Graz University of Technology

Master Thesis

Alternative Descriptions for Random Variables

conducted at the
Signal Processing and Speech Communications Laboratory
Graz University of Technology, Austria

by
Christian Knoll

Supervisors:
Dipl.-Ing. Bernhard Geiger

Assessors/Examiners:
Univ.-Prof. Dipl.-Ing. Dr.techn. Gernot Kubin

Graz, December 10, 2013

Abstract:

This work introduces and investigates an alternative way of describing random variables. Instead of looking at the cumulants, that are the coefficients of the Taylor expansion of the logarithm of the characteristic function we estimate the coefficients of the corresponding Fourier expansion. Such coefficients are similar to coefficients introduced in cepstrum analysis.

Discrete random variables have periodic characteristic functions, therefore the Fourier expansion lends itself for approximating these functions. As opposed to the Taylor expansion, the Fourier series is suited for periodic functions.

We expand the idea of the existing cepstrum theory and apply it for describing random variables. Some basic and advanced properties are presented for this special application of the cepstrum. Subsequently, we study the connection between cumulants and the cepstrum.

In various experiments the cepstrum is applied to certain distributions and the resulting advantages are discussed. Finally we show the benefits of using cepstral coefficients for hypothesis testing.

This thesis investigates the idea of bringing probability theory and cepstrum analysis together, and gives an overview of the corresponding possibilities. It should serve as a foundation for further research.

Kurzfassung:

Diese Masterarbeit beschäftigt sich mit einer alternativen Möglichkeit Zufallsvariablen zu beschreiben. Bildet man die Taylorreihe des Logarithmus der characteristischen Funktion und betrachtet die entsprechenden Koeffizienten so erhält man die Kumlanten. In dieser Arbeit wird stattdesen die entsprechende Fourierreihenentwicklung betrachtet. Die so bestimmten Koeffizienten entsprechen zufälligerweise jenen des Cepstrums.

Diskrete Zufallsvariablen haben eine periodisch fortgesetzte characteristische Funktion. Da die Fourierreihe eine gute Näherung für periodische Funktionen darstellt, ist es naheliegend diese zu verwenden, um diskrete Zufallsvariablen zu beschreiben.

Die bestehenden Theorien des Cepstrums werden eingehend betrachtet, entsprechend wird das Cepstrum angewandt um diskrete Zufallsvariablen zu beschreiben. Einige Eigenschaften für diese spezielle Anwendung werden präsentiert und angewandt. Weiters wird der direkte Zusammenhang zwischen den Kumulanten und den Cepstralkoeffizienten untersucht.

Das Cepstrum wird in einigen Experimenten auf verschiedene Zufallsvariablen angewandt. Die Resultate sowie die entsprechende Vorteile werden untersucht und präsentiert. Ein weiterer Teil der Arbeit beschäftigt sich mit der möglichen Anwendung dieser Cepstralkoeffizienten um statistische Hypothesen zu testen. Anhand von praktischen Beispielen wird diese Möglichkeit mit gängigen Methoden verglichen.

In dieser Masterarbeit werden die Konzepte der Cepstralanalyse und der Wahrscheinlichkeitstheorie vereint und ein Überblick über die entsprechenden Anwendungsmöglichkeiten gegeben. Diese Aufbereitung des Themas soll als Grundlage für weitere Untersuchungen dienen.

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

date

(signature)

Danksagung:

Hiermit möchte ich sämtlichen Personen, die mich im Zuge dieser Masterarbeit unterstützten, meinen Dank aussprechen.

Zunächst möchte ich Gernot Kubin für die grundlegende Idee zu dieser Arbeit und sämtliche Unterstützung danken. Weiters geht mein Dank an Bernhard Geiger, der mir immer mit Rat und Tat zur Seite stand und sehr viel Engagement in die Betreuung dieses Themas steckte. Diese Unterstützung und die angeregten Diskussionen waren äußerst hilfreich für meine Arbeit. An dieser Stelle ist noch Franz Lehner zu erwähnen. Seine Bemühungen, etwaige Probleme von einem mathematischem Standpunkt aus zu betrachten, waren immer wieder eine Bereicherung.

Weiters möchte ich meiner Familie und vor allem meine Eltern danken, die mich in meinen Entscheidungen stets unterstützten und mir die Möeglichkeit gaben mich mit diesem Thema in aller Zeit und Tiefe auseinanderzusetzen. Insbesondere Ihnen verdanke ich meine schöne und interessante Studienzeit.

Zuletzt gehört meiner Freundin Katrin ein großer Dank dafür, dass sie mich in Phasen des Zweifelns immer wieder aufmunterte und mir soweit wie möglich den Rücken freihielt um mich auf meine Masterarbeit konzentrieren zu können.

Graz, Dezember 2013 Christian Knoll

Contents

1 Introduction

1.1 Motivation

Looking for a way to describe a random variable and the shape of the distribution, the moments and cumulants are widely used. These are obtained by evaluating the terms of the Taylor expansion of the characteristic function or the logarithm of the characteristic function, respectively. For a characteristic function which is k times differentiable the terms of the Taylor expansion up to order k exist. However looking at discrete random variables, the characteristic function is periodic, therefore a Taylor expansion might be problematic.

The moments and cumulants offer a good intuitive description of the shape of the distribution. Nonetheless approximating the random variable just in terms of these descriptors, we need infinitely many terms of the Taylor expansion. Since convergence is not guaranteed, this may lead to substantial problems.

To circumvent these issues a new way of describing random variables is presented. The continuity of the characteristic function for discrete random variables points at applying the Fourier expansion instead. Applying such a transform, coefficients are estimated which are similar to the one introduced in cepstrum analysis. The cepstrum is a widely used transformation throughout the field of signal processing.

This work shall expand the according concepts and investigate the existing cepstrum theory to get some insights into potential benefits of describing a random variable via its cepstral coefficients instead of its cumulants. Different definitions of the cepstrum are presented and compared. After presenting the theoretical framework, the cepstrum analysis is applied to different types of random variables. In these experiments some benefits and disadvantages of an approach using the Fourier series instead of the Taylor series will appear.

The convergence behavior of the cepstral coefficients proves to be expedient. Therefore truncating the series seems to be less of a deal, as opposed to truncating a cumulant series. Some general constraints for the behavior of the cepstrum, applied to describe random variables, are derived and presented. Further properties are given as well to ease the way for further applications in this field. Some random variables which have a particularly nice cepstrum lend themselves to introduce new hypothesis tests, which are compared to established types.

This work aims to give an overview of using the cepstrum for describing random variables. Some possible applications are marked. The presented investigations shall give a feeling of how we can bring these two different fields of research together.

1.2 Related Work

To understand and expand the given concepts for describing random variables some classic literature of probability theory was used. Therefore [PP02] served as a main source of information, [Chu01] gives a more detailed treatment on some subjects.

As a thorough understanding of the existing theory was important [Lin57] and [Rai37] shall be mentioned to understand the decomposition of Gaussian and Poisson distributions. [Pet75] deals with sums of independent random variables in general. However we shall especially mention the literature with regard to the characteristic function. [Luk70], [LL64] and [Luk72] give a detailed treatment of this field and explain the behavior of the characteristic functions. Some important constraints and conditions are presented therein. Further work which pays special attention to the cumulants can be found in [Mat99] and [RS00].

In [BM98] the approximation of random variables is presented. Some details with regard to convergence are listed and it serves as a good overview of the whole field of research. The connection between discrete and continuous random variables and the effect of quantization is described in [WKL96].

Some basic mathematic work which was of great help throughout the work is [Rud76] with regard to analysis, as was [BC09] to give a good understanding of complex variables.

The whole concept of the cepstrum and homomorphic transformations can be understood by looking at [BHT63], [Opp65] and [OS10]. Note that [OS04] is a good paper on the history of the cepstrum. Many possible applications are given and we shall therefore mention the references therein.

The work of [KY92] and [Sch81] were of great importance. These papers hint at some relationship between the cepstrum and the cumulants, respectivly the moments.

1.3 Structure of the Work

In **chapter 2** the different concepts for describing random variables are presented. Section 2.1 gives a short definition of the random variable itself to serve as a consistent foundation for the whole work. The characteristic function for discrete and continuous random variables is given in section 2.2, subsequently some important properties for a function to be the characteristic function of a valid distribution are listed.

The definition of the cumulants and the moments is shortly recaped in section 2.3, whereas the subsequent section deals with the approximation of random variables. Different types of approximations applying Hermite polynomials are presented and compared.

Chapter 3 deals with the cepstrum in a more detailed way. We will present concept of the cepstrum and some applications. Subsequently in section 3.3 the various definitions of the cepstrum are presented. These definitions are explained and connected to each other.

The idea of the cepstrum is expanded for describing random variables in **chapter 4**. To start with, the cepstral coefficients are linked to the cumulants in an analytic way in section 4.1. In the subsequent sections some basic constraints and existence conditions are given. Section 4.3 presents how distributions can be approximated by looking at the cepstrum. The knowledge we obtained so far is applied to some random variables in section 4.4. The subsequent sections investigate the cepstral coefficients for five different distributions.

In **chapter 5** we derive some properties for the cepstrum with regard to the application on random variables. Section 5.1 presents some basic symmetry properties, whereas the subsequent sections observe some basic properties for the complex cepstrum. Section 5.6 gives constraints which have to be fulfilled in order to be the cepstrum of a valid distribution. We present and investigate the resulting characteristic function in terms of its cepstral coefficients in section 5.7 and 5.8. The influence of scaling the cepstrum is presented in section 5.9.

In section 5.10 we deal with right hand sided PMFs. It is shown how a one sided distribution constrains the complex cepstrum. Afterwards we investigate if it is possible to decompose a random variable into contributions which are one sided.

Finally we apply the cepstrum analysis in **chapter 6** to perform a hypothesis test. By the means of the Poisson distribution and the geometric distribution it is shown how we can validate the hypothesis just by looking at the cepstral coefficients.

2 Characterizing Probability Distributions

This chapter aims to give a short overview of probability theory and some underlying concepts, which are applied throughout this work. In section 2.1 the concept of a random variable is presented. We introduce and present some important functions of random variables. Section 2.2 starts with defining univariate characteristic functions and explains this concept in some detail due to its importance in this work. Section 2.3 expands these ideas and gives introduction to moments and cumulants. Some problems using these descriptors are pointed out. Section 2.4 presents some expansions applying Hermite polynomials for approximating a given distribution and the corresponding applications.

2.1 The Random Variable

We start with a given sample space $\mathcal{X}$. On this set we shall define a random variable $\mathbf{X}$. Note that we will denote all random variables with bold upper case letters throughout the work. Suppose $\mathbf{X}$ to be a continuous random variable, the according probability density function (PDF) is denoted as $f_x(\xi)$ where ξ is an element of $\mathcal{X}$.

If we are looking at a discrete random variable the according probability mass function (PMF) is $f_x(\xi)$ where $\xi \in \mathcal{X}$. A PMF is said to be a lattice distribution, if ξ can be written as $a + h \cdot n$, with $h, a \in \mathbb{R},\ n \in \mathbb{Z}$ and h being the step size. That is, the support is a set of equidistant points.

We will restrict every PMF to be a lattice distribution for the remainder of this work. Furthermore we will only deal with real random variables, since every complex random variable $\mathbf{Z}$ can be written with the help of two real random variables $\mathbf{Z} = \mathbf{X} + j\mathbf{Y}$. However, the PMF of the sum of two random variables $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$ is given by the convolution of the according PMFs:

$$f_z(\xi) = (f_x * f_y)(\xi). \tag{2.1}$$

For a more detailed treatment of random variables and probability theory in general we refer to [PP02] or [Chu01].

2.2 Characteristic Functions

Since the characteristic function plays an important role in this work we will give a short overview of the motivation creating a characteristic function and some basic properties which are used later on. The original purpose for introducing the characteristic function was to investigate the behavior of limiting distributions. However since it offers the possibility to deal with various problems of probability theory in an analytical way, the characteristic function is widely used.

Eugene Lukacs was a statistician whose main interest was investigating the characteristic function and one of the first to present a thorough work on this subject and its applications. Therefore we shall reference to his classic monograph [Luk70] . However [Luk72] presents many interesting results in a condensed way without giving every proof in full detail.

According to [PP02] the characteristic function of a random variable is defined by the Fourier transform of the PDF, respectively the PMF, and can be written as follows

$$\Phi_x(\mu) = E\left\{e^{j\mu X}\right\}, \tag{2.2}$$

were $E\{\}$ denotes the expectation value. In the discrete case the characteristic function can be given as

$$\Phi_x(\mu) = \sum_{n=-\infty}^{\infty} f_x(\xi)e^{j\mu n}. \tag{2.3}$$

Since there is no proper name for it, we introduce the formal frequency parameter μ in equation (2.2)

[WKL96] deals with the statistical theory of quantization and shows how quantization corresponds to sampling of the PDF, that is assigning a corresponding PMF. From Fourier transform properties we know that sampling a function in the time domain corresponds to periodic repetition of the function in the frequency domain. Similarly a discrete random variable is described by a periodic characteristic function. Figure 2.1 recaps and presents the connection between continuous and discrete random variables.

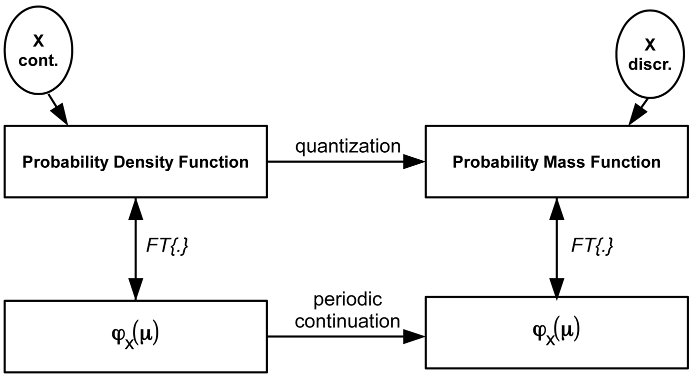

Figure 2.1: correspondence between discrete- and continuous-type random variables and their characteristic functions

Before presenting some important properties of the characteristic function we may prove the existence of the characteristic function. It can be shown that the characteristic function exists for any arbitrary random variable $\mathbf{X}$ since equation (2.2) can be rewritten, applying Euler's formula, as follows.

$$\Phi_x(\mu) = E\{cos(\mu\mathbf{X})\} + iE\{sin(\mu\mathbf{X})\} \tag{2.4}$$

Since we restrict $\mathbf{X}$ to be real and μ is always real we can state that both, the sine and the cosine, are bounded functions. As the expectation of a bounded random variable is always finite, we can conclude that the characteristic function always exists.

Subsequently some basic properties of the characteristic function are noted.

1. $\Phi_x(\mu)$ has to be uniformly continuous everywhere.
2. $\Phi_x(\mu)$ has to be 1 for $\mu = 0$.
   Note that for any discrete random variable $\Phi_x(\mu)$ is periodic and $M \neq 0$ exists such that $|\Phi_x(n \cdot M)| = 1$ for $n \in \mathbb{Z}$.
3. $\Phi_x(\mu)$ is bounded. On its entire domain $|\Phi_x(\mu)| \leq 1$ and therefore $|\Phi_x(\mu)| \leq |\Phi_x(0)|$.
4. $\Phi_x(\mu)$ is Hermitian.
5. There is a direct correspondence between $\Phi_x(\mu)$ and the distribution of $\mathbf{X}$. Two random variables are identically distributed if and only if their characteristic functions are equal.

Finally it shall be noted that looking at the sum of two random variables $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$, the convolution of the distribution functions given in equation (2.1) reduces to the multiplication

$$\Phi_z(\mu) = \Phi_x(\mu) \cdot \Phi_y(\mu), \tag{2.5}$$

due to Fourier transform properties.

2.3 Cumulants

Using a Taylor approximation of the characteristic function directly leads to the moments. Assuming that the derivative exists, the relationship between the n-th order derivative of the characteristic function and the n-th moment m_n is given by evaluating the corresponding derivative at the origin [PP02, page 154]:

$$m_n \cdot j^n = \left. \frac{d^n(\Phi_x(\mu))}{d\mu^n} \right|_{\mu=0}. \tag{2.6}$$

However there are some drawbacks of describing random variables using moments. Moments are no orthogonal descriptors, that is, higher order moments are dependent on lower order moments. Moments of the sum of two i.i.d. random variables are not additive in general.

These drawbacks can be circumvented by introducing some modified parameters. Such descriptors were first defined by T.N. Thiele, and named semi-invariants; the history regarding cumulants can be found in [Mat99].

Some properties of random variables can be better expressed with cumulants as opposed to the moments: The main advantages result, from applying the logarithm to the characteristic function before evaluating the derivatives. The n-th order cumulant κ_n is found by looking at the Taylor expansion again. Hence,

$$\kappa_n(X) \cdot j^n = \left. \frac{d^n \log(\Phi_x(\mu))}{d\mu^n} \right|_{\mu=0}. \tag{2.7}$$

Listed below are some important properties of cumulants.

1. Cumulants are orthogonal descriptors.

2. Cumulants have a homomorphism property. That is, the logarithm turns the multiplication of equation (2.5) into an addition. Hence the cumulants of the sum of two i.i.d random variables are the sum of the individual cumulants.

3. All but the first order cumulant are invariant to a shift of the random variable.

However there are some drawbacks of using a Taylor expansion for approximating the logarithm of the characteristic function. Taylor's theorem [Rud76] states that any n-times differentiable function may be approximated by a polynomial of degree $n-1$. Therefore for any characteristic function of order n, there exist cumulants up to κ_n . Thus it is necessary to estimate n cumulants to ensure a correct representation of the according random variable. Moreover it is not possible to make any assumptions regarding the behavior of such a cumulant series as there are no restrictions known, moreover the may even diverge.

For example the cumulants of the standard uniform distribution are

$$\kappa_n = \frac{B_n}{n}, \tag{2.8}$$

where B_n is the n-th order Bernoulli number. As a matter of fact the numerator in equation (2.8) is increasing faster than the denominator. Hence the cumulants for the standard uniform distribution diverge.

The Marcinkiewicz theorem deals with the existence of cumulants. Although [BH85] generalizes this theorem to the non commutative case, the classical theorem is presented in a neat way. The classical Marcinkiewicz theorem is presented in [Luk70], and states that any function, which is the exponential of a polynomial and is positive definite on $\mathbb{R}$, has a polynomial which is of degree 2 at most. However since the cumulants are estimated from the derivatives of $\log(\Phi_x(\mu))$ this statement can be rewritten as follows.

Theorem 2.3.1. *(Marcinkiewicz theorem) If the logarithm of the characteristic function is a polynomial of finite order, then it is a polynomial of order 2 at most.*

Consequently if a random variable $\boldsymbol{X}$ has cumulants κ_n, which vanish for every $n > N$, then N is 2. That is, $\boldsymbol{X}$ may either have infinitely many cumulants, be degenerated, or have cumulants up to order 2, hence being normally distributed.

It shall be mentioned that the development of the Taylor expansion in the origin is suitable for approximating the characteristic function of a continuous random variable but may be not an optimal way for approximating a periodic function.

2.4 Approximation of Random Variables

In this section some expansions, which are commonly used to reconstruct a distribution function f_x in terms of its cumulants, are investigated. To estimate the unknown distribution some sort of orthogonal polynomials, in fact Hermite polynomials, are used. Note that such polynomials may also be known as Chebyshev or Chebyshev-Hermite polynomials. For the sake of convenience we will call them Hermite polynomials from now on.

[BM98] applies various such expansions for estimating nearly Gaussian distributions to solve statistical problems of astrophysics and present the theory behind these methods. For more details the references therein shall be mentioned.

We will present and investigate the Edgeworth expansion, Gram-Charlier-A (or just Gram Charlier) - series, and Gauss Hermite series which apply Hermite polynomials in slightly different ways.

To start with, equation (2.9) and (2.10) give two different sets of Hermite polynomials. $H_n(\xi)$ denotes the Hermite polynomial of order n.

$$He_n(\xi) = (-1)^n \cdot e^{\xi^2/2} \frac{d^n}{d\xi^n} e^{-\xi^2/2} \tag{2.9}$$

$$H_n(\xi) = (-1)^n \cdot e^{\xi^2} \frac{d^n}{d\xi^n} e^{-\xi^2} \tag{2.10}$$

Obviously both polynomials are different, however the one form can easily be obtained by the other one, since

$$H_n(\xi) = 2^{n/2} \cdot He_n(\sqrt{2}\xi), \tag{2.11}$$

Hermite polynomials given in equation (2.10) are applied in the **Gauss Hermite expansion**. This expansion has particularly nice convergence behaviour as opposed to some other approximation methods. However there is no direct relationship between cumulants or moments and the coefficients in the expansion. This makes it especially hard to deal with the Gauss Hermite expansion in an analytical way.

Such direct relationship to cumulants is obviously aspired. The **Gram Charlier series** applies the Hermite series given in equation (2.9), which work well for Gaussian distributions since the PDF of a unit Gaussian random variable is given by $f_x(\xi) = \frac{1}{2\pi} e^{-\xi^2/2}$. This series is basically approximating a distribution function by evaluating some coefficients. These coefficients measure the deviation of the observed random variable from a given distribution by comparing their cumulants. The given distribution is usually a Gaussian distribution, hence the approximation apparently suits nearly Gaussian distributions well.

One major drawback is, that the Gram Charlier series tends to diverge for many practical applications, due to its sensitivity with regard to the behavior of $f_x(\xi)$ for $\xi \to \infty$. In fact it is shown in [Cra57] that the observed distribution function has to decline faster than $e^{-x^2/4}$ for the series to converge. Besides, this series is no true asymptotic expansion, that is, there is no way to estimate the error.

The **Edgeworth expansion** collects the deviation of the cumulants in a similar way as the Gram Charlier expansion but arranges the terms in a way to ensure a true asymptotic expansion. Hence both series have different behavior with regard to truncating the series. While the standard references like [Cra57] do not take higher order cumulants into account, [Pet75] presents an Edgeworth expansion for arbitrarily high order. It shall be mentioned that the convergence behavior of this series shows limitations to the application of non Gaussian distributions. In fact the Edgeworth expansion diverges if some higher order terms are included as well. However since it is possible to obtain the error contribution of each term, it is possible to truncate the expansion when the error increases.

To conclude one can state that approximations of random variables which are based on cumulants or moments do not converge in general for arbitrary distributions. Hence we state that describing random variables using cumulants, and moments respectively, is only good if the according random variable has nearly Gaussian shape.

According to theorem 2.3.1 every random variable which is not normally distributed has infinitely many cumulants. Clearly every approximation which is based on cumulants can practically only take finitely many coefficients into account. Hence the reconstruction always results in a degenerated random variable for non Gaussian distribution.

3 Cepstrum

In this chapter the cepstrum analysis is introduced. We start with presenting the main idea and underlying methods in section 3.1. Section 3.2 points out some fields where cepstrum analysis is applied and gives a short examination of the current developments. To conclude, section 3.3 lists the different types of cepstrum and gives their according definitions.

3.1 Basic Idea of the Cepstrum

The definition of the cepstrum and applications are given in [OS10], which served as the main source for applying the previous work on this field and expanding its ideas.

The cepstrum was introduced and first published in [BHT63]. The main motivation was to investigate the influence of an echo in a signal and how the echo delay can be determined by looking at the cepstrum. This work opted for applying methods commonly used in the time domain, to the frequency domain. Hence the spelling of the "spectrum" was rephrased and created the word cepstrum. [OS04] wraps up the development and the history of the cepstrum and gives an overview of the whole glossary which was invented consequently.

Around the same period Oppenheim was working on his doctoral research. His dissertation [Opp65] developed a theory for representing nonlinear systems in a way to ensure the principle of superposition. Therefore he introduced homomorphic filtering which maps signals combined through convolution into a signal space where the contribution of these signals is additive. It is shown that such systems always consist of a cascade of exactly three systems. That is, mapping convolution into addition, through applying the first operator. Secondly applying a linear system, and finally mapping addition back to convolution through the inverse system of the first one.

His doctoral research aimed for applying such methods to deconvolution. It proved that this work was coincidently about the same ideas as in [BHT63].

3.2 Applications of Cepstrum Analysis

The motivational example shall be presented to give a rough idea of the basic principles behind the cepstrum analysis. As already mentioned the question was how to determine the echo delay of a mixed signal. Note that we are working with the real cepstrum for now. Let $s[t]$ be the undistorted signal and τ be the echo delay. Since the contribution of the echo is attenuated in real life scenarios we introduce the loss parameter a. Therefore the received signal $x[t]$ is given as

$$x[t] = s[t] + s[t-\tau] \cdot a. \tag{3.1}$$

Applying a Fourier transform and the logarithm subsequently, the log-spectrum is given as

$$\log|X(e^{j\omega})| = \log|S(e^{j\omega})| + \frac{1}{2}\log\big(1 + a^2 + 2a \cdot cos(\omega\tau)\big)\,. \tag{3.2}$$

We can see in equation (3.2) that the echo introduces an additive contribution with the fundamental frequency τ in $\log|X(e^{j\omega})|$. Note that we are not really dealing with a fundamental frequency since this occurrence is already in the frequency domain. However if we opt to apply the Fourier, respectively the inverse Fourier transform again we expect a peak at exactly the echo delay.

It turns out that the cepstrum is an excellent tool for estimating the echo delay. Accordingly, this signal processing method was primarily used in the domain of speech processing. A major field where the ceptrum proved to be very helpful was for pitch detection [Nol64]; the pitch detection problem is formulated in a way very similar to echo detection.

Nevertheless the concepts of the cepstrum and homomorphic system theory were generalized and applied to various fields of problems. In [STTI75] the cepstrum is applied for blind deconvolution to restore old acoustic recordings. An application to seismology can be found in the work of [Ulr71]. Today the cepstrum plays a significant role for speech and word recognition and is widely applied [Fur86].

However due to the frequent occurrence of convolutions in the task of modeling the physical world, there are various fields where applying the cepstrum seems to be convenient. [OS04] gives a comprehensive overview of many applications, therefore we shall mention the references therein.

3.3 Types of the cepstrum

This chapter presents the various definitions for the cepstrum which can be found in the literature. The differences and the according advantages of each particular cepstrum are pointed out. Moreover it is shown how each type corresponds to the other definitions of the cepstrum.

We start with defining the **real cepstrum**, that is the definition given in [OS10]. Let $x[n]$ be a discrete time sequence, and $X(e^{j\omega})$ be its corresponding discrete time Fourier transform, $\mathcal{F}\left\{x[n]\right\}$:

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] \cdot e^{-j\omega n}; \tag{3.3}$$

Taking the magnitude of the Fourier transform and applying the logarithm subsequently we obtain $\log|X(e^{j\omega})|$.

The real cepstrum, which is denoted as $\underset{\sim}{c}_x[n]$, is estimated by the inverse discrete Fourier transform of the latter term, such that

$$\underline{c}_x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log|X(e^{j\omega})| \cdot e^{j\omega n} \, \mathrm{d}\omega. \tag{3.4}$$

For the real cepstrum to exist the energy of $\log|X(e^{j\omega})|$ has to be finite. These existence condition will be treated in further detail in section 4.2. The real cepstrum has the main advantage, that neglecting the phase term, eases the computation of the coefficients. Hence, the real cepstrum can not be inverted in general. However, certain constraints exist under which the original sequence can be obtained from looking at the real cepstrum.

In [BHT63] the original definition of the cepstrum was given by taking the the inverse Fourier transform of the squared magnitude instead. Therefore this type is often referred to as **power cepstrum**. Therefore equation (3.4) can be rewritten as follows.

$$c_x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log\left(|X(e^{j\omega})|^2\right) \cdot e^{j\omega n} \, \mathrm{d}\omega. \tag{3.5}$$

Note that the logarithm of the squared magnitude is just two times the logarithm of the absolute value. Consequently we can put the constant factor 2 in front of the integral which directly shows the connection between the real and the power cepstrum to

$$c_x[n] = 2 \cdot \underline{c}_x[n]. \tag{3.6}$$

Due to this direct relationship the advantages and restrictions of the real and the power cepstrum stay the same. However it shall be mentioned that, due to some properties which shall be mentioned later on, the power cepstrum was more suitable for the proposed applications. Hence for the rest of the work we usually opt for applying the power cepstrum, if it is sufficient to neglect the phase information.

In signal processing one often uses the complex logarithm to exploit the phase information. Note that the complex logarithm is given as

$$\log(z) = \log|z| + j \cdot \mathrm{Arg}(z). \tag{3.7}$$

with $z \in \mathbb{C}$. Obtaining the **complex cepstum** gives a more general form of the cepstrum. However it is by no means the case that the complex cepstrum does necessarily have to be complex valued. In fact for every real sequence the complex cepstrum will always have real valued coefficients.

The idea of expanding the definition and introducing a complex cepstrum is presented in [OS10]. This book really pays a lot of attention to the concept of the complex cepstrum and its properties. In this work we start by giving the corresponding definitions and present some practical properties. We will continue by applying the complex cepstrum to random variables and hope to get some further insights using the imaginary part of the logarithm as well.

Following the notation of the power cepstrum, we will denote the complex cepstrum as $\hat{c}_x[n]$, which is given as

$$\hat{c}_x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Big(\log|X(e^{j\omega})| + j \cdot \arg(X(e^{j\omega}))\Big) \cdot e^{j\omega n} \, \mathrm{d}\omega. \tag{3.8}$$

We can see that the real cepstrum is expanded with the phase information. Due to linearity of the Fourier Transform, we can directly obtain the power cepstrum by adding the complex conjugate pairs of the complex cepstrum:

$$c_x[n] = \hat{c}_x[n] + \hat{c}_x^*[n]. \tag{3.9}$$

Such a correspondence is obviously not reversible. Nonetheless, we will present some conditions under which it is feasible to obtain the complex cepstrum by looking at the power cepstrum.

Important properties in this work will be derived for the most general case, that is the complex cepstrum. Most properties can be used for the power cepstrum as well.

The complex cepstrum exists under the same constraints as the real cepstrum, that is the energy of $\log|X(e^{j\omega})|$ has to be finite. However we have to ensure that both the magnitude and phase are continuous functions. In [OS10] this is done through applying a simple phase unwrapping algorithm. We have no performance targets to achieve. Thus it is sufficient to adapt the tolerance parameter of the phase unwrapping algorithm for each simulation separately and stick to a simple algorithm as well.

However it shall be noted, that some alternatives are available. The necessity of phase unwrapping, if using the complex logarithm, and its underlying theory is presented in great detail in [BC09, pages 347-350]

A detailed study of the field of phase unwrapping problems and different ways to estimate the continuous phase function can be found in the Sc.D Thesis of Tribolet, J.M.; however, since these explanations are beyond the scope of this work, an older paper of the same author shall be mentioned [Tri77]. Therein an alternative algorithm for estimating the unwrapped phase is presented. It is pointed out how conventional phase unwrapping tackles the problem and some examples are mentioned where this approach fails. Consequently an algorithm using adaptive integration is explained and its advantages over the conventional method are presented.

Calculation of the complex cepstrum without applying any phase unwrapping method is shown in [WB85]. There the Fourier transform properties are exploited. The fact that differentiating the log function corresponds to a multiplication of the complex cepstrum with n, is applied. It is easier to derive the function in “frequency” first and estimate such an altered complex cepstrum:

$$\mathcal{IF}\left\{\frac{d}{d\mu}\log(X(e^{j\omega}))\right\} = n \cdot \hat{c}_x[n] \tag{3.10}$$

Following this estimation it is not a difficult task to divide each coefficient by n so that the complex cepstrum according to its definition is available. The algorithm is presented in detail in the according paper. Some experiments compare the performance of this algorithm to the classical approach using phase unwrapping.

4
Cepstrum Analysis for Random Variables

In the Marcinkiewicz theorem it was stated that there are infinitely many cumulants necessary to represent a non Gaussian distribution. Since the estimation of cumulants involves the logarithm of the characteristic function we can see that applying the Fourier series expansion instead of the Taylor expansion yields the cepstral coefficients of the according random variable.

Let $f_x(\xi)$ be the PMF of the random variable $\mathbf{X}$. Applying the discrete time Fourier transform we obtain the characteristic function $\mathcal{F}\{f_x(\xi)\} = \Phi_x(\mu)$ according to its definition in equation (2.2). Taking the logarithm and applying a Fourier series expansion leads to:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log(\Phi_x(\mu)) \cdot e^{j\mu n}\,\mathrm{d}\mu, \tag{4.1}$$

that coincidently corresponds directly to the definition of the complex cepstrum in equation (3.8).

Motivated by this connection we want to expand this idea and hope to get rid of some problems, which occur when working with cumulants. In section 4.1 we want to find an analytic connection between the cepstral coefficients and cumulants. The subsequent sections deal with restrictions which result, based on such a connection, and take a closer look at the cepstrum analysis and its existence conditions.

In section 4.3 we try to approximate the underlying random variable by looking at its cepstral coefficients, and find an interpretation of the cepstral coefficients. Some experiments for certain distributions are given in section 4.4.

4.1 Linking Cumulants and Cepstral Coefficients

In this section we want to give some relations between cepstral coefficients and cumulants or moments. Taking a closer look at these relations shall lead to further insights concerning the cepstral coefficients and yield some properties and constraints.

There are some papers which point at such relations, although these are mainly side results and not treated with priority and in detail. [Sch81] deals with the relations between the cepstrum and predictor coefficients. Furthermore, a connection with statistical moments and cumulants is presented. However, the complex cepstrum is defined using the z-transform instead, which has an influence on this connection. In [KY92] the moments of the cepstrum are calculated and it is shown how they correspond to the moments of the original sequence, in order to circumvent the explicit calculation of the cepstrum. In this work these cepstral moments are also linked to the cumulants for the application of echo removal and deconvolution.

Since both papers hint at a link between the cepstrum and the cumulants of the original sequence, although presenting different results due to different definitions of the cepstrum, we make further investigation.

Motivated through the work mentioned above, we will use basic Fourier transform properties to derive an explicit connection between the cepstrum and the cumulants. We start with the inverse Fourier transform as relation between the characteristic function and the cepstrum in equation 4.2. Note that we are looking at the complex cepstrum in this case, since this is the most general definition and the obtained results can be applied to the power cepstrum as well:

$$\log(\Phi_x(\mu)) = \sum_{n=-\infty}^{\infty} \hat{c}_x[n] \cdot^{-j\mu n} \tag{4.2}$$

Differentiation in the formal domain of the characteristic function corresponds to a multiplication with jn, so that

$$\frac{d^k}{d\mu^k} \log(\Phi_x(\mu)) = \sum_{n=-\infty}^{\infty} (jn)^k \hat{c}_x[n] \cdot^{-j\mu n} . \tag{4.3}$$

The exponential terms vanishes, due to evaluating the equation in the origin, so that:

$$\frac{d^k}{d\mu^k} \log(\Phi_x(\mu)) \,|_{\mu=0} = \sum_{n=-\infty}^{\infty} (jn)^k \hat{c}_x[n]. \tag{4.4}$$

We can now plug in the definition of the cumulants of order k, so that

$$j^k \kappa_k = j^k \sum_{n=-\infty}^{\infty} n^k \cdot \hat{c}_x[n]. \tag{4.5}$$

Looking at the right hand side of the equation there is a sum which looks similar to the k-th moment of a a discrete distributions. This leads to the idea of introducing a similar measurement $\hat{m}_k = \sum n^k \cdot \hat{c}_x[n]$. However, we have to consider that the definition of moments is evaluated using the probability mass function which is non negative, whereas there are no such restrictions for the cepstrum. We will call $\hat{m}_k$ the cepstral moment of order k to point at the similarities, noting that these are no moments in the conventional sense.
We can now rewrite equation 4.5 in a neat way to show the direct relation between the cumulants and the cepstral moments.

$$\kappa_k = \hat{m}_k. \tag{4.6}$$

There is some interest in estimating the cumulants as combination of moments [RS00]. Therefore the question arises if looking at the relation between cumulants and the cepstrum from a similar point of view would yield interesting results. However it is worth mentioning that looking at this connection, we can see that every single cumulant contains information of all existing cepstral coefficients. The same is valid into the other direction, that is, every single cepstral coefficient contains information regarding each existing cumulant.

Some constraints for the complex cepstrum can be derived by taking a close look at this connection, which can be seen in the subsequent section.

4.1.1 Constraints for the Cepstrum

Looking at the properties of a cumulant series we may find some resulting properties for the cepstral coefficients. It is mentioned in [OS10, page 1012] that the complex cepstrum decays at least as fast as $1/n$ in general. In particular $\lim_{n\to\infty} \hat{c}_x[n] = 0$.

However taking equation (4.6) into account we can find further restrictions for the behavior of cepstral coefficients. If κ_k exists and is finite it follows that the cepstral moment of order k exists and consequently $\sum n^k \cdot c_x[n]$ has to converge. Every right hand sided PMF has a right hand sided cepstrum, therefore we can state that the complex cepstrum of a right hand sided distribution decays with at least $1/n^k$ if the k-th order cumulant exists.

We continue with deriving a more general definition of the cepstral coefficients. Suppose that all derivatives of the characteristic function exist. Evaluating equation (3.8) leads to the following partial integral

$$\begin{aligned}\hat{c}_x[n] &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \log(\Phi_x(\mu)) \cdot e^{j\mu n} \quad \mathrm{d}\mu \qquad (4.7)\\ &= \frac{1}{2\pi}\left[\log(\Phi_x(\mu)) \cdot e^{j\mu n}\frac{1}{jn}\Big|_{-\pi}^{\pi} - \int_{-\pi}^{\pi}\frac{\mathrm{d}\log(\Phi_x(\mu))}{\mathrm{d}\mu} \cdot e^{j\mu n} \quad \mathrm{d}\mu\right]\\ &= \frac{1}{2\pi}\left[\log(\Phi_x(\mu)) \cdot e^{j\mu n}\frac{1}{jn} - \frac{\mathrm{d}\log(\Phi_x(\mu))}{\mathrm{d}\mu}\frac{e^{j\mu n}}{(jn)^2} + \frac{\mathrm{d}^2\log(\Phi_x(\mu))}{\mathrm{d}\mu^2}\frac{e^{j\mu n}}{(jn)^3} - \cdots\right]_{-\pi}^{\pi},\end{aligned}$$

which can be rearranged and rewritten as

$$\hat{c}_x[n] = \frac{1}{2\pi}\cdot\frac{-1}{jn}\sum_{k=1}^{\infty}(-1)^k\,\frac{\mathrm{d}^k\log(\Phi_x(\mu))}{\mathrm{d}\mu^k}\Big|_{\mu=-\pi}^{\pi}\cdot\frac{1}{(jn)^k}\cdot e^{j\mu n}. \qquad (4.8)$$

For distributions where $\log(\Phi_x(\mu))$ is a polynomial of order α_Φ it is sufficient to build the sum from $k = 1$ to $k = \alpha_\Phi$. As stated in the Marcinkiewicz theorem the normal distribution is the only distribution where such a polynomial has finite order, that is $\alpha_\Phi = 2$.

To conclude we can state, that for any other distribution the summands do not vanish. Therefore the cepstrum will have contributions which decay with greater power.

4.2 Existence of the Cepstrum

We shall deal with the existence and the convergence of the cepstral transformation in this section. Applying the cepstrum to random variables we know that $\sum f_x(\xi) = 1$ and is stable. Consequently we know from Parsevals theorem that the characteristic function is square integrable. Let the Fourier series be

$$\mathcal{IF}\left\{\log(\Phi_x(\mu))\right\}. \qquad (4.9)$$

For this transform to exist $\log(\Phi_x(\mu))$ must be absolutely integrable, that is:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|\log(\Phi_x(\mu))|d\mu < \infty. \qquad (4.10)$$

Using the complex logarithm some further considerations are necessary. Looking at the definition of the complex logarithm in equation (3.7), we can apply the triangle inequality, so that

$$|\log(\Phi_x(\mu))| \leq |\log|\Phi_x(\mu)|| + |\mathrm{Arg}(\Phi_x(\mu))|. \qquad (4.11)$$

The phase of the characteristic function is uniformly continuous and has compact support, that is on the interval $[-\pi, \pi]$. Thus the phase is a bounded function. Hence it is sufficient to neglect the phase term and guarantee absolute integrability, if

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |\log|\Phi_x(\mu)||d\mu < \infty, \tag{4.12}$$

is fulfilled. Since $|\Phi_x(\mu)| \leq 1$ we can change the sign and state that,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log|\Phi_x(\mu)|d\mu > -\infty, \tag{4.13}$$

which reminds us of the Paley Wiener condition. The Paley Wiener condition is fulfilled if equation (4.13) holds and if $\Phi_x(\mu)$ is square integrable. Note that $\Phi_x(\mu)$ is square integrable and therefore stable for every valid random variable.

If and only if the Paley Wiener criterion is fulfilled, then there exists a PMF which is right hand sided. Consequently we can ensure that the cepstrum exists for any right hand sided PMF, and for every PMF with bounded support.

For the remainder of the work we assume that the cepstrum exists.

4.3 Approximating Distributions applying Cepstral Coefficients

Suppose looking at the cepstral coefficients of a given random variable, we can clearly obtain the underlying PMF by applying the inverse cepstrum transformation considering all cepstral coefficients. This section compares the approximation using complex cepstral coefficients $\hat{c}_x[n]$ and power cepstral coefficients $c_x[n]$. Furthermore we are comparing the cumulants of the approximation to the cumulants of the given PMF.

Let $\Phi_x(\mu)$ be the characteristic function of a random variable $\mathbf{X}$. Looking at the according cepstral coefficients $\hat{c}_x[n]$, the logarithm of the characteristic function is obtained by applying the discrete Fourier transform $\log(\Phi_x(\mu)) = \mathcal{IF}\{\hat{c}_x[n]\}$. Note that we are looking at the complex cepstrum; however a similar inverse transformation will be found by looking at the real or the power cepstrum instead.

We can obtain the characteristic function, and therefore the PMF as well, by building the complete Fourier series and applying the exponential function subsequently, so that

$$\Phi_{\tilde{x}}(\mu) = e^{\sum_{n=-\infty}^{+\infty} \hat{c}_x[n]\cdot e^{-jn\mu}}. \tag{4.14}$$

Assume that the approximated characteristic function, $\Phi_{\tilde{x}}(\mu)$, is equal to the characteristic function $\Phi_x(\mu)$. It is obvious that the cumulants of the approximation are equal to the cumulants of $\mathbf{X}$.

Now let the approximation be based on the power cepstrum. We can apply the Fourier transform and the exponential function similar to equation (4.14) which leads to $|\Phi_{\tilde{x}}(\mu)|^2$. However since we neglected the phase for estimating the power cepstrum we cannot undo the absolute value due to its symmetry property.

Instead we we will look at $|\Phi_x(\mu)|^2$ as if it were the characteristic function $\Phi_Z(\mu)$ of a new random variable $\mathbf{Z}$. The probability mass function $f_Z(\xi)$ shall be evaluated and compared to $f_X(\xi)$.

We start by rewriting the absolute value for a complex valued variable as

$$\Phi_Z(\mu) = \Phi_x(\mu) \cdot \Phi_x(\mu)^*. \tag{4.15}$$

Using the fact that every valid characteristic function has to be Hermitian, we can further rewrite equation (4.15) to

$$\Phi_Z(\mu) = \Phi_x(\mu) \cdot \Phi_x(-\mu). \tag{4.16}$$

For further simplification we are now introducing a random variable $\mathbf{Y}$ which is independent and identical distributed (i.i.d.) to $\mathbf{X}$. Looking at the difference of these two random variables $\mathbf{X} - \mathbf{Y}$ we can rewrite their joint characteristic function to

$$\Phi(\mu) = \Phi_x(\mu) \cdot \Phi_y(\mu) = \Phi_x(\mu) \cdot \Phi_{-x}(\mu), \tag{4.17}$$

since X and Y are i.i.d. Applying Fourier transform properties [OSB99, page 59] we can now bring equation (4.16) and (4.17) together so that

$$\Phi_x(\mu) \cdot \Phi_{-x}(\mu) = \Phi_x(\mu) \cdot \Phi_x(-\mu). \tag{4.18}$$

This reveals that the random variable $\mathbf{Z}$ which follows from the power cepstral coefficients $c_x[n]$ of $\mathbf{X}$, is no more than the difference between two i.i.d. random variables.

Moreover we know from [PP02, page 182] that the distribution of the sum, respectively their difference equals to the convolution of the independent densities.

The convolution of two i.i.d. distributions is basically the autocorrelation of the density of $\mathbf{X}$. Such a cepstrum of the autocorrelation is often called autocepstrum, however it shall be noted that the autocepstrum is sometimes defined as the autocorrelation of the complex cepstrum.

Such an approximation clearly yields a different characteristic function. Consequently, $\mathbf{Z}$ will have different cumulants and moments as $\mathbf{X}$. We can assume that the moments will not give us any useful information due to their nonadditive nature, and the random variable $\mathbf{Z}$ being the difference of two i.i.d random variables. For this reason we will pay no further attention to the moments of the reconstructed characteristic function. Instead we will investigate the bahaviour of the corresponding cumulants.

The cumulants of the approximation are given to

$$\kappa_n(\mathbf{Z}) = \kappa_n(\mathbf{X}) + \kappa_n(\mathbf{Y}). \tag{4.19}$$

We know that $\mathbf{Y} = -\mathbf{X}$, so that $\kappa_n(\mathbf{Y}) = (-1)^n \cdot \kappa_n(\mathbf{X})$ and consequently equation (4.19) can be rewritten as

$$\kappa_n(\mathbf{Z}) = \kappa_n(\mathbf{X}) \cdot (1 + (-1)^n). \tag{4.20}$$

Thus

$$\kappa_n(\mathbf{Z}) = \begin{cases} 2 \cdot \kappa_n(\mathbf{X}), & \text{n even} \\ 0, & \text{n odd} \end{cases} \tag{4.21}$$

It is of great interest if it is possible to get an estimate of the distribution of $\mathbf{X}$ by just looking at the cepstral coefficients.

4.4 Experiments

In this section the knowledge which we gained so far will be applied to different types of random variables. The according power cepstrum and complex cepstrum will be presented. Furthermore it is pointed out how to reconstruct the distribution parameters given the cepstrum. As already mentioned we restricted the application of the cepstrum to discrete random variables. There are five distributions which will be investigated in this section.

The **binomial distribution** will be presented and investigated, due to its broad field of applications. We will further look at the **Poisson distribution** and investigate its cepstrum to show some unique properties which can be seen by looking at Poisson distributions. The **uniform distribution** shall be the presented subsequently. We will further investigate the **delta distribution** and the **geometric distribution**.

If not mentioned otherwise, all formulas for the according PMF or characteristic function can be found in [PP02, page 162].

4.4.1 Cepstrum of Binomial Distributions

The binomial distribution is determined with just two parameters. Let $p \in [0, 1]$ be the probability of success in each trial and consequently $q = 1 - p$ be the probability of failure. The number of trials is given as $N \in \mathbb{N}$. Note that the binomial distribution approximates the Poisson distribution for small p and large N. The according characteristic function shall be presented due to its importance in this work.

$$\Phi_x(\mu) = (pe^{j\mu} + q)^N. \tag{4.22}$$

Actually we can see that the number of trials N is just a constant scaling factor of the cepstrum. Looking at equation (3.8) and plugging in the above definition of the characteristic function $\Phi_x(\mu)$ we can bring down the exponent because of the logarithm, such that

$$\hat{c}_x[n] = \frac{N}{2\pi} \int_{-\pi}^{\pi} \log(p \cdot e^{j\mu} + q) \cdot e^{j\mu n} \, \mathrm{d}\mu. \tag{4.23}$$

The PMF and its characteristic function can be seen in figure 4.1 and figure 4.2 respectively. We can see how the PMF is skewed to the left and may approach a Poisson distribution for a relatively small success probability. Figure 4.3 shows both, the complex and the power cepstrum of a binomial distribution with a success probability of $p = 0.3$ and a number of trials $N = 50$.

One can see that the name of the complex cepstrum may be misleading indeed. Note that the complex cepstrum is real valued but takes different values than the power cepstrum because we are considering the complex part of the logarithm as well.

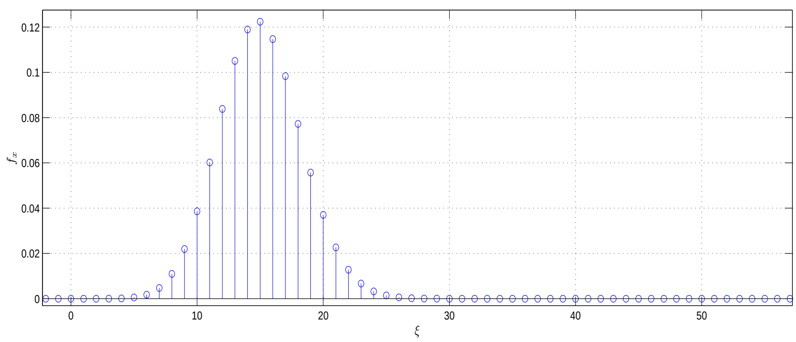

Figure 4.1: probability mass function for a binomial distribution for $p = 0.3$ and $N = 50$.

Figure 4.4 presents the cepstrum for a distribution with the same success probability $p = 0.3$. However the number of trials $N = 100$, that is two times the value of the above example. As

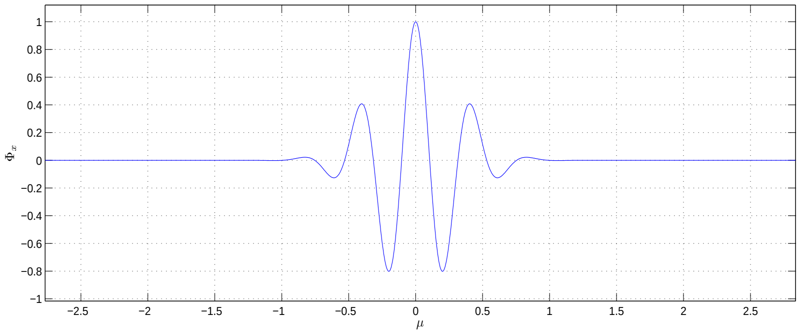

Figure 4.2: real part of the characteristic function for a binomial distribution for $p = 0.3$ and $N = 50$.

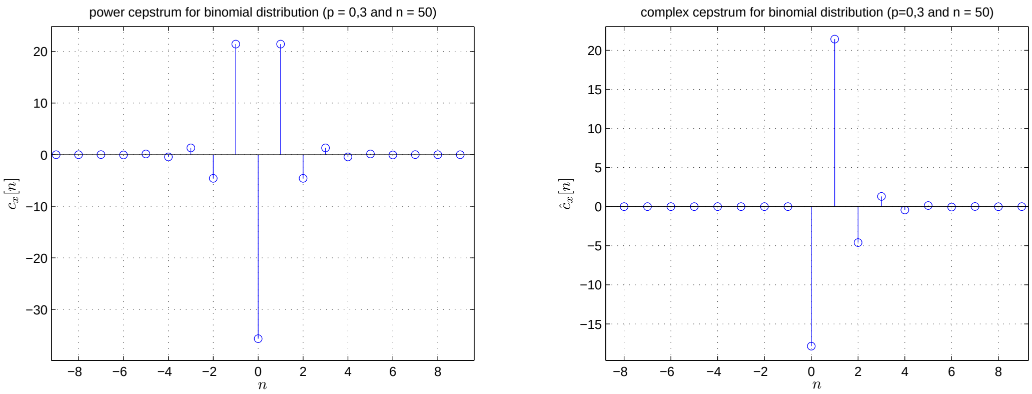

Figure 4.3: power cepstrum and complex cepstrum for a binomial distribution with success probability $p = 0.3$ and $N = 50$ number of trials.

expected one can see that the complex cepstrum as well as the power cepstrum experience a scaling with a constant factor of two.

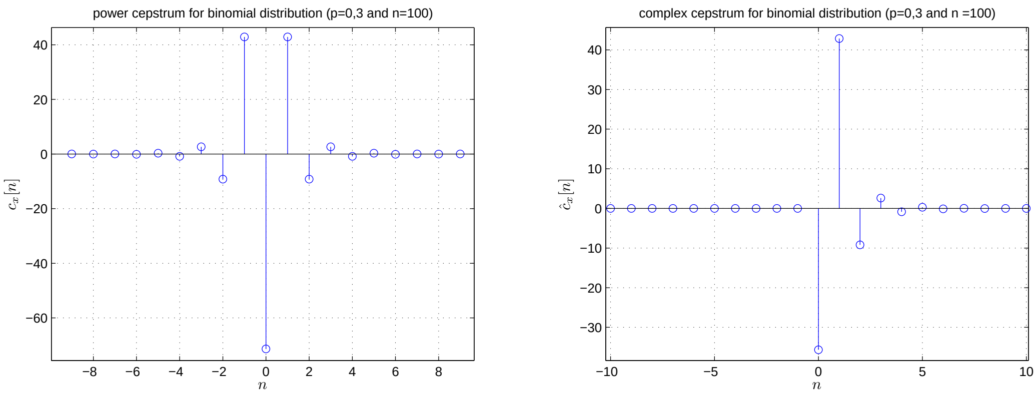

Figure 4.4: power cepstrum and complex cepstrum for a binomial distribution with success probability $p = 0.3$ and $N = 100$ number of trials. Notably the coefficients are two times the coefficients presented in figure 4.3.

Therefore we state that the behaviour of the cepstral coefficients is mainly determined by the success probability. If we choose $p = 0.45$ we can see in figure 4.5 that the cepstral coefficients decay more slowly. Consequently we can say that the more the binomial distribution approximates the normal distribution, that is for $p = 0.5$ the slower the cepstrum decays.

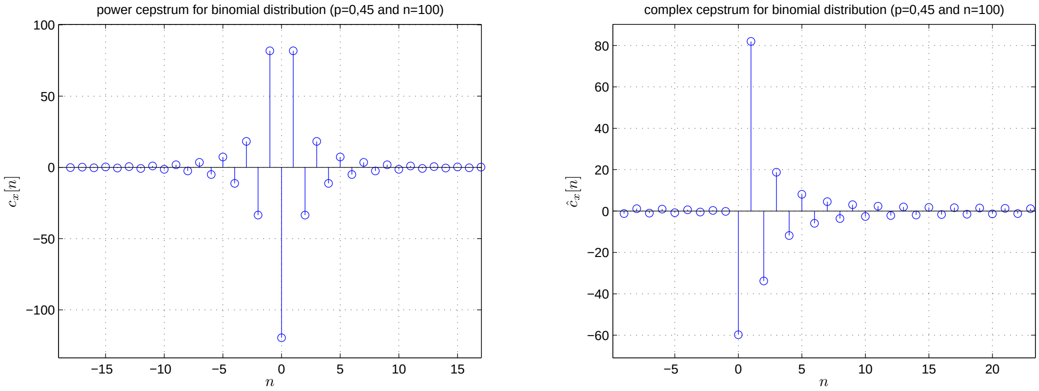

Figure 4.5: power cepstrum and complex cepstrum for a binomial distribution with success probability $p = 0.45$ and $N = 100$ number of trials. Note how there are increasingly more cepstral coefficients not equal to zero.

However it shall be noted that the cepstrum can not be estimated straight forward in this special case. Let $N = 1$ since the number of trials has no influence on the convergence behavior in any case. To highlight the problem, we rewrite equation (4.22) by applying Euler's formula to the characteristic function, such that

$$\Phi_x(\mu) = q + p \cdot \cos(\mu) + jp \cdot \sin(\mu), \tag{4.24}$$

and consequently the Fourier transform in equation (4.23) can be rewritten as

$$\hat{c}_x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log(q + p \cdot \cos(\mu) + jp \cdot \sin(\mu)) \cdot e^{j\mu n} \, \mathrm{d}\mu. \tag{4.25}$$

In order for a Fourier transform representation to converge the series may be either absolutely summable or converge in the mean squared sense if the series is square integrable. To recap, the Paley Wiener condition holds if if $|\Phi_x(\mu)| > 0$ almost everywhere. This condition is fulfilled if and only if

$$\frac{1}{2\pi} \cdot \int_{-\pi}^{\pi} \log|\Phi_x(\mu)| \, \mathrm{d}\mu > -\infty. \tag{4.26}$$

Since $\lim_{\mu \to \pi} (p \cdot cos(\mu)) = -p$ we can conclude that $\lim_{\mu \to \pi} |\Phi_x(\mu)| = q - p$. In this special case $q = p = 1/2$, such that we have a singularity for $\mu = \pm\pi$. Note that we are looking at a well behaved singularity. Thus we can still estimate the cepstrum as long as we evaluate the Fourier transform with these singularities in mind. Thus we can rewrite equation (4.25) to

$$\hat{c}_x[n] = \frac{1}{2\pi} \lim_{\epsilon \to 0} \int_{-\pi+\epsilon}^{\pi-\epsilon} \log(q + p \cdot \cos(\mu) + jp \cdot \sin(\mu)) \cdot e^{j\mu n} \, \mathrm{d}\mu, \tag{4.27}$$

such that the Paley Wiener condition holds. Consequently we can estimate the cepstral coefficients for a binomial distribution with $p = 0.5$.

Nevertheless it is of great interest, if it is possible to reconstruct the binomial distribution parameters just by looking at the cepstrum.

Furthermore we want to investigate the influence of limiting the number of cepstral coefficients which are taken into account. However it has to be noted that the whole information of the distribution is spread over all cepstral coefficients. Yet we can estimate the parameters by looking at the cumulants, that is evaluating the cepstral moments. For now we loosen the definition of the cepstral moments in (4.5) to

$$\hat{m}_k = \sum_{n=-L}^{L} n^k \cdot \hat{c}_x[n]. \tag{4.28}$$

Hence we are considering $2 \cdot L$ cepstral coefficients for estimating the cepstral moments. A good list of cumulants for the binomial distribution including higher order cumulants can be found in [Hal40]. Taking $\hat{m}_2$ and $\hat{m}_3$ into account the parameters can be estimated to:

$$\hat{m}_2 = Npq \tag{4.29}$$

$$\hat{m}_3 = Npq \cdot (p - q) \tag{4.30}$$

$$\frac{\hat{m}_3}{\hat{m}_2} = p - q \tag{4.31}$$

We substitute $q = 1 - p$ and thus finally obtain,

$$p = 1/2 + \frac{\hat{m}_3}{2 \cdot \hat{m}_2} \tag{4.32}$$

Subsequently the number of trials can be obtained directly to

$$N = \frac{\hat{m}_2}{p \cdot (1 - p)} \tag{4.33}$$

It is obviously necessary to sum up over all cepstral coefficients in order to a get a correct estimate for the binomial distribution. However it shall be stated that it is possible to get a good estimate of the binomial distribution with a finite sum, due to the strong decline of the cepstrum.

We shall present the results of a short experiment. Let us estimate the parameters of a binomial distribution with $p = 0.3$ and $N = 100$. Table 4.1 presents the obtained parameters via the cepstral moments applying equation (4.32) and (4.33) for considering $2 \cdot L + 1$ coefficients. We can see that taking very little coefficients into account the calculation fails. However for

L	p	n
2	3	-1.02
5	0.02	1052
10	0.316	96.77
15	0.299	100.1

Table 4.1: estimated parameters for the binomial distribution

taking $L = 10$ coefficients into account we can already observe a reasonable accuracy of the estimation. Taking $L = 15$ coefficients into account we can even state that the estimation of the parameters is really good.

4.4.2 Cepstrum of Poisson Distributions

In this section we will estimate the cepstrum for the Poisson distribution. We will show that the complex cepstrum has only two coefficients which are nonzero and exploit the unique position the Poisson distribution takes.

We shall present the characteristic function of the Poisson distribution first. Let $\lambda \in \mathbb{R}_+$, so the characteristic function is given as

$$\Phi_x(\mu) = e^{\lambda(e^{j\mu}-1)}. \tag{4.34}$$

As a matter of fact the Poisson has infinitely many cumulants, however it is notable that these cumulants are all exactly the same.

$$\kappa_n = \lambda, \tag{4.35}$$

for $n \in \mathbb{N}$ and λ being the mean, as well as the variance, skewness and so on.

Estimating the Cepstral Coefficients: We will use the method of equating the coefficients. Writing down the logarithm of the characteristic function yields

$$\log(e^{\lambda(e^{j\mu}-1)}) = \lambda e^{j\mu} - \lambda = \sum_{n=-\infty}^{\infty} \hat{c}_x[n] \cdot e^{-j\mu n}. \tag{4.36}$$

Hence we can directly obtain the coefficients, which are given as

$$\hat{c}_x[0] = -\lambda \tag{4.37}$$
$$\hat{c}_x[1] = \lambda \tag{4.38}$$
$$\hat{c}_x[n] = 0 \quad \text{for } n < 0 \text{ and } n \geq 2 \tag{4.39}$$

The PMF and the characteristic function of a Poisson distribution with $\lambda = 10$ are sketched in figure 4.6 and figure 4.7. The similarity between the Poisson and the Binomial distribution can be seen by looking at the according characteristic functions.

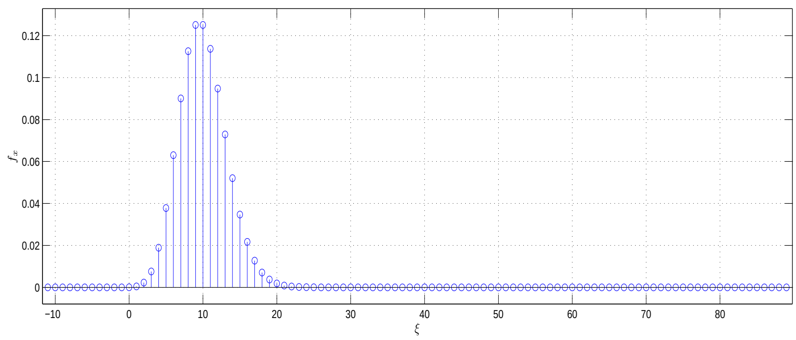

Figure 4.6: probability mass function for a poisson distribution for $\lambda = 10$.

The cepstral coefficients for a Poisson distribution with $\lambda = 10$ can be seen in figure 4.8. Looking at the complex cepstrum we can see that all but two coefficients are equal to zero. These two coefficients directly correspond to the parameter of the Poisson distribution. Therefore a reconstruction of the underlying random variable can be done by looking at the complex cepstrum.

The question arises if it is possible to construct a different distribution with just two coefficients as well. Considering the fact that the sum over all cepstral coefficients is always zero (Lemma 5.6.1) and that the zero order coefficient is always non positive (Lemma 5.6.2) it is implied that $\hat{c}_x[0] = -\hat{c}_x[1]$. Hence we can state that having just two non zero coefficients $\hat{c}_x[0]$ and $\hat{c}_x[1]$ we are always looking at a Poission distribution.

Remarkably we may refer to the complex cepstrum of the Binomial distribution in sec-

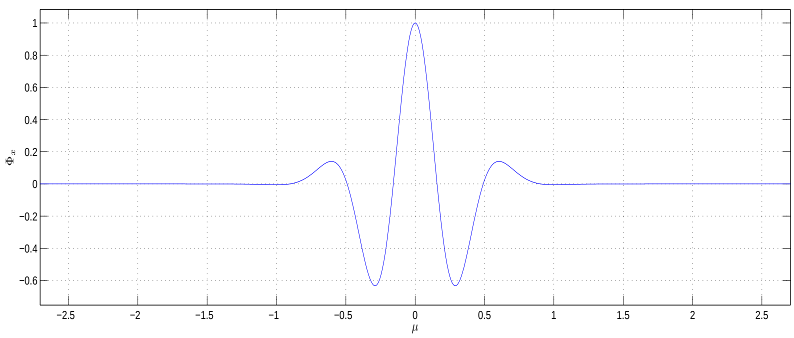

Figure 4.7: real part of the characteristic function for a poisson distribution for $\lambda = 10$.

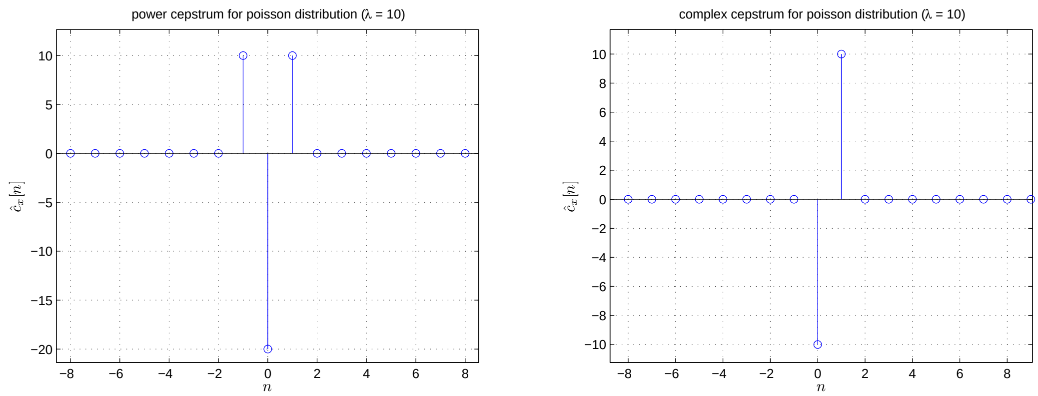

Figure 4.8: power cepstrum and complex cepstrum for a poisson distribution with $\lambda = 10$

tion 4.4.1. There we stated, that the smaller the probability of success p, the less cepstral coefficients are significant. Therefore for p being sufficient small, all but the first two coefficients of the complex cepstrum will converge to zero. In fact, this perfectly relates to theory, since such a Binomial distribution approximates to a Poisson distribution.

Subsequently we will roughly sketch the estimation of the power cepstral coefficients as well. Instead of evaluating $\log(\Phi_x(\mu))$ in equation (4.36) we use the $\log|\Phi_x(\mu)|^2$.

We split up the exponential term on the left hand side of the equation in its real and complex part. Due to the absolute value the complex value vanishes such that

$$2 \cdot \lambda \cos(\mu) - 2 \cdot \lambda = \sum_{n=-\infty}^{\infty} c_x[n] \cdot e^{-j\mu n}. \tag{4.40}$$

Rewriting the cosine in its exponential terms directly leads to the following solution.

$$c_x[\pm 1] = \lambda \tag{4.41}$$

$$c_x[0] = -2\lambda \tag{4.42}$$

$$c_x[n] = 0 \quad \text{for } |n| \geq 2 \tag{4.43}$$

Similar to the complex cepstrum there is no way of arranging three power cepstral coefficients other than in the sense of a Poisson distribution.

Consequently we want to find a condition for a random variable to have finitely many cepstral coefficients. For a characteristic function to analytically have finitely many cepstral coefficients the characteristic function has to be an exponential function in a way that,

$$\Phi_x(\mu) = e^{f(\mu)}, \tag{4.44}$$

where $f()$ is a finite order Fourier series in μ. While some distributions like the Poisson or the Cauchy distribution have according characteristic functions most of the commonly used distributions do not meet such a constraint. However, higher order coefficients may still converge to zero in a way that we can truncate the series.

4.4.3 Cepstrum of Uniform Distributions

In this section we want to investigate the cepstrum of the discrete uniform distribution. Let $a, b \in \mathbb{Z}$ be the parameters of random variable. Suppose that a marks the beginning and b the end of the support of the uniform distribution. The resulting parameter N is given to $N = b - a + 1$, and the characteristic function is given as,

$$\Phi_x(\mu) = \frac{e^{j \cdot a\mu} - e^{j\mu(b+1)}}{N(1 - e^{j\mu})} \tag{4.45}$$

We estimate the cepstrum for a uniform distribution with $a = 0$ and $b = 10$. The complex cepstrum and the power cepstrum are presented in figure 4.9. Both types of coefficients show a similar decay of the coefficients. However we notice some coefficients which take negative values. The cepstrum takes negative value every $N = 11$ coefficients. Consequently a uniform distribution with $a = 0$ can be fully reconstructed by looking at the first negative coefficient. It follows that $b = k - 1$ whereas $k = N$ is the index of this negative coefficient. Notable we can get these descriptors either by looking at the complex or the power cepstrum.

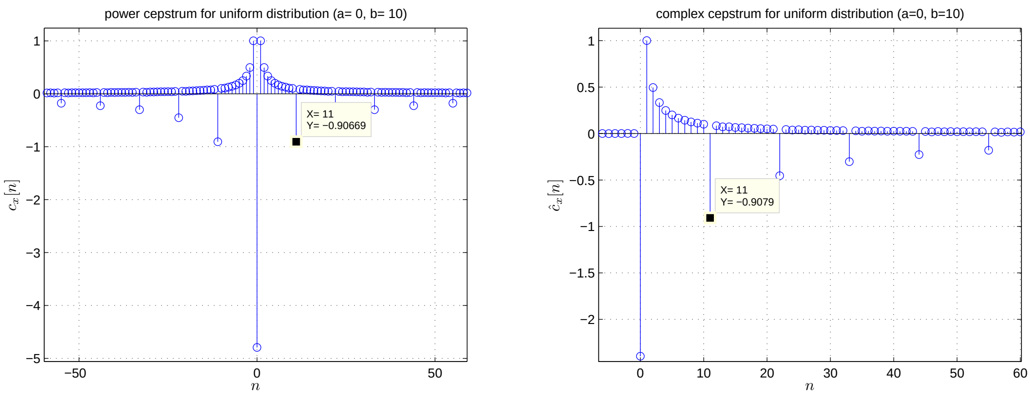

Figure 4.9: power cepstrum and complex cepstrum for a uniform distribution with $a = 0$ and $b = 10$.

Figure 4.10 shows a similar distribution with $N = 11$. However in this example $b = 0$ and $a = -10$. As the power cepstrum looses all information regarding the position of the distribution, the power cepstrum is the same for two similar distributions, which are just distinguished by a shift along the ξ-axis. Comparing figure 4.9 and 4.10 we can validate the statement.

Interestingly the coefficients of the complex cepstrum stay the same, although they are mirrored. Let $\mathbf{X}$ be the random variable of this example, and $\mathbf{Y}$ the one of the above example we can state that

$$\hat{c}_x[-n] = \hat{c}_y[n]. \tag{4.46}$$

Meanwhile this observation shall be suspended. We will make further investigation in section 5.10.1.

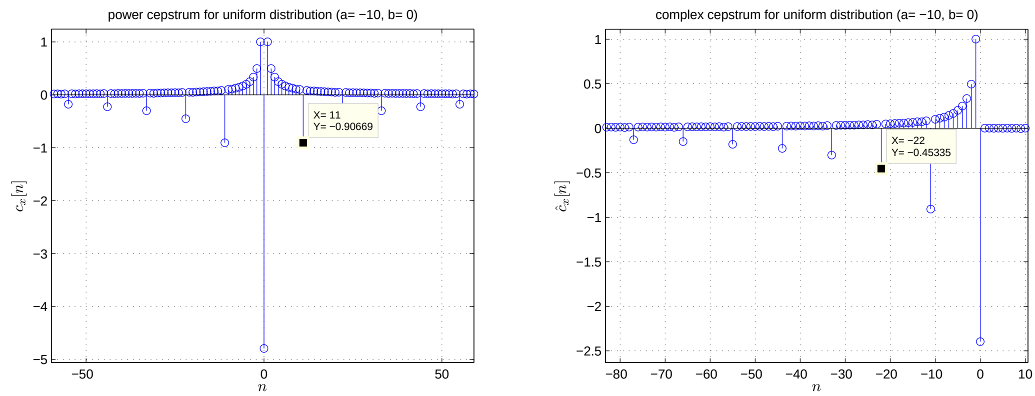

Figure 4.10: power cepstrum and complex cepstrum for a uniform distribution with $a = -10$ and $b = 0$.

We shall take a look at a third uniform distribution which is centered in the origin. Figure 4.11 gives the cepstral coefficients for a distribution with $a = -10$ and $b = 10$. We can see that the complex cepstrum decays, however no direct relationship to the distribution descriptors can be found.

We may not be able to fully reconstruct the uniform distribution by just looking at the power cepstrum, as the power cepstrum is shift invariant. Nevertheless we can directly obtain N by considering the first negative coefficient. Note that the complex cepstrum for a centered uniform distribution in figure 4.11is the sum of the of a right hand side uniform distribution and a delta distribution. See the according contribution in section 4.4.4.

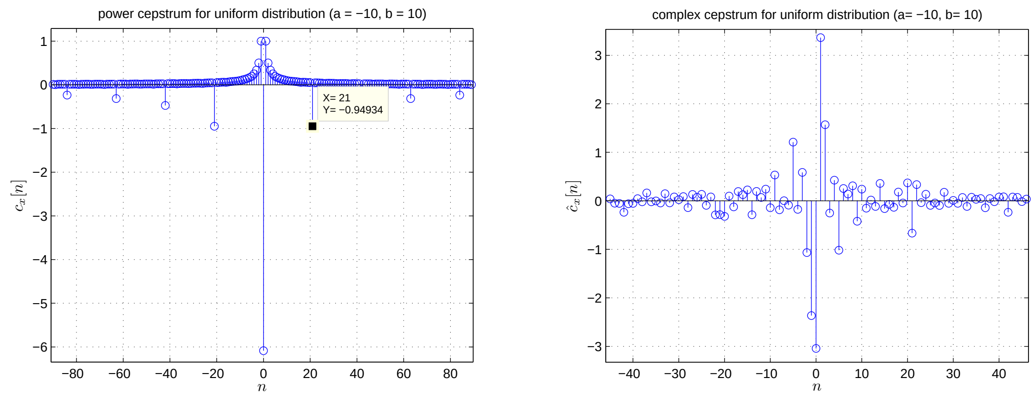

Figure 4.11: power cepstrum and complex cepstrum for a uniform distribution with $a = -10$ and $b = -10$.

4.4.4 Cepstrum of Delta Distributions

In this section we will present the cepstrum for a non standard distribution. However since these distributions are valid and will be used later on we shall present the according cepstrum. The Kronecker delta $\delta[n]$ is a valid PMF in the sense that it is non negative for all integer numbers and the sum over the whole domain is one.

Looking at the Fourier transform, the characteristic function $\Phi_x(\mu) = 1$ everywhere. Taking a step further we know that the logarithm for the characteristic function is zero everywhere and consequently all cepstral coefficients are zero as well.

Looking at a shifted delta $\delta[n-n_0]$ the characteristic function is given as $\Phi_x(\mu) = e^{-j\mu \cdot n_0}$. The absolute value of the logarithm is obviously equal to zero. Hence the power cepstrum is zero everywhere. However taking the phase into account we can estimate the complex cepstrum, which is consequently given to

$$\hat{c}_x[n] = \mathcal{F}^{-1}\left\{(j\mu n_0)\right\} = \frac{1}{2\pi} \cdot \frac{2n(\pi n \cdot \cos(\pi n) - \sin(\pi n))}{n^2} \tag{4.47}$$

We can neglect the sine contribution, since we are looking at integer values of n, and the above equation simplifies to

$$\hat{c}_x[n] = \frac{n_0}{n} \cdot \cos(\pi n) \tag{4.48}$$

$$= \frac{n_0}{n} \cdot (-1)^{n+1} \tag{4.49}$$

which is perfectly in tune with the following example.

Let $\mathbf{X}$ be a shifted delta distribution with $n_0 = 5$, whose PMF is zero everywhere except for $f_x(\xi = 5) = 1$. The complex cepstrum is consequently given by equation (4.49) and can be seen in figure 4.12.

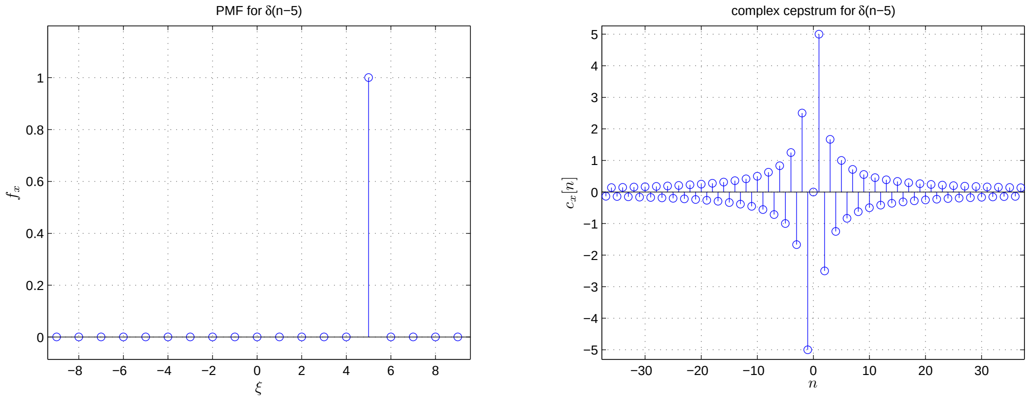

Figure 4.12: PMF and complex cepstrum for $\delta(n-5)$

4.4.5 Cepstrum of Geometric Distributions

We shall present the estimation of the complex cepstrum for the geometric distribution and how the probability of success $p \in [0, 1)$ can be estimated by looking at the cepstrum. These steps are explained by looking at a right hand sided geometric distribution. However we will generalize the results to the left hand sided and two sided geometric distribution later on. To start with, the characteristic functions for the different types of the geometric distribution shall be given. Let $\mathbf{X}_r$ be a right hand side geometric distribution, then

$$\Phi_{x,r}(\mu) = \frac{p}{1-(1-p) \cdot e^{-j\mu}}. \tag{4.50}$$

Let $\mathbf{X}_l$ be a negative geometric distribution. Note that we modify the general definition of the characteristic function by introducing $\alpha \in \mathbb{Z}$. That corresponds to the rightmost non zero value of the distribution. Let be $\alpha = 0$ to shift the distribution into the origin.

$$\Phi_{x,l}(\mu) = \frac{p \cdot e^{-\alpha j\mu}}{1 - (1-p) \cdot e^{j\mu}}. \tag{4.51}$$

Let $\mathbf{X}_t$ be a two sided geometric distribution which is constructed by adding the PMF of a positive and a negative geometric distribution. Note that we are not adding the random variables but the PMF instead. The Fourier transform preserves addition, thus we are adding the corresponding characteristic functions as well. Nonetheless we have to alter the according characteristic function to ensure a valid distribution.

$$\Phi_{x,t}(\mu) = 1/2 \cdot \left(\underbrace{\frac{p}{1-(1-p)\cdot e^{-j\mu}}}_{\text{positive geometric distribution}} + \underbrace{\frac{p\cdot(1-p)\cdot e^{1j\mu}}{1-(1-p)\cdot e^{j\mu}}}_{\text{negative geometric distribution}} \right). \tag{4.52}$$

Let the negative contribution start at $\xi = -1$, that is choosing $\alpha = -1$. Note that the numerator is $p \cdot (1-p)$, therefore the first probability of the negative geometric distribution is omitted. Thus it is virtually starting in the origin. To warrant the sum over the whole domain to be 1, $\sum_\xi f_x(\xi) = 1$, we scale the resulting characteristic function with 0.5.

The PMFs for all three types of a geometric distributions with probability of success $p = 0.2$ are given in figure 4.13.

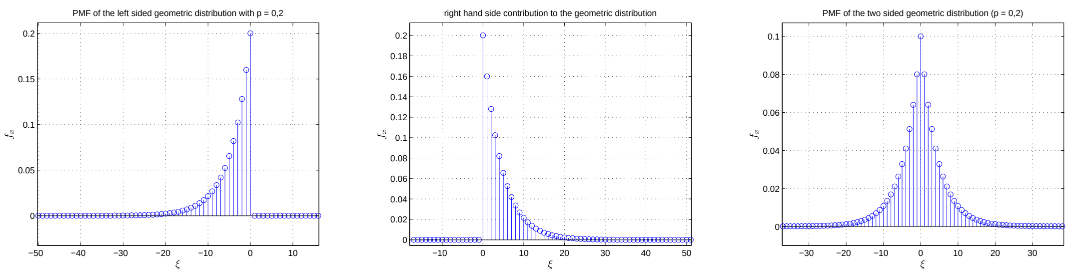

Figure 4.13: PMF for the negative, positive and two-sided geometric distribution with the given characteristic functions for p = 0.2

Taking a closer look at the complex cepstrum of the **right hand side geometric distribution** we can see that $\hat{c}_{x,r}[0]$ is taking a negative value, whereas the rest of the coefficients decays with exponential behavior. In particular by looking at $\hat{c}_{x,r}[1]$ we can find a direct relationship between the parameter p and the cepstrum. It can be seen in figure 4.14 that $\hat{c}_x[1] = 1 - p$ and therefore $\hat{c}_x[1]$ is sufficient to reconstruct the distribution.

Proof. To evaluate the first order coefficient of the complex cepstrum we plug in the characteristic function of equation (4.50), so that

$$\hat{c}_x[1] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log\left(\frac{p}{1-(1-p)\cdot e^{-j\mu}}\right) e^{j\mu}\, \mathrm{d}\mu. \tag{4.53}$$

Splitting the logarithm leads to

$$\hat{c}_x[1] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log(p)\, e^{j\mu}\, \mathrm{d}\mu - \frac{1}{2\pi} \int_{-\pi}^{\pi} \log\left(1-(1-p)\cdot e^{-j\mu}\right) e^{j\mu}\, \mathrm{d}\mu. \tag{4.54}$$

$\log(p)$ is a constant, hence the first integral integrates to zero. To recap we can write the complex

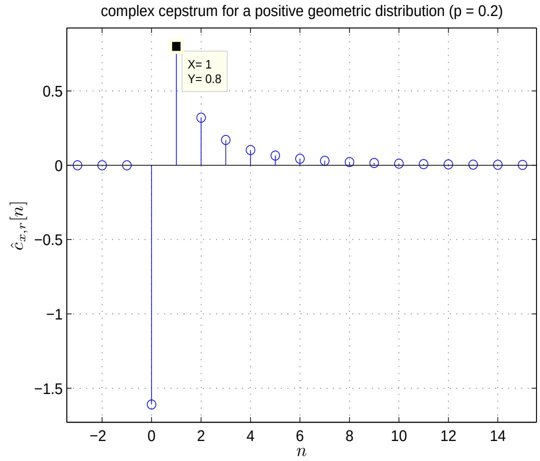

Figure 4.14: complex cepstrum for the positive geometric distribution for $p = 0.2$. $\hat{c}_{x,r}[1]$ is marked to show the connection to p

logarithm in terms of its Taylor series expansion, that is:

$$\log(z) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \cdot (z-1)^n, \tag{4.55}$$

$$= (z-1) - \frac{(z-1)^2}{2} + \frac{(z-1)^3}{3} \cdots . \tag{4.56}$$

Consequently we can rewrite equation (4.54) using the above expansion with $z = 1-(1-p)e^{-j\mu}$. We can pull the integrand into the sum, so that we are looking at a sum over infinitely many integrals:

$$\hat{c}_x[1] = -\frac{1}{2\pi}\int_{-\pi}^{\pi}(p-1)e^{-j\mu}e^{j\mu}\mathrm{d}\mu + \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{((p-1)e^{-j\mu})^2}{2}e^{j\mu}\mathrm{d}\mu - \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{((p-1)e^{-j\mu})^3}{3}e^{j\mu}\mathrm{d}\mu \cdots . \tag{4.57}$$

For the first term the exponential term vanishes and we integrate over a constant, whereas some exponential function remains for the rest of the terms. Note that we are always integrating on the unit circle, therefore all but the first term integrate to zero. Thus the first order coefficient of the complex cepstrum is given as:

$$\hat{c}_x[n] = -\frac{1}{2\pi} \cdot (p-1) \cdot 2\pi = 1-p. \tag{4.58}$$

□

As expected looking at the complex cepstrum of the **negative geometric distribution** and comparing it to $\hat{c}_{x,r}[n]$ we can seen that both cepstra are equal, except of being mirrored. That is $\hat{c}_{x,l}[n] = \hat{c}_{x,r}[-n]$. Thus p can be obtained by looking at $\hat{c}_{x,l}[-1]$. The according complex cepstrum can be seen in figure 4.15. Note that by choosing $\alpha = 0$ we choose the cepstral coefficients, to be similar to the right hand side case.

Let $\mathbf{Y}$ be a negative geometric distribution with $\alpha = -n_0$, we are in fact looking at $\mathbf{Y} = \mathbf{X}_l + n_0$. As a matter of fact the the cepstrum is given by the sum of the individual terms, that is $\hat{c}_y[n] = \hat{c}_{x,l}[n] + \hat{c}_{\delta,n_0}[n]$. Nonetheless as the cepstrum for the latter contribution is given in the preceding chapter to equation 4.49 one could still obtain $\hat{c}_{x,l}[n]$.

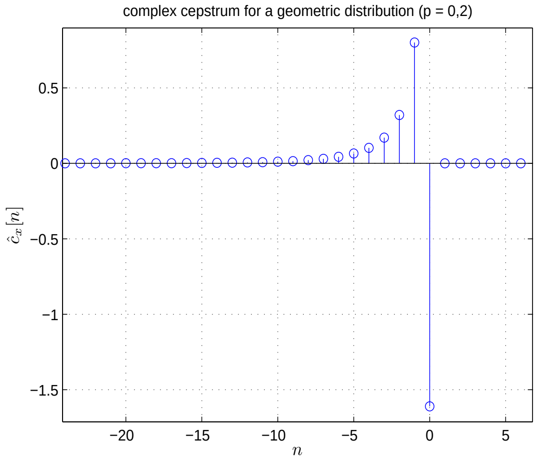

Figure 4.15: complex cepstrum for the negative geometric distribution for $p = 0.2$. Note the similiarity to figure 4.14

We can see that $\hat{c}_{x,r}[n] = 0$ for all $n < 0$ and $\hat{c}_{x,l}[n] = 0$ for all $n > 0$. Interestingly it happens that the cepstrum of the **two sided geometric distribution** is just the sum of these cepstral coefficients. Thus p can be obtained by either looking at $\hat{c}_{x,t}[1] = 1 - p$ or $\hat{c}_{x,t}[-1] = 1 - p$.

Note that $\hat{c}_{x,t} = \underline{c}_{x,t}[n]$ since $f_x(\xi)$ is symmetric in the origin and $\Phi_{x,t}(\mu)$ is real valued. Figure 4.16 shows the according complex cepstrum.

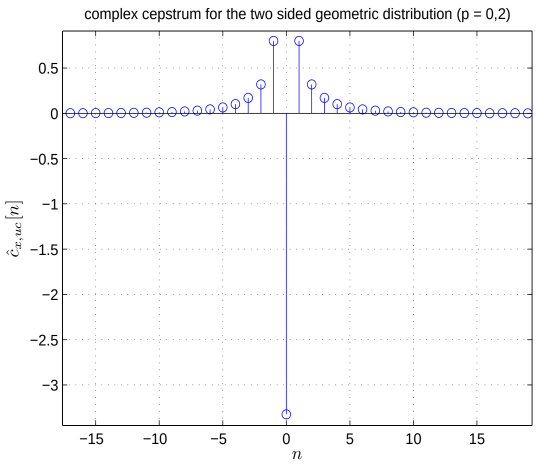

Figure 4.16: complex cepstrum for the two sided geometric distribution for $p = 0.2$.

5 Properties for the Cepstrum

In this chapter we will derive some properties similar to Fourier transform properties. We start with some symmetry properties and show how the cepstrum respond to certain changes of the random variable. Some constraints which have to be fulfilled to guarantee, that the cepstral coefficients yield a valid characteristic function are given. The properties are derived for the complex cepstrum, since this is the most general definition. However there are some properties which are especially interesting in case of estimating the power cepstrum. The according properties are presented in such cases.

An interesting connection between the complex and the power cepstrum will be pointed out in section 5.10. We will further investigate this connection and give certain constraints where it may be sufficient to neglect the phase.

Due to the importance of the Fourier transform in the estimation of the cepstrum, many properties where derived applying Fourier transform theory. A good overview of the according properties and transform pairs can be found in [OSB99].

5.1 Symmetry Properties for the Cepstrum

This section shall give a short overview of some symmetry properties.

Looking at a random variable $\mathbf{X}$, the even part of $f_x(\xi)$ transforms to the real part of the characteristic function $\Phi_x(\mu)$, whereas the odd part of $f_x(\xi)$ transforms to the imaginary part of $\Phi_x(\mu)$.

For every random variable $\mathbf{X}$ the the characteristic function $\Phi_x(\mu)$ is Hermitian. The magnitude $|\Phi_x(\mu)|$ is even symmetric, whereas the phase $\arg(\Phi_x(\mu))$ is odd symmetric. Consequently, the complex cepstrum is always real valued.

Let $\mathbf{X}$ be a random variable with $f_x(\xi)$ being symmetric in the origin. Note that any arbitrary random variable which is symmetric in the origin is in fact even symmetric, since $f_x(\xi)$ can just take positive values. Consequently $\Phi_x(\mu)$ is real and even symmetric. The logarithm does not change these properties. Applying the Time Reversal Theorem [OSB99, page 55] we know that an even function in time yields an even function in frequency and consequently, that the complex cepstrum is even symmetric in this case.

Note that the power cepstrum is obtained by looking at an autocorrelated random variable, that is a random variable which is centered in the origin. Hence the power cepstrum is always even symmetric. If we suppose that $c_x[n]$ is not an even function, we can see that the Fourier transform $\log|\Phi_x(\mu)|^2$ has to be complex, which is a contradiction. Consequently we can state the following lemma:

Lemma 5.1.1. *The power cepstrum is an even function.*

5.2 Convolution

The behavior of the cepstral coefficients with regard to convolution was one of the main motivations for introducing the cepstrum and homomorphic transformations in first hand. The sum $\tilde{\mathbf{X}}$ of two independent random variables $\mathbf{X_1}$ and $\mathbf{X_2}$ is equivalent to convolving the according PMFs. The complex cepstrum of such a convolution can be given in the following way.

$$
\begin{aligned}
f_{\tilde{x}}(\xi) &= (f_{x_1} * f_{x_2})(\xi) && (5.1)\\
\Phi_{\tilde{x}}(\mu) &= \Phi_{x_1}(\mu) \cdot \Phi_{x_2}(\mu) \\
\log(\Phi_{\tilde{x}}(\mu)) &= \log(\Phi_{x_1}(\mu)) + \log(\Phi_{x_2}(\mu)) \\
\hat{c}_{\tilde{x}}[n] &= \frac{1}{2\pi}\int_{-\pi}^{\pi} (\log(\Phi_{x1}(\mu)) + \log(\Phi_{x2}(\mu))) \cdot e^{j\mu n}\, \mathrm{d}\mu && (5.2)\\
\hat{c}_{\tilde{x}}[n] &= \hat{c}_{x_1}[n] + \hat{c}_{x_2}[n]
\end{aligned}
$$

Convolving two PMFs leads to a cepstrum equal to the sum of the individual cepstra. Looking at this property we can see that cepstral coefficients are additive descriptors, in the same way cumulants are.

5.3 Shift of the Mean

A shift of the mean corresponds to adding a constant to $\mathbf{X}$. Let m_0 be a constant which shifts the distribution. Therefore the characteristic function is given to

$$
\begin{aligned}
f_{\tilde{x}}(\xi) &= f_x(\xi - m_0) && (5.3)\\
\Phi_{\tilde{x}}(\mu) &= \Phi_x(\mu) \cdot e^{j\mu m_0}.
\end{aligned}
$$

The logarithm of the product is a sum, so that

$$
\begin{aligned}
\log(\Phi_{\tilde{x}}(\mu)) &= \log(\Phi_x(\mu)) + \log(e^{j\mu m_0}) && (5.4)\\
&= \log(\Phi_x(\mu)) + j\mu m_0.
\end{aligned}
$$

Coincidently, we already estimated the inverse Fourier transform of the second term in section 4.4.4. To recap, for $n \in \mathbb{Z}$,

$$
\frac{1}{2\pi}\int_{-\pi}^{\pi} j\mu m_0 \cdot e^{j\mu n}\, \mathrm{d}\mu = m_0 \cdot (-1)^{n+1}\frac{1}{n}. \quad (5.5)
$$

Putting the above equations together we can see that the cepstrum is extended with an additional term and is given to

$$\hat{c}_{\tilde{x}}[n] = \hat{c}_x[n] + m_0 \cdot (-1)^{n+1} \frac{1}{n}. \quad (5.6)$$

However it shall be noted that the second term contributing to the complex cepstrum is imaginary. The term vanishes when taking the logarithm of the absolute value instead. Thus the influence introduced through a shift of the mean vanishes and the power cepstrum is shift invariant.

5.4 Time Reversal

Mirroring the PMF in the origin corresponds to time reversal. If the distribution is mirrored the characteristic function becomes

$$\begin{aligned} f_{\tilde{x}}(\xi) &= f_x(-\xi) \\ \Phi_{\tilde{x}}(\mu) &= \Phi_x(-\mu), \end{aligned} \quad (5.7)$$

as is shown in [OSB99, page 60]. Considering that $\Phi_x(\mu)$ is Hermitian, equation (5.7) becomes

$$\Phi_{\tilde{x}}(\mu) = \Phi_x^*(\mu). \quad (5.8)$$

Taking the symmetry properties of the Fourier transform into account we can see that the resulting cepstrum is mirrored in the origin as well. That is,

$$\hat{c}_{\tilde{x}}[n] = \hat{c}_x[-n]. \quad (5.9)$$

Estimating the power cepstrum the phase of the characteristic function is neglected and consequently time reversal has no influence on the power cepstrum. However this is quite intuitive since the power cepstrum is even symmetric anyway.

5.5 Linearity

Although scaling a given probabilty mass function with a constant a does not yield a valid probability mass function anymore we want to show how linear scaling maps onto the cepstrum. The constant may be pulled in front of the Fourier transform so that the characteristic function is scaled as well.

$$\begin{aligned} f_{\tilde{x}} &= a \cdot f_x(\xi) \\ \Phi_{\tilde{x}}(\mu) &= a \cdot \Phi_x(\mu) \end{aligned} \quad (5.10)$$

There are two contributions to the cepstrum, separated by applying the logarithm.

$$\begin{aligned} \log(\Phi_{\tilde{x}}(\mu)) &= \log(a) + \log(\Phi_{\tilde{x}}(\mu)) \\ \hat{c}_{\tilde{x}}[n] &= \int \log(\Phi_x(\mu)) \cdot e^{j\mu n} \,\mathrm{d}\mu + \int \log(a) \cdot e^{j\mu n} \,\mathrm{d}\mu \end{aligned} \quad (5.11)$$

Since $\log(a)$ is a constant and the Fourier series expansion of a constant yields a scaled version of the discrete delta impulse we can state that

$$\hat{c}_{\tilde{x}}[n] = \hat{c}_x[n] + \delta[n] \cdot \log(a). \quad (5.12)$$

We can see that the zero order coefficient is the only one being altered.

5.6 Constraints for the Cepstral Coefficients

This section will list some important properties of the cepstral coefficients which have to be fulfilled in order to yield a valid characteristic function. To recap, some basic properties of the characteristic function are given subsequently.

- $\Phi_x(\mu)$ has to be uniformly continuous everywhere.
- $\Phi_x(\mu)$ has to be 1 for $\mu = 0$.
- $\Phi_x(\mu)$ is bounded. On its entire domain $|\Phi_x(\mu)| \leq 1$.

Lemma 5.6.1. *The sum over all cepstral coefficients has to be zero:* $\sum_n \hat{c}_x[n] = 0$

Proof. As we know the characteristic function is one in its origin. By taking the logarithm of the statement $\Phi_x(\mu = 0) = 1$ we can rewrite this constraint so that $\log(\Phi_x(\mu = 0)) = 0$.

As we are now evaluating the characteristic function only in its origin we use the fact that $\mu = 0$, hence all the exponential terms evaluate to 1 and we can rewrite the Fourier series as follows

$$\log(\Phi_x(\mu = 0)) = \sum_n \hat{c}_x[n] \cdot e^{-jn\cdot 0} \tag{5.13}$$

$$0 = \sum_n \hat{c}_x[n]. \tag{5.14}$$

□

Looking at section 5.5 we can see that scaling a given PMF does not only violate the constraints of a PMF, but alters the zero order coefficient of the cepstrum. As a consequence these coefficients do not sum up to zero which violates lemma 5.6.1.

Lemma 5.6.2. *The cepstral coefficient of order zero has to be non positive:* $\hat{c}_x[0] \leq 0$

Proof. We know that $|\Phi_x(\mu)| \leq 1$, consequently this statement resolves to $\log(\Phi_x(\mu)) \leq 0$. As we know from Fourier series theory the zero order coefficient corresponds to the mean of the function we want to approximate.

Since we want to approximate $\log(\Phi_x(\mu))$, which is always non positive the mean of this function has to be non positive as well and hence we can say that $c_x[0] \leq 0$. □

Note that lemma (5.6.1) and (5.6.2) are valid constraints for all types of cepstra.

5.7 Presenting the Characteristic Function in Terms of its Cepstrum

Due to its importance in section 5.8 we will point out how the characteristic function can be expressed in terms of its cepstral coefficients. Looking at the power cepstrum, note that $\Phi_z(\mu)$ is the characteristic function of the difference of two i.i.d random variables again.

$$\Phi_z(\mu) = e^{\log|\Phi_x(\mu)|^2}. \tag{5.15}$$

Expressing the corresponding characteristic function $\log|\Phi_x(\mu)|^2$ in terms of its Fourier series representation leads to

$$\log|\Phi_x(\mu)|^2 = \sum_{n=-\infty}^{\infty} c_x[n] \cdot e^{-jn\mu}. \tag{5.16}$$

Consequently we can bring these two equations together. Furthermore we know from lemma 5.1.1 that the power cepstrum is even symmetric and thus we can rewrite the exponential function to cosine terms, so that

$$\Phi_z(\mu) = e^{c_x[0]} \cdot \prod_{n=1}^{\infty} e^{c_x[n]\cdot 2cos(\mu n)}. \tag{5.17}$$

A similar expression shall be derived to express $\Phi_x(\mu)$ in terms of its complex cepstrum. Let the characteristic function be given as

$$\log(\Phi_x(\mu)) = \sum_{n=-\infty}^{\infty} \hat{c}_x[n] \cdot e^{-jn\mu}. \tag{5.18}$$

Since the symmetry constraint does not hold anymore we cannot rewrite $\Phi_x(\mu)$ in its cosine terms. However we may start by collecting the corresponding positive and negative coefficients such that

$$\Phi_x(\mu) = e^{\hat{c}_x[0]} \cdot e^{\hat{c}_x[1]e^{j\mu}+\hat{c}_x[-1]e^{-j\mu}} \cdot e^{\hat{c}_x[2]e^{2j\mu}+\hat{c}_x[-2]e^{-2j\mu}} \cdots \tag{5.19}$$

Applying Euler's formula and collecting the terms, $\Phi_x(\mu)$ can be expressed in terms of its magnitude and phase contributions, so that

$$\Phi_x(\mu) = e^{\hat{c}_x[0]} \cdot \prod_{n=1}^{\infty} e^{(\hat{c}_x[n]+\hat{c}_x[-n])\cdot cos(\mu n)} \cdot \prod_{n=1}^{\infty} e^{-j(\hat{c}_x[n]-\hat{c}_x[-n])\cdot sin(\mu n)}. \tag{5.20}$$

5.8 Uniform Continuity of the Characteristic Function

Lemma 5.8.1. *Let the cepstral coefficients of a real random variable be such that only finitely many are non-zero and that they satisfy the conditions in lemma 5.6.1(the sum over all coefficients is zero) and lemma 5.6.2(the zero order coefficient is non positive). Then the reconstructed characteristic function of these coefficients is uniformly continuous.*

Proof. If we take a look at equation (5.18) and how we construct the characteristic function we can see that we are a looking at an exponential function $\Phi_x(\mu) = e^{f(u)}$ whereas $f(u)$ is a weighted sum of exponential terms contributed by the Fourier series. However according to equation (5.17) we can see, that $f(u)$ is a sum of cosine terms for the power cepstrum. Looking at equation (5.20), $f(u)$ is a sum of cosine and complex sine terms for the complex cepstrum.

In the later case we may look at $\Phi_x(\mu) = e^{g(u)} \cdot e^{j\cdot h(u)}$. Note that for a complex characteristic function to be uniformly continuous both the magnitude and the phase contribution have to ensure uniform continuity. We can rewrite the characteristic in the most general form to

$$\Phi_x(\mu) = e^{\hat{c}_x[0]} \cdot e^{(\hat{c}_x[1]+\hat{c}_x[-1])cos(\mu)} \cdot e^{-j\cdot(\hat{c}_x[1]-\hat{c}_x[-1])sin(\mu)} \cdots \tag{5.21}$$

Further investigations shall consider the representation of the characteristic function in terms of its cosine and sine terms. Both, the cosine and the sinus are bounded function, so is the exponential function of them. Hence we finally obtained the characteristic function as a product of bounded uniformly continuous functions. While the product of two uniformly continuous functions is not necessarily uniformly continuous, it is if the underlying functions are bounded uniformly continuous as we show in section 5.8.1. □

Hence we can state that the characteristic function is always uniformly continuous if we have a finite number of cepstral coefficients. Note that the converse is not true, that is having infinitely many coefficients does not violate the uniform continuity necessarily.

It shall be noted that we are always taking finitely many coefficients into account in practical examples. However, suppose the inverse Fourier transform to converge either absolutely or in a mean squared sense we know that the coefficients go to zero at least asymptotically. Therefore if the cepstral transformation exists, we can always truncate the cepstrum and thus limit the amount of cepstral coefficients, guaranteeing a sufficient representation.

5.8.1 Product of Uniformly Continuous Bounded Functions

We can show that the product of two, and consequently the product of finitely many, uniformly continuous and bounded functions, is uniformly continuous again. Starting with two uniformly continuous functions f and g which are defined on $\mathcal{X}$ and both bounded by L, and the product of these functions $h = f \cdot g$ we start with looking at two points $x_1, x_2 \in \mathcal{X}$. Looking at

$$|h(x_1) - h(x_2)| = |f(x_1) \cdot g(x_1) - f(x_2) \cdot g(x_2)|, \tag{5.22}$$

and expanding the above equation with $f(x_1) \cdot g(x_2) - f(x_1) \cdot g(x_2)$ we can apply the triangle inequality which leads to

$$|h(x_1) - h(x_2)| \leq |f(x_1)| \cdot |g(x_1) - g(x_2)| + |g(x_2)| \cdot |f(x_1) - f(x_2)|. \tag{5.23}$$

Due to f being uniformly continuous we know from the definition in [Rud76, page 90] that for every $\alpha > 0$ there exists a $\delta_1 > 0$ such that $|f(x_1) - f(x_2)| < \alpha$, for each pair of points x_1, x_2 for which $|x_1 - x_2| < \delta_1$ is satisfied.

Let $\alpha = \frac{\epsilon}{2 \cdot L}$ for reasons that will become obvious later on. Thus

$$|f(x_1) - f(x_2)| < \frac{\epsilon}{2 \cdot L} \tag{5.24}$$

for which $|x_1 - x_2| < \delta_1$. In a similar fashion there exists $\delta_2 > 0$ such that

$$|g(x_1) - g(x_2)| < \frac{\epsilon}{2 \cdot L} \tag{5.25}$$

for which $|x_1 - x_2| < \delta_2$.

Let us define $\delta = \min(\delta_1, \delta_2)$. We will now consider a pair of values such that $|x_1 - x_2| < \delta$ and both, inequality (5.24) and inequality (5.25) are fulfilled. Plugging in the boundary value L for both functions we can put all these constraints together into inequality (5.23) which subsequently resolves to inequality (5.26).

$$\begin{aligned} |h(x_1) - h(x_2)| &< L \cdot \frac{\epsilon}{2 \cdot L} + L \cdot \frac{\epsilon}{2 \cdot L} \\ &< \epsilon \end{aligned} \tag{5.26}$$

To conclude, for each $\epsilon > 0$ we can find a number $\delta > 0$ such that $|x_1 - x_2| < \delta$ and $|h(x_1) - h(x_2)| < \epsilon$, for every pair of points $x_1, x_2 \in \mathcal{X}$, which is the definition of uniform continuity.

5.9 Scaling Cepstral Coefficients

The question arises what happens to the characteristic function, respectively to the PMF, if we apply linear operations to the cepstral coefficients. We start with scaling the cepstral coefficients and looking if the resulting functions are still valid according to their properties and how they are altered.

The work of Lukacs served as base for the following observations. While the publications [Luk70] and [LL64] offer a detailed treatment of the subject we mainly used [Luk72] since this paper gives a clear and short overview and presents many interesting facts without giving the proofs in full detail. Instead, the paper points out to various literature and gives a neat summary.

We start by obtaining the inverse Fourier transform of the modified cepstral coefficients, that is

$$\log(\Phi_{\tilde{x}}(\mu)) = \sum_{n=-\infty}^{+\infty} a \cdot c_x[n] \cdot e^{-jn\mu}. \tag{5.27}$$

The scaling factor a is assumed to take only positive values. Since the cepstrum clearly vanishes for $a = 0$ a non-positive value would alter the sign of the cepstral coefficients and consequently violates Theorem 5.6.2 which states that the zero-order coefficient $c_x[0]$ has to be negative. Starting with a being a natural number we will obtain some first insights, expanding the observation for the scaling factor taking non integer values later on.

Due to the linearity of the Fourier transform we can pull out the constant factor a of the sum and rewrite equation (5.27) according to the definition of the power cepstrum in equation (3.5), so that

$$\log|\Phi_{\tilde{x}}(\mu)|^2 = 2a \cdot \log|\Phi_x(\mu)|. \tag{5.28}$$

Getting rid of the logarithm we can easily obtain the following result.

$$|\Phi_{\tilde{x}}(\mu)|^2 = |\Phi_x(\mu)|^{2a} \tag{5.29}$$

However if we skip taking the absolute value and look at the complex logarithm instead, equation (5.29) rewrites as follows

$$\Phi_{\tilde{x}}(\mu) = \Phi_x(\mu)^a. \tag{5.30}$$

Note that estimating the power cepstrum is looking at $|\Phi_x(\mu)|^2$. Taking the absolute value to the power of two obviously has an influence on the scaling. Therefore we neglect the power cepstrum from now on and are investigating the influence of scaling complex cepstral coefficients. However one could apply and extend the obtained results to the power cepstrum afterwards.

If $a \in \mathbb{N}$ equation (5.30) can be rewritten as the convolution of the PMF of the corresponding i.i.d RVs:

$$f_{\tilde{x}}(\xi) = (\underbrace{f_x * f_x * f_x * \cdots * f_x}_{a})(\xi). \tag{5.31}$$

Since we are basically summing a i.i.d random variables we can state that by increasing the scaling factor of the cepstral coefficients the according PMF is approaching some limiting distribution.

We are now looking at the case where a does not take integer values, however for simplification of the following analysis we restrict a to the rational numbers. If we have a factor $a \geq 1$ we can always rewrite the resulting characteristic function as the product of two characteristic functions as $\Phi_x^a = \Phi_x^\alpha \cdot \Phi_x^\beta$ with $\alpha \in \mathbb{N}$ and $\beta \in \mathbb{Q} \cap [0, 1)$. Due to $\beta = \frac{p}{q}$ being a rational number and $p, q \in \mathbb{N}$ we can further express the characteristic function by $\Phi_x^a = \Phi_x^\alpha \cdot (\Phi_x^{(1/q)})^p$. Having already looked at the influence of scaling the cepstral coefficients with an integer value we will now pay closer attention to the effect of scaling with the factor $1/q$.

Due to q being a rational number we can rewrite this part of the characteristic function as

$$\Phi_x^{1/q} = \sqrt[q]{\Phi_x}. \tag{5.32}$$

If the root of a characteristic function corresponds to a valid distribution, and how such a distribution may look like, can be answered if we take a closer look at literature. [Chu01, pages 252-255] points out how to correctly estimate the q-th order root of a characteristic function and how this is linked to infinitely divisible characteristic functions.

Hence we will take a closer look at the class of infinitely divisible characteristic functions and its existing subclasses. Figure 5.1 gives an overview of the existing classifications. The observations which are presented in [Luk72] give an excellent overview of the whole area, however the insights for stable and self decomposable distributions are restricted to continuous random variables. In [SVH79] the whole concept is expanded and similar properties are shown for discrete random variables.

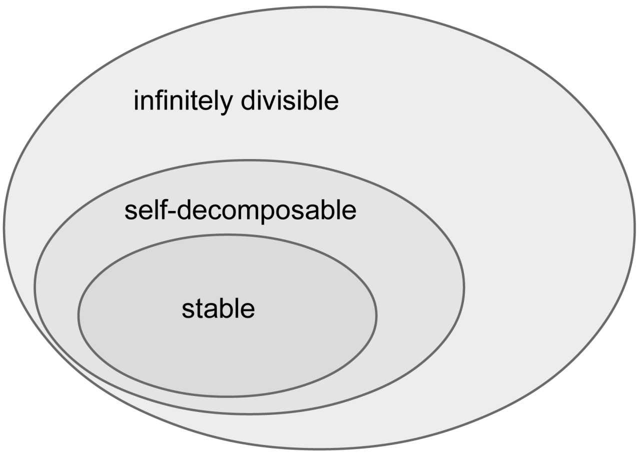

Figure 5.1: the class of infinitely divisible distributions and its subclasses

- An infinitely divisible characteristic function can be written as the product of n characteristic functions, that is

$$\Phi_x(\mu) = \Phi_{x,n}^n(\mu) \tag{5.33}$$

Similar to equation (5.30) and (5.31) we can rewrite this statement to a n-fold convolution of the corresponding PMF. Equivalently, an infinitely divisible random variable can be written as the sum of n i.i.d. random variables.

In other words if we take the n-th order root of $\Phi_x(\mu)$ the resulting characteristic function is valid for every n if and only if X is infinitely divisible. The exponential distribution, the χ^2- or the Pareto distribution are infinitely divisible for example.

- Characteristic functions which can be decomposed as is shown in equation (5.34) for $0 < c < 1$ and $\Phi_c(\mu)$ being a valid characteristic function, are called self decomposable functions.

  $$\Phi_x(\mu) = \Phi_x(\mu \cdot c) \cdot \Phi_c(\mu) \tag{5.34}$$

  This class is often named as the L-class in the literature. All distributions belonging to this class are infinitely divisible.

  Typical examples for the L-class are the Poisson- and the Frechet distribution. The discrete logarithmic distribution belongs to the class of the self-decomposable distributions as well.

- Stable distributions are distributions which have the property, that a linear combination of two random variables with a positive scaling factor and a real valued shift in location always yields the same distribution. It is known that stable distribution functions are unimodal. That is there exists exactly one point y along the ξ-axes, such that $f_x(\xi)$ is convex for $\xi < y$ and concave for $\xi \geq y$. Given the definition of a stable distribution the characteristic function can be given as follows:

  $$\Phi_x(\mu \cdot b)e^{i\cdot(c-c_1-c_2)\cdot\mu} = \Phi_x(\mu \cdot b_1) \cdot \Phi_x(\mu \cdot b_2), \tag{5.35}$$

  with $b, b_1, b_2 > 0$ and $c, c_1, c_2 \in \mathbb{R}$. It can be shown that all stable distributions are self decomposable and subsequently infinitely divisible.

  The Levy distribution, Cauchy distribution and the normal distribution are part of the stable class.

The discrete uniform distribution or the binomial distribution shall serve as examples for distributions which are not infinitely divisible.

Theorems 6.1 and 6.2 in [Luk72, page 24] present some interesting results concerning the decomposition of characteristic function which were first obtained by [Khi37]. These theorems state that every characteristic function can be decomposed into exactly two characteristic functions, whereas one function consists of indecomposable factors while the other function does not have any indecomposable factors at all. Moreover it is said that a characteristic function not having any indecomposable factor is infinitely divisible. Such functions are sometimes called to belong to the class I_0. This means that we can always rewrite a characteristic function as the product of an infinitely divisible function with another indecomposable characteristic function.

Looking back at the question of the existence of $\Phi_x^{1/q}$ as defined in equation (5.32) we can see that we are looking for a decomposition into equal functions. However, considering the above two theorems, it is a matter of fact, that $\Phi_x^{1/q}$ is a valid characteristic function if and only if Φ_x has no indecomposable factor. Subsequently Φ_x has to be infinitely divisible or the resulting characteristic function will be degenerate.

However, it has to be noted that $\Phi_x^{1/q}$ does not necessarily have to be identically distributed as Φ_x. For example one can take a look at the exponential distribution which can be approximated by a product of gamma distributions. Nevertheless the work of Levy and Cramer resulting in the Levy-Cramer-Theorem [Luk70] shows the decomposition of a normal distribution in factors which are normal distributed as well. Similar results were obtained for the Poisson distribution by Raikov [Rai37], which were concluded and generalized by Linnik [Lin57]. Linnik showed that there exist distributions like the normal or the Poisson distribution which are decomposed into factors of the same distribution type.

5.10 Right Hand Sided Probability Mass Functions

The idea of calculating minimum and maximum phase sequences is presented and many resulting properties can be found in [OS10]. In this work we try to gain some similar knowledge and investigate if it makes sense to introduce the whole terminology of this concept to the field of random variables. A sequence is minimum phase if it is

- real-valued
- stable
- causal (acausal for maximum phase)

Dealing with random variables, the PMF is always real-valued. Since a PMF sums up to one, every PMF is a stable sequence. We can see that the whole point of having a minimum phase property breaks down to the question if the PMF is right-hand sided. To emphasize that we are looking at distributions and not sequences we will not speak of minimum or maximum phase property, but left-hand and right-hand sided distributions instead.

Treating the PMF as a sequence we can reuse the fact that a minimum phase sequence leads to a right-hand sided complex cepstrum whereas a maximum phase sequence leads to a left-hand sided complex cepstrum [OS10, page 1013]. So just by looking at the complex cepstrum it is easy to check if the according PMF is right-hand sided.

Looking at the definitions in equation (3.5) and in equation (3.8) we see that the power cepstrum directly relates to the complex cepstrum via

$$c_x[n] = \hat{c}_x[n] + \hat{c}_x^*[-n], \tag{5.36}$$

where $\hat{c}_x^*[-n]$ is the complex conjugate of $\hat{c}_x[-n]$. Looking at the cepstrum of a right-hand sided PMF we know that $\hat{c}_x[n] = 0$ for n being negative. Such a simplification gives extra knowledge, so that we can reverse equation (5.36). Subsequently, the complex cepstrum is fully given by the power cepstrum in this case, where

$$\begin{aligned} \hat{c}_x[0] &= c_x[0]/2, \\ \hat{c}_x[n] &= c_x[n] \quad \text{for } n > 0, \\ \hat{c}_x[n] &= 0 \qquad\quad \text{for } n < 0, \end{aligned} \tag{5.37}$$

Hence the constraint of right-hand sidedness implies a relationship between phase and magnitude of the Fourier transform, known as the Hilbert transform relationship. Having such a relationship it is sufficient to evaluate the power cepstrum.

5.10.1 Decomposing the Complex Cepstrum

Since right hand sided random variables show nice properties, especially the close connection between the different types of cepstra, the question arises if we can decompose any given random variable $\mathbf{X}$ into two parts, where one part, $\mathbf{X}_{rhs}$ corresponds to a right hand sided PMF. As is shown in equation (5.39) such a decomposition corresponds into the addition of two RVs.

$$f_x(\xi) = (f_{x,rhs} * f_y)(\xi) \tag{5.38}$$

$$\mathbf{X} = \mathbf{X}_{rhs} + \mathbf{Y} \tag{5.39}$$

It shall be noted that such a deconvolution is on no account unique. Depending on the distribution of $\mathbf{Y}$,we may find different ways how $\mathbf{X}_{rhs}$ is distributed.

Right hand sided and left hand sided decomposition: The complex cepstrum can be decomposed into a right hand sided and a left hand sided contribution such that $\mathbf{X} = \mathbf{X}_{rhs} + \mathbf{X}_{lhs}$. This task is straightforward since we know that the contribution of the right hand sided distribution to the cepstrum is right hand sided as well, whereas the left hand sided distribution corresponds to the left hand side of the complex cepstrum. While such a decomposition is clearly always possible it shall be noted that the resulting random variables are violating some constraints in general and have degenerated distribution.

An example where such a decomposition fails can be seen by looking at a uniform distribution which is centered in the origin. Let $\mathbf{X}$ be uniformly distributed with $a = -30$ and $b = 30$. The according PMF is presented in figure 5.2. Note how the distributions takes values of $1/N = 1/61$ in the area of support and thus $\sum_\xi f_x(\xi) = 1$.

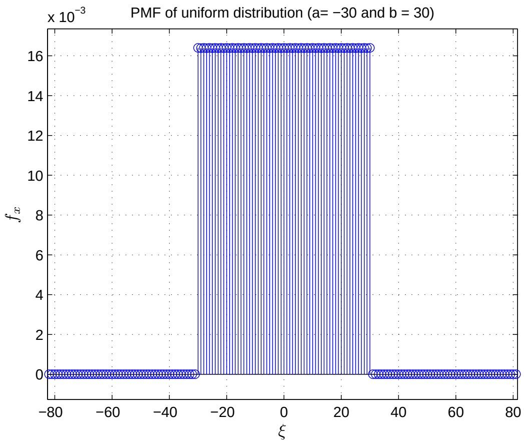

Figure 5.2: PMF of a valid uniform distribution, centered in the origin with $N = 61$

Applying the decomposition we can see in figure 5.3 that both, the left hand sided and right hand sided contributions violate the constraints for a valid PMF. Obviously these PMFs do take negative values. Moreover we can see that the constraint, that any distribution has to sum up to 1 on its entire domain, is violated. In fact the left hand sided distribution sums up to $\sum_\xi f_{x,lhs}(\xi) = 63.62$, whereas the right hand sided distribution sums up to $\sum_\xi f_{x,rhs}(\xi) = 0.16$.

Nonetheless we can see that the product of both contribution still sums up to one, that is

$$\sum_\xi f_{x,lhs}(\xi) \cdot \sum_\xi f_{x,rhs}(\xi) = 1, \tag{5.40}$$

and the convolution of these two seemingly random PMFs indeed yields the underlying uniform distribution with $a = -30$ and $b = 30$.

As a consequence it can be concluded, that a decomposition into right hand sided and left hand sided contribution is indeed not reasonable.

Hence we are trying if a **decomposition into a right hand side and an unconstrained contribution** suits our problem better. We decompose a PMF to $f_x = f_{x,rhs} * f_{x,uc}$ and consequently $\mathbf{X} = \mathbf{X}_{rhs} + \mathbf{X}_{uc}$. Note that $\mathbf{X}_{rhs}$ is differently distributed than $\mathbf{X}_{rhs}$ in the above decomposition. We start by estimating the complex cepstrum of the right hand sided contribution $\hat{c}_{x,rhs}$ first, which has the same magnitude of the Fourier transform as $\hat{c}_x$ has.

We shall treat the PMF as if it were right hand sided and estimate the corresponding cepstral coefficients $\hat{c}_{x,rhs}$ through estimating the power cepstrum first and applying the mapping introduced in section 5.10 to get the complex cepstrum. The complex cepstrum $\hat{c}_x$ is estimated as well. Since the convolution in equation (5.39) resolves to an addition in the cepstral domain,

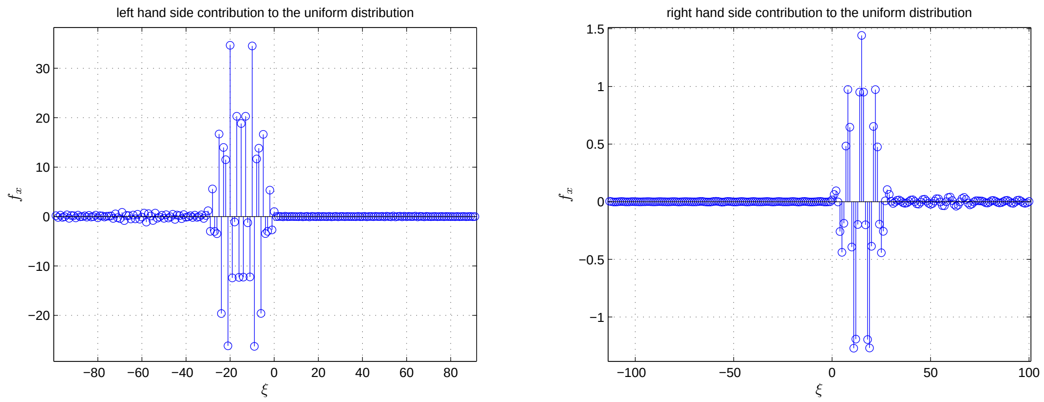

Figure 5.3: PMF of the left hand side and right hand side contribution to a centered uniform distribution with $n = 61$

the allpass contribution is given by

$$\hat{c}_{x,uc} = \hat{c}_x - \hat{c}_{x,rhs}. \tag{5.41}$$

The corresponding PMF can be estimated subsequently by applying the inverse cepstral transform. Figure 5.4 shows the principle approach to obtain the unconstrained and right hand sided contribution to the distribution.

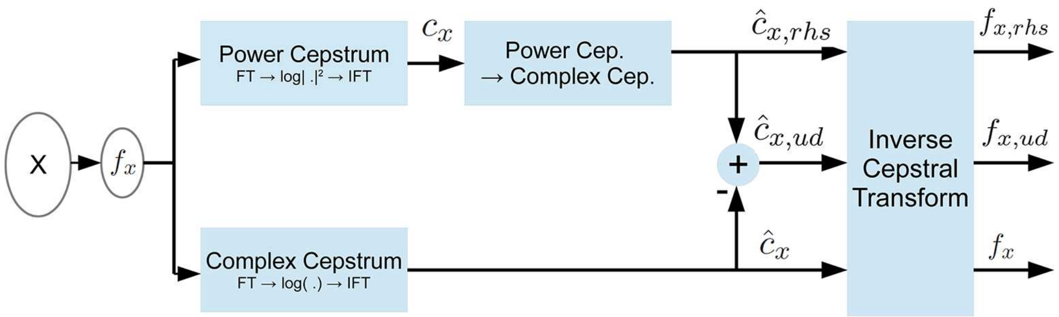

Figure 5.4: decomposition into right hand sided and unconstrained contributions, using the property to map the power cepstrum of a right hand side distribution to the complex cepstrum. The block "Power Cep.→ Complex Cep." flips the left hand side of the power cepstrum onto the right hand side to generate the complex cepstrum of a right hand side distribution. All other contributions are neglected by this procedure.

To see if this deconvolution yields two valid characteristic functions we will go through some experiments with some common distributions. First we start by applying this procedure to the **Poisson distribution**.
We start with the characteristic function which is given to equation (5.42)

$$\Phi_x(\mu) = e^{-\lambda(1-e^{-j\cdot\mu})} \tag{5.42}$$

We continue with estimating the complex cepstrum. Applying the inverse Fourier transform, we obtain

$$\hat{c}_x[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} -\lambda \cdot e^{j\mu n}\,\mathrm{d}\mu + \frac{1}{2\pi}\int_{-\pi}^{\pi} \lambda \cdot e^{j\mu n - j\mu}\,\mathrm{d}\mu, \tag{5.43}$$

whose solution is evidently zero for all but two values for n, which are given to $\hat{c}_x[0] = -\lambda$ and $\hat{c}_x[1] = \lambda$. Since the complex cepstrum is zero for $n < 0$ we see that the Poisson distribution fulfills already right hand side properties. Note that these results are in accordance to the mapping for right hand side distributions in section 5.10 and the previously made investigations regarding the Poisson distribution in section 4.4.2, since

$$\begin{aligned}
\hat{c}_x[0] &= c_x[0]/2 = -\lambda, \\
\hat{c}_x[1] &= c_x[1] \quad\;\; = \lambda, && (5.44)\\
\hat{c}_x[n] &= c_x[n] \quad\;\; = 0 \quad \forall\, n \in \mathbb{N}\backslash 0, 1 && (5.45)\\
\hat{c}_x[n] &= 0 \qquad\qquad\quad\; \forall n \in \mathbb{Z}_{<0}. && (5.46)
\end{aligned}$$

Figure 5.5 shows the power cepstrum of a Poisson distribution with $\lambda = 40$. Looking at

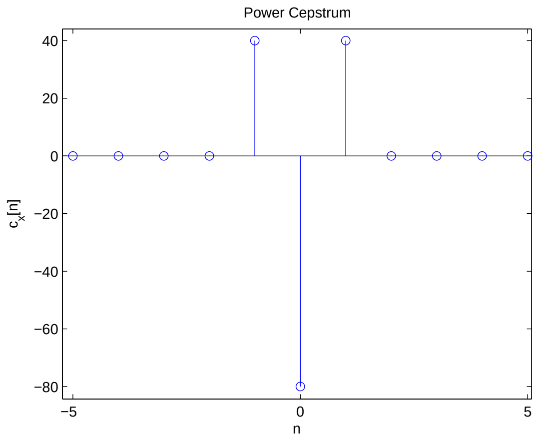

Figure 5.5: power cepstrum of a Poisson distribution with $\lambda = 40$

figure 5.6 we can see the complex cepstrum on the left side of the figure and the right hand sided complex cepstrum which was estimated applying equation (5.44). We can clearly see the connection between the power and the complex cepstrum and the obvious fact that the Poisson distribution already has right hand side property. A decomposition is clearly not necessary in this case, however applying the implementation of the algorithm shown in figure 5.4 presents the result shown in figure 5.7 for the resulting contributions. As we expected the right hand sided contribution is equal to the original PMF, whereas interestingly we can observe the unconstrained contribution to be a delta, which is apparently the neutral element with regard to the right hand side, unconstrained decomposition. It corresponds to the random variable $\mathbf{X}_{uc} \equiv 0$.

We shall refer to section 4.4.4 where some investigations with regard to the delta distribution $\delta(n)$ were made.

We know that the cepstral coefficients for $\delta(n)$ are zero for $n = 0$, such that the resulting complex cepstrum $\hat{c}_{x,poiss}[n] = \hat{c}_{x,rhs}[n] + \hat{c}_{x,uc}[n]$ resolves to $\hat{c}_{x,poiss}[n] = \hat{c}_{x,rhs}[n]$.

To conclude, we can say that $\delta(n)$ is an unconstrained distribution which apparently has a valid distribution with no influence on the complex cepstrum. Although clearly working in this special case, the question arises if such a deconvolution does yield two valid distributions for every given distribution.

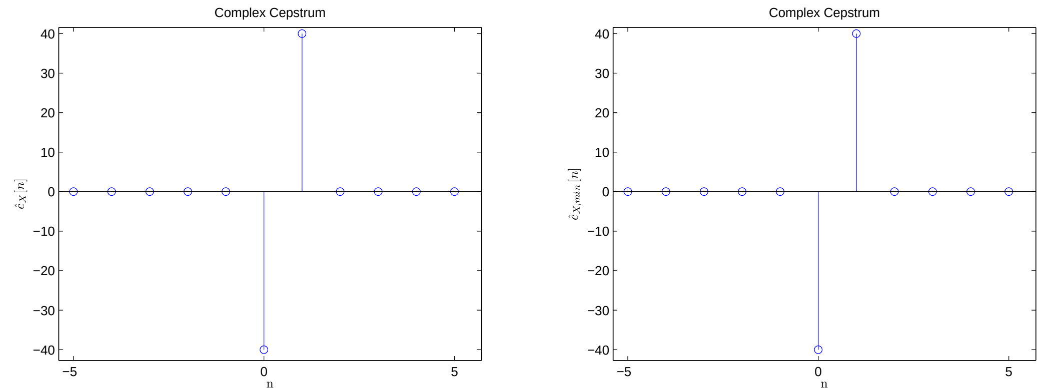

Figure 5.6: complex cepstrum of a poisson distribution and its according right hand sided contribution

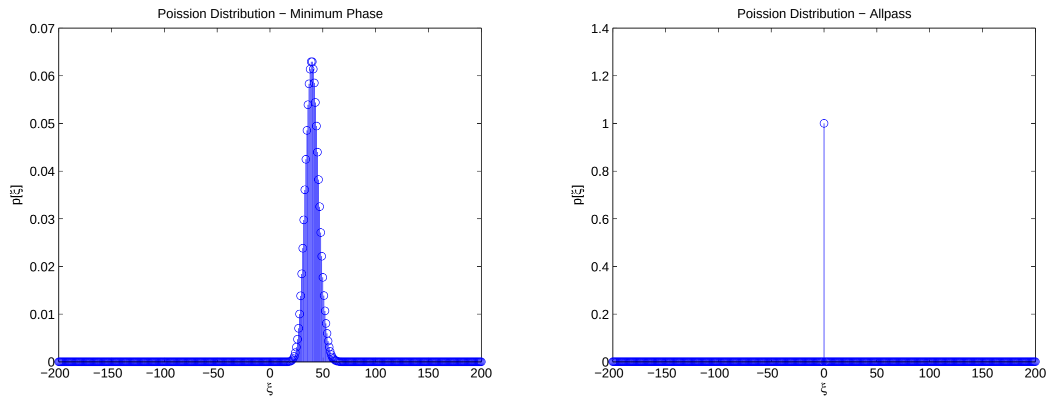

Figure 5.7: right hand side and unconstrained contributions of a poisson distribution with $\lambda = 40$

We will further apply this procedure to a distribution which is definitely not right hand sided, that is the **uniform distribution** centered in the origin. We start with the PMF shown in

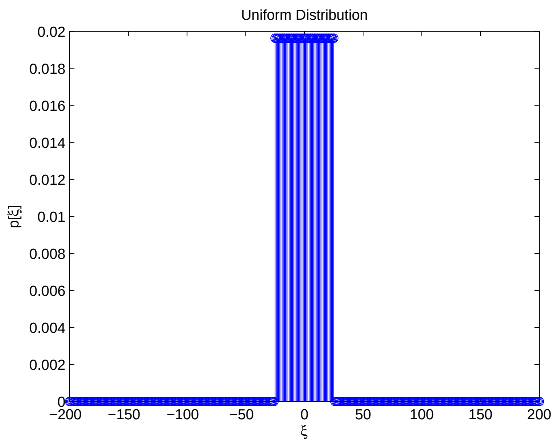

Figure 5.8: PMF of the uniform distribution with $a = -25$ *and* $b = 25$

figure 5.8 with $a = -25$ and $b = -25$ which is no right hand sided PMF since the PMF has values on the left hand side. Figure 5.9 shows the contributions after applying the right hand side unconstrained decomposition. We can see that the right hand sided contribution is just a

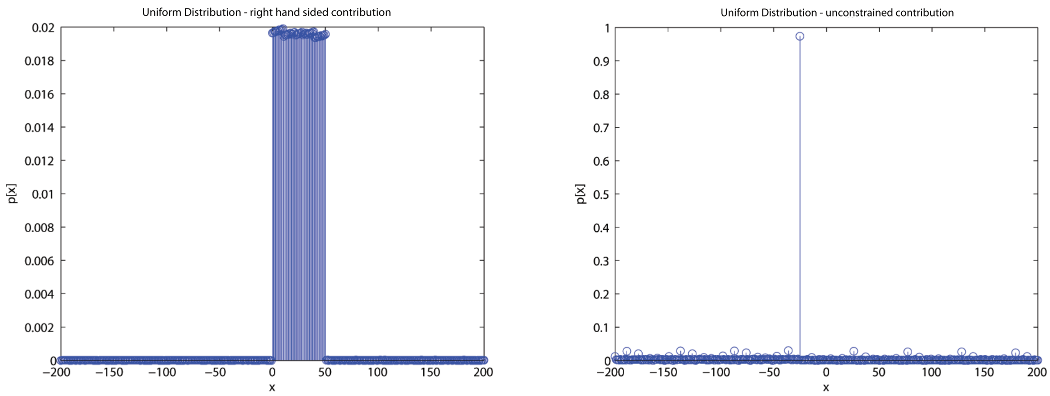

Figure 5.9: right hand side and unconstrained contribution to a uniform distribution with $a = -25$ and $b = 25$

shifted version of the original PMF so that $a = 0$ and $b = 50$. The unconstrained contribution is a delta impulse $\delta(n + 25)$, which is a valid PMF as well corresponding to the random variable $\mathbf{X}_{uc} \equiv -25$. Notable is that the shift of the impulse corresponds exactly to the shift of the right hand sided contribution.

As a consequence, it can be concluded that, in addition to all PMFs supported on the integers, the PMFs of all random variables bounded from below have an equivalent right hand sided representation.

Let $\mathbf{X}$ be a random variable bounded from below, there exists a constant C so that $\mathbf{X} \geq C$. That is equivalent to the sum of both random variables having positive support, $\mathbf{X} + C \geq 0$. As a matter of fact $\mathbf{X} + C$ is right hand sided. The according contribution by the unconstrained distribution is given to $\mathbf{X}_{uc} \equiv C$.

However the question still remains if such a decomposition can be applied to the more general case of a two-sided distribution.

A random variable which is not bounded from below, but still one sided should give some further insights. Let $\mathbf{X}$ be a left hand sided exponential decaying distribution. This is the **geometric distribution** with $0 < p \leq 1$. We shall present the corresponding characteristic function again:

$$\Phi_x(\mu) = \frac{p \cdot e^{-\alpha j \mu}}{1 - (1-p) \cdot e^{j\mu}}. \tag{5.47}$$

Thereby we modify the general definition as was already shown in section 4.4.5. Let the distribution be shifted into the origin, that is $\alpha = 0$ and $p = 0.2$.

The according distribution can be seen in figure 5.12. The power cepstrum and the resultant complex cepstrum of the right hand side distribution are given in figure 5.10. One can see how $c_x[n]$ for $n < 0$ are flipped around the origin. Figure 5.11 shows the actual complex cepstrum and the resulting complex cepstrum for the unconstrained contribution.

We can see that $\mathbf{X}_{rhs}$ is in fact just a mirrored version of $\mathbf{X}$. Consequently the complex cepstrum $\hat{c}_{x,rhs}[n] = \hat{c}_x[-n]$. $\mathbf{X}_{uc}$ seems to be some kind of left hand side exponential decaying distribution. However note that it is not a valid distribution anymore. The constraint for a PMF to be non-negative is clearly violated for $f_{x,uc}(\xi = 0)$. Both contributions can be seen in figure 5.13.

Note that the unconstrained distribution still sums up to one, so that a valid right hand side contribution is ensured.

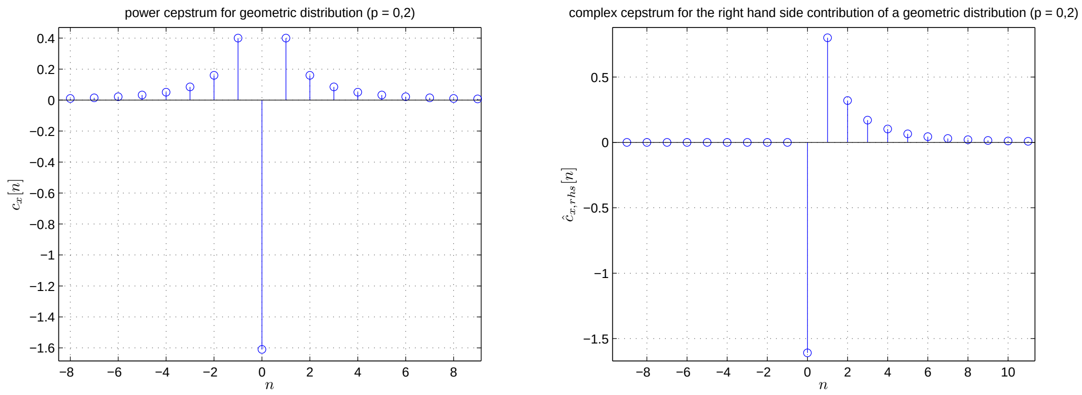

Figure 5.10: power cepstrum of the geometric distribution and the associated complex cepstrum of the right hand side contribution

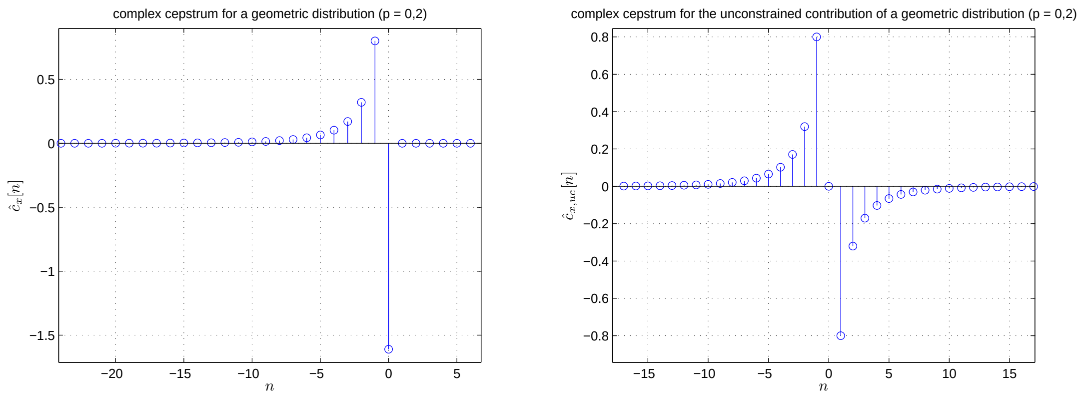

Figure 5.11: complex cepstrum of the geometric distribution and the unconstrained contribution

Nonetheless we can conclude that the decomposition of a random variable, which is not bounded from below, into its right hand side and unconstrained contribution does not lead to valid distributions in general. There may exist an unconstrained contribution $\mathbf{X}_{uc}$ which is not valid, such that validity of the right hand side contribution is ensured. Therefore, we choose $\mathbf{X}_{rhs}$ so that $\mathbf{X}_{rhs} + \mathbf{X}_{uc} = \mathbf{X}$.

Let us further investigate the characteristic function of the unconstrained contribution $\Phi_{x,uc}(\mu)$. As a matter of fact this decomposition leads to the magnitude being one everywhere, $|\Phi_{x,uc}(\mu)| = 1$, and the phase being odd symmetric. Thus $\Phi_{x,uc}(\mu)$ is solely determined by its phase. Bochners theorem gives a necessary and sufficient condition for an arbitrary function to be the characteristic function of a valid random variable. That is the characteristic function has to be positive definite and be continuous in the origin with $\Phi_x(\mu = 0) = 1$

Note that $\Phi_{x,uc}(\mu = 0) = 1$, since $\sum_n \hat{c}_{x,uc}[n] = 0$ and thus Lemma 5.6.1 is fulfilled.

$\Phi_{x,uc}(\mu)$ has to be continuous in the origin. If finitely many cepstral coefficients are considered and we take Lemma 5.8.1 into account, $\Phi_{x,uc}(\mu)$ is in fact uniformly continuous on the whole domain.

By the principle of exclusion $\Phi_{x,uc}(\mu)$ has to violate the positive definiteness consequently.

To conclude, a decomposition into right hand side and unconstrained contribution can be guaranteed to lead to valid random variables only if we are decomposing a distribution which is bounded from below.

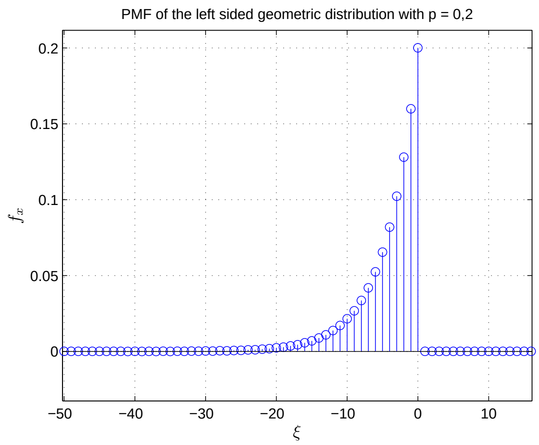

Figure 5.12: PMF of a geometric distribution with $p = 0.2$ which is bounded from above

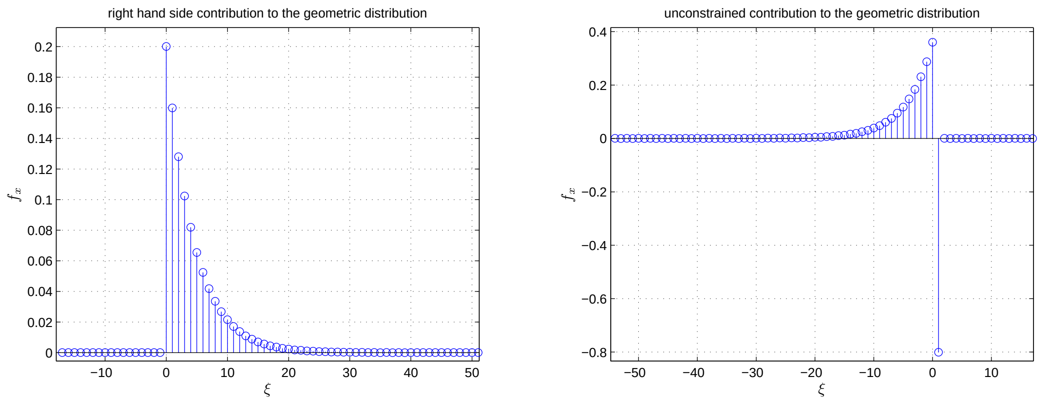

Figure 5.13: PMF of the right hand side and the unconstrained contribution to the geometric distribution. Note how the PMF of the unconstrained distribution takes negative values

6
Hypothesis Tests using Cepstral Coefficients

A possible application of the cepstrum is hypothesis testing. Looking at data obtained from a given observation or study, the question behind hypothesis testing is, if the distribution of the data occurred randomly, or if there was any statistically significant influence. A goodness of fit test is introduced in this chapter and compared with examples where the χ^2 test for goodness of fit is applied.

The χ^2 test assumes the data to be distributed according to the null hypotheses. If the level of significance exceeds a certain threshold α, which is typically chosen to be $\alpha = 0.05$, the observed data is not distributed according to the assumption.

Since the observed data is compared to the expected data we are applying a normalized mean squared error criterion. Let p_n be the observed probability for n and P_n be the expected probability for n. N being the number of bins we estimate the variation χ^2 to

$$\chi^2 = \sum_{n=1}^{N} \frac{(p_n - P_n)^2}{P_n}. \tag{6.1}$$

The χ^2 distribution is the sum of squared standard normal i.i.d random variables. Therefore it suits for describing equation (6.1). For more details regarding the χ^2 distribution and the according hypothesis test [PP02, page 89ff, page 361ff] offers thorough explanations.

6.1 Hypothesis Test for Poisson Distributions

In this section we start with giving a few examples for applications of Poisson distributions. Considering the constraints for cepstral coefficients of Poisson distributions we introduce a goodness of fit test and compare it to the well known χ^2 test.

From chapter 4 we know that for every Poisson distribution the complex cepstrum has only two coefficients which are not equal to zero, that are $\hat{c}_x[0]$ and $\hat{c}_x[1]$.

Given some real life examples for applying Poisson distributions we want to validate this fact. The Poisson distribution is used to determine the probability for a given number of events to occur in a certain time span. These events have to be independent of previous occurrences. Since we are talking about rare events, the Poisson distribution is sometimes known as law of small numbers. Estimating the probability for men of the Prussian Army to be kicked to death by horses was one of the first applications of a Poisson distribution to real life events. The following probability table is taken from [Bor98, page 66]. Table 6.1 presents the probability for every event. That is how many men were killed per year per corps. Both the observed number of occurrences p_n and the expected number of occurrences according to the Poisson distribution P_n are listed.

Number of Deaths	p_n	P_n
0	109	108.7
1	65	66.3
2	22	20.2
3	3	1
4	1	0.6
5 and more	0	0.1

Table 6.1: number of men kicked to death per year per corps

Looking at this table one can immediately see that the observations match the expectations very well. Figure 6.1 shows the cepstral coefficients for this example.

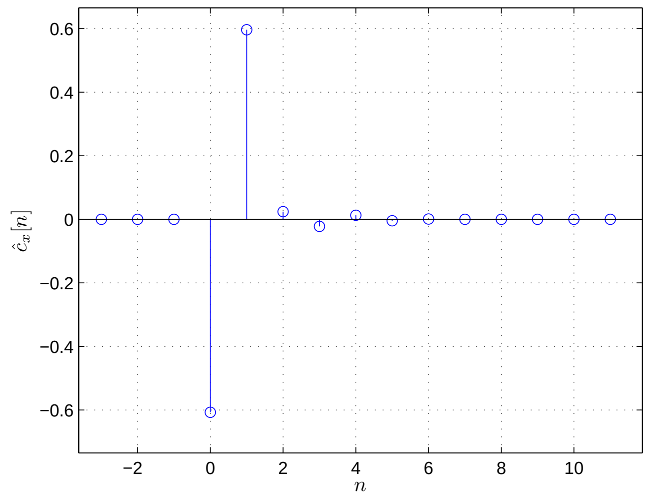

Figure 6.1: cepstral coefficients for the observation of the prussian army

One can see the presence of both coefficients. Due to non perfect fit a slight increase of $\hat{c}_x[3]$ to $\hat{c}_x[5]$ can be observed. However due to $\hat{c}_x[1]$ being at least 25 times greater than the side coefficients we may neglect them and conclude that this example is indeed a good example for a Poisson distribution. Furthermore we know from [Bor98] that the the Poisson parameter of the fitted distribution is $\lambda = 0.61$.

According to section 4.4.2 we can directly obtain this coefficient from looking at $\hat{c}_x[0]$. In this example we can obtain $\hat{c}_x[0] = 0.607$, which is indeed close to λ.

Many more applications for the Poisson distribution can be found in [Ger55] and the references therein. Some further examples shall be presented, to show the application of the cepstrum as a hypothesis test. The probability table 6.2 presents the number of storms occurring at a weather station per year and its theoretical expectations. There are some slight deviations between observations and expectations. Nevertheless applying a Poisson distribution intuitively seems to bee a good choice. However we will evaluate the χ^2 and compare it to the introduced hypothesis

Number of Rainstorms	p_n	P_n
0	102	99.3
1	114	119.1
2	74	71.6
3	28	28.7
4	10	8.6
5	2	2.0
6 and more	0	0.7

Table 6.2: occurrence of rainstorms per weather station per year

test. Evaluating equation (6.1) with the values taken from table 6.2 resolves to

$$\chi^2 = 1.317. \tag{6.2}$$

As a next step we look up a χ^2 distribution table to evaluate the goodness of fit. Note that we are looking at the the entries with 6 degrees of freedom, since we have 7 classes in our probability table. By convention the level of significance is chosen to be $\alpha = 0.05$. With this parameters the table yields $\chi^2_{0.05} = 12.059$.

Since $1.317 < 12.059$ we accept the hypothesis and state the Poisson distribution is indeed a suitable choice.

In comparison we evaluate the complex cepstrum and take a look at the coefficients in figure 6.2. An increase of $\hat{c}_x[2]$ due to the mismatch can be seen in figure 6.2 . However the first

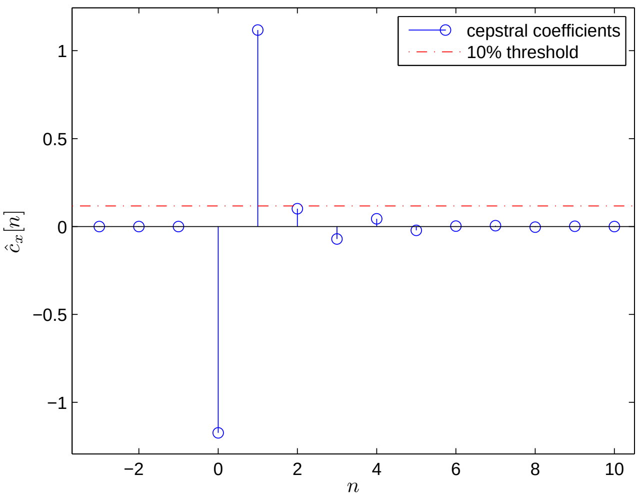

Figure 6.2: cepstral coefficients for the observation of rainstorms

two cepstral coefficients are still dominant. To enable a comparable goodness of fit criterion we introduce a threshold $\delta = 10\% \cdot \lambda$. We state that as long as all cepstral coefficient except $\hat{c}_x[0]$ and $\hat{c}_x[1]$ stay below this threshold, that is $\hat{c}_x[n] < \delta$ the observations can be modeled with a poisson distribution. However it shall be mentioned that this criterion seems to be more rigid than the χ^2 criterion.

Finally an example where the Poisson distribution is indeed a bad fit shall be presented. Table 6.3 shows the number of cars which make a right turn during a 3 minutes time interval at an observed crossroad. This traffic phenomena is described with a Poisson distribution with $\lambda = 3.89$. The according expectations are also given in table 6.3 The first inspection could

Number of cars making right turns	p_n	P_n
0	14	6.1
1	30	23.7
2	36	46.2
3	68	59.9
4	43	58.3
5	43	45.4
6	30	29.4
7	14	16.4
8	10	8.0
9	6	3.4
10	4	1.3
11	1	0.6
12	1	0.3
13 and more	0	1.0

Table 6.3: observation of a crossroad

indicate a suitable fit, however we can see that this fit is not acceptable with a significance level of $\alpha = 0.05$. The χ^2 test yields $\chi^2 = 24.7$, such that with 13 degrees of freedom $\chi^2 > \chi^2_{0.05} = 22.36$. Hence the fit seams to be not acceptable and there is a strong evidence, that the observations are not randomly occurring. Looking at figure 6.3 we can see a dramatic increase of various

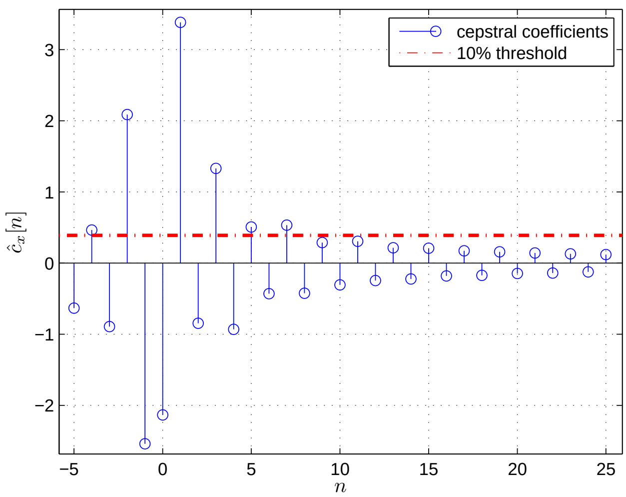

Figure 6.3: cepstral coefficients for the observation right hand turns

cepstral coefficients. The second order coefficient clearly exceeds the 10% threshold $\hat{c}_x[n] > \delta$, furthermore there is a significant increase of even higher order coefficients. Hence we can varify the result of the χ^2 test by looking at the cepstrum.

Figure 6.4 shows the observations against the expectations. However it shall be noted, that the observations may seem non-random. Changing the observation time interval can change the distributions remarkably, so that a good fit could possibly be achieved with choosing a different interval.

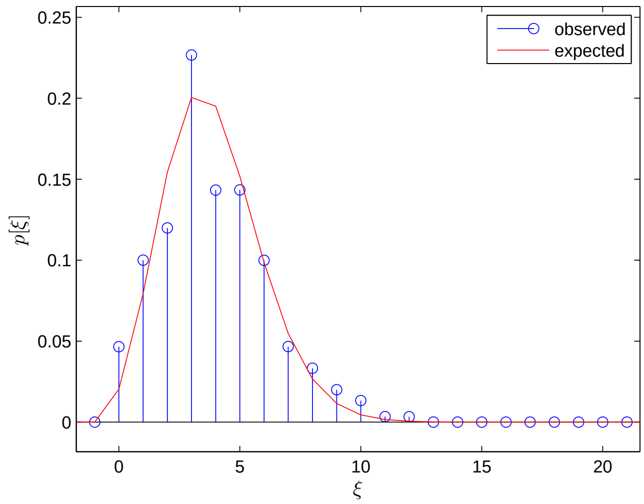

Figure 6.4: the poisson distribution with $\lambda = 3.893$ is plotted in red, whereas the observations are stemmed in blue

6.2 Hypothesis Test for Geometric Distributions

We will describe the main field of applications for the geometric distribution in this section. Some experiments are presented with the according probability tables. Based on these tables we will consider the constraints for the cepstral coefficients of a geometric distribution.

It is necessary to shift the distribution into the origin, or subtract the influence of the delta distribution, for the complex cepstrum to be well behaved (see section 4.4.5). We opt for presenting a general applicable procedure. Thus we shall use the power cepstrum throughout this section.

A power cepstrum is associated with a geometric distribution if and only if $c_x[0] < 0$ and all other coefficients $c_x[n] > 0$ are exponentially decaying. The χ^2 seems to be not appropriate for evaluating the goodness of fit. In [KBB92] the goodness of fit for geometric distributions is primarily evaluated by looking at the histograms of the data. Some tests for the goodness of fit of a geometric distribution are evaluated in [BCG11], however they are rather complex and beyond the scope of this work. Therefore we try to give an examples where the geometric distribution certainly does fit very well and an examples where the fit is rather poor. Subsequently the cepstral coefficients will be evaluated and compared to the PMF.

Let us look at the classic example of tossing a coin. Suppose a coin is tossed 1000 times. Table 6.4 presents the observations p_n of how often the coin lands on its tail for the first time after exactly n trials. Note that the probability for success $p = 0.5$ is the same for every trial. P_n lists the expected value assuming that a geometric distribution with $p = 0.5$ is the underlying distribution.

Table 6.4 shows some slight deviations between observations and expectations due to the limited amount of repetitions. However the PMF in figure 6.5 shows roughly the behavior of a geometric distribution. Looking at the power cepstrum in figure 6.5 we can validate this assumption since the mentioned constraints for the cepstrum to be associated to a geometric distribution are met. Furthermore we can see that $p = c_x[1] = 0.47$ which matches the theoretical value of $p = 0.5$ quite well.

To give a counterexample let us think of rolling the dice. The question is how many tries it takes to roll a one. Thus the probability is given to be $p(1) = 1/6$. This experiment shall be

Number of Tosses until Head	p_n	P_n
1	505	500
2	237	250
3	131	125
4	60	62.5
5	33	31.25
6	13	15.63
7	13	7.81
8	5	3.91
9 and more	3	1.15

Table 6.4: tossing a coin until heads

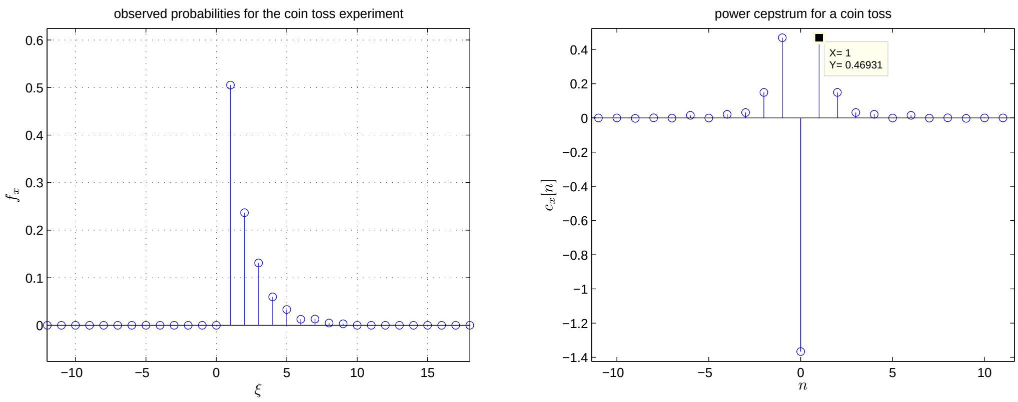

Figure 6.5: PMF and the according power cepstrum for the coin toss experiment

repeated 10000 times. However, let us suppose that the participant always uses a fake dice for his third roll, where the probability to throw a one is given to $p(1) = 2/6$. The observations p_n and the expectations for a fair game P_n are given in table 6.5. Note that we group together the occurrences of 15 and more trials in table 6.4. However in the simulation we treat each case separately.

The outcome of the experiment is presented in figure 6.6. Looking at the cepstral coefficients in this figure we can see how some coefficients take negative values. Thus we can assume that we are not looking at a geometric distribution.

Throughout some experiments a **threshold** $\delta = -0.1$ seemed to be a sufficient criterion for the goodness of fit. Thus we state that as soon as any $c_x[n \neq 0] < \delta$ the observation should not be modeled with a geometric distribution.

Number of Rolls until 1	p_n	P_n
1	1736	1666.7
2	1386	1388.9
3	2306	1157.4
4	960	964.5
5	769	803.8
6	666	669.8
7	543	558.2
8	480	465.1
9	400	387.6
10	355	323
11	264	269.2
12	190	224.3
13	200	186.9
14	135	155.8
15 and more	763	778.9

Table 6.5: roll of a dice

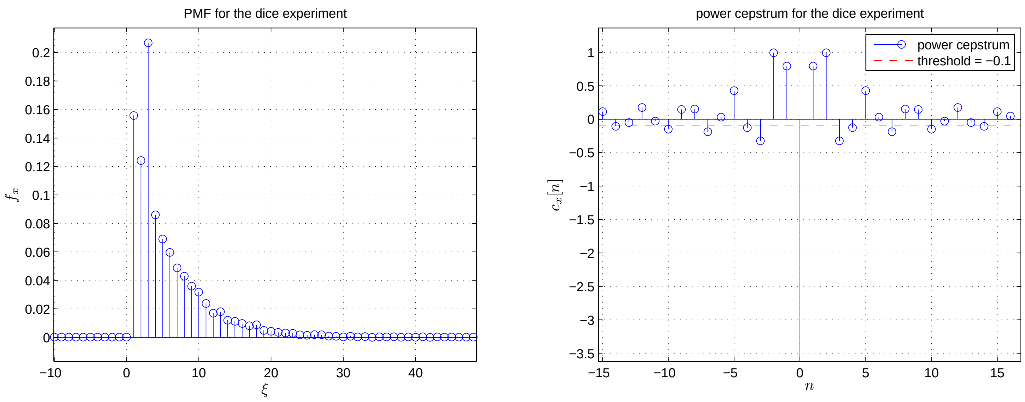

Figure 6.6: PMF and the according power cepstrum for the dice experiment

7
Conclusion and Outlook

In this work we have given the motivation to introduce and apply the cepstrum for describing random variables. We started by giving a short overview of moments and cumulants and some approximations using these existing descriptors. The cepstrum was introduced subsequently. After presenting the different definitions of the cepstrum we opted for using the complex cepstrum for the bigger part of the work.

The cepstrum analysis was applied for describing random variables. A direct relation between the cepstral moments and the cumulants was given to $\hat{m}_k = \kappa_k$. Finally we discussed the cepstral coefficients for some distributions. It is remarkable how some of these random variables can be interpreted by looking at the cepstrum.

We continued by giving basic and advanced properties for the cepstrum in the context of applying it to random variables.

It could be shown that the cepstrum can be a practical tool for hypothesis testing, given the fact that one knows how the cepstrum is constrained for the according random variable. Basically a detailed survey was made to get a good understanding of probability theory and of homomorphic transformations. The benefits and disadvantages of bringing these two fields of research together are presented in this work. To conclude, the cepstrum is not superior to the cumulants, however some applications were given where the cepstrum has certain advantages.

As there are still many open questions, we shall give a short outlook. We have already given constraints for the cepstrum in this work. As there is very limited knowledge with regard to the behavior of cumulant series, one could apply the direct relationship between cumulants and the cepstral moments to develop some constraints for the cumulants.

Taking a closer looker at the investigation with regard to scaling of the cepstrum we shall note that every infinitely divisible characteristic function can be represented in Levy Khinchine form. It seems feasible that we could obtain some necessary and sufficient conditions for a cepstrum to belong to a infinitely divisible random variable.

There seems to be a relationship between the cepstral moments and Good's formula, see [Goo75]. If we could prove this connection we could possibly get further insights into the cepstrum.

In this work we have only been looking at discrete random variables which seems suitable, given the basic motivation for looking at the Fourier series to exploit the periodicity of its characteristic functions. However it could be interesting to generalize the cepstrum for continuous random variables. Note that the cepstrum will be a continuous function as well. Sampling such a continuous cepstrum would therefore correspond to quantization of the random variable itself.

It would be of great interest to further investigate the effect of truncating the cepstral coefficients. We could possibly get a better representation of a given random variable considering a certain amount of cepstral coefficients, as opposed to the same amount of cumulants. However it is a matter of fact that the effect of truncating the series will also depend on the observed random variable.

In particular, based on the cepstral coefficients of an observed distribution, one could fit a Poisson, a geometric, or any other distribution to the observations. It would be interesting to see if such approximations are optimal in a well-defined sense.

A crucial question, however, is the existence of the cepstrum for two sided probability mass functions. It would be of great interest to define some general classes of characteristic functions, for which existence of the cepstrum is guaranteed.

Bibliography

[BC09] Brown, J. W. ; Churchill, R. V.: *Complex Variables and Applications*. McGraw Hill, 2009

[BCG11] Bracquemond, C. ; Cretois, E. ; Gaudoin, O.: A comparative study of goodness-of-fit tests for the geometric distribution and application to discrete time reliability. (2011)

[BH85] Baumann, K. ; Hegerfeldt, G. C.: A Noncommutative Marcinkiewicz Theorem. In: *Publications of the Research Institute for Mathematical Sciences* 21 (1985), S. 191–204

[BHT63] Bogert, B. P. ; Healy, M. J. R. ; Tukey, J. W.: The quefrency analysis of time series for echoes: cepstrum, pseudo-autocovariance, cross-cepstrum, and saphe cracking. In: Rosenblatt, M. (Hrsg.): *Proceedings Symp. Time Series Analysis*, 1963, S. 209–243

[BM98] Blinnikov, S. ; Moessner, R.: Expansions for nearly Gaussian Distr. In: *Astronomy and Astrophysics Supplement Series* 130 (1998), S. 193–205

[Bor98] Bortkiewicz, L.: *Das Gesetz der kleinen Zahlen*. BG Teubner, 1898

[Chu01] Chung, K. L.: *A Course in Probability Theory*. Academic Press, 2001

[Cra57] Cramer, H.: *Mathematical Methods of Statistics*. Princeton University Press, 1957

[Fur86] Furui, S.: Speaker-independent isolated word recognition based on dynamics-emphasized cepstrum. In: *IEICE Transactions* 69 (1986), S. 1310–1317

[Ger55] Gerlough, D. L.: Use of Poisson Distribution in Highway Traffic. In: *Eno Foundation for Highway Traffic Control* (1955)

[Goo75] Good, I. J.: A new formula for cumulants. In: *Mathematical Proceedings of the Cambridge Philosophical Society* 78 (1975), S. 333–337

[Hal40] Haldane, J. B. S.: The Cumulants and Moments of the Binomial Distribution, and the Cumulants of chi-squared for a (n x 2)-Fold Table. In: *Biometrika* 31 (1940), S. 392–396

[KBB92] Kaminsky, F. C. ; Benneyan, R. D. ; Burke, R. J.: Statistical control charts based on a geometric distribution. In: *Journal of Quality Technology* 24.2 (1992), S. 63–69

[Khi37] Khinchine, A.: Contribution a l'arithmetique des lois de distribution. In: *Bull. Math. Univ. Moscou* 1 (1937), S. 6–17

[KY92] Khare, Anil ; Yoshikawa, Toshinori: Moment of cepstrum and its applications. In: *IEEE Transactions on Signal Processing,* 40 (1992), Nr. 11, S. 2692–2702

[Lin57] Linnik, Y.V.: On the Decomposition of the Convolution of Gaussian and Poissonian Laws. In: *Theory of Probability & Its Applications* 2 (1957), S. 31–57

[LL64] Lukacs, E. ; Laha, R. G.: *Applications of Characteristic Functions.* Griffin, 1964

[Luk70] Lukacs, E.: *Characteristic functions. 2nd rev.* Griffin, 1970

[Luk72] Lukacs, E.: A Survey of the Theory of Characteristic Functions. In: *Advances in Applied Probability* 4 (1972), S. 1–38

[Mat99] Mattner, L.: What are Cumulants. In: *Documenta Mathematica* 4 (1999), S. 601–622

[Nol64] Noll, A. M.: Short Time Spectrum and Cepstrum Techniques for Vocal Pitch Detection. In: *The Journal of the Acoustical Society of America* 36 (1964), S. 296

[Opp65] Oppenheim, A. V.: *Superposition in a class of nonlinear systems.*, Massachusetts Inst of Tech Camridge Research Lab of Electronics, Diss., 1965

[OS04] Oppenheim, A. V. ; Schafer, R.W.: From Frequency to Quefrency: A History of the Cepstrum. In: *IEEE Signal Processing Magazine* 21 (2004), Nr. 5, S. 95–106

[OS10] Oppenheim, A. V. ; Schafer, R. W.: *Discrete-Time Signal Processing.* Bd. 3rd edtition. Pearson Education, Limited, 2010

[OSB99] Oppenheim, A. V. ; Schafer, R. W. ; Buck, J. R.: *Prentice-Hall signal processing series.* Bd. 2nd edtition: *Discrete-time signal processing*. Prentice Hall, 1999

[Pet75] Petrov, V. V.: *Sums of independent random variables.* Springer Verlag, 1975

[PP02] Papoulis, A. ; Pillai, S. U.: *McGraw-Hill electrical and electronic engineering series.* Bd. 4th edition: *Probability, random variables, and stochastic processes*. McGraw-Hill, 2002

[Rai37] Raikov, D.: On the decomposition of Gauss and Poisson laws. In: *Academy of Sciences of USSR* 14 (1937), S. 9–11

[RS00] Rota, G. C. ; Shen, J.: On the combinatorics of cumulants. In: *Journal of Combinatorial Theory* Series A 91.1 (2000), S. 283–304

[Rud76] Rudin, W.: *Principles of mathematical analysis.* Third. New York : McGraw-Hill Book Co., 1976. – x+342 S. – International Series in Pure and Applied Mathematics

[Sch81] Schroeder, M. R.: Direct (Nonrecursive) Relations Between Cepstrum and Predictor Coefficients. In: *Speech and Signal Processing, IEEE Transactions on Acoustics* 29 (1981), S. 297–301

[STTI75] Stockham, Jr ; Thomas, G ; Thomas, M. C. ; Ingebretsen, R. B.: Blind deconvolution through digital signal processing. In: *Proceedings of the IEEE* 63 (1975), S. 678–629

[SVH79] Steutel, F.W. ; Van Harn, K.: Discrete Analogous of Self-Decomposobility and Stability. In: *The Annals of Probability* 7 (1979), S. 893–899

[Tri77] Tribolet, J. M.: A new phase unwrapping algorithm. In: *IEEE Transactions on Acoustics, Speech and Signal Processing* 2 (1977), S. 170–177

[Ulr71] ULRYCH, T. J.: Application of homomorphic deconvolution to seismology. In: *Geophysics* 36(4) (1971), S. 650–660

[WB85] WATT, T. L. ; BEDNAR, J. B.: Calculating the Complex Cepstrum without Phase Unwrapping or Integration. In: *IEEE Transactions on Acoustics, Speech and Signal Processing* 4 (1985), S. 1014–1017

[WKL96] WIDROW, B. ; KOLLAR, I. ; LIU, M. C.: Statistical theory of quantization. In: *Instrumentation and Measurement, IEEE Transactions on* 45 (1996), S. 353–361