

An invitation to formal power series*

Benjamin Sambale[†]

November 29, 2022

Dedicated to the memory of Christine Bessenrodt.

Abstract

This is a lecture on the theory of formal power series developed entirely without any analytic machinery. Combining ideas from various authors we are able to prove Newton's binomial theorem, Jacobi's triple product, the Rogers–Ramanujan identities and many other prominent results. We apply these methods to derive several combinatorial theorems including Ramanujan's partition congruences, generating functions of Stirling numbers and Jacobi's four-square theorem. We further discuss formal Laurent series and multivariate power series and end with a proof of MacMahon's master theorem.

Keywords: formal power series; Jacobi's triple product; partitions; Ramanujan; Stirling numbers; MacMahon's master theorem

AMS classification: 13F25, 16W60, 11D88, 11P84, 05A15, 05A17

Contents

1	Introduction	2
2	Definitions and basic properties	3
3	The toolkit	8
4	The main theorems	14
5	Applications to combinatorics	21
6	Stirling numbers	27
7	Laurent series	35
8	Multivariate power series	39
9	MacMahon's master theorem	46

*This version differs slightly from the published article at Jahresbericht DMV

[†]Institut für Algebra, Zahlentheorie und Diskrete Mathematik, Leibniz Universität Hannover, Welfengarten 1, 30167 Hannover, Germany, sambale@math.uni-hannover.de

1 Introduction

In a first course on abstract algebra students learn the difference between polynomial (real-valued) functions familiar from high school and formal polynomials defined over arbitrary fields. In courses on analysis they learn further that certain “well-behaved” functions possess a Taylor series expansion, i. e. a power series which converges in a neighborhood of a point. On the other hand, only specialized courses cover the formal world of power series where no convergence questions are asked.

The purpose of these expository notes is to give a far-reaching introduction to formal power series without appealing to any analytic machinery (we only use an elementary discrete metric). In doing so, we go well beyond a dated account undertaken by Niven [30] in 1969 (for instance, Niven cites Euler’s pentagonal theorem without proof). An alternative approach with different emphases can be found in Tutte [39, 40]. To illustrate the usefulness of formal power series we offer several combinatorial applications including some deep partition identities due to Ramanujan and others. This challenges the statement “While the formal analogies with ordinary calculus are undeniably beautiful, strictly speaking one can’t go much beyond Euler that way...” from the introduction of the recent book by Johnson [19]. While most proofs presented here are not new, they are scattered in the literature spanning five decades and cannot be found in a unified treatment to my knowledge. Our main source of inspiration is the accessible book by Hirschhorn [16] (albeit based on analytic reasoning) in combination with numerous articles cited when appropriate. The work on these notes was initiated by lectures on combinatorics and discrete mathematics at the universities of Jena and Hannover. I hope that the present notes may serve as the basis of seminars for undergraduate and graduate students alike. The prerequisites do not go beyond a basic abstract algebra course (from Section 7 on, some knowledge of algebraic and transcendental field extensions is assumed).

The material is organized as follows: In the upcoming section we define the ring of formal power series over an arbitrary field and discuss its basis properties. Thereafter, we introduce our toolkit consisting of compositions, derivations and exponentiations of power series. In Section 4 we first establish the binomial theorems of Newton and Gauss and later obtain Jacobi’s famous triple product identity, Euler’s pentagonal number theorem and the Rogers–Ramanujan identities. In the subsequent section we apply the methods to combinatorial problems to obtain a number of generating functions. Most notably, we prove Ramanujan’s partitions congruences (modulo 5 and 7) as well as his so-called “most beautiful” formula. Another section deals with Stirling numbers, permutations, Faulhaber’s formula and the Lagrange–Jacobi four-square theorem. In Section 7 we consider formal Laurent series in order to prove the Lagrange–Bürmann inversion formula and Puiseux’ theorem on the algebraic closure. In the following section, multivariate power series enter the picture. We give proofs of identities of Vieta, Girard–Newton and Waring on symmetric polynomials. We continue by developing multivariate versions of Leibniz’ differentiation rule, Faá di Bruno’s rule and the inverse function theorem. In the final section we go somewhat deeper by taking matrices into account. After establishing the Lagrange–Good inversion formula, we culminate by proving MacMahon’s master theorem. Along the way we indicate analytic counterparts, connections to other areas and insert a few exercises.

2 Definitions and basic properties

The sets of positive and non-negative integers are denoted by $\mathbb{N} = \{1, 2, \dots\}$ and $\mathbb{N}_0 = \{0, 1, \dots\}$ respectively.

Definition 2.1.

- (i) The letter K will always denote a (commutative) field. In this section there are no requirements on K , but at later stages we need that K has characteristic 0 or contains some roots of unity. At this point we often replace K by \mathbb{C} for convenience (and not for making analytic arguments available). This is not much loss of generality, since our theorems always involve at most countable many field elements, say a_1, a_2, \dots , and $\mathbb{Q}(a_1, a_2, \dots)$ can be embedded into \mathbb{C} .
- (ii) A (formal) *power series* over K is just an infinite sequence $\alpha = (a_0, a_1, \dots)$ with *coefficients* $a_0, a_1, \dots \in K$. The set of power series forms a K -vector space denoted by $K[[X]]$ with respect to the familiar componentwise operations:

$$\alpha + \beta := (a_0 + b_0, a_1 + b_1, \dots), \quad \lambda\alpha := (\lambda a_0, \lambda a_1, \dots),$$

where $\beta = (b_0, b_1, \dots) \in K[[X]]$ and $\lambda \in K$. We identify the elements of $a \in K$ with the *constant* power series $(a, 0, \dots)$. In general we call a_0 the *constant term* of α and set $\text{inf}(\alpha) := \inf\{n \in \mathbb{N}_0 : a_n \neq 0\}$ with $\text{inf}(0) = \text{inf} \emptyset = \infty$ (as a group theorist I avoid calling $\text{inf}(\alpha)$ the order of α as in many sources).

- (iii) To motivate a multiplication on $K[[X]]$ we introduce an *indeterminant* X and its powers

$$X^0 := 1 = (1, 0, \dots), \quad X = X^1 = (0, 1, 0, \dots), \quad X^2 = (0, 0, 1, 0, \dots), \quad \dots$$

We can now formally write

$$\alpha = \sum_{n=0}^{\infty} a_n X^n.$$

If there exists some $d \in \mathbb{N}_0$ with $a_n = 0$ for all $n > d$, then α is called a (formal) *polynomial*. The smallest d with this property is the *degree* $\text{deg}(\alpha)$ of α (by convention $\text{deg}(0) = -\infty$). In this case, $a_{\text{deg}(\alpha)}$ is the *leading coefficient* and α is called *monic* if $a_{\text{deg}(\alpha)} = 1$. The set of polynomials (inside $K[[X]]$) is denoted by $K[X]$.

- (iv) We borrow from the usual multiplication of polynomials (sometimes called *Cauchy product* or *discrete convolution*) to define

$$\alpha \cdot \beta := \sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_k b_{n-k} \right) X^n$$

for arbitrary $\alpha, \beta \in K[[X]]$ as above.

Note that $1, X, X^2, \dots$ is a K -basis of $K[X]$, but not of $K[[X]]$. Indeed, $K[[X]]$ has no countable basis. Opposed to a popular trend to rename X to q (as in [16]), we always keep X as “formal” as possible.

Lemma 2.2. *With the above defined addition and multiplication $(K[[X]], +, \cdot)$ is an integral domain with identity 1, i. e. $K[[X]]$ is a commutative ring such that $\alpha \cdot \beta \neq 0$ for all $\alpha, \beta \in K[[X]] \setminus \{0\}$. Moreover, K and $K[X]$ are subrings of $K[[X]]$.*

Proof. Most axioms follows from the definition in a straight-forward manner. To prove the associativity of \cdot , let $\alpha = (a_0, \dots)$, $\beta = (b_0, \dots)$ and $\gamma = (c_0, \dots)$ be power series. The n -th coefficient of $\alpha \cdot (\beta \cdot \gamma)$ is

$$\sum_{i=0}^n a_i \sum_{j=0}^{n-i} b_j c_{n-i-j} = \sum_{i+j+k=n} a_i b_j c_k = \sum_{i=0}^n \left(\sum_{j=0}^i a_j b_{i-j} \right) c_{n-i},$$

which happens to be the n -th coefficient of $(\alpha \cdot \beta) \cdot \gamma$.

Now let $\alpha \neq 0 \neq \beta$ with $k := \inf(\alpha)$ and $l := \inf(\beta)$. Then the $(k+l)$ -th coefficient of $\alpha \cdot \beta$ is $\sum_{i=0}^{k+l} a_i b_{k+l-i} = a_k b_l \neq 0$. In particular, $\inf(\alpha \cdot \beta) = \inf(\alpha) + \inf(\beta)$ and $\alpha \cdot \beta \neq 0$.

Since $K \subseteq K[X] \subseteq K[[X]]$ and the operations agree in these rings, it is clear that K and $K[X]$ are subrings of $K[[X]]$ (with the same neutral elements). \square

The above proof does not require K to be a field. It works more generally for integral domains and this is needed later in Definition 8.1. From now on we will usually omit the multiplication symbol \cdot and apply multiplications always before additions. For example, $\alpha\beta - \gamma$ is shorthand for $(\alpha \cdot \beta) + (-\gamma)$. Moreover, we often omit the summation index in writing $\sum a_n X^n$ if it is clear from the context. The scalar multiplication is compatible with the ring multiplication, i. e. $\lambda(\alpha\beta) = (\lambda\alpha)\beta = \alpha(\lambda\beta)$ for $\alpha, \beta \in K[[X]]$ and $\lambda \in K$. This turns $K[[X]]$ into a K -algebra.

Example 2.3.

(i) The following power series can be defined for any K :

$$1 - X, \quad \sum_{n=0}^{\infty} X^n, \quad \sum nX^n, \quad \sum (-1)^n X^n.$$

We compute

$$(1 - X) \sum_{n=0}^{\infty} X^n = \sum_{n=0}^{\infty} X^n - \sum_{n=1}^{\infty} X^n = 1.$$

(ii) For a field K of characteristic 0 (like $K = \mathbb{Q}, \mathbb{R}$ or \mathbb{C}) we can define the *formal exponential series*

$$\exp(X) := \sum_{n=0}^{\infty} \frac{X^n}{n!} = 1 + X + \frac{X^2}{2} + \frac{X^3}{6} + \dots \in K[[X]].$$

We will never write e^X for the exponential series, since Euler's number e simply does not live in the formal world.

Definition 2.4.

- (i) We call $\alpha \in K[[X]]$ *invertible* if there exists some $\beta \in K[[X]]$ such that $\alpha\beta = 1$. As usual, β is uniquely determined and we write $\alpha^{-1} := 1/\alpha := \beta$. As in any ring, the invertible elements form the group of units denoted by $K[[X]]^\times$.
- (ii) For $\alpha, \beta, \gamma \in K[[X]]$ we write more generally $\alpha = \frac{\beta}{\gamma}$ if $\alpha\gamma = \beta$ (regardless whether γ is invertible or not). For $k \in \mathbb{N}_0$ let $\alpha^k := \alpha \dots \alpha$ with k factors and $\alpha^{-k} := (\alpha^{-1})^k$ if $\alpha \in K[[X]]^\times$.
- (iii) For $\alpha \in K[[X]]$ let $(\alpha) := \{\alpha\beta : \beta \in K[[X]]\}$ the principal ideal generated by α .

The reader may know that every ideal of $K[X]$ is principal. The next lemma implies that every proper ideal of $K[[X]]$ is a power of (X) (hence, $K[[X]]$ is a discrete valuation ring with unique maximal ideal (X)).

Lemma 2.5. *Let $\alpha = \sum a_n X^n \in K[[X]]$. Then the following holds*

- (i) α is invertible if and only if $a_0 \neq 0$. Hence, $K[[X]]^\times = K[[X]] \setminus (X)$.
- (ii) If there exists some $m \in \mathbb{N}$ with $\alpha^m = 1$, then $\alpha \in K$. Hence, the elements of finite order in $K[[X]]^\times$ lie in K^\times .

Proof.

- (i) Let $\beta = \sum b_n X^n \in K[[X]]$ such that $\alpha\beta = 1$. Then $a_0 b_0 = 1$ and $a_0 \neq 0$. Assume conversely that $a_0 \neq 0$. We define $b_0, b_1, \dots \in K$ recursively by $b_0 := 1/a_0$ and

$$b_k := -\frac{1}{a_0} \sum_{i=1}^k a_i b_{k-i} \in K$$

for $k \geq 1$. Then

$$\sum_{i=0}^k a_i b_{k-i} = \begin{cases} 1 & \text{if } k = 0, \\ 0 & \text{if } k > 0. \end{cases}$$

Hence, $\alpha\beta = 1$ where $\beta := \sum b_n X^n$.

- (ii) We may assume that $m > 1$. For any prime divisor p of m it holds that $(\alpha^{m/p})^p = 1$. Thus, by induction on m , we may assume that $m = p$. By way of contradiction, suppose $\alpha \notin K$ and let $n := \min\{k \geq 1 : a_k \neq 0\}$. The n -th coefficient of $\alpha^p = 1$ is $pa_0^{p-1}a_n = 0$. Since α is invertible (indeed $\alpha^{-1} = \alpha^{m-1}$), we know $a_0 \neq 0$ and conclude that $p = 0$ in K (i. e. K has characteristic p). Now we investigate the coefficient of X^{np} in α^p . Obviously, it only depends on a_0, \dots, a_{np} . Since p divides $\binom{p}{k} = \frac{p(p-1)\dots(p-k+1)}{k!}$ for $0 < k < p$, the binomial theorem yields $(a_0 + a_1 X)^p = a_0^p + a_1^p X^p$. This familiar rule extends inductively to any finite number of summands. Hence,

$$(a_0 + \dots + a_{np} X^{np})^p = a_0^p + a_n^p X^{np} + a_{n+1}^p X^{(n+1)p} + \dots + a_{np}^p X^{np^2}.$$

In particular, the np -th coefficient of α^p is $a_n^p \neq 0$; a contradiction to $\alpha^p = 1$. □

Example 2.6.

- (i) By Example 2.3 we obtain the familiar formula for the *formal geometric series*

$$\frac{1}{1-X} = \sum X^n$$

- (ii) For any $\alpha \in K[[X]] \setminus \{1\}$ and $n \in \mathbb{N}$ an easy induction yields

$$\sum_{k=0}^{n-1} \alpha^k = \frac{1 - \alpha^n}{1 - \alpha}.$$

- (iii) For distinct $a, b \in K \setminus \{0\}$ one has the *partial fraction decomposition*

$$\frac{1}{(a+X)(b+X)} = \frac{1}{b-a} \left(\frac{1}{a+X} - \frac{1}{b+X} \right), \tag{2.1}$$

which can be generalized depending on the algebraic properties of K .

We now start forming infinite sums of power series. To justify this process we introduce a discrete norm, which behaves much simpler than the euclidean norm on \mathbb{C} , for instance.

Definition 2.7. For $\alpha = \sum a_n X^n \in K[[X]]$ let

$$|\alpha| := 2^{-\inf(\alpha)} \in \mathbb{R}$$

be the *norm* of α with the convention $|0| = 2^{-\infty} = 0$.

The number 2 in Definition 2.7 can of course be replaced by any real number greater than 1. Note that α is invertible if and only if $|\alpha| = 1$. The following lemma turns $K[[X]]$ into an ultrametric space.

Lemma 2.8. For $\alpha, \beta \in K[[X]]$ we have

- (i) $|\alpha| \geq 0$ with equality if and only if $\alpha = 0$,
- (ii) $|\alpha\beta| = |\alpha||\beta|$,
- (iii) $|\alpha + \beta| \leq \max\{|\alpha|, |\beta|\}$ with equality if $|\alpha| \neq |\beta|$.

Proof.

- (i) This follows from the definition.
- (ii) Without loss of generality, let $\alpha \neq 0 \neq \beta$. We have already seen in the proof of Lemma 2.2 that $\inf(\alpha\beta) = \inf(\alpha) + \inf(\beta)$.
- (iii) From $a_n + b_n \neq 0$ we obtain $a_n \neq 0$ or $b_n \neq 0$. It follows that $\inf(\alpha + \beta) \geq \min\{\inf(\alpha), \inf(\beta)\}$. This turns into the ultrametric inequality $|\alpha + \beta| \leq \max\{|\alpha|, |\beta|\}$. If $\inf(\alpha) > \inf(\beta)$, then clearly $\inf(\alpha + \beta) = \inf(\beta)$. \square

Theorem 2.9. The norm $\alpha \mapsto |\alpha|$ turns $K[[X]]$ into a Banach space.

Proof. Since $|\alpha + \beta| \leq \max\{|\alpha|, |\beta|\} \leq |\alpha| + |\beta|$ for $\alpha, \beta \in K[[X]]$, the map $\alpha \mapsto |\alpha|$ fulfills the properties of a norm in the analytic sense by Lemma 2.8. Now let $\alpha_1, \alpha_2, \dots \in K[[X]]$ be a Cauchy sequence with $\alpha_m = \sum a_{m,n} X^n$ for $m \geq 1$. For every $k \geq 0$ there exists some $M_k \geq 1$ such that $|\alpha_m - \alpha_{M_k}| < 2^{-k}$ for all $m \geq M_k$. This shows $a_{m,n} = a_{M_k,n}$ for all $m \geq M_k$ and $n \leq k$. Without loss of generality, we may assume that $M_0 \leq M_1 \leq \dots$. We define

$$a_k := a_{M_k, k}$$

and $\alpha = \sum a_k X^k$. Then $|\alpha - \alpha_m| < 2^{-k}$ for all $m \geq M_k$, i. e. $\lim_{m \rightarrow \infty} \alpha_m = \alpha$. Therefore, $K[[X]]$ is complete with respect to $|\cdot|$. \square

Note that $K[[X]]$ is the completion of $K[X]$ with respect to $|\cdot|$. In other words: power series can be regarded as Cauchy series of polynomials. For convergent sequences $(\alpha_k)_k$ and $(\beta_k)_k$ we have (as in any metric space with multiplication)

$$\lim_{k \rightarrow \infty} (\alpha_k + \beta_k) = \lim_{k \rightarrow \infty} \alpha_k + \lim_{k \rightarrow \infty} \beta_k, \quad \lim_{k \rightarrow \infty} (\alpha_k \beta_k) = \lim_{k \rightarrow \infty} \alpha_k \cdot \lim_{k \rightarrow \infty} \beta_k.$$

The infinite sum

$$\sum_{k=1}^{\infty} \alpha_k := \lim_{n \rightarrow \infty} \sum_{k=1}^n \alpha_k$$

can only converge if $(\alpha_k)_k$ is a *null sequence*, that is, $\lim_{k \rightarrow \infty} |\alpha_k| = 0$. Surprisingly and in stark contrast to euclidean spaces, the converse is also true as we are about to see. This crucial fact makes the arithmetic of formal power series much simpler than the analytic counterpart.

Lemma 2.10. *For every null sequence $\alpha_1, \alpha_2, \dots \in K[[X]]$ the series $\sum_{k=1}^{\infty} \alpha_k$ and $\prod_{k=1}^{\infty} (1 + \alpha_k)$ converge, i. e. they are well-defined in $K[[X]]$.*

Proof. By Theorem 2.9 it suffices to show that the partial sums form Cauchy sequences. For $\epsilon > 0$ let $N \geq 0$ such that $|\alpha_k| < \epsilon$ for all $k \geq N$. Then, for $k > l \geq N$, we have

$$\begin{aligned} \left| \sum_{i=1}^k \alpha_i - \sum_{i=1}^l \alpha_i \right| &= \left| \sum_{i=l+1}^k \alpha_i \right| \stackrel{2.8}{\leq} \max\{|\alpha_i| : i = l+1, \dots, k\} < \epsilon, \\ \left| \prod_{i=1}^k (1 + \alpha_i) - \prod_{i=1}^l (1 + \alpha_i) \right| &= \prod_{i=1}^l \underbrace{|1 + \alpha_i|}_{\leq 1} \left| \prod_{i=l+1}^k (1 + \alpha_i) - 1 \right| \leq \left| \sum_{\emptyset \neq I \subseteq \{l+1, \dots, k\}} \prod_{i \in I} \alpha_i \right| \\ &\leq \max\{|\alpha_i| : i = l+1, \dots, k\} < \epsilon. \quad \square \end{aligned}$$

We often regard finite sequences as null sequences by extending them silently. Let $\alpha_1, \alpha_2, \dots \in K[[X]]$ be a null sequence and $\alpha_k = \sum a_{k,n} X^n$ for $k \geq 1$. For every $n \geq 0$ only finitely many of the coefficients $a_{1,n}, a_{2,n}, \dots$ are non-zero. This shows that the coefficient of X^n in

$$\sum_{k=1}^{\infty} \alpha_k = \sum_{n=0}^{\infty} \left(\sum_{k=1}^{\infty} a_{k,n} \right) X^n \quad (2.2)$$

depends on only finitely many terms. The same reasoning applies to the $\prod_{k=1}^{\infty} (1 + \alpha_k)$.

For $\gamma \in K[[X]]$ and null sequences $(\alpha_k), (\beta_k)$ it holds that $\sum \alpha_k + \sum \beta_k = \sum (\alpha_k + \beta_k)$ and $\gamma \sum \alpha_k = \sum \gamma \alpha_k$ as expected. Moreover, a convergent sum does not depend on the order of summation. In fact, for every bijection $\pi : \mathbb{N} \rightarrow \mathbb{N}$ and $n \in \mathbb{N}$ there exists some $N \in \mathbb{N}$ such that $\pi(k) > n$ for all $k > N$. Hence,

$$\left| \sum_{k=1}^N \alpha_k - \sum_{k=1}^N \alpha_{\pi(k)} \right| \leq \max\{|\alpha_k| : k > n\} \rightarrow 0.$$

For similar reasons we can freely interchange the order of summations. More precisely, let $\alpha_{k,n} \in K[[X]]$ such that $\lim_{k+n \rightarrow \infty} \alpha_{k,n} = 0$. Then

$$\sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \alpha_{k,n} = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \alpha_{k,n},$$

because

$$\left| \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \alpha_{k,n} - \sum_{n=1}^N \sum_{k=1}^{\infty} \alpha_{k,n} \right| = \left| \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \alpha_{k,n} - \sum_{k=1}^{\infty} \sum_{n=1}^N \alpha_{k,n} \right| = \left| \sum_{k=1}^{\infty} \sum_{n=N+1}^{\infty} \alpha_{k,n} \right| \xrightarrow{N \rightarrow \infty} 0.$$

Example 2.11.

- (i) For $\alpha \in (X)$ we have $|\alpha^n| = |\alpha|^n \leq 2^{-n} \rightarrow 0$ and therefore $\sum \alpha^n = \frac{1}{1-\alpha}$. So we have substituted X by α in the geometric series. This will be generalized in Definition 3.1.

(ii) Since every non-negative integer has a unique 2-adic expansion, we obtain

$$\prod_{k=0}^{\infty} (1 + X^{2^k}) = 1 + X + X^2 + \dots = \frac{1}{1 - X}.$$

Equivalently,

$$\prod_{k=0}^{\infty} (1 + X^{2^k}) = \prod \frac{(1 + X^{2^k})(1 - X^{2^k})}{1 - X^{2^k}} = \prod \frac{1 - X^{2^{k+1}}}{1 - X^{2^k}} = \frac{1}{1 - X}.$$

More interesting series will be discussed in Section 5.

3 The toolkit

Definition 3.1. Let $\alpha = \sum a_n X^n \in K[[X]]$ and $\beta \in K[[X]]$ such that $\alpha \in K[X]$ or $\beta \in (X)$. We define

$$\alpha \circ \beta := \alpha(\beta) := \sum_{n=0}^{\infty} a_n \beta^n.$$

If α is a polynomial, it is clear that $\alpha(\beta)$ is a valid power series, while for $\beta \in (X)$ the convergence of $\alpha(\beta)$ is guaranteed by Lemma 2.10. In the following we will silently assume that one of these conditions is fulfilled. Observe that $|\alpha(\beta)| \leq |\alpha|$ if $\beta \in (X)$.

Example 3.2. For $\alpha = \sum a_n X^n \in K[[X]]$ we have $\alpha(0) = a_0$ and $\alpha(X^2) = \sum a_n X^{2n}$. On the other hand for $\alpha = \sum X^n$ we are not allowed to form $\alpha(1)$.

Lemma 3.3. For $\alpha, \beta, \gamma \in (X)$ and every null sequence $\alpha_1, \alpha_2, \dots \in K[[X]]$ we have

$$\left(\sum \alpha_k \right) \circ \beta = \sum \alpha_k(\beta), \quad (3.1)$$

$$\left(\prod (1 + \alpha_k) \right) \circ \beta = \prod (1 + \alpha_k(\beta)), \quad (3.2)$$

$$\alpha \circ (\beta \circ \gamma) = (\alpha \circ \beta) \circ \gamma. \quad (3.3)$$

Proof. Since $|\alpha_k(\beta)| \leq |\alpha_k| \rightarrow 0$ for $k \rightarrow \infty$, all series are well-defined. Using the notation from (2.2) we deduce:

$$\left(\sum \alpha_k \right) \circ \beta = \sum_{n=0}^{\infty} \left(\sum_{k=1}^{\infty} a_{k,n} \right) \beta^n = \sum_{k=1}^{\infty} \left(\sum_{n=0}^{\infty} a_{k,n} \beta^n \right) = \sum \alpha_k(\beta).$$

We begin proving (3.2) with only two factors, say $\alpha_1 = \sum a_n X^n$ and $\alpha_2 = \sum b_n X^n$:

$$(\alpha_1 \alpha_2) \circ \beta = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_k b_{n-k} \right) \beta^n = \sum_{n=0}^{\infty} \sum_{k=0}^n (a_k \beta^k) (b_{n-k} \beta^{n-k}) = (\alpha_1 \circ \beta) (\alpha_2 \circ \beta).$$

Inductively, (3.2) holds for finitely many factors. Hence,

$$\begin{aligned} \left| \left(\prod (1 + \alpha_k) \right) \circ \beta - \prod (1 + \alpha_k(\beta)) \right| &= \left| \left(\prod_{k=1}^{\infty} (1 + \alpha_k) - \prod_{k=1}^n (1 + \alpha_k) \right) \circ \beta \right| \\ &\leq \left| \prod_{k=1}^{\infty} (1 + \alpha_k) - \prod_{k=1}^n (1 + \alpha_k) \right| \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Finally, setting $\alpha = \sum a_n X^n$ we have

$$\alpha \circ (\beta \circ \gamma) = \sum a_n (\beta \circ \gamma)^n \stackrel{(3.2)}{=} \sum a_n (\beta^n \circ \gamma) \stackrel{(3.1)}{=} \left(\sum a_n \beta^n \right) \circ \gamma = (\alpha \circ \beta) \circ \gamma. \quad \square$$

We warn the reader that in general

$$\alpha \circ \beta \neq \beta \circ \alpha, \quad \alpha \circ (\beta \gamma) \neq (\alpha \circ \beta)(\alpha \circ \gamma), \quad \alpha \circ (\beta + \gamma) \neq \alpha \circ \beta + \alpha \circ \gamma.$$

Nevertheless, the last statement can be corrected for the exponential series (Lemma 3.6).

Theorem 3.4. *The set $K[[X]]^\circ := (X) \setminus (X^2) \subseteq K[[X]]$ forms a group with respect to \circ .*

Proof. Let $\alpha, \beta \in K[[X]]^\circ$. Then $\alpha(\beta) \in K[[X]]^\circ$, i. e. $K[[X]]^\circ$ is closed under \circ . The associativity holds by (3.3). By definition, $X \in K[[X]]^\circ$ and $X \circ \alpha = \alpha = \alpha \circ X$.

To construct inverses we argue as in Lemma 2.5. Let $\alpha^k = \sum_{n=0}^{\infty} a_{k,n} X^n$ for $k \in \mathbb{N}_0$. Since $a_{1,0} = 0$, also $a_{k,n} = 0$ for $n < k$ and $a_{n,n} = a_{1,1}^n \neq 0$. We define recursively $b_0 := 0$, $b_1 := \frac{1}{a_{1,1}} \neq 0$ and

$$b_n := -\frac{1}{a_{n,n}} \sum_{k=1}^{n-1} b_k a_{k,n}$$

for $n \geq 2$. Setting $\beta := \sum b_n X^n \in K[[X]]^\circ$, we obtain

$$\beta(\alpha) = \sum_{k=0}^{\infty} b_k \alpha^k = \sum_{k=1}^{\infty} \sum_{n=0}^{\infty} b_k a_{k,n} X^n = \sum_{n=0}^{\infty} \left(\sum_{k=1}^n b_k a_{k,n} \right) X^n = X.$$

Now replacing α by β , we find $\gamma \in K[[X]]^\circ$ such that $\gamma \circ \beta = X$. Hence,

$$\gamma = \gamma \circ X = \gamma \circ \beta \circ \alpha = X \circ \alpha = \alpha$$

and $\alpha \circ \beta = X$. □

For $\alpha \in K[[X]]^\circ$, we call the unique $\beta \in K[[X]]^\circ$ with $\alpha(\beta) = X = \beta(\alpha)$ the *reverse* of α . To avoid confusion with the inverse α^{-1} (which is not defined here), we refrain from introducing a symbol for the reverse.

Example 3.5.

(i) Let α be the reverse of $X + X^2 + \dots = \frac{X}{1-X}$. Then

$$X = \frac{\alpha}{1-\alpha}$$

and it follows that $\alpha = \frac{X}{1+X} = X - X^2 + X^3 - \dots$. This is an example of a *Möbius transformation*. In general, it is much harder to find a closed-form expression for the reverse. We do so for the exponential series with the help of formal derivatives (Example 3.11). Later we provide the explicit Lagrange–Bürmann inversion formula (Theorem 7.5) using the machinery of Laurent series.

(ii) For the field \mathbb{F}_p with p elements (where p is a prime), the subgroup $N_p := X + (X^2)$ of $\mathbb{F}_p[[X]]^\circ$ is called *Nottingham group*. One can show that every finite p -group is a subgroup of N_p , so it must have a very rich structure.

Lemma 3.6 (Functional equation). *For every null sequence $\alpha_1, \alpha_2, \dots \in (X) \subseteq \mathbb{C}[[X]]$,*

$$\boxed{\exp\left(\sum \alpha_k\right) = \prod \exp(\alpha_k).} \quad (3.4)$$

In particular, $\exp(kX) = \exp(X)^k$ for $k \in \mathbb{Z}$.

Proof. Since $\sum \alpha_k \in (X)$ and $\exp(\alpha_k) \in 1 + \alpha_k + \frac{\alpha_k^2}{2} + \dots$, both sides of (3.4) are well-defined. For two summands $\alpha, \beta \in (X)$ we compute

$$\begin{aligned} \exp(\alpha + \beta) &= \sum \frac{(\alpha + \beta)^n}{n!} = \sum_{n=0}^{\infty} \sum_{k=0}^n \binom{n}{k} \frac{\alpha^k \beta^{n-k}}{n!} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{\alpha^k \beta^{n-k}}{k!(n-k)!} = \sum \frac{\alpha^n}{n!} \cdot \sum \frac{\beta^n}{n!} = \exp(\alpha) \exp(\beta). \end{aligned}$$

By induction we obtain (3.4) for finitely many summands. This implies

$$\begin{aligned} \left| \exp\left(\sum \alpha_k\right) - \prod_{k=1}^n \exp(\alpha_k) \right| &= \left| \exp\left(\sum_{k=1}^n \alpha_k + \sum_{k=n+1}^{\infty} \alpha_k\right) - \exp\left(\sum_{k=1}^n \alpha_k\right) \right| \\ &= \left| \exp\left(\sum_{k=1}^n \alpha_k\right) \left| \exp\left(\sum_{k=n+1}^{\infty} \alpha_k\right) - 1 \right| \right| \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

For the second claim let $k \in \mathbb{N}_0$. Then $\exp(kX) = \exp(X + \dots + X) = \exp(X)^k$. Since

$$\exp(kX) \exp(-kX) = \exp(kX - kX) = \exp(0) = 1,$$

we also have $\exp(-kX) = \exp(kX)^{-1} = \exp(X)^{-k}$. □

Definition 3.7. For $\alpha = \sum a_n X^n \in K[[X]]$ we call

$$\boxed{\alpha' := \sum_{n=1}^{\infty} n a_n X^{n-1} \in K[[X]]}$$

the (formal) *derivative* of α . Moreover, let $\alpha^{(0)} := \alpha$ and $\alpha^{(n)} := (\alpha^{(n-1)})'$ the n -th derivative for $n \in \mathbb{N}$.

It seems natural to define formal *integrals* as counterparts, but this is less useful, since in characteristic 0 we have $\alpha = \beta$ if and only if $\alpha' = \beta'$ and $\alpha(0) = \beta(0)$.

Example 3.8. As expected we have $1' = 0$, $X' = 1$ as well as

$$\exp(X)' = \sum_{n=1}^{\infty} n \frac{X^{n-1}}{n!} = \sum_{n=0}^{\infty} \frac{X^n}{n!} = \exp(X).$$

Note however, that $(X^p)' = 0$ if K has characteristic p .

In characteristic 0, derivatives provide a convenient way to extract coefficients of power series. For $\alpha = \sum a_n X^n \in \mathbb{C}[[X]]$ we see that $\alpha^{(0)}(0) = \alpha(0) = a_0$, $\alpha'(0) = a_1$, $\alpha''(0) = 2a_2, \dots, \alpha^{(n)}(0) = n!a_n$. Hence, *Taylor's theorem* (more precisely, the *Maclaurin series*) holds

$$\boxed{\alpha = \sum_{n=0}^{\infty} \frac{\alpha^{(n)}(0)}{n!} X^n.} \quad (3.5)$$

Over arbitrary fields we are not allowed to divide by $n!$. Alternatively, one may use the k -th *Hasse derivative* defined by

$$H^k(\alpha) := \sum_{n=k}^{\infty} \binom{n}{k} a_n X^{n-k}$$

(the integer $\binom{n}{k}$ can be embedded in any field). Note that $H^k(\alpha) = k! \alpha^{(k)}$ and $\alpha = \sum_{n=0}^{\infty} H^n(\alpha)(0) X^n$. In the following we restrict ourselves to complex power series.

Lemma 3.9. *For $\alpha, \beta \in \mathbb{C}[[X]]$ and every null sequence $\alpha_1, \alpha_2, \dots \in \mathbb{C}[[X]]$ the following rules hold:*

$$\begin{aligned} \left(\sum \alpha_k\right)' &= \sum \alpha'_k && \text{(sum rule),} \\ (\alpha\beta)' &= \alpha'\beta + \alpha\beta' && \text{((finite) product rule),} \\ \left(\prod(1 + \alpha_k)\right)' &= \prod(1 + \alpha_k) \sum \frac{\alpha'_k}{1 + \alpha_k}, && \text{((infinite) product rule),} \\ \left(\frac{\alpha}{\beta}\right)' &= \frac{\alpha'\beta - \alpha\beta'}{\beta^2} && \text{(quotient rule),} \\ (\alpha \circ \beta)' &= \alpha'(\beta)\beta' && \text{(chain rule).} \end{aligned}$$

Proof.

(i) Using the notation from (2.2), we have

$$\left(\sum \alpha_k\right)' = \left(\sum_{n=0}^{\infty} \sum_{k=1}^{\infty} a_{k,n} X^n\right)' = \sum_{n=0}^{\infty} \sum_{k=1}^{\infty} n a_{k,n} X^{n-1} = \sum_{k=1}^{\infty} \left(\sum_{n=0}^{\infty} n a_{k,n} X^{n-1}\right) = \sum \alpha'_k.$$

(ii) By (i) we may assume $\alpha = X^k$ and $\beta = X^l$. In this case,

$$(\alpha\beta)' = (X^{k+l})' = (k+l)X^{k+l-1} = kX^{k-1}X^l + lX^{l-1}X^k = \alpha'\beta + \beta'\alpha.$$

(iii) Without loss of generality, suppose $\alpha_k \neq -1$ for all $k \in \mathbb{N}$ (otherwise both sides vanish). Let $|\alpha_k| < 2^{-N-1}$ for all $k > n$. The coefficient of X^N on both sides of the equation depends only on $\alpha_1, \dots, \alpha_n$. From (ii) we verify inductively:

$$\left(\prod_{k=1}^n (1 + \alpha_k)\right)' = \prod_{k=1}^n (1 + \alpha_k) \sum_{l=1}^n \frac{\alpha'_l}{1 + \alpha_l}$$

for all $n \in \mathbb{N}$. Now the claim follows with $N \rightarrow \infty$.

(iv) By (ii),

$$\alpha' = \left(\frac{\alpha}{\beta}\beta\right)' = \left(\frac{\alpha}{\beta}\right)'\beta + \frac{\alpha\beta'}{\beta}.$$

(v) By (iii), the *power rule* $(\alpha^n)' = n\alpha^{n-1}\alpha'$ holds for $n \in \mathbb{N}_0$. The sum rule implies

$$(\alpha \circ \beta)' = \left(\sum a_n \beta^n \right)' = \sum a_n (\beta^n)' = \sum_{n=1}^{\infty} n a_n \beta^{n-1} \beta' = \alpha'(\beta) \beta'. \quad \square$$

The product rule implies the rather trivial *factor rule* $(\lambda\alpha)' = \lambda\alpha'$ as well as *Leibniz' rule*

$$(\alpha\beta)^{(n)} = \sum_{k=0}^n \binom{n}{k} \alpha^{(k)} \beta^{(n-k)}$$

for $\lambda \in \mathbb{C}$ and $\alpha, \beta \in \mathbb{C}[[X]]$. A generalized version of the latter and a chain rule for higher derivatives are proven in Section 8.

Exercise 3.10. Let $\alpha, \beta \in (X)$ such that $\beta \notin (X^2)$. Prove *L'Hôpital's rule* $\frac{\alpha}{\beta}(0) = \frac{\alpha'(0)}{\beta'(0)}$.

Example 3.11. Define the (formal) *logarithm* by the *Mercator series*

$$\log(1 + X) := \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} X^n = X - \frac{X^2}{2} + \frac{X^3}{3} \mp \dots \in \mathbb{C}[[X]].$$

By Theorem 3.4, $\alpha := \exp(X) - 1$ possesses a reverse and $\log(\exp(X)) = \log(1 + \alpha) \in \mathbb{C}[[X]]^\circ$. Since

$$\log(1 + X)' = 1 - X + X^2 \mp \dots = \sum (-X)^n = \frac{1}{1 + X},$$

the chain rule yields

$$\log(1 + \alpha)' = \frac{\alpha'}{1 + \alpha} = \frac{\exp(X)}{\exp(X)} = 1.$$

This shows that $\log(\exp(X)) = X$. Therefore, $\log(1 + X)$ is the reverse of $\alpha = \exp(X) - 1$ as expected from analysis. Moreover, $\log(1 - X) = -\sum_{n=1}^{\infty} \frac{X^n}{n}$.

The only reason why we called the power series $\log(1 + X)$ instead of $\log(X)$ or just \log is to keep the analogy to the natural logarithm (as an analytic function).

Lemma 3.12 (Functional equation). *For every null sequence $\alpha_1, \alpha_2, \dots \in (X) \subseteq \mathbb{C}[[X]]$,*

$$\log\left(\prod (1 + \alpha_k)\right) = \sum \log(1 + \alpha_k). \quad (3.6)$$

Proof.

$$\begin{aligned} \log\left(\prod (1 + \alpha_k)\right) &= \log\left(\prod \exp(\log(1 + \alpha_k))\right) \stackrel{(3.4)}{=} \log\left(\exp\left(\sum \log(1 + \alpha_k)\right)\right) \\ &= \sum \log(1 + \alpha_k). \end{aligned} \quad \square$$

Example 3.13. By (3.6),

$$\log\left(\frac{1}{1 - X}\right) = -\log(1 - X) = \sum_{n=1}^{\infty} \frac{X^n}{n}.$$

Definition 3.14. For $c \in \mathbb{C}$ and $\alpha \in (X)$ let

$$(1 + \alpha)^c := \exp(c \log(1 + \alpha)).$$

If $c = 1/k$ for some $k \in \mathbb{N}$, we write more customary $\sqrt[k]{1 + \alpha} := (1 + \alpha)^{1/k}$ and in particular $\sqrt{1 + \alpha} := \sqrt[2]{1 + \alpha}$.

By Lemma 3.6,

$$(1 + \alpha)^c (1 + \alpha)^d = \exp(c \log(1 + \alpha) + d \log(1 + \alpha)) = (1 + \alpha)^{c+d}$$

for every $c, d \in \mathbb{C}$ as expected. Consequently, $\sqrt[k]{1 + \alpha}^k = 1 + \alpha$ for $k \in \mathbb{N}$, i. e. $\sqrt[k]{1 + \alpha}$ is a k -th root of $1 + \alpha$ with constant term 1. Suppose that $\beta \in \mathbb{C}[[X]]$ also satisfies $\beta^k = 1 + \alpha$. Then $\beta^{-1} \sqrt[k]{1 + \alpha}$ has order $\leq k$ in $\mathbb{C}[[X]]^\times$. From Lemma 2.5 we conclude that $\beta^{-1} \sqrt[k]{1 + \alpha}$ is constant, i. e. $\beta = \beta(0) \sqrt[k]{1 + \alpha}$. Consequently, $\sqrt[k]{1 + \alpha}$ is the unique k -th of $1 + \alpha$ with constant term 1.

The inexperienced reader may find the following exercise helpful.

Exercise 3.15. Check that the following power series in $\mathbb{C}[[X]]$ are well-defined:

$$\begin{aligned} \sin(X) &:= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} X^{2n+1}, & \cos(X) &:= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} X^{2n}, \\ \tan(X) &:= \frac{\sin(X)}{\cos(X)}, & \sinh(X) &:= \sum_{k=0}^{\infty} \frac{X^{2k+1}}{(2k+1)!}, \\ \arcsin(X) &:= \sum_{n=0}^{\infty} \frac{(2n)!}{(2^n n!)^2} \frac{X^{2n+1}}{2n+1}, & \arctan(X) &:= \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} X^{2k+1}. \end{aligned}$$

Show that

(a) (EULER's formula) $\exp(iX) = \cos(X) + i \sin(X)$ where $i = \sqrt{-1} \in \mathbb{C}$.

(b) $\sin(2X) = 2 \sin(X) \cos(X)$ and $\cos(2X) = \cos(X)^2 - \sin(X)^2$.

Hint: Use (a) and separate real from non-real coefficients.

(c) (PYTHAGOREAN identity) $\cos(X)^2 + \sin(X)^2 = 1$.

(d) $\sinh(X) = \frac{1}{2}(\exp(X) - \exp(-X))$.

(e) $\sin(X)' = \cos(X)$ and $\cos(X)' = -\sin(X)$.

(f) $\arctan \circ \tan = X$.

Hint: Mimic the argument for $\log(1 + X)$.

(g) $\arctan(X) = \frac{i}{2} \log\left(\frac{i+X}{i-X}\right)$.

(h) $\arcsin(X)' = \frac{1}{\sqrt{1-X^2}}$.

(i) $\arcsin \circ \sin = X$.

4 The main theorems

For $c \in \mathbb{C}$ and $k \in \mathbb{N}$ we extend the definition of usual binomial coefficient by

$$\binom{c}{k} := \frac{c(c-1)\dots(c-k+1)}{k!} \in \mathbb{C}$$

(it is useful to know that numerator and denominator both have exactly k factors). The next theorem is a vast generalization of the binomial theorem (take $c \in \mathbb{N}$) and the geometric series (take $c = -1$).

Theorem 4.1 (NEWTON's binomial theorem). *For $\alpha \in (X)$ and $c \in \mathbb{C}$ the following holds*

$$\boxed{(1 + \alpha)^c = \sum_{k=0}^{\infty} \binom{c}{k} \alpha^k.} \quad (4.1)$$

Proof. It suffices to prove the equation for $\alpha = X$ (we may substitute X by α afterwards). By the chain rule,

$$((1 + X)^c)' = \exp(c \log(1 + X))' = c \frac{(1 + X)^c}{1 + X} = c(1 + X)^{c-1}$$

and inductively, $((1 + X)^c)^{(k)} = c(c-1)\dots(c-k+1)(1 + X)^{c-k}$. Now the claim follows from Taylor's theorem (3.5). \square

A striking application of Theorem 4.1 will be given in Theorem 6.12.

Example 4.2. Let $\zeta \in \mathbb{C}$ be an n -th root of unity and let $\alpha := (1 + X)^\zeta - 1 \in (X)$. Then

$$\alpha \circ \alpha = (1 + (1 + X)^\zeta - 1)^\zeta - 1 = (1 + X)^{\zeta^2} - 1$$

and inductively $\alpha \circ \dots \circ \alpha = (1 + X)^{\zeta^n} - 1 = X$. In particular, the order of α in the group $\mathbb{C}[[X]]^\circ$ divides n . Thus in contrast to the group $K[[X]]^\times$ studied in Lemma 2.5, the group $\mathbb{C}[[X]]^\circ$ possesses "interesting" elements of finite order.

Since we do not call our indeterminate q (as in many sources), it makes no sense to introduce the q -Pochhammer symbol $(q; q)_n$. Instead we devise a non-standard notation in reminiscence of the binomial coefficient.

Definition 4.3. For $n \in \mathbb{N}_0$ let $X^{n!} := (1 - X)(1 - X^2)\dots(1 - X^n)$. For $0 \leq k \leq n$ we call

$$\left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle := \frac{X^{n!}}{X^k! X^{n-k}!} = \frac{1 - X^n}{1 - X^k} \cdots \frac{1 - X^{n-k+1}}{1 - X} \in \mathbb{C}[[X]]$$

a *Gaussian coefficient*. If $k < 0$ or $k > n$ let $\left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle := 0$.

As for the binomial coefficients, we have $\left\langle \begin{matrix} n \\ 0 \end{matrix} \right\rangle = \left\langle \begin{matrix} n \\ n \end{matrix} \right\rangle = 1$ and $\left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle = \left\langle \begin{matrix} n \\ n-k \end{matrix} \right\rangle$ for all $n \in \mathbb{N}_0$ and $k \in \mathbb{Z}$. Moreover, $\left\langle \begin{matrix} n \\ 1 \end{matrix} \right\rangle = \frac{1 - X^n}{1 - X} = 1 + X + \dots + X^{n-1}$. The familiar recurrence formula for binomial coefficients needs to be altered as follows.

Lemma 4.4. *For $n \in \mathbb{N}_0$ and $k \in \mathbb{Z}$,*

$$\left\langle \begin{matrix} n+1 \\ k \end{matrix} \right\rangle = X^k \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle + \left\langle \begin{matrix} n \\ k-1 \end{matrix} \right\rangle = \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle + X^{n+1-k} \left\langle \begin{matrix} n \\ k-1 \end{matrix} \right\rangle. \quad (4.2)$$

Proof. For $k > n + 1$ or $k < 0$ all parts are 0. Similarly, for $k = n + 1$ or $k = 0$ both sides equal 1. Finally, for $1 \leq k \leq n$ it holds that

$$\begin{aligned} X^k \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle + \left\langle \begin{matrix} n \\ k-1 \end{matrix} \right\rangle &= \left(X^k \frac{1 - X^{n-k+1}}{1 - X^k} + 1 \right) \frac{X^n!}{X^{k-1}! X^{n-k+1}!} = \frac{1 - X^{n+1}}{1 - X^k} \frac{X^n!}{X^{k-1}! X^{n+1-k}!} \\ &= \left\langle \begin{matrix} n+1 \\ k \end{matrix} \right\rangle = \left\langle \begin{matrix} n+1 \\ n+1-k \end{matrix} \right\rangle = X^{n+1-k} \left\langle \begin{matrix} n \\ n+1-k \end{matrix} \right\rangle + \left\langle \begin{matrix} n \\ n-k \end{matrix} \right\rangle \\ &= \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle + X^{n+1-k} \left\langle \begin{matrix} n \\ k-1 \end{matrix} \right\rangle. \end{aligned} \quad \square$$

Since $\left\langle \begin{matrix} n \\ 0 \end{matrix} \right\rangle$ and $\left\langle \begin{matrix} n \\ 1 \end{matrix} \right\rangle$ are polynomials, (4.2) shows inductively that all Gaussian coefficients are polynomials. We may therefore evaluate $\left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle$ at $X = 1$. Indeed (4.2) becomes the recurrence for the binomial coefficients if $X = 1$. Hence $\left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle(1) = \binom{n}{k}$. This can be seen more directly by writing

$$\left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle = \frac{\frac{1-X^n}{1-X} \cdots \frac{1-X^{n-k+1}}{1-X}}{\frac{1-X^k}{1-X} \cdots \frac{1-X}{1-X}} = \frac{(1+X+\dots+X^{n-1}) \cdots (1+X+\dots+X^{n-k})}{(1+X+\dots+X^{k-1}) \cdots (1+X)1}.$$

We will interpret the coefficients of $\left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle$ in Theorem 5.4.

Example 4.5.

$$\left\langle \begin{matrix} 4 \\ 2 \end{matrix} \right\rangle = X^2 \left\langle \begin{matrix} 3 \\ 2 \end{matrix} \right\rangle + \left\langle \begin{matrix} 3 \\ 1 \end{matrix} \right\rangle = X^2(1+X+X^2) + (1+X+X^2) = 1+X+2X^2+X^3+X^4.$$

The next formulas resemble $(1+X)^n = \sum_{k=0}^n \binom{n}{k} X^k$ and $(1-X)^{-n} = \sum_{k=0}^{\infty} \binom{n+k-1}{k} X^k$ from Newton's binomial theorem. The first is a finite sum, the second an infinite sum arising from some sort of inverses. We will encounter many more such "dual pairs" in Theorem 6.9, Theorem 8.4 and (9.3), (9.4).

Theorem 4.6 (GAUSS' binomial theorem). *For $n \in \mathbb{N}$ and $\alpha \in \mathbb{C}[[X]]$ the following holds*

$$\boxed{\begin{aligned} \prod_{k=0}^{n-1} (1 + \alpha X^k) &= \sum_{k=0}^n \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle \alpha^k X^{\binom{k}{2}}, \\ \prod_{k=1}^n \frac{1}{1 - \alpha X^k} &= \sum_{k=0}^{\infty} \left\langle \begin{matrix} n+k-1 \\ k \end{matrix} \right\rangle \alpha^k X^k. \end{aligned}}$$

Proof. We argue by induction on n .

- (i) For $n = 1$ both sides become $1 + \alpha$. For the induction step we let all sums run from $-\infty$ to ∞ (this will not change their value, but makes index shifts much more transparent):

$$\begin{aligned} \prod_{k=0}^n (1 + \alpha X^k) &= (1 + \alpha X^n) \sum_{k=-\infty}^{\infty} \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle \alpha^k X^{\binom{k}{2}} \\ &= \sum \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle \alpha^k X^{\binom{k}{2}} + \sum \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle \alpha^{k+1} X^{n-k} X^{\binom{k}{2} + \binom{k+1}{2}} \\ &= \sum \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle \alpha^k X^{\binom{k}{2}} + \sum X^{n+1-k} \left\langle \begin{matrix} n \\ k-1 \end{matrix} \right\rangle \alpha^k X^{\binom{k}{2}} \\ &\stackrel{(4.2)}{=} \sum \left\langle \begin{matrix} n+1 \\ k \end{matrix} \right\rangle \alpha^k X^{\binom{k}{2}}. \end{aligned}$$

(ii) Here, $n = 1$ is the geometric series $\frac{1}{1-\alpha X} = \sum_{k=0}^{\infty} \alpha^k X^k$. In general:

$$\begin{aligned}
(1 - \alpha X^{n+1}) \sum_{k=0}^{\infty} \binom{n+k}{k} \alpha^k X^k &= \sum_{k=0}^{\infty} \binom{n+k}{k} \alpha^k X^k - X^n \sum_{k=0}^{\infty} \binom{n+k}{k} \alpha^{k+1} X^{k+1} \\
&= \sum_{k=0}^{\infty} \left(\binom{n+k}{k} - X^n \binom{n+k-1}{k-1} \right) \alpha^k X^k \\
&\stackrel{(4.2)}{=} \sum_{k=0}^{\infty} \binom{n+k-1}{k} \alpha^k X^k = \prod_{k=1}^n \frac{1}{1 - \alpha X^k}. \quad \square
\end{aligned}$$

Remark 4.7. We emphasize that in the proof of Theorem 4.6, α is treated as a variable independent of X . The proof and the statement are therefore still valid if we substitute X by some $\beta \in (X)$ *without* changing α to $\alpha(\beta)$.

Using

$$\lim_{n \rightarrow \infty} \binom{n}{k} = \frac{1}{X^k!} \lim_{n \rightarrow \infty} (1 - X^{n-k+1}) \dots (1 - X^n) = \frac{1}{X^k!}, \quad (4.3)$$

we obtain an infinite variant of Gauss' theorem:

Corollary 4.8 (EULER). *For all $\alpha \in \mathbb{C}[[X]]$,*

$$\prod_{k=0}^{\infty} (1 + \alpha X^k) = \sum_{k=0}^{\infty} \frac{\alpha^k X^{\binom{k}{2}}}{X^k!}, \quad (4.4)$$

$$\prod_{k=1}^{\infty} \frac{1}{1 - \alpha X^k} = \sum_{k=0}^{\infty} \frac{\alpha^k X^k}{X^k!}. \quad (4.5)$$

If $\alpha \in (X)$, we can apply (4.5) with αX^{-1} to obtain

$$\prod_{k=0}^{\infty} \frac{1}{1 - \alpha X^k} = \sum_{k=0}^{\infty} \frac{\alpha^k}{X^k!}. \quad (4.6)$$

We are now in a position to derive one of the most powerful theorems on power series.

Theorem 4.9 (JACOBI's triple product identity). *For every $\alpha \in \mathbb{C}[[X]] \setminus (X^2)$ the following holds*

$$\boxed{\prod_{k=1}^{\infty} (1 - X^{2k})(1 + \alpha X^{2k-1})(1 + \alpha^{-1} X^{2k-1}) = \sum_{k=-\infty}^{\infty} \alpha^k X^{k^2}.}$$

Proof. We follow Andrews [2]. Observe that $\alpha^{-1} X^{2k-1}$ is a power series for all $k \geq 1$, since $\alpha \notin (X^2)$. Also, both sides of the equation are well-defined. According to Remark 4.7 we are allowed to substitute X by X^2 and simultaneously α by $\alpha^{-1} X$ in (4.4):

$$\begin{aligned}
\prod_{k=1}^{\infty} (1 + \alpha^{-1} X^{2k-1}) &= \prod_{k=0}^{\infty} (1 + \alpha^{-1} X^{2k+1}) = \sum_{k=0}^{\infty} \frac{\alpha^{-k} X^{k^2}}{(1 - X^2) \dots (1 - X^{2k})} \\
&= \prod_{k=1}^{\infty} \frac{1}{1 - X^{2k}} \sum_{k=0}^{\infty} \alpha^{-k} X^{k^2} \prod_{l=0}^{\infty} (1 - X^{2l+2k+2}).
\end{aligned}$$

Again, $\alpha^{-k} X^{k^2}$ is still a power series. Since for negative k the inner product vanishes, we may sum over $k \in \mathbb{Z}$. A second application of (4.4) with X^2 instead of X and $-X^{2k+2}$ in the role of α shows

$$\begin{aligned} \prod_{k=1}^{\infty} (1 + \alpha^{-1} X^{2k-1})(1 - X^{2k}) &= \sum_{k=-\infty}^{\infty} \alpha^{-k} X^{k^2} \sum_{l=0}^{\infty} \frac{(-1)^l X^{l^2+l+2kl}}{(1 - X^2) \dots (1 - X^{2l})} \\ &= \sum_{l=0}^{\infty} \frac{(-\alpha X)^l}{(1 - X^2) \dots (1 - X^{2l})} \sum_{k=-\infty}^{\infty} X^{(k+l)^2} \alpha^{-k-l}. \end{aligned}$$

After the index shift $k \mapsto -k - l$, the inner sum does not depend on l anymore. We then apply (4.6) on the first sum with X replaced by X^2 and $-\alpha X \in (X)$ instead of α :

$$\prod_{k=1}^{\infty} (1 + \alpha^{-1} X^{2k-1})(1 - X^{2k}) = \prod_{k=0}^{\infty} \frac{1}{1 + \alpha X^{2k+1}} \sum_{k=-\infty}^{\infty} X^{k^2} \alpha^k = \prod_{k=1}^{\infty} \frac{1}{1 + \alpha X^{2k-1}} \sum_{k=-\infty}^{\infty} X^{k^2} \alpha^k.$$

We are done by rearranging terms. □

Remark 4.10. Since the above proof is just a combination of Euler's identities, we are still allowed to replace X and α individually. Furthermore, we have required the condition $\alpha \notin (X^2)$ only to guarantee that $\alpha^{-1} X$ is well-defined. If we replace X by X^m , say, the proof is still sound for $\alpha \in \mathbb{C}[[X]] \setminus (X^{m+1})$.

A (somewhat analytical) proof only making use of (4.4) can be found in [44]. There are numerous purely combinatorial proofs like [22, 26, 37, 38, 43, 45], which are meaningful for formal power series.

Example 4.11.

(i) Choosing $\alpha \in \{\pm 1, X\}$ in Theorem 4.9 reveals the following elegant identities:

$$\prod_{k=1}^{\infty} (1 - X^{2k})(1 + X^{2k-1})^2 = \sum_{k=-\infty}^{\infty} X^{k^2}, \quad (4.7)$$

$$\prod_{k=1}^{\infty} \frac{(1 - X^k)^2}{1 - X^{2k}} = \prod_{k=1}^{\infty} (1 - X^{2k})(1 - X^{2k-1})^2 = \sum_{k=-\infty}^{\infty} (-1)^k X^{k^2}, \quad (4.8)$$

$$\prod_{k=1}^{\infty} (1 - X^{2k})(1 + X^{2k})^2 = \frac{1}{2} \sum_{k=-\infty}^{\infty} X^{k^2+k} = \sum_{k=0}^{\infty} X^{k^2+k}, \quad (4.9)$$

where in (4.9) we made use of the bijection $k \mapsto -k - 1$ on \mathbb{Z} . These formulas are needed in the proof of Theorem 6.17. In (4.9) we find X only to even powers. By equating the corresponding coefficients, we may replace X^2 by X to obtain

$$\prod_{k=1}^{\infty} (1 - X^{2k})(1 + X^k) = \prod_{k=1}^{\infty} (1 - X^k)(1 + X^k)^2 = \sum_{k=0}^{\infty} X^{\frac{k^2+k}{2}}.$$

A very similar identity will be proved in Theorem 4.13.

(ii) Relying on Remark 4.10, we can replace X by X^3 and α by $-X$ at the same time in Theorem 4.9. This leads to

$$\prod_{k=1}^{\infty} (1 - X^{6k})(1 - X^{6k-2})(1 - X^{6k-4}) = \sum_{k=-\infty}^{\infty} (-1)^k X^{3k^2+k}.$$

Substituting X^2 by X provides Euler's celebrated *pentagonal number theorem*:

$$\boxed{\prod_{k=1}^{\infty} (1 - X^k) = \prod_{k=1}^{\infty} (1 - X^{3k})(1 - X^{3k-1})(1 - X^{3k-2}) = \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{3k^2+k}{2}}.} \quad (4.10)$$

There is a well-known combinatorial proof of (4.10) by Franklin, which is reproduced in the influential book by Hardy–Wright [12, Section 19.11].

- (iii) The following formulas arise in a similar manner by substituting X by X^5 and selecting $\alpha \in \{-X, -X^3\}$ afterwards (this is allowed by Remark 4.10):

$$\prod_{k=1}^{\infty} (1 - X^{5k})(1 - X^{5k-2})(1 - X^{5k-3}) = \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2+k}{2}}, \quad (4.11)$$

$$\prod_{k=1}^{\infty} (1 - X^{5k})(1 - X^{5k-1})(1 - X^{5k-4}) = \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2+3k}{2}}. \quad (4.12)$$

This will be used in the proof of Theorem 4.15.

To obtain yet another triple product identity, we first consider a finite version due to Hirschhorn [14].

Lemma 4.12. *For all $n \in \mathbb{N}_0$,*

$$\prod_{k=1}^n (1 - X^k)^2 = \sum_{k=0}^n (-1)^k (2k+1) X^{\frac{k^2+k}{2}} \left\langle \begin{matrix} 2n+1 \\ n-k \end{matrix} \right\rangle. \quad (4.13)$$

Proof. The proof is by induction on n : Both sides are 1 if $n = 0$. So assume $n \geq 1$ and let Q_n be the right hand side of (4.13). The summands of Q_n are invariant under the index shift $k \mapsto -k - 1$ and vanish for $|k| > n$. Hence, we may sum over $k \in \mathbb{Z}$ and divide by 2. A threefold application of (4.2) gives:

$$\begin{aligned} Q_n &= X^n \frac{1}{2} \sum_{k=-\infty}^{\infty} (-1)^k (2k+1) X^{\frac{k^2-k}{2}} \left\langle \begin{matrix} 2n \\ n-k \end{matrix} \right\rangle + \frac{1}{2} \sum_{k=-\infty}^{\infty} (-1)^k (2k+1) X^{\frac{k^2+k}{2}} \left\langle \begin{matrix} 2n \\ n-k-1 \end{matrix} \right\rangle \\ &= X^n \frac{1}{2} \sum_{k=-\infty}^{\infty} (-1)^k (2k+1) X^{\frac{k^2-k}{2}} \left\langle \begin{matrix} 2n-1 \\ n-k \end{matrix} \right\rangle + X^{2n} \frac{1}{2} \sum_{k=-\infty}^{\infty} (-1)^k (2k+1) X^{\frac{k^2+k}{2}} \left\langle \begin{matrix} 2n-1 \\ n-k-1 \end{matrix} \right\rangle \\ &\quad + \frac{1}{2} \sum_{k=-\infty}^{\infty} (-1)^k (2k+1) X^{\frac{k^2+k}{2}} \left\langle \begin{matrix} 2n-1 \\ n-k-1 \end{matrix} \right\rangle + X^n \frac{1}{2} \sum_{k=-\infty}^{\infty} (-1)^k (2k+1) X^{\frac{k^2+3k+2}{2}} \left\langle \begin{matrix} 2n-1 \\ n-k-2 \end{matrix} \right\rangle. \end{aligned}$$

The second and third sum amount to $(1 + X^{2n})Q_{n-1}$. We apply the transformations $k \mapsto k + 1$ and $k \mapsto k - 1$ in the first sum and fourth sum respectively:

$$\begin{aligned} Q_n &= (1 + X^{2n})Q_{n-1} - X^n \frac{1}{2} \sum_{k=-\infty}^{\infty} (-1)^k \left((2k+3) X^{\frac{k^2+k}{2}} \left\langle \begin{matrix} 2n-1 \\ n-k-1 \end{matrix} \right\rangle + (2k-1) X^{\frac{k^2+k}{2}} \left\langle \begin{matrix} 2n-1 \\ n-k-1 \end{matrix} \right\rangle \right) \\ &= (1 - X^{2n})Q_{n-1} - 2X^n Q_{n-1} = (1 - X^n)^2 Q_{n-1} = \prod_{k=1}^n (1 - X^k)^2. \quad \square \end{aligned}$$

Theorem 4.13 (JACOBI). *We have*

$$\boxed{\prod_{k=1}^{\infty} (1 - X^k)^3 = \sum_{k=0}^{\infty} (-1)^k (2k+1) X^{\frac{k^2+k}{2}}.} \quad (4.14)$$

Proof. By Lemma 4.12, we have

$$\begin{aligned}
\prod_{k=1}^{\infty} (1 - X^k)^3 &= \sum_{k=0}^{\infty} (-1)^k (2k+1) X^{\frac{k^2+k}{2}} \lim_{n \rightarrow \infty} \left\langle \begin{matrix} 2n+1 \\ n-k \end{matrix} \right\rangle \lim_{n \rightarrow \infty} \prod_{l=1}^{n+k+1} (1 - X^l) \\
&= \sum_{k=0}^{\infty} (-1)^k (2k+1) X^{\frac{k^2+k}{2}} \lim_{n \rightarrow \infty} (1 - X^{n-k+1}) \dots (1 - X^{2n+1}) \\
&= \sum_{k=0}^{\infty} (-1)^k (2k+1) X^{\frac{k^2+k}{2}}. \quad \square
\end{aligned}$$

In an analytic framework, (4.14) can be derived from Theorem 4.9 (see [12, Theorem 357]). A combinatorial proof was given in [20].

As a preparation for the infamous Rogers–Ramanujan identities [34], we start again with a finite version due to Bressoud [7]. The impatient reader may skip these technical results and start right away with the applications in Section 5 (Theorem 4.15 is only needed in Theorem 5.5(v),(vi)).

Lemma 4.14. For $n \in \mathbb{N}_0$,

$$\sum_{k=0}^{\infty} \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle X^{k^2} = \sum_{k=-\infty}^{\infty} (-1)^k \left\langle \begin{matrix} 2n \\ n+2k \end{matrix} \right\rangle X^{\frac{5k^2+k}{2}}, \quad (4.15)$$

$$\sum_{k=0}^{\infty} \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle X^{k^2+k} = \sum_{k=-\infty}^{\infty} (-1)^k \left\langle \begin{matrix} 2n+1 \\ n+2k \end{matrix} \right\rangle X^{\frac{5k^2-3k}{2}}. \quad (4.16)$$

Proof. We follow a simplified proof by Chapman [9]. Let α_n and $\tilde{\alpha}_n$ be the left and the right hand side respectively of (4.15). Similarly, let β_n and $\tilde{\beta}_n$ be the left and right hand side respectively of (4.16). Note that all four sums are actually finite. We show both equations at the same time by establishing a common recurrence relation between α_n , β_n and $\tilde{\alpha}_n$, $\tilde{\beta}_n$.

We compute $\alpha_0 = \beta_0 = \tilde{\alpha}_0 = \tilde{\beta}_0 = 1$. For $n \geq 1$,

$$\begin{aligned}
\alpha_n &\stackrel{(4.2)}{=} \sum_{k=-\infty}^{\infty} \left(\left\langle \begin{matrix} n-1 \\ k \end{matrix} \right\rangle + X^{n-k} \left\langle \begin{matrix} n-1 \\ k-1 \end{matrix} \right\rangle \right) X^{k^2} = \alpha_{n-1} + X^n \sum \left\langle \begin{matrix} n-1 \\ k-1 \end{matrix} \right\rangle X^{k(k-1)} \\
&= \alpha_{n-1} + X^n \sum \left\langle \begin{matrix} n-1 \\ k \end{matrix} \right\rangle X^{k(k+1)} = \alpha_{n-1} + X^n \beta_{n-1}, \\
\beta_n - X^n \alpha_n &= \sum_{k=-\infty}^{\infty} \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle X^{k^2+k} (1 - X^{n-k}) = \sum \frac{X^n!}{X^k! X^{n-k}!} X^{k^2+k} (1 - X^{n-k}) \\
&= (1 - X^n) \sum \left\langle \begin{matrix} n-1 \\ k \end{matrix} \right\rangle X^{k^2+k} = (1 - X^n) \beta_{n-1}.
\end{aligned}$$

These recurrences characterize α_n and β_n uniquely. The familiar index transformation $k \mapsto -k - 1$

implies $\sum (-1)^k \langle \frac{2n-2}{n+2k} \rangle X^{\frac{5(k^2+k)}{2}} = 0$. This is used in the following computation:

$$\begin{aligned}
\tilde{\alpha}_n - \tilde{\alpha}_{n-1} &= \sum_{k=-\infty}^{\infty} (-1)^k \left(\left\langle \frac{2n}{n+2k} \right\rangle - \left\langle \frac{2n-2}{n-1+2k} \right\rangle \right) X^{\frac{5k^2+k}{2}} \\
&\stackrel{(4.2)}{=} \sum (-1)^k \left(\left\langle \frac{2n-1}{n+2k} \right\rangle + X^{n-2k} \left\langle \frac{2n-1}{n+2k-1} \right\rangle - \left\langle \frac{2n-2}{n-1+2k} \right\rangle \right) X^{\frac{5k^2+k}{2}} \\
&\stackrel{(4.2)}{=} \sum (-1)^k \left(X^{n+2k} \left\langle \frac{2n-2}{n+2k} \right\rangle + X^{n-2k} \left\langle \frac{2n-1}{n+2k-1} \right\rangle \right) X^{\frac{5k^2+k}{2}} = X^n \tilde{\beta}_{n-1}, \\
\tilde{\beta}_n - X^n \tilde{\alpha}_n &= \sum_{k=-\infty}^{\infty} (-1)^k \left(\left\langle \frac{2n+1}{n+2k} \right\rangle - X^{n+2k} \left\langle \frac{2n}{n+2k} \right\rangle \right) X^{\frac{5k^2-3k}{2}} \\
&= \sum (-1)^k \left\langle \frac{2n}{n+2k-1} \right\rangle X^{\frac{5k^2-3k}{2}} \\
&= \sum (-1)^k \left(\left\langle \frac{2n-1}{n+2k-1} \right\rangle + X^{n-2k+1} \left\langle \frac{2n-1}{n+2k-2} \right\rangle \right) X^{\frac{5k^2-3k}{2}} \\
&= \tilde{\beta}_{n-1} + X^n \sum (-1)^k \left\langle \frac{2n-1}{n+2k-2} \right\rangle X^{\frac{5k^2-7k+2}{2}} \\
&= \tilde{\beta}_{n-1} + X^n \sum (-1)^{1-k} \left\langle \frac{2n-1}{n-2k} \right\rangle X^{\frac{5(1-k)^2-7(1-k)+2}{2}} \\
&= \tilde{\beta}_{n-1} - X^n \sum (-1)^k \left\langle \frac{2n-1}{n+2k-1} \right\rangle X^{\frac{5k^2-3k}{2}} = (1 - X^n) \tilde{\beta}_{n-1}.
\end{aligned}$$

By induction on n , it follows that $\alpha_n = \tilde{\alpha}_n$ and $\beta_n = \tilde{\beta}_n$ as desired. \square

Theorem 4.15 (ROGERS–RAMANUJAN identities). *We have*

$$\prod_{k=1}^{\infty} \frac{1}{(1 - X^{5k-1})(1 - X^{5k-4})} = \sum_{k=0}^{\infty} \frac{X^{k^2}}{X^{k!}}, \quad (4.17)$$

$$\prod_{k=1}^{\infty} \frac{1}{(1 - X^{5k-2})(1 - X^{5k-3})} = \sum_{k=0}^{\infty} \frac{X^{k^2+k}}{X^{k!}}. \quad (4.18)$$

Proof.

$$\begin{aligned}
\sum_{k=0}^{\infty} \frac{X^{k^2}}{X^{k!}} &\stackrel{(4.3)}{=} \sum_{k=0}^{\infty} X^{k^2} \lim_{n \rightarrow \infty} \left\langle \frac{n}{k} \right\rangle \stackrel{(4.15)}{=} \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2+k}{2}} \lim_{n \rightarrow \infty} \left\langle \frac{2n}{n+2k} \right\rangle \\
&\stackrel{(4.11)}{=} \frac{\prod_{k=1}^{\infty} (1 - X^{5k})(1 - X^{5k-2})(1 - X^{5k-3})}{\prod_{k=1}^{\infty} (1 - X^k)} = \prod_{k=1}^{\infty} \frac{1}{(1 - X^{5k-1})(1 - X^{5k-4})}, \\
\sum_{k=0}^{\infty} \frac{X^{k^2+k}}{X^{k!}} &\stackrel{(4.3)}{=} \sum_{k=0}^{\infty} X^{k^2+k} \lim_{n \rightarrow \infty} \left\langle \frac{n}{k} \right\rangle \stackrel{(4.16)}{=} \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2-3k}{2}} \lim_{n \rightarrow \infty} \left\langle \frac{2n+1}{n+2k} \right\rangle \\
&\stackrel{(4.12)}{=} \frac{\prod_{k=1}^{\infty} (1 - X^{5k})(1 - X^{5k-1})(1 - X^{5k-4})}{\prod_{k=1}^{\infty} (1 - X^k)} = \prod_{k=1}^{\infty} \frac{1}{(1 - X^{5k-2})(1 - X^{5k-3})}. \quad \square
\end{aligned}$$

The Rogers–Ramanujan identities were long believed to lie deeper within the theory of elliptic functions (Hardy [12, p. 385] wrote “No proof is really easy (and it would perhaps be unreasonable to expect an

easy proof.”; Andrews [4, p. 105] wrote “. . . no doubt it would be unreasonable to expect a really easy proof.”. Meanwhile a great number of proofs were found, some of which are combinatorial (see [3] or the recent book [35]). An interpretation of these identities is given in Theorem 5.5 below. We point out that there are many “finite identities”, like Lemma 4.14, approaching the Rogers–Ramanujan identities (as there are many rational sequences approaching $\sqrt{2}$).

One can find many more interesting identities, like the *quintuple product*, along with comprehensive references (and analytic proofs) in Johnson [19].

5 Applications to combinatorics

In this section we bring to life all of the abstract theorems and identities of the previous section. If a_0, a_1, \dots is a sequence of numbers usually arising from combinatorial context, the power series $\alpha = \sum a_n X^n$ is called the *generating function* of $(a_n)_n$. Although this seems pointless at first, power series manipulations often reveal explicit formulas for a_n , which can hardly be seen by inductive arguments. We give a first impression with the most familiar generating functions.

Example 5.1.

- (i) The number of k -element subsets of an n -element set is $\binom{n}{k}$ with generating function $(1 + X)^n$. A k -element multi-subset $\{a_1, \dots, a_k\}$ of $\{1, \dots, n\}$ with $a_1 \leq \dots \leq a_k$ (where elements are allowed to appear more than once) can be turned into a k -element subset $\{a_1, a_2 + 1, \dots, a_k + k - 1\}$ of $\{1, \dots, n + k - 1\}$ and vice versa. The number of k -element multi-subsets of an n -element set is therefore $\binom{n+k-1}{k}$ with generating function $(1 - X)^{-n}$ by Newton’s binomial theorem.
- (ii) The number of k -dimensional subspaces of an n -dimensional vector space over a finite field with $q < \infty$ elements is $\langle \binom{n}{k} \rangle$ evaluated at $X = q$ (indeed there are $(q^n - 1)(q^n - q) \dots (q^n - q^{n-k+1})$ linearly independent k -tuples and $(q^k - 1)(q^k - q) \dots (q^k - q^{k-1})$ of them span the same subspace). The generating function is closely related to Gauss’ binomial theorem.
- (iii) The *Fibonacci numbers* f_n are defined by $f_n := n$ for $n = 0, 1$ and $f_{n+1} := f_n + f_{n-1}$ for $n \geq 1$. The generating function α satisfies $\alpha = X + X^2\alpha + X\alpha$ and is therefore given by $\alpha = \frac{X}{1 - X - X^2}$. An application of the partial fraction decomposition (2.1) leads to the well-known *Binet formula*

$$f_n = \frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right)^n.$$

- (iv) The *Catalan numbers* c_n are defined by $c_n := n$ for $n = 0, 1$ and $c_n := \sum_{k=1}^{n-1} c_k c_{n-k}$ for $n \geq 2$ (most authors shift the index by 1). Its generating function α fulfills $\alpha - \alpha^2 = X$, i. e. it is the reverse of $X - X^2$. This quadratic equation has only one solution $\alpha = \frac{1}{2}(1 - \sqrt{1 - 4X})$ in $\mathbb{C}[[X]]^\circ$. Newton’s binomial theorem can be used to derive $c_{n+1} = \frac{1}{n+1} \binom{2n}{n}$. A slightly shorter proof based on Lagrange–Bürmann’s inversion formula is given in Example 7.6 below.

We now focus on combinatorial objects which defy explicit formulas.

Definition 5.2. A *partition* of $n \in \mathbb{N}$ is a sequence of positive integers $\lambda = (\lambda_1, \dots, \lambda_l)$ such that $\lambda_1 + \dots + \lambda_l = n$ and $\lambda_1 \geq \dots \geq \lambda_l$. We call $\lambda_1, \dots, \lambda_l$ the *parts* of λ . The set of partitions of n is denoted by $P(n)$ and its cardinality is $p(n) := |P(n)|$. For $k \in \mathbb{N}_0$ let $p_k(n)$ be the number of partitions of n with each part $\lambda_i \leq k$. Finally, let $p_{k,l}(n)$ be the number of partitions of n with each part $\leq k$ and at most l parts in total. Clearly, $p_1(n) = p_{n,1}(n) = 1$ and $p_n(n) = p_{n,n}(n) = p(n)$. Moreover,

$p_{k,l}(n) = 0$ whenever $n > kl$. For convenience let $p(0) = p_0(0) = p_{0,0}(0) = 1$ (0 can be interpreted as the empty sum).

Example 5.3. The partitions of $n = 7$ are

$$(7), (6, 1), (5, 2), (5, 1^2), (4, 3), (4, 2, 1), (4, 1^3), (3^2, 1), \\ (3, 2^2), (3, 2, 1^2), (3, 1^4), (2^3, 1), (2^2, 1^3), (2, 1^5), (1^7).$$

Hence, $p(7) = 15$, $p_3(7) = 8$ and $p_{3,3}(7) = 2$.

Theorem 5.4. *The generating functions of $p(n)$, $p_k(n)$ and $p_{k,l}(n)$ are given by*

$$\boxed{\sum_{n=0}^{\infty} p(n)X^n = \prod_{k=1}^{\infty} \frac{1}{1 - X^k}}, \\ \sum_{n=0}^{\infty} p_k(n)X^n = \frac{1}{X^{k!}}, \\ \sum_{n=0}^{\infty} p_{k,l}(n)X^n = \left\langle \begin{matrix} k+l \\ k \end{matrix} \right\rangle.$$

Proof. It is easy to see that $p_k(n)$ is the coefficient of X^n in

$$(1 + X^1 + X^{1+1} + \dots)(1 + X^2 + X^{2+2} + \dots) \dots (1 + X^k + X^{k+k} + \dots) \\ = \frac{1}{1 - X} \frac{1}{1 - X^2} \dots \frac{1}{1 - X^k} = \frac{1}{X^{k!}}. \quad (5.1)$$

This shows the second equation. The first follows from $p(n) = \lim_{k \rightarrow \infty} p_k(n)$. For the last claim we argue by induction on $k + l$ using (4.2). If $k = 0$ or $l = 0$, then both sides equal 1. Thus, let $k, l \geq 1$. Pick a partition $\lambda = (\lambda_1, \lambda_2, \dots)$ of n with each part $\leq k$ and at most l parts. If $\lambda_1 < k$, then all parts are $\leq k - 1$ and λ is counted by $p_{k-1,l}(n)$. If on the other hand $\lambda_1 = k$, then $(\lambda_2, \lambda_3, \dots)$ is counted by $p_{k,l-1}(n - k)$. Conversely, each partition counted by $p_{k,l-1}(n - k)$ can be extended to a partition counted by $p_{k,l}(n)$. We have proven the recurrence

$$p_{k,l}(n) = p_{k-1,l}(n) + p_{k,l-1}(n - k).$$

Induction yields

$$\sum p_{k,l}(n)X^n = \sum p_{k-1,l}(n)X^n + X^k \sum p_{k,l-1}(n)X^n \\ = \left\langle \begin{matrix} k+l-1 \\ k-1 \end{matrix} \right\rangle + X^k \left\langle \begin{matrix} k+l-1 \\ k \end{matrix} \right\rangle \stackrel{(4.2)}{=} \left\langle \begin{matrix} k+l \\ k \end{matrix} \right\rangle. \quad \square$$

Theorem 5.5. *The following assertions hold for $n, k, l \in \mathbb{N}_0$:*

- (i) $p_{k,l}(n) = p_{l,k}(n) = p_{k,l}(kl - n)$ for $n \leq kl$.
- (ii) *The number of partitions of n into exactly k parts is the number of partitions with largest part k .*
- (iii) (GLAISHER) *The number of partitions of n into parts not divisible by k equals the number of partitions with no part repeated k times (or more).*

- (iv) (EULER) *The number of partitions of n into unequal parts is the number of partitions into odd parts.*
- (v) (SCHUR) *The number of partitions of n in parts which differ by more than 1 equals the number of partitions in parts of the form $\pm 1 + 5k$.*
- (vi) (SCHUR) *The number of partitions of n in parts which differ by more than 1 and are larger than 1 equals the number of partitions into parts of the form $\pm 2 + 5k$.*

Proof.

- (i) Since $\langle \begin{smallmatrix} k+l \\ k \end{smallmatrix} \rangle = \langle \begin{smallmatrix} k+l \\ l \end{smallmatrix} \rangle$, we obtain $p_{k,l}(n) = p_{l,k}(n)$ by Theorem 5.4. Let $\lambda = (\lambda_1, \dots, \lambda_s)$ be a partition counted by $p_{k,l}(n)$. After adding zero parts if necessary, we may assume that $s = l$. Then $\bar{\lambda} := (k - \lambda_l, k - \lambda_{l-1}, \dots, k - \lambda_1)$ is a partition counted by $p_{k,l}(kl - n)$. Since $\bar{\bar{\lambda}} = \lambda$, we obtain a bijection between the partitions counted by $p_{k,l}(n)$ and $p_{k,l}(kl - n)$.
- (ii) The number of partitions of n with largest part k is $p_k(n) - p_{k-1}(n)$. The number of partitions with exactly k parts is

$$p_{n,k}(n) - p_{n,k-1}(n) \stackrel{(i)}{=} p_{k,n}(n) - p_{k-1,n}(n) = p_k(n) - p_{k-1}(n).$$

- (iii) Looking at (5.1) again, it turns out that the desired generating function is

$$\prod_{k \nmid m} \frac{1}{1 - X^m} = \prod_{m=1}^{\infty} \frac{1 - X^{km}}{1 - X^m} = (1 + X + \dots + X^{k-1})(1 + X^2 + \dots + X^{2(k-1)}) \dots$$

- (iv) Take $k = 2$ in (iii).
- (v) According to [35, Section 2.4], it was Schur, who first gave this interpretation of the Rogers–Ramanujan identities. The coefficient of X^n on the left hand side of (4.17) is the number of partitions into parts of the form $\pm 1 + 5k$. The right hand side can be rewritten (thanks to Theorem 5.4) as

$$\sum_{k=0}^{\infty} \sum_{n=0}^{\infty} p_k(n) X^{n+k^2} = \sum_{n=0}^{\infty} \sum_{k=0}^n p_k(n - k^2) X^n.$$

By (ii), $p_k(n - k^2)$ counts the partitions of $n - k^2$ with at most k parts. If $(\lambda_1, \dots, \lambda_k)$ is such a partition (allowing $\lambda_i = 0$ here), then $(\lambda_1 + 2k - 1, \lambda_2 + 2k - 3, \dots, \lambda_k + 1)$ is a partition of $n - k^2 + 1 + 3 + \dots + 2k - 1 = n$ with exactly k parts, which all differ by more than 1.

- (vi) This follows similarly using $k^2 + k = 2 + 4 + \dots + 2k$. □

There is a remarkable connection between (iii), (iv) and (v) of Theorem 5.5: Numbers not divisible by 3 are of the form $\pm 1 + 3k$, while odd numbers are of the form $\pm 1 + 4k$.

Example 5.6. For $n = 7$ the following partitions are counted by Theorem 5.5:

exactly three parts:	$(5, 1^2), (4, 2, 1), (3^2, 1), (3, 2^2)$
largest part 3:	$(3^2, 1), (3, 2^2), (3, 2, 1^2), (3, 1^4)$
unequal parts:	$(7), (6, 1), (5, 2), (4, 3), (4, 2, 1)$
odd parts:	$(7), (5, 1^2), (3^2, 1), (3, 1^4), (1^7)$
parts differ by more than 1:	$(7), (6, 1), (5, 2)$
parts of the form $\pm 1 + 5k$	$(6, 1), (4, 1^3), (1^7)$
parts ≥ 2 differ by more than 1:	$(7), (5, 2)$
parts of the form $\pm 2 + 5k$	$(7), (3, 2^2)$

Some of the statements in Theorem 5.5 permit nice combinatorial proofs utilizing *Young diagrams* (or *Ferres diagrams*). We refer the reader to the introductory book [5]. The following exercise (inspired by [5]) can be solved with formal power series.

Exercise 5.7. Prove the following statements for $n, k \in \mathbb{N}$:

- (a) The number of partition of n into even parts is the number of partitions whose parts have even multiplicity.
- (b) (LEGENDRE) If n is not of the form $\frac{1}{2}(3k^2 + k)$ with $k \in \mathbb{Z}$, then the number of partitions of n into an even number of unequal parts is the number of partitions into an odd number of unequal parts.
Hint: Where have we encountered $\frac{1}{2}(3k^2 + k)$ before?
- (c) (SUBBARAO) The number of partitions of n where each part appears 2, 3 or 5 times equals the number of partitions into parts of the form $\pm 2 + 12k, \pm 3 + 12k$ or $6 + 12k$.
- (d) (MACMAHON) The number of partitions of n where each part appears at least twice equals the number of partitions in parts not of the form $\pm 1 + 6k$.

The reader may have noticed that Euler's pentagonal number theorem (4.10) is just the inverse of the generating function of $p(n)$ from Theorem 5.4, i. e.

$$\sum_{n=0}^{\infty} p(n)X^n \cdot \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{3k^2+k}{2}} = 1$$

and therefore

$$\sum_{k=-n}^n (-1)^k p\left(n - \frac{3k^2 + k}{2}\right) = 0$$

for $n \in \mathbb{N}$, where $p(k) := 0$ whenever $k < 0$. This leads to a recurrence formula

$$\begin{aligned} p(0) &= 1, \\ p(n) &= p(n-1) + p(n-2) - p(n-5) - p(n-7) + \dots \quad (n \in \mathbb{N}). \end{aligned}$$

Example 5.8. We compute

$$\begin{aligned} p(1) &= p(0) = 1, & p(4) &= p(3) + p(2) = 3 + 2 = 5, \\ p(2) &= p(1) + p(0) = 2, & p(5) &= p(4) + p(3) - p(0) = 5 + 3 - 1 = 7, \\ p(3) &= p(2) + p(1) = 3, & p(6) &= p(5) + p(4) - p(1) = 7 + 5 - 1 = 11 \end{aligned}$$

(see <https://oeis.org/A000041> for more terms).

The generating functions we have seen so far all have integer coefficients. If $\alpha, \beta \in \mathbb{Z}[[X]]$ and $d \in \mathbb{N}$, we write $\alpha \equiv \beta \pmod{d}$, if all coefficients of $\alpha - \beta$ are divisible by d . This is compatible with the ring structure of $\mathbb{Z}[[X]]$, namely if $\alpha \equiv \beta \pmod{d}$ and $\gamma \equiv \delta \pmod{d}$, then $\alpha + \gamma \equiv \beta + \delta \pmod{d}$ and $\alpha\gamma \equiv \beta\delta \pmod{d}$. Now suppose $\alpha \in 1 + (X)$. Then the proof of Lemma 2.5 shows $\alpha^{-1} \in \mathbb{Z}[[X]]$. In this case $\alpha \equiv \beta \pmod{d}$ is equivalent to $\alpha^{-1} \equiv \beta^{-1} \pmod{d}$. If $d = p$ happens to be a prime, we have

$$(\alpha + \beta)^p = \sum_{k=0}^p \frac{p(p-1)\dots(p-k+1)}{k!} \alpha^k \beta^{p-k} \equiv \alpha^p + \beta^p \pmod{p},$$

as in any commutative ring of characteristic p .

With this preparation, we come to a remarkable discovery by Ramanujan [33].

Theorem 5.9 (RAMANUJAN). *The following congruences hold for all $n \in \mathbb{N}_0$:*

$$\boxed{p(5n+4) \equiv 0 \pmod{5}, \quad p(7n+5) \equiv 0 \pmod{7}.}$$

Proof. Let $\alpha := \prod(1 - X^k)$. By the remarks above, $\alpha^5 = \prod(1 - X^k)^5 \equiv \prod(1 - X^{5k}) \equiv \alpha(X^5) \pmod{5}$ and $\alpha^{-5} \equiv \alpha(X^5)^{-1} \pmod{5}$. For $k \in \mathbb{Z}$ we compute

$$\frac{1}{2}(k^2 + k) \equiv \begin{cases} 0 & \text{if } k \equiv 0, -1 \pmod{5}, \\ 1 & \text{if } k \equiv 1, -2 \pmod{5}, \\ 3 & \text{if } k \equiv 2 \pmod{5}. \end{cases}$$

This allows to write Jacobi's identity (4.14) in the form

$$\begin{aligned} \alpha^3 &= \sum_{k \equiv 0, -1 \pmod{5}} (-1)^k (2k+1) X^{\frac{k^2+k}{2}} + \sum_{k \equiv 1, -2 \pmod{5}} (-1)^k (2k+1) X^{\frac{k^2+k}{2}} + \sum_{k \equiv 2 \pmod{5}} (-1)^k \underbrace{(2k+1)}_{\equiv 0 \pmod{5}} X^{\frac{k^2+k}{2}} \\ &\equiv \alpha_0 + \alpha_1 \pmod{5}, \end{aligned}$$

where α_i is formed by the monomials $a_k X^k$ with $k \equiv i \pmod{5}$. Now Theorem 5.4 implies

$$\sum_{n=0}^{\infty} p(n) X^n = \alpha^{-1} = \frac{(\alpha^3)^3}{(\alpha^5)^2} \equiv \frac{(\alpha_0 + \alpha_1)^3}{\alpha(X^5)^2} \pmod{5}. \quad (5.2)$$

If we expand $(\alpha_0 + \alpha_1)^3$, then only terms X^k with $k \equiv 0, 1, 2, 3 \pmod{5}$ occur, while in $\alpha(X^5)^{-2}$ only terms X^{5k} occur. Therefore the right hand side of (5.2) contains no terms of the form X^{5k+4} . So we must have $p(5k+4) \equiv 0 \pmod{5}$.

For the congruence modulo 7 we compute similarly $\frac{1}{2}(k^2 + k) \equiv 0, 1, 3, 6 \pmod{7}$, where the last case only occurs if $k \equiv 3 \pmod{7}$ and in this case $2k+1 \equiv 0 \pmod{7}$. As before we may write $\alpha^3 \equiv \alpha_0 + \alpha_1 + \alpha_3 \pmod{7}$. Then

$$\sum_{n=0}^{\infty} p(n) X^n = \alpha^{-1} = \frac{(\alpha^3)^2}{\alpha^7} \equiv \frac{(\alpha_0 + \alpha_1 + \alpha_3)^2}{\alpha(X^7)} \pmod{7}.$$

Again X^{7k+5} does not appear on the right hand side. □

Ramanujan has also discovered the congruence $p(11n + 6) \equiv 0 \pmod{11}$ for all $n \in \mathbb{N}_0$ (the reader finds the history of this and other results in [5, 16], for instance). This was believed to be more difficult to prove, until elementary proofs were found by Marivani [28], Hirschhorn [15] and others (see also [16, Section 3.5]). The details are however extremely tedious to verify by hand.

By the Chinese remainder theorem, two congruence of coprime moduli can be combined as in

$$p(35n + 19) \equiv 0 \pmod{35}.$$

Ahlgren [1] (building on Ono [32]) has shown that in fact for every integer k coprime to 6 there is such a congruence modulo k . Unfortunately, they do not look as nice as Theorem 5.9. For instance,

$$p(11^3 \cdot 13n + 237) \equiv 0 \pmod{13}.$$

The next result explains the congruence modulo 5 and is known as Ramanujan's "most beautiful" formula (since Theorem 4.15 was first discovered by Rogers).

Theorem 5.10 (RAMANUJAN). *We have*

$$\boxed{\sum_{n=0}^{\infty} p(5n + 4)X^n = 5 \prod_{k=1}^{\infty} \frac{(1 - X^{5k})^5}{(1 - X^k)^6}.$$

Proof. The arguments are taken from [16, Chapter 5], leaving out some unessential details. This time we start with Euler's pentagonal number theorem. Since

$$\frac{3k^2 + k}{2} \equiv \begin{cases} 0 & \text{if } k \equiv 0, -2 \pmod{5}, \\ 1 & \text{if } k \equiv -1 \pmod{5}, \\ 2 & \text{if } k \equiv 1, 2 \pmod{5}, \end{cases}$$

we can write (4.10) in the form

$$\alpha := \prod (1 - X^k) = \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{3k^2+k}{2}} = \alpha_0 + \alpha_1 + \alpha_2,$$

where α_i is formed by the terms $a_k X^k$ with $k \equiv i \pmod{5}$. In fact,

$$\alpha_1 = \sum_{k=-\infty}^{\infty} (-1)^{5k-1} X^{\frac{3(5k-1)^2+5k-1}{2}} = -X \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{75k^2-25k}{2}} = -X\alpha(X^{25}). \quad (5.3)$$

On the other hand we have

$$\sum_{k=0}^{\infty} (-1)^k (2k + 1) X^{\frac{k^2+k}{2}} \stackrel{(4.14)}{=} \alpha^3 = (\alpha_0 + \alpha_1 + \alpha_2)^3.$$

When we expand the right hand side, the monomials of the form X^{5k+2} all occur in $3\alpha_0(\alpha_0\alpha_2 + \alpha_1^2)$. Since we have already realized in the proof of Theorem 5.9 that $(k^2 + k)/2 \not\equiv 2 \pmod{5}$, we conclude that

$$\alpha_1^2 = -\alpha_0\alpha_2. \quad (5.4)$$

Let $\zeta \in \mathbb{C}$ be a primitive 5-th root of unity. Using that

$$X^5 - 1 = \prod_{i=0}^4 (X - \zeta^i) = \zeta^{1+2+3+4} \prod (\zeta^{-i} X - 1) = \prod (\zeta^i X - 1),$$

we compute

$$\prod_{i=0}^4 \alpha(\zeta^i X) = \prod_{k=0}^{\infty} \prod_{i=0}^4 (1 - \zeta^{ik} X^k) = \alpha(X^5)^5 \prod_{5 \nmid k} (1 - X^{5k}) = \frac{\alpha(X^5)^6}{\alpha(X^{25})}.$$

This leads to

$$\begin{aligned} \sum p(n) X^n &= \frac{1}{\alpha} = \frac{\alpha(X^{25})}{\alpha(X^5)^6} \alpha(\zeta X) \alpha(\zeta^2 X) \alpha(\zeta^3 X) \alpha(\zeta^4 X) \\ &= \frac{\alpha(X^{25})}{\alpha(X^5)^6} (\alpha_0 + \zeta \alpha_1 + \zeta^2 \alpha_2) (\alpha_0 + \zeta^2 \alpha_1 + \zeta^4 \alpha_2) (\alpha_0 + \zeta^3 \alpha_1 + \zeta \alpha_2) (\alpha_0 + \zeta^4 \alpha_1 + \zeta^3 \alpha_2). \end{aligned} \quad (5.5)$$

We are only interested in the monomials X^{5n+4} . Those arise from the products $\alpha_0^2 \alpha_2^2$, $\alpha_0 \alpha_1^2 \alpha_2$ and α_1^4 . To facilitate the expansion of the right hand side of (5.5), we notice that the Galois automorphism γ of the cyclotomic field \mathbb{Q}_5 sending ζ to ζ^2 permutes the four factors cyclically. Whenever we obtain a product involving some ζ^i , say $\alpha_0^2 \alpha_2^2 \zeta^3$, the full orbit under $\langle \gamma \rangle$ must occur, which is $\alpha_0^2 \alpha_2^2 (\zeta + \zeta^2 + \zeta^3 + \zeta^4) = -\alpha_0^2 \alpha_2^2$. Now there are six choices to form $\alpha_0^2 \alpha_2^2$. Four of them form a Galois orbit, while the two remaining appear without ζ . The whole contribution is therefore $(1 + 1 - 1) \alpha_0^2 \alpha_2^2 = \alpha_0^2 \alpha_2^2$. In a similar manner we compute,

$$\sum p(5n+4) X^{5n+4} = \frac{\alpha(X^{25})}{\alpha(X^5)^6} (\alpha_0^2 \alpha_2^2 - 3\alpha_0 \alpha_1^2 \alpha_2 + \alpha_1^4) \stackrel{(5.4)}{=} 5 \frac{\alpha(X^{25})}{\alpha(X^5)^6} \alpha_1^4 \stackrel{(5.3)}{=} 5 X^4 \frac{\alpha(X^{25})^5}{\alpha(X^5)^6}.$$

The claim follows after dividing by X^4 and replacing X^5 by X . □

Partitions can be generalized to higher dimensions. A *plane partition* of $n \in \mathbb{N}$ is an $n \times n$ -matrix $\lambda = (\lambda_{ij})$ consisting of non-negative integers such that

- $\lambda_{i,1} \geq \lambda_{i,2} \geq \dots$ and $\lambda_{1,j} \geq \lambda_{2,j} \geq \dots$ for all i, j ,
- $\sum_{i,j=1}^n \lambda_{ij} = n$.

Ordinary partitions can be regarded as plane partitions with only one non-zero row. The number $pp(n)$ of plane partitions of n has the fascinating generating function

$$\sum_{n=0}^{\infty} pp(n) X^n = \prod_{k=1}^{\infty} \frac{1}{(1 - X^k)^k} = 1 + X + 3X^2 + 6X^3 + 13X^4 + 24X^5 + \dots$$

discovered by MacMahon (see [36, Corollary 7.20.3]).

6 Stirling numbers

We cannot resist to present a few more exciting combinatorial objects related to power series. Since there are literally hundreds of such combinatorial identities, our selection is inevitably biased by personal taste.

Definition 6.1. A set partition of $n \in \mathbb{N}$ is a disjoint union $A_1 \dot{\cup} \dots \dot{\cup} A_k = \{1, \dots, n\}$ of non-empty sets A_i in no particular order (we may require $\min A_1 < \dots < \min A_k$ to fix an order). The number of set partitions of n is called the n -th *Bell number* $b(n)$. The number of set partitions of n with exactly k parts is the *Stirling number of the second kind* $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$. In particular, $\left\{ \begin{smallmatrix} n \\ n \end{smallmatrix} \right\} = \left\{ \begin{smallmatrix} n \\ 1 \end{smallmatrix} \right\} = n$. We set $\left\{ \begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \right\} = b(0) = 1$ describing the empty partition of the empty set.

Example 6.2. The set partitions of $n = 3$ are

$$\{1, 2, 3\} = \{1\} \cup \{2, 3\} = \{1, 3\} \cup \{2\} = \{1, 2\} \cup \{3\} = \{1\} \cup \{2\} \cup \{3\}.$$

Hence, $b(3) = 5$ and $\left\{ \begin{smallmatrix} 3 \\ 2 \end{smallmatrix} \right\} = 3$.

Unlike the binomial or Gaussian coefficients the Stirling numbers do not obey a symmetry as in Pascal's triangle. While the generating functions of $b(n)$ and $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ have no particularly nice shape, there are close approximations which we are about to see.

Lemma 6.3. For $n, k \in \mathbb{N}_0$,

$$\left\{ \begin{smallmatrix} n+1 \\ k \end{smallmatrix} \right\} = k \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} + \left\{ \begin{smallmatrix} n \\ k-1 \end{smallmatrix} \right\}. \quad (6.1)$$

Proof. Without loss of generality, let $1 \leq k \leq n$. Let $A_1 \cup \dots \cup A_{k-1}$ a set partition of n with $k-1$ parts. Then $A_1 \cup \dots \cup A_{k-1} \cup \{n+1\}$ is a set partition of $n+1$ with k parts. Now let $A_1 \cup \dots \cup A_k$ be a set partition of n . We can add the number $n+1$ to each of the k sets A_1, \dots, A_k to obtain a set partition of $n+1$ with k parts. Conversely, every set partition of $n+1$ arises in precisely one of the two described ways. \square

Lemma 6.4. For $n \in \mathbb{N}_0$,

$$b(n+1) = \sum_{k=0}^n \binom{n}{k} b(k).$$

Proof. Every set partition \mathcal{A} of $n+1$ has a unique part A containing $n+1$. If $k := |A| - 1$, there are $\binom{n}{k}$ choices for A . Moreover, $\mathcal{A} \setminus A$ is a uniquely determined partition of the set $\{1, \dots, n\} \setminus A$ with $n-k$ elements. Hence, there are $b(n-k)$ possibilities for this partition. Consequently,

$$b(n+1) = \sum_{k=0}^n \binom{n}{k} b(n-k) = \sum_{k=0}^n \binom{n}{k} b(k). \quad \square$$

Theorem 6.5. For $n \in \mathbb{N}$ we have

$$\begin{aligned} \sum_{k=0}^n \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} X^k &= \exp(-X) \sum_{k=0}^{\infty} \frac{k^n}{k!} X^k, \\ \sum_{k=0}^{\infty} \frac{b(k)}{k!} X^k &= \exp(\exp(X) - 1). \end{aligned}$$

Proof. We prove both assertions by induction on n .

(i) For $n = 1$, we have

$$\exp(-X) \sum_{k=1}^{\infty} \frac{1}{(k-1)!} X^k = X \exp(-X) \exp(X) = X = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} X$$

as claimed. Assuming the claim for n , we have

$$\begin{aligned} \sum_{k=0}^{n+1} \begin{Bmatrix} n+1 \\ k \end{Bmatrix} X^k &\stackrel{(6.1)}{=} \sum k \begin{Bmatrix} n \\ k \end{Bmatrix} X^k + \sum \begin{Bmatrix} n \\ k-1 \end{Bmatrix} X^k = X \left(\sum \begin{Bmatrix} n \\ k \end{Bmatrix} X^k \right)' + X \sum \begin{Bmatrix} n \\ k \end{Bmatrix} X^k \\ &= X \left(\exp(-X) \sum \frac{k^n}{k!} X^k \right)' + X \exp(-X) \sum \frac{k^n}{k!} X^k \\ &= \exp(-X) \sum \frac{k^{n+1}}{k!} X^k. \end{aligned}$$

(ii) Since $\exp(X) - 1 \in (X)$, we can substitute X by $\exp(X) - 1$ in $\exp(X)$. Let

$$\alpha := \exp(\exp(X) - 1) = \sum \frac{a_n}{n!} X^n.$$

Then $a_0 = \exp(\exp(0) - 1) = \exp(0) = 1 = b(0)$. The chain rule gives

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{a_{n+1}}{n!} X^n &= \alpha' = \exp(X) \exp(\exp(X) - 1) \\ &= \left(\sum_{k=0}^{\infty} \frac{1}{k!} X^k \right) \left(\sum_{k=0}^{\infty} \frac{a_k}{k!} X^k \right) = \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{a_k}{k!(n-k)!} X^n. \end{aligned}$$

Therefore, $a_{n+1} = \sum_{k=0}^n \binom{n}{k} a_k$ for $n \geq 0$ and the claim follows from Lemma 6.4. \square

Now we discuss permutations.

Definition 6.6. Let S_n be the symmetric group consisting of all permutations on the set $\{1, \dots, n\}$. The number of permutations in S_n with exactly k (disjoint) cycles including fixed points is denoted by the *Stirling number of the first kind* $\begin{bmatrix} n \\ k \end{bmatrix}$. By agreement, $\begin{bmatrix} 0 \\ 0 \end{bmatrix} = 1$ (the identity on the empty set has zero cycles).

Example 6.7. There are $\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 11$ permutations in S_4 with exactly two cycles:

$$\begin{aligned} (1, 2, 3)(4), (1, 3, 2)(4), (1, 2, 4)(3), (1, 4, 2)(3), (1, 3, 4)(2), (1, 4, 3)(2), \\ (1)(2, 3, 4), (1)(2, 4, 3), (1, 2)(3, 4), (1, 3)(2, 4), (1, 4)(2, 3). \end{aligned}$$

Since $|S_n| = n!$, there is no need for a generating function of the number of permutations.

Lemma 6.8. For $k, n \in \mathbb{N}_0$,

$$\begin{bmatrix} n+1 \\ k \end{bmatrix} = \begin{bmatrix} n \\ k-1 \end{bmatrix} + n \begin{bmatrix} n \\ k \end{bmatrix}. \quad (6.2)$$

Proof. Without loss of generality, let $1 \leq k \leq n$. Let $\sigma \in S_n$ with exactly $k - 1$ cycles. By appending the 1-cycle $(n + 1)$ to σ we obtain a permutation counted by $\begin{bmatrix} n+1 \\ k \end{bmatrix}$. Now assume that σ has k cycles. When we write σ as a sequence of n numbers and $2k$ parentheses, there are n meaningful positions where we can add the digit $n + 1$. For example, there are three ways to add 4 in $\sigma = (1, 2)(3)$, namely

$$(4, 1, 2)(3), \quad (1, 4, 2)(3), \quad (1, 2)(4, 3).$$

This yields n distinct permutations counted by $\begin{bmatrix} n+1 \\ k \end{bmatrix}$. Conversely, every permutation counted by $\begin{bmatrix} n+1 \\ k \end{bmatrix}$ arises in precisely one of the described ways. \square

While the recurrences relations we have seen so far appear arbitrary, they can be explained in a unified way (see [23]).

It is time to present the next dual pair of formulas resembling Theorem 4.6.

Theorem 6.9. *The following generating functions of the Stirling numbers hold for $n \in \mathbb{N}$:*

$$\boxed{\begin{aligned} \prod_{k=0}^{n-1} (1 + kX) &= \sum_{k=0}^n \begin{bmatrix} n \\ n-k \end{bmatrix} X^k, \\ \prod_{k=1}^n \frac{1}{1 - kX} &= \sum_{k=0}^{\infty} \left\{ \begin{matrix} n+k \\ n \end{matrix} \right\} X^k. \end{aligned}}$$

Proof. This is another induction on n .

(i) The case $n = 1$ yields 1 on both sides of the equation. Assuming the claim for n , we compute

$$\begin{aligned} \prod_{k=0}^n (1 + kX) &= (1 + nX) \sum_{k=0}^n \begin{bmatrix} n \\ n-k \end{bmatrix} X^k = \sum \left(\begin{bmatrix} n \\ n-k \end{bmatrix} + n \begin{bmatrix} n \\ n-k+1 \end{bmatrix} \right) X^k \\ &\stackrel{(6.2)}{=} \sum \begin{bmatrix} n+1 \\ n+1-k \end{bmatrix} X^k. \end{aligned}$$

(ii) For $n = 1$, we get the geometric series $\frac{1}{1-X} = \sum X^k$ on both sides. Assume the claim for $n - 1$. Then

$$\begin{aligned} (1 - nX) \sum_{k=0}^{\infty} \left\{ \begin{matrix} n+k \\ n \end{matrix} \right\} X^k &= \sum \left(\left\{ \begin{matrix} n+k \\ n \end{matrix} \right\} - n \left\{ \begin{matrix} n+k-1 \\ n \end{matrix} \right\} \right) X^k \\ &\stackrel{(6.1)}{=} \sum \left\{ \begin{matrix} n-1+k \\ n-1 \end{matrix} \right\} X^k = \prod_{k=1}^{n-1} \frac{1}{1 - kX}. \quad \square \end{aligned}$$

For those who still do not have enough, the next exercise might be of interest.

Exercise 6.10.

- (a) Prove *Vandermonde's identity* $\sum_{k=0}^n \binom{a}{k} \binom{b}{n-k} = \binom{a+b}{n}$ for all $a, b \in \mathbb{C}$ by using Newton's binomial theorem.
- (b) For every prime p and $1 < k < p$, show that $\begin{bmatrix} p \\ k \end{bmatrix}$ is divisible by p (a property shared with $\binom{p}{k}$).

(c) Prove that

$$(-1)^n \frac{\log(1-X)^n}{n!} = \sum_{k=0}^{\infty} \binom{k}{n} \frac{X^k}{k!}$$

for $n \in \mathbb{N}_0$.

(d) Determine all $n \in \mathbb{N}$ such that the Catalan number c_n is odd.

Hint: Consider the generating function modulo 2.

(e) The *Bernoulli numbers* $b_n \in \mathbb{Q}$ are defined directly by their (exponential) generating function

$$\frac{X}{\exp(X) - 1} = \sum_{n=0}^{\infty} \frac{b_n}{n!} X^n.$$

Compute b_0, \dots, b_3 and show that $b_{2n+1} = 0$ for every $n \in \mathbb{N}$.

Hint: Replace X by $-X$.

The *cycle type* of a permutation $\sigma \in S_n$ is denoted by $(1^{a_1}, \dots, n^{a_n})$, meaning that σ has precisely a_k cycles of length k .

Lemma 6.11. *The number of permutations $\sigma \in S_n$ with cycle type $(1^{a_1}, \dots, n^{a_n})$ is*

$$\frac{n!}{1^{a_1} \dots n^{a_n} a_1! \dots a_n!}.$$

Proof. Each cycle of σ determines a subset of $\{1, \dots, n\}$. The number of possibilities to choose such subsets is given by the multinomial coefficient

$$\frac{n!}{(1!)^{a_1} \dots (n!)^{a_n}}.$$

Since the a_i subsets of size i can be permuted in $a_i!$ ways, each corresponding to the same permutation (as disjoint cycles commute), the number of relevant choices is only

$$\frac{n!}{(1!)^{a_1} \dots (n!)^{a_n} a_1! \dots a_n!}.$$

A given subset $\{\lambda_1, \dots, \lambda_k\} \subseteq \{1, \dots, n\}$ can be arranged in $k!$ permutations, but only $(k-1)!$ different cycles, since $(\lambda_1, \dots, \lambda_k) = (\lambda_2, \dots, \lambda_k, \lambda_1) = \dots$. Hence, the number of permutations in question is

$$\frac{n!}{(1!)^{a_1} \dots (n!)^{a_n} a_1! \dots a_n!} ((1-1)!)^{a_1} \dots ((n-1)!)^{a_n} = \frac{n!}{1^{a_1} \dots n^{a_n} a_1! \dots a_n!}. \quad \square$$

The following is a sibling to Glaisher's theorem. For a non-negative real number r we denote the largest integer $n \leq r$ by $n = \lfloor r \rfloor$.

Theorem 6.12 (ERDŐS–TURÁN). *Let $n, d \in \mathbb{N}$. The number of permutations in S_n , whose cycle lengths are not divisible by d is*

$$n! \prod_{k=1}^{\lfloor n/d \rfloor} \frac{kd-1}{kd}.$$

Proof. According to [29], the idea of the proof is credited to Pólya. We need to count permutations with cycle type $(1^{a_1}, \dots, n^{a_n})$ where $a_k = 0$ whenever $d \mid k$. By Lemma 6.11, the total number divided by $n!$ is the coefficient of X^n in

$$\begin{aligned} \prod_{\substack{k=1 \\ d \nmid k}}^{\infty} \sum_{a=0}^{\infty} \frac{1}{a!} \left(\frac{X^k}{k}\right)^a &= \prod_{d \nmid k} \exp\left(\frac{X^k}{k}\right) \stackrel{(3.4)}{=} \exp\left(\sum_{d \nmid k} \frac{X^k}{k}\right) = \exp\left(\sum_{k=1}^{\infty} \frac{X^k}{k} - \sum_{k=1}^{\infty} \frac{X^{dk}}{dk}\right) \\ &= \exp\left(-\log(1-X) + \frac{1}{d} \log(1-X^d)\right) \stackrel{(3.6)}{=} \sqrt[d]{1-X^d} \frac{1}{1-X} \\ &= \frac{1-X^d}{1-X} (1-X^d)^{\frac{1-d}{d}} \stackrel{(4.1)}{=} \left(\sum_{r=0}^{d-1} X^r\right) \left(\sum_{q=0}^{\infty} \binom{(1-d)/d}{q} (-X^d)^q\right). \end{aligned}$$

Therein, X^n appears if and only if $n = qd + r$ with $0 \leq r < d$ and $q = \lfloor n/d \rfloor$ (Euclidean division). In this case the coefficient is

$$(-1)^q \binom{(1-d)/d}{q} = (-1)^q \prod_{k=1}^q \frac{\frac{1}{d} - k}{k} = \prod_{k=1}^q \frac{kd-1}{kd}. \quad \square$$

Example 6.13. A permutation has odd order as an element of S_n if and only all its cycles have odd length. The number of such partitions is therefore

$$n! \prod_{k=1}^{\lfloor n/2 \rfloor} \frac{2k-1}{2k} = \begin{cases} 1^2 \cdot 3^2 \cdot \dots \cdot (n-1)^2 & \text{if } n \text{ is even,} \\ 1^2 \cdot 3^2 \cdot \dots \cdot (n-2)^2 \cdot n & \text{if } n \text{ is odd.} \end{cases}$$

Exercise 6.14. Find and prove a similar formula for the number of permutations $\sigma \in S_n$ whose cycle lengths are all divisible by d .

We insert a well-known application of Bernoulli numbers.

Theorem 6.15 (FAULHABER). For every $d \in \mathbb{N}$ there exists a polynomial $\alpha \in \mathbb{Q}[X]$ of degree $d+1$ such that $1^d + 2^d + \dots + n^d = \alpha(n)$ for every $n \in \mathbb{N}$.

Proof. We compute the generating function

$$\begin{aligned} \sum_{d=0}^{\infty} \left(\sum_{k=1}^{n-1} k^d\right) \frac{X^d}{d!} &= \sum_{k=1}^{n-1} \sum_{d=0}^{\infty} \frac{(kX)^d}{d!} = \sum_{k=1}^{n-1} \exp(kX) = \sum_{k=1}^{n-1} \exp(X)^k = \frac{\exp(X)^n - 1}{\exp(X) - 1} \\ &= \frac{\exp(nX) - 1}{X} \frac{X}{\exp(X) - 1} \stackrel{6.10}{=} \sum_{k=0}^{\infty} \frac{n^{k+1}}{(k+1)!} X^k \sum_{l=0}^{\infty} \frac{b_l}{l!} X^l \\ &= \sum_{d=0}^{\infty} \sum_{k=0}^d \left(\frac{n^{k+1} b_{d-k} d!}{(k+1)! (d-k)!}\right) \frac{X^d}{d!} \\ &= \sum_{d=0}^{\infty} \sum_{k=0}^d \left(\frac{1}{k+1} \binom{d}{k} b_{d-k} n^{k+1}\right) \frac{X^d}{d!} \end{aligned}$$

and define $\alpha := \sum_{k=0}^d \frac{1}{k+1} \binom{d}{k} b_{d-k} (X+1)^{k+1} \in \mathbb{Q}[X]$. Since $b_0 = 1$, α is a polynomial of degree $d+1$ with leading coefficient $\frac{1}{d+1}$. \square

Example 6.16. For $d = 3$ the formula in the proof evaluates with some effort (using Exercise 6.10) to:

$$\alpha = b_3(X+1) + \frac{3}{2}b_2(X+1)^2 + b_1(X+1)^3 + \frac{1}{4}b_0(X+1)^4 = \frac{1}{4}(X+1)^2X^2 = \binom{X+1}{2}^2.$$

This is known as *Nicomachus's identity*:

$$1^3 + 2^3 + \dots + n^3 = (1 + 2 + \dots + n)^2.$$

Even though Faulhaber's formula $1^d + 2^d + \dots + n^d = \alpha(n)$ has not much to do with power series, there still is a dual formula, again featuring Bernoulli numbers:

$$\sum_{k=1}^{\infty} \frac{1}{k^{2d}} = (-1)^{d+1} \frac{(2\pi)^{2d} b_{2d}}{2(2d)!} \quad (d \in \mathbb{N}).$$

Strangely, no such formula is known to hold for odd negative exponents (perhaps because $b_{2d+1} = 0$). In fact, it is unknown if *Apéry's constant* $\sum_{k=1}^{\infty} \frac{1}{k^3} = 1,202\dots$ is transcendent.

We end this section with a power series proof of the famous four-square theorem.

Theorem 6.17 (LAGRANGE–JACOBI). *Every positive integer is the sum of four squares. More precisely,*

$$q(n) := |\{(a, b, c, d) \in \mathbb{Z}^4 : a^2 + b^2 + c^2 + d^2 = n\}| = 8 \sum_{4 \nmid d | n} d$$

for $n \in \mathbb{N}$.

Proof. We follow [16, Section 2.4]. Obviously, it suffices to prove the second assertion (by Jacobi). Since the summands $(-1)^k(2k+1)X^{\frac{k^2+k}{2}}$ in (4.14) are invariant under the transformation $k \mapsto -k-1$, we can write

$$\prod_{k=1}^{\infty} (1 - X^k)^3 = \frac{1}{2} \sum_{k=-\infty}^{\infty} (-1)^k(2k+1)X^{\frac{k^2+k}{2}}.$$

Taking the square on both sides yields

$$\alpha := \prod_{k=1}^{\infty} (1 - X^k)^6 = \frac{1}{4} \sum_{k,l=-\infty}^{\infty} (-1)^{k+l}(2k+1)(2l+1)X^{\frac{k^2+k+l^2+l}{2}}.$$

The pairs (k, l) with $k \equiv l \pmod{2}$ are transformed by $(k, l) \mapsto (s, t) := \frac{1}{2}(k+l, k-l)$, while the pairs $k \not\equiv l \pmod{2}$ are transformed by $(s, t) := \frac{1}{2}(k-l-1, k+l+1)$. Notice that $k = s+t$ and $l = s-t$

or $l = t - s - 1$ respectively. Hence,

$$\begin{aligned}
\alpha &= \frac{1}{4} \sum_{s,t=-\infty}^{\infty} (2s+2t+1)(2s-2t+1)X^{\frac{(s+t)^2+s+t+(s-t)^2+s-t}{2}} \\
&\quad - \frac{1}{4} \sum_{s,t=-\infty}^{\infty} (2s+2t+1)(2t-2s-1)X^{\frac{(s+t)^2+s+t+(t-s-1)^2+t-s-1}{2}} \\
&= \frac{1}{4} \sum_{s,t} ((2s+1)^2 - (2t)^2)X^{s^2+s+t^2} - \frac{1}{4} \sum_{s,t} ((2t)^2 - (2s+1)^2)X^{s^2+s+t^2} \\
&= \frac{1}{2} \sum_{s,t} ((2s+1)^2 - (2t)^2)X^{s^2+s+t^2} \\
&= \frac{1}{2} \sum_{t=-\infty}^{\infty} X^{t^2} \sum_{s=-\infty}^{\infty} (2s+1)^2 X^{s^2+s} - \frac{1}{2} \sum_{s=-\infty}^{\infty} X^{s^2+s} \sum_{t=-\infty}^{\infty} (2t)^2 X^{t^2}.
\end{aligned}$$

For $\beta := \sum X^{t^2}$ and $\gamma := \frac{1}{2} \sum X^{s^2+s}$ we have $\gamma + 4X\gamma' = \frac{1}{2} \sum (2s+1)^2 X^{s^2+s}$ and therefore

$$\alpha = \beta(\gamma + 4X\gamma') - 4X\beta'\gamma = \beta\gamma + 4X(\beta\gamma' - \beta'\gamma).$$

Now we apply the infinite product rule to (4.7) and (4.9):

$$\begin{aligned}
\beta' &= \left(\prod_{k=1}^{\infty} (1 - X^{2k})(1 + X^{2k-1})^2 \right)' = \beta \sum_{k=1}^{\infty} \left(2 \frac{(2k-1)X^{2k-2}}{1 + X^{2k-1}} - \frac{2kX^{2k-1}}{1 - X^{2k}} \right) \\
\gamma' &= \left(\prod_{k=1}^{\infty} (1 - X^{2k})(1 + X^{2k})^2 \right)' = \gamma \sum_{k=1}^{\infty} \left(2 \frac{2kX^{2k-1}}{1 + X^{2k}} - \frac{2kX^{2k-1}}{1 - X^{2k}} \right)
\end{aligned}$$

We substitute:

$$\alpha = \beta\gamma \left(1 + 8 \sum_{k=1}^{\infty} \left(\frac{2kX^{2k}}{1 + X^{2k}} - \frac{(2k-1)X^{2k-1}}{1 + X^{2k-1}} \right) \right).$$

Here,

$$\beta\gamma = \prod (1 - X^{2k})^2 (1 + X^{2k-1})^2 (1 + X^{2k})^2 = \prod (1 - X^{2k})^2 (1 + X^k)^2 = \prod (1 - X^{2k})^4 (1 - X^k)^{-2}.$$

After we set this off against α , it remains

$$\left(\sum_{k=-\infty}^{\infty} (-1)^k X^{k^2} \right)^4 \stackrel{(4.8)}{=} \prod_{k=1}^{\infty} \frac{(1 - X^k)^8}{(1 - X^{2k})^4} = \frac{\alpha}{\beta\gamma} = 1 + 8 \sum_{k=1}^{\infty} \left(\frac{2kX^{2k}}{1 + X^{2k}} - \frac{(2k-1)X^{2k-1}}{1 + X^{2k-1}} \right)$$

Finally we replace X by $-X$:

$$\begin{aligned}
\sum q(n)X^n &= \left(\sum_{k=-\infty}^{\infty} X^{k^2} \right)^4 = 1 + 8 \sum_{k=1}^{\infty} \left(\frac{2kX^{2k}}{1 + X^{2k}} + \frac{(2k-1)X^{2k-1}}{1 - X^{2k-1}} \right) \\
&= 1 + 8 \sum_{k=1}^{\infty} \left(\frac{(2k-1)X^{2k-1}}{1 - X^{2k-1}} + \frac{2kX^{2k}}{1 - X^{2k}} - \frac{2kX^{2k}}{1 - X^{2k}} + \frac{2kX^{2k}}{1 + X^{2k}} \right) \\
&= 1 + 8 \sum_{k=1}^{\infty} \left(\frac{kX^k}{1 - X^k} - \frac{4kX^{4k}}{1 - X^{4k}} \right) = 1 + 8 \sum_{4 \nmid k} \frac{kX^k}{1 - X^k} \\
&= 1 + 8 \sum_{4 \nmid k} k \sum_{l=1}^{\infty} X^{kl} = 1 + 8 \sum_{n=1}^{\infty} \sum_{4 \nmid d|n} dX^n. \quad \square
\end{aligned}$$

Example 6.18. For $n = 28$ we obtain

$$\sum_{4 \nmid d | 28} d = 1 + 2 + 7 + 14 = 24.$$

Hence, there are $8 \cdot 24 = 192$ possibilities to express 28 as a sum of four squares. However, they all arise as permutations and sign-choices of

$$28 = 5^2 + 1^2 + 1^2 + 1^2 = 4^2 + 2^2 + 2^2 + 2^2 = 3^2 + 3^2 + 3^2 + 1^2.$$

Theorem 6.17 is best possible in the sense that every integers $n \equiv 7 \pmod{8}$ is not the sum of three squares since $a^2 + b^2 + c^2 \not\equiv 7 \pmod{8}$.

If $n, m \in \mathbb{N}$ are sums of four squares, so is nm by the following identity of Euler (encoding the multiplicativity of the norm in *Hamilton's quaternion* skew field):

$$(a_1^2 + a_2^2 + a_3^2 + a_4^2)(b_1^2 + b_2^2 + b_3^2 + b_4^2) = (a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4)^2 \\ + (a_1b_2 - a_2b_1 + a_3b_4 + a_4b_3)^2 + (a_1b_3 - a_3b_1 + a_4b_2 - a_2b_4)^2 + (a_1b_4 - a_4b_1 + a_2b_3 - a_3b_2)^2$$

This reduces the proof of the first assertion (Lagrange's) of Theorem 6.17 to the case where n is a prime.

Waring's problem ask for the smallest number $g(k)$ such that every positive integer is the sum of $g(k)$ non-negative k -th powers. Hilbert proved that $g(k) < \infty$ for all $k \in \mathbb{N}$. We have $g(1) = 1$, $g(2) = 4$ (Theorem 6.17), $g(3) = 9$, $g(4) = 19$ and in general it is conjectured that

$$g(k) = \left\lfloor \left(\frac{3}{2}\right)^k \right\rfloor + 2^k - 2$$

(see [24]). Curiously, only the numbers $23 = 2 \cdot 2^3 + 7 \cdot 1^3$ and $239 = 2 \cdot 4^3 + 4 \cdot 3^3 + 3 \cdot 1^3$ require nine cubes. It is even conjectured that every sufficiently large integer is a sum of only four non-negative cubes (see [10]).

7 Laurent series

Every integral domain R can be embedded into its *field of fractions* consisting of the formal fractions $\frac{r}{s}$ where $r, s \in R$ and $s \neq 0$. For our ring $K[[X]]$ these fractions have a more convenient shape.

Definition 7.1. A (formal) *Laurent series* in the indeterminant X over the field K is a sum of the form

$$\alpha = \sum_{k=m}^{\infty} a_k X^k$$

where $m \in \mathbb{Z}$ and $a_k \in K$ for $k \geq m$ (i.e. we allow X to negative powers). We often write $\alpha = \sum_{k=-\infty}^{\infty} a_k X^k$ assuming that $\inf(\alpha) = \inf\{k \in \mathbb{Z} : a_k \neq 0\}$ exists. The set of all Laurent series over K is denoted by $K((X))$. Laurent series can be added and multiplied like power series:

$$\alpha + \beta = \sum_{k=-\infty}^{\infty} (a_k + b_k) X^k, \quad \alpha\beta = \sum_{k=-\infty}^{\infty} \left(\sum_{l=-\infty}^{\infty} a_l b_{k-l} \right) X^k$$

(one should check that the inner sum is finite). Moreover, the norm $|\alpha|$ and the derivative α' are defined as for power series.

If a Laurent series is a finite sum, it is naturally called a *Laurent polynomial*. The ring of Laurent polynomials is denoted by $K[X, X^{-1}]$, but plays no role in the following. In analysis one allows double infinite sums, but then the product is no longer well defined as in $\left(\sum_{n=-\infty}^{\infty} X^n\right)^2$.

Theorem 7.2. *The field of fractions of $K[[X]]$ is naturally isomorphic to $K((X))$. In particular, $K((X))$ is a field.*

Proof. Repeating the proof of Lemma 2.2 shows that $K((X))$ is a commutative ring. Let $\alpha \in K((X)) \setminus \{0\}$ and $k := \inf(\alpha)$. By Lemma 2.5, $X^{-k}\alpha \in K[[X]]^\times$. Hence, $X^{-k}(X^{-k}\alpha)^{-1} \in K((X))$ is the inverse of α . This shows that $K((X))$ is a field. By the universal property of the field of fractions $Q(K[[X]])$, the embedding $K[[X]] \subseteq K((X))$ extends to a (unique) field monomorphism $f: Q(K[[X]]) \rightarrow K((X))$. If $k = \inf(\alpha) < 0$, then $f\left(\frac{X^{-k}\alpha}{X^{-k}}\right) = \alpha$ and f is surjective. \square

Of course, we will view $K[[X]]$ as a subring of $K((X))$. In fact, $K[[X]]$ is the *valuation ring* of $K((X))$, i. e. $K[[X]] = \{\alpha \in K((X)) : |\alpha| \leq 1\}$. The field $K((X))$ should not be confused with the field of *rational functions* $K(X)$, which is the field of fractions of $K[X]$.

If $\alpha \in K((X))$ and $\beta \in K[[X]]^\circ$, the substitute $\alpha(\beta)$ is still well-defined and Lemma 3.3 remains correct (α deviates from a power series by only finitely many terms).

Definition 7.3. The (formal) *residue* of $\alpha = \sum a_k X^k \in K((X))$ is defined by $\text{res}(\alpha) := a_{-1}$.

The residue is a K -linear map such that $\text{res}(\alpha') = 0$ for all $\alpha \in K((X))$.

Lemma 7.4. *For $\alpha, \beta \in K((X))$ we have*

$$\begin{aligned} \text{res}(\alpha'\beta) &= -\text{res}(\alpha\beta') \\ \text{res}(\alpha'/\alpha) &= \inf(\alpha) && (\alpha \neq 0) \\ \text{res}(\alpha)\inf(\beta) &= \text{res}(\alpha(\beta)\beta') && (\beta \in (X)) \end{aligned}$$

Proof.

(i) This follows from the product rule

$$0 = \text{res}((\alpha\beta)') = \text{res}(\alpha'\beta) + \text{res}(\alpha\beta').$$

(ii) Let $\alpha = X^k\gamma$ with $k = \inf(\alpha)$ and $\gamma \in K[[X]]^\times$. Then

$$\frac{\alpha'}{\alpha} = \frac{kX^{k-1}\gamma + X^k\gamma'}{X^k\gamma} = kX^{-1} + \gamma'\gamma^{-1}.$$

Since $\gamma^{-1} \in K[[X]]$, it follows that $\text{res}(\alpha'/\alpha) = k = \inf(\alpha)$.

(iii) Since res is a linear map, we may assume that $\alpha = X^k$. If $k \neq -1$, then

$$\text{res}(\alpha(\beta)\beta') = \text{res}(\beta^k\beta') = \text{res}\left(\left(\frac{\beta^{k+1}}{k+1}\right)'\right) = 0 = \text{res}(\alpha) = \text{res}(\alpha)\inf(\beta).$$

If $k = -1$, then

$$\text{res}(\alpha(\beta)\beta') = \text{res}(\beta'/\beta) \stackrel{\text{(ii)}}{=} \inf(\beta) = \text{res}(\alpha)\inf(\beta). \quad \square$$

Theorem 7.5 (LAGRANGE–BÜRMANN’s inversion formula). *The reverse of $\alpha \in \mathbb{C}[[X]]^\circ$ is*

$$\boxed{\sum_{k=1}^{\infty} \frac{\operatorname{res}(\alpha^{-k})}{k} X^k.}$$

Proof. The proof is influenced by [11]. Let $\beta \in \mathbb{C}[[X]]^\circ$ be the reverse of α , i.e. $\alpha(\beta) = X$. From $\alpha \in \mathbb{C}[[X]]^\circ$ we know that $\alpha \neq 0$. In particular, α is invertible in $\mathbb{C}((X))$. By Lemma 3.3, we have $\alpha^{-k}(\beta) = X^{-k}$. Now the coefficient of X^k in β turns out to be

$$\frac{1}{k} \operatorname{res}(kX^{-k-1}\beta) = -\frac{1}{k} \operatorname{res}((X^{-k})'\beta) = \frac{1}{k} \operatorname{res}(X^{-k}\beta') = \frac{1}{k} \operatorname{res}(\alpha^{-k}(\beta)\beta') = \frac{1}{k} \operatorname{res}(\alpha^{-k})$$

by Lemma 7.4. □

Since Theorem 7.5 is actually a statement about power series, it should be mentioned that $\operatorname{res}(\alpha^{-k})$ is just the coefficient of X^{k-1} in the power series $(X/\alpha)^k$. This interpretation will be used in our generalization to higher dimensions in Theorem 9.8.

Example 7.6. Recall from Example 5.1 that the generating function α of the Catalan numbers c_n is the reverse of $X - X^2$. Since

$$\left(\frac{X}{X - X^2}\right)^{n+1} = (1 - X)^{-n-1} \stackrel{(4.1)}{=} \sum_{k=0}^{\infty} \binom{-n-1}{k} (-1)^k X^k,$$

we compute

$$c_{n+1} = \frac{\operatorname{res}(\alpha^{-n-1})}{n+1} = \frac{1}{n+1} (-1)^n \binom{-n-1}{n} = \frac{1}{n+1} \frac{(n+1) \dots 2n}{n!} = \frac{1}{n+1} \binom{2n}{n}.$$

Our next objective is the construction of the algebraic closure of $\mathbb{C}((X))$. We need a well-known tool.

Lemma 7.7 (HENSEL). *Let $R := K[[X]]$. For a polynomial $\alpha = \sum_{k=0}^n a_k Y^k \in R[Y]$ let*

$$\bar{\alpha} := \sum a_k(0) Y^k \in K[Y].$$

Let $\alpha \in R[Y]$ be monic such that $\bar{\alpha} = \alpha_1 \alpha_2$ for some coprime monic polynomials $\alpha_1, \alpha_2 \in K[Y] \setminus K$. Then there exist monic $\beta, \gamma \in R[Y]$ such that $\bar{\beta} = \alpha_1, \bar{\gamma} = \alpha_2$ and $\alpha = \beta\gamma$.

Proof. By hypothesis, $n := \deg(\alpha) = \deg(\alpha_1) + \deg(\alpha_2) \geq 2$. Observe that $\bar{\alpha}$ is essentially the reduction of α modulo the ideal (X) . In particular, the map $R[Y] \rightarrow K[Y], \alpha \mapsto \bar{\alpha}$ is a ring homomorphism. For $\sigma, \tau \in R[Y]$ and $k \in \mathbb{N}$ we write more generally $\sigma \equiv \tau \pmod{(X^k)}$ if all coefficients of $\sigma - \tau$ lie in (X^k) . First choose any monic polynomials $\beta_1, \gamma_1 \in R[Y]$ with $\bar{\beta}_1 = \alpha_1$ and $\bar{\gamma}_1 = \alpha_2$. Then $\deg(\beta_1) = \deg(\alpha_1)$, $\deg(\gamma_1) = \deg(\alpha_2)$ and $\alpha \equiv \beta_1 \gamma_1 \pmod{(X)}$. We construct inductively monic $\beta_k, \gamma_k \in R[Y]$ for $k \geq 2$ such that

- (a) $\beta_k \equiv \beta_{k+1}$ and $\gamma_k \equiv \gamma_{k+1} \pmod{(X^k)}$,
- (b) $\alpha \equiv \beta_k \gamma_k \pmod{(X^k)}$.

Suppose that β_k, γ_k are given. Choose $\delta \in R[Y]$ such that $\alpha = \beta_k \gamma_k + X^k \delta$ and $\deg(\delta) < n$. Since α_1, α_2 are coprime in the Euclidean domain $K[Y]$, there exist $\sigma, \tau \in R[Y]$ such that $\bar{\beta}_k \bar{\sigma} + \bar{\gamma}_k \bar{\tau} = \alpha_1 \bar{\sigma} + \alpha_2 \bar{\tau} = 1$ by Bézout's lemma. Since β_k is monic, we can perform Euclidean division by β_k without leaving $R[Y]$. This yields $\rho, \nu \in R[Y]$ such that $\tau \delta = \beta_k \rho + \nu$ and $\deg(\nu) < \deg(\beta_k)$. Let $d := \deg(\gamma_1)$ and write $\sigma \delta + \gamma_k \rho = \mu + \eta Y^d$ with $\deg(\mu) < d$. Then

$$\beta_{k+1} := \beta_k + X^k \nu, \quad \gamma_{k+1} := \gamma_k + X^k \mu$$

are monic and satisfy (a). Moreover,

$$\delta \equiv (\beta_k \sigma + \gamma_k \tau) \delta \equiv \beta_k (\sigma \delta + \gamma_k \rho) + \gamma_k \nu \equiv \beta_k \mu + \beta_k \eta Y^d + \gamma_k \nu \pmod{(X)}.$$

Since the degrees of $\delta, \beta_k \mu$ and $\gamma_k \nu$ are all smaller than n and $\deg(\beta_k \eta Y^d) \geq n$, it follows that $\bar{\eta} = 0$. Therefore,

$$\beta_{k+1} \gamma_{k+1} \equiv \alpha - X^k \delta + (\beta_k \mu + \gamma_k \nu) X^k \equiv \alpha \pmod{(X^{k+1})},$$

i. e. (b) holds for $k + 1$. This completes the induction.

Let $\beta_k = \sum_{j=0}^e b_{kj} Y^j$ and $\gamma_k = \sum_{j=0}^d c_{kj} Y^j$ with $b_{ij}, c_{ij} \in R$. By construction, $|b_{kj} - b_{k+1,j}| \leq 2^{-k}$ and similarly for c_{kj} . Consequently, $b_j := \lim_k b_{kj}$ and $c_j := \lim_k c_{kj}$ converge in R . We can now define $\beta := \sum_{j=0}^e b_j Y^j$ and $\gamma := \sum_{j=0}^d c_j Y^j$. Then $\bar{\beta} = \bar{\beta}_1 = \alpha_1$ and $\bar{\gamma} = \bar{\gamma}_1 = \alpha_2$. Since $\beta \gamma \equiv \beta_k \gamma_k \equiv \alpha \pmod{(X^k)}$ for every $k \geq 1$, it follows that $\alpha = \beta \gamma$. \square

One can proceed to show that β and γ are uniquely determined in the situation of Lemma 7.7, but we do not need this in the following. Indeed, since R is a principal ideal domain, $R[Y]$ is a factorial ring (also called unique factorization domain) by Gauss' lemma. This means that every monic polynomial in $R[Y]$ has a unique factorization into monic irreducible polynomials. For every irreducible factor ω of $\alpha, \bar{\omega}$ either divides α_1 or α_2 (but not both). This determines the prime factorization of β and γ uniquely.

Example 7.8. Let $n \in \mathbb{N}, a \in (X) \subseteq R := \mathbb{C}[[X]]$ and $\alpha = Y^n - 1 - a \in R[Y]$. Then $\bar{\alpha} = Y^n - 1 = \alpha_1 \alpha_2$ with coprime monic $\alpha_1 = Y - 1$ and $\alpha_2 = Y^{n-1} + \dots + Y + 1$. By Hensel's lemma there exist monic $\beta, \gamma \in R[Y]$ such that $\bar{\beta} = Y - 1, \bar{\gamma} = \alpha_2$ and $\alpha = \beta \gamma$. We may write $\beta = Y - 1 - b$ for some $b \in (X)$. Then $(1 + b)^n = 1 + a$ and the remark after Definition 3.14 implies $1 + b = \sqrt[n]{1 + a}$. The constructive procedure in the proof above must inevitably lead to Newton's binomial theorem $1 + b = \sum_{k=0}^{\infty} \binom{1/n}{k} a^k$.

We have seen that invertible power series in $\mathbb{C}[[X]]$ have arbitrary roots. On the other hand, X does not even have a square root in $\mathbb{C}((X))$. This suggests to allow X not only to negative powers, but also to fractional powers.

Definition 7.9. A *Puiseux series* over K is defined by

$$\sum_{k=m}^{\infty} a_{\frac{k}{n}} X^{\frac{k}{n}},$$

where $m \in \mathbb{Z}, n \in \mathbb{N}$ and $a_{\frac{k}{n}} \in K$ for $k \geq m$. The set of Puiseux series is denoted by $K\{\{X\}\}$. For $\alpha, \beta \in K\{\{X\}\}$ there exists $n \in \mathbb{N}$ such that $\tilde{\alpha} := \alpha(X^n)$ and $\tilde{\beta} := \beta(X^n)$ lie in $K((X))$. We carry over the field operations from $K((X))$ via

$$\alpha + \beta := (\tilde{\alpha} + \tilde{\beta})(X^{\frac{1}{n}}), \quad \alpha \cdot \beta := (\tilde{\alpha} \tilde{\beta})(X^{\frac{1}{n}}).$$

It is straight-forward to check that $(K\{\{X\}\}, +, \cdot)$ is a field. At this point we have established the following inclusions:

$$K \subseteq K[X] \subseteq K[[X]] \subseteq K((X)) \subseteq K\{\{X\}\}.$$

Theorem 7.10 (PUISEUX). *The algebraic closure of $\mathbb{C}((X))$ is $\mathbb{C}\{\{X\}\}$.*

Proof. We follow Nowak [31]. Set $R := \mathbb{C}[[X]]$, $F := \mathbb{C}((X))$ and $\hat{F} := \mathbb{C}\{\{X\}\}$. We show first that \hat{F} is an algebraic field extension of F . Let $\alpha \in \hat{F}$ be arbitrary and $n \in \mathbb{N}$ such that $\beta := \alpha(X^n) \in F$. Let $\zeta \in \mathbb{C}$ be a primitive n -th root of unity. Define

$$\Gamma := \prod_{i=1}^n (Y - \beta(\zeta^i X)) = Y^n + \gamma_1 Y^{n-1} + \dots + \gamma_n \in F[Y].$$

Replacing X by ζX permutes the factors $Y - \beta(\zeta^i X)$ and thus leaves Γ invariant. Consequently, $\gamma_i(\zeta X) = \gamma_i$ for $i = 1, \dots, n$. This means that there exist $\tilde{\gamma}_i \in F$ such that $\gamma_i = \tilde{\gamma}_i(X^n)$. Now let

$$\tilde{\Gamma} := Y^n + \tilde{\gamma}_1 Y^{n-1} + \dots + \tilde{\gamma}_n \in F[Y].$$

Substituting X by X^n in $\tilde{\Gamma}(\alpha)$ gives $\Gamma(\beta) = 0$. Thus, also $\tilde{\Gamma}(\alpha) = 0$. This shows that α is algebraic over F and \hat{F} is an algebraic extension of F .

Now we prove that \hat{F} is algebraically closed. Let $\Gamma = Y^n + \gamma_1 Y^{n-1} + \dots + \gamma_n \in \hat{F}[Y]$ be arbitrary with $n \geq 2$. We need to show that Γ has a root in \hat{F} . Without loss of generality, $\Gamma \neq Y^n$. After applying the *Tschirnhaus transformation* $Y \mapsto Y - \frac{1}{n}\gamma_1$, we may assume that $\gamma_1 = 0$. Let

$$r := \min \left\{ \frac{1}{k} \inf(\gamma_k) : k = 1, \dots, n \right\} \in \mathbb{Q}$$

and $m \in \mathbb{N}$ such that $\gamma_k(X^m) \in F$ for $k = 1, \dots, n$ and $r = \frac{s}{m}$ for some $s \in \mathbb{Z}$. Define $\delta_0 := 1$ and $\delta_k := \gamma_k(X^m)X^{-ks} \in F$ for $k = 1, \dots, n$. Since

$$\inf(\delta_k) = m \inf(\gamma_k) - ks = m(\inf(\gamma_k) - kr) \geq 0,$$

$\Delta := Y^n + \delta_2 Y^{n-2} + \dots + \delta_n \in R[Y]$. Consider $\bar{\Delta} := Y^n + \delta_2(0)Y^{n-2} + \dots + \delta_n(0) \in \mathbb{C}[Y]$. Since $\inf(\delta_k) = 0$ for at least one $k \geq 1$, we have $\bar{\Delta} \neq Y^n$. Since $\delta_1 = 0$, also $\bar{\Delta} \neq (Y - c)^n$ for all $c \in \mathbb{C}$. Using that $\mathbb{C}[Y]$ is algebraically closed, we can decompose $\bar{\Delta} = \bar{\Delta}_1 \bar{\Delta}_2$ with coprime monic polynomials $\bar{\Delta}_1, \bar{\Delta}_2 \in \mathbb{C}[Y]$ of degree $< n$. By Hensel's lemma, there exists a corresponding factorization $\Delta = \Delta_1 \Delta_2$ with $\Delta_1, \Delta_2 \in R[Y]$. Finally, replace X by $X^{\frac{1}{m}}$ in Δ_i to obtain $\Gamma_i \in \hat{F}[Y]$. Then

$$\Gamma = X^{nr} \sum_{k=0}^n \gamma_k X^{-kr} (Y X^{-r})^{n-k} = X^{nr} \sum_{k=0}^n \delta_k (X^{\frac{1}{m}}) (Y X^{-r})^{n-k} = X^{nr} \Gamma_1(Y X^{-r}) \Gamma_2(Y X^{-r}).$$

Induction on n shows that Γ has a root and \hat{F} is algebraically closed. □

8 Multivariate power series

In Remark 4.7 and even more in the last two results it became clear that power series in more than one indeterminate make sense. We give proper definitions now.

Definition 8.1.

- (i) The ring of formal power series in n indeterminants X_1, \dots, X_n over a field K is defined inductively via

$$K[[X_1, \dots, X_n]] := K[[X_1, \dots, X_{n-1}]][[X_n]].$$

Its elements have the form

$$\alpha = \sum_{k_1, \dots, k_n \geq 0} a_{k_1, \dots, k_n} X_1^{k_1} \dots X_n^{k_n}$$

where $a_{k_1, \dots, k_n} \in K$. We (still) call $a_{0, \dots, 0}$ the *constant term* of α . Let $\inf(\alpha) := \inf\{k_1 + \dots + k_n : a_{k_1, \dots, k_n} \neq 0\}$ and

$$|\alpha| := 2^{-\inf(\alpha)}.$$

- (ii) If all but finitely many coefficients of α are zero, we call α a (formal) polynomial in X_1, \dots, X_n . In this case,

$$\deg(\alpha) := \sup\{k_1 + \dots + k_n : a_{k_1, \dots, k_n} \neq 0\}$$

is the *degree* of α , where $\deg(0) = \sup \emptyset = -\infty$. Moreover, a polynomial α is called *homogeneous* if all monomials occurring in α (with non-zero coefficient) have the same degree. The set of polynomials is denoted by $K[X_1, \dots, X_n]$.

Once we have convinced ourselves that Lemma 2.2 remains true when K is replaced by an integral domain, it becomes evident that also $K[[X_1, \dots, X_n]]$ is an integral domain. Likewise the norm still gives rise to a complete ultrametric (to prove $|\alpha\beta| = |\alpha||\beta|$ one may assume that α and β are homogeneous polynomials) and the crucial Lemma 2.10 holds in $K[[X_1, \dots, X_n]]$ too. We stress that this metric is finer than the one induced from $K[[X_1, \dots, X_{n-1}]]$ as, for example, $\lim_{k \rightarrow \infty} X_1^k X_2$ converges in the former, but not in the latter (with $n = 2$). Moreover, a power series α is invertible in $K[[X_1, \dots, X_n]]$ if and only if its constant term is non-zero. Indeed, after scaling, the constant term is 1 and

$$\alpha^{-1} = \frac{1}{1 - (1 - \alpha)} = \sum_{k=0}^{\infty} (1 - \alpha)^k$$

converges.

Unlike $K[X]$ or $K[[X]]$, the multivariate rings are not principal ideal domains, but one can show that they are still factorial (see [25, Theorem IV.9.3]). We do not require this fact in the sequel.

The degree function equips $K[X_1, \dots, X_n]$ with a *grading*, i. e. we have

$$K[X_1, \dots, X_n] = \bigoplus_{d=0}^{\infty} P_d$$

and $P_d P_e \subseteq P_{d+e}$ where P_d denotes the set of homogeneous polynomials of degree d . In the following we will restrict ourselves mostly to polynomials of a special type. Note that if $\alpha, \beta_1, \dots, \beta_n \in K[X_1, \dots, X_n]$, we can substitute X_i by β_i in α to obtain $\alpha(\beta_1, \dots, \beta_n) \in K[X_1, \dots, X_n]$. It is important that these substitutions happen simultaneously and not one after the other (more about this at the end of the section).

Definition 8.2. A polynomial $\alpha \in K[X_1, \dots, X_n]$ is called *symmetric* if

$$\alpha(X_{\pi(1)}, \dots, X_{\pi(n)}) = \alpha(X_1, \dots, X_n)$$

for all permutations $\pi \in S_n$.

It is easy to see that the symmetric polynomials form a subring of $K[X_1, \dots, X_n]$.

Example 8.3.

(i) The *elementary symmetric polynomials* are $\sigma_0 := 1$ and

$$\sigma_k := \sum_{1 \leq i_1 < \dots < i_k \leq n} X_{i_1} \dots X_{i_k} \quad (k \geq 1).$$

Note that $\sigma_k = 0$ for $k > n$ (empty sum).

(ii) The *complete symmetric polynomials* are $\tau_0 := 1$ and

$$\tau_k := \sum_{1 \leq i_1 \leq \dots \leq i_k \leq n} X_{i_1} \dots X_{i_k} \quad (k \geq 1).$$

(iii) The *power sum polynomials* are $\rho_k := X_1^k + \dots + X_n^k$ for $k \geq 0$.

Keep in mind that σ_k , τ_k and ρ_k depend on n . All three sets of polynomials are homogeneous. The elementary and complete symmetric polynomials are special instances of *Schur polynomials*, which we do not attempt to define here.

Theorem 8.4 (VIETA). *The following identities hold in $K[[X_1, \dots, X_n, Y]]$:*

$$\prod_{k=1}^n (1 + X_k Y) = \sum_{k=0}^n \sigma_k Y^k, \tag{8.1}$$

$$\prod_{k=1}^n \frac{1}{1 - X_k Y} = \sum_{k=0}^{\infty} \tau_k Y^k. \tag{8.2}$$

Proof. The first equation is only a matter of expanding the product. The second equation follows from

$$\prod_{k=1}^n \frac{1}{1 - X_k Y} = \prod_{k=1}^n \sum_{l=0}^{\infty} (X_k Y)^l = \sum_{k=0}^{\infty} \left(\sum_{l_1 + \dots + l_n = k} X_1^{l_1} \dots X_n^{l_n} \right) Y^k = \sum_{k=0}^{\infty} \tau_k Y^k. \quad \square$$

When we specialize $X_1 = \dots = X_n = 1$ in Vieta's theorem (as we may), we recover the generating functions of the binomial coefficients and the multiset counting coefficients in Example 5.1. When we substitute $X_k = k$ for $k = 1, \dots, n$, we obtain a new formula for the Stirling numbers by virtue of Theorem 6.9.

It is easy to see that the grading by degree carries over to symmetric polynomials. The following theorem shows that the elementary symmetric polynomials are the building blocks of all symmetric polynomials.

Theorem 8.5 (Fundamental theorem on symmetric polynomials). *For every symmetric polynomial $\alpha \in K[X_1, \dots, X_n]$ there exists a unique $\gamma \in K[X_1, \dots, X_n]$ such that $\alpha = \gamma(\sigma_1, \dots, \sigma_n)$.*

Proof. We first prove the *existence* of γ : Without loss of generality, let

$$\alpha = \sum_{i_1, \dots, i_n} a_{i_1, \dots, i_n} X_1^{i_1} \dots X_n^{i_n} \neq 0.$$

We order the tuples (i_1, \dots, i_n) lexicographically and argue by induction on

$$f(\alpha) := \max\{(i_1, \dots, i_n) : a_{i_1, \dots, i_n} \neq 0\}$$

(see Example 8.6 below for an illustration). If $f(\alpha) = (0, \dots, 0)$, then $\gamma := \alpha = a_{0, \dots, 0} \in K$. Now let $f(\alpha) = (d_1, \dots, d_n) > (0, \dots, 0)$. Since $\alpha = \alpha(X_{\pi(1)}, \dots, X_{\pi(n)})$ for all $\pi \in S_n$, $d_1 \geq \dots \geq d_n$. Let

$$\beta := a_{d_1, \dots, d_n} \sigma_1^{d_1 - d_2} \sigma_2^{d_2 - d_3} \dots \sigma_{n-1}^{d_{n-1} - d_n} \sigma_n^{d_n}.$$

Then we have $f(\sigma_k^{d_k - d_{k+1}}) = (d_k - d_{k+1})f(\sigma_k) = (d_k - d_{k+1}, \dots, d_k - d_{k+1}, 0, \dots, 0)$ and

$$f(\beta) = f(\sigma_1^{d_1 - d_2}) + \dots + f(\sigma_n^{d_n}) = (d_1, \dots, d_n).$$

Hence, the symmetric polynomial $\alpha - \beta$ satisfies $f(\alpha - \beta) < (d_1, \dots, d_n)$ and the existence of γ follows by induction.

Now we show the *uniqueness* of γ : Let $\gamma, \delta \in K[X_1, \dots, X_n]$ such that $\gamma(\sigma_1, \dots, \sigma_n) = \delta(\sigma_1, \dots, \sigma_n)$. For $\rho := \gamma - \delta$ it follows that $\rho(\sigma_1, \dots, \sigma_n) = 0$. We have to show that $\rho = 0$. By way of contradiction, suppose $\rho \neq 0$. Let $d_1 \geq \dots \geq d_n$ be the lexicographically largest n -tuple such that the coefficient of $X_1^{d_1 - d_2} X_2^{d_2 - d_3} \dots X_n^{d_n}$ in ρ is non-zero. As above, $f(\sigma_1^{d_1 - d_2} \dots \sigma_n^{d_n}) = (d_1, \dots, d_n)$. For every other summand $X_1^{e_1 - e_2} \dots X_n^{e_n}$ of ρ we obtain $f(\sigma_1^{e_1 - e_2} \dots \sigma_n^{e_n}) < (d_1, \dots, d_n)$. This yields $f(\rho(\sigma_1, \dots, \sigma_n)) = (d_1, \dots, d_n)$ in contradiction to $\rho(\sigma_1, \dots, \sigma_n) = 0$. \square

Example 8.6. Consider $\alpha = XY^3 + X^3Y - X - Y \in K[X, Y]$. With the notation from the proof above, $f(\alpha) = (3, 1)$ and

$$\beta := \sigma_1^2 \sigma_2 = (X + Y)^2 XY = X^3Y + 2X^2Y^2 + XY^3.$$

Thus, $\alpha - \beta = -2X^2Y^2 - X - Y$. In the next step we have $f(\alpha - \beta) = (2, 2)$ and

$$\beta_2 := -2\sigma_2^2 = -2X^2Y^2.$$

It remains: $\alpha - \beta - \beta_2 = -X - Y = -\sigma_1$. Finally,

$$\alpha = \beta + \beta_2 - \sigma_1 = \sigma_1^2 \sigma_2 - 2\sigma_2^2 - \sigma_1 = \gamma(\sigma_1, \sigma_2)$$

where $\gamma = X^2Y - 2Y^2 - X$.

From an algebraic point of view, Theorem 8.5 (applied to $\alpha = 0$) states that the elementary symmetric polynomials $\sigma_1, \dots, \sigma_n$ are algebraically independent over K , so they form a transcendence basis of $K(X_1, \dots, X_n)$ (recall that $K(X_1, \dots, X_n)$ has transcendence degree n). The identities in the next theorem express the σ_i recursively in terms of the τ_j and in terms of the ρ_j . So the latter sets of symmetric polynomials form transcendence bases too. It is no coincidence that $\deg(\sigma_k) = \deg(\tau_k) = \deg(\rho_k) = k$ for $k \leq n$. A theorem from invariant theory (in characteristic 0) implies that any algebraically independent, symmetric, homogeneous polynomials $\lambda_1, \dots, \lambda_n$ have degrees $1, \dots, n$ in some order (see [18, Proposition 3.7]).

Theorem 8.7 (GIRARD–NEWTON identities). *The following identities hold in $K[X_1, \dots, X_n]$ for all $n, k \in \mathbb{N}$:*

$$\boxed{\begin{aligned} \sum_{i=0}^k (-1)^i \sigma_i \tau_{k-i} &= 0, \\ \sum_{i=1}^k \rho_i \tau_{k-i} &= k \tau_k, \\ \sum_{i=1}^k (-1)^i \sigma_{k-i} \rho_i &= -k \sigma_k. \end{aligned}}$$

Proof. Let $\sigma = \sum (-1)^k \sigma_k Y^k = \prod (1 - X_k Y)$ and $\tau := \sum \tau_k Y^k = \prod \frac{1}{1 - X_k Y}$ as in Vieta's theorem.

(i) The claim follows by comparing coefficients of Y^k in

$$1 = \sigma \tau = \sum_{k=0}^{\infty} \left(\sum_{i=0}^k (-1)^i \sigma_i \tau_{k-i} \right) Y^k.$$

(ii) We differentiate with respect to Y using the product rule while noticing that $\left(\frac{1}{1 - X_k Y} \right)' = \frac{X_k}{(1 - X_k Y)^2}$:

$$\begin{aligned} \sum_{k=1}^{\infty} k \tau_k Y^k &= Y \tau' = \tau \sum_{k=1}^n \frac{X_k Y}{1 - X_k Y} = \tau \sum_{k=1}^n \sum_{i=1}^{\infty} (X_k Y)^i \\ &= \tau \sum_{i=1}^{\infty} \rho_i Y^i = \sum_{k=1}^{\infty} \left(\sum_{i=1}^k \rho_i \tau_{k-i} \right) Y^k. \end{aligned}$$

(iii) We differentiate again with respect to Y (this idea is often attributed to [6, p. 212]):

$$-\sum_{k=0}^{\infty} (-1)^k k \sigma_k Y^k = -Y \sigma' = \sigma \sum_{k=1}^n \frac{X_k Y}{1 - X_k Y} = \sigma \sum_{k=1}^{\infty} \rho_k Y^k = \sum_{k=1}^{\infty} \left(\sum_{i=1}^k (-1)^{k-i} \sigma_{k-i} \rho_i \right) Y^k. \quad \square$$

Now that we know that each of the σ_i , τ_i and ρ_i can be expressed by the other two sets of polynomials, it is natural to ask for explicit formulas. This is achieved by Waring's formula. Here $P(n)$ stands for the set of partitions of n as introduced in Definition 5.2.

Theorem 8.8 (WARING's formula). *The following holds in $\mathbb{C}[X_1, \dots, X_n]$ for all $n, k \in \mathbb{N}$:*

$$\boxed{\begin{aligned} \rho_k &= (-1)^k k \sum_{(1^{a_1}, \dots, k^{a_k}) \in P(k)} (-1)^{a_1 + \dots + a_k} \frac{(a_1 + \dots + a_k - 1)!}{a_1! \dots a_k!} \sigma_1^{a_1} \dots \sigma_k^{a_k}, \\ &= -k \sum_{(1^{a_1}, \dots, k^{a_k}) \in P(k)} (-1)^{a_1 + \dots + a_k} \frac{(a_1 + \dots + a_k - 1)!}{a_1! \dots a_k!} \tau_1^{a_1} \dots \tau_k^{a_k}. \end{aligned}}$$

Proof. We introduce a new variable Y and compute in $\mathbb{C}[[X_1, \dots, X_n, Y]]$. The generating function of $(-1)^k \frac{\rho_k}{k}$ is

$$\begin{aligned} \sum_{k=1}^{\infty} (-1)^k \frac{\rho_k}{k} Y^k &= - \sum_{i=1}^n \sum_{k=1}^{\infty} (-1)^{k-1} \frac{(X_i Y)^k}{k} = - \sum_{i=1}^n \log(1 + X_i Y) \stackrel{(3.6)}{=} - \log \left(\prod_{i=1}^n (1 + X_i Y) \right) \\ &\stackrel{(8.1)}{=} - \log \left(1 + \sum_{i=1}^n \sigma_i Y^i \right) = \sum_{l=1}^{\infty} \frac{(-1)^l}{l} \left(\sum_{i=1}^n \sigma_i Y^i \right)^l. \end{aligned}$$

Now we use the multinomial theorem to expand the inner sum:

$$\begin{aligned} \sum_{k=1}^{\infty} (-1)^k \frac{\rho_k}{k} Y^k &= \sum_{l=1}^{\infty} \frac{(-1)^l}{l} \sum_{a_1 + \dots + a_n = l} \frac{l!}{a_1! \dots a_n!} \sigma_1^{a_1} \dots \sigma_n^{a_n} Y^{a_1 + 2a_2 + \dots + na_n} \\ &= \sum_{k=1}^{\infty} \sum_{(1^{a_1}, \dots, k^{a_k}) \in P(k)} (-1)^{a_1 + \dots + a_k} \frac{(a_1 + \dots + a_k - 1)!}{a_1! \dots a_k!} \sigma_1^{a_1} \dots \sigma_k^{a_k} Y^k. \end{aligned}$$

Note that $\sigma_k = 0$ for $k > n$. This implies the first equation. For the second we start similarly:

$$\sum_{k=1}^{\infty} \frac{\rho_k}{k} Y^k = \sum_{i=1}^n \sum_{k=1}^{\infty} \frac{(X_i Y)^k}{k} = \sum_{i=1}^n \log((1 - X_i Y)^{-1}) = \log \left(\prod_{i=1}^n \frac{1}{1 - X_i Y} \right) \stackrel{(8.2)}{=} \log \left(1 + \sum_{i=1}^{\infty} \tau_i Y^i \right).$$

Since we are only interested in the coefficient of X^k , we can truncate the sum to

$$\log \left(1 + \sum_{i=1}^k \tau_i Y^i \right) = - \sum_{l=1}^{\infty} \frac{(-1)^l}{l} \sum_{a_1 + \dots + a_k = l} \frac{l!}{a_1! \dots a_k!} \tau_1^{a_1} \dots \tau_k^{a_k} Y^{a_1 + 2a_2 + \dots + ka_k}$$

and argue as before. □

The first instances of Waring's formula are

$$\rho_1 = \sigma_1, \quad \rho_2 = \sigma_1^2 - 2\sigma_2, \quad \rho_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3.$$

Example 8.9. Since we are dealing with polynomials, it is legitimate to replace the indeterminants by actual numbers. Let $x, y, z \in \mathbb{C}$ be the roots of

$$\alpha = X^3 + 2X^2 - 3X + 1 \in \mathbb{C}[X]$$

(guaranteed to exist by the fundamental theorem of algebra). By Vieta's theorem,

$$\sigma_1(x, y, z) = -2, \quad \sigma_2(x, y, z) = -3, \quad \sigma_3(x, y, z) = -1.$$

We compute with the first Waring formula

$$x^3 + y^3 + z^3 = \rho_3(x, y, z) = (-2)^3 - 3(-2)(-3) + 3(-1) = -29$$

without knowing what x, y, z are! Here is an alternative approach for those who like matrices. The companion matrix

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 3 \\ 0 & 1 & -2 \end{pmatrix}$$

of α has characteristic polynomial α . Hence, the eigenvalues of A^k are x^k , y^k and z^k . This shows $\rho_k(x, y, z) = \text{tr}(A^k)$.

We invite the reader to prove the other four transition formulas.

Exercise 8.10. Show that the following holds in $\mathbb{C}[X_1, \dots, X_n]$ for all $n, k \in \mathbb{N}$:

$$\begin{aligned} \sigma_k &= (-1)^k \sum_{(1^{a_1}, \dots, k^{a_k}) \in P(k)} (-1)^{a_1 + \dots + a_k} \frac{(a_1 + \dots + a_k)!}{a_1! \dots a_k!} \tau_1^{a_1} \dots \tau_k^{a_k}, \\ &= (-1)^k \sum_{(1^{a_1}, \dots, k^{a_k}) \in P(k)} \frac{(-1)^{a_1 + \dots + a_k}}{1^{a_1} a_1! \dots k^{a_k} a_k!} \rho_1^{a_1} \dots \rho_k^{a_k}, \end{aligned} \quad (8.3)$$

$$\begin{aligned} \tau_k &= (-1)^k \sum_{(1^{a_1}, \dots, k^{a_k}) \in P(k)} (-1)^{a_1 + \dots + a_k} \frac{(a_1 + \dots + a_k)!}{a_1! \dots a_k!} \sigma_1^{a_1} \dots \sigma_k^{a_k}, \\ &= \sum_{(1^{a_1}, \dots, k^{a_k}) \in P(k)} \frac{1}{1^{a_1} a_1! \dots k^{a_k} a_k!} \rho_1^{a_1} \dots \rho_k^{a_k}. \end{aligned} \quad (8.4)$$

Hint: For (8.3) and (8.4), mimic the proof of Theorem 6.12 (these are specializations of *Frobenius' formula* on Schur polynomials).

Exercise 8.11. Use Exercise 8.10 to solve the non-linear system

$$\begin{aligned} x + y + z &= 3, \\ x^2 + y^2 + z^2 &= 15, \\ x^3 + y^3 + z^3 &= 45. \end{aligned}$$

Hint: As the solution is too complicated to guess, look up *Cardano's formula*.

We leave polynomials to fully develop multivariate power series.

Definition 8.12. For $\alpha \in K[[X_1, \dots, X_n]]$ and $1 \leq i \leq n$ let $\partial_i \alpha$ be the i -th *partial derivative* with respect to X_i , i. e. we regard α as a power series in X_i with coefficients in $K[[X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n]]$ and form the usual (formal) derivative. For $k \in \mathbb{N}_0$ let $\partial_i^k \alpha$ be the k -th derivative with respect to X_i .

Note that ∂_i is a linear operator, which commutes with all ∂_j (*Schwarz' theorem*). Indeed, by linearity it suffices to check

$$\partial_i \partial_j (X_i^k X_j^l) = \partial_i (l X_i^k X_j^{l-1}) = kl X_i^{k-1} X_j^{l-1} = \partial_j (k X_i^{k-1} X_j^l) = \partial_j \partial_i (X_i^k X_j^l).$$

We need a fairly general form of the product rule.

Lemma 8.13 (LEIBNIZ' rule). *Let $\alpha_1, \dots, \alpha_s \in \mathbb{C}[[X_1, \dots, X_n]]$ and $k_1, \dots, k_n \in \mathbb{N}_0$. Then*

$$\partial_1^{k_1} \dots \partial_n^{k_n} (\alpha_1 \dots \alpha_s) = \sum_{l_{11} + \dots + l_{1s} = k_1} \dots \sum_{l_{n1} + \dots + l_{ns} = k_n} \frac{k_1! \dots k_n!}{\prod_{i,j} l_{ij}!} \prod_{t=1}^s \partial_1^{l_{1t}} \dots \partial_n^{l_{nt}} \alpha_t.$$

Proof. For $n = 1$ the claim is more or less equivalent to the familiar multinomial theorem

$$(a_1 + \dots + a_s)^k = \sum_{l_1 + \dots + l_s = k} \frac{k!}{l_1! \dots l_s!} a_1^{l_1} \dots a_s^{l_s},$$

where a_1, \dots, a_s lie in any commutative ring. With every new indeterminate we simply apply the case $n = 1$ to the formula for $n - 1$. In this way the multinomial coefficients are getting multiplied. \square

Our next goal is the multivariate chain rule for (higher) derivatives. We equip $\mathbb{C}[[X_1, \dots, X_n]]^n$ with the direct product ring structure and use the shorthand notation $\alpha := (\alpha_1, \dots, \alpha_n)$ and $0 := (0, \dots, 0)$. Write

$$\alpha \circ \beta := (\alpha_1(\beta_1, \dots, \beta_n), \dots, \alpha_n(\beta_1, \dots, \beta_n))$$

provided this is well-defined. It is not difficult to show that

$$\begin{aligned} (\alpha + \beta) \circ \gamma &= (\alpha \circ \gamma) + (\beta \circ \gamma), \\ (\alpha \cdot \beta) \circ \gamma &= (\alpha \circ \gamma) \cdot (\beta \circ \gamma) \end{aligned} \tag{8.5}$$

as in Lemma 3.3. It was remarked by M. Hardy [13] that Leibniz' rule as well as the chain rule become slightly more transparent when we give up on counting multiplicities of derivatives as follows.

Theorem 8.14 (FAÁ DI BRUNO'S rule). *Let $\alpha, \beta_1, \dots, \beta_n \in K[[X_1, \dots, X_n]]$ such that $\alpha(\beta_1, \dots, \beta_n)$ is defined. Then for $1 \leq k_1, \dots, k_s \leq n$ we have*

$$\partial_{k_1} \dots \partial_{k_s} (\alpha(\beta_1, \dots, \beta_n)) = \sum_{t=1}^s \sum_{1 \leq i_1, \dots, i_t \leq n} \sum_{\substack{A_1 \dot{\cup} \dots \dot{\cup} A_t \\ = \{1, \dots, s\}}} (\partial_{A_1} \beta_{i_1}) \dots (\partial_{A_t} \beta_{i_t}) (\partial_{i_1} \dots \partial_{i_t} \alpha)(\beta_1, \dots, \beta_n),$$

where $A_1 \dot{\cup} \dots \dot{\cup} A_t$ runs through the set partitions of s and $\partial_{A_t} := \prod_{a \in A_t} \partial_{k_a}$.

Proof. By (8.5), we may assume that $\alpha = X_1^{a_1} \dots X_n^{a_n}$. Then by the product rule,

$$\partial_k (\alpha(\beta_1, \dots, \beta_n)) = \sum_{i=1}^n (\partial_k \beta_i) a_i \beta_1^{a_1} \dots \beta_i^{a_i-1} \dots \beta_n^{a_n} = \sum_{i=1}^n (\partial_k \beta_i) (\partial_{i_1} \dots \partial_{i_t} \alpha)(\beta_1, \dots, \beta_n). \tag{8.6}$$

This settles the case $s = 1$. Now assume that the claim for some s is established. When we apply some $\partial_{k_{s+1}}$ on the right hand side of the induction hypothesis, we need the product rule again. There are two cases: either $s+1$ is added to some of the existing sets A_t or $\partial_{k_{s+1}}$ is applied to $(\partial_{i_1} \dots \partial_{i_t} \alpha)(\beta_1, \dots, \beta_n)$. In the latter case s increases to $s+1$, $A_{s+1} = \{s+1\}$ and i_{s+1} is introduced as in (8.6). \square

Example 8.15. For $n = 1$ and $K = \mathbb{C}$, Theorem 8.14 ‘‘simplifies’’ to

$$\begin{aligned} (\alpha(\beta))^{(s)} &= \sum_{t=1}^s \sum_{A_1 \dot{\cup} \dots \dot{\cup} A_t} \beta^{(|A_1|)} \dots \beta^{(|A_t|)} \alpha^{(t)}(\beta) \\ &= \sum_{(1^{a_1}, \dots, s^{a_s}) \in P(s)} \frac{s!}{(1!)^{a_1} \dots (s!)^{a_s} a_1! \dots a_s!} (\beta')^{a_1} \dots (\beta^{(s)})^{a_s} (\alpha(\beta))^{(a_1 + \dots + a_s)}, \end{aligned}$$

where $(1^{a_1}, \dots, s^{a_s})$ runs over the partitions of s and the coefficient is explained just as in Lemma 6.11.

9 MacMahon's master theorem

In this final section we enter a non-commutative world by making use of matrices. The ultimate goal is the *master theorem* found and named by MacMahon [27, Chapter II]. Since $K[[X_1, \dots, X_n]]$ can be embedded in its field of fractions, the familiar rules of linear algebra (over fields) remain valid in

the ring $K[[X_1, \dots, X_n]]^{n \times n}$ of $n \times n$ -matrices with coefficients in $K[[X_1, \dots, X_n]]$. In particular, the determinant of $A = (\alpha_{ij})_{i,j}$ can be defined by *Leibniz' formula* (not rule)

$$\det(A) := \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)}.$$

It follows that $\det(A(0)) = \det(A)(0)$ by (8.5). Recall that the *adjoint* of A is defined by $\operatorname{adj}(A) := ((-1)^{i+j} \det(A_{ji}))_{i,j}$ where A_{ji} is obtained from A by deleting the j -th row and i -th column. Then

$$A \operatorname{adj}(A) = \operatorname{adj}(A)A = \det(A)1_n,$$

where 1_n denotes the identity $n \times n$ -matrix. This shows that A is invertible if and only if $\det(A)$ is invertible in $K[[X_1, \dots, X_n]]$, i. e. $\det(A)$ has a non-zero constant term. Expanding the entries of A as $\alpha_{ij} = \sum_{k_1, \dots, k_n} a_{k_1, \dots, k_n}^{(i,j)} X_1^{k_1} \cdots X_n^{k_n}$ gives rise to a natural bijection

$$\begin{aligned} \Omega: K[[X_1, \dots, X_n]]^{n \times n} &\rightarrow K^{n \times n}[[X_1, \dots, X_n]], \\ A &\mapsto \sum_{k_1, \dots, k_n} (a_{k_1, \dots, k_n}^{(i,j)})_{i,j} X_1^{k_1} \cdots X_n^{k_n}. \end{aligned}$$

Clearly, Ω is a vector space isomorphism. To verify that it is even a ring isomorphism, it is enough to consider matrices A, B with only one non-zero entry each. But then $AB = 0$ or AB is just the multiplication in $K[[X_1, \dots, X_n]]$. So we can now freely pass from one ring to the other, keeping in mind that we are dealing with power series with non-commuting coefficients! Allowing some flexibility, we can also expand $A = \sum_i A_i X_k^i$ where k is fixed and $A_i \in K[[X_1, \dots, X_{k-1}, X_{k+1}, \dots, X_n]]^{n \times n}$. This suggests to define

$$\partial_k A := \sum_{i=1}^{\infty} i A_i X_k^{i-1} = (\partial_k \alpha_{ij})_{i,j}.$$

The sum and product differentiation rules remain correct, but the power rule $\partial_k(A^s) = s \partial_k(A) A^{s-1}$ (and in turn Leibniz' rule) does not hold in general, since A might not commute with $\partial_k A$.

The next two results are just a warm-up and are not needed later on.

Lemma 9.1. *Let $A \in \mathbb{C}[[X_1, \dots, X_n]]^{n \times n}$ and $1 \leq k \leq n$. Then $\partial_k \det(A) = \operatorname{tr}(\operatorname{adj}(A) \partial_k A)$.*

Proof. Write $A = (\alpha_{ij})$. By Leibniz' formula and the product rule, it follows that

$$\begin{aligned} \partial_k \det(A) &= \partial_k \left(\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)} \right) \\ &= \sum_{i=1}^n \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \alpha_{1\sigma(1)} \cdots \partial_k(\alpha_{i\sigma(i)}) \cdots \alpha_{n\sigma(n)} \\ &= \sum_{i=1}^n \sum_{j=1}^n \sum_{\substack{\sigma \in S_n \\ \sigma(j)=i}} \operatorname{sgn}(\sigma) \alpha_{1\sigma(1)} \cdots \partial_k(\alpha_{ji}) \cdots \alpha_{n\sigma(n)}. \end{aligned}$$

The permutations $\sigma \in S_n$ with $\sigma(j) = i$ correspond naturally to

$$\tau := (i, i+1, \dots, n)^{-1} \sigma(j, j+1, \dots, n) \in S_{n-1}$$

with $\text{sgn}(\tau) = (-1)^{i+j} \text{sgn}(\sigma)$. Hence, Leibniz' formula applied to $\det(A_{ji})$ gives

$$\sum_{j=1}^n \sum_{\substack{\sigma \in S_n \\ \sigma(j)=i}} \text{sgn}(\sigma) \alpha_{1\sigma(1)} \dots \partial_k(\alpha_{ji}) \dots \alpha_{n\sigma(n)} = \sum_{j=1}^n (-1)^{i+j} \det(A_{ji}) \partial_k(\alpha_{ji}).$$

Since this is the entry of $\text{adj}(A) \partial_k A$ at position (i, i) , the claim follows. \square

If $A \in \mathbb{C}^{n \times n}[[X_1, \dots, X_n]]$ has zero constant term, then $\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}$ converges and is even invertible since it has constant term 1_n .

Theorem 9.2 (JACOBI's determinant formula). *Let $A \in \mathbb{C}^{n \times n}[[X_1, \dots, X_n]]$ with zero constant term. Then*

$$\boxed{\det(\exp(A)) = \exp(\text{tr}(A)).}$$

Proof. We introduce a new variable Y and consider $B := \exp(AY)$. Denoting the derivative with respect to Y by $'$, we have

$$B' = \left(\sum_{k=0}^{\infty} \frac{A^k}{k!} Y^k \right)' = \sum_{k=1}^{\infty} \frac{A^k}{(k-1)!} Y^{k-1} = AB.$$

Invoking Lemma 9.1 and using that B is invertible, we compute:

$$\det(B)' = \text{tr}(\text{adj}(B)B') = \det(B) \text{tr}(B^{-1}AB) = \det(B) \text{tr}(A).$$

This is a differential equation, which can be solved as follows. Write $\det(B) = \sum_{k=0}^{\infty} B_k Y^k$ with $B_k \in \mathbb{C}[[X_1, \dots, X_n]]$. Then $B_0 = \det(B(0)) = \det(\exp(0)1_n) = \det(1_n) = 1$ and $B_{k+1} = \frac{1}{k+1} \text{tr}(A)B_k$ for $k \geq 0$. This yields

$$\det(B) = 1 + \text{tr}(A)Y + \frac{\text{tr}(A)^2}{2}Y^2 + \dots = \exp(\text{tr}(A)Y).$$

Since we already know that $\exp(A)$ converges, we are allowed to specialize $Y = 1$ in B , from which the claim follows. \square

Definition 9.3. For $\alpha = (\alpha_1, \dots, \alpha_n) \in K[[X_1, \dots, X_n]]^n$ we call

$$J(\alpha) := (\partial_j \alpha_i)_{i,j} \in K[[X_1, \dots, X_n]]^{n \times n}$$

the *Jacobi matrix* of α .

Example 9.4. The Jacobi matrix of the power sum polynomials $\rho = (\rho_1, \dots, \rho_n)$ is a deformed *Vandermonde matrix* $J(\rho) = (iX_j^{i-1})_{i,j}$ with determinant $n! \prod_{i < j} (X_j - X_i)$. The next theorem furnishes a new proof for the algebraic independence of ρ_1, \dots, ρ_n .

Theorem 9.5. *Polynomials $\alpha_1, \dots, \alpha_n \in \mathbb{C}[X_1, \dots, X_n]$ form a transcendence basis of $\mathbb{C}(X_1, \dots, X_n)$ if and only if $\det(J(\alpha)) \neq 0$.*

Proof. The proof follows [18, Proposition 3.10]. Suppose first that $\alpha_1, \dots, \alpha_n$ are algebraically dependent. Then there exists $\beta \in \mathbb{C}[X_1, \dots, X_n] \setminus \mathbb{C}$ such that $\beta(\alpha_1, \dots, \alpha_n) = 0$ and $\deg(\beta)$ is as small as possible. By (8.6),

$$\sum_{i=1}^n (\partial_k \alpha_i)(\partial_i \beta)(\alpha_1, \dots, \alpha_n) = \partial_k(\beta(\alpha_1, \dots, \alpha_n)) = 0$$

for $k = 1, \dots, n$. This is a homogeneous linear system over $\mathbb{C}(X_1, \dots, X_n)$ with coefficient matrix $J(\alpha)^t$ (the transpose of $F(\alpha)$). Since $\beta \notin \mathbb{C}$, there exists $1 \leq k \leq n$ such that $\partial_k \beta \neq 0$. Now $(\partial_k \beta)(\alpha_1, \dots, \alpha_n) \neq 0$, because $\deg(\beta)$ was chosen to be minimal. Hence, the linear system has a non-trivial solution and $\det(J(\alpha))$ must be 0.

Assume conversely, that $\alpha_1, \dots, \alpha_n$ are algebraically independent over \mathbb{C} . Since $\mathbb{C}(X_1, \dots, X_n)$ has transcendence degree n , the polynomials $X_i, \alpha_1, \dots, \alpha_n$ are algebraically dependent for each $i = 1, \dots, n$. Let $\beta_i \in \mathbb{C}[X_0, X_1, \dots, X_n] \setminus \mathbb{C}$ such that $\beta_i(X_i, \alpha_1, \dots, \alpha_n) = 0$ and $\deg(\beta_i)$ as small as possible. Again by (8.6),

$$\delta_{ik}(\partial_0 \beta_i)(X_i, \alpha_1, \dots, \alpha_n) + \sum_{j=1}^n (\partial_k \alpha_j)(\partial_j \beta_i)(X_i, \alpha_1, \dots, \alpha_n) = \partial_k(\beta_i(X_i, \alpha_1, \dots, \alpha_n)) = 0$$

for $i = 1, \dots, n$. Since $\alpha_1, \dots, \alpha_n$ are algebraically independent, X_0 must occur in every β_i . In particular, $\partial_0 \beta_i \neq 0$ has smaller degree than β_i . The choice of β_i implies $(\partial_0 \beta_i)(X_i, \alpha_1, \dots, \alpha_n) \neq 0$ for $i = 1, \dots, n$. This leads to the following matrix equation in $\mathbb{C}[X_1, \dots, X_n]$:

$$((\partial_j \beta_i)(X_i, \alpha_1, \dots, \alpha_n))_{i,j} J(\alpha) = -(\delta_{ij}(\partial_0 \beta_i)(X_i, \alpha_1, \dots, \alpha_n))_{i,j}.$$

Since the determinant of the diagonal matrix on the right hand side does not vanish, also $\det(J(\alpha))$ cannot vanish. \square

Definition 9.6. Let $C_a \subseteq K[[X_1, \dots, X_n]]$ be the set of power series with constant term $a \in K$, i. e. $\alpha \in C_a \iff \alpha(0) = a$. Let

$$K[[X_1, \dots, X_n]]^\circ := \{\alpha \in C_0^n : \det(J(\alpha)) \notin C_0\} \subseteq K[[X_1, \dots, X_n]]^n.$$

For $n = 1$ we have $\alpha \in K[[X_1, \dots, X_n]]^\circ \iff \alpha(0) = 0 \neq \alpha'(0) \iff \alpha \in (X) \setminus (X^2)$, so our notation is consistent with Theorem 3.4. The following is a multivariate analog.

Theorem 9.7 (Inverse function theorem). *The set $K[[X_1, \dots, X_n]]^\circ$ is a group with respect to \circ and*

$$K[[X_1, \dots, X_n]]^\circ \rightarrow \mathrm{GL}(n, K), \quad \alpha \mapsto J(\alpha)(0)$$

is a group epimorphism.

Proof. Let $\alpha, \beta \in K[[X_1, \dots, X_n]]^\circ$. Clearly, $\alpha \circ \beta \in C_0^n$. By (8.6),

$$\partial_j(\alpha_i(\beta)) = \sum_{k=1}^n (\partial_j \beta_k)(\partial_k \alpha_i)(\beta)$$

and $J(\alpha \circ \beta) = J(\alpha)(\beta) \cdot J(\beta)$. It follows that

$$J(\alpha \circ \beta)(0) = J(\alpha)(0)J(\beta)(0) \neq 0 \tag{9.1}$$

and $\alpha \circ \beta \in K[[X_1, \dots, X_n]]^\circ$. By fully exploiting (8.5), the associativity $(\alpha \circ \beta) \circ \gamma = \alpha \circ (\beta \circ \gamma)$ can be reduced to the easy case where $\alpha = (0, \dots, 0, X_i, 0, \dots, 0)$. The identity element of $K[[X_1, \dots, X_n]]^\circ$ is clearly (X_1, \dots, X_n) . For the construction of inverse elements, we first assume that $J(\alpha)(0) = 1_n$. Here we can adapt the proof of Theorem 8.5. We sort the n -tuples of indices lexicographically and define $\beta_{i,1} := X_i \in C_0$. For a given $\beta_{i,j}$ let $f(i, j) := (k_1, \dots, k_n)$ be the minimal tuple such that the coefficient c of $X_1^{k_1} \dots X_n^{k_n}$ in $\beta_{i,j}(\alpha_1, \dots, \alpha_n) - X_i$ is non-zero (if there is no such tuple we are done). Now let

$$\beta_{i,j+1} := \beta_{i,j} - cX_1^{k_1} \dots X_n^{k_n} \in C_0.$$

Since $(\partial_j \alpha_k)(0) = \delta_{kj}$, X_k is the unique monomial of degree 1 in α_k . Consequently, $X_1^{k_1} \dots X_n^{k_n}$ is the unique lowest degree monomial in $\alpha_1^{k_1} \dots \alpha_n^{k_n}$. Hence, going from $\beta_{i,j}(\alpha_1, \dots, \alpha_n)$ to $\beta_{i,j+1}(\alpha_1, \dots, \alpha_n)$ replaces $X_1^{k_1} \dots X_n^{k_n}$ with terms of higher degree. Consequently, $f(i, j+1) > f(i, j)$ and $\beta_i := \lim_{j \rightarrow \infty} \beta_{i,j} \in C_0$ exists with $\beta_i(\alpha_1, \dots, \alpha_n) = X_i$.

Now we consider the general case. As explained before, $\det(J(\alpha)) \notin C_0$ implies that $J(\alpha)$ is invertible. Let $S := (s_{ij}) = J(\alpha)^{-1}(0) \in K^{n \times n}$ and

$$\tilde{\alpha}_i := \sum_{j=1}^n s_{ij} \alpha_j \in C_0$$

for $i = 1, \dots, n$. Then

$$J(\tilde{\alpha})(0) = (\partial_j \tilde{\alpha}_i)_{i,j}(0) = \left(\sum_{k=1}^n s_{ik} (\partial_j \alpha_k)(0) \right)_{i,j} = SJ(\alpha)(0) = 1_n.$$

By the construction above, there exists $\tilde{\beta} \in C_0^n$ with $\tilde{\alpha} \circ \tilde{\beta} = (X_1, \dots, X_n)$. Define $\tilde{X}_i := \sum_{j=1}^n s_{ij} X_j \in C_0$ and $\beta_i := \tilde{\beta}_i(\tilde{X}_1, \dots, \tilde{X}_n) \in C_0$ for $i = 1, \dots, n$. Then

$$\sum_{j=1}^n s_{ij} \alpha_j(\beta_1, \dots, \beta_n) = \tilde{\alpha}_i \circ \beta = \tilde{\alpha}_i \circ \tilde{\beta} \circ (\tilde{X}_1, \dots, \tilde{X}_n) = \tilde{X}_i = \sum_{j=1}^n s_{ij} X_j.$$

Since S is invertible, it follows that $\alpha_i(\beta_1, \dots, \beta_n) = X_i$ for $i = 1, \dots, n$. By (9.1), $J(\beta)(0) = S$ and $\beta \in K[[X_1, \dots, X_n]]^\circ$ is the inverse of α with respect to \circ . This shows that $K[[X_1, \dots, X_n]]^\circ$ is a group. The map $\alpha \mapsto J(\alpha)(0)$ is a homomorphism by (9.1). For $A = (a_{ij}) \in \text{GL}(n, K)$ let $\alpha_i := a_{i1}X_1 + \dots + a_{in}X_n$. Then $\alpha \in C_0$ and $J(\alpha)(0) = A$. So our map is surjective. \square

If $\alpha_1, \dots, \alpha_n \in \mathbb{C}[X_1, \dots, X_n]$ are polynomials such that $\det(J(\alpha)) \in \mathbb{C}^\times$, the *Jacobi conjecture* (put forward by Keller [21] in 1939) claims that there exist polynomials β_1, \dots, β_n such that $\alpha \circ \beta = (X_1, \dots, X_n)$. This is still open even for $n = 2$ (see [41]).

An explicit formula for the reverse (i. e. the inverse with respect to \circ) is given by the following multivariate version of Theorem 7.5. To simplify the proof (which is still difficult) we restrict ourselves to those $\beta \in \mathbb{C}[[X_1, \dots, X_n]]$ such that $\beta_i \in X_i C_1 \subseteq C_0$. Note that $J(\beta)(0) = 1_n$ here.

Theorem 9.8 (LAGRANGE–GOOD’s inversion formula). *Let $\alpha \in \mathbb{C}[[X_1, \dots, X_n]]$ and $\beta_i \in X_i C_1$ for $i = 1, \dots, n$. Then*

$$\alpha = \sum_{k_1, \dots, k_n \geq 0} c_{k_1, \dots, k_n} \beta_1^{k_1} \dots \beta_n^{k_n} \tag{9.2}$$

where $c_{k_1, \dots, k_n} \in \mathbb{C}$ is the coefficient of $X_1^{k_1} \dots X_n^{k_n}$ in

$$\alpha \left(\frac{X_1}{\beta_1} \right)^{k_1+1} \dots \left(\frac{X_n}{\beta_n} \right)^{k_n+1} \det(J(\beta)).$$

Proof. The proof is taken from Hofbauer [17]. By the inverse function theorem, there exists $\gamma \in \mathbb{C}[[X_1, \dots, X_n]]^\circ$ such that $\gamma \circ \beta = (X_1, \dots, X_n)$. Replacing X_i by $\gamma_i(\beta)$ in α yields an expansion in the form (9.2) where we denote the coefficients by $\bar{c}_{k_1, \dots, k_n}$ for the moment. Observe that $\tau_i := X_i/\beta_i \in C_1$ and $\det(J(\beta)) \in C_1$. For $l_1, \dots, l_n \geq 0$ we define

$$\rho_{l_1, \dots, l_n} := \tau_1^{l_1+1} \dots \tau_n^{l_n+1} \det(J(\beta)) \in C_1.$$

Then c_{l_1, \dots, l_n} is, by definition, the coefficient of $X_1^{l_1} \dots X_n^{l_n}$ in $\alpha \rho_{l_1, \dots, l_n}$. So it also must be the coefficient of $X_1^{l_1} \dots X_n^{l_n}$ in

$$\sum_{\substack{k_1, \dots, k_n \geq 0 \\ \forall i: k_i \leq l_i}} \bar{c}_{k_1, \dots, k_n} X_1^{k_1} \dots X_n^{k_n} \rho_{l_1-k_1, \dots, l_n-k_n}.$$

It is easy to see that $c_{0, \dots, 0} = \alpha(0) = \bar{c}_{0, \dots, 0}$ as claimed. Hence, it suffices to show that $X_1^{k_1} \dots X_n^{k_n}$ does not occur in ρ_{k_1, \dots, k_n} for $(k_1, \dots, k_n) \neq (0, \dots, 0)$. By the product rule,

$$\tau_i \partial_j \beta_i = \partial_j(\beta_i \tau_i) - \beta_i \partial_j \tau_i = \delta_{ij} - X_i \frac{\partial_j \tau_i}{\tau_i}.$$

Since the (Jacobi) determinant is linear in every row, it follows that

$$\rho_{k_1, \dots, k_n} = \det(\delta_{ij} \tau_i^{k_i} - X_i \tau_i^{k_i-1} \partial_j \tau_i) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^n (\delta_{i\sigma(i)} \tau_i^{k_i} - X_i \tau_i^{k_i-1} \partial_{\sigma(i)} \tau_i).$$

By the (multivariate) Taylor series, we want to show that $(\partial_1^{k_1} \dots \partial_n^{k_n} \rho_{k_1, \dots, k_n})(0) = 0$.

Leibniz' rule applied to the inner product yields

$$P_\sigma := \sum_{l_{11} + \dots + l_{1s} = k_1} \dots \sum_{l_{n1} + \dots + l_{ns} = k_n} \frac{k_1! \dots k_n!}{\prod_{i,j} l_{ij}!} \prod_{t=1}^n \partial_1^{l_{1t}} \dots \partial_n^{l_{nt}} (\delta_{t\sigma(t)} \tau_t^{k_t} - X_t \tau_t^{k_t-1} \partial_{\sigma(t)} \tau_t)$$

Therein, we find

$$(\partial_1^{l_{1t}} \dots \partial_n^{l_{nt}} (X_t \tau_t^{k_t-1} \partial_{\sigma(t)} \tau_t))(0) = l_{tt} (\partial_1^{l_{1t}} \dots \partial_t^{l_{tt}-1} \dots \partial_n^{l_{nt}} (\tau_t^{k_t-1} \partial_{\sigma(t)} \tau_t))(0).$$

In particular, the product is zero if $\sigma(t) \neq t$ and $l_{tt} = 0$. We will disregard this case in the following. This also means that $t_{\sigma(t)\sigma(t)} < k_t$ whenever $\sigma(t) \neq t$. We set $\mu_i := \tau_i^{k_i}$ and observe that $\frac{1}{k_t} \partial_{\sigma(t)}(\mu_t) = \tau_t^{k_t-1} \partial_{\sigma(t)} \tau_t$. Hence, the inner product of $P_\sigma(0)$ takes the form

$$\prod_{t=1}^n (\delta_{t\sigma(t)} \partial_1^{l_{1t}} \dots \partial_n^{l_{nt}} \mu_t - \frac{l_{tt}}{k_t} \partial_1^{l_{1t}} \dots \partial_t^{l_{tt}-1} \dots \partial_{\sigma(t)}^{l_{\sigma(t)t}+1} \dots \partial_n^{l_{nt}} \mu_t).$$

Finally, we transform the indices via $l_{jt} \mapsto m_{jt} := l_{jt} - \delta_{jt} + \delta_{j\sigma(t)}$ (the problematic cases $l_{tt} = 0$ and $l_{\sigma(t)\sigma(t)} = k_{\sigma(t)}$ were excluded above). Note that $m_{1t} + \dots + m_{nt} = k_t$ and

$$\frac{l_{tt}}{l_{1t}! \dots l_{nt}!} = \frac{l_{\sigma(t)t} + 1}{l_{1t}! \dots (l_{tt} - 1)! \dots (l_{\sigma(t)t} + 1)! \dots l_{nt}!} = \frac{m_{\sigma(t)t}}{m_{1t}! \dots m_{nt}!}.$$

This turns $P_\sigma(0)$ into

$$P_\sigma(0) = \sum_{m_{ij}} \frac{k_1! \dots k_n!}{\prod_{i,j} m_{ij}!} \prod_{t=1}^n \partial_1^{m_{1t}} \dots \partial_n^{m_{nt}} (\mu_t)(0) \left(\delta_{t\sigma(t)} - \frac{m_{\sigma(t)t}}{k_t} \right).$$

Since only the last term actually depends on σ , we conclude

$$(\partial_1^{k_1} \dots \partial_n^{k_n} \rho_{k_1, \dots, k_n})(0) = \sum_{m_{ij}} \frac{k_1! \dots k_n!}{\prod_{i,j} m_{ij}!} \prod_{t=1}^n \partial_1^{m_{1t}} \dots \partial_n^{m_{nt}} (\mu_t)(0) \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{t=1}^n \left(\delta_{t\sigma(t)} - \frac{m_{\sigma(t)t}}{k_t} \right).$$

The final sum is the determinant of $(\delta_{ij} - m_{ji}/k_i)_{ij}$. This matrix is singular, since each column sum is $1 - \frac{1}{k_i} \sum_{j=1}^n m_{ji} = 0$. This completes the proof of $(\partial_1^{k_1} \dots \partial_n^{k_n} \rho_{k_1, \dots, k_n})(0) = 0$. \square

In an attempt to unify and generalize some dual pairs we have already found, we study the following setting. Let $A = (a_{ij}) \in \mathbb{C}^{n \times n}$ and $D = \operatorname{diag}(X_1, \dots, X_n)$. For $I \subseteq N := \{1, \dots, n\}$ let $A_I := (a_{ij})_{i,j \in I}$ and $X_I = \prod_{i \in I} X_i$. Since the determinant is linear in every row, we obtain

$$\begin{aligned} \det(1_n + DA) &= \begin{vmatrix} 1 & 0 & \dots & 0 \\ a_{21}X_2 & 1 + a_{22}X_2 & & a_{2n}X_2 \\ \vdots & & \ddots & \vdots \\ a_{n1}X_n & \dots & \dots & 1 + a_{nn}X_n \end{vmatrix} + \begin{vmatrix} a_{11} & \dots & a_{1n} \\ a_{21}X_2 & & a_{2n}X_2 \\ \vdots & & \vdots \\ a_{n1}X_n & \dots & 1 + a_{nn}X_n \end{vmatrix} X_1 \\ &= \begin{vmatrix} 1 + a_{22}X_2 & \dots & a_{2n}X_2 \\ \vdots & \ddots & \vdots \\ a_{n2}X_n & \dots & 1 + a_{nn}X_n \end{vmatrix} + \begin{vmatrix} a_{11} & \dots & \dots & a_{1n} \\ 0 & 1 & 0 & 0 \\ a_{31}X_3 & & & a_{3n}X_3 \\ \vdots & & & \vdots \\ a_{n1}X_n & \dots & \dots & 1 + a_{nn}X_n \end{vmatrix} X_1 \\ &\quad + \begin{vmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ a_{31}X_3 & \dots & a_{3n}X_3 \\ \vdots & & \vdots \\ a_{n1}X_n & \dots & 1 + a_{nn}X_n \end{vmatrix} X_1 X_2 = \dots \\ &= 1 + \sum_{i=1}^n a_{ii} X_i + \sum_{i < j} \det(A_{\{i,j\}}) X_i X_j + \dots + \det(A) X_N. \end{aligned}$$

Altogether,

$$\det(1_n + DA) = \sum_{I \subseteq N} \det(A_I) X_I, \quad (9.3)$$

where $\det(A_\emptyset) = 1$ for convenience. The dual equation, discovered by Vere-Jones [42], uses the *permanent* $\operatorname{per}(A) = \sum_{\sigma \in S_n} a_{1\sigma(1)} \dots a_{n\sigma(n)}$ of A :

$$\frac{1}{\det(1_n - DA)} = \sum_{k=0}^{\infty} \sum_{I \in N^k} \operatorname{per}(A_I) \frac{X_I}{k!}, \quad (9.4)$$

where I now runs through all tuples of elements in N (in contrast to the determinant, $\operatorname{per}(A_I)$ does not necessarily vanish if A_I has identical rows). We will derive (9.4) in Corollary 9.10 from the following result, which seems more amenable to applications.

Theorem 9.9 (MACMAHON'S Master Theorem). *Let $A = (a_{ij}) \in \mathbb{C}^{n \times n}$ and $D = \operatorname{diag}(X_1, \dots, X_n)$. Then*

$$\boxed{\frac{1}{\det(1_n - DA)} = \sum_{k_1, \dots, k_n \geq 0} c_{k_1, \dots, k_n} X_1^{k_1} \dots X_n^{k_n}}, \quad (9.5)$$

where $c_{k_1, \dots, k_n} \in \mathbb{C}$ is the coefficient of $X_1^{k_1} \dots X_n^{k_n}$ in

$$\prod_{i=1}^n (a_{i1}X_1 + \dots + a_{in}X_n)^{k_i}.$$

Proof. Let $A_i := a_{i1}X_1 + \dots + a_{in}X_n$ and $\beta_i := X_i(1 + A_i)^{-1} \in X_i C_1$ for $i = 1, \dots, n$. Let $D(\beta) := \text{diag}(\beta_1, \dots, \beta_n)$ and $\alpha := \det(1_n - D(\beta)A)^{-1}$. Since $\partial_j A_i = a_{ij}$, we obtain

$$\partial_j \beta_i = \frac{\delta_{ij}(1 + A_i) - X_i a_{ij}}{(1 + A_i)^2} = \frac{\delta_{ij} - \beta_i a_{ij}}{1 + A_i}$$

and

$$\alpha \det(J(\beta)) = \prod_{i=1}^n \frac{1}{1 + A_i}.$$

Hence, by Theorem 9.8, the coefficient of $\beta_1^{k_1} \dots \beta_n^{k_n}$ in α is the coefficient of $X_1^{k_1} \dots X_n^{k_n}$ in

$$\left(\frac{X_1}{\beta_1}\right)^{k_1+1} \dots \left(\frac{X_n}{\beta_n}\right)^{k_n+1} \prod_{i=1}^n \frac{1}{1 + A_i} = \prod_{i=1}^n (1 + a_{i1}X_1 + \dots + a_{in}X_n)^{k_i}.$$

Since the product on the right hand side has degree $k_1 + \dots + k_n$, the additional summand 1 plays no role and the desired coefficient really is c_{k_1, \dots, k_n} . By Theorem 9.7, the X_i can be substituted by some γ_i such that $\beta_1^{k_1} \dots \beta_n^{k_n}$ becomes $X_1^{k_1} \dots X_n^{k_n}$ and α becomes $\det(1_n - DA)^{-1}$. \square

A graph-theoretical proof of Theorem 9.9 was given by Foata and is presented in [8, Section 9.4]. There is also a short analytic argument which reduces the claim to the easy case where A is a triangular matrix.

Corollary 9.10. *Equation (9.4) holds.*

Proof. By the multinomial theorem we have

$$\begin{aligned} & \prod_{i=1}^n (a_{i1}X_1 + \dots + a_{in}X_n)^{k_i} \\ &= \sum_{k_{11} + \dots + k_{1n} = k_1} \dots \sum_{k_{n1} + \dots + k_{nn} = k_n} \frac{k_1! \dots k_n!}{\prod_{i,j} k_{ij}!} a_{11}^{k_{11}} a_{12}^{k_{12}} \dots a_{nn}^{k_{nn}} X_1^{k_{11} + \dots + k_{n1}} \dots X_n^{k_{1n} + \dots + k_{nn}}. \end{aligned}$$

To obtain c_{k_1, \dots, k_n} one needs to run only over those indices k_{ij} with $\sum_i k_{ij} = k_j$ for $j = 1, \dots, n$.

On the other hand, we need to sum over those tuples $I \in N^{k_1 + \dots + k_n}$ in (9.4) which contain i with multiplicity k_i for each $i = 1, \dots, n$. The number of those tuples is $\frac{(k_1 + \dots + k_n)!}{k_1! \dots k_n!}$. The factor $(k_1 + \dots + k_n)!$ cancels with $\frac{1}{k!}$ in (9.4). Since the permanent is invariant under permutations of rows and columns, we may assume that $I = (1^{k_1}, \dots, n^{k_n})$. Then A_I has the block form $A_I = (A_{ij})_{i,j}$ where

$$A_{ij} = a_{ij} \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix} \in \mathbb{C}^{k_i \times k_j}.$$

In the definition of $\text{per}(A_I)$, every permutation σ corresponds to a selection of n entries in A_I such that one entry in each row and each column is selected. Suppose that k_{ij} entries in block A_{ij} are selected.

Then $\sum_i k_{ij} = k_j$ and $\sum_j k_{ij} = k_i$. To choose the rows in each A_{ij} there are $\frac{k_1! \dots k_n!}{\prod k_{ij}!}$ possibilities. We get the same number for the selections of columns. Finally, once rows and columns are fixed, there are $\prod k_{ij}!$ choices to permute the entries in each block A_{ij} . Now the coefficient of $X_1^{k_1} \dots X_n^{k_n}$ in (9.4) turns out to be

$$\sum_{\substack{k_{ij} \\ \sum_i k_{ij} = k_j \\ \sum_j k_{ij} = k_i}} \frac{k_1! \dots k_n!}{\prod_{i,j} k_{ij}!} a_{11}^{k_{11}} a_{12}^{k_{12}} \dots a_{nn}^{k_{nn}} = c_{k_1, \dots, k_n}. \quad \square$$

We illustrate with some examples why MacMahon called Theorem 9.9 the *master* theorem (as he was a former major, I am tempted to call it the M^4 -theorem).

Example 9.11.

- (i) The expression $\det(1_n - DA)$ is reminiscent to the definition of the characteristic polynomial $\chi_A = X^n + s_{n-1}X^{n-1} + \dots + s_0 \in \mathbb{C}[X]$ of A . In fact, setting $X := X_1 = \dots = X_n$ allows us to regard $\det(1_n - XA)$ as a Laurent polynomial in X . We can then introduce X^{-1} to obtain

$$\det(1_n - XA) = X^n \det(X^{-1}1_n - A) = X^n \chi_A(X^{-1}) = 1 + s_{n-1}X + \dots + s_0X^n.$$

Now (9.3) in combination with Vieta's theorem yields

$$\sum_{\substack{I \subseteq N \\ |I|=k}} \det(A_I) = (-1)^k s_{n-k} = \sigma_k(\lambda_1, \dots, \lambda_n),$$

where $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ are the eigenvalues of A . This extends the familiar identities $\det(A) = \lambda_1 \dots \lambda_n$ and $\text{tr}(A) = \lambda_1 + \dots + \lambda_n$. With the help of Exercise 8.10, one can also express s_k in terms of $\rho_l(\lambda_1, \dots, \lambda_n) = \text{tr}(A^l)$.

- (ii) If $A = 1_n$ and $X_1 = \dots = X_n = X$, then (9.3) and (9.5) become

$$(1 + X)^n = \sum_{I \subseteq N} X^{|I|} = \sum_{k=0}^n \binom{n}{k} X^k,$$

$$(1 - X)^{-n} = \sum_{k_1, \dots, k_n \geq 0} X^{k_1 + \dots + k_n} = \sum_{k=0}^{\infty} \binom{n+k-1}{k} X^k,$$

since the k -element multisets correspond to the tuples (k_1, \dots, k_n) with $k_1 + \dots + k_n = k$ where k_i encodes the multiplicity of i .

- (iii) Taking $A = 1_n$ and $X_k = X^k$ in (9.5) recovers an equation from Theorem 5.4:

$$\prod_{k=1}^n \frac{1}{1 - X^k} = \sum_{k_1, \dots, k_n \geq 0} X^{k_1 + 2k_2 + \dots + nk_n} = \sum_{k=0}^{\infty} p_n(k) X^k.$$

Similarly, choosing $X_k = kX$ or $X_k = X_k Y$ leads more or less directly to Theorem 6.9 and Theorem 8.4 respectively.

- (iv) Take $(X_1, X_2, X_3) = (X, Y, Z)$ and

$$A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$$

in (9.5). Then by *Sarrus' rule*,

$$\begin{aligned} \frac{1}{\det(\mathbf{1}_3 - DA)} &= \frac{1}{1 + XZ + YZ + XY} = \sum_{k=0}^{\infty} (-1)^k (XY + YZ + ZX)^k \\ &= \sum_{k=0}^{\infty} (-1)^k \sum_{a+b+c=k} \frac{k!}{a!b!c!} X^{a+c} Y^{a+b} Z^{b+c}. \end{aligned}$$

The coefficient of $(XYZ)^{2n}$ is easily seen to be $(-1)^n \frac{(3n)!}{(n!)^3}$. On the other hand, the same coefficient in

$$(Y - Z)^{2n} (Z - X)^{2n} (X - Y)^{2n} = \sum_{a,b,c \geq 0} \binom{2n}{a} \binom{2n}{b} \binom{2n}{c} (-1)^{a+b+c} X^{c-b+2n} Y^{a-c+2n} Z^{b-a+2n}$$

occurs for $a = b = c$. This yields *Dixon's identity*:

$$(-1)^n \frac{(3n)!}{(n!)^3} = \sum_{k=0}^{2n} (-1)^k \binom{2n}{k}^3.$$

We end with a short outlook. Power series with an infinite set of indeterminants $\{X_i : i \in I\}$ can be defined by

$$K[[X_i : i \in I]] := \bigcup_{\substack{J \subset I \\ |J| < \infty}} K[[X_j : j \in J]].$$

Moreover, power series in non-commuting indeterminants exists and form what is sometimes called the *Magnus ring* $K\langle\langle X_1, \dots, X_n \rangle\rangle$ (the polynomial version is the *free algebra* $K\langle X_1, \dots, X_n \rangle$). The Lie bracket $[a, b] := ab - ba$ turns $K\langle\langle X_1, \dots, X_n \rangle\rangle$ into a *Lie algebra* and fulfills *Jacobi's identity*

$$[a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0.$$

The functional equation for $\exp(X)$ is replaced by the *Baker–Campbell–Hausdorff formula* in this context.

The reader might ask about formal Laurent series in multiple indeterminants. Although the field of fractions $K((X_1, \dots, X_n))$ certainly exists, its elements do not look like one might expect. For example, the inverse of $X - Y$ could be $\sum_{k=1}^{\infty} X^{-k} Y^{k-1}$ or $-\sum_{k=1}^{\infty} X^{k-1} Y^{-k}$. The first series lies in $K((X))((Y))$, but not in $K((Y))((X))$. For the second series it is the other way around.

Acknowledgment

I thank Diego García Lucas, Till Müller and Alexander Zimmermann for proofreading. The work is supported by the German Research Foundation (SA 2864/1-2 and SA 2864/3-1).

References

- [1] S. Ahlgren, *Distribution of the partition function modulo composite integers M* , Math. Ann. **318** (2000), 795–803.

- [2] G. E. Andrews, *A simple proof of Jacobi's triple product identity*, Proc. Amer. Math. Soc. **16** (1965), 333–334.
- [3] G. E. Andrews, *On the proofs of the Rogers-Ramanujan identities*, in: *q-series and partitions* (Minneapolis, MN, 1988), 1–14, IMA Vol. Math. Appl., Vol. 18, Springer, New York, 1989.
- [4] G. E. Andrews, *The theory of partitions*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1998.
- [5] G. E. Andrews and K. Eriksson, *Integer partitions*, Cambridge University Press, Cambridge, 2004.
- [6] E. Berlekamp, *Algebraic coding theory*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2015.
- [7] D. M. Bressoud, *Some identities for terminating q-series*, Math. Proc. Cambridge Philos. Soc. **89** (1981), 211–223.
- [8] R. A. Brualdi and H. J. Ryser, *Combinatorial matrix theory*, Encyclopedia of Mathematics and its Applications, Vol. 39, Cambridge University Press, New York, 2013.
- [9] R. Chapman, *A new proof of some identities of Bressoud*, Int. J. Math. Math. Sci. **32** (2002), 627–633.
- [10] J.-M. Deshouillers, F. Hennecart and B. Landreau, 7 373 170 279 850, Math. Comp. **69** (2000), 421–439.
- [11] I. M. Gessel, *Lagrange inversion*, J. Combin. Theory Ser. A **144** (2016), 212–249.
- [12] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, Sixth edition, Oxford University Press, Oxford, 2008.
- [13] M. Hardy, *Combinatorics of partial derivatives*, Electron. J. Combin. **13** (2006), Research Paper 1, 13.
- [14] M. D. Hirschhorn, *Polynomial identities which imply identities of Euler and Jacobi*, Acta Arith. **32** (1977), 73–78.
- [15] M. D. Hirschhorn, *A short and simple proof of Ramanujan's mod 11 partition congruence*, J. Number Theory **139** (2014), 205–209.
- [16] M. D. Hirschhorn, *The power of q*, Developments in Mathematics, Vol. 49, Springer, Cham, 2017.
- [17] J. Hofbauer, *A short proof of the Lagrange-Good formula*, Discrete Math. **25** (1979), 135–139.
- [18] J. E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge Studies in Advanced Mathematics, Vol. 29, Cambridge University Press, Cambridge, 1990.
- [19] W. P. Johnson, *An introduction to q-analysis*, American Mathematical Society, Providence, RI, 2020.
- [20] J. T. Joichi and D. Stanton, *An involution for Jacobi's identity*, Discrete Math. **73** (1989), 261–271.
- [21] O.-H. Keller, *Ganze Cremona-Transformationen*, Monatsh. Math. Phys. **47** (1939), 299–306.
- [22] L. W. Kolitsch and S. Kolitsch, *A combinatorial proof of Jacobi's triple product identity*, Ramanujan J. **45** (2018), 483–489.
- [23] J. Konvalina, *A unified interpretation of the binomial coefficients, the Stirling numbers, and the Gaussian coefficients*, Amer. Math. Monthly **107** (2000), 901–910.
- [24] J. M. Kubina and M. C. Wunderlich, *Extending Waring's conjecture to 471,600,000*, Math. Comp. **55** (1990), 815–820.
- [25] S. Lang, *Algebra*, Graduate Texts in Mathematics, Vol. 211, Springer-Verlag, New York, 2002.
- [26] R. P. Lewis, *A combinatorial proof of the triple product identity*, Amer. Math. Monthly **91** (1984), 420–423.
- [27] P. A. MacMahon, *Combinatory analysis Vol.1*, Cambridge University Press, Cambridge, 1915.
- [28] S. Marivani, *Another elementary proof that $p(11n + 6) \equiv 0 \pmod{11}$* , Ramanujan J. **30** (2013), 187–191.

- [29] A. Maróti, *Symmetric functions, generalized blocks, and permutations with restricted cycle structure*, European J. Combin. **28** (2007), 942–963.
- [30] I. Niven, *Formal power series*, Amer. Math. Monthly **76** (1969), 871–889.
- [31] K. J. Nowak, *Some elementary proofs of Puiseux’s theorems*, Univ. Iagel. Acta Math. (2000), 279–282.
- [32] K. Ono, *Distribution of the partition function modulo m* , Ann. of Math. (2) **151** (2000), 293–307.
- [33] S. Ramanujan, *Some properties of $p(n)$, the number of partitions of n* , Proc. Camb. Philos. Soc. **19** (1919), 207–210.
- [34] L. J. Rogers and S. Ramanujan, *Proof of certain identities in combinatory analysis*, Proc. Camb. Philos. Soc. **19** (1919), 211–216.
- [35] A. V. Sills, *An invitation to the Rogers-Ramanujan identities*, CRC Press, Boca Raton, FL, 2018.
- [36] R. P. Stanley, *Enumerative combinatorics. Vol. 2*, Cambridge Studies in Advanced Mathematics, Vol. 62, Cambridge University Press, Cambridge, 1999.
- [37] C. Sudler, *Two enumerative proofs of an identity of Jacobi*, Proc. Edinburgh Math. Soc. (2) **15** (1966), 67–71.
- [38] J. J. Sylvester and F. Franklin, *A Constructive Theory of Partitions, Arranged in Three Acts, an Interact and an Exodion*, Amer. J. Math. **5** (1882), 251–330.
- [39] W. T. Tutte, *On elementary calculus and the Good formula*, J. Combinatorial Theory Ser. B **18** (1975), 97–137.
- [40] W. T. Tutte, *Erratum: “On elementary calculus and the Good formula”*, J. Combinatorial Theory Ser. B **19** (1975), 287.
- [41] A. van den Essen, S. Kuroda and A. J. Crachiola, *Polynomial automorphisms and the Jacobian conjecture—new results from the beginning of the 21st century*, Frontiers in Mathematics, Birkhäuser/Springer, Cham, 2021.
- [42] D. Vere-Jones, *An identity involving permanents*, Linear Algebra Appl. **63** (1984), 267–270.
- [43] E. M. Wright, *An enumerative proof of an identity of Jacobi*, J. London Math. Soc. **40** (1965), 55–57.
- [44] J.-M. Zhu, *A semi-finite proof of Jacobi’s triple product identity*, Amer. Math. Monthly **122** (2015), 1008–1009.
- [45] J. Zolnowsky, *A direct combinatorial proof of the Jacobi identity*, Discrete Math. **9** (1974), 293–298.