Capacity of General Discrete Noiseless Channels

G. Bocherer
Institute for Theoretical Information Technology
RWTH Aachen University
52056 Aachen, Germany
Email: boecherer@ti.rwth-aachen.de

V.C. da Rocha Jr., C. Pimentel
Communications Research Group - CODEC
Department of Electronics and Systems, P.O. Box 7800
Federal University of Pernambuco
50711-970 Recife PE, Brazil
E-mail: {vcr,cecilio}@ufpe.br

***Abstract*—This paper concerns the capacity of the discrete noiseless channel introduced by Shannon. A sufficient condition is given for the capacity to be well-defined. For a general discrete noiseless channel allowing non-integer valued symbol weights, it is shown that the capacity—if well-defined—can be determined from the radius of convergence of its generating function, from the smallest positive pole of its generating function, or from the rightmost real singularity of its complex generating function. A generalisation is given for Pringsheim's Theorem and for the Exponential Growth Formula to generating functions of combinatorial structures with non-integer valued symbol weights.**

I. Introduction

When modelling digital communication systems, there are situations where we do not explicitly model physical noise. We rather introduce constraints on the allowed system configurations that minimise the influence of undesired effects. An example is the runlength-limited constraint in magnetic recording [1]. We consider in this paper the discrete noiseless channel (DNC) as introduced by Shannon [2]. A DNC is specified by a set of constraints imposed on strings over a certain alphabet, and only those strings that fulfil the constraints are allowed for transmission or storage. A DNC allows the specification of two types of constraints. The first constraint is on symbol constellations (for example, only binary strings with not more than two consecutive 0s are allowed), and the second constraint is on symbol weights (for example, the symbol a has to be of duration 5.53 seconds). Depending on the system we want to model, the symbol weights represent the critical resource over which we want to optimise. This can for example be duration, length or energy. We then ask the following question. What is the maximum rate of data per string weight that can be transmitted over a DNC?

This question was first answered by Shannon in [2]. In [3], the authors extend Shannon's results to DNCs with non-integer valued symbol weights. In both [2] and [3], the authors use the following approach to derive the capacity of a DNC. They restrict the class of considered DNCs to those that allow the transmission of a set of strings forming a regular language. The regularity allows then to represent the DNC by a finite state machine and results from matrix theory are applied to derive the capacity of the DNC.

Our approach is different in the following sense. We consider general DNCs with the only restriction that the capacity as defined in [2] and generalised in [3] has to be well-defined, which will turn out to be a restriction on the set of possible string weights.

This allows us then to represent the combinatorial complexity of a DNC by a generating function with a well-defined radius of convergence and we use analytical methods to derive the capacity. In this sense, our work is a generalisation of [3]. Perhaps more important, in many cases that could be treated by the techniques proposed in [3], it is much simpler to construct the generating function of the considered DNC and to use our results to derive the capacity. We give two simple examples that may serve as illustrations. In this sense, our work can also be considered as an interesting alternative to [3].

II. Definitions

We formally define a DNC and its generating function as follows.

Definition 1. A DNC $\mathcal{A} = (A, w)$ consists of a countable set A of strings accepted by the channel and an associated weight function $w\colon A \mapsto \mathbb{R}^{\oplus}$ ($\mathbb{R}^{\oplus}$ denotes the nonnegative real numbers) with the following property. If $s_1, s_2 \in A$ and $s_1 s_2 \in A$ ($s_1 s_2$ denotes the concatenation of s_1 and s_2), then

$$w(s_1 s_2) = w(s_1) + w(s_2). \tag{1}$$

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

Definition 2. Let $\mathcal{A} = (A, w)$ represent a DNC. We define the *generating function* of $\mathcal{A}$ by

$$\mathrm{G}_{\mathcal{A}}(y) = \sum_{s \in A} y^{w(s)} \qquad y \in \mathbb{R}. \tag{2}$$

We order and index the set of possible string weights $w(A)$ such that $w(A) = \{w_k\}_{k=1}^{\infty}$ with $w_1 < w_2 < \cdots$. We can then write

$$\mathrm{G}_{\mathcal{A}}(y) = \sum_{k=1}^{\infty} N[w_k] y^{w_k} \tag{3}$$

where for each $k \in \mathbb{N}$, the coefficient $N[w_k]$ is equal to the number of distinct strings of weight w_k. Since the coefficients $N[w_k]$ result from an enumeration, they are all nonnegative. Note that for any DNC $\mathcal{A}$, we have $\mathrm{G}_{\mathcal{A}}(0) = N[0] = 1$ since every DNC allows the transmission of the empty string and since there is only one empty string.

The maximum rate of data per string weight that can be transmitted over a DNC is given by its capacity. We define capacity in accordance with [2] and [3] as follows.

Definition 3. The *capacity* C of a DNC $\mathcal{A} = (A, w)$ is given by

$$C = \limsup_{k\to\infty} \frac{\ln N[w_k]}{w_k} \tag{4}$$

in nats per symbol weight. This is equivalent to the following. For all ϵ with $C > \epsilon > 0$, the following two properties hold.

1) The number $N[w_k]$ is greater than or equal to $e^{w_k(C-\epsilon)}$ infinitely often (i.o.) with respect to k.
2) The number $N[w_k]$ is less than or equal to $e^{w_k(C+\epsilon)}$ almost everywhere (a.e.) with respect to k.

We assume in the following that the number sequence $\{w_k\}_{k=1}^{\infty}$ is not too dense in the sense that for any integer $n \geq 0$

$$\max_{w_k < n} k \leq L n^K \tag{5}$$

for some constant $L \geq 0$ and some constant $K \geq 0$. Otherwise, the number of possible string weights in the interval $[n, n+1]$ increases exponentially with n. In this case, Definition 3 does not apply. We present in the following example a case where capacity is not well-defined.

Example 1. Let $N[w_k]$ denote the coefficients of the generating function of some DNC. Assume $N[w_k] = 1$ for all $k \in \mathbb{N}$ and assume

$$\max_{w_k < n} k = \lceil R^n \rceil \tag{6}$$

for some $R > 1$. According to Definition 3, the capacity of the DNC is then equal to zero because of $\ln N[w_k] = 0$ 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 $\ln R^n/n = \ln R$, which is according to the assumption greater than zero. ◁

Whenever we say that the capacity of a DNC is well-defined, we mean that (5) is fulfilled.

III. Capacity by Radius of Convergence

One way to calculate the capacity of a DNC is by determining the radius of convergence of its generating function.

Lemma 1. *Let $\mathcal{A}$ be a DNC with the generating function $\mathrm{G}_{\mathcal{A}}(y)$. If the capacity C of $\mathcal{A}$ is well-defined, then it is given by $C = -\ln R$ where R denotes the radius of convergence of $\mathrm{G}_{\mathcal{A}}(y)$.*

In the proof of this lemma, we will need the following result from [3].

Lemma 2. *If* (5) *is fulfilled and if ρ is a positive real number, then $\sum_{k=1}^{\infty} \rho^{w_k}$ converges iff $\rho < 1$.*

Proof of Lemma 1: We define $M[k] = N^{1/w_k}[w_k]$ and write $\mathrm{G}_{\mathcal{A}}(y)$ as

$$\mathrm{G}_{\mathcal{A}}(y) = \sum_{k=1}^{\infty} \bigl(M[k]y\bigr)^{w_k}. \tag{7}$$

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

$$D(y) = \bigl\{k \in \mathbb{N} \big| M[k]y < 1\bigr\} \tag{8}$$

$$E(y) = \mathbb{N} \setminus D(y) = \bigl\{k \in \mathbb{N} \big| M[k]y \geq 1\bigr\} \tag{9}$$

and write

$$\mathrm{G}_{\mathcal{A}}(y) = \sum_{k \in D(y)} \bigl(M[k]y\bigr)^{w_k} + \sum_{l \in E(y)} \bigl(M[l]y\bigr)^{w_l}. \tag{10}$$

It follows from Lemma 2 that $\mathrm{G}_{\mathcal{A}}(y)$ converges iff the set $E(y)$ is finite. The number R is the radius of convergence of $\mathrm{G}_{\mathcal{A}}(y)$, therefore, for any $\delta > 1$, the set $E(R/\delta)$ is finite. Since $D(y) = \mathbb{N} \setminus E(y)$, the finiteness of $E(R/\delta)$ is equivalent to

$$k \in D(R/\delta) \qquad \text{a.e.} \tag{11}$$

We define $\epsilon = \ln \delta$. Equation (11) is then equivalent to

$$N[w_k] < e^{w_k(-\ln R + \epsilon)} \qquad \text{a.e.} \tag{12}$$

which implies

$$N[w_k] \leq e^{w_k(-\ln R + \epsilon)} \qquad \text{a.e.} \tag{13}$$

Again since R is the radius of convergence of $\mathrm{G}_{\mathcal{A}}(y)$, for any $\delta > 1$, the set $E(R\delta)$ is infinite. For $\epsilon = \ln \delta$, this is equivalent to

$$N[w_k] \geq e^{w_k(-\ln R - \epsilon)} \qquad \text{i.o.} \tag{14}$$

It follows from (13) and (14) and Definition 3 that $-\ln R$ is equal to the capacity of $\mathcal{A}$. We therefore have $C = -\ln R$. ■

In the following example, we show how Lemma 1 applies in practice. We denote by $A \cup B$ the union of the two sets A and B, we denote by AB the set of all concatenations ab with $a \in A$ and $b \in B$, and we denote by $S^\star$ the Kleene star operation on S, which is defined as $S^\star = \epsilon \cup S \cup SS \cup \cdots$.

Example 2. We consider a DNC $\mathcal{A} = (A, w)$ with the alphabet $\{0, 1\}$ and symbol weights $w(0) = 1$ and $w(1) = \pi$. The DNC $\mathcal{A}$ does not allow strings that contain two or more consecutive 1s. We represent A by a regular expression and write $A = \{\varepsilon \cup 1\}\{0 \cup 01\}^\star$. For the generating function of $\mathcal{A}$ we get

$$\mathrm{G}_{\mathcal{A}}(y) = (1 + y^\pi) \sum_{n=0}^{\infty} (y + y^{1+\pi})^n. \tag{15}$$

The radius of convergence is given by the smallest positive solution R of the equation $y + y^{1+\pi} = 1$. We find $R = 0.72937$. According to Lemma 1, the capacity of $\mathcal{A}$ is thus given by $C = -\ln R = 0.31558$. ◁

IV. Capacity by Rightmost Real Singularity

There are cases where we derive the closed-form representation of the generating function of a DNC without explicitly using its series representation. The techniques introduced in [4] and [5] may serve as two examples. In this section, we show how the capacity of a DNC $\mathcal{A}$ can be determined from the closed-form representation of its generating function. We do this in two steps. We first identify the region of convergence (r.o.c.) of the complex generating function $\mathrm{F}_{\mathcal{A}}(e^{-s})$ with its rightmost real singularity. The complex generating function $\mathrm{F}_{\mathcal{A}}(e^{-s})$ results from evaluating the generating function $\mathrm{G}_{\mathcal{A}}(y)$ in $y = e^{-s}$, $s \in \mathbb{C}$. Second, we show that the rightmost real singularity of $\mathrm{F}_{\mathcal{A}}(e^{-s})$ determines the capacity of $\mathcal{A}$.

Theorem 1. *If the r.o.c. of* $\mathrm{F}_{\mathcal{A}}(e^{-s})$ *is determined by* $\Re\{s\} > Q$*, then* $\mathrm{F}_{\mathcal{A}}(e^{-s})$ *has a singularity in* $s = Q$*.*

Proof: Suppose in contrary that $\mathrm{F}_{\mathcal{A}}(e^{-s})$ is analytic in $s = Q$ implying that it is analytic in a disc of radius r centred at Q. We choose a number h such that $0 < h < r/3$, and we consider the Taylor expansion of $\mathrm{F}_{\mathcal{A}}(e^{-s})$ around $s_0 = Q + h$ as follows.

$$\mathrm{F}_{\mathcal{A}}(e^{-s}) = \sum_{n=0}^{\infty} \frac{\left[\mathrm{F}_{\mathcal{A}}(e^{-s_0})\right]^{(n)}}{n!}(s - s_0)^n \tag{16}$$

$$= \sum_{n=0}^{\infty} \frac{\sum_{k=1}^{\infty} N[w_k](-w_k)^n e^{-w_k s_0}}{n!}(s - s_0)^n. \tag{17}$$

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

$$\mathrm{F}_{\mathcal{A}}(e^{-Q+h}) = \sum_{n=0}^{\infty} \frac{\sum_{k=1}^{\infty} N[w_k](-w_k)^n e^{-w_k s_0}}{n!}(-2h)^n \tag{18}$$

$$= \sum_{k=1}^{\infty} N[w_k] e^{-w_k s_0} \sum_{n=0}^{\infty} \frac{w_k^n (2h)^n}{n!} \tag{19}$$

$$= \sum_{k=1}^{\infty} N[w_k] e^{-w_k s_0} e^{w_k 2h} \tag{20}$$

$$= \sum_{k=1}^{\infty} N[w_k] e^{-w_k (Q-h)}. \tag{21}$$

But convergence in the last line contradicts that the r.o.c. of $\mathrm{F}_{\mathcal{A}}(e^{-s})$ is strictly given by $\Re\{s\} > Q$. ■

We now relate the rightmost real singularity of $\mathrm{F}_{\mathcal{A}}(e^{-s})$ to the capacity of $\mathcal{A}$.

Theorem 2. *Assume that* $\mathrm{F}_{\mathcal{A}}(e^{-s})$ *has its rightmost real singularity in* $s = Q$*. The capacity of* $\mathcal{A}$ *is then given by* $C = Q$*.*

Proof: Since $\mathrm{F}_{\mathcal{A}}(e^{-s})$ has its rightmost real singularity in $s = Q$, it follows from Theorem 1 that the r.o.c. of $\mathrm{F}_{\mathcal{A}}(e^{-s})$ is determined by $\Re\{s\} > Q$. For $\mathrm{F}_{\mathcal{A}}(e^{-s})$, we have

$$\mathrm{F}_{\mathcal{A}}(e^{-s}) = \sum_{k=1}^{\infty} N[w_k] e^{-w_k s} \tag{22}$$

$$\leq \sum_{k=1}^{\infty} |N[w_k] e^{-w_k s}| \tag{23}$$

$$= \sum_{k=1}^{\infty} N[w_k] |e^{-w_k s}| \tag{24}$$

where equality in (24) holds because the coefficients $N[w_k]$ are all nonnegative and where we have equality in (23) if s is real. It follows that if the r.o.c. of $\mathrm{F}_{\mathcal{A}}(e^{-s})$ is given by $\Re\{s\} > Q$, then the radius of convergence of $\mathrm{G}_{\mathcal{A}}(y)$ is given by $R = e^{-Q}$. Using Lemma 1, we have for the capacity $C = -\ln R = Q$. ■

Note 1. With Theorem 1 and Theorem 2, we generalised Pringsheim's Theorem and the Exponential Growth Formula, see [6], to generating functions of DNCs with non-integer valued symbol weights.

V. Capacity by Smallest Positive Pole

We formulate the most important application of Theorem 2 in the following corollary:

Corollary 1. *Let* $\mathcal{A}$ *represent a DNC with a well-defined capacity* C*. Suppose that the generating function* $\mathrm{G}_{\mathcal{A}}(y)$ *can be written as*

$$\mathrm{G}_{\mathcal{A}}(y) = \frac{n_1 y^{\tau_2} + n_2 y^{\tau_2} + \cdots + n_p y_p^{\tau}}{d_1 y^{\nu_1} + d_2 y^{\nu_2} + \cdots + d_q y^{\nu_q}} \tag{25}$$

for some finite positive integers p *and* q*. The capacity* C *is then given by* $-\ln P$ *where* P *is the smallest positive pole of* $\mathrm{G}_{\mathcal{A}}(y)$*.*

Note 2. The corollary was already stated in [4, Theorem 1]. However, the proof given by the authors does not apply for the general case, which we consider in this paper.

Proof of Corollary 1: If $\mathrm{G}_{\mathcal{A}}(y)$ is of the form (25), the complex generating function $\mathrm{F}_{\mathcal{A}}(e^{-s})$ as defined in the previous section 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 Q is the rightmost real singularity of $\mathrm{F}_{\mathcal{A}}(e^{-s})$, then e^{-Q} is the smallest positive pole of $\mathrm{G}_{\mathcal{A}}(y)$. Applying Theorem 2, we get for the capacity $C = Q = -\ln P$. ■

Example 3. We consider the DNC $\mathcal{A} = (A, w)$ where A is the set of all binary strings that do not contain the substring 111 and where the symbol weights are given by $w(0) = w(1)$. We use a result from [5] in the form of [6, Proposition 1.4]. It states that the set of binary strings that do not contain a certain pattern p has the generating function

$$f(y) = \frac{c(y)}{y^k + (1 - 2y)c(y)} \tag{26}$$

where k is the length (in bits) of p and where $c(y)$ is the autocorrelation polynomial of p. It is defined as $c(y) = \sum_{i=0}^{k-1} c_i y^i$ with c_i given by

$$c_i = \delta[p_{1+i}p_{2+i}\cdots p_k, p_1 p_2 \cdots p_{k-i}] \tag{27}$$

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 $a \neq b$. For $p = 111$, we have $c(y) = 1 + y + y^2$ and $k = 3$. This yields for the generating function of $\mathcal{A}$

$$\mathrm{G}_{\mathcal{A}}(y) = \frac{1 + y + y^2}{y^3 + (1-2y)(1 + y + y^2)}. \tag{28}$$

Note that the application of the technique from [4] would have led to the same formula. For the smallest positive pole P of $\mathrm{G}_{\mathcal{A}}(y)$ we find $P = 0.54369$. According to Corollary 1, the capacity of $\mathcal{A}$ is thus given by $C = -\ln P = 0.60938$. $\triangleleft$

VI. Conclusions

For a general DNC, we identified the capacity with the characteristics of its generating function, namely the radius of convergence of its generating function, the rightmost real singularity of its complex generating function, and the smallest positive pole of its generating function. We generalised Pringsheim's Theorem and the Exponential Growth Formula as given in [6] to generating functions that allow non-integer valued symbol weights.

Representing a DNC by its generating function and not by a finite state machine has an additional advantage. Although the finite state machine allows the derivation of the correct capacity of the DNC, it says nothing about the exact number of valid strings of weight w. The generating function of a DNC provides this information. The coefficients $N[w_k]$ are equal to the number of distinct strings of length w_k that are accepted by the DNC. The coefficients can either be calculated by an algebraic expansion of the generating function or they can be approximated by means of analytic asymptotics as discussed for integer valued symbol weights in [6]. In [7], the analytic approach is extended to generating functions of DNCs with non-integer valued symbol weights.

For a regular DNC fulfilling some further restrictions, the authors in [3] define a Markov process that generates valid strings at an entropy rate equal to the capacity of the channel. Based on generating functions as introduced in this paper, it is shown in [7] that for a general DNC, any entropy rate C' smaller than the capacity C is achievable in the sense that there exists a random process that generates strings that are transmitted over the channel at an entropy rate C'.

Acknowledgement

V. C. da Rocha Jr. and C. Pimentel acknowledge partial support of this research by the Brazilian National Council for Scientific and Technological Development (CNPq) under Grants No. 305226/2003-7 and 301253/2004-8, respectively.

References

[1] B. H. Marcus, R. M. Roth, and P. H. Siegel, "An introduction to coding for constrained systems," Oct. 2001.

[2] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423 and 623–656, Jul. and Oct. 1948.

[3] 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.

[4] C. Pimentel and B. Uchoa-Filho, "A combinatorial approach to finding the capacity of the discrete noiseless channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 2024–2028, 2003.

[5] 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.

[6] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*. Cambridge University Press, 2008.

[7] G. Bocherer, "Analytic asymptotics of discrete noiseless channels," Master's thesis, ETH Zurich, 2007. [Online]. Available: http://www.georg-boecherer.de/repository/analyticAsymptotics.pdf