ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

Analytic Asymptotics of Discrete Noiseless Channels

Master's Thesis

Georg Böcherer

March 2, 2007

Professor: A. Lapidoth
Advisor: V. C. da Rocha Jr.
Co-Advisors: C. Pimentel, T. Koch

To my mother

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

Master's Thesis Winter Semester 2006/2007

The Discrete Noiseless Channel

Georg Böcherer

Advisor: Prof. Valdemar Cardoso da Rocha Junior
Co-Advisors: Prof. Cecilio José Lins Pimentel and Tobias Koch

1 Introduction

Theorems 1 and 8, from Shannon's 1948 classic paper [1] both deal with the capacity of what Shannon called the discrete noiseless channel. A discrete noiseless channel is a channel which allows the noiseless transmission of a sequence of symbols chosen from a finite alphabet $\mathcal{A}$, each symbol having a certain duration in time, possibly different for different symbols. Furthermore, there may be restrictions on the allowed sequences of symbols from $\mathcal{A}$. Suppose $\mathcal{A}$ is a q-letter alphabet, and that associated with each letter $a \in \mathcal{A}$ is a positive number $\tau(a)$ called the duration of a.

Example 1. *Shannon's telegraphy alphabet, with $\mathcal{A} = \{d, D, s, S\}$ and durations given in the following table:*

a	d	D	s	S
$\tau(a)$	2	4	3	6

where, in Shannon's terminology, d stands for dot*, D for* dash*, s for* letter space*, and S for* word space*.*

A word of length k over $\mathcal{A}$ is a finite string of k letters from $\mathcal{A}$. If $a = a_1 a_2 \ldots a_k$ is such a word, its duration is defined to be $\tau(a) = \tau(a_1) + \ldots + \tau(a_k)$. For example, 010110 is a word of length 6 and duration 6 over the standard binary alphabet, 01221022 is a word of length 8 and duration 8 over the standard 3-ary alphabet, and $dddsddddsdDsDdsDdsDDDsDdS$ is a word of length 25 and duration 74 over Shannon's telegraphy alphabet.

A language $\mathcal{L}$ over $\mathcal{A}$ is a collection of words over $\mathcal{A}$. The discrete noiseless channel associated with $\mathcal{L}$, the $\mathcal{L}$-channel for short, is the channel which is only allowed to transmit sequences from $\mathcal{L}$, although it transmits them without error.

What is the capacity of the $\mathcal{L}$-channel? For a general language $\mathcal{L}$, not much can be said, and the usual treatment of this issue is restricted to a special class of languages, as did Shannon. This study will consider only Shannon languages. A Shannon language is defined by a directed graph whose edges are labelled with letters from the alphabet $\mathcal{A}$. The corresponding language $\mathcal{L}$ is then defined to be the set of words that result by reading off the edge labels on paths of the graph. For example, if the graph consists of a single vertex v and q self-loops at v, each labeled with a different element of $\mathcal{A}$, the resulting language consists of all possible sequences over $\mathcal{A}$. Shannon gave two different definitions for the capacity of a noiseless channel corresponding to a Shannon language $\mathcal{L}$, and then showed that the two definitions gave the same value. This common value was thereby established unambiguously as the maximum rate, in bits per second, that information can be transmitted over the channel. These two definitions are called the combinatorial capacity and the probabilistic capacity [2]. If $\mathcal{L}$ is a Shannon language, let $N(\tau)$ denote the total number of words in $\mathcal{L}$ of duration τ. The combinatorial capacity of the $\mathcal{L}$-channel is defined as

$$C_{\text{comb}} = \overline{\lim_{\tau\to\infty}} \frac{1}{\tau} \log N(\tau).$$

Similarly, let $\{X_n\}$ be a stationary discrete Markov chain defined on the labeled graph defining $\mathcal{L}$. If the entropy of $\{X_n\}$ is H and the average branch duration is T, then the information rate of $\{X_n\}$, in bits per second, is H/T. The probabilistic capacity of the $\mathcal{L}$-language is defined to be the maximum of this rate, over all possible Markov chains:

$$C_{\text{prob}} = \sup_{\{X_n\}} \frac{H}{T}.$$

In his original paper, Shannon gave a simple algebraic method for computing C_{comb} (Theorem 1), and showed that the same value held for C_{prob} (Theorem 8). His proofs, however, were considered brief and in places quite cryptic [2].

2 Tasks

The following tasks may be helpful for your work. Since it cannot be foreseen how the project will evolve, you will not be evaluated exclusively on the fulfillment of these tasks, but more on the creativity that you exhibit.

- A careful treatment of the case of symbols of different durations, including non-integer durations.
- Khandekar's proof [2] which uses the *partition function* technique to find the combinatorial capacity.
- Pimentel's technique [3] to compute capacity by defining constraints in terms of forbidden strings.
- Alternative derivation of McMillans' inequality [4].

3 General Regulations

The project will be supervised by Professor Rocha and co-supervised by Professor Pimentel and Tobias Koch. You will have to hand in a report (an original and a copy, both signed) that is typeset in LaTeX. The original as well as the copy are property of the laboratory.

There will be a mandatory introduction to the lab and its facilities where further details and regulations are going to be explained.

Dates

Beginning	**:**	Monday,	September 4, 2006
End	**:**	Friday,	March 2, 2007, 12:00 pm

Addresses

Prof. A. Lapidoth	T. Koch
ETF E 107	ETF D 107
`<lapidoth@isi.ee.ethz.ch>`	`<tkoch@isi.ee.ethz.ch>`
044/632 51 92	044/632 65 87

Zurich, August 3, 2006

Prof. Amos Lapidoth

References

[1] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423 and 623–656, July and October 1948.

[2] A. Khandekar, R. McEliece, and E. Rodemich. The discrete noiseless channel revisited. In *Coding, Communications, and Broadcasting*, Research Studies Press Ltd., pages 115–137. John Wiley & Sons, 2000.

[3] C. Pimentel and B.F. Uchoa-Filho. A combinatorial approach to finding the capacity of the discrete noiseless channel. 49(8):2024–2028, August 2003.

[4] V.C. da Rocha Jr. Some information-theoretic aspects fo uniquely decodable codes. In *Coding, Communications, and Broadcasting*, Research Studies Press Ltd., pages 39–47. John Wiley & Sons, 2000.

Acknowledgement

I would like to thank Professor Amos Lapidoth and Professor Valdemar Cardoso da Rocha Junior for the possibility to do my master's thesis at the Federal University of Pernambuco, Brazil.

I wish to thank Professor Valdemar Cardoso da Rocha Junior, Professor Cecilio José Lins Pimentel, and Tobias Koch for their scientific and personal support of my work.

I would like to thank Márcio Lima for the interesting discussions and for his support.

Georg Böcherer

Abstract

Of main interest in this work is the number $N[w]$ of distinct strings of weight w that are accepted by a discrete noiseless channel. We investigate the asymptotic behavior of $N[w]$ for w getting large. We show how the number sequence $N[w]$ can precisely be represented by a generating series. We then interpret this generating series as a generating function of a complex argument, and we show how the asymptotic behavior of $N[w]$ is related to the analytic characteristics of the generating function. We generalize Pringsheim's Theorem and the Exponential Growth Formula to the case where the string weights w take non-integer values. This allows us to relate the exponential behavior of $N[w]$ to the radius of convergence of its generating function, to the leftmost real singularity of its generating function, and to the smallest positive pole of its generating function. We show how the sub-exponential behavior of $N[w]$ is determined by the main parts of the Laurent series expansions of the generating function around its poles. Finally, we apply these techniques to information theoretic problems. We define the capacity of a general discrete noiseless channel, and we show how it can be calculated from the generating function. For a general discrete noiseless channel, we prove the first part of the fundamental theorem of discrete noiseless channels, which says that for every rate C' smaller than the capacity of the considered channel, there exists a random process that generates strings accepted by the channel at an entropy rate equal to C'.

Contents

1. Introduction **1**

2. Algebraic Representation **5**
2.1. Definitions . . . 6
2.1.1. Discrete Noiseless Channel . . . 6
2.1.2. Generating Series . . . 6
2.2. Generating Series by Basic Set Operations . . . 7
2.2.1. Union and Intersection . . . 7
2.2.2. Concatenation . . . 8
2.2.3. Kleene Star . . . 9
2.3. Generating Series from a List of Forbidden Substrings . . . 11
2.3.1. Correlation Functions . . . 11
2.3.2. Forbidden Substrings . . . 12
2.3.3. Pattern Codes . . . 14
2.4. Shannon's Telegraphy Channel . . . 14

3. Asymptotic Analysis **17**
3.1. Information-Theoretic Criteria for DNCs of Interest . . . 17
3.2. Exponential Behavior . . . 20
3.2.1. Exponential Order by Radius of Convergence . . . 21
3.2.2. Exponential Order by Leftmost Singularity . . . 24
3.2.3. Exponential Order by Smallest Positive Pole . . . 27
3.3. Sub-Exponential Behavior for Integer String Weights . . . 29
3.3.1. Expansion of Rational Functions . . . 29
3.3.2. Sub-Exponential Behavior . . . 31
3.4. Shannon's Telegraphy Channel . . . 35
3.4.1. Exponential Order of Coefficients . . . 35
3.4.2. Approximation of Coefficients with Arbitrary Precision . . . 36
3.5. Combination of Two Pattern Codes . . . 43
3.6. Sub-Exponential Behavior for Non-Integer String Weights . . . 45
3.6.1. Sub-Exponential Behavior for Commensurable String Weights . . . 46
3.6.2. Sub-Exponential Behavior for Incommensurable String Weights . . . 46

4. Information Theoretic Aspects **53**
4.1. Capacity . . . 53
4.1.1. Definition . . . 53

4.1.2. Calculation . . . 55
4.1.3. Capacity of a DNC Represented by a Finite State Machine . . . 56
4.2. Channel Coding . . . 57
4.2.1. Random Walk in a Tree . . . 58
4.2.2. DNC Coding Theorem . . . 58
4.3. Uniquely Decodable Codes . . . 61
4.3.1. Capacity Achieving Distribution . . . 61
4.3.2. McMillan's Inequality . . . 62
4.4. Shannon's Telegraphy Channel . . . 63
4.4.1. Capacity . . . 63
4.4.2. Capacity Achieving Distribution . . . 65

5. Conclusions 67

A. Notation 69

B. Mathematics 71
B.1. Complex Analysis . . . 71
B.1.1. Analytic Functions . . . 71
B.1.2. Localization of Poles . . . 72
B.1.3. Analyticity of z^r . . . 72
B.2. Miscellaneous Mathematics . . . 74
B.2.1. Proof of Lemma 4 . . . 74
B.2.2. Proof of Lemma 6 . . . 75
B.2.3. Lower Bound of Factorial Function . . . 77
B.2.4. Newton's Expansion . . . 77
B.3. Equality of Operational and Combinatorial Capacity . . . 79

List of Figures

1.1. A general communication system. 1

3.1. The complex logarithm and its inversion 25
3.2. First application of argument principle 37
3.3. Second application of argument principle 41
3.4. Location of complex roots 42
3.5. The coefficients $N_{011}[k]$, $N_{0111}[k]$, and $N_F[k]$. 44
3.6. Plot of the coefficients 49
3.7. Logarithmic plot of the coefficients 50
3.8. Linear plot of the cumulated coefficients 51

4.1. Accepted strings represented by a tree 64

1. Introduction

The fundamental problem in communication engineering is to reliably transmit data from a sender to a receiver spending as least resources as possible. The data transmission is reliable, if it is possible to recover the sent data from the received data in a satisfactory way. For a fixed resource, how much data can be transmitted in the maximum? Claude E. Shannon introduced in his legendary paper [1] a measure of information, which allows us to investigate this question with mathematical rigor. The ultimate performance of the system is equal to the result of maximizing the mutual information between the sent data and the received data over all possible configurations of the system that use the fixed resource only. In many cases, this maximization problem can be divided into several independent maximization problems, depending on how we model our communication system. A typical scenario is displayed in Figure 1.1. Since for every data we transmit we have to spend resources, we first compress the data to be sent. In Figure 1.1, the corresponding procedure is denoted by "source encoding". Data compression is a wide area in information theory, the fundamentals are discussed in [2, ch. 5] and [3, ch. 2]. The remaining two maximization problems have to be read from the right. For given channel specifications, we first ask the question how much data per resource can in the maximum be reliably transferred over the channel. This questions leads to the channel capacity. Closely related to the channel capacity is the problem how to represent the data to achieve the capacity. In Figure 1.1, this problem is denoted by "channel encoding". Channel capacity and channel encoding depends highly on the channel specifications, which can be very different for different communication problems. In wired and wireless communication, it has turned out to be necessary to explicitly model the noise. Otherwise, the channel models would allow reliable communication even if the system only spends an infinitesimal small amount of resources. From an information theoretic point of view, this situation is equivalent to a channel of infinite capacity, which clearly does not model reality in a

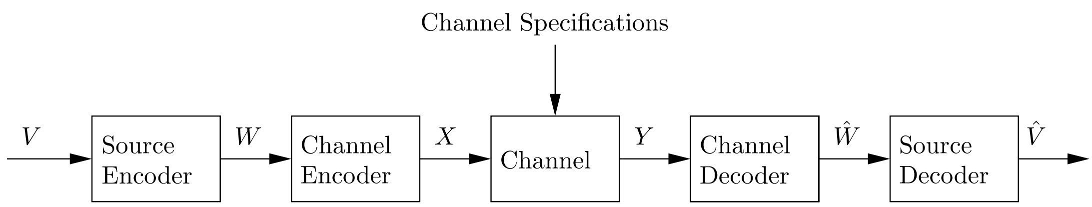

Figure 1.1.: A general communication system.

reasonable way. For an introduction to communications in additive noise see [4], and for an introduction to communication over fading channels, which are modelled by multiplicative noise in addition to additive noise, see [5] and [6]. In other situations, such as magnetic recording or data storage on hard-disks, the common modelling approach is different. As in wired or wireless communications, the physical nature of the considered storage systems implies undesired side-effects. These can be of very different kind including inter-symbol interference, perturbations from other communication systems, and limitations on the maximum resolution of our system. Instead of explicitly modeling all these effects, we formulate some constraints on how to store the data that guarantee that the undesired effects do not affect the performance of our system significantly. A very simple example would be the communication over one pin of the parallel-port of a computer. To guarantee reliable communication, we use the following two constraints. First, we only allow two symbols—a 0 represented by a voltage of 0V and a 1 represented by a voltage of 5V. Second, we only allow the symbol to be changed 1000 times in a second. These two constraints eliminate all undesired side-effects: if we send a 1 the receiver will almost surely receive a 1. As long as we fulfill the constraints, it is reasonable to model the channel without noise. It is however important to keep in mind that no physical channel is noiseless.

In this work, we investigate the performance of discrete noiseless channels. Discrete noiseless channels allow the error free transmission of strings of symbols fulfilling the channel specifications. The channel specifications of a discrete noiseless channel consist of constraints both on symbol constellations and symbol weights. If 0 and 1 are the two allowed symbols for transmission, a constraint on the symbol constellation could for example be “only strings with at most 4 consecutive 1s are allowed for transmission”. Depending on the system we model, the symbol weight can for example represent the duration in time of the symbols, the energy of the symbols or the length of the symbols in space. A possible constraint for the symbol weights could for example be “the symbol 0 has to have the energy of 1W and the symbol 1 has to have the energy of 2W”. We consider the symbol weight as the resource of our system. If the resource is energy, the maximization problem of transmission is how much data per Watt can be transmitted over the channel in the maximum.

For a general discrete noiseless channel, let $N[w]$ denote the number of strings of weight w that fulfill the constraints of the channel. The maximum data rate at which we can transmit over the channel is then, under certain conditions, given by

$$C = \limsup_{w\to\infty} \frac{\log N[w]}{w} \tag{1.1}$$

and we call C the combinatorial capacity of the channel. Here and hereafter we will denote by $\log x$ the natural logarithm. In our work, we will take a more comprehensive look at the performance of discrete noiseless channels. Our fundamental question is

> “How many strings of the weight w are accepted by the channel?”

We will answer this question with exponential and sub-exponential precision. It will turn out that for the determination of the combinatorial capacity, only exponential precision is needed.

The discrete noiseless channel was first defined by Shannon in [1, ch. 1]. Shannon considered integer valued symbol weights and allowed symbol constellations that can be represented by finite state machines. Shannon calculated the capacity and the capacity achieving data representation for this class of channels.

In [7], the authors call this first chapter of Shannon's paper a "relatively unsung part of Shannon's classic paper" and argue that the subject of discrete noiseless channels is only rudimentary or not at all treated in current textbooks on information theory. However, under different names, the discrete noiseless channel has attracted a lot of attention in the last decades. If the symbol weights are integer-valued, investigating the behavior of the numbers $N[w]$ from (1.1) for increasing w is equivalent to the problem of enumerating strings that fulfill the channel constraints. Odlyzko suggests in [8] a unified treatment of string enumerations by generating functions. In [9, ch. 6], he then uses this approach to derive formulas for the asymptotic behavior of $N[w]$ for $w \to \infty$. To use analytic methods to investigate the asymptotic behavior of string enumerations falls in the area of analytic combinatorics. To cite the Online Journal of Analytic Combinatorics, whose first issue appeared in 2006, "An exciting new branch of mathematics is emerging at the intersection of analysis, combinatorics, and number theory"[10].

In the book [11], which will probably be published in 2007, the authors provide a systematic approach called the "symbolic method" to map enumerations of strings to generating functions and they then use complex analysis on the generating functions to derive formulas for the asymptotic behavior of the enumerations with sub-exponential precision. The book can be interpreted as a detailed treatment of discrete noiseless channels with integer-valued symbol weights, however, the authors come from the areas of computer science and algorithm theory and mention information theoretic aspects of their results only in some examples.

In [12], the authors focus on the constraints on symbol constellations and consider discrete noiseless channels with symbol weights equal 1 and call these channels "constraint systems". They dedicate a whole chapter to the derivation of the capacity of these channels using results from matrix theory and they also provide capacity achieving data representations.

There are two recent publications in information theory dealing with the discrete noiseless channel. In [7], the authors focus on discrete noiseless channels which can be represented by finite automata. They generalize the concept of generating functions to non-integer valued symbol weights and use results from matrix theory to give a formula for the capacity of the channel. Again using results from matrix theory, they provide a capacity achieving data representation. In [13], the authors use combinatorial methods to derive the generating function for discrete noiseless channels where the allowed symbol constellations can be represented by a list of forbidden substrings. For the case of integer-valued symbol weights, this result was independently derived in [14].

In our work, we collect and extend the upper results with a focus on non-integer valued symbol weights. Our main contributions are the following:

- For a general discrete noiseless channel with non-integer valued symbol weights, we define the generating series as an algebraic representation of its combinatorial complexity. We then present various techniques to unambiguously construct the generating series.
- We precisely identify the asymptotic behavior of the number of strings $N[w]$ of weight w allowed by the channel with the analytic characteristics of its generating function, which results from interpreting the corresponding generating series as a function.
- We calculate the capacity C of general discrete noiseless channels and show that for every entropy rate $C' < C$, there is a data representation which can be transmitted at a rate of C' over the channel.

Our work is divided into three independent parts. In the first part, we define the generating series of discrete noiseless channels as a formal power series. We construct discrete noiseless channels by algebraic operations and show how these operations translate to the generating series. Our approach stands in contrast to the matrix approach used in [7], it can be seen in parallel with [15] or [11, ch. 1]. As alternative techniques to unambiguously construct generating series, we represent the results from [13] and [14].

In the second part, we interpret the generating series of discrete noiseless channels as functions and we define the complex generating function. For the number $N[w]$ of distinct strings of weight w accepted by some discrete noiseless channel, we show how we can predict its asymptotic exponential behavior and its asymptotic sub-exponential behavior from the characteristics of the complex generating function of the channel. We show this both for discrete noiseless channels with integer valued string weights and for discrete noiseless channels with non-integer valued string weights. This part can therefore be seen as a generalization of the results from [11, ch. 4] to non-integer valued string weights.

In the third part, we investigate information theoretic aspects of the discrete noiseless channel. We first show how to calculate the combinatorial capacity of discrete noiseless channels. Inspired by [16], we start by considering discrete noiseless channels where the set of allowed strings are generated by a uniquely decodable code. We introduce a new proof technique, which allows us to generalize McMillan's inequality to non-integer valued symbol weights and to derive the capacity achieving distribution for uniquely decodable codes. Finally, we generalize the results for uniquely decodable codes to general discrete noiseless channels and we show that for every entropy rate C' smaller than the combinatorial capacity C, there is a random process that achieves the rate C'. Referring to Figure 1.1, we show how the data X must look like such that it can be transmitted over the channel at a rate of C'. In the converse, we show that there is no random process that allows data transmission over a discrete noiseless channel with an entropy rate C' larger than the combinatorial capacity C of the considered channel. The results from this part can be seen as a generalization of the results from [7]. In contrast to other works, we do not use matrix theory in our derivations.

2. Algebraic Representation

We model data in general digital communication systems to be discrete both in time/space and value. A wide range of digital applications can directly be modeled in this way. An example of a time discrete system is the data stream from a standard audio CD where we have 2 channels with 44 100 samples per second with each sample consisting of 16 bits. To model the discrete nature in time/space we represent the data by strings resulting from concatenations of symbols. The discrete nature of the values we model by taking the symbols from a countable (possibly infinite) set. In addition, we associate a nonnegative real weight with each symbol. This weight can represent a duration in time, a distance in space or any other physical measure of interest.

For two sets A and B, we denote by $A \cup B$ the union of A and B:

$$s \in A \cup B \Leftrightarrow s \in A \text{ or } s \in B. \tag{2.1}$$

By AB, we denote the concatenation of A and B:

$$s \in AB \Leftrightarrow \exists a, b\colon\, s = ab \text{ and } a \in A, b \in B. \tag{2.2}$$

Let A denote a set of possible symbols in a discrete system. The set of all possible strings S resulting from arbitrary concatenations of symbols from A is given by

$$S = A^\star \tag{2.3}$$
$$= \varepsilon \cup A \cup AA \cup AAA \cup \dots \tag{2.4}$$

where the symbol * denotes the Kleene star, see [17, ch. 1, p. 23], and where ε denotes the empty string with $s\varepsilon = \varepsilon s = s$ for any string s. The set $A^\star$ is the set of all concatenations of zero or more symbols from A. In many cases however, the set of strings that actually are used by the system of interest is only a subset of $A^\star$. This can for example result from spectrum shaping constraints, inter symbol interference constraints or synchronization constraints on the allowed strings. To represent these constraints, Shannon introduced in [1] the notion of a discrete noiseless channel (DNC). His idea was the following: let $B \subset S = A^\star$ be the set of strings allowed by the system. We interpret the system as a channel with an input consisting of strings $s \in S$. Whenever $s \in B$, s is transmitted correctly (noiseless), and whenever $s \in S \setminus B$, the string is not transmitted by the channel. This notion is closely related to the definition of regular languages, where a string gets processed by a finite automata that accepts the string if it is part of the language and generates an error if it is not [17, ch. 1, p. 16].

2.1. Definitions

2.1.1. Discrete Noiseless Channel

We formally define the DNC as follows.

Definition 1. The *DNC* $\mathcal{A} = (A, w)$ consists of a countable set A of strings and an associated weight function $w\colon A \mapsto \mathbb{R}^{\oplus}$ with the following property: if the strings s_1 and s_2 are in A, and if the concatenation $s_1 s_2$ is also in A, then we have for the weight of the concatenation $s_1 s_2$:

$$w(s_1 s_2) = w(s_1) + w(s_2). \tag{2.5}$$

In addition, the empty string ε is always of weight zero, i.e., $w(\varepsilon) = 0$.

2.1.2. Generating Series

For a given DNC, we are interested in the number $N[w]$ of strings that are accepted by the DNC and that are of the same weight w. The concept of generating series allows us to represent the number $N[w]$ for all possible string weights w in one single algebraic expression.

Definition 2. Let $\mathcal{A} = (A, w)$ represent a DNC. The *generating series* of $\mathcal{A}$ we define by

$$\mathrm{GEN}_{\mathcal{A}}(x) = \sum_{s \in A} x^{w(s)}, \qquad x \text{ undefined.} \tag{2.6}$$

Note that the series is for now a pure algebraic concept similar to the one of formal power series as defined for example in [18] and we do not care about the convergence of the sum. We denote by $[x^w]\mathrm{GEN}_{\mathcal{A}}(x)$ the coefficient of the power x^w in $\mathrm{GEN}_{\mathcal{A}}(x)$. The comparison of two generating series we perform element-wise, i.e.,

$$\mathrm{GEN}_{\mathcal{A}}(x) \leq \mathrm{GEN}_{\mathcal{B}}(x) \tag{2.7}$$

$$\Leftrightarrow$$

$$[x^w]\mathrm{GEN}_{\mathcal{A}}(x) \leq [x^w]\mathrm{GEN}_{\mathcal{B}}(x), \qquad \forall w \geq 0. \tag{2.8}$$

Summing up the terms corresponding to strings of the same weight yields

$$\mathrm{GEN}_{\mathcal{A}}(x) = \sum_{s \in A} x^{w(s)} \tag{2.9}$$

$$= \sum_{w \in w(A)} N[w] x^w \tag{2.10}$$

where we denote by $w(A)$ the set of all possible string weights. The coefficient $[x^w]\mathrm{GEN}_{\mathcal{A}}(x)$ is thus equal to the number $N[w]$ of distinct strings of weight w accepted by the DNC $\mathcal{A}$:

$$N[w] = [x^w]\mathrm{GEN}_{\mathcal{A}}(x). \tag{2.11}$$

2.2. Generating Series by Basic Set Operations

Until now, we did not assume any structure for the set of strings accepted by a DNC. In practice, the set of strings is often defined by a short list of specifications leading to a well-structured set of accepted strings. A large class of these specifications can be translated into simple algebraic operations on basic DNC.

2.2.1. Union and Intersection

We start by defining the union of two DNCs.

Definition 3. Let $\mathcal{A} = (A, w_A)$ and $\mathcal{B} = (B, w_B)$ denote two DNCs with $w_A(s) = w_B(s)$ for all $s \in A \cap B$. We define the union of $\mathcal{A}$ and $\mathcal{B}$ by

$$\mathcal{A} \cup \mathcal{B} = (A \cup B, w_C), \tag{2.12}$$

$$w_C(s) = \begin{cases} w_A(s), & s \in A \\ w_B(s), & \text{otherwise.} \end{cases} \tag{2.13}$$

The union of two DNCs is a DNC that accepts strings fulfilling the specifications of the first DNC or the specifications of the second DNC. The two DNCs are thus driven in parallel. We next consider the intersection of two DNCs.

Definition 4. Let $\mathcal{A} = (A, w_A)$ and $\mathcal{B} = (B, w_B)$ denote two DNCs with $w_A(s) = w_B(s)$ for all $s \in A \cap B$. We define the intersection of $\mathcal{A}$ and $\mathcal{B}$ by

$$\mathcal{A} \cap \mathcal{B} = (A \cap B, w_A). \tag{2.14}$$

The intersection of two DNCs is a DNC that accepts strings that both fulfill the specifications of the first DNC and the specifications of the second DNC. The two DNCs are thus driven in series. The generating series of the intersection $\mathcal{A} \cap \mathcal{B}$ follows directly from Definition 2. For the generating series of the union $\mathcal{A} \cup \mathcal{B}$, we have the following lemma:

Lemma 1. *Let $\mathcal{A} = (A, w_A)$ and $\mathcal{B} = (B, w_B)$ denote two DNCs with $w_A = w_B$ on $A \cap B$. We then have*

i. The generating function of the union $\mathcal{A} \cup \mathcal{B}$ is given by

$$\mathrm{GEN}_{\mathcal{A}\cup\mathcal{B}}\,(x) = \mathrm{GEN}_{\mathcal{A}}\,(x) + \mathrm{GEN}_{\mathcal{B}}\,(x) - \mathrm{GEN}_{\mathcal{A}\cap\mathcal{B}}\,(x)\,. \tag{2.15}$$

ii. We can upper-bound the generating function by

$$\mathrm{GEN}_{\mathcal{A}\cup\mathcal{B}}\,(x) \leq \mathrm{GEN}_{\mathcal{A}}\,(x) + \mathrm{GEN}_{\mathcal{B}}\,(x) \tag{2.16}$$

with equality if and only if

$$A \cap B = \emptyset. \tag{2.17}$$

Proof. The lemma follows from basic set-theory. We have to consider the cardinality of the union of two sets. □

2.2.2. Concatenation

The next operation we will consider is the concatenation of two DNCs.

Definition 5. Let $\mathcal{A} = (A, w_A)$ and $\mathcal{B} = (B, w_B)$ denote two DNCs with $w_A(s) = w_B(s)$ on $s \in A \cap B$. We define the concatenation of $\mathcal{A}$ and $\mathcal{B}$ by

$$\mathcal{A}\mathcal{B} = (AB, w_C) \tag{2.18}$$

where for any $s_1 \in A$, $s_2 \in B$, the weight of $s = s_1 s_2$ is given by

$$w_C(s) = w_A(s_1) + w_B(s_2). \tag{2.19}$$

We interpret the concatenation of two DNCs in the following way. It is a DNC that uses different specifications for the first and second part of the accepted strings. For the generating series of the concatenation of two DNCs we have the following lemma.

Lemma 2. *Let $\mathcal{A} = (A, w_A)$ and $\mathcal{B} = (B, w_B)$ denote two DNCs with $w_A(s) = w_B(s)$ for all $s \in A \cap B$. The generating series of $\mathcal{A}\mathcal{B}$ can be upper-bounded by*

$$\mathrm{GEN}_{\mathcal{A}\mathcal{B}}(x) \leq \mathrm{GEN}_{\mathcal{A}}(x)\,\mathrm{GEN}_{\mathcal{B}}(x)\,, \tag{2.20}$$

with equality if and only if

$$|\bar{A}\bar{B}| = |\bar{A}|\,|\bar{B}| \tag{2.21}$$

for all finite subsets $\bar{A} \subseteq A$ and $\bar{B} \subseteq B$. By $|A|$, we denote the cardinality of the set A.

Proof. The lemma follows from basic set-theory. We have to consider the cardinalities of the involved sets. □

The condition (2.21) is equivalent to the following. Every string $s \in AB$ is generated unambiguously by the concatenation of a unique symbol from A and a unique symbol from B (No $s \in AB$ is generated "twice" when concatenating A with B). The following example illustrates the derivation of the generating series of two concatenated DNCs.

Example 1. (*Generating function of concatenated DNCs*). We consider the DNC $\mathcal{A} = (A, w)$ with $A = \{s, st\}$ and $w(s) = w(t) = 1$ and the DNC $\mathcal{B} = (B, w)$ with $B = \{t, \varepsilon\}$. We have

$$\mathrm{GEN}_{\mathcal{A}}(x) = x + x^2 \tag{2.22}$$

$$\mathrm{GEN}_{\mathcal{B}}(x) = 1 + x. \tag{2.23}$$

The concatenation of $\mathcal{A}$ and $\mathcal{B}$ results in the DNC $\mathcal{A}\mathcal{B} = (D, w)$ with $D = \{s, st, stt\}$. The generating series of $\mathcal{A}\mathcal{B}$ is thus given by

$$\mathrm{GEN}_{\mathcal{A}\mathcal{B}}(x) = x + x^2 + x^3. \tag{2.24}$$

The product of $\mathrm{GEN}_{\mathcal{A}}(x)$ and $\mathrm{GEN}_{\mathcal{B}}(x)$ would have led to

$$\mathrm{GEN}_{\mathcal{A}}(x)\,\mathrm{GEN}_{\mathcal{B}}(x) = (x + x^2)(1 + x) \tag{2.25}$$
$$= x + x^2 + x^2 + x^3 \tag{2.26}$$
$$= x + 2x^2 + x^3. \tag{2.27}$$

◁

It is important to note that the generating series of two concatenated DNCs is in general not equal to the product of the generating series of the two DNCs. In [11], the authors combine combinatorial classes not by concatenating the involved sets but by building the Cartesian product of the involved sets. We denote the Cartesian product of two sets A and B as $A \times B$. It is defined as

$$A \times B = \big\{(a, b)\big|a \in A, b \in B\big\}. \tag{2.28}$$

As we can see, every element $c \in A \times B$, can be uniquely written as a pair $c = (c_1, c_2)$ with $c_1 \in A$ and $c_2 \in B$. As a consequence, the generating series of a Cartesian product of two combinatorial classes is always equal to the product of the generating series of the combinatorial classes. For the concatenation of two sets, this is not the case. For the string $st \in AB$ from the previous example, the decomposition $st = c_1c_2$ with $c_1 \in A$ and $c_2 \in B$ is not unique. We can either write $c_1 = s$ and $c_2 = t$ or we can write $c_1 = st$ and $c_2 = \varepsilon$.

2.2.3. Kleene Star

The last operation we introduce is the Kleene star operation denoted by *, which we already defined for sets. For DNCs we define it as follows:

Definition 6. Let $\mathcal{A} = (A, w)$ be a DNC. The DNC $\mathcal{A}^\star$ is defined as

$$\mathcal{A}^\star = (A^\star, w). \tag{2.29}$$

We interpret $\mathcal{A}^\star$ as a DNC that piecewise checks the validity of the input string. If the string is a concatenation of substrings fulfilling the specifications of $\mathcal{A}$, it is valid in $\mathcal{A}^\star$. According to its definition, the DNC $\mathcal{A}^\star$ results from the union of concatenations of the basic DNC $\mathcal{A}$. To allow us to give a simple formula for the generating series of $\mathcal{A}^\star$, the concatenations have to fulfill the condition (2.21) and the unions have to fulfill the condition (2.17). Otherwise, the resulting set is generated ambiguously, and we can not give a formula for the generating series of $\mathcal{A}^\star$ in terms of the generating function of $\mathcal{A}$. It is not easy to check these conditions, as we illustrate in the following example.

Example 2. (*Concatenations and unions in* $\mathcal{A}^\star$). Consider the DNC $\mathcal{A} = (A, w)$ with $A = \{a, b, ab\}$ and $w(a) = w(b) = 1$. The set AA is given by

$$AA = \{aa, ab, aab, ba, bb, bab, aba, abb, abab\}. \tag{2.30}$$

The concatenation AA fulfills (2.21). However, $A \cap AA$ does not fulfill (2.17) since $A \cap AA = \{ab\} \neq \emptyset$. The concatenation AAA does not fulfill condition (2.21), since the string $abab$ is both generated by $a\,b\,ab$ and $ab\,a\,b$. ◁

To avoid these complications, we impose on A the restriction that every string $s \in A^\star$ can be uniquely written as a concatenation of symbols from A. This coincides with the definition of uniquely decodable codes as given in [2]. If A forms a uniquely decodable code, then all strings $s \in A^\star$ are uniquely generated. Otherwise, they are generated ambiguously. Based on this observation, we have for the generating series of $\mathcal{A}^\star$ the following lemma:

Lemma 3. *Let $\mathcal{A} = (A, w)$ denote a DNC. The generating series of the DNC $\mathcal{A}^\star$ is then upper-bounded by*

$$\mathrm{GEN}_{\mathcal{A}^\star}(x) \leq \sum_{k=0}^{\infty} [\mathrm{GEN}_{\mathcal{A}}(x)]^k \tag{2.31}$$

$$= \frac{1}{1 - \mathrm{GEN}_{\mathcal{A}}(x)} \tag{2.32}$$

with equality in (2.31) *if and only if the set A forms a uniquely decodable code.*

Proof. The lemma follows from the upper discussion. □

To check if the set A is uniquely decodable, we can use the Sardinas-Patterson test. The following paragraph is a verbatim quotation of the description of this test given in Problem 24 in [2, ch. 5]:

A code is not uniquely decodable if and only if there exists a finite sequence of code symbols which can be resolved in two different ways into sequences of codewords. That is, a situation such as

$$\begin{array}{|c|c|c|c|}
\hline
A_1 & A_2 & A_3 \quad \dots & A_m \\
\hline
B_1 \;|\; B_2 & B_3 & \dots & B_n \\
\hline
\end{array} \tag{2.33}$$

must occur where each A_i and each B_i is a codeword. Note that B_1 must be a prefix of A_1 with some resulting “dangling suffix.” Each dangling suffix must in turn be either a prefix of a codeword or have another codeword as prefix, resulting in another dangling suffix. Finally, the last dangling suffix in the sequence must also be a codeword. Thus one can set up a test for uniquely decodability (which is essentially the Sardinas-Patterson test [19]) in the following way: construct a set S of all possible dangling suffixes. The code is uniquely decodable if and only if S contains no codewords.

There is an important class of codes called prefix-free codes. A code is prefix-free if no codeword is a prefix of another codeword. Prefix-free codes are uniquely decodable, see [2]. In the same way, we can define suffix-free codes as codes where no codeword is a suffix of another codeword. Suffix-free codes are also uniquely decodable. This can be seen by reading prefix-free codes from behind.

For a DNC $\mathcal{A} = (A, w)$, we may be interested in the generating function of $\mathcal{A}^\star$ although A does not form a uniquely decodable code. In this case, we must find an alternative representation of the set $A^\star$. We illustrate this in the following example.

Example 3. (*Making a set uniquely decodable*). Let $\mathcal{A} = (A, w)$ represent a DNC with $A = \{0, 01, 10\}$ and $w(0) = w(1) = 1$. What is the generating function of $\mathcal{A}^\star$? We check if A is uniquely decodable. The concatenation of 01 with 0 results in the string 010. However, the symbol $0 \in A$ is a prefix of 010, which leads to the dangling suffix 10, which itself is a symbol from A. Thus, A does not form a uniquely decodable code and we cannot use Lemma 3 to derive the generating function of $\mathcal{A}^\star$. We therefore have to represent $A^\star$ in a different way. The symbol 1 appears in strings from $A^\star$ either isolated in the form $\ldots 010 \ldots$ or as a pair in the form $\ldots 0110 \ldots$. The symbol 0 appears in an arbitrary manner. Strings of this kind are generated by $B^\star$ with $B = \{0, 01, 011\}$. Note that B forms a uniquely decodable code, since it is suffix-free. To represent $A^\star$ using B we have to consider the border conditions. With respect to the rightmost bits, the strings in $B^\star$ are equal to the strings in $A^\star$. However, all strings in $B^\star$ start with 0, but there are strings from $A^\star$ that start with the symbol 10. Including these strings, we have the following alternative representation of $A^\star$

$$\begin{aligned} A^\star &= \{0 \cup 01 \cup 10\}^\star && (2.34)\\ &= \{\varepsilon \cup 1\}\{0 \cup 01 \cup 011\}^\star && (2.35)\\ &= \{\varepsilon \cup 1\}B^\star. && (2.36)\end{aligned}$$

Since B forms a uniquely decodable code, and since the sets $\{\epsilon \cup 1\}$ and $B^\star$ fulfill (2.21), we have for the generating series of $\mathcal{A}^\star$

$$\begin{aligned} \mathrm{GEN}_{\mathcal{A}^\star}(x) &= (1+x)\sum_{k=0}^{\infty}(x+x^2+x^3)^k && (2.37)\\ &= \frac{1+x}{1-(x+x^2+x^3)}. && (2.38)\end{aligned}$$

$\triangleleft$

2.3. Generating Series from a List of Forbidden Substrings

In many cases, the specifications of the allowed strings of a DNC consist of a finite alphabet and a finite list of forbidden substrings. All strings generated over the alphabet that do not contain any of the forbidden substrings are accepted by the channel. In [13] and [14], the authors give formulas how to directly derive the generating series for this type of DNCs. The case of non-integer valued string weights is considered in [13].

2.3.1. Correlation Functions

To derive the generating series of a DNC specified by a list of forbidden substrings, the main tool is the correlation function of the substrings from this list. We introduce the correlation function in a simple example.

Example 4. (*Correlation function*). We consider the DNC $\mathcal{A} = (A, w)$, where A is the set of all binary strings that do not contain the substring 101, and where the symbol weights are given by $w(0) = w(1)$. We use a result from [14] in the form of [11, Proposition 1.4]. It states that the set of binary strings not containing a certain pattern p has the generating series

$$f(x) = \frac{c(x)}{x^k + (1-2x)c_{p,p}(x)} \tag{2.39}$$

where k is the length (in bits) of p and where $c_{p,p}(x)$ is the correlation function representing the correlation of p with itself. It is defined as

$$c_{p,p}(x) = \sum_{i=0}^{k-1} c_i x^i \tag{2.40}$$

with c_i given by

$$c_i = \delta[p_{1+i}p_{2+i}\cdots p_k, p_1p_2\cdots p_{k-i}] \tag{2.41}$$

where p_i denotes the ith bit (from the left) of p and where $\delta[a,b] = 1$ if $a = b$ and $\delta[a,b] = 0$ if $a \neq b$. In (2.41), we compare the pattern p with the pattern p shifted to the right by i bits. This can be visualized in the following way:

$$\begin{array}{c|ccccc|c} i & 1 & 0 & 1 & & & c_i \\ \hline 0 & 1 & 0 & 1 & & & 1 \\ 1 & & 1 & 0 & 1 & & 0 \\ 2 & & & 1 & 0 & 1 & 1 \end{array} \tag{2.42}$$

For the pattern p shifted to the right by $i = 0$, the part that overlaps with p is clearly identical to p, i.e., $101 = 101$, and we therefore have $c_0 = 1$. For $i = 1$, the overlapping parts are different, i.e., $01 \neq 10$, and we have $c_1 = 0$. For $i = 2$, the overlapping parts are again identical, i.e., $1 = 1$, which yields $c_2 = 1$. Thus, the correlation function is for $p = 101$ given by

$$c_{p,p}(x) = 1 + x^2. \tag{2.43}$$

The size of p in bits is $k = 3$, according to (2.39), the generating series of $\mathcal{A}$ is therefore given by

$$\mathrm{GEN}_{\mathcal{A}}(x) = \frac{1+x^2}{x^3 + (1-2x)(1+x^2)}. \tag{2.44}$$

$\triangleleft$

2.3.2. Forbidden Substrings

We now consider the general case. Let $\mathcal{A} = (A, w)$ denote a DNC and let D denote a finite alphabet. Let L denote a finite list of finite strings. Assume that the set A contains all strings from $D^\star$ that do not have substrings in L. What is the generating series of the DNC $\mathcal{A}$? The answer is given by a formula in [13]. Here, we will only explain how to apply this formula, and we will not discuss its derivation. Let s_i and s_j denote two strings from the list of forbidden substrings L. The correlation function $c_{s_i,s_j}(x)$ is defined as the generating series of the correlation between the string s_i and the string s_j. It is difficult to give a simple representation of the form (2.42) for the correlation between strings of possibly non-integer valued weights. We therefore define the correlation between two strings in a slightly different way. Assume that s_i can be written as $s_i = ab$ and that s_j can be written as bc for a non-empty string b. We denote by E the set of all a for which such a decomposition of s_i and s_j exists. Basically, the substring b corresponds to the overlapping parts in (2.42). The correlation function $c_{s_i,s_j}(x)$ is then defined as

$$c_{s_i,s_j}(x) = \sum_{a \in E} x^{w(a)}. \tag{2.45}$$

For example, assume that the alphabet is given by $D = \{0, 1\}$ with $w(0) = 0.5$ and $w(1) = 1$ and assume that the list of forbidden strings is given by $L = \{s_1, s_2\}$ with $s_1 = 10101$ and $s_2 = 1010$. Then the correlation functions are

$$c_{s_1,s_1}(x) = 1 + x^{1.5} + x \qquad c_{s_1,s_2}(x) = x^{1.5} + x \tag{2.46}$$

$$c_{s_2,s_1}(x) = 1 + x^{1.5} \qquad c_{s_2,s_2}(x) = 1 + x^{1.5}. \tag{2.47}$$

We define the correlation matrix $\mathbf{T}$ by

$$[\mathbf{T}(x)]_{i,j} = c_{s_i,s_j}(x). \tag{2.48}$$

In our example, the correlation matrix is given by

$$\mathbf{T}(x) = \begin{pmatrix} 1 + x^{1.5} + x & x^{1.5} + x \\ 1 + x^{1.5} & 1 + x^{1.5} \end{pmatrix}. \tag{2.49}$$

Let $D(x)$ denote the generating series of D. In our example, $D(x)$ is equal

$$D(x) = x^{0.5} + x. \tag{2.50}$$

We define the $L \times L$ matrix $\mathbf{\Phi}(x)$ with the diagonal elements defined as

$$[\mathbf{\Phi}(x)]_{i,i} = x^{w(s_i)} \tag{2.51}$$

and the off-diagonal elements equal zero. The diagonal entries can be interpreted as the generating series of the forbidden substrings from L. In our example, $\mathbf{\Phi}(x)$ is equal

$$\mathbf{\Phi}(x) = \begin{pmatrix} x^4 & 0 \\ 0 & x^3 \end{pmatrix}. \tag{2.52}$$

According to [13], the generating series of the DNC $\mathcal{A}$ is now given by

$$\mathrm{GEN}_{\mathcal{A}}(x) = \frac{1}{1 - D(x) + \mathbf{1}^T \mathbf{T}^{-1}(x) \mathbf{\Phi}(x) \mathbf{1}} \tag{2.53}$$

where $\mathbf{1}$ denotes a column vector with all components equal to one.

2.3.3. Pattern Codes

In the case where we only have one forbidden string p, (2.53) becomes

$$\mathrm{GEN}_{\mathcal{A}}\,(x) = \frac{1}{1 - D(x) + c(x)^{-1}x^{w(p)}} \tag{2.54}$$

$$= \frac{c(x)}{[1 - D(x)]c(x) + x^{w(p)}} \tag{2.55}$$

where we denote the correlation function of p by $c(x) = c_{p,p}(x)$. A variation of the DNC $\mathcal{A}$ is the DNC $\mathcal{B} = (B, w)$ that allows strings that end with the string p but do not contain it elsewhere. The generating series of the DNC $\mathcal{B}$ is given by

$$\mathrm{GEN}_{\mathcal{B}}\,(x) = \frac{x^{w(p)}}{[1 - D(x)]c(x) + x^{w(p)}}. \tag{2.56}$$

For a derivation of this formula based on [14] see [11, p. 57]. A concatenation of strings from B can be uniquely decomposed into the strings from B by looking for the substring p. The set B therefore forms a pattern code. For a discussion of this type of codes, see [20].

2.4. Shannon's Telegraphy Channel

Claude Shannon started his legendary paper [1] with the consideration of a telegraphy channel. Many authors have since then used Shannon's telegraphy channel to illustrate their results about discrete noiseless channels. To follow this tradition, we will illustrate our results by applying them to the telegraphy channel. Shannon's telegraphy channel is a DNC and we denote it by $\mathcal{T} = (T, w)$. In this section, we derive the generating series of $\mathcal{T}$. The set T consists of strings over the alphabet $\{d, D, s, S\}$. The symbol s represents a letter-space and the symbol S represents a word-space. The symbol weights are given by

$$w(d) = 2 \qquad w(D) = 4 \qquad w(s) = 3 \qquad w(S) = 6. \tag{2.57}$$

The specifications of $\mathcal{T}$ are:

1. The strings cannot contain more than one consecutive s and the strings cannot contain more than one consecutive S.
2. The symbol s cannot be followed by the symbol S and vice versa.

For example, if we encode a dot by the symbol d, a dash by the symbol D, a letter-space by the symbol s and a word-space by the symbol S, we can transmit a Morse-message over the channel $\mathcal{T}$. We denote by T_d the set of allowed strings starting with d or D and we denote by T_s the set of allowed strings starting with s or S. We define the DNCs $\mathcal{T}_d = (T_d, w)$ and $\mathcal{T}_s = (T_s, w)$. For the telegraphy channel $\mathcal{T}$ we then have $\mathcal{T} = \mathcal{T}_d \cup \mathcal{T}_s$ and since $T_d \cap T_s = \emptyset$, we have, according to Lemma 1, for the generating series of $\mathcal{T}$

$$\mathrm{GEN}_{\mathcal{T}}\,(x) = \mathrm{GEN}_{\mathcal{T}_d}\,(x) + \mathrm{GEN}_{\mathcal{T}_s}\,(x)\,. \tag{2.58}$$

The set T_d results from concatenating symbols from the set

$$\{d, D, ds, Ds, dS, DS\} \tag{2.59}$$

which is uniquely decodable since it is suffix-free. We therefore know from Lemma 3 that the set T_d is uniquely generated by

$$T_d = \{d, D, ds, Ds, dS, DS\}^\star \tag{2.60}$$

and we get for the generating series of $\mathcal{T}_d$

$$\mathrm{GEN}_{\mathcal{T}_d}(x) = \sum_{k=0}^{\infty} \left(x^2 + x^4 + x^5 + x^7 + x^8 + x^{10}\right)^k . \tag{2.61}$$

Since $T_s = \{s \cup S\} T_d$, the generating series of $\mathcal{T}_s$ is given by

$$\mathrm{GEN}_{\mathcal{T}_s}(x) = (x^3 + x^6)\mathrm{GEN}_{\mathcal{T}_d}(x) . \tag{2.62}$$

For the generating function of the telegraphy channel $\mathcal{T}$ we finally get

$$\begin{aligned}
\mathrm{GEN}_{\mathcal{T}}(x) &= \mathrm{GEN}_{\mathcal{T}_d}(x) + \mathrm{GEN}_{\mathcal{T}_s}(x) && (2.63)\\
&= \mathrm{GEN}_{\mathcal{T}_d}(x) + (x^3 + x^6)\mathrm{GEN}_{\mathcal{T}_d}(x) && (2.64)\\
&= (1 + x^3 + x^6) \sum_{k=0}^{\infty} \left(x^2 + x^4 + x^5 + x^7 + x^8 + x^{10}\right)^k && (2.65)\\
&= \frac{1 + x^3 + x^6}{1 - x^2 - x^4 - x^5 - x^7 - x^8 - x^{10}} . && (2.66)
\end{aligned}$$

Note that this results coincides with the generating function of Shannon's telegraphy channel as given in [13].

3. Asymptotic Analysis

We return to the fundamental question of our work. How many distinct strings of the same weight are accepted by a DNC? In the last chapter, we showed that the exact answer to this question is given by the coefficients $N[w]$ of the generating series and that these coefficients can be obtained by algebraic means. In many cases however, we are not interested in the exact number $N[w]$ of distinct strings of the same weight—we are rather interested in the asymptotic behavior of $N[w]$ for w getting large. In this chapter, we will introduce the concept of generating functions and we will show that approximations of $N[w]$ can be derived from the generating function by analytic means. How to quantify combinatorial structures by mathematical analysis is in detail discussed in [11]. In this chapter, we generalize the main results from [11, ch. 4] to the case where the strings accepted by a DNC take non-integer values.

3.1. Information-Theoretic Criteria for DNCs of Interest

We will in detail discuss information-theoretic aspects of DNCs in the next chapter. Here, we specify the class of DNCs of interest by requiring that their combinatorial capacity is well-defined.

According to Definition 2, the generating series of a DNC $\mathcal{A} = (A, w)$ is given by

$$\mathrm{GEN}_{\mathcal{A}}(x) = \sum_{s \in A} x^{w(s)}. \tag{3.1}$$

We order and index the weights $w \in w(A)$ such that

$$w(A) = \{w_k\}_{k=1}^{\infty} \tag{3.2}$$

and

$$w_1 < w_2 < w_3 < \cdots . \tag{3.3}$$

We can then write the generating series as

$$\mathrm{GEN}_{\mathcal{A}}(x) = \sum_{k=1}^{\infty} N[w_k] x^{w_k}. \tag{3.4}$$

The generating series of a DNC was introduced as an algebraic representation. We did not impose any restriction on the set of possible symbol weights $\{w_k\}_{k=1}^{\infty}$ and we did not impose any restriction on the set of the numbers of distinct strings of the same weight $\{N[w_k]\}_{k=1}^{\infty}$.

The DNC was originally introduced as a model for a class of communication channels. An important performance measure for a general communication channel is its capacity, see [1]. The capacity of a DNC was in [1] and [7] defined as

$$C = \limsup_{k\to\infty} \frac{\log N[w_k]}{w_k}. \tag{3.5}$$

From its definition, the capacity of a DNC is of pure combinatorial nature. In [7], the authors show that the capacity of a DNC is under certain conditions equal to the maximum rate of information per string weight at which data can be transmitted over a DNC. We will generalize this property in Chapter 4. We restrict our interest to DNCs that have their capacity actually given by (3.5). This implies for the number sequence of possible string weights $\{w_k\}_{k=1}^{\infty}$ that it increases with at most polynomial speed in the sense that for any integer $n \geq 0$

$$\max_{w_k<n} k < Ln^K \tag{3.6}$$

for some constant $K > 0$ and some constant $L > 0$. Otherwise, the density of $\{w_k\}_{k=1}^{\infty}$ increases exponentially fast and the capacity of the considered DNC is no longer given by the combinatorial capacity as defined in (3.5). We illustrate this in the following example.

Example 5. (*Too dense string weights*). Let $\{N[w_k]\}_{k=1}^{\infty}$ denote the coefficients of the generating series of some DNC. Assume $N[w_k] = 1$ for all $k \in \mathbb{N}$ and assume

$$\max_{w_k<n} k = \lceil R^n \rceil, \qquad \forall n \in \mathbb{N} \tag{3.7}$$

for some $R > 0$. According to (3.5), the combinatorial capacity of the considered DNC is equal 0 because of $\log N[w_k] = 0$ for all $k \in \mathbb{N}$. However, the channel accepts R^n distinct strings of weight smaller than n. The average amount of data per string weight that we can transmit over the channel is thus lower-bounded by $\log R^n/n = R$ which is, according to the assumption, larger than 0.

We still assume $N[w_k] = 1$ for all $k \in \mathbb{N}$, but we assume for the possible string weights

$$\max_{w_k<n} k < Ln^K \tag{3.8}$$

for some positive and constant K and L. The average amount of data per string weight that we can transmit over the channel is now upper-bounded by

$$\frac{\log Ln^K}{an} \tag{3.9}$$

since there are at most Ln^K distinct strings of weight smaller or equal n and since the average weight of strings of weight smaller or equal n increases linearly with n. For $n \to \infty$, (3.9) goes to zero, which is in accordance with the value we obtained for the considered DNC by applying (3.5). ◁

We conclude from the example that the capacity of a DNC with a set of possible string weights $\{w_k\}_{k=1}^{\infty}$ increasing exponentially in density is not given by its capacity as defined in (3.5).

Note 1. The authors of [21] give a definition for the capacity of DNCs that would not lead to the problem we encountered in Example 5. However, the technique we will introduce in the following does not apply in general when using the definition of capacity from [21].

The restriction we have on the set of the numbers of distinct strings of the same weight $\{N[w_k]\}_{k=1}^{\infty}$ is that it increases at most exponentially in w_k such that for some finite R

$$N[w_k] < R^{w_k} \tag{3.10}$$

almost everywhere. Otherwise, the limit in (3.5) would not exist. There are combinatorial structures that do not fulfill this restriction. We illustrate this in the following example.

Example 6. (*Infinite combinatorial capacity*). We start by considering the DNC $\mathcal{A} = (\{0,1\}^\star, w)$ with $w(0) = w(1) = 1$. For the coefficients of the generating series we have $N[k] = 2^k$. Let s be a string in $\{0,1\}^\star$ of weight $|s| = k$. We assign labels to the k bits forming s. We do this by assigning the integer numbers from 1 to k to the k bits in an arbitrary order, e.g., for $s = 01001$ a possible assignment is

$$0_5 1_4 0_1 0_3 1_2. \tag{3.11}$$

The number $M[k]$ of distinct labeled strings of weight k is now given by $M[k] = 2^k k!$. For the combinatorial capacity as defined in (3.5) we get

$$\begin{aligned}
\frac{\log M[k]}{k} &= \frac{\log 2^k k!}{k} && (3.12)\\
&= \frac{k \log 2}{k} + \frac{\log k!}{k} && (3.13)\\
&\geq \log 2 + \frac{\log k^k e^{-k}}{k} && (3.14)\\
&\geq \frac{k \log k - k \log e}{k} && (3.15)\\
&= \log k - \log e. && (3.16)
\end{aligned}$$

where we used in (3.14) the lower bound $(n/e)^n$ for the factorial function $n!$. A derivation of this bound can be found in Section B.2.3 in the appendix. The term in (3.16) goes to infinity for $k \to \infty$. It follows that the limit in (3.5) does not exist for the considered combinatorial structure. $\triangleleft$

Note 2. The asymptotic behavior of the complexity of combinatorial structures as the one presented in Example 6 is in detail discussed in [11]. The key-concepts are labelled structures and exponential generating functions.

For the rest of our work, we will only consider DNCs that fulfill (3.6) and (3.10). This restriction is reasonable, since in practice, we almost exclusively consider DNCs that accept strings that result from the concatenation of symbols from a finite set. DNCs of this kind automatically fulfill (3.6) and (3.10). We state this in the following lemma.

Lemma 4. *Let $\mathcal{A} = (A, w)$ represent a DNC with the set of accepted strings A resulting from the concatenations of symbols from a finite set. Let*

$$\mathrm{GEN}_{\mathcal{A}}(x) = \sum_{k=1}^{\infty} N[w_k]x^k \tag{3.17}$$

denote the generating series of $\mathcal{A}$.

i. For any integer $n \geq 0$

$$\max_{w_k<n} < Ln^K \tag{3.18}$$

for some constant $K > 0$ and some constant $L > 0$.

ii. There exist some constant $R > 0$ and some constant $M > 0$ such that

$$N[w_k] < MR^{w_k} \tag{3.19}$$

almost everywhere with respect to k.

Proof. See Appendix B.2.1. □

3.2. Exponential Behavior

A first approximation of the asymptotic behavior of $N[w_k]$ can be obtained by determining its exponential order.

Definition 7. [11, p. 230] Let $\{w_k\}_{k=1}^{\infty}$ be a strictly increasing sequence of nonnegative real numbers. We say that a sequence of nonnegative numbers $\{N[w_k]\}_{k=1}^{\infty}$ is of *exponential order* R^{w_k} if and only if

$$\limsup_{k\to\infty} N[w_k]^{\frac{1}{w_k}} = R. \tag{3.20}$$

This is equivalent to the following: for all ϵ with $R > \epsilon > 0$, the following two properties hold.

i. The number $N[w_k]$ is larger or equal $(R-\epsilon)^{w_k}$ infinitely often (i.o.) with respect to k:

$$N[w_k] \geq (R-\epsilon)^{w_k}, \qquad \text{i.o.} \tag{3.21}$$

ii. The number $N[w_k]$ is smaller or equal $(R+\epsilon)^{w_k}$ almost everywhere (a.e.) with respect to k:

$$N[w_k] \leq (R+\epsilon)^{w_k}, \qquad \text{a.e.} \tag{3.22}$$

If the number sequence $\{N[w_k]\}_{k=1}^{\infty}$ is of exponential order R^{w_k}, we denote this by

$$N[w_k] \bowtie R^{w_k}. \tag{3.23}$$

It is important to take the limit superior in the definition. In general, the limit does not exist, as we can see in the following example:

Example 7. (*Existence of the limit*). Consider the DNC $\mathcal{A} = (\{a, b, c\}^\star, w)$ with $w(a) = 1$, $w(b) = \sqrt{2}$, and $w(c) = \pi$. Because the numbers 1, $\sqrt{2}$, and π are incommensurable, the number $N[w_k]$ of distinct strings of weight w_k accepted by the DNC will always be equal 1 when $w_k \in \mathbb{N}$. However, we will show in Example 11 that the supremum $\sup_{w_k<w} N[w_k]$ increases exponentially like 1.9837^w. Therefore, the limit $k \to \infty$ of $N[w_k]^{1/w_k}$ does not exist. $\triangleleft$

In the following, we will define the generating function and the complex generating function of a DNC and we will show how we can determine the exponential order of the coefficients $N[w_k]$ from these functions by analytic techniques.

Note 3. We will see in the next chapter that the exponential order of the coefficients $N[w_k]$ is equivalent to the combinatorial capacity of the considered DNC.

3.2.1. Exponential Order by Radius of Convergence

The generating series of a DNC as introduced in the last chapter is of pure algebraic nature. We now interpret it as a function of a variable taking real values.

Definition 8. We define the *generating function* of a DNC $\mathcal{A}$ as

$$\mathrm{G}_{\mathcal{A}}(y) = \mathrm{GEN}_{\mathcal{A}}(x)\big|_{x=y}, \qquad y \in \mathbb{R}. \tag{3.24}$$

The exponential order of the coefficients $N[w_k]$ is related to the radius of convergence of $\mathrm{G}_{\mathcal{A}}(y)$.

Lemma 5. *Let $\mathcal{A}$ be a DNC with the generating function*

$$\mathrm{G}_{\mathcal{A}}(y) = \sum_{k=1}^{\infty} N[w_k] y^{w_k}. \tag{3.25}$$

Then

$$N[w_k] \bowtie R^{-w_k} \tag{3.26}$$

where R denotes the radius of convergence of $\mathrm{G}_{\mathcal{A}}(y)$.

To prove this lemma, we need a property of the generating function $\mathrm{G}_{\mathcal{A}}(y)$, which follows from the restriction (3.6) we stated for the density of the possible string weights $\{w_k\}_{k=1}^{\infty}$. We give this property in the following lemma.

Lemma 6. *Assume that the strictly ordered sequence of positive real numbers* $\{w_k\}_{k=1}^{\infty}$ *is not too dense in the sense that for any integer* $n \geq 0$

$$\max_{w_k < n} k \leq L n^K \tag{3.27}$$

for some constant $L > 0$ *and some constant* $K \geq 0$*. Let* ρ *be a nonnegative real number. Then the sum*

$$\sum_{k=1}^{\infty} \rho^{w_k} \tag{3.28}$$

converges if and only if $\rho < 1$*.*

Proof. See Appendix B.2.2. □

Proof of Lemma 5. We can write the generating function(3.25) as

$$\begin{aligned} \mathrm{G}_{\mathcal{A}}(y) &= \sum_{k=1}^{\infty} N[w_k] y^{w_k} && (3.29)\\ &= \sum_{k=1}^{\infty} \left(N[w_k]^{\frac{1}{w_k}} y\right)^{w_k}. && (3.30) \end{aligned}$$

We define the two sets $D(y)$ and $E(y)$ as

$$\begin{aligned} D(y) &= \left\{k \in \mathbb{N} \middle| N[w_k]^{\frac{1}{w_k}} y < 1\right\} && (3.31)\\ E(y) &= \mathbb{N} \setminus D(y) = \left\{k \in \mathbb{N} \middle| N[w_k]^{\frac{1}{w_k}} y \geq 1\right\} && (3.32) \end{aligned}$$

and write the generating function as

$$\mathrm{G}_{\mathcal{A}}(y) = \sum_{k \in D(y)} \left(N[w_k]^{\frac{1}{w_k}} y\right)^{w_k} + \sum_{k \in E(y)} \left(N[w_k]^{\frac{1}{w_k}} y\right)^{w_k}. \tag{3.33}$$

It follows from Lemma 6 that $\mathrm{G}_{\mathcal{A}}(y)$ converges if and only if the set $E(y)$ is finite. The number R is the radius of convergence of $\mathrm{G}_{\mathcal{A}}(y)$, therefore, for any ϵ with $R > \epsilon > 0$, the set $E(R-\epsilon)$ is finite. Since $D(y) = \mathbb{N} \setminus E(y)$, the finiteness of $E(R-\epsilon)$ is equivalent to $k \in D(R-\epsilon)$ a.e.:

$$N[w_k]^{\frac{1}{w_k}} (R-\epsilon) < 1, \qquad \text{a.e.} \tag{3.34}$$

We therefore have

$$N[w_k] < (R-\epsilon)^{-w_k}, \qquad \text{a.e.} \tag{3.35}$$

which implies

$$N[w_k] \leq (R-\epsilon)^{-w_k}, \qquad \text{a.e.} \tag{3.36}$$

For any ϵ with $R > \epsilon > 0$, since R is the radius of convergence of $\mathrm{G}_{\mathcal{A}}(y)$, the set $E(R+\epsilon)$ is infinite:

$$N[w_k]^{\frac{1}{w_k}}(R+\epsilon) \geq 1, \qquad \text{i.o.} \tag{3.37}$$

We thus have shown that for any ϵ with $R > \epsilon > 0$

$$N[w_k] \leq (R-\epsilon)^{-w_k}, \qquad \text{a.e.} \tag{3.38}$$
$$\text{and} \tag{3.39}$$
$$N[w_k] \geq (R+\epsilon)^{-w_k}, \qquad \text{i.o.} \tag{3.40}$$

This is according to Definition 7 equivalent to

$$N[w_k] \bowtie R^{-w_k} \tag{3.41}$$

which concludes the proof. □

The following example illustrates how we can use this lemma to determine the exponential order of the number of distinct strings of the same weight accepted by a DNC.

Example 8. (*Radius of convergence and exponential behavior*). Let $\mathcal{A} = (A, w)$ be a DNC that allows binary strings that do not contain the substring 011. Assume $w(0) = w(1) = 1$. We write $\mathcal{A} = \mathcal{BC}$ where $\mathcal{B} = (\{1\}^\star, w)$ and $\mathcal{C} = (\{0 \cup 01\}^\star, w)$. The generating series of $\mathcal{A}$ is then given by

$$\mathrm{GEN}_{\mathcal{A}}(x) = \mathrm{GEN}_{\mathcal{B}}(x)\,\mathrm{GEN}_{\mathcal{C}}(x) \tag{3.42}$$
$$= \left(\sum_{k=0}^{\infty} x^k\right)\left(\sum_{l=0}^{\infty}(x+x^2)^l\right). \tag{3.43}$$

For the generating function of $\mathcal{A}$ we get

$$\mathrm{G}_{\mathcal{A}}(y) = \left(\sum_{k=0}^{\infty} y^k\right)\left(\sum_{l=0}^{\infty}(y+y^2)^l\right). \tag{3.44}$$

The radius of convergence of the power series is the smallest solution of the equation

$$y + y^2 = 1 \tag{3.45}$$

which is given by

$$R = \frac{-1+\sqrt{5}}{2}. \tag{3.46}$$

We thus have

$$N[k] \bowtie \left(\frac{-1+\sqrt{5}}{2}\right)^{-k} \tag{3.47}$$
$$= 0.61803^{-k}. \tag{3.48}$$

◁

3.2.2. Exponential Order by Leftmost Singularity

We have seen in the last chapter that in many cases, we can derive a closed-form representation of the generating series of a DNC without explicitly using its series representation. We will show that the exponential order of the coefficients $N[w_k]$ of the series representation can be determined directly from the closed-form representation of the generating function. To do this, we have to go to the complex plane. Singularities of analytic functions will play a key-role. To cite [22], "Und noch ein Geheimnis: In den isolierten Singularitäten einer Funktion f [...] ist globale Information über f codiert." ("One more secret: Global informations about a function f are encoded in its isolated singularities"). In this sense, functions from the complex plane to the complex plane are very different from functions with a real argument. We expect the reader to be familiar with basic concepts from complex analysis. Our reference is [22], an easy-written introduction into complex analysis with an emphasis on topics of interest for engineers. The course is written in German. Some definitions and identities from complex analysis we list in Appendix B.1. Our derivations are highly inspired by [11, ch. 4], with the important difference that we include non-integer weighted string weights in our derivations. The most important consequence is that we have to use a generating function different from the one used in [11]. In the case of integer valued string weights, the generating series of a DNC can directly be evaluated in the complex plane to obtain a well-defined analytic function of the form

$$\mathrm{G}_{\mathcal{A}}(z) = \sum_{k=0}^{\infty} N[k] z^k, \qquad z \in \mathbb{C}. \tag{3.49}$$

This series representation can be interpreted as the Taylor series expansion of $\mathrm{G}_{\mathcal{A}}(z)$ around 0, and as a consequence, powerful theorems such as Pringsheims's Theorem [11] and the Exponential Growth Formula [11] can be applied directly. However, in the case of non-integer valued symbol weights, evaluating the generating series in $\mathbb{C}$ does not result in a well-defined analytic function. Expressions of the form z^r, $z \in \mathbb{C}$ and $r \in \mathbb{R}$, are a priori not well-defined and it is not possible to define them such that the resulting function is analytic in 0. For a discussion of this problem see Appendix B.1.3. To circumvent this problem, we evaluate the generating series in $x = e^s$, $s \in \mathbb{C}$. Terms of the form x^r, $r \in \mathbb{R}$ become terms of the form e^{rs}, which are analytic in $\mathbb{C}$. We can relate the substitutions $x = z$ and $x = e^s$ to each other by defining the complex logarithm as

$$\operatorname{Log} v = \left\{ z \middle| \exp(z) = v, -\pi < \Im\{z\} \leq \pi \right\}. \tag{3.50}$$

The function $z = \operatorname{Log} v$ maps the punctured complex plane $\mathbb{C} \setminus \{0\}$ to a horizontal stripe with $-\pi < \Im\{z\} \leq \pi$. See Figure 3.1 for an illustration.

Definition 9. We define the *complex generating function* of a DNC $\mathcal{A}$ as

$$\mathrm{G}_{\mathcal{A}}(e^s) = \mathrm{GEN}_{\mathcal{A}}(x)\big|_{x=e^s}, \qquad s \in \mathbb{C}. \tag{3.51}$$

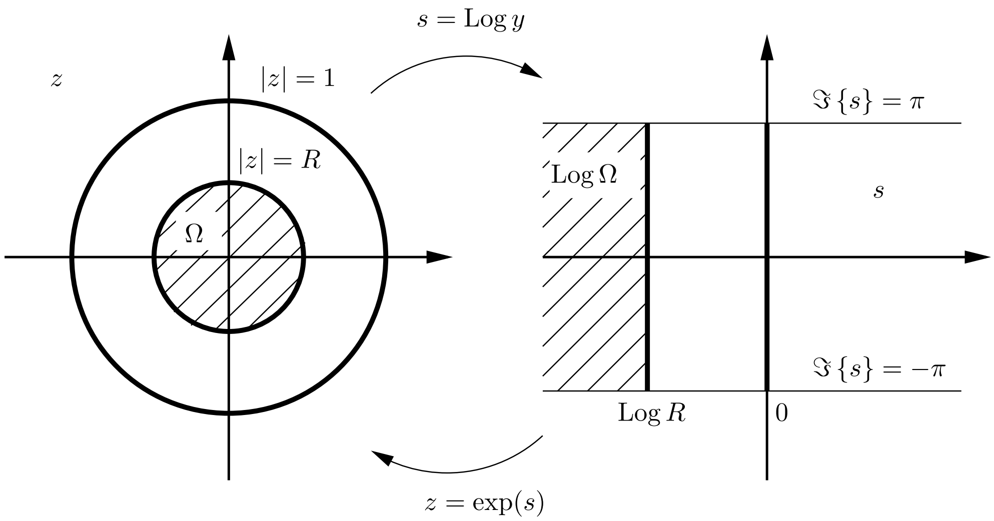

Figure 3.1.: Mapping from the punctured complex plane $z \in \mathbb{C} \setminus \{0\}$ to a horizontal stripe with $-\pi < \Im\{s\} \leq \pi$.

In literature, the generating functions of combinatorial structures appear in many forms. In [14], we find expressions of the form

$$\sum_{k=0}^{\infty} N[k] z^{-k} \tag{3.52}$$

where the minus sign is probably inspired by the Z Transform. In [11], the authors write

$$\sum_{k=0}^{\infty} N[k] z^{k} \tag{3.53}$$

and in [7] and [13], the authors chose

$$\sum_{k=1}^{\infty} N[w_k] e^{-w_k s}. \tag{3.54}$$

We defined the complex generating function in yet another way. As stated above, the forms (3.52) and (3.53) do not serve for our purpose since they are not well-defined for $z \in \mathbb{C}$. We did not use the form (3.54), since we want the coefficients to remain nonnegative when we differentiate the complex generating function (we will need this property in a later proof). In the form (3.54), the complex generating function does not have this property. This can for example be seen from the first differentiation of the summands in (3.54), which is given by

$$\frac{\mathrm{d}}{\mathrm{d}s} N[k] e^{-w_k s} = -w_k N[k] e^{-w_k s}. \tag{3.55}$$

Before we can show how the exponential order of the coefficients $N[w_k]$ is related to the complex generating function, we need the following generalization of Pringsheim's Theorem, which can be found in its original form in [11].

Theorem 1. *Let $\mathcal{A}$ be a DNC with the complex generating function*

$$\mathrm{G}_{\mathcal{A}}\left(e^s\right) = \sum_{k=1}^{\infty} N[w_k] e^{w_k s}. \tag{3.56}$$

If the region of convergence (r.o.c.) of $\mathrm{G}_{\mathcal{A}}\left(e^s\right)$ is determined by $\Re\{s\} < S$, then $\mathrm{G}_{\mathcal{A}}\left(e^s\right)$ has a singularity in $s = S$.

Proof. Suppose in contrary that $\mathrm{G}_{\mathcal{A}}\left(e^s\right)$ is analytic in $s = S$ implying that it is analytic in a disc of radius r centered at S. We choose a number h such that $0 < h < r/3$ and consider the Taylor expansion of $\mathrm{G}_{\mathcal{A}}\left(e^s\right)$ around $s_0 = S - h$:

$$\mathrm{G}_{\mathcal{A}}\left(e^s\right) = \sum_{n=0}^{\infty} \frac{\left[\mathrm{G}_{\mathcal{A}}\left(e^{s_0}\right)\right]^{(n)}}{n!}(s - s_0)^n, \tag{3.57}$$

$$= \sum_{n=0}^{\infty} \frac{\sum\limits_{k=1}^{\infty} N[w_k] w_k^n e^{w_k s_0}}{n!}(s - s_0)^n. \tag{3.58}$$

For $s = S + h$, this is according to our supposition a converging double sum with positive terms and we can reorganize it in any way we want. We thus have convergence in

$$\mathrm{G}_{\mathcal{A}}\left(e^{S+h}\right) = \sum_{n=0}^{\infty} \frac{\sum\limits_{k=1}^{\infty} N[w_k] w_k^n e^{w_k s_0}}{n!}(2h)^n \tag{3.59}$$

$$= \sum_{k=1}^{\infty} N[w_k] e^{w_k s_0} \sum_{n=0}^{\infty} \frac{w_k^n (2h)^n}{n!} \tag{3.60}$$

$$= \sum_{k=1}^{\infty} N[w_k] e^{w_k s_0} e^{w_k 2h} \tag{3.61}$$

$$= \sum_{k=1}^{\infty} N[w_k] e^{w_k (S+h)}. \tag{3.62}$$

But convergence in the last line contradicts that the r.o.c. of $\mathrm{G}_{\mathcal{A}}(s)$ is strictly given by $\Re\{s\} < S$. □

We are now ready to state the fundamental theorem of this chapter. For DNCs that accept strings with integer valued weights only, the corresponding theorem is the Exponential Growth Formula as given in [11].

Theorem 2. *Let $\mathcal{A}$ be a DNC with the complex generating function*

$$\mathrm{G}_{\mathcal{A}}\left(e^{s}\right)=\sum_{k=1}^{\infty} N[w_k]e^{w_k s}. \tag{3.63}$$

If $\mathrm{G}_{\mathcal{A}}\left(e^{s}\right)$ has its leftmost real singularity in $s=\ln Q$, then the exponential order of the coefficients $N[w_k]$ is given by

$$N[w_k] \bowtie Q^{-w_k}. \tag{3.64}$$

Proof. If $\mathrm{G}_{\mathcal{A}}\left(e^{s}\right)$ has its leftmost real singularity in $s=\ln Q$, then the r.o.c. of $\mathrm{G}_{\mathcal{A}}\left(e^{s}\right)$ is according to Theorem 1 given by $\Re\{s\}<\ln Q$. We write

$$\begin{aligned}
\mathrm{G}_{\mathcal{A}}\left(e^{s}\right)&=\sum_{k=1}^{\infty} N[w_k]e^{w_k s} &(3.65)\\
&\leq \sum_{k=1}^{\infty} |N[w_k]e^{w_k s}| &(3.66)\\
&= \sum_{k=1}^{\infty} N[w_k]|e^{w_k s}| &(3.67)
\end{aligned}$$

where we have equality in (3.67) because the coefficients $N[w_k]$ are all nonnegative, and where we have equality in (3.66) if s is real. The complex generating function $\mathrm{G}_{\mathcal{A}}\left(e^{s}\right)$ evaluated in $s=\ln y$ is equal to the generating function $\mathrm{G}_{\mathcal{A}}\left(y\right)$. It follows that if the r.o.c. of $\mathrm{G}_{\mathcal{A}}\left(e^{s}\right)$ is given by $\Re\{s\}<\ln Q$, then the radius of convergence of $\mathrm{G}_{\mathcal{A}}\left(y\right)$ is given by $R=Q$. Using Lemma 5, we have for the exponential order of $N[w_k]$

$$\begin{aligned}
N[w_k] &\bowtie R^{-w_k} &(3.68)\\
&= Q^{-w_k} &(3.69)
\end{aligned}$$

which concludes the proof. □

3.2.3. Exponential Order by Smallest Positive Pole

The most important application of Theorem 2 we formulate in the following corollary.

Corollary 1. *Suppose that the generating series of a DNC $\mathcal{A}$ can be written as*

$$\mathrm{GEN}_{\mathcal{A}}\left(x\right)=\frac{n_1x^{\tau_1}+n_2x^{\tau_2}+\cdots+n_px^{\tau_p}}{d_1x^{\nu_1}+d_2x^{\nu_2}+\cdots+d_qx^{\nu_q}}, \qquad \tau_1,\ldots,\tau_p,\nu_1,\ldots,\nu_q\in\mathbb{R}^{\oplus} \tag{3.70}$$

for some finite p and q. The exponential order of the coefficients $N[w_k]$ is then given by

$$N[w_k] \bowtie P^{-w_k} \tag{3.71}$$

where P is the smallest positive pole of $\mathrm{G}_{\mathcal{A}}\left(y\right)$, which results from evaluating $\mathrm{GEN}_{\mathcal{A}}\left(x\right)$ in $y=x$, $y\in\mathbb{R}$.

Note 4. The corollary was already stated in [13, Theorem 1]. At an important step in the proof, which states that the smallest positive pole of $\mathrm{G}_{\mathcal{A}}(y)$ determines the region of convergence of the series representation of $\mathrm{G}_{\mathcal{A}}(y)$, the authors refer to [7]. It is not clear to us if this reference applies, since it refers to a statement about the convergence of matrix power series of the form

$$\sum_{k=0}^{\infty} \mathbf{A}^k \tag{3.72}$$

where $\mathbf{A}$ is a square matrix. This is different from the context considered in our work and it is also different from the context considered in [13]. We believe that the generalization of Pringsheim's Theorem we gave in Theorem 1 is essential for the complete proof of Corollary 1.

Proof of Corollary 1. If the generating series $\mathrm{GEN}_{\mathcal{A}}(x)$ is of the form (3.70), the complex generating function $\mathrm{G}_{\mathcal{A}}(e^s)$, which results from evaluating $\mathrm{GEN}_{\mathcal{A}}(x)$ in $x = e^s$, is meromorphic, which implies that all its singularities are poles. The substitution $y = e^s$, for s real, is a one-to-one mapping from the real axis to the positive real axis. Therefore, if $\ln Q$ is the leftmost real pole of $\mathrm{G}_{\mathcal{A}}(e^s)$, then $P = Q$ is the smallest positive pole of the generating function $\mathrm{G}_{\mathcal{A}}(y)$. Applying Theorem 2, we get for the exponential order of $N[w_k]$

$$\begin{aligned} N[w_k] &\bowtie Q^{-w_k} && (3.73)\\ &= P^{-w_k} && (3.74) \end{aligned}$$

which concludes the proof. □

We illustrate the application of this corollary by using it to solve the problem from Example 8 in a different way.

Example 9. (*Smallest positive pole and exponential behavior*). As in Example 8, let $\mathcal{A} = (A, w)$ be a DNC that allows binary strings that do not contain the substring 011. Assume $w(0) = w(1) = 1$. We know from (2.55) that the generating series of binary strings that do not contain a certain pattern p is given by

$$\mathrm{GEN}_{\mathcal{A}}(x) = \frac{c(x)}{x^l + (1-2x)c(x)} \tag{3.75}$$

where l is the weight of the pattern p and where $c(x)$ is the autocorrelation function of p. For $p = 011$ we get $l = 3$ and $c(x) = 1$. We evaluate $\mathrm{GEN}_{\mathcal{A}}(x)$ in $x = y$, $y \in \mathbb{R}$ and get for the generating function of $\mathcal{A}$

$$\mathrm{G}_{\mathcal{A}}(y) = \frac{1}{y^3 + (1-2y)}. \tag{3.76}$$

The smallest positive pole of $\mathrm{G}_{\mathcal{A}}(y)$ is given by $P = 0.61803$. According to Corollary 1, we have

$$N[k] \bowtie P^{-k} \tag{3.77}$$
$$= 0.61803^{-k}. \tag{3.78}$$

We note that this result coincides with the result from Example 8. $\triangleleft$

3.3. Sub-Exponential Behavior for Integer String Weights

In the last section, we have seen how we can derive the exponential order of the coefficients $N[w_k]$ of the generating series of a DNC from the corresponding (complex) generating function of the DNC by analytic methods. In the following, we will extend this approach and we will show how we can derive asymptotic approximations of $N[w_k]$ with sub-exponential precision. The simplest case is when the considered DNC allows strings of integer valued weights only and has a generating function that is rational. We will therefore start this section with a summary of some properties of rational functions important to our work. A more detailed discussion of rational functions can be found in [11].

3.3.1. Expansion of Rational Functions

A rational function $f : \mathbb{C} \mapsto \mathbb{C}$ is a function that can be written as

$$f(z) = \frac{N(z)}{D(z)} \tag{3.79}$$

where $N(z)$ and $D(z)$ are polynomials in z. In the following, we consider rational functions f that are analytic in zero, which is equivalent to $|f(0)| < \infty$. This is guaranteed by

$$[z^0]D(z) \neq 0 \tag{3.80}$$

which implies that f has no pole in 0. As in Chapter 2, $[z^k]D(z)$ denotes the coefficient of the exponential term z^k in $D(z)$.

Note 5. If a function $f(z)$ is analytic in zero, then it has according to Theorem 10 a Taylor series expansion around zero given by

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} z^k. \tag{3.81}$$

It follows from this equation that $[z^k]f(z) = f^{(k)}(0)/k!$. We call the number sequence $\{f^{(k)}(0)/k!\}_{k=0}^{\infty}$ the *coefficients* of $f(z)$, which is short for "the coefficients of the Taylor expansion of $f(z)$ around zero".

Partial Fraction Expansion

Any rational function $f(z) = N(z)/D(z)$ that is analytic in zero has a partial fraction expansion of the form

$$f(z) = Q(z) + \sum_{\alpha} \sum_{r=1}^{r_\alpha} \frac{c_{\alpha,r}}{(z-\alpha)^r} \tag{3.82}$$

where $Q(z)$ is a polynomial of degree $k_0 = \deg(N) - \deg(D)$. We assign to α the poles of $f(z)$ and for each pole, and we let r take the values $1, \dots, r_\alpha$ where r_α denotes the multiplicity of the pole α. The coefficients $c_{\alpha,r}$ take complex values. We are finally interested in the coefficients $[z^k]f(z)$ of the expansion of $f(z)$ around zero for large k. The term $Q(z)$ does not contribute to the value of $[z^k]f(z)$ for $k > k_0$. Without loss of generality, we can therefore assume in the following that $\deg(D) > \deg(N)$. In the term on the right-hand side of (3.82), we have for each summand

$$\frac{c_{\alpha,r}}{(z-\alpha)^r} = \frac{c_{\alpha,r}}{(-\alpha)^r(1-\frac{z}{\alpha})^r} \tag{3.83}$$

$$= \frac{c_{\alpha,r}}{(-\alpha)^r} \sum_{k=0}^{\infty} \binom{k+r-1}{r-1} \left(\frac{z}{\alpha}\right)^k \tag{3.84}$$

where equality in the last line follows from Newton's expansion as given in Theorem 14 in the appendix. We thus have

$$[z^k]\frac{c_{\alpha,r}}{(z-\alpha)^r} = c_{\alpha,r}(-1)^r \binom{k+r-1}{r-1} \alpha^{-k-r}. \tag{3.85}$$

This leads to the following theorem:

Theorem 3. [11, p. 243] (Expansion of rational functions). *If $f(z)$ is a rational function that is analytic in zero and has poles at points $\alpha_1, \alpha_2, \dots, \alpha_m$, then its coefficients are a sum of exponential polynomials. There exist m polynomials $\{\Pi_j(k)\}_{j=1}^m$ in k such that, for k larger than some fixed k_0,*

$$[z^k]f(z) = \sum_{j=1}^{m} \Pi_j(k)\alpha_j^{-k}. \tag{3.86}$$

Furthermore, the degree of Π_j is equal to $r_j - 1$ where r_j is the multiplicity of the pole of $f(z)$ at α_j.

Note 6. The partial fraction expansion is a tool frequently used in engineering. For example, it is used to find the inverse Laplace transform of transfer functions of dynamic systems, see [23]. In the same way, it also helps when inverting the Z transform of time discrete signals, see [24]. The main task when performing the partial fraction expansion is the determination of the coefficients $c_{\alpha,r}$ in (3.82). This can either be done by hand applying an appropriate algorithm or it can be done by using the built-in functions of mathematical computer programs such as MATHEMATICA or MAPLE. Here, we will only treat explicitly the simplest case where all poles are of multiplicity one.

Laurent Series Expansion

Let $f(z) = N(z)/D(z)$ be a rational function with $\deg(N) < \deg(D)$. Equation (3.82) then becomes

$$f(z) = \sum_{\alpha} \sum_{r=1}^{r_\alpha} \frac{c_{\alpha,r}}{(z-\alpha)^r}. \tag{3.87}$$

For every pole α, the inner sum is equal to the main part of the Laurent series expansion of $f(z)$ around α, see Theorem 11. The coefficient $c_{\alpha,r}$ is thus equal to the coefficient c_{-r} given by the formula in Theorem 11. We conclude that a rational function $f(z) = N(z)/D(z)$ with $\deg(N) < \deg(D)$ can be written as the sum of the main parts of its Laurent series expansions around its poles. For the calculation of the coefficient $c_{\alpha,1}$ in the case where the pole α is of multiplicity 1, we have the following lemma.

Lemma 7. [22, p. 112]. *Let f be analytic in a region $\Omega \setminus a$. Denote by c_{-1} the coefficient of the term $1/(z-a)$ in the Laurent series expansion of $f(z)$ near a. Then*

i. If f has a pole of multiplicity one in $z = a$, then

$$c_{-1} = \lim_{z \to a} (z-a) f(z). \tag{3.88}$$

ii. If f can be written as

$$f(z) = \frac{p(z)}{q(z)} \tag{3.89}$$

where p and q are analytic in Ω and q has a zero of multiplicity one in $z = a$, then

$$c_{-1} = \frac{p(a)}{q'(a)}. \tag{3.90}$$

3.3.2. Sub-Exponential Behavior

We now use the properties of rational functions to determine the sub-exponential behavior of a general DNC $\mathcal{A}$ that has a generating series of the form

$$\mathrm{GEN}_{\mathcal{A}}(x) = \frac{N(x)}{D(x)} \tag{3.91}$$

where $N(x)$ and $D(x)$ are polynomials in the indeterminate x. Since expressions of the form

$$z^k, \qquad z \in \mathbb{C},\ k \in \mathbb{N} \tag{3.92}$$

are well-defined and analytic in $\mathbb{C}$, we can define the generating function of the DNC $\mathcal{A}$ as

$$\mathrm{H}_{\mathcal{A}}(z) = \mathrm{GEN}_{\mathcal{A}}(x)\big|_{x=z}, \qquad z \in \mathbb{C}. \tag{3.93}$$

The generating function $\mathrm{H}_{\mathcal{A}}(z)$ has a closed-form representation of the form

$$\mathrm{H}_{\mathcal{A}}(z) = \frac{N(z)}{D(z)} \tag{3.94}$$

with $N(z)$ and $D(z)$ being polynomials in z, and it also has a series representation of the form

$$\mathrm{H}_{\mathcal{A}}(z) = \sum_{k=1}^{\infty} N[k]z^k. \tag{3.95}$$

Our goal is to determine the sub-exponential behavior of the coefficients $N[k]$ in (3.95) from the analytic characteristics of (3.94). Since there is only one empty string and since every DNC accepts the empty string, we have $N[0] = 1$, which implies $\mathrm{H}_{\mathcal{A}}(0) = 1$. The rational function $\mathrm{H}_{\mathcal{A}}(z)$ is thus finite in zero, which shows that it is analytic in zero. We remember that $[z^k]\mathrm{H}_{\mathcal{A}}(z) = N[k]$. According to Theorem 3 we thus have

$$N[k] = \sum_{j=1}^{m} \Pi_j(k)\alpha_j^{-k}. \tag{3.96}$$

Without loss of generality, we assume for the poles

$$P = |\alpha_1| = |\alpha_2| = \cdots = |\alpha_q| < |\alpha_{q+1}| \leq \cdots \leq |\alpha_m| \tag{3.97}$$

and write

$$N[k] = \sum_{j=1}^{q} \Pi_j(k)\alpha_j^{-k} + \mathrm{O}(\alpha_{q+1}^{-k}). \tag{3.98}$$

For a real function $f(x)$, we mean by

$$f(x) = \mathrm{O}\big(g(x)\big) \tag{3.99}$$

that the ratio $f(x)/g(x)$ stays bounded for $x \to \infty$, i.e., there exists some constant x_0 and some $K > 0$ such that

$$|f(x)| \leq |Kg(x)|, \qquad \forall x > x_0. \tag{3.100}$$

As we can see from (3.98), Theorem 3 allows us to approximate the coefficients $N[k]$ and characterize the error term. The dominant term consists of the exponential terms α_j^{-k} with $|\alpha_j| = P$ and the sub-exponential factors $\Pi_j(k)$, with $j = 1, \ldots, q$. When there is only one pole with amount equal P, which implies $q = 1$ in (3.97), then this pole is according to Corollary 1 positive and real and therefore equal P. The equation (3.98) is then of the simple form

$$N[k] = \theta(k)P^{-k} + \mathrm{O}(\alpha_2^{-k}), \qquad \theta(k) = \Pi_1(k). \tag{3.101}$$

All examples considered in this work will have an approximation of this form. For the exponential order we repeat the fundamental observation stated in Corollary 1.

Note 7. *(First principle of coefficient asymptotics, see [11, p. 215]).* The *location* of the dominant poles of a rational function dictates the exponential growth P^{-k} of its coefficients.

The error term in (3.98) is exponentially smaller than the dominant term for $k \to \infty$. In most cases, when characterizing the asymptotic behavior of $N[k]$, we do not need to take into account the poles α_j with $j = q+1, \ldots, m$ explicitly. However, these poles have an influence on the sub-exponential factors $\Pi_j(k)$, $j = 1, \ldots, q$ of the dominant exponential terms in (3.98). For generating functions with only one dominant pole we state this observation in the following note.

Note 8. *(Second principle of coefficient asymptotics, see [11, p. 215]).* The *nature* of the poles of a rational function determines the sub-exponential factor $\theta(k)$.

We can use the knowledge of the sub-exponential factor $\theta(k)$ to compare two DNCs with coefficients that have the same exponential order.

Example 10. (*Comparison of two DNCs*). Let $\mathcal{A} = (A, w)$ represent a DNC that allows binary strings that do not contain the substring 11 and let $\mathcal{B} = (B, w)$ represent a DNC that allows binary strings not containing the substring 011. Assume $w(0) = w(1) = 1$. The sets A and B are then given by the regular expressions

$$A = \{\varepsilon \cup 1\}\{0 \cup 01\}^\star \qquad \text{and} \qquad B = \{1\}^\star\{0 \cup 01\}^\star. \tag{3.102}$$

For the generating function of $\mathcal{A}$ we get

$$\mathrm{H}_{\mathcal{A}}(z) = \sum_{k=0}^{\infty} N_{\mathcal{A}}[k] z^k \tag{3.103}$$

$$= (1+z) \sum_{k=0}^{\infty} (z+z^2)^k \tag{3.104}$$

$$= \frac{1+z}{1-z-z^2} \tag{3.105}$$

and for the generating function of $\mathcal{B}$ we get

$$\mathrm{H}_{\mathcal{B}}(z) = \sum_{k=0}^{\infty} N_{\mathcal{B}}[k] z^k \tag{3.106}$$

$$\left(\sum_{l=0}^{\infty} z^l\right)\left(\sum_{m=0}^{\infty} (z+z^2)^m\right) \tag{3.107}$$

$$= \frac{1}{(1-z)} \frac{1}{(1-z-z^2)}. \tag{3.108}$$

As we can see, $\mathrm{H}_{\mathcal{B}}(z)$ has in addition to the poles of $\mathrm{H}_{\mathcal{A}}(z)$ a pole in $z = 1$. Does the DNC $\mathcal{B}$ allow more strings of the same weight than the DNC $\mathcal{A}$? We find the poles

$$z_1 = 0.61803 \qquad z_2 = 1 \qquad z_3 = -1.6180. \tag{3.109}$$

According to (3.82) we can write

$$\mathrm{H}_{\mathcal{A}}(z) = \frac{a_1}{z - z_1} + \frac{a_2}{z - z_3} \tag{3.110}$$

$$\mathrm{H}_{\mathcal{B}}(z) = \frac{b_1}{z - z_1} + \frac{b_2}{z - z_2} + \frac{b_3}{z - z_3}. \tag{3.111}$$

We apply (3.85) and get for the coefficients

$$N_{\mathcal{A}}[k] = -a_1 z_1^{-k-1} - a_2 z_3^{-k-1} \tag{3.112}$$

$$N_{\mathcal{B}}[k] = -b_1 z_1^{-k-1} - b_2 z_2^{-k-1} - b_3 z_3^{-k-1}. \tag{3.113}$$

We see that the exponential order of $N_{\mathcal{A}}[k]$ and $N_{\mathcal{B}}[k]$ is determined by the smallest positive pole, which is given by $z_1 = 0.61803$. This is in accordance with Corollary 1. We only take into account the dominant pole z_1 and write for the coefficients

$$N_{\mathcal{A}}[k] = -a_1 z_1^{-k-1} + \mathrm{O}\left(z_3^{-k}\right) \tag{3.114}$$

$$N_{\mathcal{B}}[k] = -b_1 z_1^{-k-1} + \mathrm{O}\left(z_2^{-k}\right). \tag{3.115}$$

To calculate the sub-exponential factors a_1 and b_1 we use Lemma 7. For a_1 we have

$$a_1 = \left.\frac{1+z}{\frac{\mathrm{d}}{\mathrm{d}z}(1 - z - z^2)}\right|_{z=z_1} \tag{3.116}$$

$$= \frac{1+z_1}{-1-2z_1} \tag{3.117}$$

$$= -0.72361. \tag{3.118}$$

In the same way, we get for b_1

$$b_1 = \left.\frac{1}{\frac{\mathrm{d}}{\mathrm{d}z}(1-z)(1 - z - z^2)}\right|_{z=z_1} \tag{3.119}$$

$$= \frac{1}{-2+3z_1^2} \tag{3.120}$$

$$= -1.1708. \tag{3.121}$$

The sub-exponential factor b_1 is larger than a_1 by a factor of

$$\frac{b_1}{a_1} = \frac{-1.1708}{-0.72361} \tag{3.122}$$

$$= 1.6180. \tag{3.123}$$

We conclude that the asymptotic number of strings accepted by the DNC $\mathcal{B}$ is larger than the number of strings accepted by the DNC $\mathcal{A}$ by a factor of ≈ 1.6. ◁

The results we derived for DNCs that have integer valued symbol weights and a rational generating function can be slightly generalized by dropping the second condition. We know from (3.98) that the error term of our approximation of $N[k]$ is determined by the pole that lies closest to the origin among all poles that we did not explicitly include in our approximation. If an error term $\mathrm{O}(R^{-k})$ is acceptable, we only consider the poles of the generating function that lie inside the disc $|z| < R$. The following theorem shows that this approach indeed leads to an approximation of the coefficients $N[k]$ with a quantifiable error term.

Theorem 4. [11, p. 255](Expansion of meromorphic functions). *Let $f(z)$ be a function meromorphic for $|z| \leq R$ with poles at points $\alpha_1, \alpha_2, \ldots, \alpha_m$ inside the disc $|z| \leq R$. Assume further that $f(z)$ is analytic at all points of $|z| = R$ and at $z = 0$. Then there exist m polynomials $\{\Pi_j(k)\}_{j=1}^m$ such that*

$$[z^k]f(z) = \sum_{j=1}^{m} \Pi_j(k)\alpha_j^{-k} + \mathrm{O}(R^{-k}). \tag{3.124}$$

Furthermore, the degree of Π_j is equal to $r_j - 1$ where r_j is equal to the multiplicity of the pole of $f(z)$ at α_j.

3.4. Shannon's Telegraphy Channel

We continue with the example from Section 2.4. Until now, we derived the generating series of the DNC $\mathcal{T}$, which represents Shannon's telegraphy channel. It is given by

$$\mathrm{GEN}_{\mathcal{T}}(x) = \sum_{k=1}^{\infty} N[k]x^k \tag{3.125}$$

$$= \frac{1 + x^3 + x^6}{1 - x^2 - x^4 - x^5 - x^7 - x^8 - x^{10}}. \tag{3.126}$$

In this section, we will investigate the asymptotic behavior of the number of distinct strings $N[k]$ of weight k accepted by the DNC $\mathcal{T}$. We will do this by investigating the rational generating function $\mathrm{H}_{\mathcal{T}}(z)$ that results from evaluating $\mathrm{GEN}_{\mathcal{T}}(x)$ in the complex plane. It is given by

$$\mathrm{H}_{\mathcal{T}}(z) = \mathrm{GEN}_{\mathcal{T}}(x)\big|_{x=z}, \qquad z \in \mathbb{C} \tag{3.127}$$

$$= \frac{1 + z^3 + z^6}{1 - z^2 - z^4 - z^5 - z^7 - z^8 - z^{10}}. \tag{3.128}$$

3.4.1. Exponential Order of Coefficients

We start by determining the exponential order of the coefficients $N[k]$. From Corollary 1, we know that the exponential order of $N[k]$ is given by

$$N[k] \bowtie P^{-k} \tag{3.129}$$

where P is equal to the smallest positive real pole of $\mathrm{H}_{\mathcal{T}}(z)$. We write

$$\mathrm{H}_{\mathcal{T}}(z) = \frac{N(z)}{D(z)}. \tag{3.130}$$

The pole P is thus equal to the smallest postitive real zero of $D(z)$. We apply a simple algorithm, called Newton's method [25] to find P. We get

$$P = 0.68\,827\,830\,840\,767\,. \tag{3.131}$$

The exponential order of the coefficients is thus given by

$$N[k] \bowtie 0.68\,827\,830\,840\,767^{-k}. \tag{3.132}$$

3.4.2. Approximation of Coefficients with Arbitrary Precision

In the following, we will show that we can predict the values $N[k]$ with arbitrary precision from the expansion of $\mathrm{H}_{\mathcal{T}}(z)$ around its poles. For sake of illustration, we try to predict $N[k]$ from its exponential order. We guess

$$N[k] \approx P^{-k} \tag{3.133}$$

where P is the smallest positive pole of the generating function $\mathrm{H}_{\mathcal{T}}(z)$ as derived above. For $k = 10$, we get from an algebraic expansion of (2.66)

$$N[10] = 17 \tag{3.134}$$

but our guess leads to

$$P^{-10} = 41.91\,400\,621 \tag{3.135}$$

which shows that this approach would lead to a very poor approximation of $N[k]$. As we learned in this chapter, better approximations can be found by expanding the function $\mathrm{H}_{\mathcal{T}}(z)$ around its poles and using the fact that $N[k] = [z^k]\mathrm{H}_{\mathcal{T}}(z)$ for $\mathrm{H}_{\mathcal{T}}(z)$ expanded around zero. We will do this in the following.

First Approximation

For a first approximation of the asymptotic behavior of $N[k]$, we take into account the dominant poles only. We know from (3.98) that we can in this case write

$$N[k] = \sum_{j=1}^{q} \Pi_j(k)\alpha_j^{-k} + \mathrm{O}(\alpha_{q+1}^{-k}) \tag{3.136}$$

where α_i, $i = 1, \ldots, q$, denote the dominant poles with $|\alpha_i| = P$, and where the factors $\Pi_j(k)$ are polynomials in k of degree $r_j - 1$ with r_j being the multiplicity of the pole α_j. The error term $\mathrm{O}(\alpha_{q+1}^{-k})$ is determined by the pole α_{q+1}. It is equal to the pole α

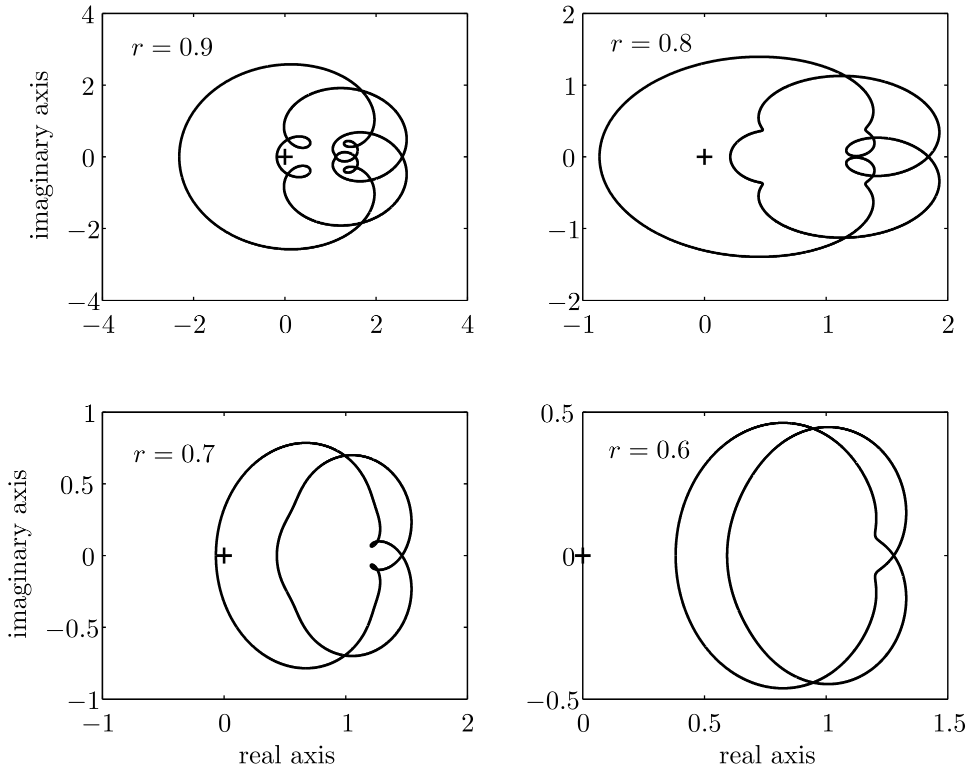

Figure 3.2.: We determine the number and multiplicity of zeros applying the argument principle. Plots for the denominator of the function $\mathrm{H}_{\mathcal{T}}(z)$ evaluated in $\gamma(t) = r\exp(it)$, $t = 0..2\pi$ for $r = 0.9$, $r = 0.8$, $r = 0.7$ and $r = 0.6$.

with $|\alpha| > P$ that lies closest to the origin. Before we can proceed, we need to determine the number q of dominant poles and their multiplicities. We do this by applying the argument principle on the denominator $D(z)$ of $\mathrm{H}_{\mathcal{T}}(z)$, see Theorem 12. We define the curve $\gamma(t)$ as

$$\gamma(t) = re^{it}, \qquad 0 \leq t < 2\pi. \tag{3.137}$$

Since $P \approx 0.69$, we choose for r the values $r = 0.6$ and $r = 0.7$. We see in Figure 3.2 that the winding number is equal one for $r = 0.7$ and that it is equal zero for $r = 0.6$. Since the zeros get counted with their multiplicities, we conclude that $\mathrm{H}_{\mathcal{T}}(z)$ has one dominant pole and that its multiplicity is equal one. We can now use the Laurent series expansion of $\mathrm{H}_{\mathcal{T}}(z)$ around $\alpha_1 = P$ to derive an approximation of the coefficients $N[k]$ of the form (3.136). The main part of the Laurent series expansion of $\mathrm{H}_{\mathcal{T}}(z)$ around α_1 is given by

$$\frac{A}{z - \alpha_1} \tag{3.138}$$

where the coefficient A is, according to Lemma 7, given by

$$\begin{aligned}
A &= \frac{N(\alpha_1)}{D'(\alpha_1)} &(3.139)\\
&= \frac{1 + \alpha_1^3 + \alpha_1^6}{-2\alpha_1 - 4\alpha_1^3 - 5\alpha_1^4 - 7\alpha_1^6 - 8\alpha_1^7 - 10\alpha_1^9} &(3.140)\\
&\approx -0.26\,142\,528\,220\,985\,. &(3.141)
\end{aligned}$$

We define (3.138) as the first approximation $\mathcal{T}_1(z)$ of $\mathrm{H}_{\mathcal{T}}(z)$. From (3.84), we get for the coefficients $N[k]$ the approximation

$$\begin{aligned}
N[k] &= [z^k]\mathrm{H}_{\mathcal{T}}(z) &(3.142)\\
&\approx [z^k]\mathcal{T}_1(z) &(3.143)\\
&= -A\alpha_1^{-k-1} &(3.144)\\
&= 0.261\,425\,28\,220\,985 \cdot 0.68\,827\,830\,840\,767^{-k-1}. &(3.145)
\end{aligned}$$

As we can see in Figure 3.2, the pole α_2 next closest to the origin is of multiplicity one and its amount lies in the interval $0.9 > |\alpha_2| > 0.8$. We thus can quantify the error of $\mathcal{T}_1(z)$ as

$$\mathrm{H}_{\mathcal{T}}(z) = \mathcal{T}_1(z) + \mathrm{O}(\alpha_2^{-k}). \tag{3.146}$$

Since $0.9 < 1$, we expect an absolute error that grows exponentially with k. To compare the approximation with the correct values, we calculate the first coefficients $N[k]$ of

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$N_1[k]$	0	1	1	1	2	2	4	5	8	11	16	23	34	49	71	103
$N[k]$	1	0	1	1	2	2	4	5	8	10	17	22	35	47	73	101

k	30	31	32	33	34	35	50
$N_1[k]$	27 968	40 635	59 038	85 777	124 625	181 068	49 133 689
$N[k]$	27 986	40 614	59 062	85 749	124 657	181 030	49 134 036

Table 3.1.: Comparison of the coefficient approximation with the correct values.

(2.66) algebraically. We have

$$N[k] = [z^k]\mathrm{H}_{\mathcal{T}}(z) \tag{3.147}$$

$$= [z^k](1+z^3+z^6)\sum_{k=0}^{\infty}\left(z^2+z^4+z^5+z^7+z^8+z^{10}\right)^k \tag{3.148}$$

$$= [z^k](1+z^3+z^6)\sum_{k=0}^{\lfloor \frac{k}{2} \rfloor}\left(z^2+z^4+z^5+z^7+z^8+z^{10}\right)^k . \tag{3.149}$$

The last line is a finite sum, which can easily be calculated. We approximate $N[k]$ by

$$N_1[k] = \mathrm{round}\left(-A\alpha_1^{-k-1}\right) \tag{3.150}$$

where $y = \mathrm{round}(x)$ assigns to y the integer that lies closest to x. For A and α_1, we use the values from (3.145). In Table 3.1, we list the approximated coefficients $N_1[k]$ and the correct coefficients $N[k]$ for $k = 0, ..., 15$, $k = 30, ..., 35$ and $k = 50$. As we predicted, we have an absolute error that increases with k.

Second Approximation

The approximation $N_1[k]$ is a bit smaller than the correct value when k is even and a bit larger when k is odd. We therefore expect the pole α_2, which lies next closest to the origin, to be negative. According to Theorem 13, complex roots of polynomials with real coefficients only appear in complex conjugated pairs. As we can see in Figure 3.2, there is only one pole with its amount lying in the interval $(0.8, 0.9)$. Therefore, the pole α_2 is real, and we can look for it along the real axis. We find

$$\alpha_2 = -0.86\,274\,326\,605\,638 . \tag{3.151}$$

The main part of the Laurent series expansion of $\mathrm{H}_{\mathcal{T}}(z)$ around α_2 is equal

$$\frac{B}{z-\alpha_2} \tag{3.152}$$

where B is according to Lemma 7 given by

$$B = \frac{N(\alpha_2)}{D'(\alpha_2)} \tag{3.153}$$

$$= 0.18\,641\,508\,614\,227 . \tag{3.154}$$

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$N_2[k]$	1	0	1	1	2	2	4	5	8	10	17	22	35	47	73	101
$N[k]$	1	0	1	1	2	2	4	5	8	10	17	22	35	47	73	101

k	30	31	32	33	34	35	50
$N_2[k]$	27 986	40 614	59 063	85 748	124 658	181 030	49 134 036
$N[k]$	27 986	40 614	59 062	85 749	124 657	181 030	49 134 036

Table 3.2.: Comparison of the new coefficient approximation with the correct values.

We define $\mathcal{T}_2(z)$ as a second approximation of $\mathrm{H}_{\mathcal{T}}(z)$ by

$$\mathcal{T}_2(z) = \frac{A}{z-\alpha_1} + \frac{B}{z-\alpha_2}. \tag{3.155}$$

From (3.84), we get for the coefficients $N[k]$ the new approximation

$$N[k] \approx [z^k]\mathcal{T}_2(z) \tag{3.156}$$

$$- A\alpha_1^{-k-1} - B\alpha_2^{-k-1}. \tag{3.157}$$

We now can approximate $N[k]$ by

$$N_2[k] = \mathrm{round}\Big(-AR^{-k-1} - BP^{-k-1}\Big). \tag{3.158}$$

In Table 3.2, we display the first values we find for the new approximation $N_2[k]$. Only for $k = 33$, $k = 34$ and $k = 35$ the values of $N_2[k]$ differ by 1 from the correct values. For the other k, the new approximation is correct. We note that the value for $k = 50$ is correct. To check if we can expect $N_2[k]$ to be still correct for large k we again apply the argument principle. In Figure 3.3, we display the contour of $\mathcal{T}(\gamma)$ for $r = 1$ and $r = 0.98$. In both cases, the winding number is 4. We remember that the winding number was 2 for $r = 0.9$. This means that $\mathrm{H}_{\mathcal{T}}(z)$ has two poles with multiplicity 1 or one pole with multiplicity 2 with the modulus lying in the interval $(0.9, 0.98)$. As a consequence, $N_2[k]$ still has an error, which will grow exponentially with k because of $0.98 < 1$. Evaluating $N_2[k]$ and $N[k]$ in $k = 75$ yields

$$C_{75}^{(2)} = 558\,826\,885\,811 \tag{3.159}$$

$$C_{75} = 558\,826\,885\,813. \tag{3.160}$$

As we predicted, there is still an error for large k.

Efficient Algorithm to Calculate $N[k]$

We now use MAPLE to find all poles with modulus smaller than 1, derive the main part of the Laurent series expansion of $\mathrm{H}_{\mathcal{T}}(z)$ around them and add each main part to our approximation. MAPLE finds all 10 complex poles of $\mathrm{H}_{\mathcal{T}}(z)$. We display them in Figure 3.4. We define a new approximation $\mathcal{T}_4(z)$ by adding the contribution of the two

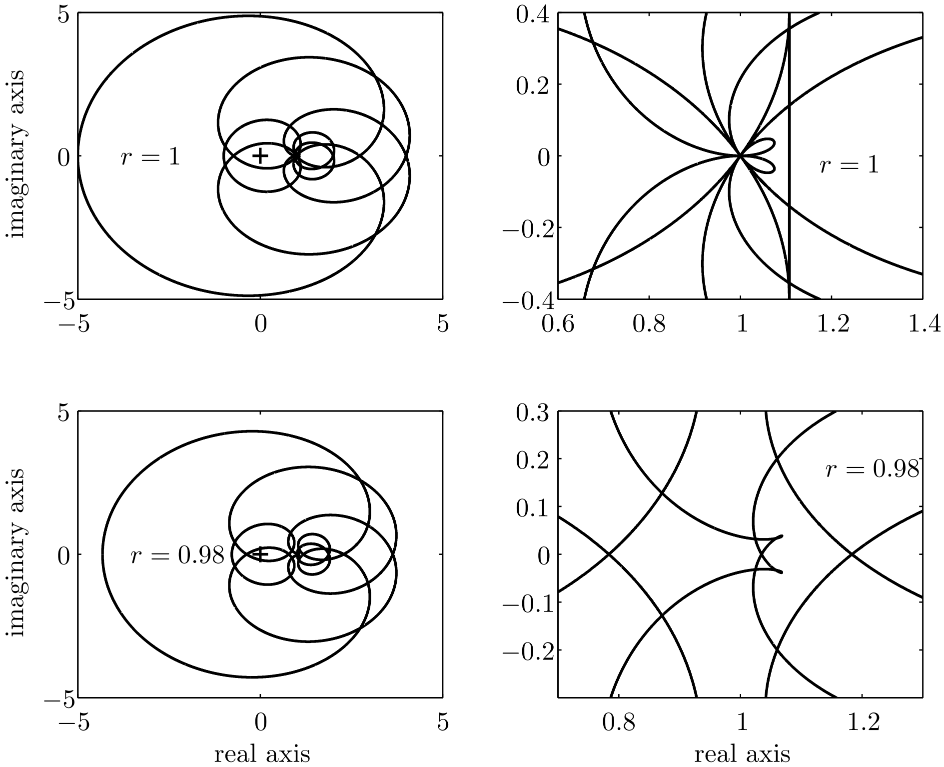

Figure 3.3.: Plots for the denominator of the generating function evaluated in $\gamma(t) = r\exp(it)$, $t = 0\dots 2\pi$ for $r = 1$ and $r = 0.98$.

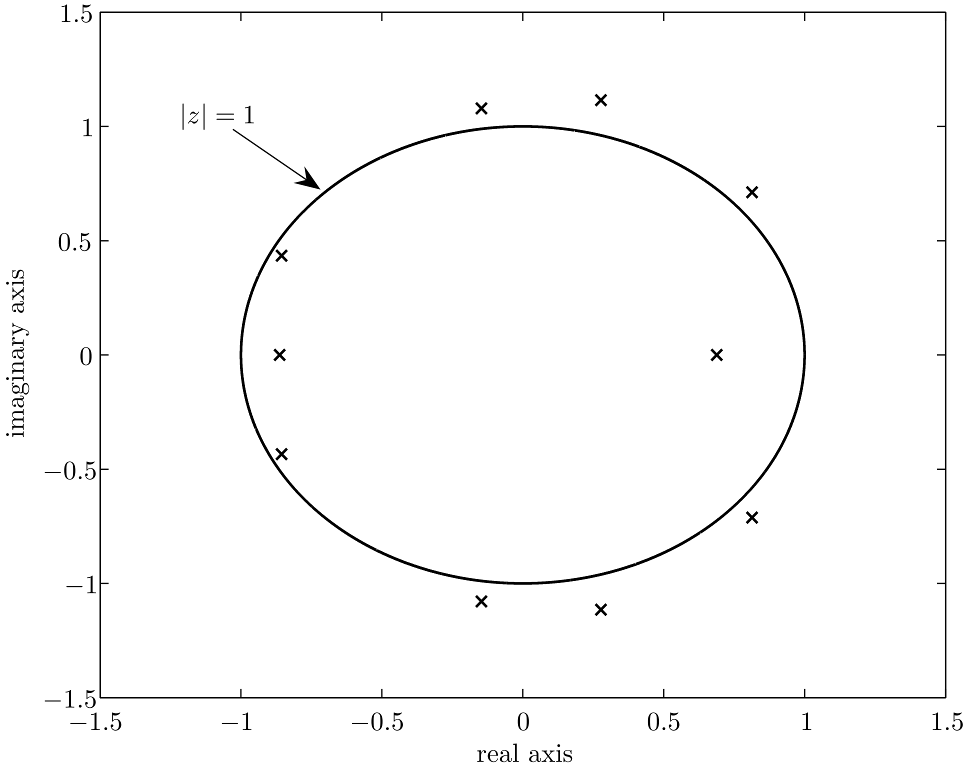

Figure 3.4.: All 10 complex roots of $\mathrm{H}_{\mathcal{T}}(z)$. Note hat there are only four roots with their modulus being smaller than one.

complex conjugated poles with modulus smaller than 1 to our last approximation $\mathcal{T}_2(z)$. We then have

$$N[k] = [z^k]\mathcal{T}_4(z) + \mathrm{O}(\alpha_5^{-k}), \qquad |\alpha_5| > 1 \tag{3.161}$$

which leads to an approximation with an absolute error vanishing for $k \to \infty$. Thus, for some k_0, the error term will be less than 1/2 for all $k > k_0$. We therefore approximate $N[k]$ by

$$N_4[k] = \mathrm{round}\Big([z^k]\mathcal{T}_4(z)\Big)\,. \tag{3.162}$$

This provides an analytic algorithm to calculate the correct values of $N[k]$ for $k > k_0$. Neglecting possible optimizations, the complexity of this algorithm is linear in k since we have to perform k multiplications, whereas the complexity of the algebraic algorithm given in (3.149) is exponential in k (to calculate the last term of the sum we have to perform $6^{k/2}$ multiplications). We determine the correct value of $N[500]$. In MAPLE, it takes 175.655 to determine $N[500]$, whereas it only takes 1.812 seconds to calculate $N[500]$ by the analytic algorithm. Both algorithms lead to the same value for $N[500]$.

3.5. Combination of Two Pattern Codes

An important problem in coding theory is the encoding of arbitrary integer sequences. A possible way to solve this problem is the application of pattern codes as we introduced in Section 2.3.3. These codes can be specified by a pattern p, which only appears in the rightmost bits of the codewords. Among all pattern codes, codes with a pattern of the form $p = 011\cdots1$ are optimal, see [20]. Here, we will only consider the code F_{011} with the pattern 011 and the code F_{0111} with the pattern 0111. The codes could be represented by the two DNCs (F_{011}, w) and (F_{0111}, w), with the weight function given by $w(0) = w(1) = 1$. We can write the set F_{011} as

$$F_{011} = \{1\}^\star\{0, 01\}^\star\{011\} \tag{3.163}$$

which yields for the corresponding generating function

$$\mathrm{G}_{011}\,(y) = \left(\sum_{k=0}^{\infty} y^k\right)\left(\sum_{l=0}^{\infty}(y + y^2)^l\right) y^3 \tag{3.164}$$

$$= \frac{y^3}{(1-y)[1-(y+y^2)]}. \tag{3.165}$$

In the same way, we have for the set F_{0111}

$$F_{0111} = \{1\}^\star\{0, 01, 011\}^\star\{0111\} \tag{3.166}$$

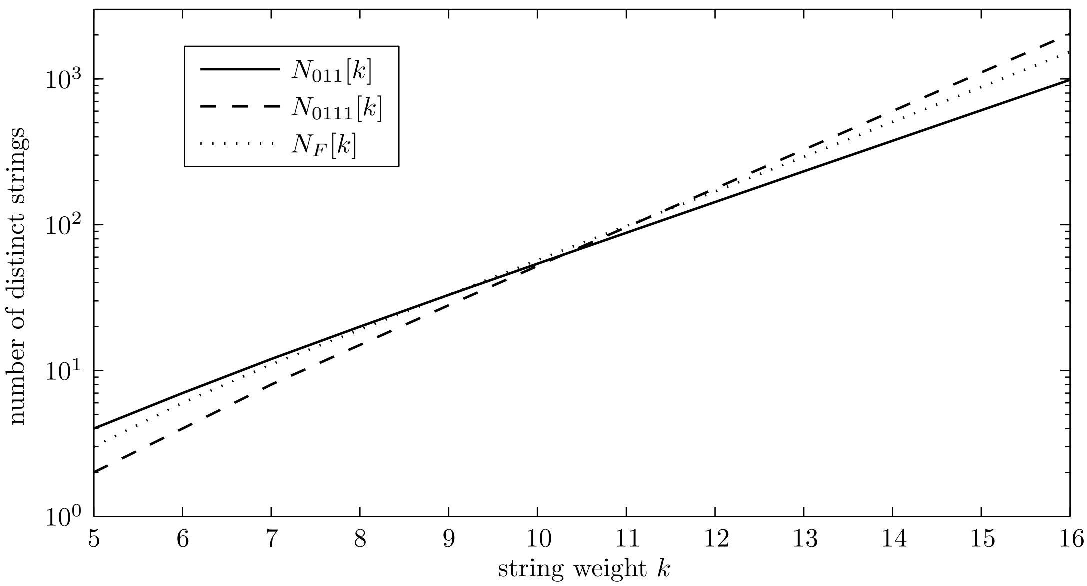

Figure 3.5.: The coefficients $N_{011}[k]$, $N_{0111}[k]$, and $N_F[k]$.

and for the corresponding generating function, we get

$$\mathrm{G}_{0111}(y) = \left(\sum_{k=0}^{\infty} y^k\right)\left(\sum_{l=0}^{\infty}(y+y^2+y^3)^l\right)y^4 \tag{3.167}$$

$$= \frac{y^4}{(1-y)[1-(y+y^2+y^3)]}. \tag{3.168}$$

We use Theorem 3 together with (3.85) to calculate the correct numbers of distinct codewords of length k for the two codes. For F_{011}, we denote this number by $N_{011}[k]$ and for F_{0111}, we denote it by $N_{0111}[k]$. We display the two number sequences in Figure 3.5. The number sequence $N_F[k]$ we will define later. As we can see, $N_{011}[k]$ is larger than $N_{0111}[k]$ for $k \leq 10$. However, $N_{0111}[k]$ is larger than $N_{011}[k]$ for $k > 10$, and it has a better asymptotic exponential behavior than $N_{011}[k]$. Based on an idea of Márcio Lima from the Federal University of Pernambuco in Recife, Brazil, we define a new codes, which combines the advantages of $N_{011}[k]$ for small k with the advantages of $N_{0111}[k]$ for large k. Let $F_{1,011} \subset F_{011}$ denote the set that contains all codewords from F_{011} that start with 1. It is given by

$$F_{1,011} = \{1\}\{1\}^\star\{0, 01\}^\star\{011\} \tag{3.169}$$

which yields for the corresponding generating function

$$\mathrm{G}_{1,011}(y) = y\mathrm{G}_{011}(y). \tag{3.170}$$

We denote by $F_{1,0111} \subset F_{0111}$ the set of all codewords from F_{0111} that start with 1. We have

$$F_{1,0111} = \{1\}\{1\}^{\star}\{0, 01, 011\}^{\star}\{0111\} \tag{3.171}$$

and for the corresponding generating function we get

$$\mathrm{G}_{1,0111}(y) = y\mathrm{G}_{0111}(y). \tag{3.172}$$

Let $F_{0,0111} \subset F_{0111}$ denote the set of all codeword from F_{0111} that start with 0. Since

$$F_{0,0111} \cup F_{1,0111} = F_{0111} \qquad \text{and} \qquad F_{0,0111} \cap F_{1,0111} = \emptyset \tag{3.173}$$

we have

$$F_{0,0111} = F_{0111} \setminus F_{1,0111}. \tag{3.174}$$

Because of $F_{0,0111} \cap F_{1,0111} = \emptyset$, this implies for the corresponding generating function

$$\mathrm{G}_{0,0111}(y) = \mathrm{G}_{0111}(y) - y\mathrm{G}_{0111}(y). \tag{3.175}$$

We now consider the set $F = F_{1,011} \cup F_{0,0111}$. Since $F_{1,011} \cap F_{0,0111} = \emptyset$, we have for the generating function

$$\mathrm{G}_F(y) = \mathrm{G}_{1,011}(y) + \mathrm{G}_{0,0111}(y) \tag{3.176}$$

$$= y\mathrm{G}_{011}(y) + \mathrm{G}_{0111}(y) - y\mathrm{G}_{0111}(y). \tag{3.177}$$

The multiplication of a generating function by y corresponds to a shift of the corresponding coefficients by -1, e.g., if $N[k]$ denotes the coefficients of $\mathrm{G}(y)$, then the coefficients $M[k]$ of $y\mathrm{G}(y)$ are given by $M[k] = N[k-1]$. The coefficients $N_F[k]$ of $\mathrm{G}_F(y)$ are therefore given by

$$N_F[k] = N_{011}[k-1] + N_{0111}[k] - N_{0111}[k-1]. \tag{3.178}$$

We display $N_F[k]$ together with $N_{011}[k]$ and $N_{0111}[k]$ in Figure 3.5. As we can see, the code F combines the advantage of the code F_{011} for small k with the advantage of the code F_{0111} for large k. In particular, $N_F[k]$ has the same exponential behavior as $N_{0111}[k]$.

3.6. Sub-Exponential Behavior for Non-Integer String Weights

In this section, we will generalize the results from Section 3.3 to DNCs with string weights taking non-integer values. Here, we have to distinguish between two cases. In the first case, the string weights take non-integer values but are commensurable and in the second case, the string weights are incommensurable. Whereas the first case can be treated in almost the same way as the case of integer valued string weights, we have to use a slightly different approach for the second case.

3.6.1. Sub-Exponential Behavior for Commensurable String Weights

We consider a DNC $\mathcal{A} = (A, w)$ accepting strings that take non-integer valued weights. We assume that the string weights are commensurable, which means that they can be written as an integer multiple of the same unit. We denote this unit by a. By scaling the weight function w of $\mathcal{A}$, we can make $\mathcal{A}$ become a DNC with integer valued string weights only. We do this by defining the DNC $\mathcal{B} = (A, v)$ with its weight function v given by

$$v(s) = \frac{w(s)}{a}, \qquad \forall s \in A. \tag{3.179}$$

The strings accepted by the DNC $\mathcal{B}$ are all of integer valued weights and we can apply the results from Section 3.3 to derive approximations with sub-exponential precision for the coefficients $N_{\mathcal{B}}[k]$ of the generating series of $\mathcal{B}$. The coefficients $N_{\mathcal{A}}[w_l]$ of the generating series of the DNC $\mathcal{A}$ are related to the coefficients $N_{\mathcal{B}}[k]$ in the following way

$$N_{\mathcal{A}}[w_l] = N_{\mathcal{B}}\left[\frac{w_l}{a}\right]. \tag{3.180}$$

For example, if the strings allowed by the DNC $\mathcal{A}$ are generated over the alphabet $\{0, 1\}$ and if the weights of these two symbols are given by $w(0) = 3\pi/2$ and $w(1) = 2\pi$, then we use as a measure $a = \pi/2$ and we write $w(0) = 3a$ and $w(1) = 4a$. The weight function of the DNC $\mathcal{B}$ is then determined by $v(0) = 3$ and $v(1) = 4$, which implies that all strings accepted by $\mathcal{B}$ will take integer valued weights only.

3.6.2. Sub-Exponential Behavior for Incommensurable String Weights

We finally consider the general case of DNCs with accepted strings possibly taking incommensurable weights. As in the case of integer valued string weights and in the case of commensurable string weights, we are interested in informations about the sub-exponential behavior of the coefficients of the generating series. Since the string weights are incommensurable, the DNCs we will consider here cannot be transformed into DNCs with integer valued string weights. We will proceed in the following way. First, we will give an alternative approach how to treat the case of integer valued string weights. Second, we will show by an example that this approach also allows to derive useful informations for the case of incommensurable string weights.

Alternative Expansions in the Case of Integer Valued String Weights

For a DNC $\mathcal{A}$ with integer valued string weights, we used a complex generating function of the form

$$\mathrm{H}_{\mathcal{A}}(x) = \sum_{k=0}^{\infty} N[k] z^k, \qquad z \in \mathbb{C}. \tag{3.181}$$

Although not necessary, we could also use the complex generating function we had to introduce for non-integer symbol weights, which is given by

$$\mathrm{G}_{\mathcal{A}}\left(e^{s}\right)=\sum_{k=0}^{\infty} N[k] e^{k s}, \qquad s \in \mathbb{C}. \tag{3.182}$$

Since $\mathrm{G}_{\mathcal{A}}\left(e^{s}\right)$ can be derived from $\mathrm{H}_{\mathcal{A}}(z)$ by the substitution $z=e^{s}$, we can transform by substitution the identities we found in Section 3.3 for $\mathrm{H}_{\mathcal{A}}(z)$ into identities for $\mathrm{G}_{\mathcal{A}}\left(e^{s}\right)$. We summarize this in the following lemma. In parenthesis, we give the corresponding concept for $\mathrm{H}_{\mathcal{A}}(z)$. For a better understanding of the substitution, we refer to Figure 3.1.

Lemma 8. *For a DNC $\mathcal{A}$ with integer-valued symbol weights, the following holds.*

i. (Partial fraction expansion) *If the generating function $\mathrm{H}_{\mathcal{A}}(z)$ is rational, then the generating function $\mathrm{G}_{\mathcal{A}}\left(e^{s}\right)$ has an expansion of the form*

$$\mathrm{G}_{\mathcal{A}}\left(e^{s}\right)=Q(e^{s})+\sum_{\alpha} \sum_{r=1}^{r_{\alpha}} \frac{c_{\alpha, r}}{\left(e^{s}-e^{\alpha}\right)^{r}}. \tag{3.183}$$

ii. (Newton expansion) *For the summands of the inner sum of* (3.183) *we have*

$$[e^{k s}] \frac{c_{\alpha, r}}{\left(e^{s}-e^{\alpha}\right)^{r}}=c_{\alpha, r}(-1)^{r}\begin{pmatrix} k+r-1 \\ r-1 \end{pmatrix} e^{-(k+r) \alpha}. \tag{3.184}$$

iii. (Expansion of meromorphic functions) *Assume that $\mathrm{G}_{\mathcal{A}}\left(e^{s}\right)$ is meromorphic for $\Re\{s\} \leq R$ and assume that $\mathrm{G}_{\mathcal{A}}\left(e^{s}\right)$ is analytic in all points $\Re\{s\}=R$. Let $\alpha_{1}, \ldots, \alpha_{m}$ denote the poles of $\mathrm{G}_{\mathcal{A}}\left(e^{s}\right)$ in $\Re\{s\}<R$ and $-\pi<\Im\{s\} \leq \pi$. Then*

$$[e^{k s}] \mathrm{G}_{\mathcal{A}}\left(e^{s}\right)=\sum_{j=1}^{m} \Pi_{j}(k) e^{-k \alpha_{j}}+\mathrm{O}(e^{-k R}) \tag{3.185}$$

where the factors $\Pi_{j}(k)$ are polynomials in k. The degree of the polynomial $\Pi_{j}(k)$ is equal to $r_{j}-1$ where r_{j} is equal to the multiplicity of the pole α_{j}.

iv. (Coefficients of Laurent series expansion) *If a pole α is of multiplicity $r=1$, then the corresponding coefficient $c_{\alpha, 1}$ in* (3.183) *is given by*

$$c_{\alpha, 1}=\lim_{s \rightarrow \alpha}\left(e^{s}-e^{\alpha}\right) \mathrm{G}_{\mathcal{A}}\left(e^{s}\right). \tag{3.186}$$

Proof. The lemma follows from substituting z by $\exp(s)$ in the corresponding identities for the generating function $\mathrm{H}_{\mathcal{A}}(z)$. □

Sub-Exponential Behavior of DNC with Incommensurable String Weights

We now use the following approach. The calculations which lead to the identities in Lemma 8 are also well-defined for generating functions with incommensurable string weights. Without investigating if equality holds in (3.183), we use in the following example the identities given in Lemma 8 to investigate the sub-exponential behavior of the coefficients of the complex generating function of a DNC with incommensurable string weights. We define an approximation of the coefficients and compare it with the correct values obtained from an algebraic expansion of the corresponding generating series.

Example 11. (*DNC with incommensurable symbol weights*). We consider the DNC $\mathcal{A} = (A^\star, w)$ with

$$A = \{a, b, c\} \tag{3.187}$$

and

$$w(a) = 1 \qquad w(b) = \sqrt{2} \qquad w(c) = \pi. \tag{3.188}$$

The generating series of $\mathcal{A}$ is given by

$$\begin{aligned} \mathrm{GEN}_{\mathcal{A}}(x) &= \sum_{k=1}^{\infty} N[w_k] x^{w_k} && (3.189)\\ &= \sum_{l=0}^{\infty} (x + x^{\sqrt{2}} + x^{\pi})^l && (3.190)\\ &= \frac{1}{1-(x + x^{\sqrt{2}} + x^{\pi})}. && (3.191) \end{aligned}$$

The weights of the symbols from A are pairwise incommensurable. This means that we cannot make the symbol weights become integer-valued by an appropriate scaling. If the symbol weights were rational numbers, this would be possible. Our goal is to make a statement about the sub-exponential asymptotic behavior of $N[w_k]$ anyhow. We will do this by applying the techniques indicated in Lemma 8. Substituting x in (3.191) by e^s, $s \in \mathbb{C}$, we get for the complex generating function of $\mathcal{A}$

$$\mathrm{G}_{\mathcal{A}}(e^s) = \frac{1}{1-(e^s + e^{\sqrt{2}s} + e^{\pi s})}. \tag{3.192}$$

For the poles of $\mathrm{G}_{\mathcal{A}}(e^s)$ in $-\pi < \Im\{s\} \leq \pi$, we find

$$p_1 = -0.68493 \qquad p_2 \approx 0.42 + 1.76i \qquad p_3 \approx 0.42 - 1.76i. \tag{3.193}$$

Since p_1 is the leftmost real singularity of $\mathrm{G}_{\mathcal{A}}(e^s)$, the exponential order of $N[w_k]$ is according to Theorem 2 given by

$$\begin{aligned} N[w_k] &\bowtie e^{-w_k p_1} && (3.194)\\ &= 0.50412^{-w_k}. && (3.195) \end{aligned}$$

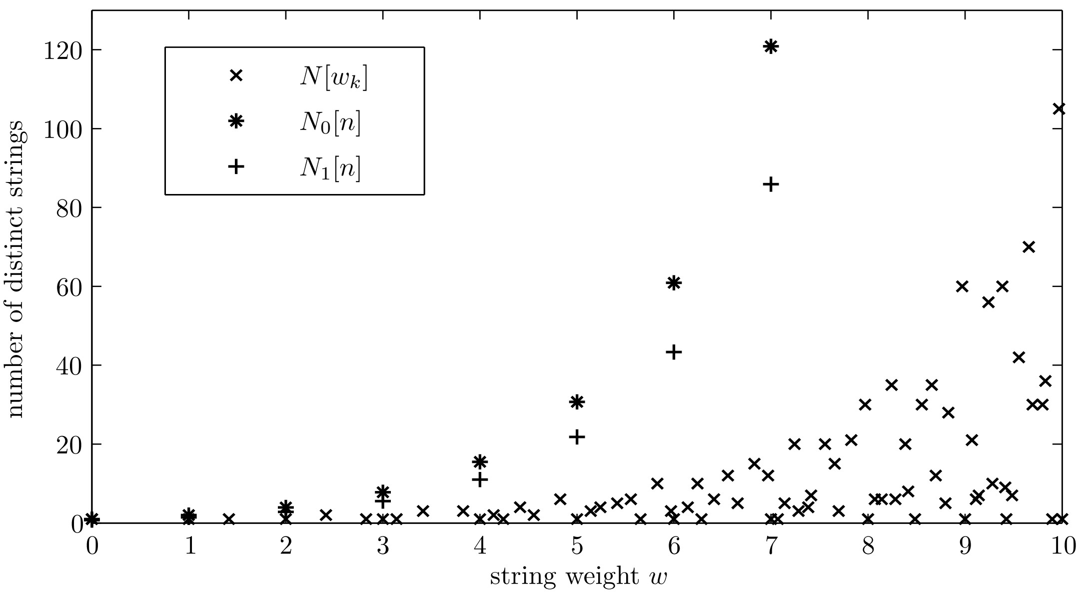

Figure 3.6.: Plot of the coefficients $N_0[n]$, $N_1[n]$ and $N[w_k]$.

Because the symbols are incommensurable, the set of possible string weights $\{w_k\}_{k=1}^{\infty}$ consists of values spread over the nonnegative real axis, see the plot of $N[w_k]$ in Figure 3.6 for an illustration. We will later define the coefficients $N_0[n]$ and $N_1[n]$, which are also displayed in the same figure. The values of $N[w_k]$ were obtained by an algebraic expansion of $\mathrm{GEN}_{\mathcal{A}}(x)$. In particular, for $w_k \in \mathbb{N}$, the number of strings allowed by $\mathcal{A}$ will always be $N[w_k] = 1$ since only concatenations of the symbol a result in strings of integer valued weight. Therefore, the number of strings of a specific weight w_k given by $N[w_k]$ does not bare much information about the performance of the considered DNC. In addition, the determination of the set $\{w_k\}_{k=1}^{\infty}$ is of exponential complexity. We solve this problem by approximating $\mathrm{G}_A(e^s)$ by a partial fraction expansion as indicated in (3.183) in Lemma 8. Note that we only showed equality in (3.183) for integer-valued symbol weights. We denote the approximation by $\mathrm{F}_{\mathcal{A}}(e^s)$. It is given by

$$\mathrm{F}_{\mathcal{A}}(e^s) = \frac{a_1}{e^s - e^{p_1}} + \frac{a_2}{e^s - e^{p_2}} + \frac{a_3}{e^s - e^{p_3}}. \tag{3.196}$$

We use (3.184) in (3.183) to expand $\mathrm{F}_{\mathcal{A}}(z)$. For the coefficients we get

$$[z^{ks}]\mathrm{F}_{\mathcal{A}}(z) = -a_1 e^{-(k+1)p_1} - a_2 e^{-(k+1)p_2} - a_3 e^{-(k+1)p_3}. \tag{3.197}$$

We define two approximations for the coefficients. For the first approximation we simply use the exponential behavior of the coefficients $[z^{ks}]\mathrm{F}_{\mathcal{A}}(z)$ and define

$$N_0[k] = e^{-kp_1}. \tag{3.198}$$

The second approximation we base on (3.197). Since $\Re\{p_2\} > 0$ and $\Re\{p_3\} > 0$, the contribution of these two poles will vanish exponentially fast for $k \to \infty$. We therefore

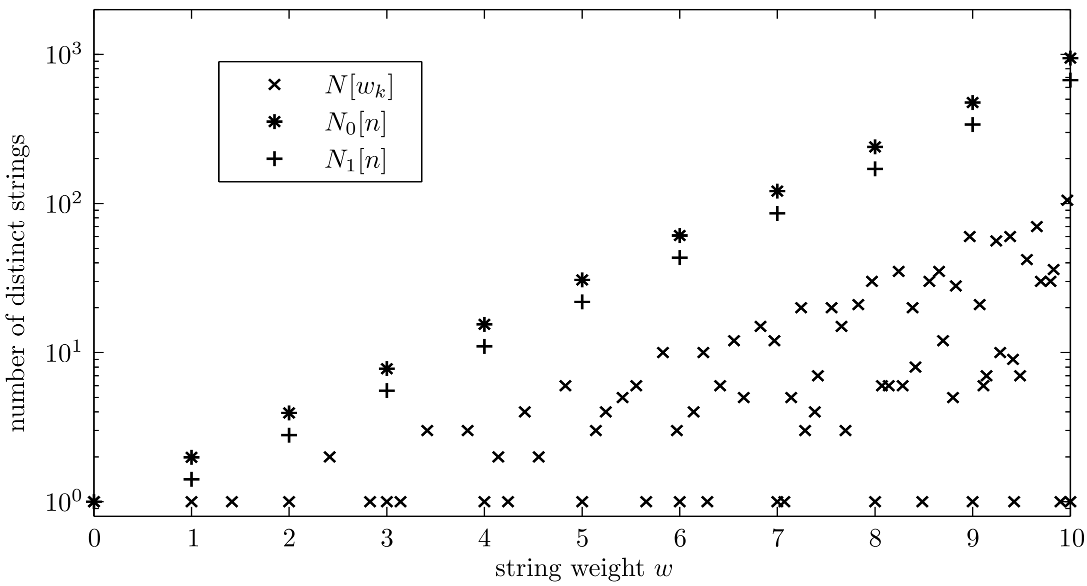

Figure 3.7.: Logarithmic plot of the coefficients $N_0[n]$, $N_1[n]$ and $N[w_k]$.

only take into account the contribution of p_1 and define the second approximation of the coefficients as

$$N_1[k] = -a_1 e^{-(k+1)p_1}. \tag{3.199}$$

For the sub-exponential factor a_1 we have from (3.186)

$$a_1 = \lim_{s\to p_1} (e^s - e^{p_1}) \frac{1}{1-(e^s + e^{\sqrt{2}s} + e^{\pi s})} \tag{3.200}$$

$$= \lim_{s\to p_1} \frac{e^s}{-(e^s + \sqrt{2}e^{\sqrt{2}s} + \pi e^{\pi z})} \tag{3.201}$$

$$= -0.35849. \tag{3.202}$$

How are the approximations $N_0[n]$ and $N_1[n]$ of the coefficients of $\mathrm{H}_{\mathcal{A}}(z)$ related to the coefficients $N[w_k]$ of $\mathrm{G}_{\mathcal{A}}(e^s)$? The plots of $N_0[n]$, $N_1[n]$ and $N[w_k]$ in Figure 3.6 look quite different. The approximations $N_0[n]$ and $N_1[n]$ take nonzero values only in $n \in \mathbb{N}$ but grow much faster than $N[w_k]$. It is not clear from Figure 3.6 which one of the approximations $N_0[n]$ and $N_1[n]$ describes $N[w_k]$ in a better way and it is also not clear if the exponential behavior of the approximations is the same as the exponential behavior of $N[w_k]$. We draw a logarithmic plot to investigate the latter question. The result is displayed in Figure 3.7. Although $N[w_k]$ is oscillating, the supremum $S[w]$ defined as

$$S[w] = \sup\left\{N[w_k]\middle|w_k \le w\right\} \tag{3.203}$$

is growing approximately with the same rate as $N_0[n]$ and $N_1[n]$. This gives a good illustration of what is meant by the exponential order, which is defined by a limit superior.

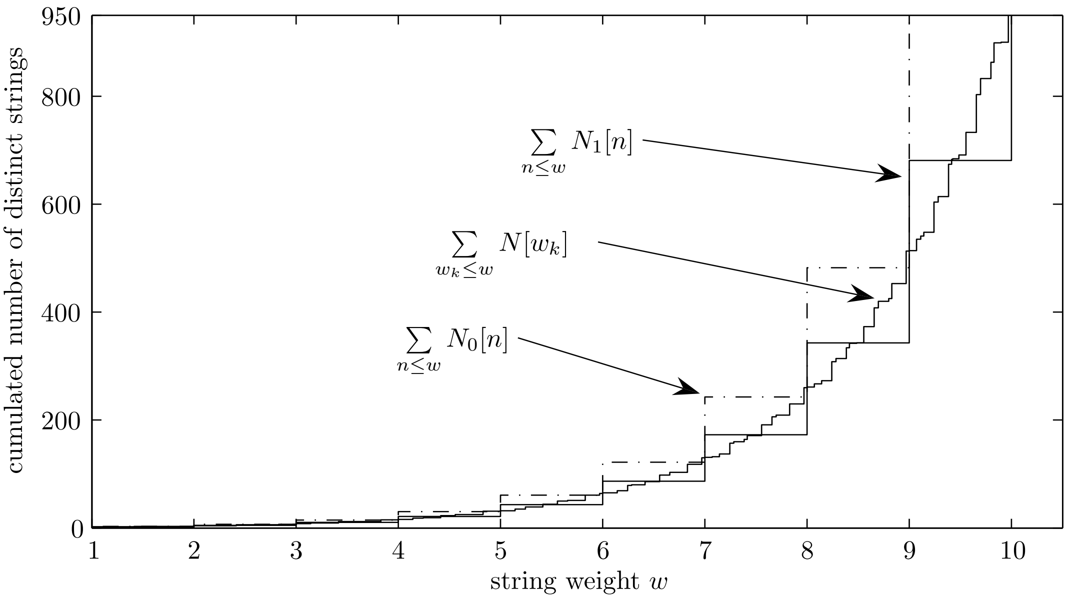

Figure 3.8.: Linear plot of the cumulated coefficients $N_0[n]$, $N_1[n]$ and $N[w_k]$.

The growths of the number sequences $N_0[n]$, $N_1[n]$ and $S[w]$ still differ by a factor. The explanation of this is the following: the coefficients $N[w_k]$ take nonzero values spread over the real axis whereas the coefficients $N_0[n]$ and $N_1[n]$ only take nonzero values in $n \in \mathbb{N}$. What is spread over the real axis in the case of $N[w_k]$ is cumulated on the integer numbers in the case of $N_0[n]$ and $N_1[n]$. We therefore consider the cumulated coefficients given by

$$\sum_{w_k \leq w} N[w_k], \qquad \sum_{n \leq w} N_0[n] \qquad \text{and} \qquad \sum_{n \leq w} N[n]. \tag{3.204}$$

The plot of the cumulated coefficients can be found in Figure 3.8. The cumulation of the approximation $N_0[n]$ is a bit to large, which could be expected since we did not take care about the sub-exponential factor when defining $N_0[n]$. We did so when defining $N_1[n]$ and the cumulation of $N_1[n]$ perfectly follows the cumulation of $N[w_k]$. We conclude that for DNCs with strings taking incommensurable weights, we can predict the cumulated number of distinct strings (the number of distinct strings of weight w or smaller that are accepted by the channel) with sub-exponential precision. $\lhd$

4. Information Theoretic Aspects

In this final chapter of our work, we will discuss information theoretic aspects of DNCs. Our objective is to relate the capacity of a DNC to the maximum rate of information per string weight at which strings generated by a random source can be transmitted over the DNC. We will proceed in the following way: first, we will define the capacity C of DNCs. We will then prove for general DNCs that every rate C' smaller than the capacity is achievable. For uniquely decodable codes, we will define a capacity achieving random source.

4.1. Capacity

4.1.1. Definition

For sake of illustration, we interpret in the following DNCs as communication channels, and we interpret the weights associated with the strings allowed by a channel as durations in time. Anyhow, our discussion applies in the same way when the considered DNC represents a storage system our any other system of interest.

We ask the following question:

> "If we can use a channel for a time t, how much information can be transmitted over the channel in the maximum?"

Following Shannon, we mean by information the logarithm of the number of choices we have when we use the channel. We interpret the time t at our disposal as follows: we can use the channel for any time w with $0 \leq w \leq t$. However, we can only use it once. We denote by $\sqcup$ when the channel is idle. For example, let DNC $\mathcal{B} = (B, v)$ with $B = \{0,1\}^\star$ and $v(0) = v(1) = 1$ represent a channel of interest. Assume further that we can use this channel for a time $t = 6$. Then 01100, $\sqcup$100, and $\sqcup$01$\sqcup$ represent valid channel usages, but $\sqcup$0$\sqcup$1 does not represent a valid channel usage. We do not differ between $\sqcup$100 and $\sqcup$100$\sqcup$.

For a general DNC $\mathcal{A} = (A, w)$, let $\{w_k\}_{k=1}^{\infty}$ with $w_1 < w_2 < \cdots$ denote the set of possible string lengths. The number of choices we have for the channel usage is then given by

$$\sum_{w_k \leq t} N[w_k] = \sum_{w_k \leq w_l} N[w_k] \tag{4.1}$$

$$= \sum_{k \leq l} N[w_k] \tag{4.2}$$

where $N[w_k]$ is equal to the number of distinct strings of length w_k that are accepted by the channel, and where w_l is defined as

$$w_l = \max_{w_k \leq t} w_k. \tag{4.3}$$

To avoid mathematical confusions in the following, we will use w_l rather than t to measure the time we have the channel at our disposal. The maximum information $I(w_l)$ that can be transmitted over the channel in the time w_l is given by

$$I(w_l) = \log\left(\sum_{k \leq l} N[w_k]\right). \tag{4.4}$$

This formula tells us exactly how much information we can transmit over the channel in the maximum when we use it for a time equal to or smaller than w_l.

A more global measure of the channel performance is the ratio of information per channel usage time, which is given by $I(w_l)/w_l$ for $w_l \to \infty$.

Definition 10. The *operational capacity* C_{op} of a DNC $\mathcal{A}$ we define as

$$C_{\mathrm{op}} = \lim_{l \to \infty} \frac{\log\left(\sum_{k \leq l} N[w_k]\right)}{w_l}. \tag{4.5}$$

where $N[w_k]$ denotes the number of distinct strings of weight w_k that are accepted by $\mathcal{A}$.

In [7], the authors gave following Shannon a different definition for the capacity of a DNC:

Definition 11. The *combinatorial capacity* C_{comb} of a DNC $\mathcal{A}$ is defined as

$$C_{\mathrm{comb}} = \lim_{k \to \infty} \sup \frac{\log N[w_k]}{w_k} \tag{4.6}$$

where $N[w_k]$ denotes the number of distinct strings of weight w_k that are accepted by $\mathcal{A}$.

Both definitions lead to the same value for the capacity of a DNC. We state this in the following Theorem:

Theorem 5. *The combinatorial capacity C_{comb} of a DNC $\mathcal{A}$ is equal to the operational capacity C_{op} of $\mathcal{A}$ and both are given by*

$$C_{\mathrm{op}} = C_{\mathrm{comb}} = \log Q \tag{4.7}$$

where Q^{w_k} is the exponential order of $N[w_k]$, which denotes the number of distinct strings of weight w_k that are accepted by $\mathcal{A}$.

Proof. For a proof, see Appendix B.3. □

In the following, we will simply speak of the capacity C of a DNC, but we keep in mind that the capacity of a DNC has an operational interpretation.

4.1.2. Calculation

Theorem 5 allows us to use the results we obtained in Section 3.2 for the exponential behavior of $N[w_k]$ to calculate the capacity of a DNC. For completeness, we will express Lemma 5 and Corollary 1 in terms of channel capacity.

Theorem 6. *(See Lemma 5). Let $\mathcal{A}$ represent a DNC with the generating function $\mathrm{G}_{\mathcal{A}}(y)$ given by*

$$\mathrm{G}_{\mathcal{A}}(y) = \sum_{k=1}^{\infty} N[w_k] y^{w_k}. \tag{4.8}$$

The capacity of $\mathcal{A}$ is then given by

$$C = -\log R \tag{4.9}$$

where R is the radius of convergence of $\mathrm{G}_{\mathcal{A}}(y)$.

Theorem 7. *(See Corollary 1). Suppose that the generating function of a DNC $\mathcal{A}$ can be written as*

$$\mathrm{G}_{\mathcal{A}}(y) = \frac{n_1 y^{\tau_1} + n_2 y^{\tau_2} + \cdots + n_p y^{\tau_p}}{d_1 y^{\nu_1} + d_2 y^{\nu_2} + \cdots + d_q y^{\nu_q}}, \qquad \tau_1, \ldots, \tau_p, \nu_1, \ldots, \nu_q \in \mathbb{R}^{\oplus} \tag{4.10}$$

for some finite positive integers p and q. The capacity of $\mathcal{A}$ is then given by

$$C = -\log P \tag{4.11}$$

where P is the smallest positive pole of $\mathrm{G}_{\mathcal{A}}(y)$.

When calculating the capacity of a DNC that results from combining two DNCs by union or concatenation, we have the following interesting properties.

Lemma 9. *Let $\mathcal{A}$ and $\mathcal{B}$ represent two DNCs with the capacities $C_{\mathcal{A}}$ and $C_{\mathcal{B}}$ respectively. Then*

i. The capacity of the DNC $\mathcal{A} \cup \mathcal{B}$ is given by

$$C = \max\{C_{\mathcal{A}}, C_{\mathcal{B}}\}. \tag{4.12}$$

ii. The capacity of the DNC $\mathcal{A}\mathcal{B}$ is given by

$$C = \max\{C_{\mathcal{A}}, C_{\mathcal{B}}\}. \tag{4.13}$$

Proof. i. Without loss of generality, we assume $C_{\mathcal{A}} \geq C_{\mathcal{B}}$. For the generating function of $\mathcal{A} \cup \mathcal{B}$ we have from Lemma 1

$$\mathrm{G}_{\mathcal{A}}(y) \leq \mathrm{G}_{\mathcal{A}\cup\mathcal{B}}(y) \leq \mathrm{G}_{\mathcal{A}}(y) + \mathrm{G}_{\mathcal{B}}(y). \tag{4.14}$$

From Lemma 5, we know that the capacity of the DNC $\mathcal{A} \cup \mathcal{B}$ is determined by the radius of convergence R of the generating function $\mathrm{G}_{\mathcal{A}\cup\mathcal{B}}(y)$. But we know from (4.14) that R is both upper-bounded and lower-bounded by the radius of convergence of $\mathrm{G}_{\mathcal{A}}(y)$. We therefore have $C = C_{\mathcal{A}}$.

ii. The proof of the second part of the lemma is almost identical to the proof of the first part, with the only difference that we use from Lemma 2 the bounds

$$\mathrm{G}_{\mathcal{A}}(y) \leq \mathrm{G}_{\mathcal{A}\mathcal{B}}(y) \leq \mathrm{G}_{\mathcal{A}}(y)\,\mathrm{G}_{\mathcal{B}}(y) \tag{4.15}$$

for the generating function of $\mathcal{A}\mathcal{B}$. □

Lemma 9 has an interesting consequence. Let $\mathcal{A} = (A, w)$ and $\mathcal{B} = (B, w)$ denote two DNCs. Even if $A \cap B \neq \emptyset$, we have

$$\mathrm{G}_{\mathcal{A}\cup\mathcal{B}}(y) =_c \mathrm{G}_{\mathcal{A}}(y) + \mathrm{G}_{\mathcal{B}}(y) \tag{4.16}$$

where $=_c$ denotes equality with respect to capacity. This stands in contrast to condition (2.17) in Lemma 1. In the same way, we have

$$\mathrm{G}_{\mathcal{A}\mathcal{B}}(y) =_c \mathrm{G}_{\mathcal{A}}(y)\,\mathrm{G}_{\mathcal{B}}(y) \tag{4.17}$$

even if we have $\left|\bar{A}\bar{B}\right| \neq \left|\bar{A}\right|\left|\bar{B}\right|$ for some finite subsets $\bar{A} \subseteq A$ and $\bar{B} \subseteq B$. This stands in contrast to condition (2.21) in Lemma 2.

4.1.3. Capacity of a DNC Represented by a Finite State Machine

As an application, we calculate the capacity of a DNC $\mathcal{A} = (A, w)$ where A is the set of strings we obtain by reading of the labels of branches in some finite state machine (FSM). Let the FSM contain L states. Let $A_{i,j}$ denote the set of paths from state i to state j. The generating function of the DNC $\mathcal{A}_{i,j} = (A_{i,j}, w)$ is then given by

$$\mathrm{G}_{\mathcal{A}_{i,j}}(y) = \left[\sum_{k=0}^{\infty} P^k(y)\right]_{i,j} \tag{4.18}$$

$$= \left\{[I - P(y)]^{-1}\right\}_{i,j} \tag{4.19}$$

where we implicitly assumed that the matrix $I - P(y)$ is invertible. The (i, j)th entry of the $L \times L$ matrix $P(y)$ is the generating function of the set of strings that result from going from state i to state j in one step. For example, if there are two branches from state 1 to state 2, one of weight 2 and one of weight π, then

$$[P(y)]_{1,2} = y^2 + y^{\pi}. \tag{4.20}$$

Since

$$\mathcal{A} = \bigcup_{i,j} \mathcal{A}_{i,j} \tag{4.21}$$

where the union is over $i, j = 1, \ldots, L$, we have for the generating function of $\mathcal{A}$

$$\mathrm{G}_{\mathcal{A}}(y) = \mathrm{G}_{\bigcup\limits_{i,j} \mathcal{A}_{i,j}}(y) \tag{4.22}$$

$$=_c \sum_{i,j} \mathrm{G}_{\mathcal{A}_{i,j}}(y) \tag{4.23}$$

$$= \sum_{i,j} \left\{ [I - P(y)]^{-1} \right\}_{i,j} \tag{4.24}$$

where equality with respect to capacity in (4.23) follows from Lemma 9. The capacity C of the DNC $\mathcal{A}$ is therefore given by

$$C = -\log P \tag{4.25}$$

where P is according to Theorem 7 the smallest positive pole of (4.24). Note that we did not use advanced results from matrix theory in our derivation. For the case of integer valued symbol weights, a similar technique is used in [11, Section V.6]. The authors call the matrix $P(y)$ the transfer matrix. Normally, the calculation of the capacity of a DNC represented by a FSM is based on matrix theory and the application of the Perron-Frobenius Theorem. In [12, Chapter 3], the authors discuss in detail the derivation of the capacity of DNCs represented by FSMs. However, the symbols they assign to the branches are all of weight 1. As a consequence, they do need to use the concept of generating functions in their derivations. The authors of [7] derive the capacity for the general case directly from the matrix series

$$Q(y) = \sum_{k=0}^{\infty} P^k(y). \tag{4.26}$$

By using arguments from matrix theory, they show that

$$C = -\log y_0 \tag{4.27}$$

where y_0 is the smallest positive real solution of the equation

$$\det[I - P(y)] = 0. \tag{4.28}$$

4.2. Channel Coding

In the previous section, we asked for the choices we have to transmit data over a DNC. This led to the notion of the capacity of a DNC, i.e., how much information per string weight we can transmit over the DNC in the maximum. Now, we actually want to transmit data over the channel. We thus consider a random source X that generates strings accepted by the channel. The maximum rate of information per symbol weight that is generated by a random source is given by its entropy rate. Our objective is to relate the maximum entropy rate of X to the capacity of the channel.

4.2.1. Random Walk in a Tree

What is the entropy rate of a random process generating strings accepted by a general DNC $\mathcal{A} = (A, w)$? To answer this question, we represent the set A by a tree T. If $s, t \in A$ are atomic (they do not result from concatenations of other strings in A), and $st \in A$, then s and t are branches in T, there is a node between the branch s and the branch t, and st is a path in T. We denote by start(s) the node where the branch s starts and we denote by end(s) the node where the branch s ends. We consider a random process generating a walk down the tree starting at the root of the tree. Conditioned on the current node, we assign probabilities to the branches below the current node. An observation of the random process results in the next branch we take in our walk. We define the entropy rate of the walk as follows:

Definition 12. We consider the DNC $\mathcal{A} = (A, w)$ and represent the set of accepted strings A by a tree. Let the random process $X = \{X_l\}_{l=1}^{\infty}$ generate a random walk down the tree starting from its root. We define the entropy rate $\bar{H}(X)$ as

$$\bar{H}(X) = \limsup_{n \to \infty} \frac{H(X_1, X_2, \ldots, X_n)}{W_n} \tag{4.29}$$

where W_n is the average weight of all paths that start at the root and consist of n branches.

4.2.2. DNC Coding Theorem

We can now state the fundamental theorem of the DNC.

Theorem 8. *Let $\mathcal{A}$ represent a DNC with capacity C. For every rate $C' < C$, there exists a random process X of entropy rate $\bar{H}(X) = C'$ that generates strings accepted by the DNC $\mathcal{A}$. Conversely, if $\bar{H}(X) > C$, then the strings generated by X cannot be transmitted over the DNC $\mathcal{A}$.*

This theorem was first stated by Shannon in [1]. In [7], it was proved for DNCs with non-integer valued string weights that can be represented by a FSM. In the following, we will prove that every rate smaller than the capacity is achievable. We will not prove the converse rigorously, but we refer to our discussion of the operational meaning of the capacity in Section 4.1, which makes it reasonable to assume that the converse holds. To prove the first part of Theorem 8, we first show that it holds for a special class of DNCs.

Lemma 10. *Let $\mathcal{A} = (A, w)$ represent a DNC with capacity C and let A be represented by a tree. Let $X = \{X_l\}_{l=1}^{\infty}$ denote a random walk down the tree with the probabilities*

$$P\left[X_{l+1} = t \middle| \mathrm{end}(X_l) = \mathrm{start}(t)\right] = e^{-w(t)C}. \tag{4.30}$$

We assume that these probabilities define a distribution, i.e., that probabilities assigned to branches that start at the same node sum up to 1. Then the entropy rate $\bar{H}(X)$ is equal C.

Proof. Let $s = s_1 s_2 \cdots s_n$ denote a path that starts at the root and consists of n branches. The probability of $(X_1, \dots, X_n) = s$ is given by

$$\begin{aligned} p(s) &= p(s_1)p(s_2|s_1)p(s_3|s_1 s_2)\cdots p(s_n|s_1\cdots s_{n-1}) && (4.31)\\ &= e^{-w(s_1)C} e^{-w(s_2)C} \cdots e^{-w(s_n)C} && (4.32)\\ &= e^{-[w(s_1)+w(s_2)+\cdots+w(s_n)]C} && (4.33)\\ &= e^{-w(s)C}. && (4.34) \end{aligned}$$

Assume that there are m paths that start at the root and consist of n branches. We denote these paths by $s^{(1)}, s^{(2)}, \dots, s^{(m)}$ respectively. We then have

$$\begin{aligned} \frac{H(X_1, X_2, \dots, X_n)}{W_n} &= \frac{-\sum\limits_{i=1}^{m} p(s^{(i)}) \log p(s^{(i)})}{\sum\limits_{i=1}^{m} p(s^{(i)}) w(s^{(i)})} && (4.35)\\ &= \frac{-\sum\limits_{i=1}^{m} e^{-w(s^{(i)})C} \log e^{-w(s^{(i)})C}}{\sum\limits_{i=1}^{m} e^{-w(s^{(i)})C} w(s^{(i)})} && (4.36)\\ &= C\,\frac{\sum\limits_{i=1}^{m} e^{-w(s^{(i)})C} w(s^{(i)})}{\sum\limits_{i=1}^{m} e^{-w(s^{(i)})C} w(s^{(i)})} && (4.37)\\ &= C. && (4.38) \end{aligned}$$

We thus have

$$\begin{aligned} \bar{H}(X) &= \lim_{n\to\infty} \frac{H(X_1, X_2, \dots, X_n)}{W_n} && (4.39)\\ &= C && (4.40) \end{aligned}$$

which concludes the proof. □

We are now ready to prove the first part of Theorem 8. The only difference to the Lemma 10 is that we do no longer assume a priori that the probabilities of branches that start at the same node sum up to 1.

Proof of the first part of Theorem 8. Let $\mathcal{A} = (A, w)$ denote a general DNC. As in the proof of Lemma 10, we represent the set A by a tree. To every branch s, we assign the probability

$$P\left[X_{l+1} = s \middle| \mathrm{end}(X_l) = \mathrm{start}(s)\right] = e^{-w(s)C'}. \quad (4.41)$$

For branches starting at the same node, we investigate the sum of their probabilities, which is given by

$$\sum_{s\colon \mathrm{end}(X_l)=\mathrm{start}(s)} P\left[s\middle|X_l\right]. \tag{4.42}$$

There are three possible scenarios:

1. The sum is equal 1. In this case we leave the probabilities unchanged.

2. The sum is larger than 1. We have two possibilities to correct this:

   - We successively delete branches (including the whole subtrees below them) from the tree as long as the sum remains larger or equal 1.
   - We change the value C' in the formula for the probabilities of the current branches to some $\bar{C}$ with $C' < \bar{C} \leq C$.

   If we cannot make the sum equal 1 by these two techniques, we consider the branches concatenated with their sub-branches as the new branches starting at the current node. We continue with this until a depth d_1 where we will succeed to make the sum equal 1.

3. The sum is smaller than 1. As in scenario 2., we concatenate the sub-branches to the branches until a depth d_2 where the sum gets larger than 1.

To guarantee that this algorithm works, d_1 and d_2 have to be bounded. Since the considered DNC is of capacity C, the number of accepted strings of weight w increases with an exponential order equal to e^{wC}, see Theorem 5. The probability of strings of weight w decreases with an exponential order equal to $2^{-wC'}$. Since $C' < C$, the number of strings increases exponentially faster than the probability of strings decreases. This guarantees that the sum over the probabilities will be larger than 1 for some bounded d. We use this observation in scenario 2. and scenario 3. It remains to comment why it is always possible to make the sum become exactly equal 1. As we increase the size of the branches by concatenating them to their sub-branches, we decrease their probabilities. Because of this, the deletion of branches together with the variation of $\bar{C}$ between C' and C will allow us to adjust the sum exactly to 1.

The deletion of branches from the tree decreases the entropy rate, but not below C' (consider the remaining tree together with Lemma 10). Concatenating branches to their sub-branches does not change the entropy rate (consider the resulting tree together with Lemma 10). The variation of $\bar{C}$ between C' and C only increases the entropy rate. Together we have

$$\bar{H}(X) \geq C' \tag{4.43}$$

which concludes the proof. □

4.3. Uniquely Decodable Codes

4.3.1. Capacity Achieving Distribution

There are DNCs where it is possible to define a capacity achieving random source . In [7], the authors showed that such a source exists for all DNCs that can be represented by a strongly connected FSM. Here, we will consider DNCs that are of the form $\mathcal{A} = (A^\star, w)$ where $A = \{a_1, a_2 \dots, a_m\}$ forms a uniquely decodable code. The generating function of $\mathcal{A}$ is given by

$$\mathrm{G}_{\mathcal{A}}(y) = \sum_{k=0}^{\infty} \left[y^{w(a_1)} + y^{w(a_2)} + \cdots + y^{w(a_m)} \right]^k. \tag{4.44}$$

According to Theorem 6, the capacity of $\mathcal{A}$ is given by $C = -\log R$, where R is the radius of convergence of $\mathrm{G}_{\mathcal{A}}(x)$. The number R is in our case given by the smallest positive solution of

$$y^{w(a_1)} + y^{w(a_2)} + \cdots + y^{w(a_m)} = 1. \tag{4.45}$$

Let X denote a random process that generates strings accepted by $\mathcal{A}$. We now want to find the capacity achieving distribution of X. To do this, we represent $A^\star$ by a tree. This leads to a periodic tree that is of the same form below every node, i.e., the set of branches that start at the same node is identical for every node. Conditioned on the current node, Lemma 10 suggest for the branch a_i the probability

$$p(a_i) = e^{-w(a_i)C}. \tag{4.46}$$

We assign to the branches in our tree the probabilities suggested by Lemma 10. If the probabilities of branches starting at the same node sum up to one, then they define a distribution, and according to Lemma 10, this distribution is capacity achieving. For the sum we have

$$\begin{aligned}
\sum_{i=1}^{m} p(a_i) &= \sum_{i=1}^{m} e^{-w(a_i)C} && (4.47)\\
&= \sum_{i=1}^{m} e^{-w(a_i)(-\log R)} && (4.48)\\
&= \sum_{i=1}^{m} e^{w(a_i)\log R} && (4.49)\\
&= \sum_{i=1}^{m} R^{w(a_i)} && (4.50)\\
&= 1 && (4.51)
\end{aligned}$$

where equality in (4.51) follows from the fact that R is a solution of (4.45). The random process $X = \{X_l\}_{l=1}^{\infty}$ with X_l independent and identically distributed (i.i.d.) according to

$$p(a_i) = e^{-w(a_i)C}, \qquad i = 1, \dots, m \tag{4.52}$$

is therefore capacity achieving.

4.3.2. McMillan's Inequality

We will now use this result to generalize McMillan's inequality, which gives a necessary condition for a set to form a uniquely decodable code. From (4.51), we know that the capacity C of $\mathcal{A}$ satisfies the equation

$$\sum_{i=1}^{\infty} e^{-w(a_i)C} = 1. \tag{4.53}$$

We will now derive an upper bound on the capacity C and we will plug it into (4.53). This will lead to a generalized form of McMillan's inequality. Since the distribution of X is capacity achieving, we know that the entropy rate of X is equal to the capacity C. Let W_n denote the average weight of paths that start at the root and consist of n branches. We then have

$$\begin{aligned}
C &= \bar{H}(X && (4.54)\\
&= \lim_{n\to\infty} \frac{H(X_1,\dots,X_n)}{W_n} && (4.55)\\
&= \lim_{n\to\infty} \frac{\sum_{l=1}^{n} H(X_l)}{W_n} && (4.56)\\
&= \lim_{n\to\infty} \frac{nH(X_1)}{nW_1} && (4.57)\\
&= \frac{H(X_1)}{W_1} && (4.58)\\
&\leq \frac{\log|A|}{W_1} && (4.59)\\
&\leq \frac{\log|A|}{\min\limits_{a_i\in A} w(a_i)} && (4.60)
\end{aligned}$$

where we have equality in (4.56) and (4.57) since the random variables X_l are i.i.d. The inequality in (4.59) follows from a known upper-bound on entropy, see [2]. We have inequality in (4.60) since the average weight of the branches that start at the same node is lower bounded by the smallest branch weight. We define

$$w_{\min} = \min_{a_i\in A} w(a_i) \tag{4.61}$$

and plug (4.60) into (4.53). We then have

$$1 = \sum_{i=1}^{\infty} e^{-w(a_i)C} \tag{4.62}$$

$$\geq \sum_{i=1}^{\infty} e^{-w(a_i)\frac{\log|A|}{w_{\min}}} \tag{4.63}$$

$$= \sum_{i=1}^{\infty} |A|^{-\frac{w(a_i)}{w_{\min}}}. \tag{4.64}$$

We formulate this result in the following theorem:

Theorem 9. *Let $A = \{a_1, \ldots, A_m\}$ denote a set of strings with the associated positive weights $w(A) = \{w(a_1), \ldots, w(a_m)\}$. Let $w_{\min}$ denote the smallest string weight that occurs in $w(A)$. If the pair (A, w) forms a uniquely decodable code, then the following inequality holds*

$$\sum_{i=1}^{\infty} |A|^{-\frac{w(a_i)}{w_{\min}}} \leq 1. \tag{4.65}$$

4.4. Shannon's Telegraphy Channel

In this section, we calculate the capacity and the capacity achieving distribution for Shannon's telegraphy channel $\mathcal{T}$. To illustrate the results from Section 4.2, our derivation will be based on the representation of the channel $\mathcal{T}$ by a tree. Note that the authors of [7] derived the same capacity achieving distribution by means of matrix theory.

4.4.1. Capacity

In Section 3.4, we calculated for the generating function of $\mathcal{T}$ the exponential order of its coefficients $N[w_k]$. We did this by determining the smallest positive solution P of the equation

$$1 - (z^2 + z^4 + z^5 + z^7 + z^8 + z^{10}) = 0. \tag{4.66}$$

According to Theorem 7, the capacity of $\mathcal{T}$ is then given by

$$C = -\log P. \tag{4.67}$$

Equation (4.66) can also be put into the form

$$z^2 + z^4 + z^5 + z^7 + z^8 + z^{10} = 1 \tag{4.68}$$

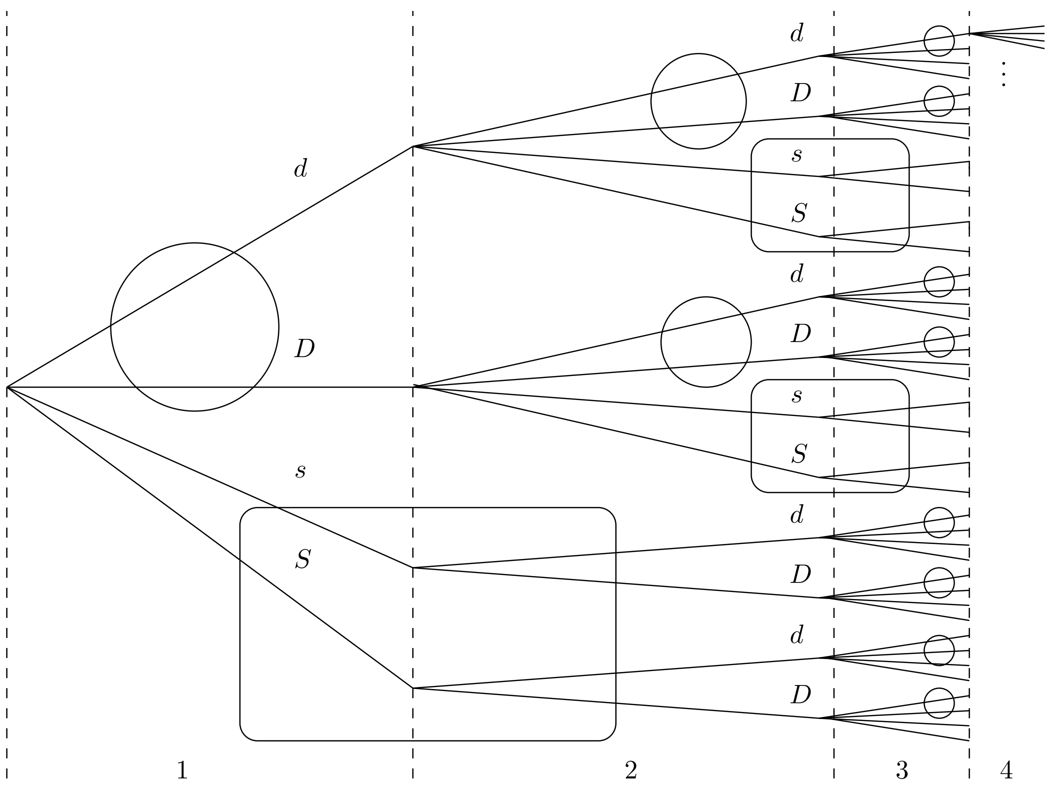

Figure 4.1.: Strings accepted by the telegraphy channel represented by a tree.

4.4.2. Capacity Achieving Distribution

We represent the strings allowed by $\mathcal{T}$ by a tree, see Figure 4.1. The random process $X = \{X_l\}_{l=1}^{\infty}$ generates a walk down the tree. In the lth step, the random variable X_l takes a value in

$$\{d, D, s, S\} \tag{4.69}$$

according to some distribution. To ensure that this distribution is capacity achieving, we have to guarantee that the string $s = X_1 X_2 \dots$ can be divided into finite substrings $s = s_1 s_2 s_3 \cdots$ such that the conditioned probability

$$p_i(a) = P[s_i = a | s_1, s_2, \dots, s_{i-1}] \tag{4.70}$$

is given by

$$p_i(a) = e^{-w(a)C}. \tag{4.71}$$

We take a closer look at Figure 4.1 and see that the tree consists of the periodic repetition of two different subtrees. One is marked by a circle and the other is marked by a box. Because of the periodic repetition, we have to guarantee (4.71) for the paths through these subtrees. For the encircled subtree we therefore have

$$p_c(d) = e^{-2C} \tag{4.72}$$

$$p_c(D) = e^{-4C} \tag{4.73}$$

and for the boxed subtree we have

$$p_b(sd) = e^{-(3+2)C} = e^{-5C} \tag{4.74}$$

$$p_b(sD) = e^{-(3+4)C} = e^{-7C} \tag{4.75}$$

$$p_b(Sd) = e^{-(6+2)C} = e^{-8C} \tag{4.76}$$

$$p_b(SD) = e^{-(6+4)C} = e^{-10C}. \tag{4.77}$$

These assignments are only valid if the probabilities of the paths of depth 2 that start at the root sum up to 1. We do not know the probabilities of the paths of depth 2 that start with d or D, but we know that the sum of their probabilities is given by

$$p_c(d) + p_c(D) = e^{-2C} + e^{-4C}. \tag{4.78}$$

We thus have to check if the sum

$$e^{-2C} + e^{-4C} + e^{-5C} + e^{-7C} + e^{-8C} + e^{-10C} \tag{4.79}$$

is equal to one. Substituting in the left-hand side of (4.68) the variable z by the solution e^{-C} results in (4.79). Since the right-hand side of (4.68) is equal to one, we conclude that the sum (4.79) is equal to one. Our probability assignments are therefore valid. The remaining probabilities are given by

$$p_b(s) = p_b(sd) + p_b(sD) = e^{-5C} + e^{-7C} \tag{4.80}$$

$$p_b(S) = p_b(Sd) + p_b(SD) = e^{-8C} + e^{-10C} \tag{4.81}$$

$$p_b(d) = \frac{p_b(sd)}{p_b(s)} = \frac{p_b(Sd)}{p_b(S)} = \frac{1}{1+e^{-2C}} \tag{4.82}$$

$$p_b(D) = \frac{p_b(sD)}{p_b(s)} = \frac{p_b(SD)}{p_b(S)} = \frac{1}{1+e^{2C}}. \tag{4.83}$$

Inside the boxed subtree, the amount of information per symbol weight varies around C. However, for the paths through the boxed subtree, the amount of information per string weight is exactly equal C. For the encircled subtree, the amount of information per string weight is also exactly equal C. It follows from our results in Section 4.2 that the distribution as defined by the upper equations is capacity achieving.

5. Conclusions

In this work, we developed a technique to investigate the asymptotic behavior of combinatorial structures of exponentially increasing complexity by analytic methods. Our results for the exponential behavior apply in the general case where incommensurable string weights are allowed, and where the considered structure can possibly not be represented by a FSM. In this way, we generalized the corresponding results from [1] and [7]. From an engineering point of view, it is questionable if there will ever be a need for these generalizations in practice. First, all our observations of the real world result in a finite set of rational numbers and rational numbers from a finite set are always commensurable. Second, if a system specification cannot be represented by a FSM, then it is in general of infinite memory. We therefore see the analytic method rather as an alternative to the method used in [1] and [7], which is based on matrix theory. We believe that our method together with the results from [14] and [13] will in many cases allow a simpler and more elegant solution of problems in the practice.

In our derivations, we generalized Pringsheim's Theorem to the case of non-integer valued string weights. Pringsheim's Theorem in its original form states that if the Taylor series expansion of an analytic function $f(z)$ around the origin has nonnegative coefficients, and if the radius of convergence of the Taylor series expansion is equal R, then $f(z)$ has a singularity in $z = R$. We conjecture that Pringsheim's Theorem can be strengthened, and that $f(z)$ has under the upper conditions a pole in $z = R$. If our conjecture holds, then we can calculate the exponential behavior of a general DNC from its real generating function by looking for its smallest positive pole. This would be a powerful result.

For the number $N[w]$ of distinct strings of weight w that are accepted by a DNC, we showed how we can predict the sub-exponential behavior from the analytic characteristics of the corresponding generating function with in many cases arbitrary precision. It would be interesting to find a practical problem where the knowledge of the sub-exponential behavior is necessary. In the most part of problems in communications, it suffices to determine the exponential behavior of the system of interest. Perhaps the sub-exponential behavior is of interest when investigating the performance of universal codes and pattern codes.

By an example, we showed that for incommensurable string weights, not the sub-exponential behavior of $N[w]$, but the sub-exponential behavior of the sum $\sum_{v\leq w} N[v]$ can be obtained from the analytic characteristics of the corresponding generating function. It would be interesting to investigate this relationship with mathematical rigor.

Another open problem is the converse of the fundamental theorem of DNCs. In [7], it was proved for the case where the considered DNC can be represented by a strongly connected FSM. It seems clear that the converse also holds for the general case, a rigorous mathematical proof for general DNCs is however still missing.

A. Notation

$\mathcal{A}$,$\mathcal{B}$	discrete noiseless channel
a.e.	almost everywhere
DNC	discrete noiseless channel
FSM	finite state machine
GEN (x)	generating series
G (y)	real generating function
G (z)	complex generating function
G (e^s)	complex generating function
i.i.d.	independent and identically distributed
i.o.	infinitely often
log	natural logarithm
r.o.c.	region of convergence

B. Mathematics

B.1. Complex Analysis

B.1.1. Analytic Functions

Definition 13. Let Ω be a region in $\mathbb{C}$ (an open and connected subset of $\mathbb{C}$). A function $f\colon \Omega \mapsto \mathbb{C}$ is called *analytic* in Ω if the limes

$$\lim_{z\to z_0} \frac{f(z)-f(z_0)}{z-z_0} \tag{B.1}$$

exists for all $z_0 \in \Omega$.

Lemma 11. [22, p. 35]. *Let f, g be functions analytic in Ω. For all $z \in \Omega$ we have:*

i. Linearity: For arbitrary $\lambda, \mu \in \mathbb{C}$ it holds that

$$(\lambda f + \mu g)'(z) = \lambda f'(z) + \mu g'(z). \tag{B.2}$$

ii. Product- and quotient-rule:

$$(fg)'(z) = f'(z)g(z) + f(z)g'(z) \tag{B.3}$$

$$(f/g)'(z) = \frac{f'(z)g(z) - f(z)g'(z)}{g^2(z)}. \tag{B.4}$$

iii. Chain-rule: If f is analytic in z and g is analytic in $w = f(z)$, then $g \circ f$ is differentiable in z and we have

$$(g \circ f)'(z) = g'(f(z))f'(z). \tag{B.5}$$

Theorem 10. [22, p. 85] (Taylor series expansion). *Let Ω be a region in $\mathbb{C}$. Let $f\colon \Omega \mapsto \mathbb{C}$ be a function analytic in Ω. Let a be a point in Ω with distance from the boundary of Ω equal δ. Then*

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(z-a)^k \tag{B.6}$$

for all $z \in D_\delta(a)$. The region $D_\delta(a)$ is an open disc around a. The boundary of $D_\delta(a)$ is of radius δ.

Theorem 11. [22, p. 99] (Laurent series expansion). *Let f be analytic in*

$$\Omega \supset \{z \in \mathbb{C} \big| a < |z| < b\}, \qquad 0 \leq a < b \leq \infty. \tag{B.7}$$

Then

$$f(z) = \sum_{k=-\infty}^{\infty} c_k z^k, \qquad a < |z| < b \tag{B.8}$$

where the coefficients c_k are for $a < r < b$ given by

$$c_k = \frac{1}{2\pi i} \int_{\partial D_r(0)} \frac{f(\zeta)}{\zeta^k} \frac{d\zeta}{\zeta}, \qquad k \in \mathbb{Z}. \tag{B.9}$$

The integral is in counter-clockwise direction along the boundary $\partial D_r(0)$ of the disc $D_r(0)$ of radius r centered at zero. The part of the sum in (B.8) *with k negative is called the* main part *of the Laurent series expansion of $f(z)$ around zero.*

Definition 14. Let $f\colon \Omega \setminus a \mapsto \bar{\mathcal{C}}$ be a function analytic in the region $\Omega \setminus a$. If the limit $z \to a$ exists and is equal ∞, then a is called a *pole* of f and we can represent f by a Laurent series expansion around a with some smallest positive $n < \infty$ such that

$$f(z) = \sum_{k=-n}^{\infty} \frac{c_k}{(z-a)^k}. \tag{B.10}$$

We call n the *multiplicity* of a.

Definition 15. Let Ω be a region in $\mathbb{C}$. Functions $f\colon \Omega \mapsto \bar{\mathbb{C}}$ that are analytic in Ω except for a countable number of poles (where they take the value ∞) are called meromorphic in Ω.

B.1.2. Localization of Poles

Theorem 12. [11, p. 256] (Argument Principle). *Let $f(z)$ be an analytic function in a region Ω and let γ be a simple closed curve interior to Ω, and on which f is assumed to have no zeros. The number of zeros of $f(z)$ (counted with multiplicity) inside γ equals the winding number of the transformed contour $f(\gamma)$ around the origin.*

Theorem 13. *Suppose $f(x)$ is a polynomial with real coefficients and z is a complex root of $f(x)$. Then the conjugate of z is also a root of the polynomial.*

B.1.3. Analyticity of z^r

Making z^τ Analytic in $z_0 \neq 0$

Lemma 12. *For an arbitrary point $z_0 \in \mathbb{C} \setminus \{0\}$, a term of the form*

$$az^\tau, \qquad z \in \mathbb{C},\ a, \tau \in \mathbb{R} \tag{B.11}$$

can be turned into a function of z analytic in z_0.

Proof. If $z_0 \notin \mathbb{R}^{\ominus}$, we define

$$\operatorname{Log} z = \big\{ w \big| \exp(w) = z, -\pi < \Im\{w\} < \pi \big\}. \tag{B.12}$$

The function $\operatorname{Log} z$ is analytic in $\mathbb{C} \setminus \mathbb{R}^{\ominus}$. We define

$$\begin{aligned} g(z) &= a z^{\tau} && \text{(B.13)} \\ &= a \exp(\operatorname{Log} z)^{\tau} && \text{(B.14)} \\ &= a \exp(\tau \operatorname{Log} z). && \text{(B.15)} \end{aligned}$$

The function $h(z) = a \exp(\tau w)$ is analytic in $\mathbb{C}$. We write $g(z) = (h \circ \operatorname{Log})(z)$ and because of the chain-rule for analytic functions given in Lemma 11(iii.), we made $g(z)$ analytic in z_0. If $z_0 \in \mathbb{R}^{\ominus}$, we define

$$\operatorname{Log} z = \big\{ w \big| \exp(w) = z, 0 < \Im\{w\} < 2\pi \big\}. \tag{B.16}$$

The function $\operatorname{Log} z$ is then analytic in $\mathbb{C} \setminus \mathbb{R}^{\oplus}$ and we can again make $f(z)$ analytic in z_0. This concludes the proof. □

Why $\sqrt{z}$ Cannot Be Made Analytic In Zero

In $\mathbb{C}$, $\sqrt{z}$ has two distinct solutions. This can be seen in the following way: we write $z = |z| \exp(i \arg z)$. The two solutions of $\sqrt{z}$ can then be written as

$$z_k = \sqrt{|z|} \exp\left[i \left(\frac{\arg z}{2} + k \frac{2\pi}{2} \right) \right], \qquad k \in \{0, 1\}. \tag{B.17}$$

To make $\sqrt{z}$ analytic in zero we have to define a function $f(z)$ that fulfills $f^2(z) = z$ and is complex differentiable in a region Ω around zero. This implies that the function $f(z)$ is continuous in Ω. We consider $f(z)$ on a circle $w(t) \in \Omega$ defined as

$$t \mapsto w(t) = a \exp(it), \qquad a \in \mathbb{R}^+, 0 \le t \le 2\pi. \tag{B.18}$$

We define $f\big(w(0)\big)$ as

$$\begin{aligned} f\big(w(0)\big) &= \sqrt{|w(0)|} && \text{(B.19)} \\ &= \sqrt{a}. && \text{(B.20)} \end{aligned}$$

It follows from (B.17) that the only continuous expansion of f on $w(t)$ is

$$f\big(w(t)\big) = \sqrt{a} \exp\left(\frac{it}{2} \right). \tag{B.21}$$

However, for $t = 2\pi$ we have $w(2\pi) = w(0)$ but

$$\begin{aligned} f\big(w(t)\big) &= -\sqrt{a} && \text{(B.22)} \\ &\ne f\big(w(0)\big). && \text{(B.23)} \end{aligned}$$

We conclude that $\sqrt{z}$ cannot be made analytic in 0.

B.2. Miscellaneous Mathematics

B.2.1. Proof of Lemma 4

Lemma 6. *Let $\mathcal{A} = (A, w)$ represent a DNC with the set of accepted strings A resulting from the concatenations of symbols from the finite set D. Let*

$$\mathrm{GEN}_{\mathcal{A}}(x) = \sum_{k=1}^{\infty} N[w_k]x^k \tag{B.24}$$

denote the generating series of $\mathcal{A}$.

i. For any integer $n \geq 0$

$$\max_{w_k<n} k < Ln^K \tag{B.25}$$

for some constant $K > 0$ and some constant $L > 0$.

ii. There exist some constant $R > 0$ and some constant $M > 0$ such that

$$N[w_k] < MR^{w_k}, \qquad a.e. \tag{B.26}$$

with respect to k.

Proof. The weight function $w : A \mapsto \mathbb{R}^{\oplus}$ of $\mathcal{A}$ is completely defined by the weights of the symbols from the finite alphabet D. We assume $|D| = q$ and denote by $\nu_1 \leq \cdots \leq \nu_q$ the weights of the symbols from D. For every symbol $s \in A$ there exist some nonnegative integers $n_{s,1}, \ldots, n_{s,q}$ such that

$$w(s) = n_{s,1}\nu_1 + \cdots + n_{s,q}\nu_q. \tag{B.27}$$

We therefore have

$$w(A) \subseteq \left\{n_1\nu_1 + \cdots + n_q\nu_q \middle| n_1, \ldots, n_q \in \mathbb{Z}^{\oplus}\right\}. \tag{B.28}$$

We calculate an upper-bound for the left-hand side of (B.25).

$$\begin{aligned}
\max_{w_k<n} k &= |w \in w(A)\colon w < n| && \text{(B.29)}\\
&\leq \left|\left\{n_1, \ldots, n_q \in \mathbb{Z}^{\oplus} \middle| n_1\nu_1 + \cdots + n_q\nu_q < n\right\}\right| && \text{(B.30)}\\
&\leq \left|\left\{n_1, \ldots, n_q \in \mathbb{Z}^{\oplus} \middle| (n_1 + \cdots + n_q)\nu_1 < n\right\}\right| && \text{(B.31)}\\
&\leq \left|\left\{n_1, \ldots, n_q \in \mathbb{Z}^{\oplus} \middle| n_1, \ldots, n_q < \frac{n}{\nu_1}\right\}\right| && \text{(B.32)}\\
&\leq \left(\frac{n}{\nu_1}\right)^q && \text{(B.33)}
\end{aligned}$$

which shows that (B.25) is fulfilled for $K = q$ and $L = \nu_1^{-q}$. This proves *i.*

To prove *ii.*, we write

$$\begin{aligned}
N[w_k] &= [x^{w_k}]\mathrm{GEN}_{\mathcal{A}}(x) && \text{(B.34)}\\
&\leq [x^{w_k}]\left(\sum_{l=0}^{\infty}(x^{\nu_1} + \cdots + x^{\nu_q})^l\right) && \text{(B.35)}\\
&\leq \sum_{w \leq w_k}[x^w]\left(\sum_{l=0}^{\infty}(x^{\nu_1} + \cdots + x^{\nu_q})^l\right) && \text{(B.36)}\\
&\leq \sum_{w \leq w_k}[x^w]\left(\sum_{l=0}^{\infty}(x^{\nu_1} + x^{\nu_1} + \cdots + x^{\nu_1})^l\right) && \text{(B.37)}\\
&= \sum_{w \leq w_k}[x^w]\left(\sum_{l=0}^{\infty}q^l x^{l\nu_1}\right) && \text{(B.38)}\\
&\leq \sum_{l=0}^{\left\lceil \frac{w_k}{\nu_1}\right\rceil} q^l && \text{(B.39)}
\end{aligned}$$

where (B.35) follows from $A \subseteq D^\star$ and where the inequality in (B.37) holds because $\nu_1 \leq \nu_i$, $i = 1, \ldots, q$. We further get

$$\begin{aligned}
\sum_{l=0}^{\left\lceil \frac{w_k}{\nu_1}\right\rceil} q^l &= \frac{q^{\left\lceil \frac{w_k}{\nu_1}\right\rceil} - 1}{q-1} && \text{(B.40)}\\
&\leq \frac{q^{\frac{w_k}{\nu_1}+1} - 1}{q-1} && \text{(B.41)}\\
&\leq \frac{q^{\frac{w_k}{\nu_1}+1}}{q-1} && \text{(B.42)}\\
&= \frac{q}{q-1}\left(q^{\frac{1}{\nu_1}}\right)^{w_k} && \text{(B.43)}
\end{aligned}$$

which shows that (B.26) is fulfilled for $M = q/(q-1)$ and $R = q^{1/\nu_1}$. This proves *ii.* □

B.2.2. Proof of Lemma 6

Lemma 6. *Assume that $\{w_k\}_{k=1}^{\infty}$ is a strictly ordered sequence of positive real numbers that is not too dense in the sense that for any integer $n \geq 0$*

$$\max_{w_k < n} k \leq Ln^K \tag{B.44}$$

for some constant $L > 0$ and some constant $K \geq 0$. Let ρ be a nonnegative real number. Then the sum

$$\sum_{k=1}^{\infty} \rho^{w_k} \tag{B.45}$$

converges if and only if $\rho < 1$.

Proof. We have to show that the sum (B.45) diverges for $\rho \geq 1$ and converges for $0 < \rho < 1$.

Assume $\rho \geq 1$. We then have

$$\sum_{k=1}^{n} \rho^{w_k} \geq \sum_{k=1}^{n} 1^{w_k} \tag{B.46}$$

$$= n. \tag{B.47}$$

The left side of this equality becomes the sum (B.45) for $n \to \infty$. However, the right-hand side of the inequality goes to infinity for $n \to \infty$. This implies that the sum (B.45) diverges for any $\rho \geq 1$.

Assume now that $0 < \rho < 1$. In this case we have

$$\sum_{k=1}^{\infty} \rho^{w_k} \leq \sum_{k=1}^{\infty} \rho^{\lfloor w_k \rfloor} \tag{B.48}$$

$$= \sum_{n=0}^{\infty} \rho^n \left(\max_{w_k < n+1} k - \max_{w_k < n} k \right) \tag{B.49}$$

$$\leq \sum_{n=0}^{\infty} \rho^n \max_{w_k < n+1} k \tag{B.50}$$

$$\leq \sum_{n=0}^{\infty} \rho^n L (n+1)^K \tag{B.51}$$

$$= \sum_{n=0}^{\infty} \frac{1}{\rho} \rho^{n+1} L n^K \tag{B.52}$$

$$= \frac{L}{\rho} \sum_{n=1}^{\infty} \left(\rho n^{\frac{K}{n}} \right)^n . \tag{B.53}$$

The sum in the last line converges if there exists for every ρ with $0 < \rho < 1$ a natural number n_0 such that for every $n > n_0$

$$\rho n^{\frac{K}{n}} < 1. \tag{B.54}$$

We prove this by showing that for any fixed $K > 0$, $n^{\frac{K}{n}}$ goes to 1 for $n \to \infty$. We have

$$\lim_{n\to\infty} n^{\frac{K}{n}} = \lim_{n\to\infty} \exp\left(\frac{K}{n}\ln n\right) \tag{B.55}$$

$$= \exp\left(\lim_{n\to\infty} \frac{K}{n}\ln n\right) \tag{B.56}$$

$$= \exp\left(\lim_{n\to\infty} \frac{K\frac{1}{n}}{1}\right) \tag{B.57}$$

$$= \exp(0) \tag{B.58}$$

$$= 1 \tag{B.59}$$

where equality in (B.56) holds since the function $\exp(z)$ is monotonically increasing, and where we used in (B.57) the rule of l'Hospital. □

B.2.3. Lower Bound of Factorial Function

From the derivation of Stirling's approximation given in [26] we know that

$$n! = \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n+\frac{\theta}{12n}} \qquad 0 < \theta < 1. \tag{B.60}$$

We thus have the following lower bound on $n!$:

$$n! = \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n+\frac{\theta}{12n}} \tag{B.61}$$

$$= \sqrt{2\pi n} e^{\frac{\theta}{12n}} \left(\frac{n}{e}\right)^n \tag{B.62}$$

$$\geq e^{\frac{\theta}{12n}} \left(\frac{n}{e}\right)^n \tag{B.63}$$

$$\geq e^{\frac{0}{12n}} \left(\frac{n}{e}\right)^n \tag{B.64}$$

$$= \left(\frac{n}{e}\right)^n. \tag{B.65}$$

B.2.4. Newton's Expansion

Theorem 14. (Newton's expansion). *The Taylor series expansion around* 0 *of the function $f : \mathbb{C} \to \mathbb{C}$,*

$$f(z) = \frac{1}{(1-z)^r}, \qquad r \in \mathbb{N},\ |z| < 1 \tag{B.66}$$

is given by

$$f(z) = \sum_{n=0}^{\infty} \binom{n+r-1}{r-1} z^n. \tag{B.67}$$

We present two proofs for this theorem.

Combinatorial proof of Theorem 14. We have

$$f(z) = \frac{1}{(1-z)^r} \tag{B.68}$$
$$= \left(\frac{1}{1-z}\right)^r \tag{B.69}$$
$$= \left(\sum_{k=0}^{\infty} z^k\right)^r \tag{B.70}$$

where we used in the last line the formula for geometric series. We denote by $[z^n]f(z)$ the nth coefficient of the Taylor series expansion of $f(z)$ around 0. We have

$$[z^n]f(z) = [z^n]\left(\sum_{k=0}^{\infty} z^k\right)^r \tag{B.71}$$
$$= [z^{n+r}]z^r\left(\sum_{k=0}^{\infty} z^k\right)^r \tag{B.72}$$
$$= [z^{n+r}]\left(\sum_{k=1}^{\infty} z^k\right)^r \tag{B.73}$$
$$= [z^{n+r}]\left(\sum_{k=1}^{n+1} z^k\right)^r . \tag{B.74}$$

The problem is now equivalent to the partition of the integer $n+r$ into r integers x_i, $1 \leq x_i \leq n+1$, such that

$$n + r = \sum_{i=1}^{r} x_i, \qquad x_i \in \mathbb{N}. \tag{B.75}$$

We write $n+r$ as an unary string $\bullet\bullet\bullet\cdots\bullet$ consisting of $n+r$ circles. We then partition this string by inserting a barrier $|$ at an arbitrary position in the string, for example $\bullet\bullet|\bullet\cdots\bullet$. We do this $r-1$ times and end up with r partitions. We have $(n+r-1)!/n!$ possibilities to insert barriers in the string, and the number of distinct partitions is

$$\frac{(n+r-1)!}{n!(r-1)!} = \binom{n+r-1}{r-1} \tag{B.76}$$

since the order in which we insert the barriers does not matter. We can now write the Taylor series expansion of $f(z)$ around 0 as

$$f(z) = \frac{1}{(1-z)^r} \tag{B.77}$$
$$= \sum_{n=0}^{\infty} \binom{n+r-1}{r-1} z^n \tag{B.78}$$

which concludes the proof. □

Analytic proof of Theorem 14. From the geometric series, we know that

$$\frac{1}{1-z} = \sum_{k=0}^{\infty} z^k. \tag{B.79}$$

We derivate the left-hand side m times and get

$$\frac{\mathrm{d}^m}{\mathrm{d}z^m}\left(\frac{1}{1-z}\right) = \frac{m!}{(1-z)^{m+1}}. \tag{B.80}$$

The mth derivation of the right-hand side is given by

$$\frac{\mathrm{d}^m}{\mathrm{d}z^m}\left(\sum_{k=0}^{\infty} z^k\right) = \sum_{k=m}^{\infty} \frac{k!}{(k-m)!} z^{k-m}. \tag{B.81}$$

Together we have

$$\frac{m!}{(1-z)^{m+1}} = \sum_{k=m}^{\infty} \frac{k!}{(k-m)!} z^{k-m}. \tag{B.82}$$

We substitue m in (B.82) by $r-1$. We then have

$$\frac{(r-1)!}{(1-z)^r} = \sum_{k=r-1}^{\infty} \frac{k!}{(k-r+1)!} z^{k-r+1} \tag{B.83}$$

$$\Leftrightarrow$$

$$\frac{1}{(1-z)^r} = \sum_{k=r-1}^{\infty} \frac{k!}{(k-r+1)!(r-1)!} z^{k-r+1} \tag{B.84}$$

$$= \sum_{n=0}^{\infty} \frac{(n+r-1)!}{n!(r-1)!} z^n \tag{B.85}$$

$$= \sum_{n=0}^{\infty} \binom{n+r-1}{r-1} z^n \tag{B.86}$$

which concludes the proof. □

B.3. Equality of Operational and Combinatorial Capacity

We consider in the following a general DNC $\mathcal{A} = (A, w)$. We write the set of possible string weights as $\{w_k\}_{k=1}^{\infty}$ with $w_1 < w_2 < \cdots$, and we assume that it is not too dense in the sense that

$$\max_{w_k \le n} k < Ln^K \tag{B.87}$$

for some constant $K > 0$ and some constant $L > 0$. For the sequence $\{N[w_k]\}_{k=1}^{\infty}$ of the number of distinct strings of the weight w_k, we assume that it is of exponential order

$$N[w_k] \bowtie Q^{w_k} \tag{B.88}$$

for some positive and finite Q.

Definition 10. The *operational capacity* C_{op} of $\mathcal{A}$ we define as

$$C_{\mathrm{op}} = \lim_{l\to\infty} \frac{\log\Bigl(\sum_{k\le l} N[w_k]\Bigr)}{w_l}. \tag{B.89}$$

Definition 11. The *combinatorial capacity* C_{comb} of $\mathcal{A}$ we define as

$$C_{\mathrm{comb}} = \limsup_{k\to\infty} \frac{\log N[w_k]}{w_k}. \tag{B.90}$$

Theorem 5. *The combinatorial capacity C_{comb} of $\mathcal{A}$ is equal to the operational capacity C_{op} of $\mathcal{A}$ and both are given by*

$$C_{\mathrm{op}} = C_{\mathrm{comb}} = \log Q \tag{B.91}$$

where Q^{w_k} is the exponential order of $N[w_k]$.

Proof. The Theorem follows directly from Lemma 13 and Lemma 14. □

Lemma 13. *The combinatorial capacity C_{comb} of $\mathcal{A}$ is given by*

$$C_{\mathrm{comb}} = \log Q \tag{B.92}$$

where Q^{w_k} is the exponential order of $N[w_k]$.

Proof. From Definition 7, $N[w_k]$ is of exponential order Q^{w_k} if and only if

$$\limsup_{k\to\infty} N[w_k]^{\frac{1}{w_k}} = Q. \tag{B.93}$$

We take the logarithm on both sides and get

$$\log Q = \log\left(\limsup_{k\to\infty} N[w_k]^{\frac{1}{w_k}}\right) \tag{B.94}$$

$$= \limsup_{k\to\infty} \log\left(N[w_k]^{\frac{1}{w_k}}\right) \tag{B.95}$$

$$= \limsup_{k\to\infty} \frac{\log N[w_k]}{w_k} \tag{B.96}$$

$$= C_{\mathrm{comb}} \tag{B.97}$$

where we have equality in (B.95) since the logarithm is a strictly monotonically increasing function. □

Lemma 14. *The operational capacity C_{op} of $\mathcal{A}$ is given by*

$$C_{\mathrm{op}} = \log Q \tag{B.98}$$

where Q^{w_k} is the exponential order of $N[w_k]$.

Proof. We define $I(w_l)$ as

$$I(w_l) = \log\left(\sum_{k\leq l} N[w_k]\right). \tag{B.99}$$

We can directly give the following lower-bound on $I(w_l)$

$$I(w_l) \geq \log\left(\max_{k\leq l} N[w_k]\right) \tag{B.100}$$

which implies for the operational capacity

$$\begin{aligned} C_{\mathrm{op}} &= \lim_{l\to\infty} \frac{I(w_l)}{w_l} && \text{(B.101)}\\ &\geq \lim_{l\to\infty} \frac{\log\left(\max_{k\leq l} N[w_k]\right)}{w_l}. && \text{(B.102)} \end{aligned}$$

For an upper-bound, we have

$$\begin{aligned} I(w_l) &= \log\left(\sum_{k\leq l} N[w_k]\right) && \text{(B.103)}\\ &\leq \log\left(L w_l^K \max_{k\leq l} N[w_k]\right) && \text{(B.104)}\\ &= \log L + K \log w_l + \log\left(\max_{k\leq l} N[w_k]\right) && \text{(B.105)} \end{aligned}$$

where (B.104) follows from assumption (B.87). For the operational capacity we have

$$\begin{aligned} C_{\mathrm{op}} &= \lim_{l\to\infty} \frac{I(w_l)}{w_l} && \text{(B.106)}\\ &\leq \lim_{l\to\infty} \left(\frac{\log L}{w_l} + \frac{K\log w_l}{w_l} + \frac{\log\left(\max_{k\leq l} N[w_k]\right)}{w_l}\right) && \text{(B.107)}\\ &= \lim_{l\to\infty} \frac{\log\left(\max_{k\leq l} N[w_k]\right)}{w_l}. && \text{(B.108)} \end{aligned}$$

Putting (B.102) and (B.108) together yields

$$C_{\text{op}} = \lim_{l\to\infty} \frac{\log\left(\max\limits_{k\le l} N[w_k]\right)}{w_l}. \tag{B.109}$$

For the right-hand side, we further get

$$\lim_{l\to\infty} \frac{\log\left(\max\limits_{k\le l} N[w_k]\right)}{w_l} = \lim_{l\to\infty} \frac{\log M[w_l]}{w_l} \tag{B.110}$$

$$= \lim_{l\to\infty} \log M[w_l]^{\frac{1}{w_l}} \tag{B.111}$$

$$= \log\left(\lim_{l\to\infty} M[w_l]^{\frac{1}{w_l}}\right) \tag{B.112}$$

where we used in (B.110) the substitution $M[w_l] = \max_{w_k \le w_l} N[w_k]$, and where we have equality in (B.112) since the logarithm is strictly monotonically increasing. Because of assumption (B.88), $N[w_k]$ is of exponential order Q^{w_k}, or equivalently: for all ϵ with $Q > \epsilon > 0$, the following holds:

$$N[w_k] \ge (Q-\epsilon)^{w_k}, \qquad \text{i.o.} \tag{B.113}$$

$$\text{and}$$

$$N[w_k] \le (Q+\epsilon)^{w_k}, \qquad \text{a.e..} \tag{B.114}$$

But since $M[w_k] \ge N[w_k]$, for all $k \in \mathbb{N}$, (B.115) also holds for $M[w_k]$:

$$M[w_k] \ge (Q-\epsilon)^{w_k}, \qquad \text{i.o.} \tag{B.115}$$

If inequality (B.114) is fulfilled almost everywhere, then it is also fulfilled for $M[w_k]$ almost everywhere:

$$M[w_k] \le (Q+\epsilon)^{w_k}, \qquad \text{a.e.} \tag{B.116}$$

This can be seen as follows: assume that

$$N[w_m] > (Q+\epsilon)^{w_m} \tag{B.117}$$

holds for some m, and assume further that $M[w_l] = N[w_m]$ for $m \le l \le n$ for some arbitrary large n. Then

$$M[w_l] > (Q+\epsilon)^{w_l} \tag{B.118}$$

is only fulfilled for w_l in some finite interval, since the right side increases exponentially. But according to our assumption in (B.87), the set $\{w_k\}_{k=1}^{\infty}$ is not too dense, which implies that for any finite interval, the number of elements from $\{w_k\}_{k=1}^{\infty}$ that lie in this interval is finite. It follows that $N[w_m] > (W+\epsilon)^{w_m}$ for some m implies $M[w_l] > (Q+\epsilon)^{w_l}$ only for a finite number of l.

It follows from (B.115) and (B.116) together with Definition 7 that $M[w_k]$ is of exponential order Q^{w_k}, i.e.,

$$\lim_{l\to\infty} M[w_l]^{\frac{1}{w_l}} = Q. \tag{B.119}$$

Putting our results together, we have shown that

$$\begin{aligned} C_{\mathrm{op}} &= \lim_{l\to\infty} \frac{\log\left(\max_{k\leq l} N[w_k]\right)}{w_l} && \text{(B.120)} \\ &= \log\left(\lim_{l\to\infty} M[w_l]^{\frac{1}{w_l}}\right) && \text{(B.121)} \\ &= \log Q && \text{(B.122)} \end{aligned}$$

which concludes the proof. □

Bibliography

[1] C. E. Shannon and W. Weaver, *The Mathematical Theory Of Communication.* Urbana and Chicago: University of Illinois Press, 1963.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* John Wiley & Sons, Inc, 1991.

[3] J. L. Massey, *Applied Digital Information Theory.* lecture notes, ETHZ, 1998.

[4] A. Lapidoth, *Information Transfer.* lecture notes, ETHZ, 2005.

[5] H. Bölcskei, *Fundamentals of Wireless Communication.* lecture notes, ETHZ, 2005.

[6] D. N. C. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[7] A. Khandekar, R. McEliece, and E. Rodemich, "The discrete noiseless channel revisited," in *Coding, Communications, and Broadcasting.* Research Studies Press Ltd., 2000, pp. 115–137.

[8] A. M. Odlyzko, "Enumeration of strings," AT&T Bell Laboritories, Murray Hill, New Jersey 07974 USA, Tech. Rep., 1984.

[9] ——, "Asymptotic enumeration methods," AT&T Bell Laboritories, Murray Hill, New Jersey 07974 USA, Tech. Rep., 19??

[10] "Message from the founding editors," Online Journal of Analytic Combinatorics, accessed 26.10.06. [Online]. Available: http://www.ojac.org/message.html

[11] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Oktober 2006, version "Oktoberfest". [Online]. Available: http://algo.inria.fr/flajolet/Publications/books.html

[12] B. H. Marcus, R. M. Roth, and P. H. Siegel, *An Introduction to Coding for Constrained Systems*, 5th ed., October 2001. [Online]. Available: http://vivaldi.ucsd.edu:8080/~psiegel/Book/Book_Draft.html

[13] C. Pimentel and B. F. Uchoa-Filho, "A combinatorial approach to finding the capacity of the discrete noiseless channel," *Information Theory, IEEE Transactions on*, vol. 49, no. 8, pp. 2024–2028, 2003.

[14] L. J. Guibas and A. M. Odlyzko, "String overlaps, pattern matching, and nontransitive games," *Journal of Combinatorial Theory*, vol. A 30, pp. 183–208, 1981.

[15] C. Pimentel, "Generating series and capacity for constrained sequences," *Revista da Sociedade Brasileira de Telecomunicaes*, vol. 12, no. 2, pp. 104–107, 1997.

[16] V. C. da Rocha, "Some information-theoretic aspects of uniquely decodable codes," *Coding, Communications, and Broadcasting*, pp. 39–47, July 2000.

[17] R. Wattenhofer, *Discrete Event Systems*. lecture notes, ETHZ, 2005. [Online]. Available: http://dcg.ethz.ch/lectures/ws0506/eventsystems/index.html

[18] H.-A. Loeliger, *Algebra, Codes, and Signal Processing*. lecture notes, ETHZ, 2005.

[19] A. A. Sardinas and G. W. Patterson, "A necessary and sufficient condition for the unique decomposition of coded messages," *IRE Convention Record, Part 8*, pp. 104–108, 1953.

[20] E. Gilbert, "Synchronization of binary messages," *Information Theory, IEEE Transactions on*, vol. 6, no. 4, pp. 470–477, 1960. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1057587

[21] S. A. Savari, "A probabilistic approach to some asymptotics in noiseless communication," *Information Theory, IEEE Transactions on*, vol. 46, no. 4, pp. 1246–1262, 2000.

[22] C. Blatter, *Komplexe Analysis, Fourier- und Laplace-Transformation fuer Ingenieure*. lecture notes, ETHZ, January 2006. [Online]. Available: http://www.math.ethz.ch/~blatter/

[23] G. F. Franklin, J. D. Powell, and A. Emami-Naeini, *Feedback Control of Dynamic Systems*, 4th ed. Prentice Hall, 2002.

[24] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, 2nd ed. Prentice Hall, 1999.

[25] "Newton's method," Wikipedia, accessed 06.11.06. [Online]. Available: http://en.wikipedia.org/wiki/Newton's_method

[26] "Stirling's approximation," PlanetMath, accessed 12.02.07. [Online]. Available: http://planetmath.org/encyclopedia/StirlingsApproximationFormula.html