Lagrange Inversion

Diplomarbeit

Verfaßt von

Markus Rosenkranz

Matr.-Nr. 9057164
Lindenstraße 12
4600 Wels

und eingereicht bei

Univ.-Doz. Dr. Peter Paule

Institut für Mathematik / RISC
Technisch-Naturwissenschaftliche Fakultät
Johannes Kepler Universität Linz

im

Februar 1997

zur Erlangung des akademischen Grades
eines Diplom-Ingenieurs in der
Studienrichtung Technische Mathematik.

Dedicated to my dear family.

Contents

For which of you, intending to build a tower, sitteth not down first, and counteth the cost, whether he have sufficient to finish it? Lest haply, after he hath laid the foundation, and is not able to finish it, all that behold it begin to mock him.
Luke 14:28–29

Preface

This diploma thesis analyzes the *phenomenon of Lagrange inversion*, also known as Lagrange's theorem, first published by the French mathematician Joseph Louis Lagrange in 1869. Its development in the following 125 years has shown to be enormously fertile in numerous branches of mathematics, revealing deep connections in the proofs and applications thereof. Trying to exhaust this depth would be an utterly futile attempt. Therefore our goal is to simply lead the reader to some of the most beautiful spots so he can have a glimpse of this fine mesh and taste the intellectual joy of understanding the underlying structures.

In *Chapter 1*, we set the stage by sketching the combinatorial framework into which Lagrange's theorem is embedded. The theorem itself is presented and investigated from various sides in *Chapter 2*. It is put to work in *Chapter 3* for solving a vast array of different problems, ranging from short stand-alone problems to an extensive theory of certain transformations.

All *formal units* (definition, notation, lemma, proposition, theorem, corollary, proof, example) are numbered in a single sequence. As labels, the numbers appear in the margin so that they can be located easily. For reasons of style, however, they are placed behind when used as references, like in "Definition 5". All formal units are terminated by a □ symbol, which is also put in the margin for better readability.

Occasionally I introduce some *special names* that are not found in the literature. Such terms are marked by a circle-superscript° when they first appear.

The *chapter mottos* are taken from the Bible (King James Version). The Scriptures often use metaphors from the natural world in order to illustrate spiritual truths. We quote these verses here because they contain universal principles that are also valid in the world of mathematics.

He is like a man which built an house, and digged deep, and laid the foundation on a rock: and when the flood arose, the stream beat vehemently upon that house and could not shake it: for it was founded upon a rock.
Luke 6:48

Chapter 1

Introduction

Before we are able to prove or even state Lagrange's famous theorem, some *preparatory work* is necessary. In Section 1.1 the reader is made familiar with the wide domain of combinatorics, probably the richest field linked with the theorem. In combinatorics, more than in other branches of mathematics, it is important to think about the concept of a solution; this is done in Section 1.2. Carefully defining the relevant concepts, we can finally lay a firm foundation for the proofs given in the next chapter. For practical reasons, the general concepts in Section 1.3 are separated from the combinatorial ones in Section 1.4.

1.1 The Combinatorial Context

In order to understand the great relevance of Lagrange inversion, we shall give a short informal overview of where it arises in the field of *combinatorics*. It seems to be rather difficult, though, to precisely define what combinatorics itself is, as some mathematicians like Halder and Heise [6] quite pessimistically point out. But apparently there are some promising approaches, especially the one taken by Aigner [1, p. vii–xiii] in his comprehensive work. He defines combinatorics as "counting and ordering morphisms", thus dividing it into the theory of counting (enumerative combinatorics) and the theory of ordering (existential and constructive combinatorics). In any case, finite structures (like permutations, graphs, Latin squares, projective geometries—just to mention a few) are considered: for the simpler ones the combinatorialist aims at counting them; for the more complicated ones he is usually content when he knows about their existence or can construct certain patterns on them.

It should also be mentioned that in the past years *computer methods* have become more and more important. On the one hand, the aspect of algorithmic construction (e.g. by Polya theory) has brought many significant advances in constructive combiatorics. On the other hand, the so-called symbolic part of enumerative combinatorics (dealing with formula manipulation) is extensively discussed nowadays. It is one of the reasons why the concept of a solution has to be reconsidered; see Section 1.2.

Our subject matter mostly shows up in enumerative combinatorics, which could also be called the *art of counting*. The typical elementary counting problems are: In how many ways can I arrange the letters A through Z ("permutations")? How many lotto drawings of 6 numbers out of 49 are there ("combinations")? Or how many numbers are there with 4 even digits ("variations")? These are the classical and also the simplest counting functions. Most of them depend upon one or two parameters; here the general terms are: permutations of n objects, combinations of k out of n objects, and variations of k out of n objects. It also makes a difference whether one allows repetitions or not. Then there is a huge variety of more complicated questions like: How many necklaces of 12 beads are there, using 3 red ones, 5 blue ones, and 4 green ones? All of these questions can be reduced to counting the number of functions (also called morphisms) from some domain N to some range R, imposing various restrictions on them and possibly considering certain elements of N or R, specified by suitable groups, as indistinguishable. A very nice illustrating example is provided by Rota's Twelvefold Way, see [25]. For more details, look up e.g. [1].

Since the number of such objects is always a quantity which depends on one or more parameters they can be viewed as sequences in one or more indices. So for the permutations of n objects we might write p_n, for the combinations of k out of n objects c_{nk}, and so on. For computing these numbers, one usually derives some connection between one particular instance and its predecessors by counting how many objects can be composed of smaller ones. The resulting relation is often called *recurrence*. Since this is only a relative characterization, one has to include enough sample values (called initial conditions) to uniquely fix the sequence.

1 Example (Recurrence for the Combinations)
The combinations, as an example, satisfy the recurrence

$$\begin{aligned}[\forall n \geq 0]\ c_{n0} &= 1,\\ [\forall k \geq 1]\ c_{0k} &= 0,\\ [\forall n, k \geq 1]\ c_{nk} &= c_{n-1,k} + c_{n-1,k-1}.\end{aligned}$$

The initial conditions are evident since there is always precisely one way to choose an empty set (first line) and it is impossible to choose any elements *from* an empty set (second line). The recurrence (third line) can be obtained by the following consideration. How many combinations of k out of n elements are there? We mark one of the n elements and call it the "special member". Without this element, there are $c_{n-1,k}$ possible combinations because this is like choosing k from the $n-1$ non-special elements. If we include the special member, we can construct $c_{n-1,k-1}$ combinations by choosing the other $k-1$ elements, again from the $n-1$ non-special ones. Since there are no other cases, this makes a total of $c_{n-1,k} + c_{n-1,k-1}$ combinations of k out of n elements. □

This has been the state of affairs until the end of the 18th century. Many recurrences turned up, some of them could be solved in explicit terms, others resisted all efforts. At any rate, mathematicians did not have a generic method for this type of problem. Not until Laplace set forth the fundamental idea of a *generating function* in his work on probability theory. The crucial idea is to consider the sequence in question as the coefficients of a power series. Any operation on the sequence can then be interpreted as acting on the power series, thereby transforming the recurrence into a functional or differential equation for the power series. Often the series can be computed easily (compared to the recurrence) and one can calculate the original sequence from it.

2 Example (A Generating Function for the Combinations)
In our example, the number of combinations c_{nk}, we can decide on which index to sum for the representation as a power series. We could take both and define $F(x, y) = \sum_{n,k=0}^{\infty} c_{nk} x^n y^k$, or keep n as a free parameter and define $G_n(y) = \sum_{k=0}^{\infty} c_{nk} y^k$, or else leave k free and define $H_k(x) = \sum_{n=0}^{\infty} c_{nk} x^n$. In this example, all three of them lead to the solution; we choose $G_n(y)$.

Now c_{0k} is always 0, except for $c_{00} = 1$, therefore we have

$$G_0(y) = \sum_{k=0}^{\infty} c_{0k}\, y^k = 1.$$

In order to obtain a functional equation for $G_n(y)$, we simply apply the recurrence to its coefficient and get

$$[\forall n \geq 1]\ G_n(y) = \sum_{k=0}^{\infty} c_{nk}\, y^k = 1 + \sum_{k=1}^{\infty} (c_{n-1,k} + c_{n-1,k-1})\, y^k$$
$$= 1 + \sum_{k=1}^{\infty} c_{n-1,k}\, y^k + \sum_{k=0}^{\infty} c_{n-1,k}\, y^{k+1} = G_{n-1}(y) + y\, G_{n-1}(y).$$

This equation, together with $G_0(y) = 1$, can be solved in explicit terms and yields

$$[\forall n \geq 0]\ G_n(y) = (1+y)\, G_{n-1}(y) = \ldots = (1+y)^n\, G_0(y) = (1+y)^n.$$

In principle, we have found the desired sequence c_{nk} now (see Section 1.2)! But it is somehow disguised in our representation of the generating function $G_n(y)$. We can unmask it by applying the binomial theorem to expand $G_n(y)$ into the underlying power series

$$[\forall n \geq 0]\ G_n(y) = (1+y)^n = \sum_{k=0}^{n} \binom{n}{k} y^k,$$

□ which gives us the desired explicit form $c_{nk} = \binom{n}{k}$.

In our example, to every operator on the sequence c_{nk}, there corresponds an operator on the power series $G_n(y)$: The sequence operator $c_{nk} \to c_{n,k-1}$ has become $G_n(y) \to yG_n(y)$ on the power-series side; addition remained addition. By such correspondences, one can translate virtually any recurrence into an adequate functional equation for the generating function. In order to facilitate the task of such translations, Wilf [27, p. 33–39] has collected a useful set of rules, which he calls the *calculus of formal power series*. He is so fanatic about this approach that he even coined the monster name "generatingfunctionology" for it (also the title of his book).

It turns out that for "real-life applications", one needs a well-equipped toolbox for handling power series. One type of key problem occurring frequently (and many problems can be transformed into it) is the *inversion* of power series. This means if $f(z)$ is a power series, its inverse is $g(z)$ iff $f(g(z)) = g(f(z)) = z$. For a precise formal definition, please refer to Section 1.3. As a trivial example (see Example 78), the inverse of the series $e^z - 1$ is the $\ln(1+z)$. Often one cannot find such a neat closed form, though, and it is only possible to calculate the coefficients of the inverse series.

This is where *Lagrange inversion* (also known as Lagrange's theorem) enters the stage: it allows to nail down these coefficients in some way. But not only that, it can even solve the more general problem of determining the coefficients c_k for expanding a given power series $f(z)$ in terms of another given power series $g(z)$ such that

$$f(z) = \sum_{k=0}^{\infty} c_k \, g(z)^k.$$

Inversion turns out to be the special case $f(z) = z$. In Example 73, we will present a classical problem where one needs to resort to Lagrange's Theorem for such a generalized inversion.

Over the years, combinatorialists have singled out a variety of number sequences of special combinatorial interest like the combinations c_{nk} in Example 2. Often they have also attached some special symbols to them; the binomial $\binom{n}{k}$ for c_{nk} is such an instance. Most of these numbers (also called counting functions) turn out to be unrepresentable in terms of "standard functions". (Confer to Section 1.2 for a brief discussion of this topic.) Therefore, it is all the more desirable to know some *combinatorial identities* interrelating them. Section 1.2 for a brief discussion of this topic.

3 Example (An Identity for Second-Kind Stirling Numbers)
Let us consider two simple counting functions, the falling factorials and the ordinary powers. We will establish a combinatorial identity relating the two by yet another counting function.

The *falling factorials* arise when we count the number of possible words built with k out of l letters without repetition. For the first letter, we can choose any of the l available ones, for the second there are $l-1$ choices left, ..., for the last letter we have got $l-k+1$ possibilities. This gives a total of $l(l-1)\dots(l-k+1)$ choices. The latter expression is known as falling factorial, written $l^{\underline{k}}$ and spoken "l to the k falling". (For reasons of convenience, one additionally set $l^{\underline{0}} = 1$.) Relaxing the restriction on repetitions, we arrive at the *ordinary powers*: there are l^k possible words using k out of l letters, since we can choose any of the l available letters at any of the k places.

For finding the relation between $l^{\underline{k}}$ and l^k, we have to introduce one more counting function, namely $\left\{ {n \atop k} \right\}$, the number of k-partitions (partitions having k blocks) of an n-set (a set with n elements). This $\left\{ {n \atop k} \right\}$ is called *Stirling* *number of the second kind* and we read it "n subset k." As an example, the 3-partitions of the set $\{1, 2, 3, 4\}$ are

$$\{1, 2|3|4\}, \{1, 3|2|4\}, \{1, 4|2|3\}, \{2, 3|1|4\}, \{2, 4|1|3\}, \{3, 4|1|2\},$$

thus we see that $\left\{ {4 \atop 3} \right\} = 6$. There is no closed form for the $\left\{ {n \atop k} \right\}$ available; only formulas involving summation are known. Here we do not need any explicit result, though. It is sufficient that they are uniquely defined and one could in principle compute them for fixed n and k.

The desired relation can now easily be obtained by constructing all of the l^n possible n-words on l letters in the following way: First we decide on a set of k distinct letters, which can be done in $\binom{l}{k}$ different ways. For each of these k letters we choose a position between 1 and n; this gives $k! \left\{ {n \atop k} \right\}$ different selections, since any of them can be viewed as an ordered k-partition of $\{1, \ldots, n\}$. Summing over all admissible values for k and utilizing that $\binom{l}{k} = l^{\underline{k}}/k!$, we finally arrive at

$$l^n = \sum_{k=0}^{n} \binom{l}{k} \left\{ {n \atop k} \right\} k! = \sum_{k=0}^{n} \left\{ {n \atop k} \right\} l^{\underline{k}},$$

which is an identity between the ordinary powers and the second-kind Stirling numbers. □

The method used in this example for finding the identity is known as a combinatorial or *bijective proof* because it bases on some combinatorial interpretations for the occurring numbers and then establishes some bijection among them. In Section 3.4, we will learn of a method for finding such relations without using any combinatorial interpretation. The falling factorials turn up again there, when we derive an interesting identity connecting them with the so-called rising factorials (which will be introduced later) in Example 100.

Viewing these identities, one question arises just naturally: Given such a relation, is it possible to find the *inverse relation*? This is often desirable because the initial relation might be easy to find whereas the inverse relation could be more difficult to handle. In our example, it was not too difficult to express l^n in terms of $l^{\underline{n}}$ but it turns out to be somewhat harder to represent the falling factorials by ordinary powers. And yet this will more likely be needed, for instance, in order to represent $l^{\underline{n}}$ as a polynomial in expanded form.

The general form of a pair of inverse relations is

$$[\forall n \geq 0] \;\; a_n = \sum_{k=0}^{n} \alpha_{nk} b_k \;\Leftrightarrow\; b_n = \sum_{k=0}^{n} \beta_{nk} a_k,$$

where α_{nk} and β_{nk} are the so-called *connecting coefficients*, a_n and b_n are arbitrary sequences. We will see in Section 3.3 that we can regard the sequences as vectors and the connecting coefficients as matrices. So the task of finding the inverse relation can be seen as inverting that matrix.

1.2 What Is a Solution?

For many branches of mathematics, there seems to be no question of what is to be regarded a solution. When we look at an algebraic equation like $z^2 + p\,z + q$, it is completely evident that $z_1 = -p/2 + \sqrt{p^2/4 - q}$ and $z_2 = -p/2 - \sqrt{p^2/4 - q}$ are two solutions, which can certainly be regarded as *closed forms*. It becomes more difficult, however, when the order is higher than five: for general coefficients, there are no solutions in radicals. In that case, we must accept other representations of the solutions—series, integrals, limits, recurrences, etc. In the case of algebraic equations, the Mellin series might be a suitable choice; see Example 80.

But what precisely is a "closed form"? As an illuminating example, take the differential equation $f'(z) = f(z)$. We all know that it has the closed-form solution $f(z) = e^z$. But what is e^z really? It is a transcendental function, which means that we cannot express it by a finite number of additions and multiplications. The only reason why we have picked it out from all the other transcendental functions is that it occurs so often in the applications. Therefore we have given it the name "exponential function", and we have added it to our *repertoire of standard forms*. From a mathematical point of view, however, we have chosen them completely at random. This means that a solution is considered to be in closed form whenever we can represent it in terms of such an arbitrary repertoire.

In the *combinatorial context*, one usually chooses such functions as factorials, binomials or Stirling numbers for the standard repertoire. But the actual choice depends only on the practical demand. Imagine that we want to list the first 200 values of a particular sum. Having an explicit representation, we would have to compute each value from the scratch, whereas the sum representation itself allows to pass from one value to the next by simply adding one term. So why not consider it a closed form?

Most of the problems in combinatorics have no explicit solutions. The question is not so much how to "compute" a sum like $\sum m^2/(m+1)$ or a recurrence like $a_0 = 1, a_m = a_m^2 + 2m$ for $m > 0$. The sum can be computed for any specific upper bound, and the recurrence can be used to calculate a_m for any specific index m. What one usually needs, though, is some special representation, suited for whatever purpose one has in mind. So solving really means being able to *change easily between different representations.* See Peter Paule's paper [16] for further discussion. Paule uses the fundamental notion of canonical simplification to precisely specify what it means to "simplify" a sum. The general concept of a canonical simplifier is described in the classical survey paper [2].

Having various different representations of the same object, another problem arises: How can we check whether two of them are equal? If there is a closed form, we can (in principle) transform the two expressions to that form and check equality. But what can we do if there are no closed forms? This leads us to the problem of *canonical forms.* A canonical form is a unique representation with a procedure for transforming equivalent expressions into that form. It must be clear which class of expressions is covered by a specific canonical form.

The benefit of canonical forms is clearly demonstrated by the example of *hypergeometric* functions. The standard notation (see [5]) is

$$\lambda \, {}_mF_n\binom{a_1,\ldots,a_m}{b_1,\ldots,b_n}; cz\Big) = \sum_{j=0}^{\infty} \frac{a_1^{\overline{j}} \cdots a_m^{\overline{j}}}{b_1^{\overline{j}} \cdots b_n^{\overline{j}}} \lambda \, c^j z^j,$$

where the $r^{\overline{0}} = 1, r^{\overline{j}} = r(r+1)\cdots(r+j-1)$ for $j > 0$ denote the rising factorials, verbalized by saying "r to the j rising".

Before the significance of hypergeometric-function notation (together with the so-called "hypergeometric machinery") was realized, the usual procedure for proving a combinatorial identity was more or less trying to rewrite both sides until they happen to match an entry in some huge identity table. One of the most famous tables is due to Gould, who collected more than 500 binomial-coefficient summations. By standardizing his table to the hypergeometric notation, it becomes considerably smaller and the calculation procedure is much more deterministic.

The hypergeometric functions can be generalized to another highly important class of functions. By looking at the underlying sequence of a hypergeometric function, one observes that they satisfy a homogeneous recurrence equation of order one, where all the coefficients are polynomials in the index variable. Removing the order restriction leads to the class of *P-finite* sequences. Their generating functions are called *holonomic* or *D-finite*. They can also be characterized as the solutions of linear differential equations with polynomial coefficients. This class of functions is so important for applications (especially summation) that it is also called the "holonomic universe".

1.3 Elementary Concepts

To rigorously analyze the subject matter, we must be precise on what terms we use. In this section, we will therefore introduce the basic concepts used throughout the thesis. First let us fix some notations that are not quite standard.

4 **Notation (Basic Concepts)**
As usual, the letters $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$ are used for the sets of natural, integer, rational, real, and complex *numbers*. The natural numbers $\mathbf{N}$ are assumed to include 0, otherwise we write $\mathbf{N}^*$.

Whenever dealing with several variables, *multi-index* notation is applied. So for some dimension $s \in \mathbf{N}^*$, we combine s ordinary indices to an s-fold multi-index $\mathbf{m} = (m_1, \ldots, m_s) \in \mathbf{Z}^s$. Note that we can view an ordinary index m as a multi-index $\mathbf{m} = (m)$. Therefore, we will use the term "index" to mean both ordinary and multifold indices; the same applies to all other concepts in multi-index notation.

There are some common *abbreviations for multi-indices*. We write $|\mathbf{m}|$ for the weight° $m_1 + \ldots + m_s$ of the index. The factorial $m_1! \ldots m_s!$ is abbreviated by $\mathbf{m}!$, the monomial $z_1^{m_1} \cdots z_s^{m_s}$ by $\boldsymbol{z}^{\mathbf{m}}$ for $\boldsymbol{z} = (z_1, \ldots, z_s)$. Addition and scalar multiplication of multi-indices is understood componentwise. The inequality $\mathbf{k} \leq \mathbf{l}$ with $\mathbf{k}, \mathbf{l} \in \mathbf{Z}^s$ means $k_i \leq l_i$ for all $i \in \{1, \ldots, s\}$; the strict inequality $\mathbf{k} < \mathbf{l}$ is short for $\mathbf{k} \leq \mathbf{l}$ and $\mathbf{k} \neq \mathbf{l}$. The Kronecker Delta is defined as usual:

$$\delta_{\mathbf{mn}} = \begin{cases} 1 & \text{for } \mathbf{m} = \mathbf{n}, \\ 0 & \text{else.} \end{cases}$$

Furthermore, we introduce some *multi-index constants* for convenience. We write $\mathbf{0}$ for $(0,\ldots,0)$ and $\mathbf{1}$ for $(1,\ldots,1)$. For any $i \in \{1,\ldots,s\}$, the i-th unit multi-index is defined by $\mathbf{1}_i = (\delta_{i1},\ldots,\delta_{is}) = (0,\ldots,1,\ldots,0)$ with 1 at the i-th place.

Multisequences will be written in angle brackets like $\langle a_{\mathbf{m}} \rangle_{\mathbf{m}\geq\mathbf{1}}$. If the domain is not indicated explicitly like here, it is assumed to be $\mathbf{N}^s$, so $\langle a_{\mathbf{m}} \rangle_{\mathbf{m}}$ is short for $\langle a_{\mathbf{m}} \rangle_{\mathbf{m}\in\mathbf{N}^s}$. The outer index is also dropped if it is clear which variable is the index. Thus we could have written $\langle a_{\mathbf{m}} \rangle$ for the previous sequence. The same conventions are used for sums and other quantifiers. Note that $\langle a_{\mathbf{m}} \rangle$ denotes a sequence, namely the *function* $a \colon \mathbf{Z}^s \to \mathbf{C}$, whereas $a_{\mathbf{m}}$ denotes some *value* in $\mathbf{C}$. □

As we have said in the introductory outline, we will deal with Lagrange inversion mainly in the context of *formal power series*. The term "formal" indicates that, for the time being, we are not concerned with any topological concepts like convergence or continuity. We do not even substitute any value for the variables; therefore they are called *indeterminates*. So writing something like $1 + 2z + 3z^2$ is merely an abbreviation for the sequence $\langle 1, 2, 3, 0, 0, \ldots \rangle$ or, as Wilf [27, p. 1] puts it: "A generating function is a clothesline on which we hang up a sequence of numbers for display." (We will drop the attribute "formal" from now on, since—with one exception in Chapter 2—we will not consider any analytical properties.) The following definition makes this idea precise.

5 Definition (Cauchy Algebra)
Choose an arbitrary $s \in \mathbf{N}^*$. Then the *Cauchy algebra* of s-fold multivariate power series is the vector space of s-fold multisequences over $\mathbf{C}$ (with addition and scalar multiplication as above), equipped with the so-called Cauchy product

$$\langle a_{\mathbf{m}} \rangle \langle b_{\mathbf{n}} \rangle = \left\langle \sum_{\mathbf{m}+\mathbf{n}=\mathbf{k}} a_{\mathbf{m}} b_{\mathbf{n}} \right\rangle_{\mathbf{k}} .$$

The number s is called *domain-dimension* (or just dimension).

Choose some indeterminates $z_1, \ldots, z_s$ to denote the *elementary*° *sequences* $\langle \delta_{\mathbf{m},\mathbf{1}_1} \rangle, \ldots, \langle \delta_{\mathbf{m},\mathbf{1}_s} \rangle$. Representing the elements in the basis induced by these sequences, the Cauchy algebra is denoted by the symbol $\mathbf{C}[[\boldsymbol{z}]]$ or $\mathbf{C}_s[[\boldsymbol{z}]]$ if one likes to explicate the dimension. □

From the elementary sequences, the entire *canonical basis* can be constructed. What *is* the canonical basis for $\mathbf{C}[[\boldsymbol{z}]]$, though? Recall that for the ordinary Euclidean vector space $\mathbf{C}^3$, it is $\{\langle 1,0,0\rangle, \langle 0,1,0\rangle, \langle 0,0,1\rangle\}$. This can also be written as

$$\left\{\langle\delta_{m0}\rangle_{m\in M}, \langle\delta_{m1}\rangle_{m\in M}, \langle\delta_{m2}\rangle_{m\in M}\right\}$$

with $M = \{0,1,2\}$. Likewise, the vector space $\mathbf{C}[[z]]$ of *ordinary sequences* (which is $\mathbf{C}^\infty$ so to say) has canonical basis

$$\left\{\langle\delta_{m0}\rangle_{m\in\mathbf{N}}, \langle\delta_{m1}\rangle_{m\in\mathbf{N}}, \langle\delta_{m2}\rangle_{m\in\mathbf{N}}, \dots\right\}.$$

Precisely speaking, we ought to call it a pseudo-basis: in general, we need an infinite number of basis vectors in order to represent an element. As an example, the sequence $\langle 1,1,1,\dots\rangle$ has to be written as $\langle\delta_{m0}\rangle + \langle\delta_{m1}\rangle + \langle\delta_{m2}\rangle + \dots$ in its basis representation. Note that the basis vectors can be expressed by the single elementary sequence $z = \langle\delta_{m1}\rangle$ because

$$\begin{aligned}
z^0 &= 1 = \langle\delta_{k0}\rangle, \\
z^1 &= z = \langle\delta_{k1}\rangle, \\
z^2 &= zz = \langle\delta_{m1}\rangle\langle\delta_{m1}\rangle = \left\langle\sum_{m+n=k}\delta_{m1}\delta_{n1}\right\rangle_k = \langle\delta_{k,1+1}\rangle = \langle\delta_{k2}\rangle, \\
z^3 &= zz^2 = \langle\delta_{m1}\rangle\langle\delta_{n2}\rangle = \left\langle\sum_{m+n=k}\delta_{m1}\delta_{n2}\right\rangle_k = \langle\delta_{k,1+2}\rangle = \langle\delta_{k3}\rangle,
\end{aligned}$$

and so on. The first line utilizes the convention that z^0 denotes the multiplicative neutral element in a ring.

So for every index $n \in \mathbf{N}$ there is one basis vector $\langle\delta_{mn}\rangle_m$ for $\mathbf{C}[[z]]$. Passing to the s-fold *multisequences* of $\mathbf{C}[[\boldsymbol{z}]]$, the canonical basis again must have one basis vector $\langle\delta_{\boldsymbol{mn}}\rangle_{\boldsymbol{m}}$ for every multi-index $\boldsymbol{n} \in \mathbf{N}^s$. They can be expressed by the elementary sequences $z_1, \dots, z_s$ via $\langle\delta_{\boldsymbol{mn}}\rangle_{\boldsymbol{m}} = \boldsymbol{z}^{\boldsymbol{n}} = z_1^{n_1}\dots z_s^{n_s}$ because the components of $\boldsymbol{z}$ are independent of one another. This can be seen from

$$z_i z_j = \langle\delta_{\boldsymbol{m},\mathbf{1}_i}\rangle\langle\delta_{\boldsymbol{n},\mathbf{1}_j}\rangle = \left\langle\sum_{\boldsymbol{m}+\boldsymbol{n}=\boldsymbol{k}}\delta_{\boldsymbol{m},\mathbf{1}_i}\delta_{\boldsymbol{n},\mathbf{1}_j}\right\rangle_{\boldsymbol{k}} = \langle\delta_{\boldsymbol{k},\mathbf{1}_i+\mathbf{1}_j}\rangle_{\boldsymbol{k}},$$

where $\mathbf{1}_i$ and $\mathbf{1}_j$ do not interfere unless $i = j$.

Any multisequence can be represented in this basis by

$$\langle a_{\boldsymbol{m}}\rangle = \sum a_{\boldsymbol{m}}\boldsymbol{z}^{\boldsymbol{m}}$$

so that the Cauchy product takes on the familiar form

$$\left(\sum a_m z^m\right)\left(\sum b_n z^n\right) = \sum_k \left(\sum_{m+n=k} a_m b_n\right) z^k.$$

Now the *generating function* of a sequence is simply its basis representation. We typically use the letters $f, g, h, \ldots$ for variables of that type. If we want to explicitly point to the indeterminates we write $f(z)$ for f. Speaking of "any" $f \in \mathbf{C}[[z]]$, we always mean an arbitrary s-fold sequence, whose dimension $s \in \mathbf{N}^*$ may also be arbitrary.

Sometimes it is convenient to collect various series into a so-called *vector series* just like one collects various complex numbers into a complex vector. If such a series is interpreted as an analytical function, it operates from $\mathbf{C}^s$ to $\mathbf{C}^r$. The proper notation is introduced below.

6 **Definition (Series-Vector)**
Let $r, s \in \mathbf{N}^*$ and $f_1, \ldots, f_r \in \mathbf{C}_s[[z]]$ be arbitrary. Then the *series-vector*° $\mathbf{f}$ denotes the vector $(f_1, \ldots, f_r)$ and likewise $\mathbf{f}(z)$ denotes $(f_1(z), \ldots, f_r(z))$. The number r is called *range-dimension.*

The corresponding vector space $\mathbf{C}[[z]]^r$ will be written as $\mathbf{C}^r[[z]]$ or $\mathbf{C}^r_s[[z]]$ if one wants to be specific about the domain-dimension, too. (Analogous notations will be used for the vector equivalents of other structures to be
□ defined in the following.)

One very common operation on a generating function is to single out a *specific coefficient.* As we have pointed out above, any s-fold sequence $\langle a_m \rangle$ is actually a function $a\colon \mathbf{N}^s \to \mathbf{C}$. Thus picking a coefficient with a certain index is merely applying this function to the index. It is convenient to have a special symbol for this.

7 **Definition (Coefficient Functional)**
For any $\langle a_m \rangle \in \mathbf{C}[[z]]$ and $n \in \mathbf{N}^s$, we define

$$[z^n] \sum a_m z^m = a_n$$

and call $[z^n]$ a *coefficient functional.* Precedence is ruled by defining

$$[z^n]\ f + g = ([z^n]\ f) + g,$$
$$[z^n]\ fg = [z^n]\ (fg)$$

for all $f, g \in \mathbf{C}[[z]]$.

□ For picking the constant term, we use the symbol $L = [z^0]$.

Among all others, the *finite sequences* are naturally distinguished. Since one can view them as infinite sequences with zero terms from some index on, their generating functions are polynomials, which form a ring (and a subalgebra) within the Cauchy algebra. The polynomials can be characterized by their degree, which tells us the length of the underlying sequence.

8 **Definition (Degree of Power Series)**
For an s-fold power series $f \in \mathbf{C}[[z]]$, we define its *degree* by

$$\deg f = \max\{\mathbf{m} : [z^{\mathbf{m}}]\ f(z) \neq 0 \wedge \mathbf{m} \in \mathbf{N}^s\},$$

where the maximum is to be taken componentwise. If the maximum does
□ not exist we set $\deg f = \infty$.

Now we can give a concise description of polynomials.

9 **Definition (Polynomial Ring)**
The *polynomial ring* is the subalgebra

$$\{f \in \mathbf{C}[[z]] : \deg f < \infty\}$$

□ and denoted by $\mathbf{C}[z]$.

The polynomial ring $\mathbf{C}[z]$ arises as a specialization of the Cauchy algebra $\mathbf{C}[[z]]$. Our next construct will be a generalization of it, because there is still a restriction in $\mathbf{C}[[z]]$ that can be disadvantageous at times: namely, not all elements have a *reciprocal* (an inverse element with respect to the Cauchy product). In terms of algebra, $\mathbf{C}[[z]]$ is only an integral domain but not a field; by extending it, we obtain its quotient field. But before we turn to introduce this field, let us first give a characterization of the power series that do have a reciprocal within $\mathbf{C}[[z]]$. For this purpose, we define another concept, completely analogous to the degree.

10 **Definition (Order of Power Series)**
For an s-fold power series $f \in \mathbf{C}[[z]]$, we define its *order* by

$$\operatorname{ord} f = \min\{\mathbf{m} : [z^{\mathbf{m}}]\ f(z) \neq 0 \wedge \mathbf{m} \in \mathbf{N}^s\},$$

□ where the minimum is to be taken componentwise.

It is the order of a power series f that determines whether or not a reciprocal exists.

11 Proposition (Reciprocal Condition)
A power series $f \in \mathbf{C}[[z]]$ has a reciprocal iff $\operatorname{ord} f = 0$; in this case, the reciprocal is unique. □

12 Proof (Reciprocal Condition)
Let $f(z) = \sum_m a_m z^m$ be an arbitrary power series that has a reciprocal $g(z) = \sum_m b_n z^n$. Then we have

$$fg = \sum_k \left(\sum_{m+n=k} a_m b_n \right) z^k = 1,$$

which is equivalent to

$$[\forall k > 0] \sum_{0 \le m \le k} a_m b_{k-m} = a_0 b_k + \sum_{0 < m \le k} a_m b_{k-m} = 0$$

and $a_0 b_0 = 1$. The latter implies that $a_0 \neq 0$ and hence $\operatorname{ord} f = 0$. Therefore we can uniquely determine the coefficients of the reciprocal g from the recurrence

$$\begin{aligned} b_0 &= 1/a_0, \\ [\forall k > 0] b_k &= -\frac{1}{a_0} \sum_{0 < m \le k} a_m b_{k-m}. \end{aligned}$$

For the converse, suppose $f \in \mathbf{C}[[z]]$ is an arbitrary power series with $\operatorname{ord} f = 0$. Then $a_0 \neq 0$ and hence the above recurrence again determines a unique power series $g \in \mathbf{C}[[z]]$ which is, by its definition, the reciprocal of f. □

Since the reciprocal of a power series f is unique (if it exists) we can assign a symbol for it.

13 Notation (Reciprocal Symbol)
If f is a power series with $\operatorname{ord} f = 0$ we denote its *reciprocal* by $1/f$ and also by f^{-1}. The powers of reciprocals are written as $1/f^k$ and f^{-k} accordingly. □

Now we know which elements of the algebra $\mathbf{C}[[z]]$ have a reciprocal. But how can we extend this algebra to make all elements invertible with respect to the Cauchy product? Take the power series $f(z) = z$ as an example. Since $\operatorname{ord} f = 1$ it cannot have a reciprocal within $\mathbf{C}[[z]]$. But we know that in elementary analysis there is a reciprocal $1/f(z) = z^{-1}$ which is defined in $\mathbf{R} \setminus \{0\}$. This suggests that we will have to admit negative exponents in our extended version of series, which are known as *Laurent series* ("formal Laurent series", to be precise). Indeed, we observe that

$$\operatorname{ord} fg = \operatorname{ord} f + \operatorname{ord} g$$

for any univariate Laurent series f and g. This means especially that $\operatorname{ord}(1/f) = -\operatorname{ord} f$ since $\operatorname{ord} f\,(1/f) = \operatorname{ord} 1 = 0$. So for any Laurent series f we know that either f itself or its reciprocal 1/f are ordinary power series. In the special case $\operatorname{ord} f = 0$, both are ordinary power series—as we already know from Proposition 11. From now on, we will reserve the term "power series" for the ordinary ones, namely those with nonnegative order.

Our generalization aims at expanding the index range from $\mathbf{N}^s$ to $\mathbf{Z}^s$. But actually we do not need all of $\mathbf{Z}^s$. Our goal is to include the reciprocals of power series. Now every power series has a finite, nonnegative order. Therefore the Laurent series will also have a finite order, but no restriction on the sign. We are now ready to put this concept into formal terms.

14 Definition (Laurent Field)
An s-fold Laurent series is a multisequence $\langle a_{\mathbf{m}} \rangle_{\mathbf{m} \geq \mathbf{n}}$ for some $\mathbf{n} \in \mathbf{Z}^s$. The set of all such series with the Cauchy product as in Definition 5 is called *Laurent field* and written as $\mathbf{C}((\boldsymbol{z}))$ when using the basis representation as introduced above.

All the terms introduced so far (like the coefficient functional or the order of a series) are adopted accordingly from the above definitions and notations by replacing $\mathbf{C}[[\boldsymbol{z}]]$ by $\mathbf{C}((\boldsymbol{z}))$. □

Just as we have defined L to pick out the constant term of a power series, we will define an operator on Laurent series to pick out the coefficient of order $-\mathbf{1}$. Just like in complex analysis, it will turn out to be highly useful later on.

15 Notation (Residue Symbol)
The coefficient functional $[\boldsymbol{z}^{-\mathbf{1}}]$ on $\mathbf{C}((\boldsymbol{z}))$ is called the *residue* and written as M. □

Note the following *absorption property*° of the coefficient functional, which allows us to rewrite the residue symbol: For any $f \in \mathbf{C}((\boldsymbol{z}))$ and any $\mathbf{k}, \mathbf{l} \in \mathbf{Z}^s$, we have $[\boldsymbol{z}^{\mathbf{l}}]\, f(\boldsymbol{z}) = [\boldsymbol{z}^{\mathbf{k}+\mathbf{l}}]\, \boldsymbol{z}^{\mathbf{k}} f(\boldsymbol{z})$ and especially $\mathrm{M} f(\boldsymbol{z}) = [\boldsymbol{z}^{\mathbf{k}-\mathbf{1}}]\, \boldsymbol{z}^{\mathbf{k}} f(\boldsymbol{z})$ when we set $\mathbf{l} = -\mathbf{1}$.

There is one more structure to introduce, which will complete our collection. We have arrived at $\mathbf{C}((\boldsymbol{z}))$ by extending the ring $\mathbf{C}[[\boldsymbol{z}]]$ to a field.

Likewise we can extend the polynomial ring $\mathbf{C}[z]$ to a field, which is denoted by $\mathbf{C}(z)$. Since all the elements are quotients of polynomials we recognize them as *rational functions*. (To be exact, one ought to call them "formal rational functions".)

16 Definition (Rational Function Field)
The *rational function field* is the subalgebra

$$\left\{ f \in \mathbf{C}((z)) \colon [\exists g, h \in \mathbf{C}[z]]\ f = \frac{g}{h} \right\}$$

□ and denoted by $\mathbf{C}(z)$.

Note that writing $\mathbf{C}(z)$ is perfectly consistent with common use in algebra where a pair of brackets stands for ring adjoints and a pair of parentheses for field adjoints. Doubling them refers to the transition from finite to infinite expressions.

Of course, all of these definitions are not restricted to the *ground field* $\mathbf{C}$ of complex numbers. One can proceed from any commutative field F with characteristic 0 and define $F[[z]]$ and the other structures accordingly. We will not need these generalizations, though, and stay with the field $\mathbf{C}$ in what follows.

As for the algebraic domains, this will be enough for our purposes. Next we need to have available some more *operations* on them. Recall our initial idea: any operation on the sequence should somehow be mirrored on the generating function side. This allows us to introduce differentiation as an operation on formal power series. In classical analysis, any analytical function can be differentiated by

$$\left(\sum a_m z^m\right)' = \sum m a_m z^{m-1}$$

within its circle of convergence. Interpreting the power series as the generating function of the series $\langle a_m \rangle$, we see that the operator $\partial/\partial z$ corresponds to the action $a_m \longrightarrow m a_m$ on the sequence. This concept can be generalized to multisequences in the following way.

17 Definition (Differentiation Operator)
Choose some $s \in \mathbf{N}^*$ and $i \in \{1, \dots, s\}$. Then the linear operator D_i on $\mathbf{C}((z))$ is defined by

$$D_i \sum a_{\mathbf{m}} \mathbf{z}^{\mathbf{m}} = \sum m_i a_{\mathbf{m}} \mathbf{z}^{\mathbf{m} - \mathbf{1}_i}$$

for any formal Laurent series. It is called partial *differentiation operator*
□ and also written as $\partial/\partial z_i$.

18 Notation (Derivative Constructs)
The usual multi-index convention is extended to derivatives by regarding $\mathbf{D}$ as the operator-vector $(D_1, \dots, D_s)$ for some $s \in \mathbf{N}^*$. So for any $\mathbf{k} \in \mathbf{N}^s$, *multiple derivatives* are obtained by

$$\mathbf{D}^{\mathbf{k}} = \frac{\partial^{|\mathbf{k}|}}{\partial z_1^{k_1} \cdots \partial z_s^{k_s}},$$

which represents the operator iteration $D_1^{k_1} \dots D_s^{k_s}$. As usual, D^0 means the identity operator. In the univariate case, D is also indicated by the familiar prime-notation.

Analogous to analysis, we define the *gradient* of an s-fold series $f \in \mathbf{C}((\boldsymbol{z}))$ as

$$f'(\boldsymbol{z}) = \frac{df}{d\boldsymbol{z}} = \left(\frac{\partial f}{\partial z_1}, \dots, \frac{\partial f}{\partial z_s}\right).$$

The gradient operator $d/d\boldsymbol{z}$ makes a series-vector out of a scalar one. When applied to a series-vector $\mathbf{f}(\boldsymbol{z}) \in \mathbf{C}_s^r((\boldsymbol{z}))$, the result is a vector of series-vector. This is again a series-vector (consisting of rs components) but is usually regarded as an $r \times s$ matrix

$$\mathbf{f}'(\boldsymbol{z}) = \frac{d\mathbf{f}}{d\boldsymbol{z}} = \begin{pmatrix} \frac{\partial f_1}{\partial z_1} & \cdots & \frac{\partial f_1}{\partial z_s} \\ \vdots & & \vdots \\ \frac{\partial f_r}{\partial z_1} & \cdots & \frac{\partial f_r}{\partial z_s} \end{pmatrix},$$

called the *Jacobian* of $\mathbf{f}(\boldsymbol{z})$. Its determinant, the so-called *Jacobian determinant*, will be denoted by $|\mathbf{f}'(\boldsymbol{z})|$. □

There is an important relation between the derivation operator and the coefficient functional. If $f(\boldsymbol{z}) \in \mathbf{C}_s[[\boldsymbol{z}]]$ is given, it can be developed about the origin by using the Taylor expansion

$$f(\boldsymbol{z}) = \sum [L\mathbf{D}^{\mathbf{m}} f(\boldsymbol{z})] \frac{\boldsymbol{z}^{\mathbf{m}}}{\mathbf{m}!},$$

which means that $[\boldsymbol{z}^{\mathbf{m}}]\, f(\boldsymbol{z}) = (1/\mathbf{m}!)\, L\mathbf{D}^{\mathbf{m}} f(\boldsymbol{z})$ or $\mathbf{m}!\, [\boldsymbol{z}^{\mathbf{m}}] = L\mathbf{D}^{\mathbf{m}}$, if written in operator form. We call this the *Taylor representation*° of the coefficient functional. Let us illustrate its usefulness by developing the powers of the geometric series.

19 **Example (Powers of Geometric Series)**
Choose an arbitrary $c \in \mathbf{N}$. We want to find the expansion of the series $1/(1-z)^c = \sum a_m z^m$. For any $m \in \mathbf{N}$, the coefficient a_m is given by the Taylor representation

$$\begin{aligned} m!\, a_m &= LD^m\,(1-z)^{-c} = LD^{m-1}\,c\,(1-z)^{-c-1} \\ &= LD^{m-2}\,c(c+1)\,(1-z)^{-c-2} = \ldots \\ &= c(c+1)\cdots(c+m-1)\; L(1-z)^{-c-m}. \end{aligned}$$

Here we have used some well-known derivative properties from classical analysis, but we will see in Proposition 26 that they are also valid in the Cauchy algebra of formal power series.

Since $1/(1-z)^{m+c} = (1+z+z^2+\ldots)^{m+c}$, we have $L\,(1-z)^{-m-c} = 1$ and thus $a_m = (c+m-1)^{\underline{m}}/m! = \binom{c+m-1}{m}$ for all $m \in \mathbf{N}$. So the desired Taylor expansion is

$$\frac{1}{(1-z)^c} = \sum \binom{c+m-1}{m} z^m.$$

For some applications, one needs a slight generalization of this identity, having a parameter a instead of the 1 in the denominator. We arrive at

$$\frac{1}{(z+a)^c} = \sum \binom{c+m-1}{m} \frac{(-1)^m}{a^{c+m}}\, z^m$$

□ by substituting $z \to -z/a$ and dividing through a^c.

One of the reasons why the residue is so useful is due to the fact that this coefficient vanishes in every derivative

$$\left(\sum_{m\geq n} a_m z^m\right)' = \sum_{m\geq n} m a_m z^{m-1} = 0 a_0\, z^{-1} + \ldots,$$

which will turn out to be very important later on.

Before we come to Lagrange's Theorem we have to establish one final notion. Recall that Lagrange inversion, in its simplest instance, deals with finding the *inverse* of a given power series. By this term we mean the inverse element with respect to composition whereas the reciprocal is the inverse element for the Cauchy product. But we have not yet defined what we mean by composing two power series. As pointed out, we must not substitute any value for the indeterminates since this would involve topological concepts. When we formally substitute an "interior" power series, $g = \sum b_m z^m$, into an "exterior" one, $f = \sum a_n z^n$, this gives

$$f(g(z)) = \sum_m a_m\, g(z)^m = \sum_m a_m \Big(\sum_n b_n z^n\Big)^m.$$

If this expression is to have any algebraic significance, it must be possible to rearrange the terms in such a way that they make up another power series. Therefore we must stipulate that $b_0 = 0$. Otherwise any term $b_n z^n$ may contribute when we collect the coefficients of some power. This means that we would possibly have infinite coefficient sums (unless f is a polynomial), which are not purely algebraic objects and therefore inadmissible.

For the same reason, we can also substitute a *multivariate series* g into a univariate f as long as $\operatorname{ord} g \geq 1$. On the other hand, if f is multivariate, then $\mathbf{g}$ must be a series-vector for structural reasons (where also $\mathbf{g}$ may be multivariate and $\mathbf{f}$ may be a series-vector) and the above order condition must be satisfied in each component. Borrowing the terminology of analysis, this means

$$\mathbf{C}^r \overset{\mathbf{g}(\boldsymbol{z})}{\longrightarrow} \mathbf{C}^s \overset{\mathbf{f}(\boldsymbol{z})}{\longrightarrow} \mathbf{C}^t$$

for arbitrary $\mathbf{g}(\boldsymbol{z}) \in \mathbf{C}^s_r[[\boldsymbol{z}]]$ and $\mathbf{f}(\boldsymbol{z}) \in \mathbf{C}^t_s[[\boldsymbol{z}]]$.

20 Definition (Composition of Power Series)
First let $f(z) = \sum a_m z^m$, $g(\boldsymbol{z}) = \sum b_{\mathbf{n}} \boldsymbol{z}^{\mathbf{n}}$ be power series. If $\operatorname{ord} g \geq 1$, then g is called *substitutable*. In this case (or with $f \in \mathbf{C}[z]$ alternatively), the *composition* of f with g is defined as the power series

$$\sum_m a_m \Big(\sum_{\mathbf{n}} b_{\mathbf{n}} \boldsymbol{z}^{\mathbf{n}}\Big)^m$$

and written $f \circ g$. For indicating the indeterminate, one can use the alternative form $f(g(\boldsymbol{z}))$.

Next let $\mathbf{g}(\boldsymbol{z}) \in \mathbf{C}^s_r[[\boldsymbol{z}]]$ and $\mathbf{f}(\boldsymbol{z}) \in \mathbf{C}^t_s[[\boldsymbol{z}]]$ be arbitrary. We call $\mathbf{g}$ *substitutable* if all its components $g_1, \ldots, g_s$ are. In this case (or with $\mathbf{f} \in \mathbf{C}^t[z]$ alternatively), the *composition* of $\mathbf{f}$ with $\mathbf{g}$ is defined as the series-vector

$$\Big(\, f_1\big(g_1(\boldsymbol{z}), \ldots, g_s(\boldsymbol{z})\big),\ \ldots,\ f_t\big(g_1(\boldsymbol{z}), \ldots, g_s(\boldsymbol{z})\big) \,\Big)$$

and written $\mathbf{f} \circ \mathbf{g}$. Again, one can use the alternative form $\mathbf{f}(\mathbf{g}(\boldsymbol{z}))$ for indicating the indeterminates. □

By collecting the coefficients of the various powers, we obtain the "explicit" form (restricting ourselves to the univariate and scalar case)

$$\sum_m a_m \Big(\sum_{n\geq 1} b_n z^n\Big)^m = a_0 + \sum_k \left(\sum_{m=1}^{k} a_m \sum_{n_1+\cdots+n_m=k} b_{n_1}\dots b_{n_m}\right) z^k.$$

Observe that all the sums appearing in the coefficient of z^k are finite for each k. For specifying the range of the rightmost sum a little clearer, we introduce a new symbol, which is also common in other areas of combinatorics.

21 Notation (Integer-Composition Symbol)
For any $n \in \mathbf{N}^*$, $(n_1, \dots, n_m)$ is called an *integer composition* of n into m parts iff $n = n_1 + \cdots + n_m$. We shall abbreviate this by the symbol
□ $(n_1, \dots, n_m) \models n$.

Using this notation, we can write the coefficients of the above series composition as

$$[z^0]\ f(g(z)) = a_0,$$

$$[\forall n > 0]\ [z^n]\ f(g(z)) = \sum_{m=1}^{n} \sum_{(n_1,\dots,n_m)\models n} a_m b_{n_1}\dots b_{n_m}.$$

Having thus analyzed the concept of composition, we can now give the formal definition of the inverse.

22 Definition (Inverse of a Power Series)
Let f, g be power series with $\operatorname{ord} f, \operatorname{ord} g \geq 1$. Then f is called the *inverse*
□ of g iff $f(g(z)) = z = g(f(z))$. In that case, f and g are called *invertible*.

23 Notation (Inverse Symbol, Identity Series)
The inverse of a power series f is written as f^*. The symbol id denotes the
□ identity power series $\operatorname{id}(z) = z$.

As we have seen above, substitutable power series must have at least order 1. As the next proposition reveals, those power series which are inverses of others are uniquely characterized by having order *precisely* 1.

24 Proposition (Inverse Condition)
Let $f \in \mathbf{C}[[z]]$ be arbitrary. Then we have the characterization: f has a
□ unique inverse f^* iff $\operatorname{ord} f = 1$.

25 Proof (Inverse Condition)

First assume f is invertible with the inverse $f^* \in \mathbf{C}[[z]]$ so that $f^* \circ f = \mathrm{id} = f \circ f^*$. Defining $r = \operatorname{ord} f$ and $s = \operatorname{ord} g$, we have $\operatorname{ord}(f^* \circ f) = rs = \operatorname{ord}(\mathrm{id}) = 1$, which implies $r = s = 1$.

For the converse, suppose we are given some $f = \sum a_m z^m$ with $\operatorname{ord} f = 1$. We find its inverse $f^* = \sum b_n z^n$ by solving the equation $f^* \circ f = \mathrm{id}$. Since $\left[z^k\right] z = \delta_{k1}$, we know that $b_0 = 0$ and therefore $\operatorname{ord} f^* \geq 1$. The above equation for the k-th coefficient leads us to the recurrence

$$b_1 = \frac{1}{a_1},$$
$$[\forall k > 1] \sum_{n=1}^{k} b_n \sum_{(m_1,\ldots,m_n) \models k} a_{m_1} \ldots a_{m_n} = 0,$$

which has a unique solution $\langle b_n \rangle_{n \geq 1}$ because $a_1 \neq 0$ and the recursive equation involves only finite sums. From the equation $f^* \circ f = \mathrm{id}$ we obtain $f \circ (f^* \circ f) = (f \circ f^*) \circ f = f \circ \mathrm{id} = f$, which implies $f \circ f^* = \mathrm{id}$, so the proof is complete. (We have used the associativity of the $\circ$ operation and the uniqueness of the neutral element id; these two properties are immediately apparent.) □

We have defined *differentiation* somewhat earlier but we have waited with analyzing it until now because the concept of composition has yet been missing. Formal differentiation shares *various useful properties* with its analytical counterpart. It will be sufficient to formulate them for the univariate case since a partial derivative is the same as an ordinary one that fixes all but one of the indeterminates.

26 Proposition (Derivative Properties)

The following properties

1. $[\forall \alpha, \beta \in \mathbf{C}]\ (\alpha f + \beta g)' = \alpha f' + \beta g'$,
2. $(fg)' = f'g + fg'$,
3. $(f^{-1})' = -f'/f^2$,
4. $(f/g)' = (f'g - fg')/g^2$,
5. $[\forall k \in \mathbf{N}^*]\ \ (g^k)' = k\, g^{k-1} g'$,
6. $[\forall \operatorname{ord} g \geq 1]\ \ (f \circ g)' = (f' \circ g)\, g'$,

7. $[\forall \operatorname{ord} f = 1]\ (f^*)' = (f' \circ f^*)^{-1}$

□ are true for all $f, g \in \mathbf{C}((z))$.

27 Proof (Derivative Properties)

1. Linearity is obvious since D is defined as the linear completion of an operator on basis elements.

2. Let

$$f = \sum_{n \geq n_0} a_n z^n, \quad g = \sum_{k \geq k_0} b_k z^k$$

the two Laurent series in question. By the definition of the convolution, we have then

$$(fg)' = D \sum_{m \geq n_0 + k_0} \left(\sum_{n+k=m} a_n b_k \right) z^m$$
$$= \sum_{m \geq n_0 + k_0} \left(\sum_{n+k=m} m\, a_n b_k \right) z^{m-1}.$$

The inner sum can be rearranged as

$$\sum_{n+k=m} m\, a_n b_k = \sum_{n+k=m} (n+k)\, a_n b_k$$
$$= \sum_{n+k=m} (n a_n)\, b_k + \sum_{n+k=m} a_n\, (k b_k),$$

which leads to the claimed result.

3. By the preceding product rule we know that

$$0 = 1' = (ff^{-1})' = f'f^{-1} + f(f^{-1})'$$

and thus $f(f^{-1})' = -f'f^{-1}$, yielding the desired expression for $(f^{-1})'$.

4. This is just rewriting the product rule

$$(f/g)' = (fg^{-1})' = f'g^{-1} + f(-g'/g^2) = (f'g - fg')/g^2$$

and using the foregoing representation of the reciprocal.

5. We apply induction on k. The induction basis is $(g^1)' = 1g^0g'$, which is of course true. The induction step is obtained by the product rule in

$$\begin{aligned}(g^k)' &= (g^{k-1}g)' = (k-1)\,g^{k-2}g'g + g^{k-1}g' \\ &= ((k-1)+1)\,g^{k-1}g' = k\,g^{k-1}g',\end{aligned}$$

where we have used the induction hypothesis for replacing $(g^{k-1})'$.

6. Writing

$$f = \sum_{n \geq n_0} a_n z^n, \quad g = \sum_{k \geq 1} b_k z^k,$$

we shall calculate $[z^m]\ f(g(z))'$ for arbitrary $m \in \mathbf{N}$. Now by definition,

$$f(g(z)) = \sum_{n=n_0}^{\infty} a_n \left(b_1 z + b_2 z^2 + \dots\right)^n,$$

which contains the power z^{m+1} only for $n \leq m+1$. Therefore both $f(g(z))'$ and $f'(g(z))$ contain the power z^m only for these $n \leq m+1$ and

$$[z^m]\ f(g(z))' = [z^m]\ \left(\sum_{n=n_0}^{m+1} a_n g(z)^n\right)'$$

$$= [z^m] \sum_{n=n_0}^{m+1} n a_n\, g(z)^{n-1} g'(z)$$

$$= [z^m] \left(\sum_{n=n_0}^{\infty} n a_n\, g(z)^{n-1}\right) g'(z) = [z^m]\ f'(g(z))\ g'(z).$$

This concludes the proof because power series are equal iff all their coefficients agree.

7. This is obtained immediately from differentiating the defining identity $f \circ f^* = \mathrm{id}$ of the inverse series. Using the chain rule from the previous point, we get $(f' \circ f^*)(f^*)' = 1$, which implies the claimed formula. □

1.4 Combinatorial Concepts

As we have already pointed out after Example 3, the combinatorial approach to proofs consists basically in comparing entities (like partitions, trees, graphs) amongst themselves by some bijective relation. The entity to be counted is thus gradually broken down to smaller ones where the enumeration is known. Our first goal will be to give a precise meaning to such ideas as entities and bijective relations, following closely Labelle's footsteps [12], [13]. The underlying conceptual apparatus is known as *species theory.* It was introduced by André Joyal in 1981 and is presented in full detail in his paper [10].

Before we set out to present its crucial elements, we must fix some notations regarding *sets* because they are the basic building blocks of species theory.

28 Definition (Notation for Sets)
Let A be any finite set.

1. The number of elements in the set A is denoted by $|A|$ and called its *cardinality.* If $|A| = n$, we also say that A is an n-set.

2. If $A_1, A_2, \dots$ are non-empty and mutually disjoint sets whose union is A, we call $\mathbf{A} = \{A_1, A_2, \dots\}$ a *set partition* (represented by boldface uppercase letters) consisting of the blocks $A_1, A_2, \dots,$ and we write $\mathbf{A} \triangleright A$ for short. The number of blocks in $\mathbf{A}$ with cardinality k is denoted by $|\mathbf{A}|_k$. The collection of all such numbers into a vector $\operatorname{typ} \mathbf{A} = (|\mathbf{A}|_1, |\mathbf{A}|_2, \dots)$ is called the *type* of the partition $\mathbf{A}$.

3. As a short notation for a *disjoint union*, we define $A = A_1 \,\dot\cup\, A_2 \,\dot\cup\, \cdots$ iff $A = A_1 \cup A_2 \cup \dots$ and the $A_1, A_2, \dots$ are mutually disjoint or are made so by "coloring" the elements according to which set they belong to.

4. Let $f\colon D \to R$ be any function. If $S \subseteq R$, we write $f^*(S)$ for the (possibly empty) *preimage set* of S. For any $s \in R$, the specific preimage set $f^*(\{s\})$ is abbreviated by $f^*(s)$ and called the *fibre* of s.

In the context of this section, the use of "$\dots$" is somewhat different from the common practice. It signifies some sequence of *unknown but finite cardinality.* For example in point 3 above, $A = A_1 \,\dot\cup\, A_2 \,\dot\cup\, \cdots$ denotes the disjoint union of the sets $A_1, A_2, \dots$ whose number we do not know. □

An "entity" such as a partition or a tree is called a *structure*, and the collection of all such structures a *species* $\mathfrak{f}$. (We will use a Gothic font to denote species, reserving lowercase letters for variables and uppercase letters for constants.) What is the essential ingredient of any such structure?

First, we should be able to construct them on any finite set $\mathfrak{U}$. For example, on $\mathfrak{U} = \{1,2,3\}$ we can build the partitions $\mathfrak{f}(\mathfrak{U}) = \{\{1,2,3\}, \{1|2,3\}, \{2|1,3\}, \{3|1,2\}, \{1|2|3\}\}$. But this assignment $\mathfrak{U} \mapsto \mathfrak{f}(\mathfrak{U})$ must not be arbitrary, it has to respect the structural information of the species. If we decide to rename the elements of $\mathfrak{U}$ in our partition example into $\mathfrak{U}' = \{a,b,c\}$ by $1 \mapsto a, 2 \mapsto b, 3 \mapsto c$, we should expect to get $\mathfrak{f}(\mathfrak{U}') = \{\{a,b,c\}, \{a|b,c\}, \{b|a,c\}, \{c|a,b\}, \{a|b|c\}\}$. The structural information of the species is respected if the composition of such relabelings commutes with the assignment $\mathfrak{U} \mapsto \mathfrak{f}(\mathfrak{U})$.

29 Definition (Species)
A combinatorial *species* $\mathfrak{f}$ is a rule that associates to any finite set $\mathfrak{U}$ a finite set $\mathfrak{f}(\mathfrak{U})$ and to any bijection $\alpha\colon \mathfrak{U} \to V$ between finite sets $\mathfrak{U}$ and V a corresponding bijection $\mathfrak{f}(\alpha)\colon \mathfrak{f}(\mathfrak{U}) \to \mathfrak{f}(V)$. All such bijections α, β must satisfy $\mathfrak{f}(\beta \circ \alpha) = \mathfrak{f}(\beta) \circ \mathfrak{f}(\alpha)$ whenever $\beta \circ \alpha$ is defined. Furthermore, all finite sets $\mathfrak{U}$ must satisfy $\mathfrak{f}(1_{\mathfrak{U}}) = 1_{\mathfrak{f}(\mathfrak{U})}$ where 1_A denotes the identity function on A.

The sets $\mathfrak{U}$ and $\mathfrak{f}(\mathfrak{U})$ are called label set° and structure set°, respectively; the elements of the latter are the structures (more precisely: $\mathfrak{f}$-structures). The bijections α are known as relabelings°, their images $\mathfrak{f}(\alpha)$ as transportations (more precisely: transportations of $\mathfrak{f}$-structures along α).

Technically speaking, $\mathfrak{f}$ is thus an endofunctor on the category of finite sets: its objects are the finite sets (label sets and structure sets) and its arrows are the bijections (relabelings and transportations). □

Having available now the precise notion of a species, let us list a few elementary examples. Some of these will also be taken up later in the text, which will also reveal the motivation for some of the names.

30 Example (Some Species)
The most trivial example is the *empty species* $\mathfrak{o}$, defined by $\mathfrak{o}(\mathfrak{U}) = \emptyset$ for all finite sets $\mathfrak{U}$: no structures can be constructed on any label set.

Only slightly more complex is the *empty-set species* $\mathfrak{1}$, defined by $\mathfrak{1}(\emptyset) = \{\emptyset\}$ and $\mathfrak{1}(\mathfrak{U}) = \emptyset$ for all other finite sets $\mathfrak{U}$. In this case, a "structure" can only be built on the empty label set.

The *singleton species* $\mathfrak{Z}$ is introduced as $\mathfrak{Z}(\mathfrak{U}) = \{\mathfrak{U}\}$ if $|\mathfrak{U}| = 1$ and $\mathfrak{Z}(\mathfrak{U}) = \emptyset$ else. This species allows to build structures only on singleton label sets, the structures being just their label (wrapped in a set).

Next the *set species* $\mathfrak{E}$ is defined by $\mathfrak{E}(\mathfrak{U}) = \{\mathfrak{U}\}$ for all finite sets $\mathfrak{U}$. Thus for any label set, the structures are merely their copies (wrapped in another set). The set species provides a very neutral structure on the labels (actually none at all, in the common sense of the word)—they do no more than preserve them. For this reason, $\mathfrak{E}$ is also called the *uniform species*.

As a less trivial example, we introduce the species $\mathfrak{S}$ of all *permutations* (forming the symmetric group), where $\mathfrak{S}(\mathfrak{U})$ consists of all bijections on $\mathfrak{U}$. A typical element can be written as

$$\mathfrak{s} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix},$$

in this case $\mathfrak{U}$ being the set $\{1, 2, 3, 4\}$. For $\mathfrak{U} = \emptyset$, there is precisely one permutation, called the empty permutation.

Finally, we introduce the species $\mathfrak{L}$ of *linear orderings* with $\mathfrak{L}(\mathfrak{U})$ containing all arrangements $(\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \dots)$ of the labels $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \dots \in \mathfrak{U}$. On $\mathfrak{U} = \emptyset$,
□ there is again precisely one empty ordering.

Sometimes it is convenient to exclude empty cases occurring in such structures as $\mathfrak{S}$ or $\mathfrak{L}$. This process is known as *regularization* and denoted as follows.

31 Definition (Regularization)
If $\mathfrak{f}$ is any species, we define its *regularization* $\mathfrak{f}^*$ by setting $\mathfrak{f}^*(\mathfrak{U}) = \mathfrak{f}(\mathfrak{U})$
□ for all $\mathfrak{U} \neq \emptyset$ and $\mathfrak{f}^*(\emptyset) = \emptyset$.

The next concept to be made precise is that of a "bijection" between two species. In fact, there are two concepts reflecting that idea. If the two species in question can be identified with one another in a consistent way (regarding the transportations), they contain the same structural information. In that case, they are said to be *isomorphic*. But for the purpose of counting, it will usually suffice that the corresponding structure sets have the same cardinality; they are called *equipotent* then.

32 Definition (Isomorphism, Equipotence)
The two species $\mathfrak{f}, \mathfrak{g}$ are said to be *equipotent* if there exists a bijection (called an equipotence) $\sigma_{\mathfrak{U}}\colon \mathfrak{f}(\mathfrak{U}) \to \mathfrak{g}(\mathfrak{U})$ for all finite sets $\mathfrak{U}$.

They are *isomorphic* if the equipotence (or rather isomorphism) commutes with all transportations. This means that the diagram on the right is commutative for all finite sets U, V and all relabelings $\alpha : U \to V$. Technically speaking, σ is a natural transformation between the categories $\mathfrak{f}$ and $\mathfrak{g}$.

$$\begin{array}{ccc} \mathfrak{f}(U) & \xrightarrow{\mathfrak{f}(\alpha)} & \mathfrak{f}(V) \\ {\scriptstyle\sigma_U}\downarrow & \sharp & \downarrow{\scriptstyle\sigma_V} \\ \mathfrak{g}(U) & \xrightarrow[\mathfrak{g}(\alpha)]{} & \mathfrak{g}(V) \end{array}$$

For denoting the equipotence or isomorphism of the species $\mathfrak{f}$ and $\mathfrak{g}$, we write $\mathfrak{f} \sim \mathfrak{g}$ or $\mathfrak{f} \simeq \mathfrak{g}$, respectively. □

Two species that are isomorphic actually cannot be distinguished; it is somehow the same thing described in two different languages. Therefore, the notion of isomorphism can be seen to be the "right" concept of equality between species. Some authors also write $\mathfrak{f} = \mathfrak{g}$ instead of $\mathfrak{f} \simeq \mathfrak{g}$ for that reason. At any rate, such a relation is usually called a *combinatorial equation.*

Obviously, isomorphism is *stronger* than equipotence. As a simple example, consider the species $\mathfrak{S}$ of permutations and the species $\mathfrak{L}$ of linear orderings. They are obviously equipotent since any ordering of n points on a line can be seen as a permutation and vice versa. For seeing that they are *not* isomorphic, it will suffice to give one concrete counter-example. Take as label sets $U = V = \{1, 2, 3\}$ and as a relabeling $\alpha \colon U \to U$, defined by $1 \mapsto 1, 2 \mapsto 3, 3 \mapsto 2$. The only natural candidate for an isomorphism between $\mathfrak{S}$ and $\mathfrak{L}$ is given by

$$\sigma_U \colon \mathfrak{S}(U) \to \mathfrak{L}(U), \quad [\forall a, b, c \in U] \ \sigma \begin{pmatrix} 1 & 2 & 3 \\ a & b & c \end{pmatrix} = (a, b, c).$$

For $\mathfrak{S} \simeq \mathfrak{L}$, the diagram in Definition 32 would have to commute also for our choice of U, V, and α, which means

$$[\forall s \in \mathfrak{S}(U)] \ \mathfrak{L}(\alpha)\, \sigma_U\, s = \sigma_U\, \mathfrak{S}(\alpha)\, s.$$

But choosing

$$s = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

gives $\sigma_U\, s = (1, 3, 2)$ with $\mathfrak{L}(\alpha)\, \sigma_U\, s = (1, 2, 3)$ on the one hand and

$$\mathfrak{S}(\alpha)\, s = \begin{pmatrix} 1 & 3 & 2 \\ 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

with $\sigma_U\, \mathfrak{S}(\alpha)\, s = (1,3,2)$ on the other hand. Therefore the species $\mathfrak{S}$ and $\mathfrak{L}$ cannot be isomorphic.

In Section 1.1, we have pointed out the great innovation brought forth by the idea of *generating functions*. They are indeed a powerful instrument for solving various combinatorial problems. Species theory now allows to clarify this idea even more, because generating functions can be associated canonically with species. In fact, they provide all the information about counting the number $\mathfrak{f}_m$ of structures of a species $\mathfrak{f}$, namely $|\mathfrak{f}(U)|$ with U running through all finite sets of cardinality m. Since the specific nature of the labels is of no importance for enumeration, we may have U run through $\emptyset$ and the standard label sets° $\{1,\ldots,m\}$ with $m \in \mathbf{N}$. Now it might seem obvious to define the generating function of the species to be just that sequence $\langle \mathfrak{f}_m \rangle$, as made precise in Definition 5—this would be its *ordinary* generating function. But for various reasons, it turns out to be more useful to choose the *exponential* generating function, which is the sequence $\langle |\mathfrak{f}_m|/m! \rangle$. Thus we are led to the following definition.

33 Definition (Generating Function of a Species)
Let $\mathfrak{f}$ be any species. Then its *generating functions* is defined as

$$\mathfrak{f}(z) = \sum \mathfrak{f}_m \frac{z^m}{m!} = \sum |\mathfrak{f}(U_m)| \frac{z^m}{m!},$$

□ where $U_m = \{1,\ldots,m\}$ for all $m \in \mathbf{N}^*$ and $U_0 = \emptyset$.

By definition, the generating functions of equipotent species are equal. In order to provide some *examples*, we list the generating functions of the species in Example 30; their derivation is immediately obvious.

34 Example (Some Generating Functions)
The power series

$$\mathfrak{o}(z) = 0,$$
$$\mathfrak{1}(z) = 1,$$
$$\mathfrak{Z}(z) = z,$$
$$\mathfrak{E}(z) = e^z,$$
$$\mathfrak{S}(z) = \mathfrak{L}(z) = (1-z)^{-1}$$

□ are the generating functions of the species $\mathfrak{o}, \mathfrak{1}, \mathfrak{Z}, \mathfrak{E}, \mathfrak{S}, \mathfrak{L}$, respectively.

Of course, e^z must be understood here as a purely *symbolic* notation for the formal power series $\sum z^m/m!$, so named because it is the Taylor expansion of the analytical function e^z. We will take this liberty throughout the thesis since power series are always assumed to be formal (unless explicitly stated otherwise).

As we have mentioned in Section 1.3, it is also highly desirable to have methods for building complicated structures from simpler ones. At the same time, this allows to break down complex structures into small and transparent parts. As we will shortly see, all of these *operations* are mirrored in the generating functions. This gives us a very powerful tool for enumerating the structures.

35 Definition (Operations on Species)
Let $\mathfrak{f}, \mathfrak{g}$ be two arbitrary species. Then we can perform the following *operations* on $\mathfrak{f}, \mathfrak{g}$ for obtaining new species.

1. The *sum* of $\mathfrak{f}$ and $\mathfrak{g}$ is defined as
$$(\mathfrak{f}+\mathfrak{g})\,(U) = \mathfrak{f}(U)\,\dot{\cup}\,\mathfrak{g}(U).$$
2. The *product* of $\mathfrak{f}$ and $\mathfrak{g}$ is defined as
$$(\mathfrak{f}\,\mathfrak{g})\,(U) = \{(f,g) \mid f \in \mathfrak{f}(U_1) \wedge g \in \mathfrak{g}(U_2) \wedge U_1 \,\dot{\cup}\, U_2 = U\}.$$
3. If $\mathfrak{g}(\emptyset) = \emptyset$, the *composition* of $\mathfrak{f}$ and $\mathfrak{g}$ is defined as
$$\begin{aligned}(\mathfrak{f}\circ\mathfrak{g})\,(U) = \{\,(f;\{g_1,g_2,\ldots\}) \mid f \in \mathfrak{f}(\{U_1,U_2,\ldots\}) \wedge \\ g_1 \in \mathfrak{g}(U_1) \wedge g_2 \in \mathfrak{g}(U_2) \wedge \ldots \wedge \\ U_1 \,\dot{\cup}\, U_2 \,\dot{\cup}\, \cdots = U\,\}.\end{aligned}$$
4. The *derivative* of $\mathfrak{f}$ is defined as
$$\mathfrak{f}'(U) = \mathfrak{f}(U \cup \{U\}).$$
5. The *pointing* of $\mathfrak{f}$ is defined as
$$\mathfrak{f}^{\bullet}(U) = \{(f,u) \mid f \in \mathfrak{f}(U) \wedge u \in U\}.$$

As expected, the processes for obtaining these structures are named adding, multiplying, substituting, differentiating, and pointing, respectively. □

Roughly speaking, the sum amounts to merging structures, the product to gluing them, and the composition to nesting them. Likewise, a pointing results in marking ("pointing at") an element, a derivative in adding one. We should also note the following details.

- The $\dot\cup$-operator in the definition of the sum prohibits the fusing of like structures. This means that the elements of $\mathfrak{f}(U)$ and $\mathfrak{g}(U)$ are regarded to have different "colors". As an example, let $M = \{a, b, c\}$. Then we have $M \cup M = \{a, b, c\}$, whereas the disjoint union gives something like $M \dot\cup M = \{a_1, b_1, c_1, a_2, b_2, c_2\}$.

- The condition $\mathfrak{g}(\emptyset) = \emptyset$, imposed on the composition of species, ensures a finite number of $(\mathfrak{f} \circ \mathfrak{g})$-structures. If this condition were not demanded, we could have infinitely many decompositions of the form

$$\begin{aligned} U &= U \dot\cup \emptyset, \\ U &= U \dot\cup \emptyset \dot\cup \emptyset, \\ U &= U \dot\cup \emptyset \dot\cup \emptyset \dot\cup \emptyset, \\ &\dots \end{aligned}$$

  Each of them would contribute at least one element to $(\mathfrak{f} \circ \mathfrak{g})(U)$, making it thus an infinite set.

- The set $U \cup \{U\}$, occurring in the definition of the derivative, means that U is extended by a new element (since $\{U\}$ certainly cannot be an element of U itself). Sometimes this set is also written as something like $U \cup \{*\}$ with $*$ as a symbol for the new element.

As we have announced before, these operations turn out to be compatible with the corresponding operations on the generating functions.

36 Proposition (Operations on Generating Functions)
Let $\mathfrak{f}$ and $\mathfrak{g}$ be two arbitrary species. Then we have the compatibility properties

1. $(\mathfrak{f} + \mathfrak{g})(z) = \mathfrak{f}(z) + \mathfrak{g}(z)$,
2. $(\mathfrak{f}\,\mathfrak{g})(z) = \mathfrak{f}(z)\,\mathfrak{g}(z)$,
3. $(\mathfrak{f} \circ \mathfrak{g})(z) = \mathfrak{f}(\mathfrak{g}(z))$,
4. $\mathfrak{f}'(z) = D\,\mathfrak{f}(z)$,

5. $\mathfrak{f}^\bullet(z) = z\, D\, \mathfrak{f}(z)$

□ for the species operations introduced in Definition 35.

37 Proof (Operations on Generating Functions)
Let us fix some arbitary label set U of any cardinality $n \in \mathbf{N}$ and look at the corresponding coefficients of the generating functions.

1. Since we have a disjoint union of $\mathfrak{f}(U)$ and $\mathfrak{g}(U)$, the number of structures in $(\mathfrak{f}+\mathfrak{g})\,(U)$ is indeed just $\mathfrak{f}_n + \mathfrak{g}_n$.

2. A structure of $(\mathfrak{f}\,\mathfrak{g})\,(U)$ consists of an $\mathfrak{f}$-structure on U_1 and a $\mathfrak{g}$-structure on U_2. Here U_1 may be an arbitary subset of U so that $U_2 = U \setminus U_1$. Since there are $\binom{n}{k}$ subsets U_1 of cardinality k, with k varying between 0 and n, we have

$$(\mathfrak{f}\,\mathfrak{g})_n = \sum_{k=0}^{n} \binom{n}{k} \mathfrak{f}_k\, \mathfrak{g}_{n-k},$$

which surely is the coefficient of $z^n/n!$ in

$$\begin{aligned}\mathfrak{f}(z)\,\mathfrak{g}(z) &= \left(\sum \mathfrak{f}_n\, z^n/n!\right)\left(\sum \mathfrak{g}_n\, z^n/n!\right)\\ &= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\, \mathfrak{f}_k\, \mathfrak{g}_{n-k}\right) z^n/n!\,.\end{aligned}$$

3. First we observe that $\mathfrak{f}(\mathfrak{g}(z))$ is defined because of the condition $\mathfrak{g}(\emptyset) = \emptyset$, which is equivalent to $\mathfrak{g}_0 = 0$ or $\operatorname{ord} \mathfrak{g}(z) \geq 1$. The latter means that $\mathfrak{g}(z)$ is substitutable according to Definition 20.

   The number of structures in $(\mathfrak{f} \circ \mathfrak{g})\,(U)$ is

$$(\mathfrak{f} \circ \mathfrak{g})_n = \sum_{\mathbf{U} \rhd U} \mathfrak{f}_{|\mathbf{U}|}\, \mathfrak{g}_1^{|\mathbf{U}|_1} \cdots \mathfrak{g}_n^{|\mathbf{U}|_n}.$$

   How many partitions $\mathbf{U}$ of type $(k_1, \ldots, k_n)$ are there? Any of the $n!$ permutations of U can be viewed as a partition. Going from left to right, we split it into: one segment containing k_1 blocks of 1 element, one segment containing k_2 blocks of 2 elements, and so on. Of course, many partitions will be generated several times. Firstly, the order of the elements within the blocks is irrelevant—there are $1!^{k_1} \cdots n!^{k_n}$ such orders. Secondly, the order of the blocks within the segments is also irrelevant—there are $k_1! \cdots k_n!$ such orders. Hence the number of partitions $\mathbf{U}$ of type $(k_1, \ldots, k_n)$ is

$$\frac{n!}{1!^{k_1} k_1! \cdots n!^{k_n} k_n!}.$$

Now the total number of blocks $m = |\mathbf{U}|$ may vary between 1 and n. Then the partition types must satisfy the simultaneous conditions $(k_1, \ldots, k_n) \in \mathbf{N}^n$ and $k_1 + \cdots + k_n = m$ and $1k_1 + \cdots + nk_n = n$. Abbreviating the set of all such $(k_1, \ldots, k_n)$ by K_n, the above sum can be rewritten into

$$\begin{aligned}(f \circ g)_n &= \sum_{m=1}^{n} \sum_{(k_1,\ldots,k_n)\in K_n} \frac{n!}{1!^{k_1} k_1! \cdots n!^{k_n} k_n!} f_m \, g_1^{k_1} \cdots g_n^{k_n} \\ &= n! \sum_{m=1}^{n} \sum_{(k_1,\ldots,k_n)\in K_n} \frac{m!}{k_1! \cdots k_n!} \frac{f_m}{m!} \left(\frac{g_1}{1!}\right)^{k_1} \cdots \left(\frac{g_n}{n!}\right)^{k_n}.\end{aligned}$$

In order to recognize the familiar structure of series composition in this sum, we have to realize that every type $(k_1, \ldots, k_n) \in K_n$ can be interpreted as an integer partition $(n_1, \ldots, n_m)$ of n into m parts: the numbers $k_1, \ldots, k_n$ just count the occurrences of the parts $1, \ldots, n$. In that representation, the powers $(g_1/1!)^{k_1} \cdots (g_n/n!)^{k_n}$ are spread out into $(g_{n_1}/n_1!) \cdots (g_{n_m}/n_m!)$. But since this multiplication is commutative, we can as well sum over integer compositions instead of integer partitions. There are $m!/(k_1! \cdots k_n!)$ integer compositions of n into m parts, corresponding to one integer partition of n into m parts with the part 1 occurring k_1 times, the part 2 occurring k_2 times, and so on. (Every partition $n_1 + \cdots + n_m$ can be permuted in $m!$ ways. But without changing it as an integer composition, the k_1 occurrences of the part 1 may be arranged in $k_1!$ ways, the k_2 occurrences of the part 2 in $k_2!$ ways, and so on.) So passing from integer partitions to compositions, the sum takes on its final form

$$\frac{(f \circ g)_n}{n!} = \sum_{m=1}^{n} \sum_{(n_1,\ldots,n_m)\models n} \frac{f_m}{m!} \frac{g_{n_1}}{n_1!} \cdots \frac{g_{n_m}}{n_m!}.$$

Comparing this with the coefficient representation of composed series as given after Notation 21, we realize that $(f \circ g)_n/n!$ is indeed just $[z^n]\, f(g(z))$.

4. Since any $\mathfrak{f}'$-structure on the n-set U is actually an $\mathfrak{f}$-structure on the $(n+1)$-set $U \cup \{U\}$, we have the simple relation $\mathfrak{f}'_n = \mathfrak{f}_{n+1}$. This fits perfectly with the corresponding generating functions

$$\begin{aligned} D\,\mathfrak{f}(z) &= D \sum_{m \geq 0} \mathfrak{f}_m \frac{z^m}{m!} = \sum_{m \geq 1} \mathfrak{f}_m \frac{z^{m-1}}{(m-1)!} = \sum_{m \geq 0} \mathfrak{f}_{m+1} \frac{z^m}{m!} \\ &= \sum_{m \geq 0} \mathfrak{f}'_m \frac{z^m}{m!} = \mathfrak{f}'(z). \end{aligned}$$

5. For pointing the $\mathfrak{f}_n$ structures of $\mathfrak{f}(U)$, we have to mark any of its n elements. Hence we obtain $\mathfrak{f}^\bullet_n = n\,\mathfrak{f}_n$ structures of $\mathfrak{f}^\bullet(U)$ altogether. Looking at the generating functions, this means

$$z\,D\,\mathfrak{f}(z) = z\,D \sum \mathfrak{f}_m \frac{z^m}{m!} = \sum m\,\mathfrak{f}_m \frac{z^m}{m!} = \sum \mathfrak{f}^\bullet_m \frac{z^m}{m!} = \mathfrak{f}^\bullet(z).$$

□ Thus we are finished with the proof.

This proposition yields an extremely important insight as a by-product: at least any finite segment of a formal power series (in other words, any polynomial) can be considered as the generating function of some species. This principle of *combinatorial back-interpretation*° holds because the trivial power series z is modelled in the singleton species $\mathfrak{Z}$, the powers z^n are obtained by species multiplication, and all other polynomials are just linear combinations of them. (The multiplication of a species $\mathfrak{f}$ with a "scalar" like 17 may be viewed as a species multiplication $\mathfrak{17}\,\mathfrak{f}$, where the species $\mathfrak{17}$ is defined as the species sum $\mathfrak{1} + \cdots + \mathfrak{1}$ containing 17 copies of the empty-set species $\mathfrak{1}$.)

Now it is time to give a little example of how the mighty apparatus of species theory can be put into action. We will look at the famous *problème des rencontres* (mentioned in [15]), which P. R. Montmort formulated in 1708. He considered an urn containing n balls numbered 1 through n. By drawing the balls from the urn without replacement, we obtain a permutation of the numbers. A coincidence occurs if for some $k \in \{1, \ldots, n\}$, the k-th ball drawn is labeled k. Montmort computed the probability of drawing a *derangement*: a permutation without such a coincidence. If the number of derangements on $\{1, \ldots, n\}$ is denoted by D_n, the desired probability is $D_n/n!$. But what is D_n? By conventional means, the solution requires either some ingenious idea or a tedious calculation. Using species theory (confer p. 10 in Joyal's paper [10]), we obtain it almost instantly and in a rather elegant way.

38 Example (Derangements)
Since derangements are permutations with 0 fixed points, let us call this species $\mathfrak{S}_0(z)$. Any ordinary permutation can be decomposed into a set of fixed points and a derangement (either of them possibly empty). Thus we obtain the combinatorial equation $\mathfrak{S} \simeq \mathfrak{S}_0\,\mathfrak{E}$, which in turn gives rise to the generating-function identity

$$\frac{1}{1-z} = \mathfrak{S}_0(z)\,e^z$$

or $\mathfrak{S}_0(z) = e^{-z}/(1-z)$. Therefore we have

$$\frac{D_n}{n!} = [z^n]\,\mathfrak{S}_0(z) = [z^n]\,\frac{e^{-z}}{1-z} = [z^n]\left(\sum \frac{(-1)^n}{n!}z^n\right)\left(\sum z^n\right)$$
$$= [z^n]\sum_n \left(\sum_{k=0}^{n} \frac{(-1)^k}{k!}\right) z^n = \sum_{k=0}^{n} \frac{(-1)^k}{k!},$$

which converges to e^{-1} when n approaches infinity. So when the dancing pairs of a big wedding party are arranged at random, the odds are $1 : e$ or approximately $10 : 27$ that nobody will dance with the usual partner. □

There is another classical combinatorial concept that can be integrated into the framework of species theory in an elegant way. It is the concept of a *coloring*, together with a very natural generalization. Since we shall need it for the combinatorial proof of Lagrange's Theorem, we introduce it here.

We have pointed out above that multiplying species means to "glue" structures. If the factors are all equal (of one primary species), then every product structure is represented by a partition of the label set (possibly supplemented by empty blocks), with one primary structure imposed on each block. The partition induces a coloring on the label set—the block numbers being the "colors". In addition to that, each color block bears a structure. The following definition introduces this notion of an *enriched coloring*.

39 Definition (Enriched Coloring)
Let U be any finite set and $\lambda \in \mathbf{N}^*$. Then $(u; r_1, \ldots, r_\lambda)$ is a *coloring* of U in λ colors, *enriched* by a species $\mathfrak{r}$ (briefly, an $\mathfrak{r}$-enriched λ-coloring) iff u is a function (called the coloring) from U into $\{1, \ldots, \lambda\}$ and $r_k \in \mathfrak{r}(u^*(k))$ for every $k \in \{1, \ldots, \lambda\}$.

For $\lambda = 0$, there is no enriched coloring unless $U = \emptyset$ (empty coloring). □

Note that some colors may be unused because the coloring function is not required to be surjective—in other words, some color blocks can be empty. By choosing the uniform species $\mathfrak{E}$ for $\mathfrak{r}$, we regain a conventional (non-enriched) coloring. But of course our focus will be on enriched colorings. As mentioned above, they can be regarded as structures from a *power species*.

40 Proposition (Power Species)

For any $\lambda \in \mathbf{N}$ and any species $\mathfrak{r}$, the set of $\mathfrak{r}$-enriched λ-colorings can be identified with the *power species* $\mathfrak{r}^\lambda$. In particular, the number of $\mathfrak{r}$-enriched λ-colorings on an n-set is $(\mathfrak{r}^\lambda)_n$. □

41 Proof (Power Species)

Let U be an arbitrary label set of cardinality $n \in \mathbf{N}$. By Definition 35, Point 2, a structure of $\mathfrak{r}^2(U)$ is of the form (r_1, r_2) with r_1 and r_2 being $\mathfrak{r}$-structures on U_1 and U_2, respectively, fulfilling $U_1 \dot\cup U_2 = U$. By induction, we infer that for every $\lambda \in \mathbf{N}^*$, the species $\mathfrak{r}^\lambda(U)$ contains structures of the form $(r_1, \dots, r_\lambda)$. Every r_i is an $\mathfrak{r}$-structure on some label set U_i for all $i \in \{1, \dots, \lambda\}$, and the label sets fulfil $U_1 \dot\cup \cdots \dot\cup U_\lambda = U$.

Now we fix one particular element $(r_1, \dots, r_\lambda)$ of $\mathfrak{r}(U)$, and we define a coloring function $u \colon U \to \{1, \dots, \lambda\}$ by setting

$$[\forall l \in U] \;\; u(l) = i \Leftrightarrow l \in U_i.$$

By this method, we can canonically assign a coloring $(u; r_1, \dots, r_\lambda)$ to every element $(r_1, \dots, r_\lambda)$ of $\mathfrak{r}^\lambda(U)$. Since this assignment is bijective, we can identify the $\mathfrak{r}$-enriched λ-colorings with the $\mathfrak{r}^\lambda$-structures. In particular, the bijectivity implies that the total number of $\mathfrak{r}$-enriched λ-colorings on the n-set U is indeed $(\mathfrak{r}^\lambda)_n$.

In the special case $\lambda = 0$, there are no colors to assign to the labels of U. So if U is not empty, there can be no coloring. For $U = \emptyset$, however, we have the empty coloring (see end of Definition 39). Therefore, the 0-colorings can be identified with the empty-set species $\imath = \mathfrak{r}^0$. □

In this section, we have set up a powerful toolbox of combinatorial concepts, adding to the elementary concepts developed in the previous section. With these sharpened instruments at hand, we are now in the position to state, prove, and apply Lagrange's Theorem.

Through wisdom is an house builded; and by understanding it is established: and by knowledge shall the chambers be filled with all precious and pleasant riches.
Proverbs 24:3–4

Chapter 2

Lagrange's Theorem

The stately edifice of Lagrange's theorem will now be erected on the foundation laid in the previous chapter, using knowledge from various branches of mathematics. In Section 2.1, we shall *state* Lagrange's Theorem together with some variants, while the remaining sections will present a variety of *different proofs*. The proof in Section 2.2 uses only elementary concepts (as introduced in Section 1.3). In Section 2.3, we apply results from classical analysis, and in Section 2.4 from linear algebra. Combinatorial concepts (as introduced in Section 1.4) are employed for the proof in Section 2.5. Finally in Section 2.6, we prove the multivariate generalization of Lagrange's theorem, again using only elementary concepts.

2.1 The Theorem

The *object* of Lagrange's Theorem is to expand a given power series f in terms of another given power series g. Thus we want to find that power series h which gives $h \circ g = f$. Note that we can solve the inversion problem in the literal sense by the specific choice $f = \mathrm{id}$. In the univariate case, the coefficients of h can be written in two different forms. The next theorem is actually the core piece of this diploma thesis and is proved in the following sections.

42 **Theorem (Lagrange Inversion)**
Let $f, g \in \mathbf{C}[[z]]$ with $\operatorname{ord} g = 1$. Then there exists an $h \in \mathbf{C}[[z]]$ such that

$h \circ g = f$, and its coefficients c_m can be found as

$$c_0 = \mathrm{L}\, f(z),$$
$$[\forall m > 0]\ \ c_m = \frac{1}{m}\,\mathrm{M}\, f'(z)\, g(z)^{-m} = \mathrm{M}\, f(z)\, g'(z)\, g(z)^{-m-1}.$$

□ These equations are known as *Lagrange inversion formulas.*

For some applications, it is more convenient to write the series $g(z)$ as a *fraction* $z/e(z)$ with an appropriate $e(z)$. For that, we can apply the following equivalent variant of Lagrange's Theorem.

43 Corollary (Lagrange Inversion, Fraction Variant)
Let $f, e \in \mathbf{C}[[z]]$ with $\operatorname{ord} e = 0$ and set $g(z) = z/e(z)$. Then there exists an $h \in \mathbf{C}[[z]]$ such that $h(g(z)) = f(z)$, and its coefficients c_m can be found as

$$c_0 = \left[z^0\right]\ f(z),$$
$$[\forall m > 0]\ \ c_m = \frac{1}{m}\left[z^{m-1}\right]\ f'(z)\, e(z)^m = [z^m]\ f(z)\, g'(z)\, e(z)^{m+1}.$$

These equations are equivalent to the *Lagrange inversion formulas* (and
□ also called by this name).

44 Proof (Lagrange Inversion, Fraction Variant)
The equations follow immediately from their counterparts in Theorem 42 when we replace $g(z)$ by $z/e(z)$ and use the absorption property of M.

On the other hand, we can write any order-1-series $g(z)$ as $z/e(z)$. We can do this by pulling out the indeterminate, which leaves a power series of order 0. This series can therefore be written as the reciprocal of some other power series $e(z)$ of order 0. With this substitution, the formulas of
□ Theorem 42 again follow by the absorption property as before.

Since Lagrange inversion is often needed for *solving functional equations*, we shall present another equivalent variant of Lagrange's Theorem, appropriate for such a situation. We shall state the inversion formulas only in their first form, though, because the other one seems uncommon in literature.

45 Corollary (Lagrange Inversion, Implicit Form)
Let $G \in \mathbf{C}[[z]]$ be the solution of the functional equation $G(z) = z\, e(G(z))$, where e is a given power series of order 0. Then we can find the expansion

coefficients $c_m = [z^m]\ f(G(z))$ by

$$c_0 = \left[z^0\right]\ f,$$
$$[\forall m > 0]\ c_m = \frac{1}{m}\ \left[z^{m-1}\right]\ f'(z)\, e(z)^m$$

□ for every $f \in \mathbf{C}[[z]]$.

46 Proof (Lagrange Inversion, Implicit Form)

First let us prove the implicit form by using Theorem 42. By defining $g(z) = z/e(z)$, we can rearrange the given functional equation as $g(G(z)) = z$ or $g \circ G = \mathrm{id}$, so g is actually the left-inverse of G. Since any left-inverse is also a right-inverse in $\mathbf{C}[[z]]$, we also have the identity $G \circ g = \mathrm{id}$. Applying the power series f on both sides, this yields $(f \circ G) \circ g = f$ or $h(z/e(z)) = f(z)$ with $h := f \circ G$. Now we can use Corollary 43 (which is equivalent to Theorem 42) to obtain the desired coefficients c_m.

Now let us deduce Corollary 43 (equivalent to Theorem 42) from the implicit form. We are given $f, e \in \mathbf{C}[[z]]$ with $\operatorname{ord} e = 0$ and $g(z) = z/e(z)$. Substituting these assignments into the functional equation $G(z) = z\, e(G(z))$, again gives us the identity $g(G(z)) = z$ and hence $G = g^*$. Applying the implicit form of Lagrange inversion yields the coefficients c_m in the development $f(G(z)) = \sum c_m z^m$, which is equivalent to $f(z) = \sum c_m\, g(z)^m$ because of $G = g^*$. But the latter equation means that the c_m are at the same time the coefficients of a power series h defined by $h(g(z)) = f(z)$,
□ just as required in Corollary 43.

It turns out that the inversion formulas in their second form can be generalized to the *multivariate case* in a very natural way. (For the first form of the inversion formulas, the multivariate generalizations are very complicated.) We are given some $f(\boldsymbol{z}) \in \mathbf{C}_s[[\boldsymbol{z}]]$, which is to be developed in terms of a given *series-vector* $\mathbf{g}(\boldsymbol{z}) \in \mathbf{C}_s^s[[\boldsymbol{z}]]$ such that

$$f(\boldsymbol{z}) = \sum c_{\mathbf{m}}\, \mathbf{g}(\boldsymbol{z})^{\mathbf{m}} = \sum_{m_1,\dots m_s \geq 0} c_{m_1,\dots,m_s}\, g_1(\boldsymbol{z})^{m_1} \dots g_s(\boldsymbol{z})^{m_s}.$$

We want to know the multisequence $\langle c_{\mathbf{m}} \rangle$.

For reasons of simplicity, however, we restrict ourselves to the case where the components have the *special form* $g_i(\boldsymbol{z}) = z_i/e_i(\boldsymbol{z})$ with $\mathrm{L}\, e_i(\boldsymbol{z}) = 1$ for all $i \in \{1, \dots, s\}$, shortly written as $\mathbf{g}(\boldsymbol{z}) = \boldsymbol{z}/\boldsymbol{e}(\boldsymbol{z})$. Thus the multivariate form of Lagrange inversion appears as an analog to Corollary 43. It will be proved in Section 2.6.

47 **Theorem (Multivariate Lagrange Inversion)**
Let $f(\boldsymbol{z}) \in \mathbf{C}_s[[\boldsymbol{z}]]$ and $\boldsymbol{e}(\boldsymbol{z}) \in \mathbf{C}_s^s[[\boldsymbol{z}]]$ with $\mathrm{L}\, e_i(\boldsymbol{z}) = 1$ for all $i \in \{1, \ldots, s\}$, and let $\boldsymbol{g}(\boldsymbol{z}) = \boldsymbol{z}/\boldsymbol{e}(\boldsymbol{z})$. Then there is an $h \in \mathbf{C}_s[[\boldsymbol{z}]]$ such that $h(\boldsymbol{g}(\boldsymbol{z})) = f(\boldsymbol{z})$, and its coefficients $c_{\boldsymbol{m}}$ are

$$c_0 = \left[\boldsymbol{z}^0\right] \ f(\boldsymbol{z}),$$
$$[\forall \boldsymbol{m} > 0] \ c_{\boldsymbol{m}} = [\boldsymbol{z}^{\boldsymbol{m}}] \ f(\boldsymbol{z}) \, |\boldsymbol{g}'(\boldsymbol{z})| \, \boldsymbol{e}(\boldsymbol{z})^{\boldsymbol{m}+1}.$$

□ These equations are known as *Jacobi-Good formulas*.

2.2 Elementary Proof

The following proof is taken from *Hofbauer*'s paper [8]. But it is remarkable that in [7] even Henrici, writing on complex analysis, resorts to formal methods. Hofbauer's proof is essentially the same as Henrici's but the latter uses the Laurent series only in their circles of convergence. We begin the proof with a lemma.

48 **Lemma (Elementary° Residues)**
Let $g \in \mathbf{C}[[z]]$ with $\operatorname{ord} g = 1$. Then the residues of the elementary Laurent series g'/g^n are

$$\mathrm{M}\, \frac{g'(z)}{g(z)^{n+1}} = \delta_{n0}$$

□ for all $n \in \mathbf{Z}$.

49 **Proof (Elementary Residues)**
For $n \neq 0$, we have $g'(z)\, g(z)^{-n-1} = -\frac{1}{n}(g(z)^{-n})'$, which has residue zero because it is the derivative of some Laurent series.

For $n = 0$, we write $g(z) = z/e(z)$ with $\operatorname{ord} e = 0$, as in Corollary 43. With this substitution, we have

$$\frac{g'(z)}{g(z)} = \frac{1}{z} - \frac{e'(z)}{e(z)}.$$

The second term has residue zero because $1/e$ is a power series. Therefore
□ $\mathrm{M}\,(1/z) = 1$ is all that is left.

This lemma is the core piece in the elementary proof of Lagrange's Theorem as presented below.

50 **Proof (Lagrange Inversion, Elementary Approach)**
By definition of h, we have

$$f(z) = \sum_{n\geq 0} c_n \, g(z)^n.$$

For obtaining the first version, we determine the derivative

$$f'(z) = \sum_{n\geq 1} nc_n \, g(z)^{n-1} g'(z),$$

divide by $g(z)^m$ (with $m > 0$) and take the residue

$$[\forall m > 0] \;\; \mathrm{M}\, f'(z)\, g(z)^{-m} = \mathrm{M} \sum_{n\geq 1} nc_n \frac{g'(z)}{g(z)^{m-n+1}} = mc_m$$

by Lemma 48. For deducing the second version, we simply multiply $f(z)$ by $g'(z)\, g(z)^{-m-1}$ (even for $m \leq 0$) and again take the residue

$$[\forall m > 0] \;\; \mathrm{M}\, f(z)\, g'(z)\, g(z)^{-m-1} = \sum_{n\geq 0} c_n \, \mathrm{M} \frac{g'(z)}{g(z)^{m-n+1}} = c_m,$$

applying Lemma 48 once again.

The special case $c_0 = \mathrm{L}\, f$ is immediately apparent from the definition of h: only $c_0\, g(z)^0$ can contribute to the constant term, since $\operatorname{ord} g = 1$. □

2.3 Analytical Proof

The analytical proof presented here can be found in *Hurwitz* and *Courant*'s classical work on complex analysis, [9, p. 135–137]. Its aim is to find the inverse function h of an analytical function g by determining its Taylor coefficients. It turns out that the problem can be solved economically by considering the more general case of finding h such that $h \circ g = f$. The solution of this problem is called *Bürmann-Lagrange series* in their book.

Following Hurwitz and Courant's approach, we shall consider all our formal concepts in their *analytical context* for the scope of this section. Thus speaking of power series, we now mean a holomorphic mapping on $\mathbf{C}$, composing two power series means substitution in the analytical sense, etc. Furthermore, all the functions are to be considered only within their circles of convergence.

We first state a lemma that guarantees a unique solution h to the inversion problem $h \circ g = \mathrm{id}$ for given g but says nothing about its analyticity.

51 **Lemma (Local Uniqueness of Inverse Series)**
Let $D \subset \mathbf{C}$ be some open disk around the origin, $\overline{D}$ its closure and $D' := D \setminus \{0\}$. If g is a holomorphic function on $\overline{D}$ such that $g(0) = 0$ and $g(z) \neq 0$ for all $z \in D'$, then there is an open disk $E \subset \mathbf{C}$ around the origin such that g is bijective from D to E. □

52 **Proof (Local Uniqueness of Inverse Series)**
Since $g(z) \neq 0$ for all $z \in D'$ we know that $M := \min\{|g(\zeta)| : \zeta \in \partial D\}$ is not zero. (By ∂D we mean the boundary of D.) So we define E as the open disk around zero with radius M. Now choose some $w \in E$. Then, by definition of E and M, we have

$$[\forall \zeta \in \partial D] \;\; |w| < M \leq |g(\zeta)|.$$

It is an elementary fact (p. 107 in Hurwitz and Courant's book) that $\psi + \psi_0$ has the same number of zeroes as ψ in some disk whenever both ψ and ψ_0 are holomorphic on that disk and $|\psi_0| < |\psi|$ on the entire boundary. We can apply this principle here with $\psi(z) := g(z)$ and $\psi_0(z) := -w$ being a constant function. But we know that g has precisely one zero, namely $z = 0$. Thus there is also one solution $z \in D$ to the equation $g(z) = w$. Since $w \in E$ was arbitrary we have shown that g can be inverted on the whole disk E. □

With this lemma, we can finally present the analytical proof of Lagrange's Theorem. Comparing this to the elementary Proof 50, one can see how much *more laborious* it is to work with the analytical objects. But the result is also stronger since analytical power series represent actual complex functions, which can be evaluated at any point within their circle of convergence.

53 **Proof (Lagrange Inversion, Analytical Approach)**
For the present consideration, we must assume that the given power series f and g have an analytical meaning. Therefore let $D \subset \mathbf{C}$ be a disk like in the preceding lemma such that f and g are holomorphic on $\overline{D}$. Then we know that there exists a disk $E \subset D$ around the origin such that g is bijective from D to E. With C being the contour of D, oriented counter-clockwise, we define

$$[\forall w \in E] \;\; \tilde{h}(w) = \frac{1}{2\pi i} \oint_C f(\zeta) \, \frac{g'(\zeta)\, d\zeta}{g(\zeta) - w}.$$

By the residue theorem of complex analysis, the value of this integral equals the sum of the residues in the singularities. But fixing some $w \in E$, there is only one singularity $z := g^{-1}(w)$ because g is bijective from D to E. Since z is a first-order pole, we can compute the residue by

$$\begin{aligned}\tilde{h}(w) &= \lim_{\zeta \to z} \left[(\zeta - z) \left(f(\zeta) \, \frac{g'(\zeta)}{g(\zeta) - w} \right) \right] \\ &= \lim_{\zeta \to z} \left[f(\zeta) \, g'(\zeta) \Big/ \frac{g(\zeta) - g(z)}{\zeta - z} \right] = f(z),\end{aligned}$$

just using the definition of complex differentiation. Writing w as $g(z)$, this means that $\tilde{h}(g(z)) = f(z)$, which holds for all $z \in D$ because g^{-1} maps E bijectively to D. On D, $\tilde{h}$ thus meets the defining property of h, namely $h \circ g = f$. From our analytical viewpoint, we can therefore conclude $\tilde{h} = h$.

From the preceding lemma we know that $|w/g(\zeta)| < 1$ for all $w \in E$ and $\zeta \in \partial D$. Therefore we can expand the above integrand into a convergent geometric series on ∂D,

$$f(\zeta) \, \frac{g'(\zeta)}{g(\zeta)} \, \frac{1}{1 - w/g(\zeta)} = f(\zeta) \, \frac{g'(\zeta)}{g(\zeta)} \left(1 + \frac{w}{g(\zeta)} + \frac{w^2}{g(\zeta)^2} + \dots \right).$$

This series can be integrated by its summands because it converges uniformly (the part in parentheses is a Neumann series). Thus we get the power-series representation

$$h(w) = c_0 + c_1 \, w + c_2 \, w^2 + \dots$$

with the desired coefficients

$$[\forall m \in \mathbf{N}] \;\; c_m = \frac{1}{2\pi i} \oint_C f(\zeta) \, \frac{g'(\zeta)}{g(\zeta)^{m+1}} \, d\zeta.$$

The only singularity of this integral is $\zeta = 0$, which is the only zero of $g(\zeta)$. Therefore we can use the residue theorem again, obtaining the result

$$[\forall m \in \mathbf{N}] \;\; c_m = M \, f(z) \, g'(z) \, g(z)^{-m-1}$$

where M must now be interpreted analytically as the residue of the singular point 0, i.e. as the coefficient of z^{-1} in the Laurent series expanded around the origin (understood in the classical sense of complex analysis). This expression for c_m is precisely the Lagrange inversion formula in its second version. (The coefficient c_0 can also be written as L f by the same reason as given in the elementary Proof 50.)

For deriving the first version of the Lagrange inversion formula, we apply integration by parts on the above integral representation of the coefficients.

Excluding the case $m = 0$ (which is settled by $c_0 = L\,f$ as said above), we get

$$[\forall m > 0]\;\; c_m = -\frac{1}{2\pi i\, m} \oint_C f(\zeta)\,(g(\zeta)^{-m})'\,d\zeta$$
$$= \frac{1}{2\pi i\, m} \oint_C f'(\zeta)\,g(\zeta)^{-m}\,d\zeta.$$

The remaining contour integral can again be considered as residue

$$[\forall m > 0]\;\; c_m = \frac{1}{m} M\, f'(z)\,g(z)^{-m},$$

□ which is indeed the Lagrange inversion formula in its first version.

In [27, p. 167–168], *Wilf* gives a different analytical proof, basically using the substitution rule for complex integrals. His proof is short and elegant but only covers the implicit variant of Lagrange's Theorem (and hence only the first form of the inversion formulas).

2.4 Algebraic Proof

In his beautiful account on Lagrange inversion, Hofbauer not only gives the usual elementary proof as presented in Section 2.2 but in [8, p. 24–26] also sketches an elegant algebraic approach to the second form of the inversion formula. The idea is the same as in many other classical branches of mathematics: sometimes one manages to transfer a problem from analysis to algebra (linear algebra, to be precise) by viewing it in an appropriate vector space. For example boundary-value problems are often solved by finding the *eigenvectors* of their differential operators. If these are known, the boundary conditions can be satisfied by expanding the boundary function with respect to these eigenvectors—which is a purely algebraic problem.

Therefore one is led to considering *eigenvalue problems* $U g_m = \lambda_m g_m$ where U is a linear operator (like the differential operator from above) and g_m are its eigenvectors (like Hermite polynomials, Bessel functions or Legendre polynomials, just to mention a few). If the boundary function is f, one seeks the developing coefficients c_m in $f = \sum c_m g_m$. Usually one chooses a bilinear form $(.|.)$ such that the g_m are orthonormal; then we have $c_m = (f|g_m)$. Such bilinear forms of functions f and g usually are

$$(f|g) = \int_D f(x) g(x)\, w(x)\, dx$$

where D is a suitable domain and w some weight function. But if the basis $\{g_m\}$ is not orthonormal with respect to $(.|.)$, one has to resort to its *biorthonormal basis* $\{g_m^*\}$ so that now $c_m = (f|g_m^*)$. In certain cases, the biorthonormal base turns out to be the eigenbasis of the adjoint eigenvalue problem as will be shown below.

In the light of these considerations, we recognize Lagrange inversion as the special case of developing $f(z)$ with respect to $g_m(z) = g(z)^m$. In fact, we shall present a method which even allows these *more general expansions* as long as the $g_m(z)$ meet certain conditions. We will generalize the classical inversion formulas in yet another way: f (and therefore also the g_m) may be arbitrary Laurent series rather than power series. But since we will not rely on analytical concepts (as in the previous section) we first have to fix an appropriate bilinear form on the vector space.

54 Definition (Bilinear Form)
The vector space $\mathbf{C}((z))$ is equipped with the *bilinear form*

$$(f|g) = \mathrm{L}\, fg.$$

□ for all $f, g \in \mathbf{C}((z))$.

Bilinearity and also symmetry of this form are obvious. It must be noted, however, that $(.|.)$ is *not* a normal scalar product since it is neither positive nor definite. The former is violated for instance by $(z-1/z|z-1/z) = \mathrm{L}\,(z^2-2+1/z^2) = -2$, the latter by $(z|z) = \mathrm{L}\, z^2 = 0$. Note that in complex vector spaces, scalar products are usually defined as *sesquilinear forms*. This guarantees positiveness but spoils symmetry. As an example, for all $\mathbf{x}, \mathbf{y} \in \mathbf{C}^n$ one defines $\mathbf{xy} = \sum x_i \bar{y}_i$ where the bar denotes the complex-conjugate. Then we have $\mathbf{xx} \geq 0$ but $\mathbf{xy} = \overline{\mathbf{yx}}$. Since our bilinear form $(.|.)$ is not positive in the first place, we can leave it symmetric. This is the reason why we did not define it as $\mathrm{L}\, f\bar{g}$.

Having available a bilinear form, we can introduce the *adjoint* of a linear operator in the usual way—except for the complex-conjugate. (Since Schmidt-orthonormalization also works for non-positive, non-definite bilinear forms we can expand a vector with respect to orthonormal bases just as if we had an ordinary scalar product. Thus for defining an operator on the whole space, it is sufficient to fix all the bilinear forms.)

55 Definition (Adjoint Operator)
Let S be any operator on $\mathbf{C}((z))$. Then its *adjoint* S^* is defined by

$$[\forall f, g \in \mathbf{C}((z))]\ \ (Sf|g) = (f|S^*g)\,,$$

being again an operator on $\mathbf{C}((z))$. (From now on, any operator is understood as a linear mapping on $\mathbf{C}((z))$ unless otherwise stated.) □

These concepts are all we need for the algebraic generalization of Lagrange inversion. We like to mention, though, that Hofbauer [8, p. 25] actually allows a slightly more general setting by considering $\tilde{U} g_m = \lambda_m \tilde{V} g_m$. There $\tilde{U}$ is an arbitrary linear operator and $\tilde{V}$ a bijective one. Our eigenvalue problem is the special case $\tilde{V} = I$. Of course the general problem can immediately be transformed to the restricted version $U g_m = \lambda_m g_m$ by setting $U = \tilde{U}\tilde{V}^{-1}$. We have chosen the restricted version because it seems to make things more transparent.

56 Theorem (Generalized Lagrange Inversion)
Let $f \in \mathbf{C}((z))$ be arbitrary, $m_0 = \operatorname{ord} f$, and $g_m \in \mathbf{C}((z))$ such that we have $\operatorname{ord} g_m = m$ for all $m \geq m_0$. Choose an arbitrary linear operator U such that the g_m satisfy the eigenvalue problem $U g_m = \lambda_m g_m$ with distinct eigenvalues $\lambda_m \in \mathbf{C}$. If the coefficients c_m are now defined by

$$f(z) = \sum_{m \geq m_0} c_m \, g_m(z),$$

their values can be found from

$$[\forall m \geq m_0] \;\; c_m = (f | g_m^*)$$

with g_m^* being the eigenvectors of the adjoint problem $U^* g_m^* = \lambda_m g_m^*$. □

57 Proof (Generalized Lagrange Inversion)
Let $m, n \geq m_0$ be arbitrary but fixed. Multiplying $U g_m = \lambda_m g_m$ bilinearly with g_n^*, we get $(U g_m | g_n^*) = \lambda_m \, (g_m | g_n^*)$; multiplying $U^* g_n^* = \lambda_n g_n^*$ with g_m, we get $(g_m | U^* g_n^*) = \lambda_n \, (g_m | g_n^*)$. By definition of the adjoint operator, the left sides are equal. Then we also have $\lambda_m \, (g_m | g_n^*) = \lambda_n \, (g_m | g_n^*)$. This means that $(g_m | g_n^*) = \delta_{mn}$ because the eigenvalues are all distinct. In other words: the eigenbasis $\{g_m^*\}$ of the adjoint problem is indeed biorthonormal with respect to the eigenbasis $\{g_m\}$ of the original problem.

The desired coefficients can now easily be obtained by multiplying g_m^* bilinearly on their defining equation, which yields

$$[\forall m \geq m_0] \;\; (f | g_m^*) = \sum_{n \geq m_0} c_n \, (g_n | g_m^*) = \sum_{n \geq m_0} c_n \delta_{nm} = c_m$$

as claimed in the theorem. □

For deriving classical Lagrange inversion from this, we just have to find a suitable operator U in the sense of the foregoing theorem. Rewriting the coefficient relation will bring us back to the familiar inversion formula.

58 **Proof (Lagrange Inversion, Algebraic Approach)**
Let $m_0 = \operatorname{ord} f$. Then for all $m \geq m_0$, we define $g_m(z) = g(z)^m$ so that $\operatorname{ord} g_m = m$ as required in Theorem 56. We observe that $m_0 \geq 0$ since f is assumed to be a power series. The desired coefficients of Theorem 42 are defined by $h \circ g = f$ or, in the form of Theorem 56,

$$\sum_{m \geq m_0} c_m \, g_m(z) = f(z)$$

because $m_0 = \operatorname{ord} f = \operatorname{ord}(h \circ g) = (\operatorname{ord} h)\ (\operatorname{ord} g) = \operatorname{ord} h$, using the input condition $\operatorname{ord} g = 1$ imposed on g in Theorem 42.

In order to find an appropriate operator U for the g_m, we differentiate their defining equation so that

$$[\forall m \geq m_0]\ \ g_m'(z) = (g(z)^m)' = m\, g(z)^{m-1}\, g'(z) = m\, g_m(z)\, \frac{g'(z)}{g(z)}.$$

Defining $Ua = (g/g')\, a'$ for all $a \in \mathbf{C}((z))$ and $\lambda_m = m$ for all $m \geq m_0$, this reads $Ug_m = \lambda_m g_m$. Therefore Theorem 56 is applicable and yields the desired coefficients $c_m = (f|g_m^*)$ with $\{g_m^*\}$ being the eigenbasis of the adjoint problem.

For finding the adjoint operator, we need some auxiliary calculations. First we observe that any multiplication operator M is self-adjoint. For if it is defined by $[\forall a \in \mathbf{C}((z))]\ Ma(z) = m(z)\, a(z)$ with some $m(z) \in \mathbf{C}[[z]]$, we have

$$[\forall a, b \in \mathbf{C}((z))]\ \ (Ma|b) = L\,(ma)b = L\, a(mb) = (a|Mb)\,.$$

Second we compute the adjoint of the normalized° differential operator $\bar{D}$, defined by $[\forall a \in \mathbf{C}((z))]\ \bar{D}a(z) = za'(z)$. Fix arbitrary $a, b \in \mathbf{C}((z))$. By the product rule, we know that $(a(z)b(z))' = a'(z)b(z) + a(z)b'(z)$. Multiplying with z, this gives $\bar{D}(ab) = (\bar{D}a)b + a(\bar{D}b)$. The residue of any derivative vanishes; therefore in the normalized derivative there is no constant term and we have

$$0 = L\,\bar{D}(ab) = L\,(\bar{D}a)b + L\, a(\bar{D}b) = \big(\bar{D}a|b\big) + \big(a|\bar{D}b\big)$$

or $\big(\bar{D}a|b\big) = \big(a| -\bar{D}b\big)$, which means $\bar{D}^* = -\bar{D}$.

The operator U can be written as $M\bar{D}$ after defining $[\forall a \in \mathbf{C}((z))]\ Ma(z) = [g(z)/zg'(z)]\, a(z)$. Therefore its adjoint is

$$U^* = (M\bar{D})^* = \bar{D}^* M^* = -\bar{D}M,$$

using the familiar rule for the adjoint of products. Now we can determine the g_m^* from the adjoint eigenvalue problem $-\bar{D}M\, g_m^*(z) = m\, g_m^*(z)$. Via the substitution $\tilde{g}_m = M g_m^*$, this differential equation can be transformed into the simpler form

$$(\tilde{g}_m(z))' = -m\, \tilde{g}_m(z)\, \frac{g'(z)}{g(z)}.$$

By comparison with with the differential equation for g_m, we find the solution $\tilde{g}_m(z) = g(z)^{-m}$ and thus $g_m^*(z) = M^{-1} g(z)^{-m} = z\, g(z)^{-m-1} g'(z)$. Note that the biorthonormality relation $[\forall m, n \geq m_0]\ (g_m | g_n^*) = \delta_{mn}$ leads back to the Elementary-Residue Lemma 48.

The coefficients c_m can now be computed as

$$\begin{aligned} [\forall m \geq m_0]\ c_m = (f|g_m^*) &= L\, z\, f(z)\, g'(z)\, g(z)^{-m-1} \\ &= M\, f(z)\, g'(z)\, g(z)^{-m-1}, \end{aligned}$$

□ which is precisely the Lagrange inversion formula in its second version.

Compared to the elementary proof, it is quite a long way to the final result. What we gain, though, is not only an algebraic interpretation but also a rather beautiful generalization of the classical inversion formulas. The main effort of the proof was to apply this generalization to the special case $g_m(z) = g(z)^m$ for recovering the inversion formulas in their original form.

2.5 Combinatorial Proof

Lagrange's Theorem can be proved combinatorially in a very elegant way by using Joyal's *species theory* as sketched in Section 1.4. For that, we will follow the proof in Labelle's beautiful paper [13], which is also the source of the figures in this section. We are grateful for his generous permission to use them in our thesis.

The *main idea* of the proof is based on the observation that Lagrange's inversion formula is equivalent to a profound generalization of Cayley's tree formula. In its classical version, this formula states that there are n^{n-1} labeled rooted trees on n labels; see for example [1, p. 331]. The connection to Lagrange's Theorem is not surprising because Pólya's classical proof of this formula derives a functional equation for the corresponding generating function. Using Lagrange inversion, this equation is solved and thus the desired n^{n-1} are computed (see Example 75). Since one can apply other methods—bypassing the use of Lagrange inversion—for proving Cayley's formula, there must also be a way to recover Lagrange's inversion formula from it. And indeed there is.

But first we have to introduce some *important species* that we shall need frequently in this section—the afore-mentioned rooted trees and the so-called endofunctions. (From now on, we shall drop the attribute "labeled" for all tree-like species because in species theory, this is the standard case.)

59 Definition (Rooted Trees, Endofunctions)
We introduce the species $\mathfrak{T}$ of rooted trees with $\mathfrak{T}(U)$ containing all connected, acyclic graphs on a label set U, where one element $r \in U$ is distinguished (called *root*). If U is empty, there is no rooted tree.

The *species of endofunctions* $\mathfrak{F}$ is defined as $\mathfrak{F}(U) = \{f \mid f\colon U \to U\}$ for all label sets U. This also includes the empty function (from $\emptyset$ to $\emptyset$). □

The presence of a root in a tree allows two natural orientations of the edges—either diverging from the root or converging toward it: see Figure 1, where the star represents the root. Following Labelle [13], we choose the second alternative as *standard orientation* (right picture in Figure 1). If a tree node a points to another tree node b, then a will be called the predecessor of b and likewise b the successor of a.

Figure 1: Tree Orientations

If a rooted tree has another distinguished label (called tail) besides the root (here called head), it is called a *vertebrate*. This name is due to the vertebral column, uniquely connecting head and tail in a vertebrate (see Figure 2, where the vertebral column is represented by a dashed line). Since a vertebrate emerges from a rooted tree by marking an additional label, this new species is really just $\mathfrak{T}^\bullet$.

A set of λ rooted trees is called a *rooted* λ*-forest*. In the language of species theory, we can construct such a forest by $\mathfrak{T}^\lambda/\lambda!$. The species $\mathfrak{T}^\lambda$ consists of ordered arrangements of rooted trees; since there are $\lambda!$ different orders for every λ-set, we have to divide by this number. We will write

$$(\mathfrak{T}^\lambda/\lambda!)\,(z) = \frac{\mathfrak{T}(z)^\lambda}{\lambda!} = \sum_{m\geq\lambda} \mathfrak{T}^\lambda_m \frac{z^m}{m!}$$

for the generating function.

Figure 2: Vertebral Column

As mentioned in the beginning of the present section, we have to generalize the above concept of a rooted tree and also that of an endofunction. The idea is basically the same as with the enriched colorings, described in Section 1.4. *Enriching* a rooted tree or an endofunction simply means imposing an additional structure on each of its fibres. (Observe that a rooted tree t can also be considered a function by interpreting its arrows as a mapping: the fibre $t^*(k)$ of a tree node k is the set of its predecessors.)

60 Definition (Enriched Rooted Trees, Enriched Endofunctions)
If $\mathfrak{r}$ is an arbitrary base species°, the species $\mathfrak{T}_\mathfrak{r}$ of $\mathfrak{r}$*-enriched rooted trees* is defined as follows. If U is any label set, then a generic structure of $\mathfrak{T}_\mathfrak{r}(U)$ is a rooted tree $t \in \mathfrak{T}(U)$, bearing a structure $r_l \in \mathfrak{r}(t^*(l))$ for every label $l \in U$.

For any base species $\mathfrak{r}$, the species $\mathfrak{F}_\mathfrak{r}$ of $\mathfrak{r}$*-enriched endofunctions* is defined in a similar fashion. If U is any label set, then a generic structure of $\mathfrak{F}_\mathfrak{r}(U)$ is an endofunction $f \in \mathfrak{F}(U)$, bearing a structure $r_l \in \mathfrak{r}(f^*(l))$ for every label $l \in U$. □

A *typical example* of an enriched rooted tree and an enriched endofunction is shown in Figure 3 and Figure 4, respectively. The structures imposed on the fibres are symbolized by arcs. The dashed arrows in Figure 4 represent the cycles of the depicted endofunction.

Figure 3: Tree

Figure 4: Endofunction

Just as with the colorings, the conventional non-enriched species $\mathfrak{T}$ and $\mathfrak{F}$ can be regained by choosing $\mathfrak{r} = \mathfrak{E}$. There are some other common *types of rooted trees* that can be obtained by an appropriate choice of $\mathfrak{r}$, for example (see [12, p. 221]): cyclic rooted trees (take $\mathfrak{r}$ as the species of cyclic permutations), binary rooted trees ($\mathfrak{r} = \mathfrak{1} + \mathfrak{Z}^2/2$), plane rooted trees ($\mathfrak{r} = \mathfrak{L}$), and permutation rooted trees ($\mathfrak{r} = \mathfrak{S}$).

The structures from $\mathfrak{T}_\mathfrak{r}^\bullet$ are called $\mathfrak{r}$-*enriched vertebrates*. They are like $\mathfrak{r}$-enriched rooted trees, but they have two distinguished labels (head and tail); see Figure 5. Likewise, the concept of forests is transferred from the non-enriched case by calling $\mathfrak{T}_\mathfrak{r}^\lambda/\lambda!$ the species of $\mathfrak{r}$-*enriched rooted* λ-*forests*. A structure of this species is simply a set of λ disjoint $\mathfrak{r}$-enriched rooted trees.

Figure 5: Vertebrate

The *enriched contractions* are the last species we like to introduce before we present the main lemma, which establishes the relations between $\mathfrak{T}_\mathfrak{r}$, $\mathfrak{F}_\mathfrak{r}$, and $\mathfrak{r}$. This time we choose to define the new species entirely in terms of other species, but we shall reveal its combinatorial meaning while proving the main lemma.

61 Definition (Enriched Contractions)
If $\mathfrak{r}$ is an arbitrary base species, we define the species of $\mathfrak{r}$-*enriched contractions* as $\mathfrak{Q}_\mathfrak{r} = \mathfrak{Z}\,(\mathfrak{r}' \circ \mathfrak{T}_\mathfrak{r})$. □

62 Lemma (Relations Enriched Species / Base Species)
Let $\mathfrak{r}$ be any base species. Then the enriched species $\mathfrak{T}_\mathfrak{r}$ and $\mathfrak{F}_\mathfrak{r}$ satisfy the

following properties.

1. The species $\mathfrak{T}_{\mathfrak{r}}$ is characterized by the isomorphism
$$\mathfrak{T}_{\mathfrak{r}} \simeq \mathfrak{Z}\,(\mathfrak{r} \circ \mathfrak{T}_{\mathfrak{r}}).$$
This characterization is unique up to equipotence.

2. The species $\mathfrak{T}_{\mathfrak{r}}$ and $\mathfrak{F}_{\mathfrak{r}}$ are linked with the species $\mathfrak{Q}_{\mathfrak{r}}$ by the isomorphisms
   (a) $\mathfrak{T}_{\mathfrak{r}}^{\bullet} \simeq (\mathfrak{L} \circ \mathfrak{Q}_{\mathfrak{r}})\,\mathfrak{T}_{\mathfrak{r}}$,
   (b) $\mathfrak{F}_{\mathfrak{r}} \simeq \mathfrak{S} \circ \mathfrak{Q}_{\mathfrak{r}}$.

3. The species $\mathfrak{T}_{\mathfrak{r}}$ and $\mathfrak{F}_{\mathfrak{r}}$ are related by the equipotence
$$\mathfrak{T}_{\mathfrak{r}}^{\bullet} \sim \mathfrak{T}_{\mathfrak{r}}\,\mathfrak{F}_{\mathfrak{r}}.$$

These are the basic structural relations between the enriched species $\mathfrak{T}_{\mathfrak{r}}$ and $\mathfrak{F}_{\mathfrak{r}}$ and the base species $\mathfrak{r}$. □

Figure 6: Tree Shape

Figure 7: Contraction

63 Proof (Relations Enriched Species / Base Species)
We prove the lemma statement by statement.

1. The claimed isomorphism is immediately apparent from a glance at Figure 6. The species product is graphically represented by the ragged boundary line between the root and the branches; the species composition is represented by the circled A's. (Labelle's paper is written in French, where rooted trees are named "arborescences".)

   Note that according to Definition 59, we have $\mathfrak{T}_\mathfrak{r}(\emptyset) = \emptyset$, this property being inherited from the base species $\mathfrak{T}$. Thus the composition $\mathfrak{r} \circ \mathfrak{T}_\mathfrak{r}$ is feasible by Definition 35, point 3.

   Now if U is an appropriate label set, the generic structure $t \in (\mathfrak{Z}\,(\mathfrak{r} \circ \mathfrak{T}_\mathfrak{r}))\,(U)$ depicted in Figure 6 is

   $$\Big(z, (r; \{t_1, t_2, t_3, t_4\})\Big).$$

   Here we have $t_i \in \mathfrak{T}_\mathfrak{r}(U_i)$ for $i \in \{1, 2, 3, 4\}$, and $r \in \mathfrak{r}(\{U_1, U_2, U_3, U_4\})$ with $\{z\} \dot\cup U_1 \dot\cup U_2 \dot\cup U_3 \dot\cup U_4 = U$, as required. On the other hand, the structure t can also be viewed an $\mathfrak{r}$-enriched rooted tree, having its root in z.

   For proving the uniqueness property, consider any given species fulfilling the isomorphism of point 1. Then its generating function must satisfy the same functional equation as $\mathfrak{T}_\mathfrak{r}(z)$. Its solution is unique, because the coefficients can be computed recursively from the functional equation. Hence the given species is equipotent to $\mathfrak{T}_\mathfrak{r}$.

2. (a) Figure 5 shows a generic vertebrate from $\mathfrak{T}_\mathfrak{r}^\bullet$, having its head in the star-shaped label and its tail in the black-square one. It can be decomposed canonically along the vertebral column (Figure 8) and then cut up into a chain of new structures (Figure 9). Except for the last, which is a pure $\mathfrak{T}_\mathfrak{r}$-structure, they are all of the form depicted in Figure 7. But this is just a generic $\mathfrak{Q}_\mathfrak{r}$-structure q. Again note that the composition $\mathfrak{r}' \circ \mathfrak{T}_\mathfrak{r}$ occurring in the definition

Figure 8: Vertebrate Decomposed

Figure 9: Vertebrate Cut Up

of $\mathfrak{Q}_\mathfrak{r}$ is permitted by the same reason as in point 1. The species $\mathfrak{Q}_\mathfrak{r}$ itself is also substitutable because $\mathfrak{Q}_\mathfrak{r}(\emptyset) = \emptyset$. If that was not so, the generating function $\mathfrak{Q}_\mathfrak{r}(z) = z\,\mathfrak{r}'(\mathfrak{T}_\mathfrak{r}(z))$ would have a constant term, which is clearly impossible.

If U is a suitable label set, then the depicted contraction $q \in \mathfrak{Q}_\mathfrak{r}(U) = (\mathfrak{Z}\,(\mathfrak{r}' \circ \mathfrak{T}_\mathfrak{r}))\,(U)$ is the structure

$$\Big(z, (r'; \{t_1, t_2, t_3\}) \Big),$$

where $t_1 \in \mathfrak{T}_\mathfrak{r}(U_1), t_2 \in \mathfrak{T}_\mathfrak{r}(U_2), t_3 \in \mathfrak{T}_\mathfrak{r}(U_3)$ and furthermore $r' \in \mathfrak{r}(\{U_1, U_2, U_3, \{U\}\})$ with $\{z\} \mathbin{\dot\cup} U_1 \mathbin{\dot\cup} U_2 \mathbin{\dot\cup} U_3 = U$, as required. The "new" element $\{U\}$, added to U_1, U_2, U_3 as a label for r' is the formalization of the dashed arrow in Figure 7.

Thus a $\mathfrak{T}_\mathfrak{r}^\bullet$-structure corresponds to a linear ordering of $\mathfrak{Q}_\mathfrak{r}$-structures together with one $\mathfrak{T}_\mathfrak{r}$-structure. This is precisely what the claimed isomorphism says.

(b) A generic enriched endofunction is shown in Figure 10, where the enriched rooted trees occurring as parts in Figure 4 have been abbreviated. They are arranged in two cycles—one bearing three enriched rooted trees, the other one two. In the same way, any enriched endofunction can be viewed as a permutation of enriched rooted trees (canonically decomposed in cycles), hence we have $\mathfrak{F}_\mathfrak{r} \simeq \mathfrak{S} \circ \mathfrak{Q}_\mathfrak{r}$.

At this point, we are able to understand the combinatorial meaning of enriched contractions. Looking at the enriched endofunction of Figure 4, we see that every label is eventually mapped into one of the cycles; there it is mapped around circularly. If there is only one cycle consisting of a single label l, then every label will be mapped eventually onto l. An enriched contraction is merely such an eventually-constant enriched endofunction. This

Figure 10: Endofunction Abbreviated

Figure 11: Auxiliary Species

can be seen from the generic representant in Figure 7 if the dashed arrow is canonically replaced by a circular arrow.

3. As pointed out after Definition 32, the species $\mathfrak{S}$ and $\mathfrak{L}$ are equipotent (though not isomorphic). Together with point 2a and 2b, this implies $\mathfrak{T}_\mathfrak{r}^\bullet \simeq (\mathfrak{L} \circ \mathfrak{Q}_\mathfrak{r})\, \mathfrak{T}_\mathfrak{r} \sim (\mathfrak{S} \circ \mathfrak{Q}_\mathfrak{r})\, \mathfrak{T}_\mathfrak{r} \simeq \mathfrak{F}_\mathfrak{r}\, \mathfrak{T}_\mathfrak{r}$, as claimed in the lemma.

□ This finishes the proof.

As mentioned at the beginning of the present section, the Lagrange inversion formula turns out to be equivalent to a generalization of *Cayley's tree formula*: it allows to count enriched rooted forests (not only ordinary ones and not only rooted trees) as the following proposition states.

64 Proposition (Cayley's Generalized Tree Formula)
Let $\mathfrak{r}$ be an arbitrary species and $\lambda \in \mathbf{N}^*$. Then there are

$$\frac{\lambda}{n}\binom{n}{\lambda}(\mathfrak{r}^n)_{n-\lambda}$$

□ different $\mathfrak{r}$-enriched rooted λ-forests on $n \geq \lambda$ labels.

65 Proof (Cayley's Generalized Tree Formula)
First let us define the following auxiliary species $\mathfrak{f}_\mathfrak{r}^\lambda$. If U is a finite label set of arbitrary cardinality $n \in \mathbf{N}^*$, a generic structure from $\mathfrak{f}_\mathfrak{r}^\lambda(U)$ shall be given by

- a subset $\Lambda \subset U$ of cardinality λ,
- a distinguished point $\omega \in \Lambda$, and
- a function $f\colon U \backslash \Lambda \to U$ with an $\mathfrak{r}$-structure on each (possibly empty) fibre.

Figure 12: Pointed Forest

Figure 13: Equipotent Species

Figure 11 shows a typical representant of this species. We will now show that $(\mathfrak{T}_\mathfrak{r}^\lambda/\lambda!)^\bullet \sim \mathfrak{f}_\mathfrak{r}^\lambda$. Consider the $\mathfrak{T}_\mathfrak{r}^\lambda/\lambda!^\bullet$-structure in Figure 12, where the pointed element is symbolized by a boxed label. By pointing, one of the λ trees of the forest has become a vertebrate (with the dashed line in Figure 12 representing the vertebral column). According to the equipotence of Lemma 62, point 3, let us replace it by a rooted tree combined with an endofunction (see Figure 13). The structure in the figure can obviously be regarded as an $\mathfrak{f}_\mathfrak{r}^\lambda$-structure, Λ being the set of roots (represented by stars in the figure) and ω being the new root coming from the equipotence (the boxed label in the figure). For the generating functions, this means

$$(\mathfrak{T}_\mathfrak{r}^\lambda/\lambda!)^\bullet(z) = \sum_{m\geq\lambda} m\,(\mathfrak{T}_\mathfrak{r}^\lambda)_m \frac{z^m}{m!} = \sum_{m\geq\lambda} (\mathfrak{f}_\mathfrak{r}^\lambda)_m \frac{z^m}{m!} = \mathfrak{f}_\mathfrak{r}^\lambda(z).$$

Thus for finding the $(\mathfrak{T}_\mathfrak{r}^\lambda)_m$, it suffices to compute the $(\mathfrak{f}_\mathfrak{r}^\lambda)_m$. For constructing an arbitrary $\mathfrak{f}_\mathfrak{r}^\lambda$-structure on U, there are

- first $\binom{n}{\lambda}$ choices for the subset Λ,
- then λ choices for the distinguished point ω,
- finally $(\mathfrak{r}^n)_{n-\lambda}$ choices for the function $f\colon U\backslash\Lambda \to U$ with $\mathfrak{r}$-structures on the fibres (since it can be seen as an $\mathfrak{r}$-enriched n-coloring of $U\backslash\Lambda$).

This gives a total of $n\,(\mathfrak{T}_\mathfrak{r}^\lambda)_n = (\mathfrak{f}_\mathfrak{r}^\lambda)_n = \binom{n}{\lambda}\,\lambda\,(\mathfrak{r}^n)_{n-\lambda}$ different $\mathfrak{f}_\mathfrak{r}^\lambda$-structures on U and thus $(\mathfrak{T}_\mathfrak{r}^\lambda)_n = (\lambda/n)\binom{n}{\lambda}(\mathfrak{r}^n)_{n-\lambda}$. □

Of course we come back to the *classical version* of Cayley's tree formula by setting $\lambda = 1$ and by choosing $\mathfrak{r}$ to be the uniform species $\mathfrak{E}$, which actually means to drop the enriching, thus counting ordinary rooted trees. Because of

$$(e^z)^n = e^{nz} = \sum_{m\geq 0} n^m \frac{z^m}{m!},$$

we obtain $\mathfrak{T}_n = (1/n)\binom{n}{1}(\mathfrak{E}^n)_{n-1} = n^{n-1}$, which is indeed the classical result of Cayley.

Now it is only a little step (though a bit of writing) to the *implicit form of Lagrange inversion*, given by Corollary 45. As we have shown in Section 2.1, the latter is equivalent to the classical version of Theorem 42 but only gives the first form of the inversion formulas.

66 Proof (Lagrange Inversion, Combinatorial Approach)
We are confronted with the functional equation $G(z) = z\,e(G(z))$ to be solved for $G(z)$, with $e(z)$ being a given power series of order 0. Then we have to find the expansion coefficients c_m of $f(G(z))$ for an arbitrary $f(z) \in \mathbf{C}[[z]]$.

Actually it suffices, though, to consider only series of the form $f(z) = z^\lambda/\lambda!$ with λ running through $\mathbf{N}$, because they form a basis of the vector space $\mathbf{C}[z]$. This applies the *principle of potential infinity*°: For showing an identity between two formal power series (the desired coefficient formula can be seen as such an identity), we truncate the series at some index. This leaves us with an identity between two polynomials; it must be valid, because we have proved it on a basis. Since we can truncate the series at an arbitrary index (so the set of these indices is potentially infinite), the identity between the series is also valid. So choose an arbitrary $\lambda \in \mathbf{N}^*$ and set $f(z) = z^\lambda/\lambda!$ for the rest of the proof. (The case $\lambda = 0$ of course gives the constant term of f.)

At first, let us assume that the coefficients e_m in $e(z) = \sum e_m z^m$ are from $\mathbf{N}$. Then it follows by induction that the coefficients g_m in $G(z) = \sum g_m z^m$ are also from $\mathbf{N}$: By looking at the functional equation, we realize that $\operatorname{ord} G = 1$, so induction starts at $m = 1$. For this initial value, we have $g_1 = e_0 \in \mathbf{N}$, hence the induction basis is shown. Next we assume for any $k \in \mathbf{N}^*$ that $g_1, \ldots, g_k$ are from $\mathbf{N}$. The functional equation tells us that $g_{k+1} = [z^k]\; e(G(z))$, therefore only $g_1, \ldots, g_k$ can contribute to g_{k+1}, and all occurring coefficients are from $\mathbf{N}$. As a consequence, g_{k+1} also must be from $\mathbf{N}$, which concludes the induction step.

With the above assumption, we know that both $e(z)$ and $G(z)$ have coefficients from $\mathbf{N}$ when regarded as ordinary generating functions. A fortiori, this must be true when they are regarded as exponential generating functions (as in species theory) because the n-th coefficient of the latter is $n!$ times greater than its ordinary counterpart. Employing the principle of combinatorial back-interpretation (stated after Proof 37), we can regard

$e(z)$ and $G(z)$ as the generating functions of some species, which we shall call $\mathfrak{r}$ and $\mathfrak{T}_{\mathfrak{r}}$, respectively. (Of course, combinatorial back-interpretation can only be applied to polynomials. But this is enough, because all the identities can be reduced to polynomials by truncating at some index. By the principle of potential infinity, validity extends to the full series.)

In the language of species theory, the given functional equation now reads $\mathfrak{T}_{\mathfrak{r}} \simeq \mathfrak{Z}\,\mathfrak{r}(\mathfrak{T}_{\mathfrak{r}})$, which happens to be the species isomorphism characterizing the $\mathfrak{r}$-enriched trees; confer Lemma 62, point 1. This identification is unique only up to equipotence—but this is all we will need.

By these observations, our initial problem has changed into a combinatorial one. As announced above, finding the coefficients c_m of $f(G(z)) = \mathfrak{T}_{\mathfrak{r}}^{\lambda}(z)/\lambda!$ really amounts to counting the number of $\mathfrak{r}$-enriched λ-forests on m labels. And this question has been answered in Proposition 64 so that

$$[\forall m \geq \lambda]\;\; m!\, c_m = \frac{\lambda}{m} \binom{m}{\lambda} (\mathfrak{r}^m)_{m-\lambda}.$$

Since the generating function of $\mathfrak{r}^m$ (like of any species) is of exponential nature we have $(\mathfrak{r}^m)_{m-\lambda} = (m-\lambda)!\,\left[z^{m-\lambda}\right] e(z)^m$. Furthermore, using the absorption property of the coefficient functional yields $\left[z^{m-\lambda}\right] e(z)^m = \left[z^{m-1}\right]\, z^{\lambda-1}\, e(z)^m$. With these rewritings, the above formula becomes

$$\begin{aligned}[\forall m \geq \lambda]\;\; c_m &= \frac{1}{m!}\, \frac{\lambda}{m}\, \frac{m!}{\lambda!\,(m-\lambda)!}\, (m-\lambda)!\, \left[z^{m-1}\right] z^{\lambda-1}\, e(z)^m \\ &= \frac{1}{m} \left[z^{m-1}\right] \frac{z^{\lambda-1}}{(\lambda-1)!}\, e(z)^m = \frac{1}{m} \left[z^{m-1}\right] f'(z)\, e(z),\end{aligned}$$

which is indeed the formula given in Corollory 45. For $m < \lambda$, our combinatorial interpretation tells us that $c_m = 0$ since there are no λ-forests if there are less than λ labels. This is also in accordance with Corollary 45, which gives

$$\begin{aligned}[\forall m < \lambda]\;\; c_m &= \frac{1}{m} \left[z^{m-1}\right] f'(z)\, e(z)^m = \frac{1}{m} \left[z^{m-1}\right] \frac{z^{\lambda-1}}{(\lambda-1)!}\, e(z)^m \\ &= \frac{1}{m\,(\lambda-1)!} \left[z^{m-\lambda}\right] e(z)^m = 0\end{aligned}$$

in this case because $\operatorname{ord} e(z) = 0$.

At this point, we have finished the proof for the case of nonnegative integer coefficients in $e(z) = \sum e_m\, z^m$. Strictly speaking, this is all the combinatorial approach can do. But usually (and especially in the present situation), one can remove the integer condition by applying the well-known *polynomial argument* to the coefficient formulas of Corollary 45. For $m = 0$, no extension is necessary, because $e(z)$ does not occur at all.

If $m > 0$ is fixed, the right-hand side of the coefficient formula can be viewed as a polynomial in in the m variables $e_0, \dots, e_{m-1}$. The same is true for the left-hand side

$$c_m = [z^m]\, f(G(z)) = [z^m]\, \frac{G(z)^\lambda}{\lambda!} = \frac{1}{\lambda!}\, [z^m]\; z^\lambda\, e(G(z))^\lambda.$$

Because of $\operatorname{ord} G(z) = 1$, every e_n with $n \geq m$ contributes to that expression only powers higher than $z^\lambda z^{n\lambda}$, which vanish under $[z^m]$. (This also holds for the degenerate case $\lambda = 0$, where $c_m = [z^m]\, 1 = 0$.)

For every fixed $m > 0$, the coefficient formula thus turns out to be an identity between two polynomials in $e_0, \dots, e_{m-1}$, valid for all nonnegative integers. But now we can extend the range of every such e_i (with i between 0 and $m-1$) from $\mathbf{N}$ to $\mathbf{C}$. Regarding the other variables as constants, the coefficient formula states that two polynomials in the variable e_i are equal at all the points $e_i \in \mathbf{N}$. By the Fundamental Theorem of Algebra, this implies that the polynomials are equal at all points $e_i \in \mathbf{C}$. (It would even be sufficient to have equality in $(\deg(g-h)+1)$ points, if the two polynomials are g and h.) □

The proof and interpretation of Lagrange inversion as given in this section only scratches the surface of a large body consisting of diverse combinatorial views, generalizations, and connections. As a primary example, consider *Strehl's parametrized inversion formula*, presented in his habilitation thesis [26]. It is a generalization of the implicit form of Lagrange inversion, Corollary 45, allowing the development of

$$\frac{f(G(z))}{[\,1 - z\, e'(G(z))\,]^{\lambda+1}},$$

not only of $f(G(z))$ as in the mentioned corollary (which corresponds to $\lambda = -1$), by determining the coefficients c_m as

$$[\forall m \in \mathbf{N}]\; c_m = \sum_{k+l=m} \frac{\lambda^{\overline{k}}}{k!}\, [z^l]\; F(z)\, e'(z)^k\, e(z)^l.$$

The symbol $\lambda^{\overline{k}}$ is defined as $\lambda\,(\lambda+1)\cdots(\lambda+k-1)$ for $k > 0$ and as 1 for $k = 0$. In analogy to the falling factorials introduced in Example 3, they are known as rising factorials and spoken "λ to the k rising".

Note that Strehl's parametrized Lagrange formula not only extends the range of applicability but also offers a combinatorial interpretation of the new parameter λ. And this is only the beginning of his deep analysis pertaining to cycle enumeration of so-called partial functions. As in all adequate and proper combinatorial generalizations of Lagrange inversion, it reveals a rich complex of *new aspects* as well as a broad range of applications.

2.6 Multivariate Proof

The multivariate Lagrange inversion formulas of Theorem 47 will be proved by an *elementary approach* similar to the one of Section 2.2. We follow Hofbauer's paper [8, p. 6–9] for this purpose, but we will dispense with the use of alternating differential forms.

Just as in the univariate case, we will first prove a lemma from which the theorem itself will easily follow. In order to find the appropriate generalization of Lemma 48, we first write it in an equivalent form which is somehow reminiscent of the *substitution rule* for integrals.

67 **Lemma (Substitution of Power Series)**
Let $f \in \mathbf{C}((z))$ and $g \in \mathbf{C}[[z]]$ with $\operatorname{ord} g = 1$. Then we have the relation
□ $\mathrm{M}\, f(g(z))\, g'(z) = \mathrm{M}\, f(z)$, which is equivalent to Lemma 48.

68 **Proof (Substitution of Power Series)**
Writing $\sum f_m z^m = f(z)$, we can utilize Lemma 48 to infer

$$\begin{aligned} \mathrm{M}\, f(g(z))\, g'(z) &= \mathrm{M} \sum f_m\, g(z)^m\, g'(z) = \sum f_m\, \mathrm{M}\, g(z)^m\, g'(z) \\ &= \sum f_m\, \delta_{m,-1} = f_{-1} = \mathrm{M}\, f(z). \end{aligned}$$

For proving the other direction of the equivalence, we choose an $n \in \mathbf{N}$ and define $f(z) = z^{-n-1}$. By the present lemma, we have $\mathrm{M}\, g(z)^{-n-1} g'(z) =$
□ $\mathrm{M}\, z^{-n-1} = \delta_{n0}$, which is precisely what Lemma 48 says.

The similarity to the substitution rule for integrals becomes clear by recalling that for holomorphic functions $f\colon \mathbf{C} \to \mathbf{C}$, the Residue Theorem of complex analysis tells us that

$$\mathrm{M}\, f(z) = \frac{1}{2\pi i} \oint_C f(\zeta)\, d\zeta,$$

where M means the residue of f in 0 and C is a simply closed curve around 0. Now we know that the results of complex analysis are just those special cases of formal relations where the series converge. This suggests that the desired multidimensional analog of Lemma 67 can be regarded as the multidimensional substitution rule for integrals. Therefore the only necessary change for generalizing Lemma 48 should be to replace the derivative $g'(z)$ by the Jacobian determinant. And indeed it is so.

69 Lemma (Elementary Residues, Multivariate Form)
Let $\boldsymbol{e}(\boldsymbol{z}) \in \mathbf{C}_s^s[[\boldsymbol{z}]]$ with $\operatorname{L} e_i(\boldsymbol{z}) = 1$ for all $i \in \{1, \ldots, s\}$, and let $\mathbf{g}(\boldsymbol{z}) = \boldsymbol{z}/\boldsymbol{e}(\boldsymbol{z})$. Then the elementary residues are found to be

$$\operatorname{M} |\mathbf{g}'(\boldsymbol{z})| \, \mathbf{g}(\boldsymbol{z})^{-\mathbf{n}-1} = \delta_{\mathbf{n}0}$$

□ for all $\mathbf{n} \in \mathbf{Z}^s$.

70 Proof (Elementary Residues, Multivariate Form)
First we shall prove that $\operatorname{M} |\mathbf{h}'(\boldsymbol{z})| = 0$ for arbitrary vector Laurent series $\mathbf{h}(\boldsymbol{z}) \in \mathbf{C}_s^s[[\boldsymbol{z}]]$. Using the principle of potential infinity (stated in Proof 66), we only need to consider the basis elements $\mathbf{h}(\boldsymbol{z}) = (\boldsymbol{z}^{\mathbf{k}_1}, \ldots, \boldsymbol{z}^{\mathbf{k}_s})$ with arbitrary multi-indices $\mathbf{k}_1, \ldots, \mathbf{k}_s \in \mathbf{Z}^s$. Writing $\mathbf{k}_i = (k_{i1}, \ldots, k_{is})$ for $i \in \{1, \ldots, s\}$, we have $\partial h_i/\partial z_j = k_{ij}\, z^{\mathbf{k}_i}/z_j$ with $i, j \in \{1, \ldots, s\}$ and hence

$$|\mathbf{h}'(\boldsymbol{z})| = z^{\mathbf{k}_1 + \cdots + \mathbf{k}_s} \left| \frac{k_{ij}}{z_j} \right| = z^{\mathbf{k}_1 + \cdots + \mathbf{k}_s - \mathbf{1}} \, |k_{ij}|,$$

taking advantage of the linearity of the determinant—first in all the rows, then in all the columns. This expression has non-zero residue only if $\mathbf{k}_1 + \cdots + \mathbf{k}_s = \mathbf{0}$. But this means that all column sums in the matrix (k_{ij}) are 0 and therefore the determinant vanishes.

Next we shall evaluate the residue of $|\mathbf{h}'(\boldsymbol{z})| \, \mathbf{h}(\boldsymbol{z})^{-1}$, which is the non-zero case of the present lemma. The determinant $|\mathbf{h}'(\boldsymbol{z})|$ can be calculated directly from its definition in linear algebra: if (a_{ij}) is an $n \times n$ matrix, its determinant is

$$|(a_{ij})| = \sum_{\pi \in S_n} \epsilon_\pi \, a_{1,\pi(1)} \cdots a_{n,\pi(n)},$$

where S_n is the symmetric group of degree n and ϵ_π denotes the parity of the permutation π. The determinant in question has the entries

$$\frac{\partial h_i}{\partial z_j} = \frac{\partial}{\partial z_j} \frac{z_i}{e_i} = \frac{\delta_{ij}}{e_i} - \frac{z_i}{e_i^2} \frac{\partial e_i}{\partial z_j} = \frac{z_i}{e_i} \left(\frac{\delta_{ij}}{z_i} - \frac{1}{e_i} \frac{\partial e_i}{\partial z_j} \right)$$
$$= h_i \left(\frac{\delta_{ij}}{z_i} - \frac{\partial}{\partial z_j} \log\left(1 + (e_i - 1)\right) \right).$$

Note that it is possible to substitute $e_i(z) - 1$ into $\log(1+z)$, because $\operatorname{ord} e_i = 0$ and hence $\operatorname{ord}(e_i - 1) = 1$. We shall write $\log e_i$ as an abbreviation for $\log\big(1 + (e_i - 1)\big)$. Altogether we obtain

$$
\begin{aligned}
M\, |\mathbf{h}'(\boldsymbol{z})|\, \mathbf{h}(\boldsymbol{z})^{-1} &= M \left(\sum_{\pi \in S_s} \epsilon_\pi \frac{\partial h_1}{\partial z_{\pi(1)}} \cdots \frac{\partial h_s}{\partial z_{\pi(s)}} \right) \mathbf{h}(\boldsymbol{z})^{-1} \\
&= M \sum_{\pi \in S_s} \epsilon_\pi \frac{1}{h_1} \frac{\partial h_1}{\partial z_{\pi(1)}} \cdots \frac{1}{h_s} \frac{\partial h_s}{\partial z_{\pi(s)}} \\
&= M \sum_{\pi \in S_s} \epsilon_\pi \left(\frac{\delta_{1,\pi(1)}}{z_1} - \frac{\partial}{\partial z_{\pi(1)}} \log e_1 \right) \cdots \left(\frac{\delta_{s,\pi(s)}}{z_s} - \frac{\partial}{\partial z_{\pi(s)}} \log e_s \right)
\end{aligned}
$$

In this sum, all terms containing one or more $\partial/\partial z_{\pi(i)} \log e_i$ as factors do not contribute because all derivatives have zero residue. The only term that remains is

$$
M \sum_{\pi \in S_s} \epsilon_\pi \frac{\delta_{1,\pi(1)}}{z_1} \cdots \frac{\delta_{s,\pi(s)}}{z_s} = M \frac{1}{z_1 \ldots z_s} = 1.
$$

The sum has been reduced to a single term because all permutations except for the identical one have at least one δ-factor being 0. This concludes the proof for the case $\mathbf{n} = \mathbf{0}$.

In the other case, there are indices $n_i \neq 0$. But using the chain rule gives

$$
\frac{\partial h_i^{-n_i}}{\partial z_{\pi(i)}} = -\, n_i\, h_i^{-n_i - 1} \frac{\partial h_i}{\partial z_{\pi(i)}}
$$

and allows us to write

$$
h_i^{-n_i - 1} \frac{\partial h_i}{\partial z_{\pi(i)}} = \frac{\partial \tilde{h}_i}{\partial z_{\pi(i)}}
$$

if we define $\tilde{h}_i = -\, h_i^{-n_i}/n_i$. So in the big sum above, we have to replace the factor $(\delta_{i,\pi(i)}/z_i - \partial/\partial z_{\pi(i)} \log e_i)$ by such an $\tilde{h}_i$ for all $n_i \neq 0$. As all derivatives, these $\tilde{h}_i$ have no residue and therefore no term in the sum will contain the factor z_i^{-1}. Consequently, the overall residue is zero and the proof for $\mathbf{n} \neq \mathbf{0}$ is complete. □

Now the inversion theorem itself can be deduced easily, just as in the univariate case. The basic difference is only that we have a Jacobian determinant instead of a simple derivative.

71 **Proof (Multivariate Lagrange Inversion)**
By definition, we have

$$f(\boldsymbol{z}) = h(\mathbf{g}(\boldsymbol{z})) = \sum_{\mathbf{n} \geq 0} c_{\mathbf{n}}\, \mathbf{g}(\boldsymbol{z})^{\mathbf{n}}.$$

Multiplying this equation by $|\mathbf{g}'(\boldsymbol{z})|\, \mathbf{g}(\boldsymbol{z})^{-\mathbf{m}-1}$ and taking the residue gives

$$[\forall \mathbf{m} > 0] \;\; \mathrm{M}\, f(\boldsymbol{z})\, |\mathbf{g}'(\boldsymbol{z})|\, \mathbf{g}(\boldsymbol{z})^{-\mathbf{m}-1} = \sum_{\mathbf{n} \geq 0} c_{\mathbf{n}}\, \mathrm{M}\, |\mathbf{g}'(\boldsymbol{z})|\, \mathbf{g}(\boldsymbol{z})^{\mathbf{n}-\mathbf{m}-1} = c_{\mathbf{m}},$$

applying Lemma 69 in the last step. If we rewrite the above formula by the substitution $\mathbf{g}(\boldsymbol{z}) = \boldsymbol{z}/\boldsymbol{e}(\boldsymbol{z})$ and the absorption property on $\mathrm{M} = \left[\boldsymbol{z}^{-1}\right]$, we immediately obtain the formula given in the theorem.

The special case $c_0 = \left[\boldsymbol{z}^{0}\right] f(\boldsymbol{z})$ is obvious from the definition of $\mathbf{h}(\boldsymbol{z})$ because only $c_0\, \mathbf{g}(\boldsymbol{z})^{0}$ can contribute to the constant term since $\operatorname{ord} g_i(\boldsymbol{z}) = 1$ for all $i \in \{1, \ldots, s\}$. □

And [ye] are built upon the foundation of the apostles and prophets, Jesus Christ himself being the chief corner stone; in whom all the building fitly framed together groweth unto an holy temple in the Lord.
Ephesians 2:20–21

Chapter 3

Applications

The time is ripe now to place our results into the noble temple of mathematics. In the first two sections, we shall be concerned with various short applications of Lagrange's Theorem, coming both from the *combinatorial world* in Section 3.1 and from the *analytical world* in Section 3.2. (Of course, such a division between combinatorics and analysis must never be orthodox. There are few combinatorial results having absolutely no analytical meaning, and vice versa. But we usually can tell whether a particular problem has mainly a combinatorial or analytical flavor.) Section 3.3 will deal with the abundant material of *inverse relations*. As a last example from the numerous applications of the inversion formulas, we shall develop the beautiful theory of *binomial sequences* in Section 3.4.

3.1 Some Combinatorial Results

Let us start with the classical example of *Catalan numbers* C_n, described in the 1830s by Eugène Catalan and recapitulated by numerous authors like [6, p. 115–118], [8, p. 4–5], [19], [27, p. 43–44]. There are several different combinatorial questions leading to them, for example: How many triangulations exist in an $(n+1)$-gon? In how many ways can we parenthesize a product of n symbols? How many binary trees are there with n leaves? What is the number of lattice paths in $\mathbf{N}^2$, connecting $(0,0)$ and $(2n-2,0)$? We shall choose the second formulation for our example.

72 Example (Catalan Numbers)
So we are concerned with counting the number C_n of different products built from the n symbols $x_1,\ldots,x_n$, assuming the multiplication is com-

mutative but not associative. As an example, take $n = 4$. Then there are the $C_4 = 5$ products

$$((x_1x_2)(x_3x_4)),(((x_1x_2)x_3)x_4),((x_1(x_2x_3))x_4),$$
$$(x_1((x_2x_3)x_4)),(x_1(x_2(x_3x_4))).$$

Including the redundant pair of exterior parentheses as in the above example, every admissible product with $n > 1$ has the form (X_1X_2), with X_1 and X_2 being again such admissible products or single symbols. If X_1 consists of k symbols, then X_1 must consist of $n-k$ symbols, k ranging from 0 to n. For the single symbols, we have to arrange $C_1 = 1$ and for the empty product $C_0 = 0$. This leaves us with the recurrence

$$C_0 = 0, C_1 = 1,$$
$$[\forall n > 1]\ \ C_n = \sum_{k=0}^{n} C_k\, C_{n-k}.$$

We can now proceed with the usual recipe of transforming the recurrence into a functional equation for the underlying generating function $C(z) = \sum C_m z^m$. This is established very easily by

$$\begin{aligned} C(z) &= 0 + z + \sum_{m\geq 2} C_m z^m = z + \sum_{m\geq 2}\left(\sum_{k=0}^{m} C_k\, C_{m-k}\right) z^m \\ &= z + \sum_{m\geq 0}\left(\sum_{k=0}^{m} C_k\, C_{m-k}\right) z^m = z + C(z)^2, \end{aligned}$$

observing that the recurrence for C_n yields 0 if $n = 0$ or $n = 1$.

Since this is a simple functional equation, one could solve it by elementary methods and then expand $C(z)$ into a power series. But even though dispensable, Lagrange inversion is indeed much more elegant for computing the desired coefficients C_n. In order to bring the equation into the implicit form suitable for Corollary 45, we perform the substitution $C(z) = z\,(G(z)+1)$ so that the functional equation now reads $z\,G(z) + z = z + z^2\,(G(z)+1)^2$ or $G(z) = z\,e(G(z))$ after defining $e(z) = (z+1)^2$. Taking $f = \mathrm{id}$, the corollary immediately tells us that the coefficients of $G(z)$ are $G_0 = \left[z^0\right] f = 0$ and

$$[\forall m > 0]\ \ G_m = \frac{1}{m}\ \left[z^{m-1}\right](z+1)^{2m} = \frac{1}{m}\binom{2m}{m-1}.$$

Our substitution $C(z) = z\,(G(z) + 1)$ implies

$$\begin{aligned}\sum C_m z^m &= z + z \sum G_m z^m = z + \sum G_m z^{m+1} \\ &= z + \sum_{m>0} G_{m-1} z^m = 0 + z + \sum_{m>1} G_{m-1} z^m,\end{aligned}$$

so we regain the initial conditions $C_0 = 0$ and $C_1 = 1$ and also the neat solution

$$\begin{aligned}[\forall m > 1]\ C_m = G_{m-1} &= \frac{1}{m-1}\binom{2m-2}{m-2} = \frac{1}{m-1}\,\frac{(2m-2)!}{(m-2)!\,m!} \\ &= \frac{1}{m}\,\frac{(2m-2)!}{(m-1)!\,(m-1)!} = \frac{1}{m}\binom{2m-2}{m-1}.\end{aligned}$$

□ of our recurrence problem.

Lagrange inversion (in its general form with $f \neq \mathrm{id}$) also seems to be a short and elegant way to *Abel's identity*, which is a famous standard example of combinatorics. It is named after the mathematician Niels Hendrik Abel, who found it in the year 1826. Our exposition will follow [8, p. 3].

73 Example (Abel's Identity)
The kernel of this identity is the expansion $e^{xz} = \sum c_m\,(z/e^{az})^m$. Choosing $f(z) = e^{xz}$ and $e(z) = e^{az}$, the inversion formulas of Corollary 43 yield the coefficients

$$c_0 = \left[z^0\right] f(z) = \left[z^0\right] \sum \frac{x^m}{m!} z^m = 1 = \frac{x(x + a0)^{0-1}}{0!},$$

$$\begin{aligned}[\forall m > 0]\ c_m &= \frac{1}{m}\,\left[z^{m-1}\right] f'(z)\,e(z)^m = \frac{1}{m}\,\left[z^{m-1}\right] x\,e^{xz}\,e^{amz} \\ &= \frac{x}{m}\,\left[z^{m-1}\right] e^{(x+am)z} = \frac{x}{m}\,\left[z^{m-1}\right] \sum \frac{(x+am)^k}{k!} z^k \\ &= \frac{x(x+am)^{m-1}}{m!}\end{aligned}$$

for the coefficients of that development so that

$$e^{xz} = \sum \frac{x(x+am)^{m-1}}{m!}\,e^{-amz}\,z^m.$$

Multiplying this equation by e^{yz} changes it into

$$\begin{aligned}e^{(x+y)z} &= \sum \frac{(x+y)^m}{m!} z^m = \sum \frac{x(x+am)^{m-1}}{m!}\,e^{(y-am)z}\,z^m \\ &= \sum_m \left(\sum_{n=0}^{m} \frac{x(x+an)^{n-1}}{n!}\,\frac{(y-an)^{m-n}}{(m-n)!} \right) z^m.\end{aligned}$$

Comparing the coefficients of z^m finally gives rise to

$$[\forall m \in \mathbf{N}] \; (x+y)^m = \sum_{n=0}^{m} \binom{m}{n} x(x+an)^{n-1}(y-an)^{m-n},$$

□ which is known as Abel's identity.

Next let us consider a simple *inversion problem*, a classical example taken from [8, p. 2]. It is mainly of combinatorial interest, although the series does converge for certain parameter values.

74 Example (A Simple Inverse)
Given the series $g(z) = z\,e^{-az}$, we want to find its inverse $g^*(z) = \sum c_m z^m$. We can immediately employ Corollory 43 with $e(z) = e^{az}$ and $f(z) = z$ to find the coefficients $c_0 = [z^0]\; z = 0$ and

$$\begin{aligned}[\forall m > 0] \; c_m &= \frac{1}{m}\,[z^{m-1}]\,e^{amz} = \frac{1}{m}\,[z^{m-1}] \sum_n \frac{(am)^n}{n!} z^n \\ &= (am)^{m-1}/m!\,.\end{aligned}$$

By the way: knowing the inverse series, we can in principle solve the transcendental equation $e^{az} = z$, finding the solution

$$g^*(1) = \sum_{m>0} \frac{(am)^{m-1}}{m!},$$

provided the series converges. It does so, for example, if we take $a = -1/100$. This will yield the correct solution 0.9901473843595011 after only
□ 10 terms!

In Section 2.5, we have applied a generalization of *Cayley's tree formula* for proving Lagrange's theorem combinatorially. As announced there, we shall now go the opposite way of finding Cayley's formula by Lagrange inversion. This can be found in most textbooks of classical combinatorics, like [3] or [27].

75 Example (Cayley's Tree Formula)
We will employ the terminology of species theory just as in Section 2.5. So $\mathfrak{T}$, $\mathfrak{Z}$, and $\mathfrak{E}$ will denote the species of rooted trees, singletons, and sets, respectively. Now it is a trivial fact that $\mathfrak{T} \simeq \mathfrak{Z}\,(\mathfrak{E} \circ \mathfrak{T})$, since any rooted tree can be decomposed canonically into the root (which forms a singleton) and its descendant trees (whose roots are the successors of the original

root). Actually, this isomorphism is a special case of Lemma 62, point 1, if we choose $\mathfrak{r} = \mathfrak{E}$. (It means that the structures of $\mathfrak{T}$ have a uniform enrichment, thus being "normal" rooted trees.)

Of course, this combinatorial relation between the species $\mathfrak{T}$, $\mathfrak{Z}$ and $\mathfrak{E}$ immediately gives rise to the functional equation

$$\mathfrak{T}(z) = z\, e^{\mathfrak{T}(z)}$$

connecting their generating functions. It is amazing that it is already in the form suitable for Corollary 45. Let us set $e(z) = e^z$ and $f(z) = z$ to find the coefficients $\mathfrak{T}_0 = \left[z^0\right] z = 0$ and

$$[\forall m > 0] \quad \frac{\mathfrak{T}_m}{m!} = \frac{1}{m}\left[z^{m-1}\right] e^{mz} = \frac{1}{m}\left[z^{m-1}\right] \sum_k \frac{m^k}{k!} z^k = \frac{1}{m} \frac{m^{m-1}}{(m-1)!}.$$

Thus the desired number of rooted trees on $m > 0$ labels is $\mathfrak{T}_m = m^{m-1}$, which certainly agrees with the result of Proposition 64 (specialized to the classical case after the proof). □

Our last application in this section will utilize mulitvariate Lagrange inversion for proving *MacMahon's famous Master Theorem.* It can be used for deriving a great deal of complicated identities, especially such ones dealing with summation. It also should be noted that most of the other proofs are much more intricate if they do not apply Lagrange's theorem. We follow Hofbauer's paper [8, p. 9–10].

76 Theorem (MacMahon)

Let (a_{ij}) be an arbitrary $s \times s$ matrix with complex entries. Then we have the identity

$$[\forall \mathbf{m} \in \mathbf{N}^s] \quad [\mathbf{z}^{\mathbf{m}}]\, \frac{1}{D(\mathbf{z})} = [\mathbf{z}^{\mathbf{m}}] \prod_{i=1}^{s} \left(\sum_{j=1}^{s} a_{ij}\, z_j \right)^{m_i},$$

where $D(\mathbf{z})$ is defined as $\det(\delta_{ij} - a_{ij}\, z_j)$. □

77 Proof (MacMahon)

We define $e_i(\mathbf{z}) = 1 + \sum_{j=1}^{s} a_{ij}\, z_j$ for all $i \in \{1, \ldots, s\}$ and $\mathbf{g}(\mathbf{z}) = \mathbf{z}/\mathbf{e}(\mathbf{z})$ as appropriate for the multivariate inversion formulas of Theorem 47. Then the derivatives are $\partial e_i/\partial z_j = a_{ij}$ and by the quotient rule (see Proposition 26, point 4)

$$\frac{\partial g_i}{\partial z_j} = \frac{\partial}{\partial z_j} \frac{z_i}{e_i} = \frac{\delta_{ij}\, e_i - z_i\, a_{ij}}{e_i^2} = \frac{\delta_{ij} - a_{ij}\, g_i}{e_i}.$$

By using multilinear and symmetry properties, we can rewrite the determinant of the numerator as

$$\begin{aligned}\det(\delta_{ij} - a_{ij}\, g_i) &= (g_1 \cdots g_s)\, \det(\delta_{ij}/g_i - a_{ij}) \\ &= (g_1 \cdots g_s)\, \det(\delta_{ij}/g_j - a_{ij}) = \det(\delta_{ij} - a_{ij}\, g_j) = D(\boldsymbol{g}(\boldsymbol{z})).\end{aligned}$$

Thus the Jacobian determinant of $\boldsymbol{g}(\boldsymbol{z})$ is

$$|\boldsymbol{g}'(\boldsymbol{z})| = (e_1^{-1} \cdots e_s^{-1})\, \det(\delta_{ij} - a_{ij}\, g_i) = \boldsymbol{e}(\boldsymbol{z})^{-1}\, D(\boldsymbol{g}(\boldsymbol{z})).$$

If we define $f(\boldsymbol{z}) = D(\boldsymbol{g}(\boldsymbol{z}))^{-1}$, applying the inversion formulas now yields

$$\begin{aligned}c_{\mathbf{m}} &= [\boldsymbol{z}^{\mathbf{m}}]\ D(\boldsymbol{g}(\boldsymbol{z}))^{-1}\, \boldsymbol{e}(\boldsymbol{z})^{-1}\, D(\boldsymbol{g}(\boldsymbol{z}))\, \boldsymbol{e}(\boldsymbol{z})^{\mathbf{m}+1} = [\boldsymbol{z}^{\mathbf{m}}]\ \boldsymbol{e}(\boldsymbol{z})^{\mathbf{m}} \\ &= [\boldsymbol{z}^{\mathbf{m}}] \prod_{i=1}^{s} (1 + \sum_{j=1}^{s} a_{ij}\, z_j)^{m_i} = [\boldsymbol{z}^{\mathbf{m}}] \prod_{i=1}^{s} (\sum_{j=1}^{s} a_{ij}\, z_j)^{m_i}\end{aligned}$$

for every expansion coefficient of $\boldsymbol{g}(\boldsymbol{z})^{\mathbf{m}}$ in the series $1/D(\boldsymbol{g}(\boldsymbol{z}))$, which is
□ of course the same as the coefficient of $\boldsymbol{z}^{\mathbf{m}}$ in $1/D(\boldsymbol{z})$.

3.2 Solving Analytical Problems

As a starter, we shall *invert the exponential function*, probably the most famous function of both real and complex analysis. In our framework, finding the inverse means to determine the series coefficients of the natural logarithm (denoted by "log" here). As a formal power series, we cannot invert e^z itself, though, since it has order 0. According to Definition 20, composition is only defined for series with order greater than 1. Therefore we will rather compute the inverse of $e^z - 1$, namely $\log(z+1)$.

78 Example (Exponential Series Inversion)
Applying Theorem 42 to $g(z) = e^z - 1$ and $f(z) = z$ yields the coefficients $c_0 = L\, z = 0$ and

$$[\forall m > 0]\ c_m = \frac{1}{m}\, M\, \frac{1}{(e^z - 1)^m}.$$

Now fix any $m > 0$. It is surprising that the c_m are difficult to compute in a straight-forward approach. Experimenting with a good symbolic-algebra package, though, one soon suspects that $M\,(e^z - 1)^{-m} = (-1)^{m+1}$ or, equivalently, $M\,(1 - e^z)^m = -1$.

We will use induction for proving this conjecture, following an idea of Peter Paule. The base case $m = 1$ is apparent from

$$\frac{1}{1-e^z} = (-z - \frac{z^2}{2!} - \frac{z^3}{3!} - \dots)^{-1} = -z^{-1}\,(1 + \frac{z}{2!} + \frac{z^2}{3!} + \dots)^{-1},$$

where only the first factor $-z^{-1}$ contributes to the residue. The second is the reciprocal of an order-0-series and has the constant term $1/1 = 1$. (This can be seen from the recurrence scheme for reciprocals, given in Proof 12.)

The induction step uses the observation

$$\begin{aligned} -\frac{1}{m}\,D\,(1-e^z)^{-m} &= (1-e^z)^{-m-1}\,(-e^z) \\ &= (1-e^z)^{-m-1}\,((1-e^z)-1) = (1-e^z)^{-m} - (1-e^z)^{-m-1}, \end{aligned}$$

which implies that

$$\begin{aligned} M\,(1-e^z)^{-(m+1)} &= M\,(1-e^z)^{-m} + \frac{1}{m}\,MD\,(1-e^z)^{-m} \\ &= M\,(1-e^z)^{-m}, \end{aligned}$$

so the conjecture is proved and the coefficients of $g^*(z)$ are given by $c_0 = 0$ and $c_m = (-1)^{m+1}/m$ for all $m > 0$. Borrowing the name from analysis, let us write $\log(1+z)$ for the inverse series, so that we have the expansion

$$\log(1+z) = \sum_{m>0} \frac{(-1)^{m+1}}{m}\,z^m,$$

which is of course the well-known Taylor expansion of the natural logarithm.

Before we turn to the next example, though, we like to add a brief remark on this proof. At the first glance, it seems a bit unsatisfactory that such a mighty instrument as Lagrange inversion does not yield the desired inverse in a straight-forward way. (An induction proof always presupposes the right solution—gained by ingenuity and empirical data—in order to verify it.) We can learn two things from that. First: A powerful method, capable of tackling the toughest problems, is not necessarily all the more effective on simpler problems. Second: If a problem appears difficult on one approach but we can solve it on another, then we usually gain some new insight as a by-product. In our case, we come across an unexpected relation to the so-called Bernoulli series

$$B(z) = \frac{z}{e^z - 1} = \sum \frac{B_m}{m!}\,z^m.$$

The Bernoulli numbers B_m can be computed by a recurrence, which is deduced from the above quotient representation (their actual computation is irrelevant for us). They turn up in numerous corners of analysis, especially in the power series of such functions as $\tan z, \cot z, \tanh z, \coth z$.

We have proved above that

$$M\,(e^z-1)^{-m} = \left[z^{m-1}\right] B(z)^m = (-1)^{m+1}.$$

Expanding the convolution $B(z)^m$ leads to

$$[\forall m > 0] \sum_{(n_1,\ldots,n_m)\models m-1} \frac{B_{n_1}}{n_1!}\cdots\frac{B_{n_m}}{n_m!} = (-1)^{m+1},$$

□ which is an interesting identity for Bernoulli numbers.

Our first "substantial" example will be a numerical application, dealing with *cubic equations*, which emerge in all branches of physics. Their general form is $x^3+ax^2+bx+c=0$ with complex coefficients a, b, c. As can be looked up in any formula handbook, this equation can always be reduced to its normal form $z^3+3pz+2q=0$ by the substitution $z = x + a/3$. The new constants are $p = b/3 + a^2/9$ and $q = a^3/27 - ab/6 + c/2$. The Italian mathematician Cardano found the explicit formula

$$z_0 = \sqrt[3]{-q+\sqrt{p^3+q^2}} + \sqrt[3]{-q-\sqrt{p^3+q^2}}$$

for one of the solutions as early as in 1545. In most of the practical applications, the coefficients a, b, c (therefore also p, q) are real so that at least one of the solutions is real. Physicists usually are only interested in non-complex solutions and it will suffice to have one of them. (The others can be found by splitting off a linear factor.) It turns out that Cardano's solution z_0 meets these conditions.

For fast computations (for instance in iteration schemes), Cardano's formula is too complicated, because it requires the evaluation of one square root and two cubic roots. So one has to resort to *numerical approximations*. Due to the nested roots, though, a Taylor expansion will involve a rather messy calculation, whereas Lagrange inversion quickly leads to an elegant solution.

79 Example (Cardano's Formula)
Solving the equation $z^3+3pz+2q=0$ is equivalent to solving $g(z) = -2q$, if we set $g(z) = z^3 + 3pz$. The desired solution is $z_0 = g^*(-2q)$, so we

must find the inverse series $g^*(z) = \sum c_m z^m$. Choosing $f(z) = z$, the inversion formulas of Theorem 42 tell us that the coefficients c_m are given by $c_0 = \mathrm{L}\, z = 0$ and

$$[\forall m > 0]\ c_m = \frac{1}{m}\,\mathrm{M}\,(z^3 + 3pz)^{-m} = \frac{1}{m}\,\left[z^{m-1}\right]\,\frac{1}{(z^2 + 3p)^m}.$$

The coefficient $c_0 = 0$ is trivial since all inverse series have order 1. Therefore let us fix an arbitrary $m > 0$. Using Example 19 with $(z, a, c) \to (z^2, 3p, m)$, we know that

$$\frac{1}{(z^2 + 3p)^m} = \sum_n \binom{m+n-1}{n} (-1)^n \frac{z^{2n}}{(3p)^{n+m}}.$$

So for even indices m, the coefficient c_m vanishes, whereas for odd indices $m = 2k+1$ (with some $k \in \mathbf{N}$), it is given by

$$c_{2k+1} = \binom{3k}{k} \frac{(-1)^k}{2k+1} \frac{1}{(3p)^{3k+1}}.$$

The solution can now be written in the form

$$z_0 = g^*(-2q) = \sum_k \binom{3k}{k} \frac{(-1)^{k+1}}{2k+1} \frac{(2q)^{2k+1}}{(3p)^{3k+1}},$$

which is a neat series in the two parameters p and q. Although it is not difficult to determine the precise region of convergence (since the series is hypergeometric), this should not be our concern here. It does seem to be a very good approximation for solutions with $|z_0| < 1$. (We can always force this by the transformation $z \to 1/z$.) As sample values, let us take $(p, q) = (3, 2)$, which yields the solution $z_0 = -0.435281$ with 6 significant figures if we truncate the series after only 7 terms! □

Finally let us consider the solution of certain more general *algebraic equations*. It is a well-known result of Galois theory that there are no explicit solutions ("in radicals") for a general algebraic equation of order higher than 4. But for some special cases, there are solutions of a certain form. The simplest special case occurs when there is only a constant term besides the highest power: the equation $z^n + a_0 = 0$ has the trivial solution $\hat{z} = \sqrt[n]{-a_0}$. The next complex case is when we add a linear term: that gives the equation $z^n + a_1 z + a_0 = 0$. By Lagrange inversion, we can find its solution in the form of a series. This is accurate in the analytical sense and also useful from the numerical point of view.

80 **Example (An Algebraic Equation)**
We first consider the algebraic equation $(z+a)^n = b\,z$ containing the three parameters $n \in \mathbf{N}^*$ and $a, b \in \mathbf{C}$. Defining $g(z) = z/(z+a)^n$, this can be written as a functional equation $g(z) = 1/b$. Thus the solution next to the origin is $\hat{z} = g^*(1/b)$. The coefficients c_m of $g^*(z)$ are delivered by Corollary 43 as $c_0 = \left[z^0\right] z = 0$ and

$$[\forall m > 0]\ c_m = \frac{1}{m}\,\left[z^{m-1}\right](z+a)^{mn} = \frac{1}{m}\binom{mn}{m-1}a^{mn-m+1}.$$

Therefore the inverse series is

$$g^*(z) = \sum_{m>0}\binom{mn}{m-1}\frac{a}{m}\,a^{(n-1)m}z^m,$$

and the desired solution is

$$\hat{z} = g^*(1/b) = \sum_{m>0}\binom{mn}{m-1}\frac{a}{m}\left(\frac{a^{n-1}}{b}\right)^m.$$

It is not the place here to settle questions of convergence. Let us just mention that parameters $|a| < 1$ and $|b| > 1$ are more likely to succeed and that other values can usually be adjusted by suitable transformations. As an example, choose $(a, b, n) = (1/3, 10, 7)$, which corresponds to solving a 7th-order equation. Summing the first 10 terms of the series yields the solution 0.000045768703190211 with 14 significant figures.

Given the general type $z^n + a_1 z + a_0 = 0$, we only have to apply the substitution $a_0 = -ab, a_1 = -b, z = w - a$. This changes the given equation into $(w-a)^n - b\,(w-a) - ab = (w-a)^n - b\,w + ab - ab = 0$, which is equivalent to $(w-a)^n = b\,w$. Thus we have reduced the general type to the one treated above. □

Lagrange inversion can even be used to find the so-called principal solution (and hence all the solutions) of the *general algebraic equation* $z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 = 0$. Like in our example, though more complicated, it has the form of a power series in the coefficients $a_0, \ldots, a_{n-1}$, known as Mellin series. For a derivation using Lagrange's theorem, see Peter Paule's concise paper [18].

3.3 Inverse Relations

As we have mentioned in the introduction chapter, the *general form* of a pair of inverse relations (briefly, an inverse pair) is

$$[\forall m \in \mathbf{N}] \;\; a_m = \sum_{n=0}^{m} \alpha_{mn} b_n \;\Leftrightarrow\; b_m = \sum_{n=0}^{m} \beta_{mn} a_n.$$

The essential information is contained in the connecting coefficients α_{mn} and β_{mn}, which can be regarded as two infinite lower-triangular matrices $A = (\alpha_{mn})$ and $B = (\beta_{mn})$. So the first part of the pair tells us that A maps the vector $\mathbf{b} = (b_0, \ldots, b_N) \in \mathbf{C}^{N+1}$ onto the vector $\mathbf{a} = (a_0, \ldots, a_N) \in \mathbf{C}^{N+1}$, whereas the second tells us that B maps $\mathbf{a}$ back onto $\mathbf{b}$. Here $N \in \mathbf{N}^*$ is an arbitrary parameter, which tells us where to "cut off" the dimension.

In other words, the matrices A and B are inverse to each other. So there is another way of stating a pair of inverse relations, which is *independent of the random choice for* $\mathbf{a}$ *and* $\mathbf{b}$, namely

$$\sum_k \alpha_{mk} \beta_{kn} = \delta_{mn},$$

which just says that $AB = I$. There is an analogous sum, which corresponds to $BA = I$. The sums are always finite because A and B are lower-triangular matrices.

The interesting thing about these relations is that they turn up virtually everywhere in the realm of combinatorics, hence *extensive study* has been devoted to them. And yet they seem to be very resistant against all kinds of systematizations and solution methods. Up to now, we do not know how to symbolically invert a given relation! This would mean that we have one half of an inverse pair and we seek the other half: given the matrix A, we look for its inverse B in closed form. Of course we cannot really expect to invert any matrix in closed form. The first step would rather be to define a subclass of nontrivial expressions which allow closed-form inversion. But we are still quite far from such a result.

On the other hand, it must be noted that progress has been made. One of the most famous attempts to classify inverse relations has been undertaken by John *Riordan* in 1968. In his book [22] he investigates approximately 60 such relations and still has to confess in the end that "the age-old dream of putting order in this chaos is doomed to failure." His book actually appears like an amorphous heap of identities, although their derivation is certainly very sophisticated. The intermediate calculations, however, involve only pattern matching for the most part: this cries for algorithmization. The following theorem, taken from [8, p. 11], is a first step in that direction. Although it does not allow to *solve* a given relation in the sense of a closed form, it does enable us to *construct* a vast class of inverse pairs. In particular, it covers all of Riordan's examples, indeed putting order in the chaos of his book.

81 **Theorem (Lagrange Construction of Inverse Pairs)**
Let $f(z)$ be a power series of order 0, and $g(z), h(z)$ power series of order 1. Then

$$\begin{aligned} \alpha_{mn} &= \mathrm{M}\, f(z) \quad g(z)^n\, h(z)^{-m-1}\, h'(z), \\ \beta_{mn} &= \mathrm{M}\, f(z)^{-1}\, h(z)^n\, g(z)^{-m-1}\, g'(z) \end{aligned}$$

□ form an inverse pair.

82 **Proof (Lagrange Construction of Inverse Pairs)**
The α_{mn} and β_{mn} appear as the expansion coefficients in

$$\begin{aligned} f(z)\, g(z)^n &= \sum_m \alpha_{mn}\, h(z)^m, \\ f(z)^{-1}\, h(z)^n &= \sum_m \beta_{mn}\, g(z)^m, \end{aligned}$$

as we can immediately verify by applying the second form of the Lagrange inversion formulas in Theorem 42. We just have to make the appropriate substitutions $f(z) \to f(z)\, g(z)^n, g(z) \to h(z), c_m \to \alpha_{mn}$ for obtaining the first formula and $f(z) \to f(z)^{-1}\, h(z)^n, c_m \to \beta_{mn}$ for the second. (Note that the conditions $\operatorname{ord} h(z) = 1$ and $\operatorname{ord} g(z) = 1$ are satisfied as needed for the first and second inversion, respectively.)

The case $m = 0$ needs a little further inspection. On the right-hand side of the above expansions, these terms are the only contributions of order 0, since both $h(z)$ and $g(z)$ have order 1. But the left-hand side always has order 1 except for $n = 0$. Thus all the α_{0n} and β_{0n} with $n > 0$ vanish, which must be so because α_{mn} and β_{mn} are to be lower-triangular matrices. Setting $m = n = 0$, we come to $\alpha_{00} = \mathrm{L}\, f(z)$ and $\beta_{00} = \mathrm{L}\, f(z)^{-1}$, which is just what the Lagrange inversion formulas produce in this case.

For proving the inversion property, we multiply the first expansion by β_{nk} and sum on n. This gives

$$f(z) \sum_n \beta_{nk}\, g(z)^n = f(z)\, f(z)^{-1} h(z)^k = h(z)^k$$

for the left-hand side, where we have substituted the second development for the sum. The right-hand side gives

$$\sum_m \left(\sum_n \alpha_{mn} \, \beta_{nk} \right) h(z)^m,$$

which implies that the expression in parentheses must be δ_{mk}, since both sides are equal.

In the same way we can derive the other inversion identity if we multiply the second development by α_{nk} and sum on n, which yields

$$f(z)^{-1} \sum_n \alpha_{nk} \, h(z)^n = f(z)^{-1} \, f(z) \, g(z)^k = g(z)^k$$

for the left-hand side and

$$\sum_m \left(\sum_n \beta_{mn} \, \alpha_{nk} \right) g(z)^m$$

for the right-hand side. Equating the two sides again means that the sum in parentheses is δ_{mk}. □

As we have already mentioned, there is a host of *examples*. We pick a classical one, which can also be found in [8, p. 11].

83 Example (Binomial Inversion)
We would like to find the inverse of $\alpha_{mn} = \binom{m+p}{n+p}$, where p is some additional parameter(or an indeterminate). Because of

$$\alpha_{mn} = \binom{m+p}{n+p} = \binom{m+p}{m-n} = M \, (1+z)^{m+p} \, z^{-m+n-1},$$

we are led to choosing $h(z) = (1+z)^{-1} z$ and $g(z) = z$. Observing that $h'(z) = (1+z)^{-2}$ makes us set $f(z) = (1+z)^{p+1}$. Now we can immediately calculate the inverse

$$\begin{aligned} \beta_{mn} &= M \, (1+z)^{-p-1} \, (1+z)^{-n} \, z^n \, z^{-m-1} \, 1 \\ &= M \, (1+z)^{-n-p-1} \, z^{n-m-1} = \binom{-n-p-1}{m-n} \end{aligned}$$

in a raw form.

But negative indices in binomials are usually avoided, so we rearrange

$$\begin{aligned}\beta_{mn} &= \frac{(-n-p-1)(-n-p-2)\cdots(-p-m)}{(m-n)!} \\ &= (-1)^{m-n}\frac{(m+p)(m+p+1)\cdots(n+p+1)}{(m-n)!} \\ &= (-1)^{m-n}\frac{(m+p)^{\underline{m-n}}}{(m-n)!} = (-1)^{m-n}\binom{m+p}{m-n}.\end{aligned}$$

The special case $p = 0$ yields

$$\sum_k (-1)^{k-n}\binom{m}{k}\binom{k}{n} = \delta_{mn},$$

□ which is probably one of the most famous inversions at all.

Our second example deals with a whole *class of inverse pairs*, all generated by a suitable power series. It is taken from Riordan's book [22, p. 99], demonstrating his effort of systematization of identities. Besides this type, he also considers a number of other inverse pairs that can be constructed by power series. But they all cover only a very small portion of the known identities: time was not mature by then. The following example reveals that his "generalization" is only a simple special case of Lagrange construction.

84 Example (An Inverse-Pair Class of Riordan)
In Section 3.3 of his book [22], Riordan states that the numbers c_{m-n} and $\bar{c}_{m-n}$ form an inverse pair whenever the generating functions $c(z) = \sum c_m z^m$ and $\bar{c}(z) = \sum \bar{c}_m z^m$ fulfill $c(z)\,\bar{c}(z) = 1$. So in order to construct an inverse pair, one chooses an arbitrary power series $c(z)$, then computes its reciprocal $\bar{c}(z)$, and their coefficients will make up an inverse pair as stated.

This class of inverse relations can be generated easily be Theorem 81 if we choose $f(z) = c(z)$ and $g(z) = h(z) = z$. For these settings, the formulas yield

$$\begin{aligned}\alpha_{mn} &= \mathrm{M}\, c(z)\, z^{n-m-1} = c_{m-n}, \\ \beta_{mn} &= \mathrm{M}\, \bar{c}(z)\, z^{n-m-1} = \bar{c}_{m-n},\end{aligned}$$

□ which is what we have claimed.

As mentioned above, Riordan considers other cases that are far more complicated, but they all fall short of the vast generality contained in Theorem 81—the previous example has demonstrated this in a very lucid way. No considerable progress was made until Georgy *Egorichev* discovered in 1984 that most of the known identities (especially all the ones in Riordan's book) can be constructed by the formulas stated above. In his book [4], he classifies the identities according to this principle by using integration methods. The main idea is that one can express the residue functional by a complex contour integral as in Section 2.3.

This is more or less the state of the art, except for some *glimmers of hope*. Let us just mention one of them, found in the article [15] of Stephen Milne and Gaurav Bhatnagar. They have observed and proved that α_{mn} and β_{mn} form an inverse pair iff there are sequences $a_k, b_k \colon \mathbf{Z} \to \mathbf{C}$ and a number $K \in \mathbf{N}$ such that

$$\sum_{k \geq -K} \Big(a_{-k}(m-k)\, \alpha_{m-k,n} + b_k(n+k)\, \alpha_{m,n+k} \Big) = 0,$$
$$\sum_{k \geq -K} \Big(a_{-k}(n)\, \beta_{m,n+k} + b_k(m)\, \beta_{m-k,n} \Big) = 0$$

holds together with enough initial conditions for determining the α_{mn} and β_{mn} uniquely.

It is evident that much further research is needed on this topic, since we do not have a unified approach for solving inverse relations. There is a closely *related problem*, though, that has been dealt with in an exhausting way: the calculation of the connecting coefficients for given sequences $\mathbf{a}$ and $\mathbf{b}$, treated in the next section. For this problem, we have a general solution, provided that the sequences $\mathbf{a}$ and $\mathbf{b}$ meet certain conditions (which is often the case in practical applications). This suggests that some inversion problems can be solved if one finds suitable sequences $\mathbf{a}$ and $\mathbf{b}$ and then computes their connecting coefficients by these methods.

3.4 Binomial Sequences

This section deals with the famous problem of finding the *connecting coefficients* between two polynomial sequences. These terms will be defined very soon, so let us just briefly state the problem. Given the polynomial sequences $\langle p_m(z)\rangle$ and $\langle q_m(z)\rangle$, we are to determine the numbers c_{mn} in

$$[\forall m \in \mathbf{N}]\ \ q_m(z) = \sum_{n=0}^{m} c_{mn}\, p_n(z).$$

In other words, we ask for the transformation matrix, changing from one basis $\langle p_m(z)\rangle$ to another basis $\langle q_m(z)\rangle$.

For general sequences, the situation is very complicated and beyond the scope. For the subclass of so-called *binomial sequences*, though, there are precise results. These are usually stated in terms of *operator calculus*, which will therefore be introduced as far as needed.

Our presentation will roughly follow [1, p. 170–190]; for a generalization of this approach, see Roberto Pirastu's PhD thesis [20]. We will go our own way, however, when it comes to prove the key result in Proposition 96. Aigner's proof uses what Peter Paule calls the "principle of omniscience" in so far as it merely verifies the basis formula°. Our proof will utilize *Lagrange inversion* for deriving these formulas in a much more natural way. So to speak, it goes opposite to Cigler's proof of Lagrange's theorem in [3, p. 75–81]. While he uses the basis formula (with a proof similar to Aigner's) for deducing Lagrange's theorem, we shall use Lagrange's theorem for deducing the basis formula. But first let us begin with the necessary definitions.

85 **Definition (Polynomial Sequence, Binomial Sequence)**
If $p_m(z) \in \mathbf{C}[z]$ and $\deg p_m(z) = m$ for all $m \in \mathbf{N}$, then $\langle p_m(z)\rangle$ is called a *polynomial sequence.*

A polynomial sequence fulfilling the generalized binomial theorem

$$p_0(z) = 1,$$
$$[\forall m > 0]\ \ p_m(z_1 + z_2) = \sum_{n=0}^{m} \binom{m}{n} p_n(z_1)\, p_{m-n}(z_2)$$

□ is called a *binomial sequence.*

Note that within the strict theory of formal power series, the term $p_m(z_1 + z_2)$ cannot be interpreted as the value of p_m at the point $z_1 + z_2$. It must be regarded as the composition of the univariate polynomial $p_m(z)$ and the bivariate polynomial $z_1 + z_2$. Analogously, polynomials like the $p(z + a)$ appearing below are to be understood as the composition of the polynomial $p(z)$ and the polynomial $z + a$, which has constant term a.

As we have mentioned above, the transformation problem of this section can be attacked successfully by *operator methods*. An operator is a linear transformation of $\mathbf{C}[z]$ onto itself. We will use sans-serif letters like P and Q to denote them. Observe that L and M are not operators in this sense because they map into $\mathbf{C}$; this is why they are called functionals. The composition of two operators P and Q is written as $\mathsf{PQ}\,p(z) = \mathsf{P}\,(\mathsf{Q}\,p(z))$ and the iteration as $\mathsf{P}^m = \mathsf{P}\cdots\mathsf{P}$. Precedence is ruled by setting $\mathsf{P}\,p(q(z)) = \mathsf{P}\,(p(q(z)))$.

Some well-known *examples* of operators are the identity $\mathsf{I}\,p(z) = p(z)$, the differentiation $\mathsf{D}\,p(z) = p'(z)$, and the a-shift $\mathsf{E}^a\,p(z) = p(z+a)$ where $a \in \mathbf{C}$. Other examples will emerge in the course of this section. The key idea for applying operator calculus to the transformation problem is to associate a unique operator with each polynomial sequence.

86 Definition (Basis Operator, Basis Sequence)
For any polynomial sequence $\langle p_m(z)\rangle$, the operator P defined by

$$\begin{aligned} \mathsf{P}\,p_0(z) &= 0, \\ [\forall m > 0]\ \ \mathsf{P}\,p_m(z) &= m\,p_{m-1}(z) \end{aligned}$$

is called its *basis operator*. Conversely, any sequence $\langle p_m(z)\rangle$ fulfilling this property for a given operator P is called its *basis sequence*. □

Existence and uniqueness of the operator is clear because any polynomial sequence forms a basis (hence the names of basis sequence and basis operator) of the vector space $\mathbf{C}[z]$, and it suffices to define an operator on the basis elements. The following proposition gives an *expansion formula*, which uses the basis operator of a binomial sequence to represent an arbitrary polynomial in terms of this sequence. (Actually it would be sufficient to have a so-called normalized sequence; see [1] for details.)

87 Proposition (Expansion in Basis Polynomials)
Let $\langle p_m(z)\rangle$ be a binomial sequence with basis operator P. Then we have the expansion

$$q(z) = \sum_m \frac{\mathrm{L}\mathsf{P}^m\,q(z)}{m!}\,p_m(z)$$

for arbitrary $q(z) \in \mathbf{C}[z]$. □

88 **Proof (Expansion in Basis Polynomials)**
We have to prove an identity over all polynomials $q(z)$ in the vector space $\mathbf{C}[z]$. It is sufficient to prove it for the polynomials $p_m(z)$ because they constitute a basis of $\mathbf{C}[z]$.

Let us first show that any binomial sequence is normalized, which means that $L\,p_m(z) = \delta_{m0}$ for all $m \in \mathbf{N}$. The case $m = 0$ is obvious from the first part in the definition of binomial sequences. The second part yields

$$[\forall m > 0]\ L\,p_m(z) = \sum_{n=0}^{m} \binom{m}{n} L\,p_n(z)\ L\,p_{m-n}(z)$$

after applying the functional L with respect to both z_1 and z_2. This relation allows a simple induction proof. The basis $m = 1$ is clear because $L\,p_1(z) = L\,p_0(z)\ L\,p_1(z) + L\,p_1(z)\ L\,p_0(z) = 2\,L\,p_1(z)$ immediately gives $L\,p_1(z) = 0$. So let us assume $L\,p_n(z) = 0$ for all $n \in \{1, \ldots, m-1\}, m > 1$. Then the above formula tells us that $L\,p_m(z) = L\,p_0(z)\ L\,p_m(z) + L\,p_m(z)\ L\,p_0(z) = 2\,L\,p_m(z)$, so $L\,p_m(z) = 0$, and the induction step is completed.

Now choose an arbitrary $n \in \mathbf{N}$. Iterating the definition of the basis operator, we get $P^m\,p_n(z) = n^{\underline{m}}\,p_{n-m}(z)$ for all $m \leq n$ and $P^m\,p_n(z) = 0$ for all $m > n$. Since the sequence $\langle p_m(z)\rangle$ is binomial, it is also normalized, as we have just shown. Taking the constant term of the iteration thus gives $LP^n\,p_n(z) = n!$ and $LP^m\,p_n(z) = 0$ for all $m \neq n$. Therefore we can write

$$p_n(z) = \sum_m \frac{LP^m\,p_n(z)}{m!}\,p_m(z),$$

□ which concludes the proof.

Although a basis operator can be defined for any polynomial sequence, our focus will be on binomial sequences. Their basis operators constitute a very important operator class, the so-called *delta operators*.

89 **Definition (Shift-Invariant Operators, Delta Operators)**
An operator P is called *shift-invariant*, if it satisfies $PE^a = E^aP$ for every $a \in \mathbf{C}$. If P has the additional property $P\,z \in \mathbf{C} \setminus \{0\}$, it is called a *delta*
□ *operator*.

The classical example of a delta operator is the *differentiation* D. It is shift-invariant because $D\,p(z+a) = (Dp)\,(z+a)$ and a delta operator because $D\,z = 1$. Other examples include its discrete analoga, the *forward* and *backward difference operators*, defined by $\Delta = \mathsf{E} - \mathsf{I}$ and $\nabla = \mathsf{I} - \mathsf{E}^{-1}$, respectively. They are shift-invariant because of

$$\begin{aligned}\Delta\, p(z+a) &= p(z+a+1) - p(z+a) = (\Delta\, p)\,(z+a),\\ \nabla\, p(z+a) &= p(z+a) - p(z+a-1) = (\nabla\, p)\,(z+a)\end{aligned}$$

and delta operators because of $\Delta\, z = (z+1) - z = 1, \nabla\, z = z - (z-1) = 1$. The next proposition asserts the claimed correspondence between binomial sequences and delta operators.

90 Proposition (Binomial Sequences and Delta Operators)
The polynomial sequence $\langle p_m(z)\rangle$ is binomial iff its basis operator P is a
□ delta operator.

91 Proof (Binomial Sequences and Delta Operators)
First let us assume that $\langle p_m(z)\rangle$ is binomial. Then we can infer $\mathrm{L}\, p_1(z) = 0$, as shown in the previous proof. Since we also know that $\deg p_1(z) = 1$, the polynomial must have the form $p_1(z) = \alpha z$ with some $\alpha \in \mathbf{C} \setminus \{0\}$. The definition of the basis operator P now yields $1 = p_0(z) = \mathsf{P}\, p_1(z) = \alpha\,(\mathsf{P}\, z)$, which implies $\mathsf{P}\, z = 1/\alpha \in \mathbf{C} \setminus \{0\}$. Thus it remains to show that the basis operator P is shift-invariant.

Once again, it suffices to show $\mathsf{P}\mathsf{E}^a\, p_m(z) = \mathsf{E}^a\mathsf{P}\, p_m(z)$ for all $m \in \mathbf{N}$ because $\langle p_m(z)\rangle$ forms a basis of $\mathbf{C}[z]$. Choose an arbitrary $n > 0$ (the case $n = 0$ is trivial). As in the previous proof, iterating the basis operator yields $\mathsf{P}^m\, p_n(z) = n^{\underline{m}}\, p_{n-m}(z)$ for all $m \le n$. This can be substituted in the binomial expansion

$$\begin{aligned}\mathsf{E}^a\, p_n(z) = p_n(z+a) &= \sum_{m=0}^{n} \binom{n}{m}\, p_{n-m}(a)\, p_m(z)\\ &= \sum_{m=0}^{n} \frac{\mathrm{L}\mathsf{E}^a\mathsf{P}^m\, p_n(z)}{m!}\, p_m(z).\end{aligned}$$

Replacing $p_n(z)$ by $\mathsf{P}\, p_n(z) = n\, p_{n-1}(z)$ brings us to the desired result

$$\begin{aligned}\mathsf{E}^a\mathsf{P}\, p_n(z) &= \sum_{m=0}^{n-1} \frac{\mathrm{L}\mathsf{E}^a\mathsf{P}^m\, n\, p_{n-1}(z)}{m!}\, p_m(z)\\ &= \sum_{m=0}^{n-1} \frac{\mathrm{L}\mathsf{E}^a\mathsf{P}^{m+1}\, p_n(z)}{m!}\, p_m(z) = \sum_{m=0}^{n} \frac{\mathrm{L}\mathsf{E}^a\mathsf{P}^m\, p_n(z)}{m!}\, m\, p_{m-1}(z)\\ &= \mathsf{P}\left(\sum_{m=0}^{n} \frac{\mathrm{L}\mathsf{E}^a\mathsf{P}^m\, p_n(z)}{m!}\, p_m(z)\right) = \mathsf{P}\mathsf{E}^a\, p_n(z).\end{aligned}$$

Now let us prove the other direction. Assuming that P is a delta operator, we have to show that its basis sequence $\langle p_m(z)\rangle$ is binomial. For that let us use the expansion of Proposition 87 with $q(z) = p_n(z+a)$ to obtain

$$\begin{aligned} p_n(z+a) &= \sum_{m=0}^{n} \frac{LP^m\, p_n(z+a)}{m!}\, p_m(z) = \sum_{m=0}^{n} \frac{LP^m E^a\, p_n(z)}{m!}\, p_m(z) \\ &= \sum_{m=0}^{n} \frac{LE^a P^m\, p_n(z)}{m!}\, p_m(z) = \sum_{m=0}^{n} \frac{n^{\underline{m}}\, p_{n-m}(a)}{m!}\, p_m(z) \\ &= \sum_{m=0}^{n} \binom{n}{m}\, p_m(z)\, p_{n-m}(a) \end{aligned}$$

for every $n \in \mathbf{N}$. This is of course the generalized binomial theorem, if we regard a as a new indeterminate y. Choosing specifically $n = 0$, this becomes $p_0(z+a) = p_0(z)\,p_0(a)$. or simply $\alpha = \alpha^2$ if we write $p_0(z) = \alpha$. Therefore α must either be 0 or 1. It cannot be 0 because by the binomial expansion, this would imply $p_1(z+a) = p_0(z)\,p_0(a) + p_0(a)\,p_0(z) = 0$, in contradiction to $\deg p_1(z) = 1$. Thus we have $p_0(z) = \alpha = 1$, and the whole sequence $\langle p_m(z)\rangle$ is binomial. □

Now we come to the last fundamental result of operator calculus—the very reason for it being such a powerful tool. It turns out that every shift-invariant operator (especially every delta operator) has an *expansion in terms of* D.

Obviously, the sum of two shift-invariant operators is shift-invariant again. The same it true for the product of two shift-invariant operators P and Q, because we have $PQE^a = PE^aQ = E^aPQ$. Together, this means that they form an algebra, and the operators $\langle D^m\rangle$ form the canonical pseudo-basis—just like the powers $\langle z^m\rangle$ in the Cauchy algebra $\mathbf{C}[[z]]$. The following proposition shows that the two algebras are actually *isomorphic*, and it allows to calculate the expansion coefficients of the series representation.

92 Proposition (Expansion of Shift-Invariant Operators)
The correspondence $\sum a_m z^m \leftrightarrow \sum a_m D^m$ is an isomorphism between the algebra of formal power series and the algebra of shift-invariant operators. Furthermore, we have the expansion

$$P = \sum \frac{LP\, z^m}{m!}\, D^m.$$

for every shift-invariant operator P. □

93 Proof (Expansion of Shift-Invariant Operators)
Let us begin with the expansion formula, which is an identity between operators. So we only need prove it on the basis $\langle z^m \rangle$, because this immediately implies validity on the whole vector space. For arbitrary $n \in \mathbf{N}$, we have

$$\mathsf{P}\mathsf{E}^u z^n = \mathsf{P}\,(z+u)^n = \sum_{m=0}^{n} \binom{n}{m} u^{n-m}\,(\mathsf{P}\,z^m).$$

This is not only a polynomial in z but also in u, hence we can regard u as a second indeterminate. For the rest of this proof, we will use an index on $\mathsf{P}, \mathsf{E}, \mathsf{D}$ and L for signifying which indeterminate they operate on. Picking the constant term with respect to z gives

$$\begin{aligned}\mathrm{L}_z\mathsf{P}_z\mathsf{E}_z^u z^n &= \sum_{m=0}^{n} \binom{n}{m} u^{n-m}\,(\mathrm{L}_z\mathsf{P}_z\,z^m) = \sum_{m=0}^{n} \frac{\mathrm{L}_u\mathsf{P}_u\,u^m}{m!}\, n^{\underline{m}} u^{n-m} \\ &= \sum_{m=0}^{n} \frac{\mathrm{L}_u\mathsf{P}_u\,u^m}{m!}\,\mathsf{D}_u^m\,u^n.\end{aligned}$$

Here we have used that $\mathrm{L}_x f(x) = \mathrm{L}_y f(y)$ for every power series f, which is nothing but renaming a bound variable.

On the other hand, we have $\mathrm{L}_z\mathsf{E}_z^u\mathsf{P}_z\,z^n = \mathsf{P}_u\,u^n$, because $\mathrm{L}_z\mathsf{E}_z^u$ has the effect of replacing z by u. But we know that P is a shift-invariant operator, so we have gained the identity

$$\mathsf{P}_u\,u^n = \left(\sum_{m=0}^{n} \frac{\mathrm{L}_u\mathsf{P}_u\,u^m}{m!}\,\mathsf{D}_u^m\right) u^n,$$

which concludes the proof of the expansion formula.

Next let us show that $\Phi\colon \sum a_m z^m \to \sum a_m \mathsf{D}^m$ is an isomorphism. Injectivity is trivial, and surjectivity follows from the expansion formula we have just proved. So it remains to show that $\Phi(f)\,\Phi(g) = \Phi(fg)$ for arbitrary power series $f(z) = \sum f_m z^m$ and $g(z) = \sum g_m z^m$. The left-hand side is $\sum a_m \mathsf{D}^m$, where the coefficients can be evaluated by means of the expansion formula as

$$\begin{aligned}[\forall m \in \mathbf{N}]\ \ m!\,a_m &= \mathrm{L}\,\Phi(f)\,\Phi(g)\,z^m = \mathrm{L}\left(\sum_n f_n\mathsf{D}^n \sum_k g_k\mathsf{D}^k\right) z^m \\ &= \mathrm{L}\left(\sum_n \sum_k f_n g_k\,\mathsf{D}^{n+k}\right) z^m = \sum_{n=0}^{m} f_n g_{m-n}\,m!\,.\end{aligned}$$

The right-hand side is

$$\Phi\Bigl(\sum_m \Bigl(\sum_{n=0}^{m} f_n g_{m-n}\Bigr) z^m\Bigr) = \sum_m \Bigl(\sum_{n=0}^{m} f_n g_{m-n}\Bigr) D^m,$$

□ so the isomorphism is proved.

The series notation of operators allows to derive a little but useful lemma, which often comes in handy for *rewriting operator expressions*.

94 Lemma (Switching under the Constant-Functional)
We have the identity $L\, f(D)\, p(z) = L\, p(D)\, f(z)$ for every $f \in \mathbf{C}[[z]]$ and
□ every $p \in \mathbf{C}[z]$.

95 Proof (Switching under the Constant-Functional)
There are two vector spaces involved, namely $\mathbf{C}[[z]]$ and $\mathbf{C}[z]$. For the latter, it is sufficient to consider the powers $p(z) = z^m$ for all $m \in \mathbf{N}$, because they form a basis there—and not just a pseudo-basis as in $\mathbf{C}[[z]]$. So fix an arbitrary $m \in \mathbf{N}$ and an arbitrary power series $f(z) = \sum f_n z^n$. Then we must show $L\, f(D)\, z^m = LD^m\, f(z)$. The left-hand side is

$$L\, \Bigl(\sum_n f_n D^n\Bigr) z^m = L\, \Bigl(\sum_n f_n m^{\underline{n}}\, z^{m-n}\Bigr) = f_m\, m!,$$

□ and the right-hand side is also $LD^m \sum_n f_n z^n = f_m\, m!$.

Writing operators as series allows to give a direct *characterization of delta operators*: A shift-invariant operator $P = \sum a_m D^m$ fulfilling $P\,z \in \mathbf{C} \setminus \{0\}$ must have $a_0 = 0$; otherwise it would cause a linear term when applied to z. Furthermore, it must have $a_1 \neq 0$; otherwise $P\,z$ would vanish. This reveals that delta operators correspond to power series of order 1.

In the same way, we can *characterize invertible operators*. As expected, the inverse T^{-1} of an operator T is defined by $TT^{-1} = I = T^{-1}T$; if such a T^{-1} exists, the operator T is called invertible. As an abbreviation, we write T^{-m} for $(T^{-1})^m$. By the isomorphism to the Cauchy algebra, we know that invertible operators correspond to power series of order 0. It follows that every delta operator P can be written as DT with an invertible operator T.

Finally we come to the *core piece* of this section—the basis formula, which allows to calculate the basis sequence of an arbitrary delta operator. And this is where Lagrange inversion enters the stage.

96 Proposition (Basis Formula)
The basis sequence $\langle p_m(z) \rangle$ can be calculated by

$$p_0(z) = 1,$$
$$[\forall m > 0]\ \ p_m(z) = z\,(T^{-m}\,z^{m-1})$$

□ for any delta operator $P = DT$.

97 Proof (Basis Formula)
According to Proposition 92, we can write the given delta operator as $P = g(D)$, where g is a suitable power series of order 1. Our goal is to involve this series into a operator equation containing the basis polynomials $p_m(z)$. Interpreting such an equation in the Cauchy algebra, it will become an identity between power series, where the basis polynomials occur as coefficients. These can then be computed by Lagrange inversion.

The most natural candidate for an operator equation is the shift, because it induces a binomial expansion on the basis polynomials. (Note that our only information about them is the binomiality.) Choosing an arbitrary $n \in \mathbf{N}$, we have

$$\begin{aligned} E^a\,p_n(z) = p_n(z+a) &= \sum_m \binom{n}{m} p_m(a)\,p_{n-m}(z) \\ &= \sum_m \frac{p_m(a)}{m!}\,n^{\underline{m}}\,p_{n-m}(z) = \left(\sum_m \frac{p_m(a)}{m!}\,g(D)^m\right) p_n(z), \end{aligned}$$

since the iteration of the basis operator yields $P^m\,p_n(z) = g(D)^m\,p_n(z) = n^{\underline{m}}\,p_{n-m}(z)$, as already noted before. (The sums may be extended to infinity because all terms with $m > n$ vanish.)

The latter identity holds for every $p_n(z)$. As these form a basis of $\mathbf{C}[z]$, it is valid for all polynomials, which yields the desired operator equation

$$E^a = \sum_m \frac{p_m(a)}{m!}\,g(D)^m.$$

For interpreting it in the Cauchy algebra, we must determine the power series corresponding to E^a. By the expansion formula of Proposition 96, this can be done in a straight-forward way. Writing $E^a = \sum (a_m/m!)\,D^m$, the coefficients are computed by

$$[\forall m \in \mathbf{N}]\ \ a_m = LE^a\,z^m = L\,(z+a)^m = a^m,$$

so $E^a = \sum (a^m D^m)/m! = e^{aD}$. Now we can transform the operator equation into the corresponding power-series identity

$$e^{az} = \sum_m \frac{p_m(a)}{m!}\, g(z)^m,$$

whose coefficients are delivered by the Lagrange inversion inversion formulas in Corollary 43. For that we have to choose $f(z) = e^{az}$ and $e(z)$ as the power series corresponding to the operator T^{-1}. With these settings, we obtain

$$\begin{aligned} \frac{p_0(a)}{1} &= \left[z^0\right]\ e^{az} = 1, \\ [\forall m > 0]\ \frac{p_m(a)}{m!} &= \frac{1}{m}\ \left[z^{m-1}\right]\ ae^{az}\, e(z)^m. \end{aligned}$$

The formula for p_0 is trivial and in accordance with what is claimed. Then let us fix $m > 0$ for rewriting the above expression for $p_m(a)$. Using the Taylor representation of the coefficient functional (explained after Notation 18) and applying the switching property of Lemma 94, it becomes

$$\begin{aligned} p_m(a) &= (m-1)!\ \left[z^{m-1}\right]\ ae^{az}\, e(z)^m = LD^{m-1}\ ae^{az}\, e(z)^m \\ &= L\, ae^{aD}\, e(D)^m\, z^{m-1} = a\,(LE^a T^{-m}\, z^{m-1}). \end{aligned}$$

This can also be viewed as a polynomial in a, so we take a to be a second indeterminate. We noted in Proof 93 that LE^a had the effect of replacing z by a. Thus we come to $p_m(a) = a\,(T_a^{-m}\, a^{m-1})$, which is the claimed basis formula, if we substitute z for a. □

At last we can attack the *transformation problem* of this section. The desired connecting coefficients will be encoded in a polynomial sequence. This sequence is in turn encoded by its basis operator, which is given by the following theorem of Mullin and Rota. After the proof we shall calculate one example, which will clarify the actual solution procedure.

98 Theorem (Mullin and Rota)
Let $\langle p_m(z)\rangle$ and $\langle q_m(z)\rangle$ be any binomial sequences with connecting coefficients c_{mn}, so that

$$q_m(z) = \sum_{n=0}^{m} c_{mn}\, p_n(z).$$

Defining

$$r_m(z) = \sum_{n=0}^{m} c_{mn} z^n,$$

the sequence $\langle r_m(z) \rangle$ is binomial with basis operator $q(p^*(D))$, where $p(D)$ and $q(D)$ are the basis operators of $\langle p_m(z) \rangle$ and $\langle q_m(z) \rangle$, respectively. □

99 Proof (Mullin and Rota)

Let us define the operator Z by $Z z^m = p_m(z)$. This operator is obviously invertible; its inverse is merely the common representation of the polynomials $p_m(z)$ in the canonical basis. It follows that $Z r_m(z) = q_m(z)$ and $r_m(z) = Z^{-1} q_m(z)$ for all $m \in \mathbf{N}$.

Writing $P = p(D)$, we observe that

$$[\forall m > 0] \;\; ZD\, z^m = m\, Z\, z^{m-1} = m\, p_{m-1}(z) = P\, p_m(z) = PZ\, z^m$$

and trivially $ZD\,1 = 0 = PZ\,1$. This means that $P = ZDZ^{-1}$ and hence $P^m = ZD^mZ^{-1}$ for all $m \in \mathbf{N}$.

Next let us abbreviate the other operator $Q = q(D)$, further $r(z) = q(p^*(z))$, and the corresponding operator $R = r(D)$. For denoting its coefficients, we set $R = \sum \rho_m D^m$. Then we see that

$$\begin{aligned} ZRZ^{-1} &= Z\, r(D)\, Z^{-1} = Z \left(\sum \rho_m D^m \right) Z^{-1} = \sum \rho_m\, ZD^mZ^{-1} \\ &= \sum \rho_m P^m = r(P) = r(p(D)) = q(p^*(p(D))) = q(D) = Q, \end{aligned}$$

which implies $R = Z^{-1}QZ$.

As to be expected, R is indeed the basis operator of $\langle r_m(z) \rangle$. To see this, choose an arbitrary $m > 0$, and we realize that

$$R\, r_m(z) = Z^{-1}QZ\, r_m(z) = Z^{-1}Q\, q_m(z) = m\, Z^{-1}\, q_{m-1}(z) = m\, r_{m-1}(z).$$

The case $m = 0$ is settled by

$$R\, r_0(z) = Z^{-1}QZ\, r_0(z) = Z^{-1}Q\, q_0(z) = Z^{-1}\, 0 = 0.$$

Therefore R is indeed the basis operator of the sequence $\langle r_m(z) \rangle$. For proving its binomiality, we show that R is a delta operator. We know that the power series $p(z)$ and $q(z)$ have order 1 since they correspond to delta operators. The series $p^*(z)$ must also have order 1 because of Proposition 24. Composing series multiplies their order, so $\operatorname{ord} r = (\operatorname{ord} q)\,(\operatorname{ord} p^*) = 1$, which means that $R = r(D)$ is indeed a delta operator. □

There are of course numerous different applications of this theorem. Let us just pick one more or less *typical example* from [1, p. 190] for illustrating the basical procedure.

100 Example (Lah Numbers)

We want to compute the connecting coefficients between the falling and rising factorials, so for

$$z^{\overline{m}} = \sum_{n=0}^{m} c_{mn}\, z^{\underline{n}},$$

we have to determine the numbers c_{mn}.

It is a well-know fact that the falling and rising factorials are the basis sequences of the difference operators Δ and ∇, respectively—just as their continuous analogon, the differentiation operator D, has basis sequence $\langle z^m \rangle$. This can be verified readily by looking at

$$\begin{aligned}
\Delta\, z^{\underline{m}} &= (z+1)^{\underline{m}} - z^{\underline{m}} = (z+1)\, z^{\underline{m-1}} - z^{\underline{m-1}}\,(z-m+1) = m\, z^{\underline{m-1}},\\
\nabla\, z^{\overline{m}} &= z^{\overline{m}} - (z-1)^{\overline{m}} = z^{\overline{m-1}}\,(z+m-1) - (z-1)\, z^{\overline{m-1}} = m\, z^{\overline{m-1}}
\end{aligned}$$

and at the trivial differences $\Delta\, 1 = 0 = \nabla\, 1$.

The difference operators are defined by $\Delta = \mathsf{E} - \mathsf{I}$ and $\nabla = \mathsf{I} - \mathsf{E}^{-1}$, hence we have to set $p(z) = e^z - 1$ and $q(z) = 1 - e^{-z}$ for applying Theorem 98. From Example 78 we know that $p^*(z) = \log(1+z)$. Therefore the relevant composition is

$$q(p^*(z)) = 1 - \frac{1}{e^{p^*(z)}} = 1 - \frac{1}{1+z} = \frac{z}{1+z},$$

and the operator $\mathsf{D}(\mathsf{I}+\mathsf{D})^{-1}$ is the basis operator of the desired sequence $\langle r_m(z) \rangle$ as defined in the theorem. By the basis formula of Proposition 96, we can calculate this sequence via

$$\begin{aligned}
r_0(z) &= 1,\\
[\forall m > 0]\;\; r_m(z) &= z\, \left((\mathsf{I}+\mathsf{D})^m\, z^{m-1}\right) = z \sum_{n=0}^{m} \binom{m}{n} \mathsf{D}^n\, z^{m-1}.
\end{aligned}$$

For any $m > 0$, let us rewrite this as

$$\begin{aligned}
r_m(z) &= \sum_{n=0}^{m} \binom{m}{n} (m-1)^{\underline{n}}\, z^{m-n} = \sum_{n=0}^{m} \binom{m}{n} (m-1)^{\underline{m-n}}\, z^n\\
&= 0 + \sum_{n=1}^{m} \frac{m!}{n!\,(m-n)!}\, \frac{(m-1)!}{(n-1)!}\; z^n = \sum_{n=1}^{m} \frac{m!}{n!} \binom{m-1}{n-1} z^n.
\end{aligned}$$

We can now read the connecting coefficients c_{mn} directly from the $r_m(z)$ as

$$\begin{aligned} c_{00} &= 1, \\ [\forall m \in \mathbf{N}^*]\ c_{m0} &= 0, \\ [\forall m, n \in \mathbf{N}^*]\ c_{mn} &= \frac{m!}{n!}\binom{m-1}{n-1}. \end{aligned}$$

In the literature, these numbers are known as “signless Lah numbers”, see for example [1, p. 152] for details. □

In a similar way, we can—in principle—find the connecting coefficients for any other pair of binomial sequences. So the field of binomial sequences is indeed a wonderful example of a *fully developed mathematical theory*.

And further, by these, my son, be admonished: of making many books there is no end; and much study is a weariness of the flesh. Let us hear the conclusion of the whole matter: Fear God, and keep his commandments: for this is the whole duty of man.
Ecclesiastes 12:12–14

Bibliography

[1] Martin Aigner
Kombinatorik, Band I
Springer, 1975

[2] Bruno Buchberger, R. Loos
Algebraic Simplification
In: *Computer Algebra—Symbolic and Algebraic Summation*
Springer, 1983 (2nd edition)

[3] Johann Cigler
Kombinatorik, Lecture Notes
Universität Wien, 1981

[4] Georgy P. Egorychev
Integral Representation and the Computation of Combinatorial Sums, Vol. 59 of *Translations of Mathematical Monographs*
American Mathematical Society, 1984

[5] Ronald L. Graham, Donald E. Knuth, Oren Patashnik
Concrete Mathematics
Addison-Wesley, 1994 (2nd edition)

[6] Heinz-Richard Halder, Werner Heise
Einführung in die Kombinatorik
Carl Hanser, München, 1976

[7] Peter Henrici
Applied and Computational Complex Analysis
Wiley, 1974 (vol. 1), 1977 (vol. 2)

[8] Joseph Hofbauer
Lagrange-Inversion
Actes du Seminaire Lothringien de Combinatoire **6** (1982)

[9] Adolf Hurwitz, Richard Courant
Vorlesungen über allgemeine Funktionentheorie
Springer, 1964 (4th edition)

[10] André Joyal
Une théorie combinatoire des séries formelles
Adv. Math. **42** (1981), 1–82

[11] Christian Krattenthaler
Operator Methods and Lagrange Inversion:
A Unified Approach to Lagrange Formulas
Trans. Amer. Math. Soc. **305** (1988), 431–465

[12] Gilbert Labelle
Counting Asymmetric Enriched Trees
Journal of Symbolic Computation **14** (1992), 211–242

[13] Gilbert Labelle
Une nouvelle demonstration combinatoire
des formules d'inversion de Lagrange
Adv. Math. **42** (1981), 217–247

[14] Joseph Louis Lagrange
Nouvelle Méthode pour Résoudre les Équations Littérales
Gauthier-Villars, Paris, 1869
In: *Œvres de Lagrange*

[15] Stephen C. Milne, Gaurav Bhatnagar
A Characterization of Inverse Relations
Preprint, 1995

[16] Peter Paule, Istvan Nemes
A Canonical Form Guide to Symbolic Summation
To appear in:
Advances in the Design of Symbolic Computation Systems
Texts and Monographs in Symbolic Computation
A. Miola and M. Temperini (eds.), Springer, Wien/New York, 1997

[17] Peter Paule
Ein neuer Weg zur q-Lagrange Inversion
Bayreuther Mathematische Schriften **18** (1985), 1–37

[18] Peter Paule
Mellin's Series from One-Variable Lagrange Inversion
RISC Report Series 4.0, Linz, 1989

[19] Peter Paule
Analytische Kombinatorik, Lecture Notes
Johannes Kepler Universität, Linz, 1995

[20] Roberto Pirastu
On Combinatorial Identities: Symbolic Summation and Umbral Calculus, PhD Thesis
Johannes Kepler Universität, Linz, 1996.

[21] George Pólya, Robert E. Tarjan, Donald R. Woods
Notes on Introductory Combinatorics
Birkhäuser, Boston, 1983

[22] John Riordan
Combinatorial Identities
John Wiley & Sons, 1968

[23] John Riordan
An Introduction to Combinatorial Analysis
John Wiley & Sons, 1958

[24] John Riordan
Inverse Relations and Combinatorial Identities
Amer. Math. Monthly **71** (1964), 485–498

[25] Richard P. Stanley
Enumerative Combinatorics I
Wadsworth & Brooks, 1986

[26] Volker Strehl
Zykel-Enumeration bei lokal-strukturierten Funktionen, Habilitation Thesis
Universität Erlangen-Nürnberg, 1990

[27] Herbert S. Wilf
Generatingfunctionology
Academic Press, 1994

In every thing give thanks: for this is the will of God in Christ Jesus concerning you.
I. Thessalonians 5:18

Acknowledgements

The thesis having come to its end, I look back on an interesting year of work. Among the numerous charming fields of mathematics, combinatorics has a very special appeal of its own. Therefore my first thanks undoubtedly go to my advisor *Peter Paule* for introducing me to this noble realm and for being my wise tutor throughout the time.

Many people have contributed to this work in one way or another. Of course, I cannot mention them all here, but I like to thank them all. I am especially indebted to *Gilbert Labelle* for his generous permission to use his figures from [13]; without them it would be virtually impossible to understand the combinatorial proof of Section 2.5. The poor quality of my reproduction is due to the spoiling process of copying, scanning and converting.

Let me also express my gratitude to *James Samland* for his kind assistance in questions of style and grammar. *Renata Mühlbachler* was so nice as to let me print the thesis in Prof. Cooper's office, and *Waltraud Kratzer* did me the favor of scanning Labelle's figures. Thanks also to *Christian Domscha*, who had the good idea of taking Bible verses as chapter mottos.

Last but certainly not least, I owe thanks to my *Mom and Dad* for all of their support. They have made it possible for me to study mathematics, and they have ever since been loyal and encouraging parents to me.