* $\mathbf{Q}$ *is admissible. Assume that we can write* $Q_i = D_i P_i$*, where* P_i *is an invertible shift invariant operator. Let* $(p_{\mathbf{n}}(\mathbf{x}))$ *be the basic sequence of the* *summable collection* $\mathbf{Q}$. *Let* $J = \{i \in \mathfrak{I} : n_i \neq 0\}$. *Then we have*

$$p_{\mathbf{n}}(\mathbf{x}) = \det\left(\delta_{i,j} \cdot P_i^{-n_i} - D_j\partial_j(P_i^{-1})P_i^{-n_i+1}\right)_{i,j\in J} \mathbf{x}^{\mathbf{n}}.$$

Moreover, this formula is equivalent to Good's inversion formula.

Proof: First we will prove that Good's inversion formula implies the general transfer formula. Let $h_i(\mathbf{t})$ be the indicator series of Q_i. Since Q_i is a delta operator we know that $h_i^{\langle -1\rangle}(\mathbf{t}) = f_i(\mathbf{t})$ exists. Moreover since P_i is invertible let $G_i(\mathbf{t})$ be the indicator series of P_i^{-1}. Hence $G_i(\mathbf{D}) = P_i^{-1}$. Thus we have

$$f_i^{\langle -1\rangle}(\mathbf{t}) = t_i \cdot G_i^{-1}(\mathbf{t}).$$

This equation is equivalent to

$$\mathbf{f}_i^{\langle -1\rangle}(\mathbf{t}) = \mathbf{t} \cdot \mathbf{G}^{-1}(\mathbf{t}),$$

which can be written as

$$\mathbf{f}(\mathbf{t}) = \mathbf{t} \cdot (\mathbf{G} \circ \mathbf{f})(\mathbf{t}).$$

Thus $f_i(\mathbf{t}) = t_i \cdot (G_i \circ \mathbf{f})(\mathbf{t})$. Let $(p_{\mathbf{n}}(\mathbf{x}))$ be the basic sequence associated with $\mathbf{Q}$. Thus by Lemma 2.12.2, by Good's inversion formula, and by Lemma 2.12.1, we have that

$$\begin{aligned}
p_{\mathbf{n}}(\mathbf{x}) &= \sum_{\mathbf{k}} \mathbf{x}^{\mathbf{k}} \frac{\mathbf{n}!}{\mathbf{k}!} [\mathbf{t}^{\mathbf{n}}] \left(\mathbf{t}^{\mathbf{k}} \circ \mathbf{f}\right)(\mathbf{t}) \\
&= \sum_{\mathbf{k}} \mathbf{x}^{\mathbf{k}} \frac{\mathbf{n}!}{\mathbf{k}!} [\mathbf{x}^{\mathbf{n}}] \left(\mathbf{x}^{\mathbf{k}} \circ \mathbf{f}\right)(\mathbf{x}) \\
&= \sum_{\mathbf{k}} \mathbf{x}^{\mathbf{k}} \frac{\mathbf{n}!}{\mathbf{k}!} \left[\mathbf{x}^{\mathbf{n}-\mathbf{k}}\right] \det\left(\delta_{i,j} \cdot G_i(\mathbf{x})^{n_i} - x_j \frac{\partial G_i(\mathbf{x})}{\partial x_j} G_i(\mathbf{x})^{n_i-1}\right)_{i,j\in J} \\
&= \sum_{\mathbf{k}} \mathbf{x}^{\mathbf{k}} \left[\mathbf{x}^{\mathbf{k}}\right] \det\left(\delta_{i,j} \cdot G_i(\mathbf{D})^{n_i} - D_j\partial_j(G_i(\mathbf{D}))G_i(\mathbf{D})^{n_i-1}\right)_{i,j\in J} \mathbf{x}^{\mathbf{n}} \\
&= \det\left(\delta_{i,j} \cdot G_i(\mathbf{D})^{n_i} - D_j\partial_j(G_i(\mathbf{D}))G_i(\mathbf{D})^{n_i-1}\right)_{i,j\in J} \mathbf{x}^{\mathbf{n}}
\end{aligned}$$

$$= \det\left(\delta_{i,j} \cdot P_i^{-n_i} - D_j \partial_j (P_i^{-1}) P_i^{-n_i+1}\right)_{i,j \in J} \mathbf{x}^{\mathbf{n}}$$

which proves the general transfer formula.

It is easy to see that the general transfer formula implies Good's inversion formula. First assume that $G_i(\mathbf{x})^{-1}$ exists for all $i \in \mathfrak{J}$. Since $f_i(\mathbf{x}) = x_i \cdot (G \circ \mathbf{f})(\mathbf{x})$, we know that $\mathbf{f}^{\langle -1 \rangle}(\mathbf{x})$ exist. Let Q be the plethystic delta operator with indicator series $f^{\langle -1 \rangle}(\mathbf{t})$. Then by the same list of equalities as above we conclude that

$$\sum_{\mathbf{k}} \mathbf{x}^{\mathbf{k}} \frac{\mathbf{n}!}{\mathbf{k}!} [\mathbf{x}^{\mathbf{n}}] \left(\mathbf{x}^{\mathbf{k}} \circ \mathbf{f}\right)(\mathbf{x})$$

$$= \sum_{\mathbf{k}} \mathbf{x}^{\mathbf{k}} \frac{\mathbf{n}!}{\mathbf{k}!} \left[\mathbf{x}^{\mathbf{n}-\mathbf{k}}\right] \det\left(\delta_{i,j} \cdot G_i(\mathbf{x})^{n_i} - x_j \frac{\partial G_i(\mathbf{x})}{\partial x_j} G_i(\mathbf{x})^{n_i - 1}\right)_{i,j \in J}$$

Take the coefficient of $\mathbf{x}^{\mathbf{k}}$ on both sides and we obtain Lagrange inversion formula,

$$[\mathbf{x}^{\mathbf{n}}] \left(\mathbf{x}^{\mathbf{k}} \circ \mathbf{f}\right)(\mathbf{x})$$

$$= \left[\mathbf{x}^{\mathbf{n}-\mathbf{k}}\right] \det\left(\delta_{i,j} \cdot G_i(\mathbf{x})^{n_i} - x_j \frac{\partial G_i(\mathbf{x})}{\partial x_j} G_i(\mathbf{x})^{n_i - 1}\right)_{i,j \in J}$$

Thus the implication is proved in the case when $G_i(\mathbf{x})^{-1}$ exists for all $i \in \mathfrak{J}$.

To complete the proof of the implication, consider the coefficients $b_{\mathbf{n},m}$ of the collection $\mathbf{G}(\mathbf{x}) = (G_m(\mathbf{x}))_{m \in \mathfrak{J}}$ as indeterminants. That is $G_m(\mathbf{x}) = \sum_{\mathbf{n}} b_{\mathbf{n},m} \mathbf{x}^{\mathbf{n}}$. Let now

$$a_{\mathbf{n},m} = \left[\mathbf{x}^{\mathbf{n}-\mathbf{e}_m}\right] \det\left(\delta_{i,j} \cdot G_i(\mathbf{x})^{n_i} - x_j \frac{\partial G_i(\mathbf{x})}{\partial x_j} G_i(\mathbf{x})^{n_i - 1}\right)_{i,j \in J}.$$

Observe that $a_{\mathbf{n},m}$ is by the above equation expressed as a polynomial in the indeterminants $b_{\mathbf{n},m}$. Let $\overline{f}_m(\mathbf{x}) = \sum_{\mathbf{n}} a_{\mathbf{n},m} \mathbf{x}^{\mathbf{n}}$. We claim that $\overline{f}_m(\mathbf{x}) = x_m \cdot (G_m \circ \overline{\mathbf{f}})(\mathbf{x})$. Compare coefficients of both sides. The equations that arises are polynomial identities in the indeterminants $b_{\mathbf{n},m}$. But we have shown above these identities are true in the case when $b_{\mathbf{0},m}$ is nonzero for each $m \in \mathfrak{J}$. By the principle of extension of algebraic identities, the polynomial identities follow and the claim is proved. By the uniqness of the collection

$\mathbf{f}$ we conclude that $f_m(\mathbf{x}) = \overline{f}_m(\mathbf{x})$. Now to prove Good's inversion formula, use that it is a polynomial identity in $b_{\mathbf{n},m}$ and apply again the principle of extension of algebraic identities. This argument finishes up the implication and thus the equivalence is proved. □

2.13 The plethystic umbral calculus

Assume now that our index set $\mathfrak{I}$ is given the structure of a c-monoid.

Definition 2.13.1 *The* Frobenius operator *of a shift invariant operator* $T = \sum_{\mathbf{n}} a_{\mathbf{n}} \frac{\mathbf{D}^{\mathbf{n}}}{\mathbf{n}!}$ *is defined to be*

$$\mathcal{F}_i(T) = \sum_{\mathbf{n}} a_{\mathbf{n}} \frac{\mathbf{D}^{\nu_i(\mathbf{n})}}{\mathbf{n}!}.$$

Thus we have that

$$\mathcal{F}_i(D_j) \;=\; D_{i\cdot j}.$$

Definition 2.13.2 *A* plethystic delta operator *is a shift invariant operator* Q *such that* $Q1 = 0$ *and* Qx_1 *is a nonzero constant.*

Observe that $\mathbf{Q} = (\mathcal{F}_i(Q))_{i\in\mathfrak{I}}$ is a summable sequence of delta operators. Hence by Proposition 2.11.2 there exist a unique polynomial sequence $p_{\mathbf{n}}(\mathbf{x})$, index by multi indices $\mathbf{n}$, such that

$$\mathcal{F}_i(Q)p_{\mathbf{n}}(\mathbf{x}) = n_i p_{\mathbf{n}-\mathbf{e}_i}(\mathbf{x}),$$

and

$$p_{\mathbf{n}}(\mathbf{0}) = \delta_{\mathbf{n},\mathbf{0}}.$$

A such sequence is called a *plethystic basic sequence* of the plethystic delta operator Q.

2.14 The plethystic transfer formula and Lagrange inversion formula

Choose a linear ordering $(\mathfrak{J}, \preceq)$ of the c-monoid $\mathfrak{J}$, which is compatible with the divisibility ordering $(\mathfrak{J}, \leq)$. That is $(\mathfrak{J}, \preceq)$ is a total order such that $i \leq j$ implies $i \preceq j$.

Example 2.14.1 *In the c-monoid of positive integers under multiplication,* $(\mathbf{P}, \cdot, 1)$ *we can choose the linear ordering* $(\mathbf{P}, \preceq)$ *to be the natural linear ordering on positive integers. Notice that this linear ordering is compatible with the divisibility ordering.*

We employ the following rule when multiplying noncommutative products over index set J, where J is a finite subset of $\mathfrak{J}$. Multiply the factors in the order given by $(\mathfrak{J}, \preceq)$. That is

$$\prod_{i \in J} A_i = A_{i_1} \cdot A_{i_2} \cdots A_{i_m},$$

where $J = \{i_1, i_2, \ldots, i_m\}$ and $i_1 \prec i_2 \prec \cdots \prec i_m$.

Notice that the plethystic inverse of $f(\mathbf{x})$ exists, that is $f^{\langle -1 \rangle}(\mathbf{x})$, is equivalent to that the inverse of $G(\mathbf{x})$ exists, which is $G^{-1}(\mathbf{x})$.

Proposition 2.14.1 *Let* $G(\mathbf{x})$ *be a invertible formal power series, and suppose* $f(\mathbf{x}) = x_1 \cdot (G \circ f)(\mathbf{x})$. *Let* P *a shift invariant operator with indicator series* $G^{-1}(\mathbf{x})$. *Then*

$$\left(\prod_{i \in J} \mathbf{x}_i P_i^{-n_i} \mathbf{x}_i^{n_i - 1}\right) 1 = \left(\prod_{i \in J} H_i(\mathbf{D})\right) \mathbf{x}^{\mathbf{n}},$$

where

$$H_i(\mathbf{x}) = \mathcal{F}_i \left(G(\mathbf{x})^{n_i} - x_1 \cdot \frac{\partial G(\mathbf{x})}{\partial x_1} \cdot G(\mathbf{x})^{n_i - 1}\right).$$

Proof: By the definition of Pincherle derivate, we have that

$$\prod_{i \in J} \mathbf{x}_i P_i^{-n_i} \mathbf{x}_i^{n_i - 1} = \prod_{i \in J} \left(P_i^{-n_i} \mathbf{x}_i^{n_i} - \partial_i \left(P_i^{-n_i}\right) \mathbf{x}_i^{n_i - 1}\right). \quad (2.1)$$

Let

$$K_{i,I} = \begin{cases} \partial_i \left(P_i^{-n_i}\right) & \text{if} \quad i \in I \\ P_i^{-n_i} & \text{if} \quad i \in J - I \end{cases},$$

and

$$L_{i,I} = \begin{cases} \underline{\mathbf{x}}_i^{n_i - 1} & \text{if} \quad i \in I \\ \underline{\mathbf{x}}_i^{n_i} & \text{if} \quad i \in J - I \end{cases}.$$

Observe that if $i \prec j$, then $i \not\succeq j$, so the operators $L_{i,I}$ and $K_{j,I}$ commutes. That is

$$L_{i,I} K_{j,I} = K_{j,I} L_{i,I}.$$

Notice also that the $K_{i,I}$ and $K_{j,I}$ commutes and that the $L_{i,I}$ and $L_{j,I}$ commutes. Thus we can expand the right hand side of (2.1), and using these commuting relations

$$\begin{aligned} \prod_{i \in J} \underline{\mathbf{x}}_i P_i^{-n_i} \underline{\mathbf{x}}_i^{n_i - 1} &= \sum_{I \subseteq J} (-1)^{|I|} \prod_{i \in J} K_{i,I} L_{i,I} \\ &= \sum_{I \subseteq J} (-1)^{|I|} \prod_{i \in J} K_{i,I} \cdot \prod_{i \in J} L_{i,I} \\ &= \sum_{I \subseteq J} (-1)^{|I|} \prod_{i \in I} K_{i,I} \prod_{i \in J - I} K_{i,I} \cdot \prod_{i \in J} L_{i,I}. \end{aligned}$$

Apply now the above operator identity to the polynomial 1, and we obtain

$$\left(\prod_{i \in J} \underline{\mathbf{x}}_i P_i^{-n_i} \underline{\mathbf{x}}_i^{n_i - 1} \right) 1 = \left(\sum_{I \subseteq J} (-1)^{|I|} \prod_{i \in I} K_{i,I} \prod_{i \in J - I} K_{i,I} \right) \mathbf{x}^{\mathbf{n} - \mathbf{e}_I}. \qquad (2.2)$$

For $i \in I$ then we have that

$$\begin{aligned} K_{i,I} &= \partial_i \left(P_i^{-n_i}\right) \\ &= \mathcal{F}_i \left(\partial_1 \left(P^{-n_i}\right)\right) \\ &= \mathcal{F}_i \left(\partial_1 \left(G(\mathbf{D})^{n_i}\right)\right) \\ &= \mathcal{F}_i \left(n_i G'(\mathbf{D}) G(\mathbf{D})^{n_i - 1}\right). \end{aligned}$$

Similarly for $i \in J - I$,

$$K_{i,I} = \mathcal{F}_i(G(\mathbf{D})^{n_i}).$$

Apply this now to equation (2.2).

$$\begin{aligned}
&\left(\prod_{i\in J} \underline{\mathbf{x}}_i P_i^{-n_i} \underline{\mathbf{x}}_i^{n_i-1}\right) 1 \\
&= \left(\sum_{I\subseteq J}(-1)^{|I|}\prod_{i\in I}\mathcal{F}_i\left(n_i G'(\mathbf{D})G(\mathbf{D})^{n_i-1}\right)\cdot\prod_{i\in J-I}\mathcal{F}_i(G(\mathbf{D})^{n_i})\right)\mathbf{x}^{\mathbf{n}-\mathbf{e}_I} \\
&= \left(\sum_{I\subseteq J}(-1)^{|I|}\prod_{i\in I}\mathcal{F}_i\left(n_i G'(\mathbf{D})G(\mathbf{D})^{n_i-1}\right)\cdot\prod_{i\in J-I}\mathcal{F}_i(G(\mathbf{D})^{n_i})\cdot\prod_{i\in I}\frac{1}{n_i}D_i\right)\mathbf{x}^{\mathbf{n}} \\
&= \left(\sum_{I\subseteq J}(-1)^{|I|}\prod_{i\in I}\mathcal{F}_i\left(D_1 G'(\mathbf{D})G(\mathbf{D})^{n_i-1}\right)\cdot\prod_{i\in J-I}\mathcal{F}_i(G(\mathbf{D})^{n_i})\right)\mathbf{x}^{\mathbf{n}} \\
&= \left(\prod_{i\in J}\mathcal{F}_i\left(G(\mathbf{D})^{n_i} - D_1 G'(\mathbf{D})G(\mathbf{D})^{n_i-1}\right)\right)\mathbf{x}^{\mathbf{n}} \\
&= \left(\prod_{i\in J} H_i(\mathbf{D})\right)\mathbf{x}^{\mathbf{n}}.
\end{aligned}$$

That completes the proof of the proposition. □

William Chen obtains in **[C1]** the plethystic transfer formula.

Theorem 8 (The plethystic transfer formula) *Let P be an invertible shift invariant operator, and let $(p_{\mathbf{n}}(\mathbf{x}))$ be the plethystic basic sequence of the plethystic delta operator $Q = D_1 P$. Let P_i denote $\mathcal{F}_i(P)$. Then we have*

$$p_{\mathbf{n}}(\mathbf{x}) = \left(\prod_{i\in J} \underline{\mathbf{x}}_i P_i^{-n_i} \underline{\mathbf{x}}_i^{n_i-1}\right) 1.$$

Theorem 9 *The plethystic transfer formula and the plethystic version of Lagrange inversion formula are equivalent.*

Proof: Apply the plethystic case to Theorem 7. Observe that $i \not\leq j$ implies that $\partial_j(\mathcal{F}_i(P)^{-1}) = 0$. Hence the determinant will be upper triangular. Recall that the determinant of an upper triangular matrix is equal to the product of the elements on the main diagonal. Thus we know that the plethystic Lagrange inversion formula is equivalent to the formula

$$\begin{aligned} p_{\mathbf{n}}(\mathbf{x}) &= \prod_{i\in J}\left(P_i^{-n_i} - D_i\partial_i(P_i^{-1})P_i^{-n_i+1}\right)\ \mathbf{x}^{\mathbf{n}} \\ &= \prod_{i\in J}\mathcal{F}_i\left(P^{-n_i} - D_1\partial_1(P^{-1})P^{-n_i+1}\right)\ \mathbf{x}^{\mathbf{n}} \\ &= \prod_{i\in J}\mathcal{F}_i\left(G(\mathbf{D})^{n_i} - D_1G'(\mathbf{D})G(\mathbf{D})^{n_i-1}\right)\ \mathbf{x}^{\mathbf{n}} \\ &= \prod_{i\in J}H_i(\mathbf{D})\ \mathbf{x}^{\mathbf{n}}. \end{aligned}$$

Thus by the identity in Proposition 2.14.1 we conclude that the plethystic Lagrange inversion formula is equivalent to the plethystic transfer formula. □

2.15 Enumerating plethystic trees

We will now obtain formulas that enumerate different plethystic trees. These formulas will be used in the two last proofs of the plethystic Lagrange inversion formula Theorem 4. We begin by recalling the definition of a labeled plethystic tree.

Definition 2.15.1 *A* labeled plethystic tree (forest) *on a* **n** *set* (E, f) *is a tree (forest) on E such that if* $b \in E$ *is the the son* $a \in E$ *then* $f(a) \leq f(b)$. *The degree of a vertex a, which we denote* $\delta(a)$, *is a multi index* $\vec{\delta}(a) = (\delta(a,i))_{i\in J}$ *where* $\delta(a,i)$ *is the number of sons of the element a of color* $i \cdot f(a)$.

Lemma 2.15.1 *Let B and C be disjoint sets and let C be nonempty. Assume that each element $a \in B \dot\cup C$ has a nonnegative integer $\delta(a)$ associated with it, such that*

$$\sum_{a \in B \dot\cup C} \delta(a) = |B \dot\cup C| - (|B| + k).$$

Then the number of forests with $|B| + k$ components on $B \dot\cup C$ such that $\delta(a)$ is the degree of vertex a in the forest and every element of B is a root is

$$\frac{(|C|-1)! \cdot (|C| - \sum_{c \in C} \delta(c))}{k! \cdot \prod_{a \in B \dot\cup C} \delta(a)!}.$$

Proof: First count the number forests on C with $\sum_{b \in B} \delta(b) + k$ components. By Theorem 5.3.4 in **[St]** the number of such forests is

$$\binom{|C|-1}{\sum_{b \in B} \delta(b) + k - 1} \cdot \binom{|C| - \sum_{b \in B} \delta(b) - k}{\{\delta(c)\}_{c \in C}}.$$

Now to make a such forest on C into a forest on $B \dot\cup C$, we need to connect the roots in C to the set B. There are

$$\binom{\sum_{b \in B} \delta(b) + k}{k, \{\delta(b)\}_{b \in B}}$$

possible ways to do this. Thus the number of trees we are looking for is

$$\begin{aligned} & \binom{|C|-1}{\sum_{b \in B} \delta(b) + k - 1} \cdot \binom{|C| - \sum_{b \in B} \delta(b) - k}{\{\delta(c)\}_{c \in C}} \cdot \binom{\sum_{b \in B} \delta(b) + k}{k, \{\delta(b)\}_{b \in B}} \\ = & \frac{(|C|-1)! \cdot (\sum_{b \in B} \delta(b) + k)}{k! \cdot \prod_{a \in B \dot\cup C} \delta(a)!} \\ = & \frac{(|C|-1)! \cdot (|C| - \sum_{c \in C} \delta(c))}{k! \cdot \prod_{a \in B \dot\cup C} \delta(a)!}. \end{aligned}$$

The last equality follows since

$$\sum_{b\in B}\delta(b)+\sum_{c\in C}\delta(c)=|C|-k.$$

□

Proposition 2.15.2 *Given an* $\mathbf{n}$ *set* (E,f). *Let* $J=\{i\in\mathfrak{I} : n_i\neq 0\}$. *Assume that each element* $a\in E$ *has a given multi-index* $\vec{\delta}(a)$ *such that*

$$\sum_{a\in E}\mathcal{V}_{f(a)}(\vec{\delta}(a))=\mathbf{n}-\mathbf{k}.$$

Then the number of plethystic forests on the colored set (E,f) *such that the degree of vertex* a *is given by* $\vec{\delta}(a)$ *and there are* k_i *roots of color* i *is given by*

$$\prod_{i\in J}\frac{n_i!}{k_i!}\left(1-\frac{1}{n_i}\sum_{f(a)=i}\delta(a,1)\right)\cdot\frac{1}{\prod_{a\in E}\vec{\delta}(a)!}.$$

Proof: Let

$$\begin{aligned}
B_i &= \{a\in E : f(a)<i\}\\
C_i &= \{a\in E : f(a)=i\}\\
\delta_i(a) &= \delta(a,j) \quad \text{where } j\cdot f(a)=i
\end{aligned}$$

Note that k_i is the number of roots of color i, and thus the number of roots in C_i. Now apply Lemma 2.15.1 to B_i, C_i and δ_i, and multiply for all $i\in\mathfrak{I}$ such that $n_i\neq 0$. That is, multiply over $i\in J$. Thus the number of forests is

$$\prod_{i\in J}\frac{(|C_i|-1)!\cdot\left(|C_i|-\sum_{c\in C_i}\delta_i(c)\right)}{k_i!\cdot\prod_{a\in B_i\dot\cup C_i}\delta_i(a)!}=\prod_{i\in J}\frac{(n_i-1)!\cdot\left(n_i-\sum_{f(a)=i}\delta(a,1)\right)}{k_i!\cdot\prod_{j\cdot f(a)=i}\delta(a,j)!}$$

$$= \prod_{i \in J} \frac{n_i!}{k_i!} \left(1 - \frac{1}{n_i} \sum_{f(a)=i} \delta(a,1)\right) \cdot \frac{1}{\prod_{a \in E} \vec{\delta}(a)!}.$$

□

Proposition 2.15.3 *Given an* $\mathbf{n}$ *set* (E, f) *and a sequence of functions* $\{r_i\}_{i \in \mathfrak{I}}$ *where each function goes from multi-index to nonnegative integers. Let* $J = \{i \in \mathfrak{I} : n_i \neq 0\}$ *and let*

$$\vec{\rho}_i = \sum_{\mathbf{m}} r_i(\mathbf{m}) \cdot \mathbf{m},$$

and assume that

$$\sum_{\mathbf{m}} r_i(\mathbf{m}) = n_i \tag{2.3}$$

$$\sum_{i \in \mathfrak{I}} \mathcal{V}_i(\vec{\rho}_i) = \mathbf{n} - \mathbf{k} \tag{2.4}$$

Then the number of labeled plethystic forests on the colored set (E, f) *such that the number of vertices of color* i *and degree* $\mathbf{m}$ *is* $r_i(\mathbf{m})$*, and there are* k_i *roots of color* i *is given by*

$$\prod_{i \in J} \frac{n_i!}{k_i!} \left(1 - \frac{1}{n_i} \rho_{i,1}\right) \binom{n_i}{\{r_i(\mathbf{m})\}_{\mathbf{m}}} \frac{1}{\prod_{\mathbf{m}} \mathbf{m}!^{r_i(\mathbf{m})}}.$$

Proof: Observe that there are

$$\binom{n_i}{\{r_i(\mathbf{m})\}_{\mathbf{m}}}$$

possible ways to assign degrees to the vertices of color i. Note also that

$$\rho_{i,1} = \sum_{\mathbf{m}} r_i(\mathbf{m}) \cdot m_1 = \sum_{f(a)=i} \delta(a, 1),$$

and

$$\prod_{a \in E} \vec{\delta}(a)! = \prod_{i \in J} \prod_{\mathbf{m}} \mathbf{m}!^{r_i(\mathbf{m})}.$$

Thus the number of forests is

$$\prod_{i\in J}\binom{n_i}{\{r_i(\mathbf{m})\}_{\mathbf{m}}}\frac{n_i!}{k_i!}\left(1-\frac{1}{n_i}\rho_{i,1}\right)\frac{1}{\prod_{\mathbf{m}}\mathbf{m}!^{r_i(\mathbf{m})}}.$$

□

Let $(t_{\mathbf{m}})_{\mathbf{m}}$ be independent indeterminates. Let π be a plethystic forest on a $\mathbf{n}$ set (E,f). Define the weight of the plethystic forest π to be

$$\mathbf{t}^{\pi}=\prod_{a\in E}t_{\delta(a)}.$$

Theorem 10 *Given an* $\mathbf{n}$ *set* (E,f), *and let* $J=\{i\in\mathfrak{I}:n_i\neq 0\}$. *The sum of the weights of labeled plethystic forests on the colored set* (E,f) *such that there are* k_i *roots of color* i *is given by*

$$\sum_{(r_i)_{i\in\mathfrak{I}}}\prod_{i\in J}\frac{n_i!}{k_i!}\left(1-\frac{1}{n_i}\rho_{i,1}\right)\binom{n_i}{\{r_i(\mathbf{m})\}_{\mathbf{m}}}\prod_{\mathbf{m}}\left(\frac{t_{\mathbf{m}}}{\mathbf{m}!}\right)^{r_i(\mathbf{m})},$$

where $(r_i)_{i\in\mathfrak{I}}$ *sum over the conditions (2.3) and (2.4).*

Proof: The proof follows directly from Proposition 2.15.3. □

In the third proof of the plethystic Lagrange inversion formula that we will present, we will make use of unlabeled plane plethystic trees. These trees are the plethystic analogue of unlabeled plane trees.

Definition 2.15.2 *A* unlabeled plane plethystic tree (forest) *on a colored* $\mathbf{n}$ *set* (E,f) *is a unlabeled tree (forest) on* E *such that if* $b\in E$ *is the the son* $a\in E$ *then* $f(a)\leq f(b)$. *Moreover for each node and each color, there is a linear order on this nodes set of sons of this color. Observe that there is also a linear order among roots of the same color.*

As before we define degree of a vertex a, which we denote $\vec{\delta}(a)$, is a multi index $\vec{\delta}(a) = (\delta(a,i))_{i\in\mathcal{I}}$ where $\delta(a,i)$ is the number of sons of the element a of color $i \cdot f(a)$. Moreover let $r_i(\mathbf{m})$ be the number of nodes of color i and degree $\mathbf{m}$.

Lemma 2.15.4 *Let r be a function from multi indices to nonnegative integers, such that*

$$\begin{aligned} \sum_{\mathbf{m}} r(\mathbf{m}) &= n, \\ \sum_{\mathbf{m}} r(\mathbf{m}) \cdot \mathbf{m} &= \vec{\rho}. \end{aligned}$$

Then the number of plane forests of a set of cardinality n where a node is labeled by its multi index degree $\vec{\delta}(a)$, but its actually degree in the tree is $\delta(a,1)$, is given by

$$\left(1 - \frac{\rho_1}{n}\right) \binom{n}{\{r(\mathbf{m})\}_{\mathbf{m}}}.$$

Proof: Let

$$r_i = \sum_{\mathbf{m}: m_1 = i} r(\mathbf{m}).$$

Then r_i is the number of nodes with actual degree i. Thus by Theorem 5.3.10 in **[St]**, we know there is

$$\frac{k}{n}\binom{n}{\{r_i\}_i},$$

number of plane trees with r_i nodes of degree i, and where

$$n - k = \sum_{i \geq 0} i \cdot r_i = \rho_1.$$

The number of ways to add the extra information to the nodes of degree i is

$$\binom{r_i}{\{r_{\mathbf{m}}\}_{\mathbf{m}: m_1 = i}}.$$

Hence multiply these values together and the result follows. □

Instead of this proof, one can use a very similar method to the proof of the Theorem 5.3.10 in **[St]**.

Proposition 2.15.5 *Given an* $\mathbf{n}$ *set* (E, f) *and a sequence of functions* $(r_i)_{i\in\mathfrak{I}}$ *where each function goes from multi indices to nonnegative integers. Let* $J = \{i \in \mathfrak{I} : n_i \neq 0\}$, *and let*

$$\vec{\rho}_i = \sum_{\mathbf{m}} r_i(\mathbf{m}) \cdot \mathbf{m},$$

and assume that

$$\sum_{\mathbf{m}} r_i(\mathbf{m}) = n_i \qquad (2.5)$$

$$\sum_{i\in\mathfrak{I}} \mathcal{V}_i(\vec{\rho}_i) = \mathbf{n} - \mathbf{k} \qquad (2.6)$$

Then the number of unlabeled plane plethystic forests on the colored set (E, f) *such that the number of vertices of color* i *and degree* $\mathbf{m}$ *is* $r_i(\mathbf{m})$, *and there are* k_i *roots of color* i *is given by*

$$\prod_{i\in J}\left(1 - \frac{\rho_{i,1}}{n_i}\right)\binom{n_i}{\{r_i(\mathbf{m})\}_{\mathbf{m}}}.$$

Proof: Count the number of unlabeled plane forests by counting what happens inside each color class. By Lemma 2.15.4 the number of forests on nodes of color i is

$$\left(1 - \frac{\rho_{i,1}}{n_i}\right)\binom{n_i}{\{r_i(\mathbf{m})\}_{\mathbf{m}}}.$$

Since the linear ordering on the roots of the same color, there is a unique way to connect all the forests together to a unlabeled plane plethystic tree. Hence we need only to multiply over $i \in \mathfrak{I}$, such that $n_i \neq 0$ and the proposition follows. □

Let $(u_{\mathbf{m}})_{\mathbf{m}}$ be independent indeterminates. Let π be a unlabeled plane plethystic forest on a $\mathbf{n}$ set (E, f). Define the weight of the plethystic forest π to be

$$\mathbf{u}^{\pi} = \prod_{a \in E} u_{\delta(a)}.$$

Theorem 11 *Given an* $\mathbf{n}$ *set* (E, f). *Let* $J = \{i \in \mathfrak{I} : n_i \neq 0\}$. *The sum of the weights of unlabeled plethystic forests on the colored set* (E, f) *such that there are* k_i *roots of color* i *is given by*

$$\sum_{(r_i)_{i \in \mathfrak{I}}} \prod_{i \in J} \left(1 - \frac{1}{n_i} \rho_{i,1}\right) \binom{n_i}{\{r_i(\mathbf{m})\}_{\mathbf{m}}} \prod_{\mathbf{m}} u_{\mathbf{m}}^{r_i(\mathbf{m})},$$

where $(r_i)_{i \in \mathfrak{I}}$ *sum over the conditions (2.5) and (2.6).*

Proof: The proof follows directly from Proposition 2.15.5. □

2.16 Enumerative proofs of plethystic Lagrange inversion formula

Let us recall the plethystic Lagrange inversion formula Theorem 4.

Theorem 4 (Lagrange inversion formula) *Let* $f(\mathbf{x})$ *and* $G(\mathbf{x})$ *be power series in the variables* $(x_i)_{i \in \mathfrak{I}}$ *such that*

$$f(\mathbf{x}) = x_1 \cdot (G * f)(\mathbf{x}).$$

Assume that $\mathbf{n} \geq \mathbf{k}$, *and let* $J = \{i \in \mathfrak{I} : n_i \neq 0\}$. *Then*

$$[\mathbf{x}^{\mathbf{n}}] \left(\mathbf{x}^{\mathbf{k}} * f\right)(\mathbf{x}) = \left[\mathbf{x}^{\mathbf{n}-\mathbf{k}}\right] \prod_{i \in J} H_i(\mathbf{x}).$$

where

$$H_i(\mathbf{x}) = \mathcal{F}_i \left(G(\mathbf{x})^{n_i} - x_1 \cdot \frac{\partial G(\mathbf{x})}{\partial x_1} \cdot G(\mathbf{x})^{n_i - 1} \right).$$

The first proof will make use of enumeration of labeled plethystic trees in Theorem 10.

Second proof: Write the power series $G(\mathbf{x})$ as

$$G(\mathbf{x}) = \sum_{\mathbf{n}} t_{\mathbf{n}} \frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!},$$

where $t_{\mathbf{n}}$ are the coefficients of $G(\mathbf{x})$, where we view $G(\mathbf{x})$ as an exponential generating function. We will use the $t_{\mathbf{n}}$'s as the weights in the definition of the weight of a plethystic forest.

We claim that

$$f(\mathbf{x}) = \sum_{\mathbf{n}} \left(\sum_{\tau} \mathbf{t}^{\tau} \right) \frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!}, \tag{2.7}$$

where τ ranges over all rooted trees on a $\mathbf{n}$ set, and the root has color 1. We have then that

$$\frac{f(\mathbf{x})^{k_i}}{k_i!} = \sum_{\mathbf{n}} \left(\sum_{\pi} \mathbf{t}^{\pi} \right) \frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!},$$

where π ranges over all rooted forests on a $\mathbf{n}$ set, and the forests have k_i roots of color 1. Thus

$$\frac{\mathcal{F}_i(f(\mathbf{x}))^{k_i}}{k_i!} = \sum_{\mathbf{n}} \left(\sum_{\pi} \mathbf{t}^{\pi} \right) \frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!},$$

where π ranges over all rooted forests on a $\mathbf{n}$ set, and the forests have k_i roots of color i. Hence by multiplying

$$\prod_{i \in \mathfrak{I}} \frac{\mathcal{F}_i(f(\mathbf{x}))^{k_i}}{k_i!} = \sum_{\mathbf{n}} \left(\sum_{\pi} \mathbf{t}^{\pi} \right) \frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!}, \tag{2.8}$$

where π ranges over all rooted forests on a $\mathbf{n}$ set, and the forests have k_i roots of color i, for $i \in \mathfrak{I}$. Thus

$$t_{\mathbf{k}} x_1 \prod_{i \in \mathfrak{I}} \frac{\mathcal{F}_i(f(\mathbf{x}))^{k_i}}{k_i!} = \sum_{\mathbf{n}} \left(\sum_{\tau} \mathbf{t}^{\tau} \right) \frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!},$$

where τ ranges over all rooted trees on a $\mathbf{n}$ set, and the root has color 1 and degree $\mathbf{k}$. By summing over all $\mathbf{k}$, we conclude that

$$\sum_{\mathbf{k}} t_{\mathbf{k}} x_1 \prod_{i \in \mathfrak{I}} \frac{\mathcal{F}_i(f(\mathbf{x}))^{k_i}}{k_i!} = f(\mathbf{x}).$$

This can be written as

$$x_1 \cdot G\left((\mathcal{F}_i(f(\mathbf{x})))_{i\in\mathfrak{I}}\right) = f(\mathbf{x}),$$

or

$$x_1 \cdot (G \circ f)(\mathbf{x}) = f(\mathbf{x}), \tag{2.9}$$

by using plethystic composition.

By uniqueness of the plethystic inverse, we know that there is only one $f(\mathbf{x})$ that satisfies the above equation. Hence our claim that $f(\mathbf{x})$ is given by equation (2.7), is justified.

Look now at the coefficient of $\left[\frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!}\right]$ in equation (2.8). That is

$$\left[\frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!}\right] \prod_{i\in\mathfrak{I}} \frac{\mathcal{F}_i(f(\mathbf{x}))^{k_i}}{k_i!} = \sum_{\pi} \mathbf{t}^{\pi}.$$

By Theorem 10 we can rewrite the right hand side, thus

$$\left[\frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!}\right] \prod_{i\in\mathfrak{I}} \frac{\mathcal{F}_i(f(\mathbf{x}))^{k_i}}{k_i!} = \sum_{(r_i)_{i\in\mathfrak{I}}} \prod_{i\in J} \frac{n_i!}{k_i!} \left(1 - \frac{\rho_{i,1}}{n_i}\right) \binom{n_i}{r_i(\mathbf{m})} \prod_{\mathbf{m}} \left(\frac{t_{\mathbf{m}}}{\mathbf{m}!}\right)^{r_i(\mathbf{m})},$$

where the $(r_i)_{i\in I}$ in the sum fulfill the conditions (2.3) and (2.4). Recall that $\rho_{i,1}$ is the first component of $\vec{\rho}_i$. We can rewrite the above equation to

$$[\mathbf{x}^{\mathbf{n}}] \prod_{i\in\mathfrak{I}} \mathcal{F}_i(f(\mathbf{x}))^{k_i} = \sum_{(r_i)_{i\in\mathfrak{I}}} \prod_{i\in J} \left(1 - \frac{\rho_{i,1}}{n_i}\right) \binom{n_i}{r_i(\mathbf{m})} \prod_{\mathbf{m}} \left(\frac{t_{\mathbf{m}}}{\mathbf{m}!}\right)^{r_i(\mathbf{m})}. \tag{2.10}$$

For $i \in J$, that is, $n_i > 0$, we have by the multinomial theorem that

$$G(\mathbf{x})^{n_i} = \sum_{s_i} \binom{n_i}{s_i(\mathbf{m})} \prod_{\mathbf{m}} \left(\frac{t_{\mathbf{m}}}{\mathbf{m}!}\right)^{s_i(\mathbf{m})} \cdot \mathbf{x}^{\vec{\sigma}_i},$$

where $\sum_{\mathbf{m}} s_i(\mathbf{m}) = n_i$. Thus

$$x_1 \cdot \frac{\partial}{\partial x_1} G(\mathbf{x})^{n_i} = \sum_{s_i} \binom{n_i}{s_i(\mathbf{m})} \prod_{\mathbf{m}} \left(\frac{t_{\mathbf{m}}}{\mathbf{m}!}\right)^{s_i(\mathbf{m})} \cdot \sigma_{i,1} \mathbf{x}^{\vec{\sigma}_i}.$$

Hence

$$G(\mathbf{x})^{n_i} - \frac{x_1}{n_i} \cdot \frac{\partial}{\partial x_1} G(\mathbf{x})^{n_i} = \sum_{s_i} \left(1 - \frac{\sigma_{i,1}}{n_i}\right) \binom{n_i}{s_i(\mathbf{m})} \prod_{\mathbf{m}} \left(\frac{t_{\mathbf{m}}}{\mathbf{m}!}\right)^{s_i(\mathbf{m})} \mathbf{x}^{\vec{\sigma}_i}.$$

Apply now the Frobenius operator on both sides.

$$\mathcal{F}_i \left(G(\mathbf{x})^{n_i} - x_1 \cdot \frac{\partial G(\mathbf{x})}{\partial x_1} \cdot G(\mathbf{x})^{n_i - 1} \right) = \sum_{s_i} \left(1 - \frac{\sigma_{i,1}}{n_i}\right) \binom{n_i}{s_i(\mathbf{m})} \prod_{\mathbf{m}} \left(\frac{t_{\mathbf{m}}}{\mathbf{m}!}\right)^{s_i(\mathbf{m})} \mathbf{x}^{\mathcal{V}_i(\vec{\sigma}_i)}.$$

But the left hand side is equal to $H_i(\mathbf{x})$. Thus we have

$$H_i(\mathbf{x}) = \sum_{s_i} \left(1 - \frac{\sigma_{i,1}}{n_i}\right) \binom{n_i}{s_i(\mathbf{m})} \prod_{\mathbf{m}} \left(\frac{t_{\mathbf{m}}}{\mathbf{m}!}\right)^{s_i(\mathbf{m})} \mathbf{x}^{\mathcal{V}_i(\vec{\sigma}_i)}.$$

Look now at the coefficient of $\mathbf{x}^{\mathcal{V}_i(\vec{\sigma}_i)}$. Thus the above equation will be

$$\left[\mathbf{x}^{\mathcal{V}_i(\vec{\sigma}_i)}\right] H_i(\mathbf{x}) = \sum_{s_i} \left(1 - \frac{\sigma_{i,1}}{n_i}\right) \binom{n_i}{s_i(\mathbf{m})} \prod_{\mathbf{m}} \left(\frac{t_{\mathbf{m}}}{\mathbf{m}!}\right)^{s_i(\mathbf{m})},$$

where we are summing over s_i which satisfies $\sum_{\mathbf{m}} s_i(\mathbf{m}) \cdot \mathbf{m} = \vec{\sigma}_i$.

Multiply the above identity for all $i \in J$, and we receive that

$$\prod_{i \in J} \left[\mathbf{x}^{\mathcal{V}_i(\vec{\sigma}_i)}\right] H_i(\mathbf{x}) = \prod_{i \in J} \sum_{s_i} \left(1 - \frac{\sigma_{i,1}}{n_i}\right) \binom{n_i}{s_i(\mathbf{m})} \prod_{\mathbf{m}} \left(\frac{t_{\mathbf{m}}}{\mathbf{m}!}\right)^{s_i(\mathbf{m})}.$$

Sum now both sides over all $(\vec{\sigma}_1)_{i \in \mathfrak{I}}$ such that $\sum_{i \in \mathfrak{I}} \mathcal{V}_i(\vec{\sigma}) = \mathbf{n} - \mathbf{k}$.

$$\sum_{(\vec{\sigma}_i)_{i \in \mathfrak{I}}} \prod_{i \in J} \left[\mathbf{x}^{\mathcal{V}_i(\vec{\sigma}_i)}\right] H_i(\mathbf{x}) = \sum_{(\vec{\sigma}_i)_{i \in \mathfrak{I}}} \prod_{i \in J} \sum_{s_i} \left(1 - \frac{\sigma_{i,1}}{n_i}\right) \binom{n_i}{s_i(\mathbf{m})} \prod_{\mathbf{m}} \left(\frac{t_{\mathbf{m}}}{\mathbf{m}!}\right)^{s_i(\mathbf{m})}.$$

We can rewrite this as

$$\begin{aligned} \left[\mathbf{x}^{\mathbf{n}-\mathbf{k}}\right] \prod_{i \in J} H_i(\mathbf{x}) &= \sum_{(\vec{\sigma}_i)_{i \in \mathfrak{I}}} \sum_{(s_i)_{i \in \mathfrak{I}}} \prod_{i \in J} \left(1 - \frac{\sigma_{i,1}}{n_i}\right) \binom{n_i}{s_i(\mathbf{m})} \prod_{\mathbf{m}} \left(\frac{t_{\mathbf{m}}}{\mathbf{m}!}\right)^{s_i(\mathbf{m})} \\ &= \sum_{(s_i)_{i \in \mathfrak{I}}} \prod_{i \in J} \left(1 - \frac{\sigma_{i,1}}{n_i}\right) \binom{n_i}{s_i(\mathbf{m})} \prod_{\mathbf{m}} \left(\frac{t_{\mathbf{m}}}{\mathbf{m}!}\right)^{s_i(\mathbf{m})}, \end{aligned}$$

where the condition on the last sums is $\sum_{i\in\mathfrak{I}} \mathcal{V}_i(\vec{\sigma}_i) = \mathbf{n} - \mathbf{k}$.

By now combining equation (2.10) and equation (2.11) we obtain

$$\left[\mathbf{x}^{\mathbf{n}-\mathbf{k}}\right] \prod_{i\in J} H_i(\mathbf{x}) = [\mathbf{x}^{\mathbf{n}}] \prod_{i\in\mathfrak{I}} \mathcal{F}_i(f(\mathbf{x}))^{k_i}.$$

Since equation (2.9) holds and the above formula holds for all coefficients $t_{\mathbf{m}}$, the plethystic Lagrange inversion formula follows. □

We prove now the theorem analogous by applying Theorem 11, which enumerates unlabeled plane plethystic forests. This proof is very similar to the first proof in this section. Thus we will leave out some details and calculations that is very similar to the previous proof.

Third proof: Assume that

$$G(\mathbf{x}) = \sum_{\mathbf{n}} u_{\mathbf{n}} \mathbf{x}^{\mathbf{n}}.$$

where $t_{\mathbf{n}}$ are the coefficients of $G(\mathbf{x})$, where we view $G(\mathbf{x})$ as an ordinary generating function. Let the $u_{\mathbf{n}}$'s be the weights in the definition of the weight of an unlabeled plane plethystic forest.

We claim that

$$f(\mathbf{x}) = \sum_{\mathbf{n}} \left(\sum_{\tau} \mathbf{u}^{\tau}\right) \mathbf{x}^{\mathbf{n}}, \tag{2.11}$$

where τ ranges over all unlabeled plane plethystic rooted trees on a $\mathbf{n}$ set, and the root has color 1. Hence

$$f(\mathbf{x})^{k_i} = \sum_{\mathbf{n}} \left(\sum_{\pi} \mathbf{u}^{\pi}\right) \mathbf{x}^{\mathbf{n}},$$

where π ranges over all unlabeled plane plethystic rooted forests on a $\mathbf{n}$ set, and the forest .ave k_i roots of color 1. By applying the Frobenius operator we get

$$\mathcal{F}_i(f(\mathbf{x}))^{k_i} = \sum_{\mathbf{n}} \left(\sum_{\pi} \mathbf{u}^{\pi}\right) \mathbf{x}^{\mathbf{n}},$$

where π ranges over all unlabeled plane plethystic rooted forests on a $\mathbf{n}$ set, and the forests have k_i roots of color i. Hence by multiplying

$$\prod_{i\in\mathfrak{I}} \mathcal{F}_i(f(\mathbf{x}))^{k_i} = \sum_{\mathbf{n}} \left(\sum_{\pi} \mathbf{u}^{\pi} \right) \mathbf{x}^{\mathbf{n}}, \tag{2.12}$$

where π ranges over all unlabeled plane plethystic rooted forests on a $\mathbf{n}$ set, and the forests have k_i roots of color i, for $i \in \mathfrak{I}$. Thus

$$t_{\mathbf{k}} x_1 \prod_{i\in\mathfrak{I}} \mathcal{F}_i(f(\mathbf{x}))^{k_i} = \sum_{\mathbf{n}} \left(\sum_{\tau} \mathbf{u}^{\tau} \right) \mathbf{x}^{\mathbf{n}},$$

where τ ranges over all unlabeled plane plethystic rooted trees on a $\mathbf{n}$ set, and the root has color 1 and degree $\mathbf{k}$. By summing over all $\mathbf{k}$, we conclude that

$$\sum_{\mathbf{k}} t_{\mathbf{k}} x_1 \prod_{i\in\mathfrak{I}} \mathcal{F}_i(f(\mathbf{x}))^{k_i} = f(\mathbf{x}).$$

This is equivalent to

$$x_1 \cdot (G \circ f)(\mathbf{x}) = f(\mathbf{x}). \tag{2.13}$$

Since the above equation has a unique solution, we know that the solution is given by (2.11).

Equation (2.12) can be written as

$$[\mathbf{x}^{\mathbf{n}}] \prod_{i\in\mathfrak{I}} \mathcal{F}_i(f(\mathbf{x}))^{k_i} = \sum_{\pi} \mathbf{u}^{\pi}.$$

By Theorem 11 we can rewrite the right hand side, thus

$$[\mathbf{x}^{\mathbf{n}}] \prod_{i\in\mathfrak{I}} \mathcal{F}_i(f(\mathbf{x}))^{k_i} = \sum_{(r_i)_{i\in\mathfrak{I}}} \prod_{i\in J} \left(1 - \frac{\rho_{i,1}}{n_i}\right) \binom{n_i}{r_i(\mathbf{m})} \prod_{\mathbf{m}} t_{\mathbf{m}}^{r_i(\mathbf{m})}, \tag{2.14}$$

where the $(r_i)_{i\in I}$ in the sum fulfill the conditions (2.5) and (2.6). Recall that $\rho_{i,1}$ is the first component of $\vec{\rho}_i$.

By a very similar computation as in previous proof, we obtain

$$\left[\mathbf{x}^{\mathbf{n}-\mathbf{k}}\right] \prod_{i \in J} H_i(\mathbf{x}) = \sum_{(s_i)_{i \in J}} \prod_{i \in J} \left(1 - \frac{\sigma_{i,1}}{n_i}\right) \binom{n_i}{s_i(\mathbf{m})} \prod_{\mathbf{m}} t_{\mathbf{m}}^{s_i(\mathbf{m})},$$

where the condition on the last sums is $\sum_{i \in J} \mathcal{V}_i(\vec{\sigma}_i) = \mathbf{n} - \mathbf{k}$. Thus it follows that

$$\left[\mathbf{x}^{\mathbf{n}-\mathbf{k}}\right] \prod_{i \in J} H_i(\mathbf{x}) = [\mathbf{x}^{\mathbf{n}}] \prod_{i \in J} \mathcal{F}_i(f(\mathbf{x}))^{k_i},$$

and the plethystic Lagrange inversion formula follows. □

Bibliography

[B] F. Bergeron, Une combinatoire du pléthysme, *J. Combin. Theory ser. A* **46**, No. 2 (1987), 291-305.

[Bj-Lo] A. Björner and L. Lovász, Pseudomodular lattices and continuous matroids, *Acta Scientiarum mathematicarum Szeged,* vol 51 (1987), 295-308.

[Bl-Lam] A. Blokhuis and C. W. H. Lam, More Coverings by Rook Domains, *J. Combin. Theory ser. A* **36**, No. 2 (1984), 240-244.

[B-R-S-V] F. Bonetti, G.-C. Rota, D. Senato and A. Venezia, On the foundations of combinatorial theory X: a categorical setting for symmetric functions, to appear.

[C1] W.Y.C. Chen, Ph. D. thesis, M.I.T. (1991).

[C2] W.Y.C. Chen, Context-free Grammars, Differential Operators and Formal Power Series, *Theoretical Computer Science*, 1993.

[Cohn] P.M. Cohn, thesis, "Integral modules, Lie-rings and free groups", Trinity College, Cambridge, 1951.

[E-R1] R. Ehrenborg and G.-C. Rota, Apolarity and Canonical Forms for Homogeneous Polynomials, *European Journal of Combinatorics* **14**, No. 3, May 1993.

[E-R2] R. Ehrenborg and G.-C. Rota, Lectures on Polynomials, to appear (1992).

[F-R] H. Fernandes and E. Rechtschaffen, The Football Pool Problem for 7 and 8 Matches, *J. Combin. Theory ser. A* **35**, No. 1 (1983), 109-114.

[Ge] I. M. Gessel, A Combinatorial Proof of the Multivariable Lagrange Inversion Formula, *J. Combin. Theory ser. A* **45**, No. 2, (1987), 178-195.

[Ge-Lab] I. M. Gessel and G. Labelle, Lagrange inversion for species, to be published.

[Go] I. J. Good, Generalization to several variables of Lagrange's expansion, *Proc., Cambridge Philo. Soc.* **56**, (1960) 367-380.

[G-Y] J. H. Grace and A. Young, "The algebra of invariants," Chelsea, New York, 1903.

[G-R-S] F. D. Grosshans, G.-C. Rota, J. A. Stein, "Invariant Theory and Superalgebras," AMS 69.

[Grö] W. Gröbner, "Moderne algebraische Geometrie," Springer Verlag, Vienna, 1949.

[H-S] M. Haiman and W. Schmitt, Incidence Algebra antipodes and Lagrange inversion in one and several variables, *J. Combin. Theory ser. A* **50** (1989), 172-185.

[I] A. W. Ingleton, Representations of Matroids, "Combinatorial Mathematics and its Applications," edited by D. J. A. Welsh, Academic Press London 1971.

[Joni] S. A. Joni, Lagrange Inversion in Higher Dimensions and Umbral Operators, *Linear and Multilinear Algebra*, Vol. 6, (1978), 111-121.

[J] A. Joyal, Une théorie combinatorie des séries formelles, *Adv. in Math.* **42** (1981), 1-82.

[K-L] H. J. L. Kamps and J. H. van Lint, The Football Pool Problem for 5 Matches, *J. Combin. Theory ser. A* **3** (1967), 315-325.

[K-R] J. P. S. Kung and G.-C. Rota, The Invariant Theory of Binary Forms, *Bulletin of the American Mathematical Society* **10**, (Jan. 1984), 27-85.

[Lab1] G. Labelle, Une nouvelle démostration combinatoire des formules d'inversion de Lagrange, *Adv. in Math.* **42** (1981), 217-247.

[Lab2] G. Labelle, Some new computational methods in the theory species, dans Combinatoire Énumerative, dans **37**, 192-209.

[Lab-Y] G. Labelle and Y. N. Yeh, The relation between Burnside Rings and Combinatorial Species, *J. Combin. Theory ser. A* **50** (1989), 269-284.

[Lag] L. de Lagrange, Nouvelle méthode pour résoudre des équations littérales par le moyen des séries, *Mém. Acad. Roy. Sci. Belles-Lettres de Berlin* **24** (1770).

[Las] E. Lasker, Zur Theorie der kanonische Formen, *Mathematische Annalen*, vol 58, (1904), pp. 434-440.

[Lind] B. Lindström, A desarguesian theorem for algebraic combinatorial geometries, *Combinatorica,* **5** (1985), 237-239.

[Lint] J. H. van Lint, "Introduction to Coding Theory," Springer-Verlag, 1992.

[Lo] G. Losey, Note on a Theorem of Zaremba, *J. Combin. Theory ser. A* **6** (1969), 208-209.

[Ma] S. MacLane, A Lattice formulation for transcedence degrees and p-bases, *Duke Math. J.* **4** (1938), 455-468.

[M-N] M. Mendez and O. Nava, Colored Species, C-monoids and Plethysm (I), to appear in *J. Combin. Theory ser. A*.

[M-Y] M. Mendez and J. Yang, Möbius species, *Adv. in Math.* **85** (1991), 83-128.

[Me] F. Meyer, "Apolarität und rationale Curven," Franz Fues, Tübingen, 1883.

[N] O. Nava, On the combinatorics of plethysm, *J. Combin. Theory ser. A* **46** (1987), 212-251.

[N-R] O. Nava and G.-C. Rota, Plethysm, categories, and combinatorics, *Adv. in Math.* **58** (1985), 86-88.

[Ra] G. N. Raney, Functional composition patterns and power series reversion, *Trans. Amer. Math. Soc.* **94** (1960), 441-451.

[Re] B. Reznick, Sums of Even Powers of Real Linear Forms, *Memoirs of the American Mathematical Society,* vol. 96, no. 463, March 1992.

[Ri] H. W. Richmond, On Canonical Forms, *Quarterly Journal of Math.*, Vol. 33 (1902), 331-340.

[Ro] E. R. Rodemich, Coverings by Rock Domains, *J. Combin. Theory ser. A* **9** (1970), 117-128.

[Ry] H. J. Ryser, "Combinatorial Mathematics," The Carus Mathematical Monographs, No. 14.

[Se1] B. Segre, "The Non-Singular Cubic Surfaces," Oxford University Press, 1942.

[Se2] B. Segre, "Prodromi di geometria algebrica", Cremonese, Roma, 1971.

[St] R. P. Stanley, "Enumerative Combinatorics," vol. 2. To appear.

[T] H. W. Turnbull, "The theory of determinants, matrices and invariants," Dover, New York, 1960.

[Wa] E. K. Wakeford, On Canonical Forms, *Proceedings of the London Mathematical Society*, 2, 18 (1918-1919), pp. 403-410.

[We] E. W. Weber, On the Football Pool Problem for 6 Matches: A New Upper Bound, *J. Combin. Theory ser. A* **35**, No. 1 (1983), 106-108.

[Wi] L. T. Wille, The Football Pool Problem for 6 Matches: A New Upper Bound Obtained by Simulated Annealing, *J. Combin. Theory ser. A* **45**, No. 2 (1987), 171-177.