Analytic Combinatorics

Generating Functions and their Applications in Algebra and Geometry

Luca Wellmeier

Bachelor's Thesis
First supervisor: Prof. Dr. Michael Baake
Second supervisor: Prof. Dr. Christopher Voll

Bielefeld University, 3rd April 2020

Contents

Introduction

Analytic combinatorics can be described as the process of encoding discrete structures as *generating functions*, i.e. formal series in formal variables (formal power series, for instance), to exploit their analytic properties and thereby quantify interesting aspects of the underlying structure. This paper is meant to be a quick introduction into the theory as presented in the book [FS09] by Philippe Flajolet and Robert Sedgewick, which was, in fact, its first extensive treatment.

In the first part we develop tools to actually find generating functions. Flajolet and Sedgewick summarize these as the "symbolic method": We describe the discrete structure in question in terms of elementary concepts like sets, multisets or sequences, and then use a *dictionary* to translate the description into generating functions (which could easily be carried out by a computer!). They answer the question "How many objects are in my structure of this given size?" simply by computing series coefficients. Moreover, we use notions from probability theory to compute more specific generating functions enabling us to solve problems like "All objects in this structure are built in some way from objects of a simpler structure. How many simpler objects can I expect this one to consist of?".

We introduce three types of generating function: ordinary, exponential and Dirichlet. Applying all of them to the same problem generally results in different levels of "usefulness", which is intimately connected to the nature of the underlying structure. Therefore, we will categorize combinatorial structures into rough categories (additive/multiplicative, unlabelled/labelled) to find which type to use. Readers might also be interested in Wilf's book [Wil94] for a more intuitive and elementary introduction to generating functons.

The second part of this paper is concerned with analytic properties of generating functions. The general idea is to extract asymptotic growth properties of the quantitative properties of the structures encoded in generating function using methods from complex analysis. These properties are absolutely necessary in the average-case analysis of algorithms, for example.

For ordinary and exponential generating functions we introduce the theory of singularity analysis, also significantly influenced by Flajolet's prior work. It allows to extract asymptotic information of a large class of singular functions by analyzing the nature and position of their singularities. Note, however, that there are many more tools available (the most important ones are treated in the book).

We analyze Dirichlet generating functions using the Wiener-Ikehara Tauberian theorem. The theorem also establishes a connection between nature and position of the singularity and quantitative properties of the structure.

Part I

Symbolic Constructions

Chapter 1

Combinatorial Classes and Generating Functions

This chapter is devoted to the basic notations and computation of generating function. Furthermore, we tackle the question of which generating function to use for which kind of problem.

The first two section of this chapter are based on [FS09] Chapters I and II, whereas the third is based on Section 2.6 of [Wil94].

Definition 1.1. *Let $\mathcal{C}$ be a set together with a size function $|\cdot|\colon \mathcal{C} \to \mathbb{N}_0$. We denote by $\mathcal{C}_n$ the set of all objects $\gamma \in \mathcal{C}$ such that $|\gamma| = n$ and we write $a_n = \#\mathcal{A}_n$.*

If $\mathcal{C}$ is countable and all a_n are finite we call the pair $(\mathcal{C}, |\cdot|)$ a combinatorial class or combinatorial structure and the sequence $\{a_n\}_{n\in\mathbb{N}_0}$ its counting sequence.

Two combinatorial classes $\mathcal{A}$ and $\mathcal{B}$ are called isomorphic if there is a bijective map ϕ between them, which preserves size. In that case we write $\mathcal{A} = \mathcal{B}$ ("overwriting" set equality).

Definition 1.2. *To each combinatorial class $\mathcal{C}$ with counting sequence $\{c_n\}$ we can associate a formal series in the indeterminant z, called a generating function, in the following ways:*

- *The ordinary generating function of $\mathcal{C}$ is defined by*

$$C(z) = \sum_{\gamma\in\mathcal{C}} z^{|\gamma|} = \sum_{n=0}^{\infty} c_n z^n.$$

We denote the transition from the class to its OGF by $\mathcal{C} \xrightarrow{OGF} C(z)$.

- *The exponential generating function of $\mathcal{C}$ is defined by*

$$C(z) = \sum_{\gamma\in\mathcal{C}} \frac{z^{|\gamma|}}{|\gamma|!} = \sum_{n=0}^{\infty} c_n \frac{z^n}{n!}.$$

The transition is denoted by $\mathcal{C} \xrightarrow{EGF} C(z)$.

- *If $c_0 = 0$, we define the Dirichlet generating function of $\mathcal{C}$ to be*

$$C(z) = \sum_{\gamma\in\mathcal{C}} |\gamma|^{-z} = \sum_{n=1}^{\infty} \frac{c_n}{n^z}$$

with notation $\mathcal{C} \xrightarrow{DGF} C(z)$.

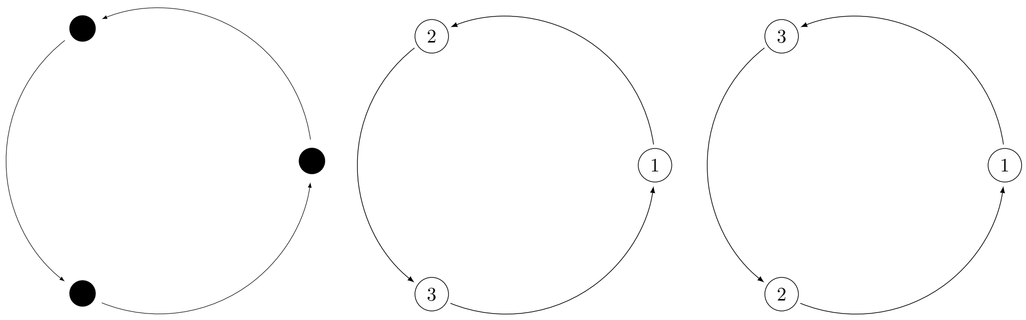


Figure 1.1: Unlabelled versus labelled structures. Three cycles: the left is unlabelled while the middle and right are labelled but differently labelled

The coefficient extraction operators $[z^n]$, $n![z^n]$ *and* $[n^{-z}]$ *map a generating function to its coefficient of* z^n, $z^n/n!$ *and* n^{-z}, *respectively.*

We will always use the same letter in different variations for combinatorial class, counting sequence and generating function ($\mathcal{A}$, a_n and $A(z)$, for instance). Also, if there are multiple classes in play, the size function will be subscripted with the letter of the belonging class, i.e. $|\cdot|_{\mathcal{A}}$.

Now we have three different ways at hand to translate a combinatorial class into a generating function. The question, which one to use in a particular problem, is treated at the beginning of each of the next sections, where the combinatorial meaning behind them is explained.

1.1 Unlabelled Classes and Ordinary Generating Functions

In all three cases it is very helpful to first look at the way the functions multiply. The reason is that most combinatorial classes can be constructed from simpler classes, thus counting all objects of a specific size can often be broken down to counting objects of smaller sizes and then merge these counts together in some way. This process frequently involves products of generating functions as we will see in this chapter.

In this case, a product of OGFs $C(z) = A(z)B(z)$ satisfies

$$[z^n]C(z) = \sum_{k=0}^{n} a_k b_{n-k}. \tag{1.1}$$

This means: The objects of the class $\mathcal{C}$ of size n are assembled from size-k-objects $\mathcal{A}$ and and size-$(n-k)$-objects of $\mathcal{B}$. The sum iterates through all possibilities. Furthermore, this formula exactly counts how many new objects appear, when two "unlabelled" structures are merged together (more detail below).

The term "unlabelled" means that we do not distinguish between the basic building blocks of the structure. This is best understood by visualizing (see Figure 1.1) and by comparison with the labelled case in the next section. We call a combinatorial class, which we intend to translate into an ordinary generating function a labelled class.

In order to construct a class using simpler objects we now introduce elementary means to do so.

Let $\mathcal{A}$ and $\mathcal{B}$ be combinatorial classes.

- A neutral class is a combinatorial class consisting of a single object, which has size 0.
- Similarly, an atom is a combinatorial class consisting of a single size-1-object.

- The combinatorial sum of two classes is again a combinatorial class defined as their disjoint union. In symbols,
$$\mathcal{A} + \mathcal{B} := \mathcal{A} \sqcup \mathcal{B}.$$
We define a size function for $\alpha \in \mathcal{A} + \mathcal{B}$ by
$$|\alpha| = \begin{cases} |\alpha|_{\mathcal{A}}, & \text{if} \quad \alpha \in \mathcal{A} \\ |\alpha|_{\mathcal{B}}, & \text{if} \quad \alpha \in \mathcal{B} \end{cases}$$

- The product of two classes is defined as their cartesian product
$$\mathcal{A} \cdot \mathcal{B} := \mathcal{A} \times \mathcal{B}.$$
Using the size function
$$|(\alpha, \beta)| = |\alpha|_{\mathcal{A}} + |\beta|_{\mathcal{B}}$$
the product becomes a combinatorial class.

- If $\mathcal{A}$ contains no objects of size 0, we define the class of all finite sequences as
$$\mathrm{Seq}(\mathcal{A}) = \mathcal{E} + \mathcal{A} + \mathcal{A} \cdot \mathcal{A} + \mathcal{A} \cdot \mathcal{A} \cdot \mathcal{A} + \dots,$$
where $\mathcal{E}$ is a neutral class.

- If $\mathcal{A}$ contains no objects of size 0, we define the combinatorial class of all finite multisets (normal sets, which allow repitions of elements) from elements of $\mathcal{A}$ as
$$\mathrm{MSet}(\mathcal{A}) = \mathrm{Seq}(\mathcal{A})/R,$$
where R is the equivalence relation which identifies two sequences whenever they have the same length and are permutations of each other.

Note that there is no need to explicitly define a size function on the last two constructions since they are derived from sum and product. On that matter, note as well that the size function on multisets is obviously well-defined.

The OGFs of neutral classes and atoms are simple as they are just 1 and z, respectively. The next result establishes a *dictionary* for the translation.

Theorem 1.3 (Unlabelled dictionary)**.** *Let $\mathcal{A}$ and $\mathcal{B}$ be combinatorial classes. The following translation rules hold for OGFs:*
$$\begin{aligned} \mathcal{A} + \mathcal{B} &\xrightarrow{OGF} A(z) + B(z) \\ \mathcal{A} \times \mathcal{B} &\xrightarrow{OGF} A(z)B(z) \\ \mathrm{Seq}(\mathcal{A}) &\xrightarrow{OGF} \frac{1}{1 - A(z)} \\ \mathrm{MSet}(\mathcal{A}) &\xrightarrow{OGF} \prod_{n \geq 1} (1 - z^n)^{-a_n} = \exp\left(\sum_{k=1}^{\infty} \frac{1}{k} A(z^k)\right). \end{aligned}$$

Proof. Sum and Product are immediately verified using the fact that the classes are disjoint for the sum and multiplication formula (1.1) of two OGFs for the product.

The OGF of sequences clearly translates as
$$\mathrm{Seq}(\mathcal{A}) \xrightarrow{OGF} A(z) = 1 + A(z) + A(z)^2 + \dots,$$
which is the multiplicative inverse of $1 - A(z)$ (in the formal power series sense).

Multisets require some more work. First, note that we can rearrange any multiset into the form $[a, \dots, a, b, \dots, b, c, \dots]$ since order does not matter. Therefore, we have the isomorphism

$$\mathrm{MSet}(\mathcal{A}) = \prod_{\alpha \in \mathcal{A}} \mathrm{Seq}(\{\alpha\}).$$

From this, the first claimed product expression follows directly and the second follows after inserting the product expression into $\exp(1 + (\log(\cdot) - 1))$, where $\exp(u)$ and $\log(1+u)$ are defined as formal power series using their usual definition. ■

One might also come across cases where the combinatorial class of interest is only expressible through an implicit equation of classes. Then, the next theorem can help.

Theorem 1.4 (Implicit dictionary for OGFs, [FS09] Theorem I.5)**.** *Let $\mathcal{A}$ and $\mathcal{B}$ be combinatorial classes. Suppose a third class $\mathcal{X}$ is unknown but specified implicitly. The following rules hold:*

$$\begin{aligned}
\mathcal{A} = \mathcal{B} + \mathcal{X} &\implies \mathcal{X} \xrightarrow{OGF} A(z) - B(z) \\
\mathcal{A} = \mathcal{B} \times \mathcal{X} &\implies \mathcal{X} \xrightarrow{OGF} \frac{A(z)}{B(z)} \\
\mathcal{A} = \mathrm{Seq}(\mathcal{X}) &\implies \mathcal{X} \xrightarrow{OGF} 1 - \frac{1}{A(z)} \\
\mathcal{A} = \mathrm{MSet}(\mathcal{X}) &\implies \mathcal{X} \xrightarrow{OGF} \sum_{k \geq 1} \frac{\mu(k)}{k} \log A(z^k).
\end{aligned}$$

(μ is the Möbius function)

Proof. The first three cases are easy manipulations of formal power series equations. Turning to multisets, we write $L(z) = \log(A(z))$. Using Theorem 1.3 we find

$$L(z) = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \frac{1}{k} x_j z^{jk} = \sum_{n=0}^{\infty} \left(\sum_{d|n} \frac{d}{n} x_d \right) z^n.$$

With $l_n = [z^n]L(z)$ we can compare coefficients and find $nl_n = \sum_{d|n} d x_d$, which satisfies Möbius inversion (see Theorem 2.9 of [Apo76]). Thus,

$$X(z) = \sum_{n=0}^{\infty} x_n z^n = \sum_{n=0}^{\infty} \sum_{d|n} \mu\left(\frac{n}{d}\right) \frac{d}{n} l_d z^n = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \frac{\mu(k)}{k} l_k z^{nk}$$

and the claim follows. ■

The tools developed in this section are very useful when it comes to actually finding ordinary generating functions of unlabelled combinatorial structures. The next two examples demonstrate this process.

Example 1.5 (Irreducible polynomials over finite fields)**.** *Consider the finite field $\mathbb{F}_p = \{0, 1, \dots, p-1\}$ for p prime. We are interested in counting various aspects of factorization in the ring of polynomials $\mathbb{F}_p[X]$. Without restriction we only need to investigate the subset of monic polynomials (i.e. polynomials with leading coefficient 1) since the results are easily extendable.*

Let $\mathcal{Z}$ be an atom. The field $\mathbb{F}_p$ can be represented as a combinatorial class by adding p copies of $\mathcal{Z}$, where each copy stands for one element of $\mathbb{F}_p$. In symbols,

$$\mathcal{F} = \underbrace{\mathcal{Z} + \cdots + \mathcal{Z}}_{p \text{ times}} \xrightarrow{OGF} F(z) = pz.$$

Now, observe that we can interpret any polynomial $\sum_{k=0}^{n} a_k X^k$ *with* $a_k \in \mathbb{F}_p$ *and* $a_n = 1$ *as a finite sequence* $(a_0, a_1, \dots, a_{n-1})$*. The empty sequence represents the constant polynomial* 1*. Therefore, denoting the class of all monic polynomials by* $\mathcal{P}$ *we can write*

$$\mathcal{P} = \mathrm{Seq}(\mathcal{F}) \xrightarrow{OGF} P(z) = \frac{1}{1 - F(z)} = \frac{1}{1 - pz}.$$

Note that the size of objects in $\mathcal{P}$ *equals exactly the degree of the polynomial, which it represents. This is a consequence of our decision to give each element of the base field the size* 1*.*

On the other hand, using the fact that $\mathbb{F}_p[X]$ *is a Euclidean ring, hence a unique factorization domain, we can factorize each polynomial into its irreducible factors. Clearly the order of the factors does not matter, but multiple copies of irreducible factors are allowed. As a consequence, we can also construct* $\mathcal{P}$ *by*

$$\mathcal{P} = \mathrm{MSet}(\mathcal{I}) \xrightarrow{OGF} P(z) = \exp\left(\sum_{k=1}^{\infty} \frac{1}{k} I(z^k)\right)$$

with $\mathcal{I} \xrightarrow{OGF} I(z)$ *as the class of all monic irreducible polynomials.*

We are now in a situation where the interesting class, $\mathcal{I}$ *in this case, is unknown, but we have an implicit equation for its OGF. Applying Theorem 1.4 yields*

$$I(z) = \sum_{k=1}^{\infty} \frac{\mu(k)}{k} \log \frac{1}{1 - pz^k}.$$

This is quite a remarkable result on its own. It enables us to compute the number of irreducible polynomials in every given degree just by extracting the corresponding coefficient in this formal power series. However, the formula still depends on factorizing natural numbers to evaluate μ*, which makes it an unhandy tool, for many practical purposes. See Example 2.9 for an extension of this example.*

Example 1.6 (Conjugacy classes of finite symmetric groups)**.** *Let* S_n *be the symmetric group of the set* $\{1, 2, \dots, n\}$*. Each permutation* $\sigma \in S_n$ *can be uniquely decomposed into a product of disjoint cycles. We assign to* σ *the cycle type* $(n_1, \dots, n_r)$ *with* $n_1 \geq \cdots \geq n_r \geq 1$ *and* $\sum_{k=0}^{r} n_k = n$*, if the factors of its decomposition have exactly these lengths. Two permutations are in the same conjugacy class, if and only if they have the same cycle type (see for example [Pro16a]). Therefore, counting conjugacy classes of* S_n *can be broken down to counting cycle types.*

To model this combinatorially we first construct the class of positive integers. Let $\mathcal{Z}$ *be an atom. We can interpret an integer* $n \geq 0$ *as class* $\mathcal{Z}^n$*, since it only contains a single object of size* n*. Consequently, we model the class in question by*

$$\mathcal{N} = \mathrm{Seq}_{\geq 1}(\mathcal{Z}) \xrightarrow{OGF} N(z) = \frac{1}{1 - z} - 1 = \frac{z}{1 - z},$$

where $\mathrm{Seq}_{\geq 1}$ *is the sequence construction without objects of size* 0*.*

Now we need to find all possibilities to decompose a positive integer into a sum. These sequences of numbers are called partitions. In a cycle the order of the disjoint factors does not matter, so neither does the order of numbers in cycle types. It follows that we can construct all partitions simply as multisets of positive integers:

$$\mathcal{P} = \mathrm{MSet}(\mathcal{N}) \xrightarrow{OGF} P(z) = \prod_{k=1}^{\infty} \frac{1}{1 - z^k}.$$

1.2 Labelled Classes and Exponential Generating Functions

The product $C(z) = A(z)B(z)$ of two exponential generating functions is again an exponential generating function with coefficients

$$n![z^n]C(z) = \sum_{n_1+n_2=n} \binom{n}{n_1, n_2} a_k b_{n-k},$$

using the multinomial coefficient $\binom{n}{n_1,\dots,n_r} = \frac{n!}{n_1!\dots n_r!}$. We interpret this as follows: We think of each object in the two classes as graphs. Assume that in both classes all objects carry distinct integer labels on each vertex. Take all such objects of size k in $\mathcal{A}$ and of size $n-k$ in $\mathcal{B}$ (i.e. graphs that consist of k and $n-k$ vertices) and "assemble" them to size-n-graphs, similar to the unlabelled case ($a_k b_{n-k}$ possibilities). Then, fix n distinct integers and assign them to each of the new objects as labels and iterate through all possibilities to do so.

Since all combinatorial structures can be thought of as graphs with certain properties, EGFs are convenient when encoding classes whose graphs are labelled (see Figure 1.1 again where the two cycles on the right are different as labelled objects, while they are not distinguishable if we took away the labels).

Now, we formalize our observations. We use the same notion of a combinatorial class, but call it a labelled class in order to clarify our intention to use exponential generating functions to encode it (i.e. we interpret its objects as labelled graphs). A neutral class is again a class containing a single object of size 0 holding no label at all, whereas an atom contains one object of size 1 labelled by **1**.

We call a labelled object well-labelled if the set of labels exactly equals $\{1, 2, \dots, n\}$, where n is its size.

As already seen above, the key concept of labelled classes in comparison to the unlabelled case is the consideration of possible relabellings of objects. Note, however, that most relabellings are equivalent, in the sense following sense:

Definition 1.7. *For some labelled object α in a labelled class (i.e. a graph) we define its reduction $\rho(\alpha)$ to be the unique labelled graph satisfying:*

1. *The labels are exactly the numbers 1, 2, $\dots$, $|\alpha|$ and*
2. *for reduced labels a', b' of a connected pair of labels $a, b \in \mathbb{Z}$ with $a < b$ we have $a' < b'$.*

We call two labelled objects α and β equivalent, if $\rho(\alpha) = \rho(\beta)$

This leads us to the (labelled) product and related constructions.

Definition 1.8. *Let $\mathcal{A}$ and $\mathcal{B}$ be labelled classes.*

- *The sum of two labelled classes is defined exactly the same way as in the unlabelled case, i.e. as disjoint union:*
$$\mathcal{A} + \mathcal{B} = \mathcal{A} \sqcup \mathcal{B}.$$
- *The labelled product of two objects (graphs) $\alpha \in \mathcal{A}, \beta \in \mathcal{B}$ is obtained by taking all inequivalent relabellings of the combined graph:*
$$\alpha * \beta = \{(\alpha', \beta') : (\alpha', \beta') \text{ is well-labelled}, \rho(\alpha') = \alpha, \rho(\beta') = \beta\}.$$
More generally, the (labelled) product of two labelled classes is defined as the labelled class
$$\mathcal{A} * \mathcal{B} = \bigcup_{\alpha \in \mathcal{A}, \beta \in \mathcal{B}} \alpha * \beta$$
with size function $|(\alpha, \beta)| = |\alpha|_{\mathcal{A}} + |\beta|_{\mathcal{B}}$.

- *The labelled class of all finite sequences is*

$$\mathrm{Seq}(\mathcal{A}) = \mathcal{E} + \mathcal{A} + \mathcal{A}\cdot\mathcal{A} + \mathcal{A}\cdot\mathcal{A}\cdot\mathcal{A} + \ldots$$

  provided that $\mathcal{A}_0 = \varnothing$*.*

- *The labelled class of all sets is in fact the same as the class of all multisets in the unlabelled class. Thus,*

$$\mathrm{Set}(\mathcal{A}) = \mathrm{Seq}(\mathcal{A})/R,$$

  where R *identifies two sequences of the same length if they are permutations of each other.*

- *The labelled class of all cycles is obtained through*

$$\mathrm{Cyc}(\mathcal{A}) = \mathrm{Seq}(\mathcal{A})/S,$$

  where S *identifies two sequences of the same length if they are cyclic permutations of each other.*

- *The labelled class of all undirected cycles is implicitely defined through*

$$\mathrm{UCyc}(\mathcal{A}) + \mathrm{UCyc}(\mathcal{A}) = \mathrm{Cyc}(\mathcal{A}).$$

Theorem 1.9 (Labelled Dictionary)**.** *Let* $\mathcal{A}$ *and* $\mathcal{B}$ *be labelled classes. The labelled constructions admit the following translations to exponential generating function:*

$$\begin{aligned}
\mathcal{A} + \mathcal{B} &\xrightarrow{EGF} A(z) + B(z) \\
\mathcal{A} * \mathcal{B} &\xrightarrow{EGF} A(z)B(z) \\
\mathrm{Seq}(\mathcal{A}) &\xrightarrow{EGF} \frac{1}{1 - A(z)} \\
\mathrm{Set}(\mathcal{A}) &\xrightarrow{EGF} \exp(A(z)) \\
\mathrm{Cyc}(\mathcal{A}) &\xrightarrow{EGF} \log\left(\frac{1}{1 - A(z)}\right) \\
\mathrm{UCyc}(\mathcal{A}) &\xrightarrow{EGF} \frac{1}{2}\log\left(\frac{1}{1 - A(z)}\right).
\end{aligned}$$

Proof. The first follow immediately from the definition. Next, we claim that the equivalence class of a set with k elements contains $k!$ sequences. Indeed, since the objects are labelled, all sequences that permute the k elements of the set are equivalent. Therefore, denoting the subclass of k-sets by $\mathrm{Set}_k(\mathcal{A})$ we obtain

$$\mathrm{Set}_k(\mathcal{A}) \xrightarrow{EGF} \frac{1}{k!}B(z)^k,$$

since $B(z)^k$ is the EGF of all sequences with k elements. It follows that $Set(\mathcal{A}) = \bigcup_k \mathrm{Set}_k(\mathcal{A}) \xrightarrow{EGF} \exp(A(z))$. The calculation for Cyc works similarly. The equivalence class of a k-cycle contains k sequences, that we get by right-shifting a sequence k times. The EGF for undirected cycles is derived from the normal cycle class by interpreting two copies of it as the two directions of edges in a cycle. ■

Example 1.10 (Point clouds in line systems)**.** *Choose a set of* n *labelled lines in the Euclidean plane such that no two lines are parallel and no three lines are concurrent (intersect in a single point). Let* P *be the set of intersection points. We call a subset*

$N \subseteq P$ a cloud if it has cardinality n and does not contain any three collinear points. How many clouds are there for given n?

Given a cloud N, if we identify the set of labelled lines with labelled vertices and choose N to be the set of edges (two points or lines are connected by an edge if their intersection is in N), we obtain a graph, within which each vertex has exactly two neighbours (due to the exclusion of collinear points). These are called 2-regular graphs.

In order to compute their generating function we apply the above construction. It is easy to see that there are none of sizes 0, 1 and 2 and all 2-regular graphs from size 3 to 5 are simply circular arrangements of the points. From size 6 on, however, we also encounter graphs which are not completely connected. Each component of these graphs must itself be 2-regular. Therefore, we find the following construction:

$$\mathcal{R} = \mathrm{Set}(\mathrm{UCyc}_{\geq 3}(\mathcal{Z})) \xrightarrow{EGF} R(z) = \frac{\exp\left(-\frac{z}{2} - \frac{z^2}{4}\right)}{\sqrt{1-z}},$$

where $\mathcal{Z}$ is a labelled atom and the $\mathrm{UCyc}_{\geq 3}$ stands for all undirected cycles of size at least 3.

1.3 Multiplicative Structures and Dirichlet Generating Functions

While the last two kinds of generating functions are convenient in handling problems of additive nature (in terms of the sizes of assembled objects), we now turn to multiplicative problems. Dirichlet generating functions are typically the tool of choice in this case. Consider the product $C(z) = A(z)B(z)$ of two DGFs. We find

$$[n^z]C(z) = \sum_{k|n} a_k b_{n/k}.$$

This looks similar to the unlabelled case with OGFs but instead of adding sizes they are multiplied.

In the following we call a combinatorial class multiplicative if it contains no object of size 0 and we intend to translate it into a DGF. The exclusion of 0 is a natural choice here since "multiplication" with such objects would void the result. Contrary to the additive case we call a class containing a single object of size 1 a neutral class and we call a class with a single prime-sized object an atom.

Unfortunately, constructing multiplicative classes using methods similar to those in the last sections is harder, because substituting Dirichlet series into one another will not always result in Dirichlet series (or at least not in a simple manner). Nevertheless, we introduce now a few basic operations on them.

- The sum of two multiplicative classes is defined exactly as in the unlabelled case, i.e. as their disjoint union, using the same size function.
- For two multiplicative classes $\mathcal{A}$ and $\mathcal{B}$ we define their product to be the cartesian product
$$\mathcal{A} \cdot \mathcal{B} = \mathcal{A} \times \mathcal{B}.$$
with size function $|(\alpha, \beta)| = |\alpha|_{\mathcal{A}} |\beta|_{\mathcal{B}}$.
- Let $\mathfrak{A} = \{\mathcal{A}^{(p)}\}_{p \text{ prime}}$ be a family of multiplicative classes, such that each $\mathcal{A}^{(p)}$ contains only objects of size p^k, $k \in \mathbb{N}_0$ We define their Euler product to be
$$\mathrm{Euler}(\mathfrak{A}) = \prod_{p \text{ prime}} \mathcal{A}^{(p)}.$$

- If a multiplicative class $\mathcal{A}$ contains no neutral objects, then we define the class of all finite sequences over $\mathcal{A}$ as

$$\operatorname{Seq}(\mathcal{A}) = \mathcal{E} + \mathcal{A} + \mathcal{A}^2 + \dots,$$

where $\mathcal{E}$ is a neutral class.

The reason to explicitly introduce the Euler product construction is that it allows to break down a construction to atoms or powers of atoms, whenever certain conditions are met:

Lemma 1.11. *Consider a multiplicative class $\mathcal{A}$. If there exist combinatorial isomorphisms $\mathcal{A}_{nm} \to \mathcal{A}_n \cdot \mathcal{A}_m$ for all pairs of coprime positive integers, then*

$$\mathcal{A} = \operatorname{Euler}\left(\left\{\sum_{k=0}^{\infty} \mathcal{A}_{p^k}\right\}_{p\ prime}\right).$$

In particular, if the condition holds for arbitrary positive integers, we have

$$\mathcal{A} = \operatorname{Euler}\left(\{\operatorname{Seq}(\mathcal{A}_p)\}_{p\ prime}\right).$$

Proof. This is a simple translation to the language of combinatorial classes of the fact, that a sum $\sum_{n=0}^{\infty} f(n)$ formally is equal to $\prod_{p \text{ prime}} \sum_{k=0}^{\infty} f(p^k)$ or $\prod_{p \text{ prime}} \frac{1}{1-f(p)}$ if f is multiplicative or totally multiplicative, respectively (easily verified with the fundamental theorem of arithmetic). ∎

Using the notation $X = p^{-z}$ we can write

$$\sum_{k=0}^{\infty} a_{p^k} (p^k)^{-z} = \sum_{k=0}^{\infty} a_{p^k} X^k,$$

which enables us to view the "local" Euler factors of DGFs as ordinary generating functions only dependant on the prime p.

Finally, we calculate the generating functions for all constructions as usual. Neutral classes translate to 1 and atoms to p^{-z} where p is the size (a prime).

Theorem 1.12. *(Multiplicative Dictionary) Using the above notation and conditions the following translations into DGFs hold:*

$$\begin{aligned}
\mathcal{A} + \mathcal{B} &\xrightarrow{DGF} A(z) + B(z) \\
\mathcal{A} \cdot \mathcal{B} &\xrightarrow{DGF} A(z)B(z) \\
\operatorname{Euler}(\mathfrak{A}) &\xrightarrow{DGF} \prod_{p\ prime} A^{(p)}(z) \\
\operatorname{Seq}(\mathcal{A}) &\xrightarrow{DGF} \frac{1}{1 - A(z)}.
\end{aligned}$$

∎

Example 1.13 (Measuring subgroup growth)**.** *Let G be a group and $\mathcal{S}$ the set of subgroups with finite index in G. If there are only finitely many subgroups of each index, $\mathcal{S}$ is a multiplicative class and its DGF $S(z)$ is often referred to as its "subgroup zeta function" (usually denoted by ζ_G in the literature). In view of Lemma 1.11 we are especially interested in groups with multiplicative counting sequences. One particular large class of groups which satisfy these conditions are nilpotent, torsionfree, finitely-generated groups, or $\mathcal{T}$-groups for short.*

A group G is called nilpotent if the lower central series $G_0 \trianglelefteq G_1 \trianglelefteq \dots$ consisting of $G_0 = G$ and commutators $G_{i+1} = [G_i, G]$ terminates, i.e. for some n we have that G_n is trivial. Note that nilpotency can be seen as a generalization of abelian: the series terminates in $n \leq 1$ if and only if the group is abelian. Besides the fact that quotients and subgroups are nilpotent again, the most important property is that finite nilpotent groups are isomorphic to the direct product of their Sylow subgroups (see for instance [Hun74] Chapter II Section 7). This entails the wanted multiplicativity:

Indeed, for any positive integer $n = p_1^{e_1} \dots p_k^{e_k}$ the subgroup $\bigcap \mathcal{S}_n$ ($\mathcal{S}_n$ is finite) is of finite index in G and contains a normal subgroup N, which is itself of finite index. The finite group G/N is nilpotent and contains the full subgroup information of G for indices less or equal n. Thus, without loss of generality we may assume that G is finite and we write

$$G \cong P_1 \times \dots \times P_k \times P',$$

where the P_i are p_i-Sylow subgroups and P' is a direct product of the remaining Sylow subgroups. A subgroup $H \leq G$ of index n corresponds to $(H \cap P_1) \times \dots \times (H \cap P_k) \times P'$ and clearly the subgroup $P_1 \times \dots \times (H \cap P_i) \times \dots \times P_k \times P'$ is of index $p_i^{e_i}$. Let H_i be the subgroup of G which corresponds to the latter. The claim then follows from the fact that the map

$$\mathcal{S}_n \to \mathcal{S}_{p_1^{e_1}} \times \dots \times \mathcal{S}_{p_k^{e_k}}, H \mapsto (H_1, \dots, H_k)$$

is a bijection.

So we have

$$\mathcal{S} = \mathrm{Euler}\left(\left\{ \sum_{k=0}^{\infty} \mathcal{S}_{p^k} \right\}_{p \ prime} \right),$$

thereby reducing the problem to computing the "local" Euler factors. This, however, is still a very hard problem. It has been proven in [GSS88] that the local Euler factors are in fact rational functions in p and p^{-z}. Moreover, in [CSV19] the author explicitely compute these rational functions in the case of nilpotency class 2.

Example 1.14 (Sublattices)**.** *A lattice is an additive subgroup of $\mathbb{R}^n$ which is isomorphic to $\mathbb{Z}^n$. The latter is a simple example of a $\mathcal{T}$-group as it is abelian. We let $\mathcal{L}$ denote the multiplicative class of subgroups of finite index in $\mathbb{Z}^n$. As explained in Example 1.13 we fix a prime p and compute the local DGFs first.*

Let L be an n-by-n matrix over $\mathbb{Z}$. We denote the $\mathbb{Z}$-span its columns of L by $\Lambda(L) \leq \mathbb{Z}^n$. Geometrically it is clear, that $\det(L) = [\mathbb{Z}^n : \Lambda(L)]$ if L is invertible. Moreover, two lattice matrices L and M generate the same lattice iff their Hermite normal forms coincide, i.e. there are invertible matrices $U, V \in \mathrm{GL}_n(\mathbb{Z})$ such that

$$LU = MV = \begin{bmatrix} d_1 & b_{12} & \dots & b_{1n} \\ 0 & d_2 & & b_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & d_n \end{bmatrix}$$

with $d_i > 0$ and $0 \leq b_{ij} < d_j$. In other words, every lattice is uniquely described by going through each dimension $1 \leq k \leq n$ and choosing a vector $v_k \in \mathbb{Z}^k_{\geq 0}$ such that the last entry is the largest. This motivates the following considerations:

Let $\mathcal{V}^{(k)}$ be the class of all such vectors of dimension k with size function set to their maximum norm (always their last entry). For the atoms we find $\mathcal{V}^{(k)}_p = \{[a_1, \dots, a_{k-1}, p] : 0 \leq a_i < p\}$, hence $v^k_p = p^{k-1}$. Note that we can regard this as the cartesian product $[0, p)^{k-1}$ of integer intervals (an integral hypercube of dimension $k-1$ with side length p). Then, for the size p^2 we find $\mathcal{V}^{(k)}_{p^2} = \{[a_1, \dots, a_{k-1}, p^2]^T : 0 \leq a_i < p^2\}$ (the scaled *hypercube with new side length* p^2*), which is combinatorially isomorphic to the set* $(\mathcal{V}_p^{(k)})^2 = \{([a_1,\ldots,a_{k-1},p]^T,[b_1,\ldots,b_{k-1},p]^T) : 0 \le a_i < p, 0 \le b_j < p\}$*. Inductively we find that the class of all* p*-power-sized objects in* $\mathcal{V}^k$ *is* $\mathrm{Seq}(\mathcal{V}_p^{(k)})$*.*

Using this representation for all dimensions $\le n$ *we finally conclude that the class* $\mathcal{Z}^{(n)}$ *of finite-index subgroups in* $\mathbb{Z}^n$ *can be described as*

$$\begin{aligned}
\mathcal{Z}^{(n)} &= \mathrm{Euler}\left(\left\{\mathrm{Seq}(\mathcal{V}_p^{(1)}) \cdot \mathrm{Seq}(\mathcal{V}_p^{(2)}) \cdot \ldots \cdot \mathrm{Seq}(\mathcal{V}_p^{(n)})\right\}_{p\ prime}\right) \\
&\xrightarrow{DGF} Z^{(n)}(z) = \prod_{p\ prime} \frac{1}{1-p^{-z}} \frac{1}{1-pp^{-z}} \cdots \frac{1}{1-p^{k-1}p^{-z}} \\
&= Z^{(1)}(z) Z^{(1)}(z-1) \ldots Z^{(1)}(z-(n-1)).
\end{aligned}$$

Chapter 2

Discrete Probability

In this chpater (based on [FS09] Chapter III and Appendix A.3) we introduce concepts to analyse combinatorial structures with respect to more specific parameters than just how many objects there are of a specific size. For example, a typical use case could be to analyse how many components of size k an object of size n is built of. More generally, one might also ask about how many components can be expected in a *random* object. Hence, we need methods from discrete probability theory, which shall be introduced in the next section.

Throughout this chapter, we use the notations $w_n \in \{1, n!\}$ and $\xrightarrow{*GF}$ in order to develop the theory for OGFs and EGFs simultaneously.

2.1 Basic concepts

Definition 2.1. *Let $\mathcal{S}$ be a finite set. We call a function $\mathbb{P} : \mathcal{S} \to \mathbb{R}$ defined by*

$$\mathbb{P}(\sigma) = p_\sigma$$

a <u>(discrete) probability measure</u> of $\mathcal{S}$ if all $p_\sigma \in \mathbb{R}_{\geq 0}$ add up to 1. The probability of a subset $\mathcal{E} \subseteq \mathcal{S}$ is measured by

$$\mathbb{P}\{\mathcal{E}\} = \sum_{\sigma \in \mathcal{E}} \mathbb{P}(\sigma).$$

In particular, if $\mathbb{P}(\sigma) = \frac{1}{\#\mathcal{S}}$ for each $\sigma \in \mathcal{S}$, we call $\mathbb{P}$ the <u>uniform probability measure</u> of $\mathcal{S}$.

Definition 2.2. *A <u>(discrete) random variable</u> of a finite set $\mathcal{S}$ with probability measure $\mathbb{P}$ is a function $X : \mathcal{S} \to \mathbb{N}_0$. We write*

$$\mathbb{P}\{X = k\} = \sum_{X(\sigma)=k} \mathbb{P}(\sigma)$$

for the probability distribution associated to X.

Definition 2.3. *Let X be a random variable of $(S, \mathbb{P})$ and let $f : \mathbb{N}_0 \to \mathbb{R}$ be a function. The <u>expectation</u> of $f(X)$ is defined by the linear functional*

$$\mathbb{E}(f(X)) = \sum_{k \in \mathbb{N}_0} \mathbb{P}\{X = k\} f(k). \tag{2.1}$$

Important examples are the <u>mean</u> with $f(X) = X$ and the <u>variance</u> $\mathbb{V}(X)$ given by $f(X) = (X - \mathbb{E}(X))^2$.

Note that the variance might also be expressed as

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(2\mu X) + \mathbb{E}(\mu^2) = \mathbb{E}(X^2) - 2\mu^2 + \mu^2 = \mathbb{E}(X^2) - \mu^2$$

with $\mu = \mathbb{E}(X)$.

2.2 Probability Generating Functions

Let us now see how these concepts integrate into the theory of generating functions. Due to the fundamentally different treatment of DGFs, only power series generating functions are considered in this section.

Definition 2.4. *A parameter of a combinatorial class $\mathcal{A}$ is a function $\chi : \mathcal{A} \to \mathbb{N}_0$. Using the extended counting sequence*

$$a_{n,k} = \#\{\alpha \in \mathcal{A} : |\alpha| = n, \chi(\alpha) = k\}$$

we define the bivariate generating function (BGF) of $(\mathcal{A}, \chi)$ to be the formal power series in two indeterminants z and u

$$A(z, u) = \sum_{n,k \geq 0} a_{n,k} \frac{z^n}{w_n} u^k.$$

In general, combinatorial classes do not need to be finite. Thus, parameters are not the same as random variables. However, since all $\mathcal{A}_n \subseteq \mathcal{A}$ are always finite, a parameter induces a random variable:

Let $\mathcal{A}$ be a combinatorial class with some parameter χ. We denote the uniform probability measure over $\mathcal{A}_n$ by $\mathbb{P}_{\mathcal{A}_n}$. Considering χ as a random variable over $\mathcal{A}_n$ we have

$$\mathbb{P}_{\mathcal{A}_n}(\chi = k) = \frac{a_{n,k}}{a_n} \tag{2.2}$$

since $a_n = \sum_{k=0}^{\infty} a_{n,k}$ ($a_{n,k}$ defined as in definition 2.4).

Definition 2.5. *Using the notations above, the probability generating function (PGF) of the parameter χ restricted to $\mathcal{A}_n$ is defined as the formal power series*

$$P_{\mathcal{A}_n}(u) = \sum_{k=0}^{\infty} \mathbb{P}_{\mathcal{A}_n}(\chi = k) u^k, \tag{2.3}$$

where u is a formal indeterminant.

The PGF clearly encodes information about the probability distribution of the parameter. In the next result we extract important pieces of it. It is a very simple consequence of equations (2.1) and (2.3).

Proposition 2.6. *Let $P_{\mathcal{A}_n}(u)$ be the PGF of the restricted parameter χ on $\mathcal{A}_n$. For an integer $k > 0$ and $f(X) = X(X-1)\dots(X-k+1)$ we find*

$$\mathbb{E}(f(\chi)) = \frac{d^k}{du^k} P_{\mathcal{A}_n}(u)\bigg|_{u=1}.$$

In particular (using the linearity of $\mathbb{E}$):

$$\mathbb{E}(\chi) = \frac{d}{du} P_{\mathcal{A}_n}(u)\bigg|_{u=1} \quad \textit{and} \quad \mathbb{E}(\chi^2) = \frac{d^2}{du^2} P_{\mathcal{A}_n}(u)\bigg|_{u=1} + \frac{d}{du} P_{\mathcal{A}_n}(u)\bigg|_{u=1}. \qquad \blacksquare$$

Furthermore, applying equation (2.2) yields:

Proposition 2.7 (Connection of BGF and PGF)**.** *Let $\mathcal{A}$ be a combinatorial class and χ a parameter. For the BGF $A(z, u)$ of $(\mathcal{A}, \chi)$ and the PGF $P_{\mathcal{A}_n}(u)$ of the restriction of χ to $\mathcal{A}_n$ we find*

$$P_{\mathcal{A}_n}(u) = \frac{[z^n]A(z, u)}{[z^n]A(z, 1)}. \qquad \blacksquare$$

In the next section we will develop a method to easily compute BGFs. Thus, the last proposition allows us to make statements about the probability distribution of the regarded parameter directly from the BGF.

2.3 Markers

As already noted above, a main part of this whole chapter is to analyze the building blocks of combinatorial constructions. In order to *mark* whatever components of a structure we would like to count, we extend the theory of combinatorial constructions for OGFs and EGFs.

Definition 2.8. *Suppose the class $\mathcal{A}$ is constructed from another class $\mathcal{B}$ by*

$$\mathcal{A} = \mathcal{K}(\mathcal{B}),$$

where $\mathcal{K} \in \{\mathrm{Seq}, \mathrm{MSet}\}$ for OGFs and $\mathcal{K} \in \{\mathrm{Seq}, \mathrm{Set}, \mathrm{Cyc}, \mathrm{UCyc}\}$ for EGFs. A marker on the components $\mathcal{B}$ in $\mathcal{A}$ is a class μ attached to $\mathcal{B}$ by product (unlabelled or labelled), which translates (in OGF or EGF sense) to the new formal indeterminant u making the GF bivariate. In symbols,

$$\mathcal{A} = \mathcal{K}(\mu\mathcal{B}) \xrightarrow{*GF} A(z,u) = \sum_{n,k=0}^{\infty} a_{n,k} \frac{z^n}{w_n} u^k.$$

Note that markers do not have any influence on the size. They give rise to parameters. Indeed, the sets $\mathcal{A}_{n,k}$ partition $\mathcal{A}$, so setting $\chi(\alpha) = k$ for all $\alpha \in \mathcal{A}_{n,k}$ yields a well-defined parameter χ on $\mathcal{A}$. With the induced random variables χ (the restrictions to $\mathcal{A}_n$) we find that $\mathbb{P}_{\mathcal{A}_n}(\chi = k)$ measures the probability of a random size-n-object in $\mathcal{A}$ having k markers, e.g. consisting of k components from $\mathcal{B}$.

Example 2.9 (Continuation of Example 1.5)**.** *An interesting question concerning polynomials over finite fields would be: how many irreducible factors can a random polynomial of degree n be expected to have? We have developed all the tools necessary to tackle this problem combinatorially.*

Let μ be marking the $\mathcal{I}$-components in $\mathcal{P}$. Written out we have,

$$\mathcal{P} = \mathrm{MSet}(\mu \times \mathcal{I}) \xrightarrow{OGF} P(z,u) = \exp\left(\sum_{k=1}^{\infty} \frac{1}{k} u^k I(z^k)\right).$$

Using Propositions 2.6 and 2.7 we find for the induced random variable χ on $\mathcal{P}_n$ that

$$\begin{aligned}\mathbb{E}_{\mathcal{P}_n}(\chi) &= \frac{\partial}{\partial u} \frac{[z^n]P(z,u)}{[z^n]P(z)}\bigg|_{u=1} = p^{-n}[z^n]\, \frac{\partial}{\partial u} P(z,u)\bigg|_{u=1} \\ &= p^{-n}[z^n]P(z)(I(z) + R_1(z))\end{aligned}$$

with $R_1(z) = \sum_{j=2}^{\infty} I(z^j)$. This gives us the mean of the number of components in $\mathcal{P}$, thus answers our question after extracting coefficients.

The variance follows similarly. First,

$$\begin{aligned}\mathbb{E}_{\mathcal{P}_n}(\chi^2) &= \mathbb{E}_{\mathcal{P}_n}(\chi(\chi-1)) + \mathbb{E}_{\mathcal{P}_n}(\chi) \\ &= p^{-n}[z^n]P(z)\,(I(z) + R_2(z)) + p^{-n}[z^n]P(z)(I(z) + R_1(z)) \\ &= p^{-n}[z^n]P(z)(2I(z) - R_1(z) + R_2(z)),\end{aligned}$$

where $R_2(z) = \sum_{j=2}^{\infty} kI(z^j)$. Plugging this into the formula of Proposition 2.6 gives us the variance. In order to get an idea of how mean and variance behave with increasing n we further pursue this in Example 3.8.

Part II

Analytic Coefficient Extraction

Chapter 3

Power Series

In the case of power series generating functions it is very natural to consider complex-analytic methods. When interpreting them as functions $\mathbb{C} \to \mathbb{C}$ they are automatically holomorphic around the origin, if their radius of convergence is positive.

Therefore, the main analytic tool of this chapter will be the Cauchy integral formula. It states that, if $f : U \to \mathbb{C}$ is analytic in a domain U, $a \in U$ and γ is a loop in U surrounding a once and with positive orientation, then

$$f^{(n)}(a) = \frac{n!}{2\pi i} \int_\gamma f(z) \frac{dz}{(z-a)^{n+1}}.$$

As a simple corollary we can express the coefficient extraction operator $[z^n]$ as a path integral using

Theorem 3.1 (Cauchy's coefficient formula)**.** *In the above setting, where now γ surrounds 0, we find*

$$[z^n]f(z) = \frac{1}{2\pi i} \int_\gamma f(z) \frac{dz}{z^{n+1}}$$

Proof. Since f coincidences locally with its Taylor series, we have

$$[z^n]f(z) = \frac{f^{(n)}(0)}{n!}.$$

Therefore applying the integral formula with $a = 0$ proves the claim. ■

Before we proceed to the technique of singularity analysis we need some properties of the Gamma function.

3.1 The Gamma Function

In this section we will omit all proofs and instead refer to the books [Lan99] Chapter XV \$2 and [WW96] Chapter 12 Section 22.

The limit

$$\Gamma(z) = \lim_{n\to\infty} \frac{n!n^z}{z(z+1)\dots(z+n)} \tag{3.1}$$

defines a mermorphic function on $\mathbb{C} \setminus \mathbb{Z}_{\leq 0}$, called the Gamma function. The poles at $-n$, $n \in \mathbb{N}_0$ are of order 1 with residues $\frac{-1}{n!}$. It satisfies the functional equation

$$\Gamma(z+1) = z\Gamma(z)$$

and, therefore, interpolates the factorial since $\Gamma(n) = (n-1)!$ for all $n \in \mathbb{N}$.

We may extend the notion of a binomial coefficient using

$$\binom{\alpha}{k} = \frac{\Gamma(\alpha+1)}{\Gamma(k+1)\Gamma(\alpha-k+1)}$$

for $\alpha \in \mathbb{C} \setminus \mathbb{Z}_{<0}$. The binomial theorem extends to this as well:

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k. \tag{3.2}$$

This series converges absolutely whenever $|x| < 1$. From equation (3.1) it is also possible to find bounds for the binomial coefficient:

$$\left|\binom{\alpha}{k}\right| \leq C \frac{1}{k^{1+\operatorname{Re}\alpha}}. \tag{3.3}$$

Other representation of the function are obtained through the Mellin transform

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt, \tag{3.4}$$

which converges for $\operatorname{Re} z > 0$, or via contour integration with a Hankel contour $\mathcal{H}$ (see Figure 3.1a):

$$\frac{1}{\Gamma(z)} = -\frac{1}{2\pi i} \int_{\mathcal{H}} (-t)^{-z} e^{-t} dt. \tag{3.5}$$

3.2 Singularity Analysis

In this section we will see that singularities of functions encode useful asymptotic information about the sequence of coefficients of their Taylor series in the origin. It is based on [FS09] Chapter VI and the paper [FO90]. In fact, the case of meromorphic functions is quite straightforward using Cauchy's coefficient formula. One is able to find asymptotic expansions with exponentially small error terms. Unfortunately though, the most interesting examples possess essential singularities. We will, thus, need to develop a theory for a much broader class of functions.

Definition 3.2. *Assume $f(z)$ has radius of convergence $R > 0$. A singularity of f with modulus R is called dominant.*

We focus on functions that have exactly one dominant singularity. An extension to a finite number is possible, but it is not necessary for our purposes.

There are two main principles that will be established:

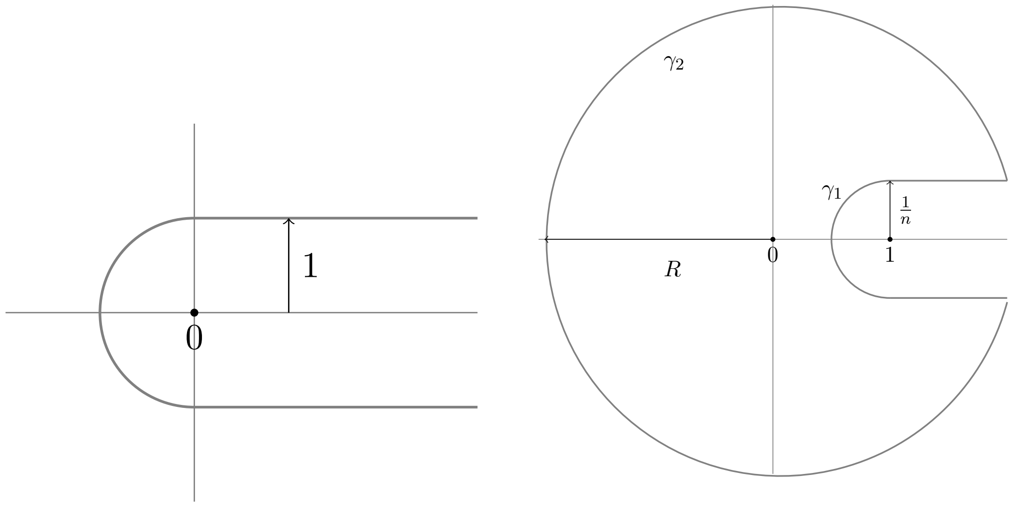
1. The exponential growth behavior of the sequence of coefficients is determined by the location of the dominant singularity.

2. The singularity type determines sub-exponential behavior.

Our first result aims to set up a catalogue of different singularity types and how they relate to the coefficients of a function possessing such a singularity. In order to simplify the proof we will first assume the dominant singularity at 1 and without any other interfering terms.

Proposition 3.3 (Function scale)**.** *Consider the function*

$$f(z) = \left(\frac{1}{1-z}\right)^\alpha \left(\frac{1}{z} \log \frac{1}{1-z}\right)^\beta$$



(a) Hankel contour $\mathcal{H}$ (b) $\gamma(n, R)$

Figure 3.1

for arbitrary $\alpha \in \mathbb{C} \setminus \mathbb{Z}_{\leq 0}$ *and* $\beta \in \mathbb{C}$. *The sequence of coefficients of the Taylor series expansion at the origin can be expanded asymptotically for large* n *as*

$$[z^n]f(z) \sim \frac{n^{\alpha-1}(\log n)^\beta}{\Gamma(\alpha)} \left(1 + O\left(\frac{1}{\log n}\right)\right).$$

To give the complex exponents and logarithms a meaning we choose the principal branch cutting the complex plane at the line $\mathbb{R}_{\leq 0}$. The factor $\frac{1}{z}$ was only introduced to make f analytic in the origin for all exponents β. In fact, it can be dropped without changing the asympotic expansion around $z = 1$ (see [FO90] Corollary 6 for details).

Proof. Let $\gamma = \gamma(n, R)$ be a simple loop, which winds around the singularity at distance $1/n$ and is closed by an arc with center 0 and radius R with subdivision $\gamma = \gamma_1 \sqcup \gamma_2$ as indicated in Figure 3.1b. Using Cauchy's coefficient formula we write

$$f_n = f_n^{(1)} + f_n^{(2)} = [z^n]f(z) = \frac{1}{2\pi i} \int_\gamma f(z) \frac{dz}{z^{n+1}},$$

where $f_n^{(j)}$ is the integral over γ_j.

Since on the surrounding circle γ_2 we have

$$f_n^{(2)} = O(R^{-n}), \tag{3.6}$$

its contribution to the sum is negligible, e.g. much smaller than the wanted precision, and we only need to consider the open contour γ_1 that is left.

Now we apply the transformation

$$z = 1 + \frac{t}{n}$$

to obtain the new contour $\gamma_1' = \gamma_1'(n, R)$, which now winds around 0 keeping distance 1 to the interval $[0, Rn]$:

$$
\begin{aligned}
f_n^{(1)} &= \frac{1}{2\pi i}\frac{1}{n}\int_{\gamma_1'} f\left(1+\frac{t}{n}\right)\left(1+\frac{t}{n}\right)^{-n-1} dt \\
&= \frac{1}{2\pi i} n^{\alpha-1}\int_{\gamma_1'} (-t)^{-\alpha} \log\left(-\frac{n}{t}\right)^{\beta}\left(1+\frac{t}{n}\right)^{-n-1-\beta} dt \\
&= \frac{1}{2\pi i} n^{\alpha-1}(\log n)^{\beta}\int_{\gamma_1'} (-t)^{-\alpha}\left(1-\frac{\log(-t)}{\log(n)}\right)^{\beta}\left(1+\frac{t}{n}\right)^{-n-1-\beta} dt \\
&= \frac{1}{2\pi i} n^{\alpha-1}(\log n)^{\beta}\int_{\gamma_1'} (-t)^{-\alpha} e^{-t}\left(1-\frac{\log(-t)}{\log(n)}\right)^{\beta} g(t,n) dt,
\end{aligned}
$$

where

$$
g(t,n) = e^{t}\left(1+\frac{t}{n}\right)^{-n-1-\beta}.
$$

Obviously, we have $g(t,n) \to 1$ as $n \to \infty$. Therefore, by dominated convergence, we can approximate the integral:

$$
f_n^{(1)} \sim \frac{1}{2\pi i} n^{\alpha-1}(\log n)^{\beta}\int_{\gamma_1'} (-t)^{-\alpha} e^{-t}\left(1-\frac{\log(-t)}{\log(n)}\right)^{\beta} dt.
$$

Next, we split the contour γ_1' into η_1 and η_2, such that η_1 contains all points t with $\operatorname{Re} t \le (\log n)^2$ and η_2 contains all points with $\operatorname{Re} t \ge (\log n)^2$. We denote their integrals by $f_n^{(1,1)}$ and $f_n^{(1,2)}$, respectively, but the latter is in fact negligible: for large n each point in η_2 is of the form $\exp\big(-\lambda(\log n)^2\big)$, $\lambda \ge 1$, hence

$$
f_n^{(1,2)} = O(\exp\big(-(\log n)^2\big)).
$$

Because of this and (3.6), we may let R tend to $+\infty$ to obtain a proper Hankel-type contour $\mathcal{H}$ as in Figure 3.1a, while still having

$$
f_n = f_n^{(1)} + f_n^{(2)} = f_n^{(1,1)} + f_n^{(1,2)} + f_n^{(2)} \sim f_n^{(1,1)}.
$$

Along η_1 we know that $|log(-t)| < \log n$ for large n (follows from $\log(|t|) \le \log\big((\log n)^2\big)$), so we may apply the generalized binomial theorem (see equation (3.2) above) to obtain the (pointwise) convergent series expansion

$$
\left(1-\frac{\log(-t)}{\log(n)}\right)^{\beta} = \sum_{k=0}^{\infty}\binom{\beta}{k}(-1)^k\left(\frac{\log(-t)}{\log(n)}\right)^{k}.
$$

After plugging this back into the integral, resulting in the form

$$
f_n^{(1,1)} \sim C\int_{\eta_1}\sum_{k=0}^{\infty} h_{n,k}(t) dt,
$$

we would like to interchange infinite summation and integration. Because

$$
h_{n,k}(t) = (-1)^k(-t)^{-\alpha}e^{-t}\binom{\beta}{k}\left(\frac{\log(-t)}{\log(n)}\right)^{k}
$$

is continuous on η_1, hence integrable, we only need to show uniform convergence of the series:

We first observe that the factor $|(-t)^{-\alpha}e^{-t}|$ is bounded by a constant $C = C(n)$ on η_1. For the next factor there is a constant D such that

$$\left|\binom{\beta}{k}\right| \le \frac{D}{k^{1+\operatorname{Re}\beta}}$$

using estimate (3.3) from the last section. Lastly, there is also a constant $E < 1$ such that for large n we get the estimate

$$|\log(-t)| \le \log\bigl((\log n)^2\bigr) < E \log n.$$

Together these three estimates show that for all $t \in \eta_1$ and n sufficiently large, we have

$$\sum_{k=0}^{\infty} h_{n,k}(t) \le CD \sum_{k=0}^{\infty} \frac{1}{k^{1+\operatorname{Re}\beta}} E^k < \infty.$$

This, in fact, satisfies the Weierstraß M-test (see [Pro16b]) giving us uniform convergence.

We may now switch integration and summation, and, after utilizing the contour integral representation 3.5 of the reciprocal gamma function derived in the last section, we finally obtain

$$\begin{aligned} f_n &\sim \frac{1}{2\pi i} n^{\alpha-1}(\log n)^{\beta} \sum_{k=0}^{\infty}(-1)^k \binom{\beta}{k} (\log n)^{-k} \int_{\mathcal{H}} (-t)^{-\alpha} e^{-t} (\log(-t))^k dt \\ &= n^{\alpha-1}(\log n)^{\beta} \sum_{k=0}^{\infty} \frac{1}{(\log(n))^k} \binom{\beta}{k} \frac{\partial^k}{\partial \alpha^k} \frac{1}{\Gamma(\alpha)}. \end{aligned} \tag{3.7}$$

■

Note that the full asymptotic expansion in (3.7) can be used, whenever more precision is needed.

This proposition hints to the second principle, which we stated earlier. However, clearly not all generating functions appearing in practice are of the required form. A tool to translate the occuring error terms is necessary.

Definition 3.4. *For parameters $\zeta \in \mathbb{C} \setminus \{0\}$, $R > |\zeta|$, $0 < \phi < \frac{\pi}{2}$, a Pacman domain $\mathcal{P}(\zeta, R, \phi)$ is defined by the domain*

$$\mathcal{P} = \{z \in \mathbb{C} : |z| < R, |\arg(z/\zeta - 1)| > \phi\}.$$

A function is called Pacman-analytic if it is analytic in some Pacman domain $\mathcal{P}(\zeta, R, \phi)$.

Proposition 3.5 (Error transfer)**.** *Let f be a $(1, R, \theta)$-Pacman-analytic function, that satisfies*

$$|f(z)| \le A \left| \left(\left(\frac{1}{1-z}\right)^{\alpha} \left(\log \frac{1}{1-z}\right)^{\beta} \right) \right|$$

inside its Pacman-domain for arbitrary $\alpha, \beta \in \mathbb{R}$. Then we can approximate its sequence of coefficients by

$$[z^n]f(z) = O(n^{\alpha-1}(\log n)^{\beta}).$$

for large n.

Proof. First, we note that due to the Pacman analyticity of f Again, we use the coefficient formula. In order to capture the properties of the critical point in 1 we choose loops $\gamma(n)$

which slowly approach it. With parameters $1 < r < R$ and $\phi < \theta < \pi/2$ the individual parts are explicitly described as follows:

$$\begin{aligned}
c_1(n) &= \left\{ |z-1| = \frac{1}{n}, |\arg(z-1)| \geq \theta \right\} \\
c_2(n) &= \{ |z| = r, |\arg(z-1)| \geq \theta \} \\
l_1(n) &= \left\{ \lambda e^{i\theta} + 1, \lambda \in [1/n, r] \right\} \\
l_2(n) &= \overline{l_1(n)}
\end{aligned}$$

We start with the inner circle part $c_1(n)$. Since $|z-1| = 1/n$ we have

$$\begin{aligned}
|f(z)| &\leq A n^\alpha (\log n)^\beta \\
|z^{-n-1}| &\leq B
\end{aligned}$$

and the contour of integration is bounded by $2\pi/n$. In total:

$$\frac{1}{2\pi i} \int_{c_1(n)} f(z) \frac{dz}{z^{n+1}} \leq C \frac{1}{n} n^\alpha (\log n)^\beta = O(n^{\alpha-1} (\log n)^\beta)$$

for constants A, B, C.

The outer circle part is even easier. Here, $|f(z)|$ is bounded since $|z| = r > 1$, while z^{-n-1} is $O(r^{-n-1})$, therefore irrelevant in the claimed scale of the problem.

Now we process $l_1(n)$. The integral over this part shall be denoted by L_n. We apply the affine-linear transformation $t = e^{-i\theta} n(z-1)$. Reading from the right, the line's starting point is first moved onto the circle with radius $1/n$, then scaled onto the unit circle (the line now has length $nr - 1$), and finally rotated onto the real line. Thus, in the integral we get

$$\begin{aligned}
L_n &= \frac{1}{2\pi i} \int_{l_1(n)} f(z) \frac{dz}{z^{n+1}} = \frac{1}{2\pi i} \int_1^{rn} f\left(1 + \frac{te^{i\theta}}{n}\right) \left(1 + \frac{te^{i\theta}}{n}\right)^{-n-1} \frac{e^{i\theta}}{n} dt \\
&\leq \frac{1}{2\pi} \int_1^{rn} \left| f\left(1 + \frac{te^{i\theta}}{n}\right) \right| \left| 1 + \frac{te^{i\theta}}{n} \right|^{-n-1} \frac{1}{n} dt. \qquad (3.8)
\end{aligned}$$

By assumption there is a constant C such that

$$\left| f\left(1 + \frac{te^{i\theta}}{n}\right) \right| \leq C \left(\frac{n}{t}\right)^\alpha \left| \log\left(\frac{n}{te^{i\theta}}\right) \right|^\beta.$$

Also, the estimate

$$\left| \log \frac{n}{te^{i\theta}} \right|^\beta \leq \max_{t \in [1, rn]} \left| \log \frac{n}{t} - \theta \right|^\beta = O((\log n)^\beta)$$

can be easily verified by distincting cases where $\beta \leq 0$ or $\beta \geq 0$ and choosing r very near to 1 (how near depends only θ, therefore possible without loss of generality). Using these considerations inside (3.8) leads to

$$L_n = O(n^{\alpha-1} (\log n)^\beta) J_n,$$

where for J_n we have

$$\begin{aligned}
J_n &= \int_1^{rn} t^{-\alpha} \left| 1 + \frac{te^{i\theta}}{n} \right|^{-n-1} dt \leq \int_1^\infty t^{-\alpha} \left(1 + \frac{t}{n} \operatorname{Re} e^{i\theta}\right)^{-n-1} dt \\
&= \int_1^\infty t^{-\alpha} \left(1 + \frac{t\cos\theta}{n}\right)^{-n-1} dt \leq \int_1^\infty t^{-\alpha} e^{-t\cos\theta} dt < \infty.
\end{aligned}$$

The second estimate can be followed from the monotone convergence theorem and the last integral exists since $0 < \cos\theta < 1$.

The remaining line works just like this, when exchanging θ for $-\theta$. Thus, the claim follows after putting all four parts together. ∎

Now we have all ingredients together to formulate the central theorem of singularity analysis (with one dominant singularity) and, thereby, give an exact formulation of the two principles from the last section.

Theorem 3.6 (Singularity analysis)**.** *Define the two sets*

$$\begin{aligned} S &= \{f(z) = \left(\frac{1}{1-z}\right)^{\alpha}\left(\log\frac{1}{1-z}\right)^{\beta} : \alpha \in \mathbb{C}\setminus\mathbb{Z}_{\leq 0}, \beta \in \mathbb{C}\} \\ E &= \{f(z) = \left(\frac{1}{1-z}\right)^{\alpha}\left(\log\frac{1}{1-z}\right)^{\beta} : \alpha, \beta \in \mathbb{R}\}. \end{aligned}$$

Let f be a Pacman analytic function eating the singularity ζ of f. Assume that there exist a finite linear combination s of functions in S and a function $e \in E$ such that

$$f(z) = s\left(\frac{z}{\zeta}\right) + O\left(e\left(\frac{z}{\zeta}\right)\right)$$

as $z \to \zeta$. Then:

$$[z^n]f(z) \sim \zeta^{-n}s^*(n) + O(\zeta^{-n}e^*(n)),$$

where s^* and e^* have their coefficients determined in the obvious way by the last two propositions.

Proof. If we define $g(z) = f(z/\zeta)$, the singularity has been moved to 1. Thus, we may apply the propositions to compute s^* and e^*. Finally, we note that $[z^n]f(z) = \zeta^{-n}[z^n]g(z)$. ∎

Example 3.7 (Continuation of Example 1.10)**.** *We found the EGF*

$$R(z) = \frac{\exp\left(-\frac{z}{2} - \frac{z^2}{4}\right)}{\sqrt{1-z}}.$$

The singularity at $z = 1$ appears in the denominator while the numerator is, in fact, an entire function. By expanding the numerator at $z = 1$ we find

$$R(z) = \frac{e^{-3/4}}{\sqrt{1-z}} + e^{-3/4}\sqrt{1-z} + O((1-z)^{3/2}).$$

Then, we apply Theorem 3.6 to obtain

$$n![z^n]R(z) = n!\left(\frac{e^{-3/4}}{\sqrt{\pi n}} - \frac{e^{-3/4}}{2\sqrt{\pi n^3}} + O\left(\frac{1}{n^{5/2}}\right)\right).$$

Example 3.8 (Continuation of Examples 1.5 and 2.9)**.** *The OGF $I(z)$ has singularities whenever $z = \sqrt[k]{\frac{1}{p}}$. Thus, there is a single dominant singularity at $z = \frac{1}{p}$. We write*

$$I(z) = \log\frac{1}{1-pz} + R(z).$$

with $R(z) = \sum_{k=2}^{\infty} \frac{\mu(k)}{k} \log \frac{1}{1-pz^k}$*, which is clearly holomorphic at* $z = 1/p$*. Thus,* $R(z) = C + O(z)$*. The expansion of the first summand is straighforward and does not require an application of the singularity analysis theorem. In total:*

$$[z^n]I(z) \sim \frac{p^n}{n} + O\left(\frac{1}{n}\right).$$

Applying singularity analysis with more expansion terms (see the proof) to the mean number of irreducible factors in a degree-n polynomial yields

$$\begin{aligned}
\mathbb{E}_{\mathcal{P}_n}(\chi) &= p^{-n}[z^n]P(z)(I(z) + R_1(z)) \\
&= p^{-n}[z^n]\frac{1}{1-pz}\left(\log\left(\frac{1}{1-pz}\right) + R(z) + R_1(z)\right) \\
&= p^{-n}[z^n]\frac{1}{1-pz}\log\left(\frac{1}{1-pz}\right) + \frac{1}{1-pz}(R(z) + R_1(z)) \\
&\sim \log n + \gamma + R(1/p) + R_1(1/p) + O((\log n)^{-1}).
\end{aligned}$$

for z *near* $1/p$*. For the variance we find similarly*

$$\mathbb{V}_{\mathcal{P}_n}(\chi) = O(\log n).$$

Example 3.9 (Continuation of Example 1.6)**.** *This example shows up the limits of this method. Even if we had developed the theory for finitely many dominant singularities, it would fail here. The reason is that the generating function of partitions*

$$P(z) = \prod_{k=1}^{\infty} \frac{1}{1-z^k}.$$

is holomorphic inside the unit disc but the singularities on the unit circle lie dense (the roots of unity). We need different methods here. (for example the saddle-point method, [FS09] Chapter VIII).

Chapter 4

Dirichlet Series

This chapter is based on [Apo76] and [Ten15].

We will now tackle coefficient extraction based on their analytic properties. However, this time we will not attempt to approximate $[n^{-z}]$ directly but instead the summatory function $\sum_{n<t} a_n$. This is due to the fact that we can express the Dirichlet series as an integral transform of this function, and are therefore in a situation, which is very roughly comparable to the power series case with Cauchy's coefficient formula.

The two principles stated in section 3.2 still hold in a sense. Again we will be interested in the domain of convergence, which is a right half-plane of $\mathbb{C}$ in this case, and in the type of the nearest singularity on the real line.

4.1 The Ikehara-Ingham-Delange Theorem

We will derive our main theorem in the next section as a consequence of the following result. It uses a simple lemma from the theory of Fourier transforms, which we w

Lemma 4.1. *Let $g : \mathbb{R} \to \mathbb{C}$ be integrable on $\mathbb{R}$ and bounded and let*

$$\hat{g}(\tau) = \int_{-\infty}^{\infty} g(t) \exp(-2\pi i \tau t) dt$$

be its Fourier transformed.

Assume there is a $T > 0$ such that

$$\sup_{x \leq y \leq x+1/T} (g(y) - g(x)) \leq K \tag{4.1}$$

for some $K < \infty$. Then,

$$\|g\|_\infty \leq 16K + 6 \int_{-T}^{T} |\hat{g}(\tau)| d\tau.$$

Proof. See [Ten15] Theorem 7.15. ■

Theorem 4.2 (Ikehara-Ingham-Delange theorem)**.** *Let A be a real-valued non-decreasing function with $A(t) = 0$ for $t \leq 0$ such that*

$$F(s) = \int_{0}^{\infty} e^{-st} A(t) dt \tag{4.2}$$

converges for all $s \in \mathbb{C}$, $\operatorname{Re} s > a > 0$ and define

$$G(s) = \frac{F(s+a)}{s+a} - \frac{c}{s^{\omega+1}}$$

with constants $c \geq 0$ and $\omega > -1$. If

$$\eta(\sigma, T) = \sigma^\omega \int_{-T}^{T} |G(2\sigma + i\tau) - G(\sigma + i\tau)| d\tau \to 0 \quad \textit{for} \quad \sigma \searrow 0 \tag{4.3}$$

for each $T > 0$, then as $x \to \infty$ we find

$$A(x) \sim \frac{c}{\Gamma(\omega + 1)} e^{ax} x^\omega.$$

Proof. Let us define a few auxiliary functions:

$$\begin{aligned}
g_\sigma(t) &= A(t) e^{-at} e^{-\sigma t}(1 - e^{-\sigma t}) \\
B(t) &= \chi_{(0,\infty)}(t) \frac{c}{\Gamma(\omega + 1)} t^\omega e^{-t}(1 - e^{-t}) \\
G_\sigma(t) &= g_\sigma(t) - \sigma^{-\omega} B(\sigma t) = \left(A(t) e^{-at} - \frac{c}{\Gamma(\omega + 1)} t^\omega \right) e^{-\sigma t}(1 - e^{-\sigma t}),
\end{aligned}$$

where $0 < \sigma \leq 1$ and χ_M is the characteristic function of the set M. The choices for them become reasonable after computing their Fourier transforms:

$$\begin{aligned}
\hat{g}_\sigma(\tau) &= G(\sigma + i\tau) - G(2\sigma + i\tau) + c((\sigma + i\tau)^{-\omega - 1} - (2\sigma + i\tau)^{-\omega - 1}) \\
\hat{G}_\sigma(\tau) &= G(\sigma + i\tau) - G(2\sigma + i\tau).
\end{aligned}$$

Both are easily verified using elementary properties of integration. For the second we also need to use the formula (3.4) for the gamma function.

Now, the idea is to apply Lemma 4.1 to G_σ in order to find that the difference $A(t)e^{-at} - t^\omega c/\Gamma(\omega + 1)$ is in fact in smaller than $Ce^{-at}t^\omega$. This will prove our claim when setting $t = x$ and $\sigma = 1/x$ because then the order of the difference is $o(e^{-ax}x^\omega)$ as $x \to \infty$. We proceed in multiple steps.

1. Next, we shall apply Lemma 4.1 to $g_\sigma(t)$. For the last term in the Fourier transformed we have

$$|c((\sigma + i\tau)^{-\omega - 1} - (2\sigma + i\tau)^{-\omega - 1})| = c(\omega + 1) \left| \int_{\sigma + i\tau}^{2\sigma + i\tau} \frac{ds}{s^{\omega + 2}} \right| \leq \frac{c(\omega + 1)\sigma}{|\sigma + i\tau|^{\omega + 2}},$$

therefore,

$$\begin{aligned}
\int_{-T}^{T} |\hat{g}_\sigma(\tau)| d\tau &\leq \frac{1}{\sigma^\omega} \eta(\sigma, T) + c(\omega + 1)\sigma \int_{-T}^{T} \frac{d\tau}{(\sigma + i\tau)^{\omega + 2}} \\
&\leq \frac{1}{\sigma^\omega} \left(\eta(\sigma, T) + c(\omega + 1)\sigma^{\omega + 1} \int_{-T}^{T} \frac{d\tau}{\max(\sigma, |\tau|)^{\omega + 2}} \right) \\
&\leq \frac{\eta(\sigma, T) + c(\omega + 3)}{\sigma^\omega}.
\end{aligned} \tag{4.4}$$

The last estimate follows from the facts that $w + 2 > 1$ and $\sigma \leq 1$ because

$$\begin{aligned}
&c(\omega + 1)\sigma^{\omega + 1} \int_{-T}^{T} \frac{d\tau}{\max(\sigma, |\tau|)^{\omega + 2}} \leq c(\omega + 1)\sigma^{\omega + 1} \int_{-\infty}^{\infty} \frac{d\tau}{\max(\sigma, |\tau|)^{\omega + 2}} \\
&= c(\omega + 1)\sigma^{\omega + 1} \left(\int_{-\infty}^{-\sigma} \frac{d\tau}{(-\tau)^{\omega + 2}} + \int_{-\sigma}^{\sigma} \frac{d\tau}{\sigma^{\omega + 2}} + \int_{\sigma}^{\infty} \frac{d\tau}{\tau^{\omega + 2}} \right) \\
&= (\omega + 1)\sigma^{\omega + 1} \frac{1}{\omega + 1} \left(\frac{2}{\sigma^{\omega + 1}} + \frac{\omega + 1}{\sigma^\omega} \right) = c(2 + \sigma(\omega + 1)) \leq c(\omega + 3).
\end{aligned}$$

In order for the lemma to be applicable we need to check the condition 4.1. Take $x \geq 0$ and $y > 0$. Using the monotonicity and positivity of A we find

$$
\begin{aligned}
g_\sigma(x+y) - g_\sigma(x) &\geq A(x)e^{-(a+\sigma)x}(e^{-(a+\sigma)y}(1-e^{-\sigma(x+y)}) - (1-e^{-\sigma x})) \\
&= A(x)e^{-(a+\sigma)x}(1-e^{-\sigma x})\left(e^{-(a+\sigma)y}\frac{1-e^{-\sigma(x+y)}}{1-e^{-\sigma x}} - 1\right) \\
&\geq A(x)e^{-(a+\sigma)x}(1-e^{-\sigma x})(e^{-(a+\sigma)y} - 1) = g_\sigma(x)(e^{-(a+\sigma)y} - 1),
\end{aligned}
$$

since the appearing fraction is smaller than 1. The second factor is negative, so we make the expression smaller by taking the supremum in x. Furthermore, $-(a+\sigma)y \leq e^{-(\sigma+a)y} - 1$ for $y \geq 0$ (easily verified by comparing values at zero and derivatives), so that we are left with

$$
g_\sigma(x+y) - g_\sigma(x) \geq -(a+\sigma)\|g_\sigma\|_\infty y. \tag{4.5}
$$

Hence, for each $T > 0$ and for $-g_\sigma$ we have

$$
\sup_{x>0, 0<y\leq 1/T} (-g_\sigma(x+y) + g_\sigma(x)) \leq \frac{(a+\sigma)\|g_\sigma\|_\infty}{T} < \infty,
$$

which satisfies condition 4.1. By specifying $T = 32(a+1)$ and setting the upper bound $K = \|g_\sigma\|_\infty/32$, then applying the lemma and the estimate (4.4), we infer

$$
\|g_\sigma\|_\infty \leq \frac{\|g_\sigma\|_\infty}{2} + 6\frac{\eta(\sigma, 32(a+1)) + c(\omega+3)}{\sigma^\omega}.
$$

We conclude this part of the proof with $M_1(\sigma) = 12(\eta(\sigma, 32(a+1)) + c(\omega+3))$ and the estimate

$$
\|g_\sigma\|_\infty \leq M_1(\sigma)\sigma^{-\omega} \tag{4.6}
$$

2. Now, we process the second summand of G_σ, namely B. The plan is to apply the lemma once again to G_σ. Thus, similar considerations as in the last part are necessary here.

   Let us compute the derivative of $B(t)$ for $t > 0$ first:

$$
B'(t) = \frac{c}{\Gamma(\omega+1)}e^{-t}t^{\omega-1}(2te^{-t} - t + \omega(1-e^{-t})).
$$

   Using the estimate $2te^{-t} - t + \omega(1-e^{-t}) \leq (\omega+1)t$ (again easily verified by the derivative test) we find

$$
B(x) = \int_0^x B'(t)dt \leq \frac{c(\omega+1)}{\Gamma(\omega+1)}\int_0^x e^{-t}t^\omega dt.
$$

   As in the last part we need to extimate $B(x+y) - B(x)$ appropriately. If $x < 0$ and $x + y \leq 0$ there is nothing to do. Assume $x < 0$ but $x + y \geq 0$. Then

$$
\begin{aligned}
B(x+y) - B(x) = B(x+y) &\leq \frac{c(\omega+1)}{\Gamma(\omega+1)}\int_0^{x+y} e^{-t}t^\omega dt \\
\leq \frac{c(\omega+1)}{\Gamma(\omega+1)}\int_0^y e^{-t}t^\omega dt &\leq \frac{c(\omega+1)}{\Gamma(\omega+1)}\int_0^y t^\omega dt = \frac{c}{\Gamma(\omega+1)}y^{\omega+1}.
\end{aligned}
$$

   Now take $x > 0$ and $x + y > 0$. In the case $-1 < \omega < 0$ we use the fact that the function $x^\omega + 1$ is concave, thus subadditive. Hence

$$
\begin{aligned}
B(x+y) - B(x) &\leq \frac{c(\omega+1)}{\Gamma(\omega+1)}\int_x^{x+y} e^{-t}t^\omega dt \leq \frac{c(\omega+1)}{\Gamma(\omega+1)}e^{-x}\int_x^{x+y} t^\omega dt \\
&= \frac{c}{\Gamma(\omega+1)}e^{-x}((x+y)^{\omega+1} - x^{\omega+1}) \leq \frac{c}{\Gamma(\omega+1)}e^{-x}y^{\omega+1}.
\end{aligned}
$$

For $\omega > 0$ we use the standard integral estimate

$$\begin{aligned} B(x+y) - B(x) &\leq \frac{c(\omega+1)}{\Gamma(\omega+1)} e^{-x} \int_x^{x+y} t^\omega dt \\ &\leq \frac{c}{\Gamma(\omega+1)} e^{-x} (\omega+1)(x+y)^\omega y. \end{aligned}$$

These three considerations shall be applied to $B(\sigma x + \sigma y) - B(\sigma x)$ for $x \in \mathbb{R}$, $0 \leq 1/T \leq 1$ and $0 < \sigma \leq 1$. The first two directly entail in their respective cases

$$B(\sigma x + \sigma y) - B(\sigma x) \leq \frac{c}{\Gamma(\omega+1)} \left(\frac{\sigma}{T}\right)^{\omega+1}.$$

In the remaining case, where $\sigma x > 0$, $\sigma x + \sigma y > 0$, the function $e^{-\sigma x}(\omega+1)\left(\sigma x + \frac{\sigma}{T}\right)$ is clearly bounded in x by some constant $C > 0$, which leads us to

$$B(\sigma x + \sigma y) - B(\sigma x) \leq \frac{c}{\Gamma(\omega+1)} C \frac{\sigma}{T}.$$

All cases considered, denoting $D = \frac{c}{\Gamma(\omega+1)}(1 + C)$ we have in total

$$B(\sigma x + \sigma y) - B(\sigma x) \leq D\left(\frac{\sigma}{T} + \left(\frac{\sigma}{T}\right)^{\omega+1}\right). \tag{4.7}$$

3. Having all ingredients in place, we can finally take on G_σ. We apply (4.5), (4.6) and (4.7) with $0 < \sigma \leq 1$, $x \in \mathbb{R}$, $y \geq 0$ and $T = (a+1)/32$ to obtain

$$\begin{aligned} G_\sigma(x+y) - G_\sigma(x) &= g_\sigma(x+y) - g_\sigma(x) - \frac{B(x+y) - B(x)}{\sigma^{-\omega}} \\ &\geq -\left(\frac{(a+\sigma)\|g_\sigma\|_\infty}{T} + \frac{B(\sigma x + \sigma y) - B(\sigma x)}{\sigma^\omega}\right) \\ &\geq -\left(\frac{\|g_\sigma\|_\infty}{32} + \frac{D(\sigma/T + (\sigma/T)^{\omega+1})}{\sigma^\omega}\right) \\ &\geq -\frac{1}{\sigma^\omega}\left(\frac{M_1(\sigma)}{32} + D\frac{\sigma}{T} + D\left(\frac{\sigma}{T}\right)^{\omega+1}\right). \end{aligned}$$

This permits us to use Lemma 4.1 once again. Thus, for each $x \in \mathbb{R}$ we find

$$|G_\sigma(x)| \leq \sigma^{-\omega}\left(\frac{M_1(\sigma)}{2} + \frac{D}{32}\left(\frac{\sigma}{T} + \left(\frac{\sigma}{T}\right)^{\omega+1}\right)\right).$$

We set $\sigma = 1/x$ and by letting $x \to \infty$ it follows that

$$\begin{aligned} |G_\sigma(x)| &= \left|A(t)e^{-ax} - \frac{c}{\Gamma(\omega+1)} t^\omega\right| e^{-1}(1 - e^{-1}) \\ &= x^\omega\left(O(1) + O\left(\frac{1}{x}\right) + O\left(\frac{1}{x^{\omega+1}}\right)\right) = O(x^\omega) = o(e^{ax}x^\omega). \end{aligned}$$

Therefore,

$$A(t) = \frac{c}{\Gamma(\omega+1)} e^{ax} x^\omega + O(x^\omega).$$

■

4.2 The Wiener-Ikehara Tauberian Theorem

Theorem 4.3 (Wiener-Ikehara Tauberian theorem)**.** *Let*

$$f(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}, \quad a_n \geq 0$$

be a Dirichlet series with abscissa of convergence at $a > 0$*. Suppose that the function*

$$f(s) - \frac{c}{(s-a)^{\omega+1}}$$

with constants $c \geq 0$*,* $\omega > -1$ *has a holomorphic continuation to a small neighbourhood of the line* $\operatorname{Re} s = a$ *with the sole exception of the point* $s = a$*. Then, as* $x \to \infty$*, the summatory function satisfies*

$$\sum_{n \leq x} a_n \sim \frac{c}{a\Gamma(\omega+1)} x^a (\log x)^\omega.$$

Proof. Define $A(x) = \sum_{\log n < x} a_n$. We may express $f(s)$ in terms of an integral in the sense of equation (4.2) by adding zeros and reordering the sum:

$$\begin{aligned} f(s) = \frac{1}{s} \sum_{n=1}^{\infty} a_n n^{-s} &= \sum_{n=1}^{\infty} (-(n+1)^{-s} + n^{-s}) \sum_{k=1}^{n} a_k \\ &= s \sum_{n=1}^{\infty} \int_{\log n}^{\log(n+1)} e^{-st} dt \sum_{k=1}^{n} a_k = s \int_0^\infty e^{-st} A(t) dt =: sF(s). \end{aligned}$$

The function

$$G(s) := \frac{F(s+a)}{s+a} - \frac{c}{s^{\omega+1}}$$

obviously satisfies the condition in (4.3) since it is holomorphic on the line $\operatorname{Re} s = a$. Thus, Theorem 4.2 tells us

$$A(t) \sim \frac{c}{a\Gamma(\omega+1)} e^{at} t^\omega.$$

Substituting $t = \log x$ proves the claim. ∎

Example 4.4 (Continuation of Example 1.14)**.** *We found that*

$$Z^{(n)}(z) = Z^{(1)}(z) Z^{(1)}(z-1) \dots Z^{(1)}(z-(n-1)).$$

The DGF for $\mathbb{Z}$ *is the Riemann Zeta function*

$$Z^{(1)}(z) = \sum_{n=1}^{\infty} \frac{1}{n^z},$$

which is continues to a meromorphic function on $\mathbb{C}$ *with a simple pole at* $z = 1$ *and residue 1. Therefore, the abszissa of convergence of* $Z^{(n)}(z)$ *is* n*, where it has a simple pole with residue* $Z^{(1)}(2)Z^{(1)}(3)\dots Z^{(1)}(n)$*. It follows for large* m*:*

$$\sum_{k \leq m} z_k^{(n)} \sim \frac{1}{n} Z^{(1)}(2) Z^{(1)}(3) \dots Z^{(1)}(n) m^n.$$

References

[FS09] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009, pp. xiv+810. ISBN: 978-0-521-89806-5. DOI: `10.1017/CBO9780511801655`. URL: `https://doi.org/10.1017/CBO9780511801655`.

[Wil94] Herbert S. Wilf. *generatingfunctionology*. Second. Academic Press Inc., Boston, MA, 1994, pp. x+228. ISBN: 0-12-751956-4.

[Apo76] Tom M. Apostol. *Introduction to analytic number theory*. Undergraduate Texts in Mathematics. Springer-Verlag, New York-Heidelberg, 1976, pp. xii+338.

[Pro16a] ProofWiki. *Conjugacy Classes of Symmetric Group*. 2016. URL: `https://proofwiki.org/wiki/Conjugacy_Classes_of_Symmetric_Group` (visited on 04/02/2020).

[Hun74] Thomas W. Hungerford. *Algebra*. Holt, Rinehart and Winston Inc., New York-Montreal, Que.-London, 1974, pp. xix+502.

[GSS88] F. J. Grunewald, D. Segal, and G. C. Smith. "Subgroups of finite index in nilpotent groups". In: *Invent. Math.* 93.1 (1988), pp. 185–223. ISSN: 0020-9910. DOI: `10.1007/BF01393692`. URL: `https://doi.org/10.1007/BF01393692`.

[CSV19] Angela Carnevale, Michael M. Schein, and Christopher Voll. *Generalized Igusa functions and ideal growth in nilpotent Lie rings*. 2019. arXiv: `1903.03090 [math.RA]`.

[Lan99] Serge Lang. *Complex analysis*. Fourth. Vol. 103. Graduate Texts in Mathematics. Springer-Verlag, New York, 1999, pp. xiv+485. ISBN: 0-387-98592-1. DOI: `10.1007/978-1-4757-3083-8`. URL: `https://doi.org/10.1007/978-1-4757-3083-8`.

[WW96] E. T. Whittaker and G. N. Watson. *A course of modern analysis*. Cambridge Mathematical Library. An introduction to the general theory of infinite processes and of analytic functions; with an account of the principal transcendental functions, Reprint of the fourth (1927) edition. Cambridge University Press, Cambridge, 1996, pp. vi+608. ISBN: 0-521-58807-3. DOI: `10.1017/CBO9780511608759`. URL: `https://doi.org/10.1017/CBO9780511608759`.

[FO90] Philippe Flajolet and Andrew M. Odlyzko. "Singularity Analysis of Generating Functions". In: *SIAM J. Discrete Math.* 3.2 (1990), pp. 216–240. DOI: `10.1137/0403019`. URL: `https://doi.org/10.1137/0403019`.

[Pro16b] ProofWiki. *Weierstrass M-Test*. 2016. URL: `https://proofwiki.org/wiki/Weierstrass_M-Test` (visited on 04/02/2020).

[Ten15] Gérald Tenenbaum. *Introduction to analytic and probabilistic number theory*. Third. Vol. 163. Graduate Studies in Mathematics. Translated from the 2008 French edition by Patrick D. F. Ion. American Mathematical Society, Providence, RI, 2015, pp. xxiv+629. ISBN: 978-0-8218-9854-3.