Analytic, numeric and algorithmic aspects of orthogonal polynomials

Hans-Jürgen Fischer

October 7, 1998

Preface

Orthogonal polynomials are traditional objects of mathematical analysis. The so-called classical orthogonal polynomials were investigated by mathematicians like Legendre, Jacobi, Hermite, Laguerre, Chebyshev and others. Many of their properties were known in the last century. For Legendre polynomials P_n, e.g., we know a Rodrigues' formula

$$P_n(x) = \frac{1}{2^n\, n!}\, \frac{d^n}{dx^n}\, (x^2-1)^n,$$

a differential equation

$$(1-x^2)P_n''(x) - 2xP_n'(x) + n(n+1)P_n(x) = 0,$$

a three-term recurrence relation

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x),$$

various generating functions and (closely related to them) integral representations. Similar formulas are known for other classical orthogonal polynomials. Especially, all of them satisfy a second-order differential equation. This makes them members of well-known families of special functions, allowing the investigation of their asymptotic properties.

The classical orthogonal polynomials can be characterized by orthogonality with respect to a inner product, which for polynomials p and q is given by the integral

$$(p, q) = \int p(x)\, q(x)\, w(x)\, dx.$$

The *weight function* w must have a special form, e.g., it is 1 on $[-1, 1]$ and zero outside that interval for Legendre polynomials.

While the classical polynomials occured first in series expansions, their orthogonality properties were found later. Of course, we can consider polynomials orthogonal with respect to this inner product for an arbitrary positive weight function (having finite moments $\mu_k = \int x^k w(x)\, dx$). The investigation of those *general orthogonal polynomials* was started by works of Gauss, Chebyshev and Stieltjes. Naturally, they didn't use the termin — the idea of a geometry of function spaces appeared much later. Instead, their starting point was the transformation of the (formal!) power series

$$\int \frac{w(x)\,dx}{z-x} \sim \sum_{k=0}^{\infty} \frac{\mu_k}{z^{k+1}}$$

into a (formally infinite) continued fraction. The denominators of the finite convergents of this continued fraction are polynomials, then. As it turned out, formulas of Rodrigues type or differential equations are valid essentially for the classical polynomials, only. The same is true for the existence of generating functions of various types. A three-term recurrence formula of the above form, however, occurs naturally from the theory of continued fractions.

This is valid not only in the case mentioned above, but extends even to an inner product given by

$$(p,q) = \sum_i \sigma_i\, p(\tau_i)\, q(\tau_i),$$

where $\{\tau_i\}$ is a finite or countable set of nodes and σ_i are the corresponding positive weights (again, all moments $\mu_k = \sum_i \sigma_i\, \tau_i^k$ have to be finite). Indeed, polynomials of this type were investigated by Chebyshev in the context of least-squares approximation of discrete data. It is remarkable that Stieltjes in his famous 1890 paper introduced a more general notion of integral – called now after him – in order to unify these two cases: He investigated the development of

$$\int \frac{d\sigma(x)}{z-x}$$

into a continued fraction, where now $\sigma(x)$ is some increasing function.

In our days, we find it natural to write a general inner product of polynomials defined on the real line in the form

$$(p,q) = \int p(x)\, q(x)\, d\sigma(x),$$

where σ is a Borel-measure. Moreover, the orthogonality usually is taken as the starting point of the theory. Nevertheless, the general case is not very well investigated. Most publications are dealing either with the discrete case or with absolutely continuous measures (often with stringent conditions about the density). This is not sufficient for applications in modern physics. The physicists Italo Guarneri and Giorgio Mantica write:

"Until recently, the traditional classification of energy spectra in pure point and absolutely continuous has proven quite sufficient to describe quantum evolution. The third element exhausting this classification, singular continuous energy spectra, has been usually considered by physicists as a mathematical curiosity and overlooked. However, this exotic kind of spectra is now attracting a growing attention: In fact, it is generic (in a mathematical sense) for one-body Schrödinger operators ... and appears fairly frequently in theoretical studies on quasiperiodic systems ..., on incommensurate structures ..., and on the electron dynamics of crystals in magnetic fields..." (see the introduction of their article [GM94]).

Actually, the singular case is difficult even for curious mathematicians. The problem of the computation of orthogonal polynomials is hard to formulate exactly: Absolutely continuous measures are given by their weight-function, discrete (or atomic) measures are given by the nodes and weights — but how a singular measure is given? However, this question can be answered in many cases: Singular measures occuring in applications have some invariance properties, which are often called *self-similarity*. This leads to recurrence relations for their moments $\mu_k = \int x^k \, d\sigma(x)$. As the above discussion shows (or the orthogonalization process of the canonical basis), the orthogonal polynomials are determined by the moments. Unfortunately, this is only theoretically useful — the condition of the underlying map grows exponentially. In the absolutely continuous case, the difficulty often can be circumvented, starting from *modified moments*, i.e. integrals of another polynomial base. It was not known, wether this is possible in the singular case. Gautschi's well-known analysis seems not to be applicable here, because we don't know enough about the distribution of zeros of those orthogonal polynomials.

We will regard the coefficients of the three-term recurrence relation as an encoding of the orthogonal polynomials, since essentially all quantities relevant for applications can be calculated from it. This will be illustrated by some (known) background information about orthogonal polynomials in the introductory Chapter 0. Let's denote the sequence of coefficients symbolically by ρ (in reality we will deal with finite sections, this will be made more precise in Chapter 3). For the sequence of modified moments we will write m. Another important object will be the Gaussian quadrature (symbolically denoted by γ) - the discretization of maximum degree of exactness (see Chapter 0). It is an important application of orthogonal polynomials and, moreover, theoretically interesting, as we will see. The relations between these three objects are shown by the following diagram:

$$\begin{tikzcd}
& \gamma \arrow[dr, "H"] & \\
m \arrow[ur, "G"] \arrow[rr, "K"] & & \rho
\end{tikzcd}$$

What we said above will determine the organization of the present paper. After we summarized some well-known (and some well-forgotten) facts from the theory, motivating the choice of ρ as *the* encoding of orthogonal polynomials, we prove some (mostly new) technical results used later (Chapter 1). In Chapter 2, we investigate asymptotic properties of ρ (convergence of means of the recurrence coefficients). Here, we make only rather weak assumptions about the measure, which are fulfilled in the self-similar singular case. Then, we close some gaps in the theory of the maps K, G and H: In Chapter 3, we propose a *direct* method to study the numerical condition of modified moment methods, i.e. of K (the well-known analysis of Walter Gautschi is based on the investigation of G, instead). In Chapter 4, we try an algorithmic implementation of G - this means the direct computation of the Gaussian quadrature formula from modified moments. The resulting iterative method, however, turns out to be slower than the methods traditionally used (essentially implementing $H^{-1} \circ K$). Additionally, we are able to calculate explicitely some polynomials occuring in Gautschi's theory of the condition of G (for Chebyshev measures). In Chapter 5, we study the condition of the map H, computing the recurrence coefficients from quadrature data. We provide some evidence for instability of Stieljes-type implementations of H in the presence of isolated mass points of the measure and propose a new algorithm (based on formulas of Paul Nevai for the addition of mass points). The method has similar stability properties as the so-called Lanczos algorithm, but it turns out to be slower. Finally, we show how to compute recurrence coefficients from self-similarity of the measure. Since the description of self-similarity is a (though rather formal) generalization of a quadrature formula, this will be an analogue of the map H. We mention here that our method (though developed independently) is very similar to the method of Mantica, which was published in 1996 (for details, see Chapter 6).

We illustrate the theorems and algorithms in this report by several examples. Though these examples have no immediate practical application, they show all typical difficulties which may occur in practice (at least in the case of compact support of the measure of orthogonality). The examples will be referred to by the number of the subsection where they are defined first.

Contents

Chapter 0

Introduction

This introductory chapter is devoted to some known basic facts about orthogonal polynomials. Another purpose of introduction is to fix notations (which differ widely in the mathematical literature). More detailed information about elementary properties of orthogonal polynomials can be found in the literature, see the references in the last section.

0.1 Orthogonal polynomials and the recurrence relation

Let σ be a given positive measure on the real line $\mathbf{R}$ with infinite support $S = S(\sigma)$ and finite moments of all orders. The termin *support* means the set of all points of growth, i.e. $S(\sigma) = \{x : \forall\delta > 0 \sigma([x-\delta, x+\delta]) > 0\}$. As we can see easily, this is always a closed set, sometimes it is called the *spectrum* of σ. Throughout this paper, we will assume that $S(\sigma)$ is compact. Some of our results do not depend on this assumption, this will be indicated where necessary.

Under the above assumptions, there exists a unique family of monic polynomials π_j with

$$\int \pi_l(x)\pi_j(x)\,d\sigma(x) = 0 \quad \text{for } l < j \text{ and } \quad \int \pi_j^2(x)\,d\sigma(x) = \gamma_j^{-2} > 0 \tag{1}$$

for $j = 0, 1, \ldots$ (γ_j is the leading coefficient of the *orthonormal* polynomial of degree j). They satisfy a three-term recurrence relation

$$\pi_{j+1}(x) = (x - \alpha_j)\pi_j(x) - \beta_j\pi_{j-1}(x) \quad \text{for} \quad j = 0, 1, \ldots, \tag{2}$$

if we set $\pi_{-1}(x) \equiv 0$.

We have to make a remark here: The simple structure of our recurrence relation is due to the special form of the inner product defining orthogonality. The inner product of two polynomials p and q is given by

$$(p, q) = \int p(x)q(x)\,d\sigma(x), \tag{3}$$

and this implies that the operator Ψ of multiplication by the independent variable (i.e. $(\Psi p)(x) = xp(x)$) is symmetric. Without this property, the recurrence relation is much more complicated (compare the short outlook to Sobolev-type orthogonal polynomials in Section 3.8).

Sometimes we are interested in *orthonormal* polynomials: There exist unique orthonormal polynomials

$$\hat{\pi}_n(x) = \gamma_n x^n + \dots \quad (\gamma_n = \gamma_n(\sigma) > 0),$$

with orthogonality relation

$$\int \hat{\pi}_m(x)\hat{\pi}_n(x)\, d\sigma(x) = \delta_{mn} \quad \text{for all} \quad m, n = 0, 1, \dots . \tag{4}$$

They are uniquely determined if we assume the highest coefficient γ_j in $\hat{\pi}_j$ to be positive. Obviously we have

$$\hat{\pi}_j = \gamma_j \pi_j$$

and from (4)

$$\gamma_j^{-2} = \int \pi_j^2(x)\, d\sigma(x).$$

This gives (see Eq. (9) below)

$$\gamma_j^2 = \frac{\gamma_{j-1}^2}{\beta_j} \tag{5}$$

for $j \geq 1$ (note that Nevai [Nev79] always writes γ_{j-1}/γ_j instead of $\beta_j^{1/2}$). The recurrence relation for these polynomials is

$$\beta_{j+1}^{1/2}\hat{\pi}_{j+1}(x) = (x - \alpha_j)\hat{\pi}_j(x) - \beta_j^{1/2}\hat{\pi}_{j-1}(x) \tag{6}$$

for $j \geq 0$.

In the relation (2) the coefficient β_0 would be arbitrary, but it is often assumed

$$\beta_0 = \sigma(S) = \sigma(\mathbf{R}),$$

since this convention unifies some formulas. One example is the normalization constant, it simply writes

$$\frac{1}{\gamma_n} = (\beta_0 \dots \beta_n)^{\frac{1}{2}} \tag{7}$$

(for these and other simple facts about general orthogonal polynomials we refer to [Chi78, Chapter I]). However, we have to admit that the convention could be criticized for mathematical reasons: The scaling properties (under a linear transformation of the support) of $\sigma(S)$ and of β_j (for $j \geq 1$) are different. Moreover, other formulas would rather suggest $\beta_0 = 0$ as a natural choice.

Following Gautschi [Gau82], we consider the problem of generating orthogonal polynomials as equivalent to the problem of calculating the coefficients α_j and β_j in their basic recurrence relation (2), and the facts cited below certainly will motivate this. In principle, simple explicit formulas for the coefficients α_j and β_j exist (due to Darboux and Stieltjes), see e.g. [Gau82][Eq. (2.2)]: We have

$$\alpha_j = \frac{\int x\pi_j^2(x)\,d\sigma(x)}{\int \pi_j^2(x)\,d\sigma(x)} \quad \text{for} \quad j = 0, 1, \ldots, \tag{8}$$

and

$$\beta_j = \frac{\int \pi_j^2(x)\,d\sigma(x)}{\int \pi_{j-1}^2(x)\,d\sigma(x)} \quad \text{for} \quad j = 1, 2, \ldots. \tag{9}$$

For later use, we formulate the following (essentially equivalent) auxiliary result.

Lemma 0.1 *The integral $\int x\pi_j(x)\pi_l(x)\,d\sigma(x)$ can be evaluated as*

$$\int x\pi_j(x)\pi_l(x)\,d\sigma(x) = \begin{cases} 0 & \textit{for} \quad l < j-1 \\ \gamma_j^{-2} & \textit{for} \quad l = j-1 \\ \alpha_j\gamma_j^{-2} & \textit{for} \quad l = j \end{cases}. \tag{10}$$

Proof: In the first case $l < j-1$, the polynomial $x\pi_l(x)$ has degree $< j$ and thus is orthogonal to $\pi_j(x)$. If $l = j-1$, then we have from Equation (2)

$$x\pi_{j-1}(x) = \pi_j(x) + \alpha_{j-1}\pi_{j-1}(x) + \beta_{j-1}\pi_{j-2}(x),$$

the second and third terms on the right-hand side of this equation being orthogonal to $\pi_j(x)$, and consequently

$$\int x\pi_j(x)\pi_{j-1}(x)\,d\sigma(x) = \int \pi_j^2(x)\,d\sigma(x) = \gamma_j^{-2}.$$

The last case $l = j$ follows analogously from the equation

$$x\pi_j(x) = \pi_{j+1}(x) + \alpha_j\pi_j(x) + \beta_j\pi_{j-1}(x).$$

Here the first and third terms of the right-hand side are orthogonal to $\pi_j(x)$, and we obtain

$$\int x\pi_j^2(x)\,d\sigma(x) = \int \alpha_j\pi_j^2(x)\,d\sigma(x) = \alpha_j\gamma_j^{-2}.\Box$$

We only remark that actually Equation (8) is just another notation of the last case in (10). Equation (9) could be derived simply by multiplying the recurrence relation (2) by $\pi_{j-1}(x)$ and integrating.

In the cases of discrete or self-similar measures (see below), these equations indeed can be used to determine recurrence coefficients.

It's rather difficult to derive properties of the sequence of recurrence coefficients from properties of the measure of orthogonality σ. Some new results will be given in Chapter 2. However, it's elementary to prove the boundedness of the coefficients under our assumption of compactness of the support $S(\sigma)$: From $S(\sigma) \subset [a,b]$ there follows immediately $\alpha_j \in [a,b]$ (this is Corollary 2 in Chapter IV of Chihara's book [Chi78], or a trivial consequence of Equation (8)). Under the same assumption we can easily prove the well-known fact that

$$\beta_j < \frac{(b-a)^2}{4} \tag{11}$$

(see [Chi78, Chapter IV-Exercise 2.12]).

The recurrence relation (2) is of second order, thus there must be a second independent solution, which we will need while analyzing the numerical properties of our recursion.

Definition 0.2 *Let π_j be the monic orthogonal polynomials with respect to some measure σ, satisfying the recurrence relation (2). The (monic) numerator polynomials κ_j are defined by*

$$\kappa_j(x) = \gamma_0^2 \int \frac{\pi_j(x) - \pi_j(t)}{x - t}\, d\sigma(t). \tag{12}$$

These polynomials (which are called sometimes *associated polynomials* or *polynomials of second kind*) satisfy the same recurrence relation (2), except for $j = 0$ (if we don't make special assumptions about β_0). Note, however, that κ_j is a monic polynomial of degree $j-1$ (compare the different notation in [Chi78]). The origin of the term *numerator polynomials* will be explained in Section 0.3. We will use the following simple property of numerator polynomials (this is just [Chi78, III-(4.4)] in our notations).

Lemma 0.3 *The identity*

$$\pi_n(x)\kappa_{n+1}(x) - \pi_{n+1}(x)\kappa_n(x) = \frac{\gamma_0^2}{\gamma_n^2} \tag{13}$$

holds for $n \geq 0$.

We will prove further properties of numerator polynomials when needed (compare the identity (4.11) in Chapter 4 or the estimates in Section 5.3).

0.2 Quadratures and discretizations

The analytical computation of integrals with respect to the measure σ is possible only in special cases. In practice, we will have to approximate the integrals by finite sums. If this approximation is exact for functions of a certain class, we call this a *quadrature formula* or *quadrature rule*. We will consider here only quadrature formulas of a special kind: The integral is approximated by a linear combination of functional values of the integrand, and the approximation is exact for polynomials up to a certain degree. This gives raise to the following definitions.

Definition 0.4 *The set of all polynomials of degree $\le d$ will be denoted by $\mathcal{P}_d$. We write* $\deg(q)$ *for the degree of a polynomial q.*

Definition 0.5 *We say that the vector*

$$\gamma = [\tau_1, \dots, \tau_n, \sigma_1, \dots, \sigma_n]^T$$

represents a quadrature rule of degree of exactness d for the measure σ if

$$\int q(x)\, d\sigma(x) = \sum_{i=1}^{n} \sigma_i\, q(\tau_i), \tag{14}$$

for any $q \in \mathcal{P}_d$.

Since a quadrature rule with n nodes and weights has $2n$ degrees of freedom, we can expect a maximal degree of exactness $2n-1$ (because $\mathcal{P}_{2n-1}$ has dimension $2n$). Indeed, this maximal degree of exactness is possible. The quadrature rule of degree of exactness $2n-1$ is called *Gaussian quadrature rule* or *Gauss-Christoffel formula.* The weights are sometimes called *Christoffel numbers.* It turns out that we have to choose the nodes τ_i as the n zeros of the orthogonal polynomial π_n. The weights σ_i can be found, e.g., from the formula

$$\sigma_i = \frac{1}{\sum_{j=0}^{n-1} \hat{\pi}_j^2(\tau_i)}$$

or from

$$\sigma_i = -\frac{\gamma_n^{-2}}{\pi_{n+1}(\tau_i)\pi_n'(\tau_i)}$$

(see Exercise 6.2 in Chapter I of Chihara's book [Chi78].

One can show that under our assumptions, all n zeros of π_n are real (see Section 5 of Chapter I in Chihara's book [Chi78]). We can see from the first formula for σ_i that all weights are positive (this is important in numerical applications).

In practice, it is not necessary to use polynomial rootfinding to compute the nodes of the Gauss formula. In fact, we can compute the Gauss rule from the recurrence coefficients! As Gautschi mentioned already in his papers [Gau90] or [Gau96], the nodes are simply the eigenvalues of a symmetric tridiagonal matrix: We define the *Jacobi matrix* by

$$J_n = \begin{bmatrix} \alpha_0 & \sqrt{\beta_1} & & & & 0 \\ \sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \sqrt{\beta_{n-1}} & \alpha_{n-1} & \sqrt{\beta_n} \\ 0 & & & & \sqrt{\beta_n} & \alpha_n \end{bmatrix}.$$

The following theorem is essentially Gautschi's Theorem 4 from [Gau96] (we adapted the notation slightly).

Theorem 0.6 *Let x_i be the eigenvalues of J_n, and u_i the corresponding normalized eigenvectors, so that*

$$J_n u_i = x_i u_i, \quad u_i^T u_i = 1, \quad i = 1, 2, \ldots, n. \tag{15}$$

Then the Gaussian nodes τ_i and weights σ_i are given by

$$\tau_i = x_i, \quad \sigma_i = \beta_0 u_{i,1}^2, \quad i = 1, 2, \ldots, n, \tag{16}$$

where $u_{i,1}$ is the first component of u_i and $\beta_0 = \sigma(\mathbf{R})$.

We only mention here that other quadrature rules known from numerical analysis, e.g. Gauss-Radau or Gauss-Lobatto quadrature rules, can be generated using similar techniques (see Sections 4.2 and 4.3 of [Gau96]).

It's useful to develop this idea a bit further. If $U = [u_1, u_2, \ldots, u_n]$ is the (orthogonal) matrix of the normalized eigenvectors of $J = J_n$, and if we set $Q = U^T$ and $D_\tau = \mathrm{diag}(\tau_1, \tau_2, \ldots, \tau_n)$, then we can write (15) and (16) in the following form:

$$\begin{bmatrix} 1 & 0^T \\ 0 & Q^T \end{bmatrix} \begin{bmatrix} 1 & \sqrt{\sigma}^T \\ \sqrt{\sigma} & D_\tau \end{bmatrix} \begin{bmatrix} 1 & 0^T \\ 0 & Q \end{bmatrix} = \begin{bmatrix} 1 & \sqrt{\beta_0}\, e_1^T \\ \sqrt{\beta_0}\, e_1 & J \end{bmatrix}, \tag{17}$$

where $e_1 = [1, 0, \ldots, 0]^T$ is the first coordinate vector and we used the symbolic notation

$$\sqrt{\sigma} = [\sqrt{\sigma_1}, \sqrt{\sigma_2}, \ldots, \sqrt{\sigma_n}]^T$$

(this is Equation (4.15) in [Gau96]). Thus, the matrix in the middle on the left (Gautschi called it 'Gauss matrix', because it represents the Gauss rule) is connected with the extended Jacobi matrix on the right by an orthogonal similarity transformation. This seems to indicate that the map from the Gauss rule to the recurrence coefficients (which we denoted by H in the Preface) is well-conditioned. However, the matrix Q depends on the initial data, as well. We will return to the numerical condition of this map in Chapter 5.

The last paragraph showed that we can look at a quadrature rule from another point of view: It can be the initial data defining our measure. In practice, of course, the Gauss-Christoffel rule for some measure σ usually will be unknown, but quadrature rules of lower (but sufficiently high) degree of exactness are often available (examples will be given in Chapter 5).

Suppose we are given *any* quadrature formula for the measure σ represented by the vector $\gamma = [\tau_1, \ldots, \tau_n, \sigma_1, \ldots, \sigma_n]^T$ having degree of exactness $2n - 1$. This means

$$\int q(x)\, d\sigma(x) = \sum_{i=1}^{N} \sigma_i\, q(\tau_i),$$

for any $q \in \mathcal{P}_{2n-1}$. The sum on the right is an integral with respect to a measure $\sigma^{(N)}$ with finite support $\{\tau_1, ..., \tau_N\}$. To make this more precise, we introduce some notations.

Definition 0.7 *For some real t let δ_t be the discrete measure defined by*

$$\int f(x)\, d\delta_t(x) = f(t)$$

for any function f (this is just a point mass of weight 1 at t).

Definition 0.8 *We say that the measure*

$$\sigma^{(N)} = \sum_{i=1}^{N} \sigma_i\, \delta_{\tau_i}$$

is a discretization of the measure σ of degree of exactness d if

$$\int q(x)\, d\sigma(x) = \int q(x)\, d\sigma^{(N)}(x) = \sum_{i=1}^{N} \sigma_i\, q(\tau_i),$$

for any $q \in \mathcal{P}_d$.

If the support of a measure is a finite set, the family of monic polynomials satisfying the orthogonality conditions (1) is also finite (compare the discussion in [Gau82][Introduction] – Gautschi calls this a *discrete measure*).

Thus, we assumed that we have a discretization of degree $2n-1$ of our measure. We will see that the replacement of σ by $\sigma^{(N)}$ will not change the first few recurrence coefficients:

Definition 0.9 *The vector of the first recursion coefficients will be denoted by*

$$\rho_n = \rho_n(\sigma) = [\alpha_0, \ldots, \alpha_{n-1}, \beta_1, \ldots, \beta_{n-1}]^T.$$

Now it is known that the vector of coefficients $\rho_n(\sigma)$ depends only on the moments of order up to $2n-1$ (see Equations (8) and (9) above) and thus the vectors $\rho_n(\sigma)$ and $\rho_n(\sigma^{(N)})$ for the measure σ and the discretization $\sigma^{(N)}$ of degree of exactness $\geq 2n-1$ coincide.

The set of monic orthogonal polynomials w.r.t. a discrete measure with support $\{\tau_1, ..., \tau_N\}$ has some pecularities (that's why we made the assumption of an infinite suppoert in Section 0.1). It consists of N polynomials $\pi_0, \ldots, \pi_{N-1}$. Formally we can calculate from the recurrence relation (2) one more polynomial π_N, but we will have

$$\int \pi_N^2(x)\, d\sigma(x) = 0$$

(see Section 5.2 below). Of course, this is equivalent to $\pi_N(\tau_i) = 0$ for $i = 0, 1, \ldots, N$. As an examination of the proof in [Gau96] shows, Theorem 0.6 and Equation (17) (with N instead of n) remain valid! Equation (17) is the starting point of an algorithm for the computation of the recurrence coefficients (or equivalently of the Jacobi matrix) for a discrete measure. It is usually named after Lanczos, though this is somewhat misleading. For a detailed description we refer to the article of Gragg and Harrod [GH84].

0.3 Continued fractions

As we mentioned in the Preface, the orthogonal polynomials appeared in the investigation of continued fractions, first.

The starting point is a formal power series. If we integrate the series

$$\frac{1}{x-z} = \sum_{k=0}^{\infty} \frac{x^k}{z^{k+1}}$$

(converging only for $|z| > |x|$) formally with respect to the measure σ, we obtain the representation

$$\int \frac{d\sigma(x)}{z-x} \sim \sum_{k=0}^{\infty} \frac{\mu_k}{z^{k+1}}$$

with the moments $\mu_k = \int x^k \, d\sigma(x)$. This is not an equality, because the series on the right may not converge for any finite value of z (in fact, this will happen if and only if the support of the measure is unbounded). However, we can formally develop the series into a continued fraction

$$\cfrac{\beta_0}{z-\alpha_0-\cfrac{\beta_1}{z-\alpha_1-\cfrac{\beta_2}{z-\alpha_2-\ddots}}} \tag{18}$$

by calculation in the field of formal power series (compare Section 4 of Chapter 2 in the book [NS88]): We begin with the formal power series

$$\phi_0(z) = \sum_{k=0}^{\infty} \frac{\mu_k}{z^{k+1}}.$$

If we choose $\beta_0 = \mu_0$, then we easily can see that

$$\beta_0/\phi_0(z) = x - \alpha_0 - \phi_1(z),$$

where α_0 is some constant and the power series $\phi_1(z)$ contains only negative powers of z. This process can be continued, giving two sequences (β_j) and (α_j).

Of course, this needs a proof, since it is not trivial that the coefficient β_j of z^{-1} in any $\phi_j(z)$ is different from zero. In fact, if μ_k are the moments of a measure with infinite support, thes e coefficients can be shown to be positive. We will not go into details here, since this approach is only of historical interest.

The infinite continued fraction (18) is a mere symbol defined by the sequences (β_j) and (α_j). If, however, we truncate it after $z - \alpha_{n-1}$, we obtain a finite continued fraction, defining a rational function $F_n(z) = A_n(z)/B_n(z)$ (we assume $F_0(z) \equiv 0$). The polynomials $A_n(z)$ and $B_n(z)$ have degrees $n-1$ and n respectively, and satisfy the general recurrence formulas for continued fractions, established by Wallis in 1655 (see Section 2 of Chapter III in [Chi78], we assume $A_0(z) \equiv 0$ and $B_0(z) \equiv 1$). For our continued fraction, the fundamental recurrence formula coincides with Equation (2)! This immediately implies $B_n(z) = \pi_n(z)$ and $A_n(z) = \beta_0\kappa_n(z)$ and justifies the termin *monic numerator polynomials* for κ_n.

The finite continued fractions obtained by truncation are called *convergents* of an infinite continued fraction (though they don't converge in general). They have a numerical value, which is $F_n(z)$ in our case. If the sequence $(F_n(z))$ converges as $n \to \infty$, we consider the limit as the numerical value of the continued fraction.

One of the first convergence results was the following:

Theorem 0.10 *If σ is a measure with compact support $S \subset [a, b]$, then the Jacobi fraction converges (uniformly on any compact $\subset \mathbf{C}$ not intersecting $[a, b]$) to the function*

$$F(z) = \int \frac{d\sigma(x)}{z - x}.$$

This is Markov's theorem, see Section 4 of Chapter III in [Chi78].

We only mention here that the convergent $F_n(z)$ (i.e. the rational function $A_n(z)/B_n(z)$) coincides with the Padé approximation $[n-1, n]$ to the power series $\phi_0(z)$ around $z = \infty$. For a thorough investigation of Padé approximations and their connection to continued fractions we refer to Chapter 2 of [NS88].

0.4 Bibliographical references

In this introductory chapter, we mainly fix notations and cite a few crucial results from literature to which we will refer most often. The standard reference book for all classical results on orthogonal polynomials is the book of Szegö [Sze75]. Another valuable source of information is the book of Freud [Fre71]. The more formal algebraic aspects of the theory as well as properties of many special systems of orthogonal polynomials can be found in Chihara's book [Chi78]. A very good introduction to various aspects of orthogonal polynomials (including Padé approximation) is the book of Nikishin and Sorokin [NS88] (in Russian).

In the present paper, we concentrate on orthogonal polynomials on compact subsets of the real line. Some properties of orthogonal polynomials extend to the case of measures given in the complex plane. The article [Saf90] is a good introduction into this area. However, many results rely on a definite inner product *and* a simple recurrence formula, this is possible only for a real support. Results for measures with noncompact support were obtained only recently using weighted potential theory in the complex plane. This theory together with many applications in approximation theory (especially orthogonal polynomials) can be found in the recent book of Saff and Totik [ST97].

Chapter 1

Technical results

1.1 Shifts and symmetrization

Here we summarize how orthogonal polynomials and their recurrence coefficients are transformed after two simple modifications of the measure of orthogonality.

Let σ be any measure with $S(\sigma) \subset [a, b]$ (i.e. we assume compactness of $S(\sigma)$). Then we can define a shifted measure ν with $S(\nu) \subset [a + \lambda, b + \lambda]$ via the equation

$$\int h(x)\, d\nu(x) = \int h(x + \lambda)\, d\sigma(x)$$

for any polynomial h. Thus, if the monic polynomials $p_j(x)$ are orthogonal with respect to ν, then the polynomials $p_j(x + \lambda)$ are monic and orthogonal with respect to σ. Consequently, the latter polynomials are identical to the monic polynomials $\pi_j(x)$ with respect to σ. This means

$$p_j(x) = \pi_j(x - \lambda),$$

and if we substitute $x - \lambda$ instead of x in the recurrence (2) we obtain immediately

$$\alpha_j(\nu) = \alpha_j(\sigma) + \lambda \quad \text{and} \quad \beta_j(\nu) = \beta_j(\sigma).$$

Let σ be any measure with $S(\sigma) \subset [0, \infty)$ (we always can achieve this by an appropriate shift). We define a measure ν by

$$\int x^{2j}\, d\nu(x) = \int x^j\, d\sigma(x) \quad \text{and} \quad \int x^{2j+1}\, d\nu(x) = 0$$

for $j \geq 0$ (we assumed $S(\sigma)$ to be a compact set, so ν is uniquely determined). Then the monic orthogonal polynomials $p_j(x)$ with respect to ν are

$$p_{2j}(x) = \pi_j(x^2) \quad \text{and} \quad p_{2j+1}(x) = x\pi_j^*(0; x^2) \quad \text{for} \quad j \geq 0,$$

where

$$\pi_j^*(0;x) = x^{-1}\left[\pi_{j+1}(x) - \frac{\pi_{j+1}(0)}{\pi_j(0)}\pi_j(x)\right] \tag{1.1}$$

(this is [Chi78, Chapter I, Theorem 8.1]). There is a simple relation between the recurrence coefficients, too: The coefficients $\alpha_j(\nu)$ are identically zero by symmetry, and the $\beta_j(\nu)$ satisfy

$$\begin{aligned} \alpha_0(\sigma) &= \beta_1(\nu), && (1.2)\\ \alpha_j(\sigma) &= \beta_{2j}(\nu) + \beta_{2j+1}(\nu), \quad j \geq 1, && (1.3)\\ \beta_j(\sigma) &= \beta_{2j-1}(\nu)\beta_{2j}(\nu), \quad j \geq 1 && (1.4) \end{aligned}$$

([Chi78, Chapter I, Theorem 9.1]).

1.2 A general result concerning dependence on a parameter

The main result of this section is a lemma describing the dependence of recurrence coefficients of orthogonal polynomials on a parameter of the measure of orthogonality. For the proof, we need a known auxiliary result:

Lemma 1.1 *The coefficient of x^{n-1} in $\pi_n(x)$ is $-\sum_{j=0}^{n-1}\alpha_j$.*

This is just [Chi78][Chapter I, Theorem 4.2 (d)].

Now we can formulate our main lemma.

Lemma 1.2 *Let l be some positive linear functional on $\mathcal{P}_{2n-1}$, differentiably depending on some parameter t, and let the functional c be defined by*

$$\frac{\partial l}{\partial t} = c.$$

Then there exist unique monic polynomials $\pi_0, \ldots, \pi_{n-1}$ with

$$l(\pi_j\pi_i) = 0 \quad \text{for} \quad 0 \leq i < j \leq n-1$$

and

$$l(\pi_j^2) = \gamma_j^{-2} > 0 \quad \text{for} \quad 0 \leq j \leq n-1,$$

fulfilling the recurrence relation (2). The partial derivatives of the recursion coefficients with respect to t are

$$\frac{\partial \alpha_j}{\partial t} = \gamma_j^2 c(\pi_j\pi_{j+1}) - \gamma_{j-1}^2 c(\pi_{j-1}\pi_j) \tag{1.5}$$

and

$$\frac{\partial \beta_j}{\partial t} = \beta_j[\gamma_j^2 c(\pi_j^2) - \gamma_{j-1}^2 c(\pi_{j-1}^2)]. \tag{1.6}$$

Proof: The assertion concerning the existence of the polynomials and the recurrence relation is clear, see [Chi78, Chapter I,§§3 and 4]. The coefficients α_j and β_j can be calculated from

$$\alpha_j = \frac{l(x\pi_j^2(x))}{l(\pi_j^2)}$$

and

$$\beta_j = \frac{l(\pi_j^2)}{l(\pi_{j-1}^2)} = \frac{\gamma_{j-1}^2}{\gamma_j^2} \tag{1.7}$$

(these equations are just (8) and (9) in another notation). Since our polynomials are monic, i.e. their highest coefficient 1 is independent of t, we have $\dfrac{\partial \pi_j}{\partial t} \in \mathcal{P}_{j-1}$. In order to calculate this polynomial, we expand it into a Fourier series:

$$\frac{\partial \pi_j}{\partial t} = \sum_{i=0}^{j-1} c_{ji}\pi_i$$

with

$$c_{ji} = \frac{l\left(\dfrac{\partial \pi_j}{\partial t}\pi_i\right)}{l(\pi_i^2)} = \gamma_i^2 l\left(\frac{\partial \pi_j}{\partial t}\pi_i\right).$$

These coefficients can be obtained by differentiating the orthogonality relations with respect to t. From the elementary identity

$$\frac{\partial}{\partial t} l(q) = l\left(\frac{\partial q}{\partial t}\right) + \frac{\partial l}{\partial t}(q) = l\left(\frac{\partial q}{\partial t}\right) + c(q) \tag{1.8}$$

we have

$$0 = \frac{\partial}{\partial t} l(\pi_j\pi_i) = l\left(\frac{\partial \pi_j}{\partial t}\pi_i\right) + l\left(\pi_j\frac{\partial \pi_i}{\partial t}\right) + c(\pi_j\pi_i) \quad \text{for} \quad i < j,$$

but

$$l\left(\pi_j\frac{\partial \pi_i}{\partial t}\right) = 0$$

due to orthogonality. Thus we have

$$c_{ji} = -\gamma_i^2 c(\pi_j\pi_i)$$

and consequently

$$\frac{\partial \pi_j}{\partial t} = -\sum_{i=0}^{j-1} \gamma_i^2 c(\pi_j\pi_i)\pi_i. \tag{1.9}$$

From

$$l(\pi_j^2) = \gamma_j^{-2}$$

and (1.8) we obtain

$$\frac{\partial \gamma_j^{-2}}{\partial t} = l\left(2\pi_j \frac{\partial \pi_j}{\partial t}\right) + c(\pi_j^2) = c(\pi_j^2)$$

due to orthogonality again (recall $\frac{\partial \pi_j}{\partial t} \in \mathcal{P}_{j-1}$). Now Equation (1.6) is an immediate consequence of (1.7) and the last equation.

By virtue of Lemma 1.1, the coefficient of x^{j-1} in π_j is $-\sum_{i=0}^{j-1} \alpha_i$. Consequently, the coefficient of x^{j-1} in $\frac{\partial \pi_j}{\partial t}$ is

$$-\frac{\partial}{\partial t}\sum_{i=0}^{j-1} \alpha_i = -\gamma_{j-1}^2 c(\pi_j \pi_{j-1})$$

due to Equation (1.9). Taking differences we arrive at Equation (1.5).□

1.3 Inequalities I

In sufficiently regular cases, the normalized polynomials $\hat{\pi}_n$ grow only polynomially in n on $S(\sigma)$. This depends on the measure and its support. The following definition is taken from [Jon94].

Definition 1.3 *Let $1 \leq p \leq \infty$, let $S \subset \mathbf{R}$ be a compact set and let σ be a measure with support S satisfying, for some constants $c_0 > 0$ and $s > 0$,*

$$\sigma(B(x,r)) \geq c_0 r^s, \qquad x \in S, \quad 0 < r \leq 1, \tag{1.10}$$

where $B(x,r)$ denotes the closed ball with center x and radius r. The set S preserves Markov's inequality in $L^p(\sigma)$ if there exist constants $M > 0$ and $\alpha > 0$ such that

$$\|q'\|_p \leq M n^\alpha \|q\|_p \quad \text{for} \quad q \in \mathcal{P}_n$$

(and for this we write $S \in \mathcal{M}^p(M,\alpha)$). The norm $\|\cdot\|_p$ is defined as usually:

$$\|q\|_p = \left(\int |q(x)|^p \, d\sigma(x)\right)^{1/p}$$

for $p < \infty$ and

$$\|q\|_\infty = \sup_{x \in S} |q(x)|.$$

For our purposes, of course, only the cases $p = 2$ and $p = \infty$ are of interest. The following result is implicit in the paper of Jonsson [Jon94], though surprisingly it is not stated as a corollary:

Proposition 1.4 *If the measure σ and its support satisfy the conditions of Definition 1.3, then*

$$\|\hat{\pi}_n\|_\infty \leq A_1 n^{\alpha s/2}$$

with some constant $A_1 = A_1(M, p, c_0, s)$.

Proof: This is an immediate consequence of Lemma 1(a) in [Jon94], claiming that under our assumptions we have the inequality

$$\|P\|_\infty \leq A_1 n^{\alpha s/p} \|P\|_p$$

for $p \geq 1$, $P \in \mathcal{P}_n$. Now for $p = 2$ and $P = \hat{\pi}_n$ we trivially have $\|P\|_p = 1$.□

In fact, the proof of Jonsson's Lemma 1 shows even more: We have $|\hat{\pi}(x)| \leq B_1 n^{\alpha s/2}$ in any $O(n^{-\alpha})$-neighborhood of S. However, usually we will not have a sharp value of α in Markov's inequality, so this generalization is of little use.

1.4 A simple error bound for special Padé approximations

The following error bound will be used only in an example in Chapter 6. The result, however, could be useful in other situations, as well. For the formulation of the proof, we need a simple definition:

Definition 1.5 *The distance of a point $x \in \mathbf{R}$ from a set $S \subset \mathbf{R}$ will be denoted by dist(x, B), i.e.*

$$dist(x, B) = \inf_{y \in B} |x - y|,$$

the distance between two sets A and B is defined by

$$dist(A, B) = \inf_{x \in A, y \in B} |x - y|.$$

Strictly speaking, we would have to write $\mathrm{dist}(\{x\}, B)$ in the first case, but we prefer the simpler notation. Now we state our result.

Lemma 1.6 *Let σ be a measure with support $\subset \Delta = [a, b]$, let*

$$F(z) = \int \frac{d\sigma(x)}{z - x} \quad \textit{for} \quad z \in \mathbf{C} \setminus \Delta$$

and let $F_n(z)$ be the Padé approximation $[n-1,n]_F$ to $F(z)$ (around $z=\infty$, see Section 0.3). Then we have the bound

$$|F(z)-F_n(z)| \leq \frac{8e^2 n(n+1)}{b-a}\gamma_0^{-2}|\psi_\Delta(z)|^{-2n} \quad \textit{for} \quad z \textit{ outside } L_r,$$

where $r = 1+\frac{1}{n}$,

$$\psi_\Delta(z) = \psi\left(\frac{z-(a+b)/2}{(b-a)/2}\right),$$

$$\psi(z) = z + \sqrt{z^2-1}$$

(the branch of the root is chosen so that $\sqrt{z^2-1}>0$ for real $z>1$) and L_r is the ellipse defined by

$$L_r = \{z \in \mathbf{C} : |\psi_\Delta(z)| = r\}$$

(which is really well-defined for $r>1$).

Proof: We follow closely the proof of Theorem 6.2 of Chapter 2 in [NS88]. The function $[F(z)-F_n(z)]\psi_\Delta(z)^{2n}$ is analytic in the exterior of L_r (i.e. in the connected component of $(\mathbf{C}\cup\{\infty\})\setminus L_r$ containing ∞) and thus the maximum of the modulus is attained at the boundary L_r. As in [NS88], we have

$$|F(z)-F_n(z)| \leq \frac{2\gamma_0^{-2}}{c(r)} \quad \text{for} \quad z \in L_r,$$

where (correcting a typographical error in [NS88])

$$c(r) = \operatorname{dist}(L_r, \Delta) = \frac{(r-1)^2}{2r}\cdot\frac{b-a}{2}.$$

Observing $r = 1+\frac{1}{n}$ and

$$|\psi_\Delta(z)^{2n}| = r^{2n} = \left(1+\frac{1}{n}\right)^{2n} < e^2 \quad \text{for} \quad z \in L_r,$$

we obtain our proposition. □

1.5 Polynomial bases and polynomial interpolation

Under the assumption

$$\deg(p_j) = j \tag{1.11}$$

the set $\{p_j\}_{j=0,\dots,N}$ (where we will be especially interested in the cases $N = n-1$ and $N = 2n-1$) forms a base in the space $\mathcal{P}_N$ of all polynomials of degree $\le N$. Consequently, there is a unique representation

$$q(x) = \sum_{i=0}^{N} c_j p_j(x) \quad \text{for any} \quad q \in \mathcal{P}_N \tag{1.12}$$

with coordinates $c_j = c_j(q)$. The coordinate functionals c_j are linear and form a base in the dual space $\mathcal{P}_N^*$. Clearly, the equation $c_i(p_j) = \delta_{ij}$ is satisfied, and thus we will write sometimes p_j^* instead of c_j. All arithmetic operations with polynomials such as addition or multiplication can principially be reduced to more or less complex manipulations of their coordinate vectors.

Usually, it is assumed that the polynomials p_j satisfy a three-term recurrence relation of the form

$$xp_j(x) = \frac{K_j}{K_{j+1}} p_{j+1}(x) + \alpha_j p_i(x) + \frac{K_j}{K_{j-1}} \beta_j p_{j-1}(x) \quad \text{for} \quad j = 0, 1, \dots, \tag{1.13}$$

where $p_{-1} \equiv 0$ and $p_0 \equiv K_0$. This includes all sets of orthogonal polynomials (compare the recurrence relation (2)). It is, however, more general: The β_j are not necessarily positive, and the K_j (the leading coefficients of our polynomials) can be chosen arbitrarily. This is convenient, if one deals with classical orthogonal polynomials, having special standardizations for historical reasons (e.g. $P_n(1) = 1$ for Legendre polynomials P_n, see Table 22.2 in the book [AS84]).

If the polynomial base satisfies the recurrence relation (1.13), then it's easy to multiply an arbitrary polynomial represented in this base by the independent variable: The polynomial q can be identified with the coordinate vector $[p_0^*(q), \dots, p_N^*(q)]$, and the recurrence relation shows that the coordinate vector of $xq(x)$ can be obtained by multiplication by a triangular matrix. Obviously, this can be done in $O(N)$ arithmetic operations.

In the general case of a polynomial base, the following statement is obvious: If the polynomial q of degree $< N$ is represented by its coordinate vector $[p_0^*(q), \dots, p_N^*(q)]$, then the coordinate vector of $xq(x)$ can be obtained from

$$[p_0^*(xq(x)), \dots, p_N^*(xq(x))] = \Psi\left([p_0^*(q), \dots, p_N^*(q)]\right), \tag{1.14}$$

where Ψ is some linear map.

Definition 1.7 *If in a polynomial base the map Ψ defined by equation (1.14) can be implemented in at most $O(N)$ arithmetic operations, i.e. there is an algorithm for Ψ with this operation count, then we will say that the base has* **Property X**.

Of course, (1.11) is only a sufficient condition for (1.12) to hold. An important example for a base not satisfying (1.11) are the fundamental polynomials of Lagrange interpolation. Let $\{\tau_i\}_{i=1,\dots,n}$ be a set of distinct numbers, and define

$$l_i(x) = \frac{\omega_n(x)}{\omega_n'(\tau_i)(x - \tau_i)} \quad \text{for} \quad i = 1, \dots, n, \tag{1.15}$$

where $\omega_n(x) = (x - \tau_1)\dots(x - \tau_n)$. The well-known Lagrange interpolation formula

$$q(x) = \sum_{i=1}^{n} q(\tau_i)\, l_i(x) \quad \text{for} \quad q \in \mathcal{P}_{n-1} \tag{1.16}$$

just says that $\{l_i\}_{i=1,\dots,n}$ forms a base of $\mathcal{P}_N$ with $N = n - 1$ and that the dual base has the explicit form $l_i^*(q) = q(\tau_i)$.

The following result is well known, too.

Lemma 1.8 *Let*

$$h_i(x) = \left[1 - \frac{\omega_n''(\tau_i)}{\omega_n'(\tau_i)}(x - \tau_i)\right] l_i^2(x) \quad \textit{and} \quad k_i(x) = (x - \tau_i)\, l_i^2(x). \tag{1.17}$$

Then we have the equation

$$q(x) = \sum_{i=1}^{n} q(\tau_i)\, h_i(x) + \sum_{i=1}^{n} q'(\tau_i)\, k_i(x) \quad \textit{for} \quad q \in \mathcal{P}_{2n-1}. \tag{1.18}$$

1.6 The arithmetic of orthogonal polynomials

In many cases, we must have algorithms for arithmetical operations with polynomials. Of course, we assume that any polynomial is given as a linear combination of our basic polynomials. Thus the polynomials are encoded as vectors, and it is clear how to add or subtract or multiply by a scalar factor. Multiplication and division require some new ideas. We assume that our basic polynomials satisfy the recurrence relation (1.13), first. Moreover, in this section we assume that the polynomials are *monic* (i.e. $K_j \equiv 1$) — this assumption saves lots of multiplications. How the general case $K_j \not\equiv 1$ can be handled, is trivial. In which cases the weaker Property X of a base is sufficient, will be indicated after the derivation of our algorithms.

Before we explain how to multiply or divide by polynomials, we look at the problem of evaluation in some point t. The starting point is the recurrence relation (1.13). Now for the evaluation of the polynomial

$$q(x) = \sum_{j=0}^{n} c_j p_j(x)$$

at $x = t$ we have Clenshaw's algorithm (see Algorithm 1.1).

As we explained in Section 1.5, the recurrence (1.13) allows to multiply a polynomial by the independent variable x. Thus, we can use Clenshaw's algorithm not only with some numerical value t, but with the independent variable x, assuming the quantities u_j to be polynomials. This seems to be a rather strange idea, since we just reconstruct the polynomial $q(x)$. Consequently, this is an

```
u_n ← c_n,    u_{n+1} ← 0,
for j = n − 1 downto 0 do
  u_j ← (t − α_j) × u_{j+1} − β_{j+1} × u_{j+2} + c_j
end for
q(t) ← u_0
```

Algorithm 1.1: Calculate $q(t)$

implementation of the identity map in $O(n^2)$ arithmetical operations. However, this can be very practical!

First, we can perform the operations of the algorithm in *another* orthogonal base satisfying another recurrence of type (2). This gives a conversion algorithm from one orthogonal base to another (or even to any base having Property X) in $O(n^2)$ arithmetical operations.

Second, we can assume that instead of the numerical values c_j we have the polynomials $c_j r(x)$, where $r(x)$ is some given polynomial. Then the algorithm obviously calculates

$$\sum_{j=0}^{n} c_j r(x) p_j(x) = q(x) r(x).$$

Again, we formulate this as Algorithm 1.2 (where we simply write $c \times u$ for the multiplication of vector u by a number c). If the degree of $r(x)$ is m, then the

```
u_n ← c_n × r,    u_{n+1} ← 0, {u_j and r are polynomials}
for j = n − 1 downto 0 do
  u_j ← Ψ(u_{j+1}) − α_j × u_{j+1} − β_{j+1} × u_{j+2} + c_j × r {Ψ multiplies by x}
end for
s ← u_0 {the result is a polynomial}
```

Algorithm 1.2: Calculate $s(x) = q(x)r(x)$

operation count of our algorithm obviously is $O((m + n/2)n)$. Here the order of factors is crucial — we should guarantee $n \leq m$ for efficiency. However, we should have in mind that the polynomials u_j are only required to be represented in a base with Property X, the product will then be obtained in this base. Naturally, this should be the base in which the second factor r is given (because u_n is $c_n \times r$). Thus, the target base has to have only Property X, and this is the base in which the second factor is given — in such cases the order of factors can't be changed.

Now we introduce some notations.

Definition 1.9 *Let q and r be polynomials. There are unique polynomials q and t with $\deg(t) < \deg(r)$ and $q(x) = r(x)\,s(x) + t(x)$, and we will write $s = q \div r$ and $t = q \bmod r$.*

Interestingly, the computation of $q \div r$ and $q \bmod r$ can be performed along the same lines. In fact, any polynomial u_j in the polynomial version of Clenshaw's algorithm 1.1 can be witten uniquely as

$$u_j = v_j r + w_j \quad \text{with} \quad \deg(w_j) < \deg(r).$$

For $\deg(u_j) < \deg(r)$ we have $v_j \equiv 0$ and $w_j = u_j$, this is true for the initial values u_{n+1} and u_n. Obviously, addition, subtraction and multiplication by a constant of the representations (v_j, w_j) of u_j can be performed componentwise. The multiplication by the variable x is not difficult, if we first multiply the "lower" component w_j by x and then reduce the product mod r to obtain the "carry" for the higher part v_j. Since $\deg(\Psi(w_j)) = \deg(w_j) + 1 \leq \deg(r)$, this carry is always a number d_j. Now, a careful analysis of the operations performed with the polynomials v_j shows that this is exactly the symbolic version of Clenshaw's algorithm. Thus, it reconstructs the polynomial from its coefficients, and this means that the numbers d_j are the coefficients of $q \div r$! These observations give an elegant algorithm for polynomial division in the base $\{p_j\}$ (see Algorithm 1.3). Again, we can use the algorithm in a slightly more general situation: In fact, to

$w_n \leftarrow c_n, \quad w_{n+1} \leftarrow 0$, $\{w_j$ are polynomials$\}$
for $j = n-1$ downto 0 **do**
  $w_j \leftarrow \Psi(w_{j+1}) - \alpha_j \times w_{j+1} - \beta_{j+1} \times w_{j+2} + c_j$ $\{\Psi$ multiplies by $x\}$
  $d_j \leftarrow p^*_m(w_j)/p^*_m(r)$ {quotient of highest coefficients, $m = \deg(r)$}
  $w_j \leftarrow w_j - d_j \times r$
end for
$s \leftarrow [d_0, \ldots, d_{n-m}]$
$t \leftarrow w_0$ {the result is a polynomial}

Algorithm 1.3: Calculate $s = q \div r$ and $t = q \bmod r$

multiply the polynomials w_j by the variable x efficiently, we need only Property X. However, for the computation of the numbers d_j as the quotient of the highest coefficients we need the assumption on the degree (1.11), as well.

1.7 Bibliographical references

The inequality in Proposition 1.4, obviously, was stated explicitly for the first time (though it is implicit in [Jon94] and thus Jonsson's result). We remark here that not only intervals, but also Cantor-type sets preserve Markov's inequality. This was shown by Białas and Volberg in [BV93] and by Totik [Tot95]. Clearly, the methods of proof extend to the more general case of self-similar supports, which we will consider in the Chapter 6.

The idea of the proof of Lemma 1.2 goes back to our article [Fis96], where we analyzed the sensitivity of recurrence coefficients to perturbations in the modified moments with essentially the same method (compare [Fis96, Theorem 1], where the result is formulated in terms of orthonormal polynomials, however). The method can be refined to be applicable to inner products of Sobolev type (see Section 3.8).

The algorithms for polynomial multiplication and division (Algorithms 1.2 and 1.3) seem to be new.

Chapter 2

Asymptotics of the recurrence coefficients

The asymptotic behavior of the coefficients α_n and β_n is very interesting. A well-known result of Rakhmanov states:

If $S \subset [0,1]$ and

$$\frac{d\sigma(x)}{dx} > 0 \quad \text{a.e. on} \quad [0,1],$$

then

$$\lim_{n\to\infty} \alpha_n = \frac{1}{2} \quad \text{and} \quad \lim_{n\to\infty} \beta_n = \frac{1}{16} \tag{2.1}$$

(the assumptions, of course, imply $S = [0,1]$). For the singular measures considered in [Bes90], the coefficients do not converge. However, in [GW93] it is suggested to investigate the behavior of the arithmetic means

$$\frac{1}{n}\sum_{k=1}^{n} \beta_k \tag{2.2}$$

(though in view of (7) the geometric mean would be more natural. In the case of an almost periodic sequence the arithmetic means, of course, converge).

In this chapter, we show that under appropriate assumptions the limits

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n} \alpha_k, \quad \lim_{n\to\infty} \frac{1}{n}\left(\sum_{k=0}^{n} \alpha_k^2 + 2\sum_{k=1}^{n} \beta_k\right) \quad \text{and} \quad \lim_{n\to\infty} (\beta_1 \dots \beta_n)^{\frac{1}{n}} \tag{2.3}$$

exist. In the case $S = [0,1]$, we can prove a weaker version of (2.1), which we could call *convergence in density*.

Our results in some sense complement the investigations of Mercer [Mer91], where the reverse problem is investigated: If we are given the mean behavior of the sequences (α_n) and (β_n) — what can we say about the limiting distribution of the zeros of $\pi_n(x)$?

2.1 Connection with zero distribution

There is a simple relation between two of the above expressions in (2.3) and the zeros of the polynomial $\pi_n(x)$:

Lemma 2.1 *Let $\tau_1, , \dots, \tau_n$ be the zeros of $\pi_n(x)$. Then*

$$\sum_{k=1}^{n} \tau_k = \sum_{k=0}^{n-1} \alpha_k \tag{2.4}$$

and

$$\sum_{k=1}^{n} \tau_k^2 = \sum_{k=0}^{n-1} \alpha_k^2 + 2\sum_{k=1}^{n-1} \beta_k. \tag{2.5}$$

Proof: We have

$$\pi_n(x) = (x - \tau_1)\dots(x - \tau_n).$$

Thus the coefficient of x^{n-1} in $\pi_n(x)$ is $-(\tau_1 + \dots + \tau_n)$. On the other hand, this coefficient is equal to $-(\alpha_0 + \dots + \alpha_{n-1})$ (see Lemma 1.1), and this proves (2.4). The coefficient of x^{n-2} is

$$\sum_{1\le i<j\le n} \tau_i\tau_j,$$

and must be equal to

$$\sum_{0\le i<j\le n-1} \alpha_i\alpha_j - \sum_{k=1}^{n-1} \beta_k$$

([Chi78, p. 24]). From this we have the elementary identity

$$\begin{aligned}
\sum_{k=1}^{n} \tau_k^2 &= \left(\sum_{k=1}^{n} \tau_k\right)^2 - 2\sum_{1\le i<j\le n} \tau_i\tau_j \\
&= \left(\sum_{k=}^{n-1} \alpha_k\right)^2 - 2\left(\sum_{0\le i<j\le n-1} \alpha_i\alpha_j - \sum_{k=1}^{n-1} \beta_k\right) \\
&= \sum_{k=0}^{n-1} \alpha_k^2 + 2\sum_{k=1}^{n-1} \beta_k,
\end{aligned}$$

and (2.5) is verified.□

Remark: It would be difficult (though possible) to write down the corresponding identities for

$$\sum_{k=1}^{n} \tau_k^m \quad \text{with} \quad m = 3, 4, \dots$$

In fact, we don't need the exact equalities to study limits of mean values, since the following lemma of P. Nevai holds (this is [Nev79, Lemma 5.1]):

Lemma 2.2 *If $S(\sigma)$ is compact and f is continuous then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \left\{ f(\tau_k) - \int f(x) \hat{\pi}_{k-1}^2(x)\, d\sigma(x) \right\} = 0.$$

This could replace (with $f(x) = x$ or $f(x) = x^2$) the equations (2.4) and (2.5), but the above derivation is simpler.

The existence of the limits

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \tau_k \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \tau_k^2$$

or, more general,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(\tau_k)$$

for any continuous function f on $S(\sigma)$, depends on wether the zeros of $\hat{\pi}_n(x)$ have a regular limiting behavior in the sense of Ullman [Ull72], i.e. wether σ is a *regular measure* in the sense of [Ull72] or [ST90]. Some sufficient conditions for regularity will be cited in the next section. We only mention that absolute continuity of σ (and moreover positivity a.e. of the density) is not necessary, in fact already in [Ull72] an example of an atomic regular measure is given.

2.2 Regular measures

For the following we need some facts from potential theory (see [ST90] or [Saf90]):

For a probability measure ν with compact support $S(\nu)$ in the complex plane $\mathbf{C}$ the logarithmic energy of ν is defined as

$$I(\nu) = \int\int \log \frac{1}{|z-t|}\, d\nu(z)\, d\nu(t).$$

If K is a compact subset of $\mathbf{C}$, then with

$$V_K = \inf_{S(\nu)\subset K} I(\nu) \tag{2.6}$$

the (logarithmic) capacity of K is defined as

$$\operatorname{cap}(K) = \exp(-V_K).$$

If $\operatorname{cap}(K) > 0$, then there uniquely exists a probability measure ω_K, called the *equilibrium distribution* of K, at which the infimum in (2.6) is attained. The notion of capacity can be extended to any Borel set B: $\operatorname{cap}(B)$ denotes the supremum of the capacities of compact subsets of B. We say that a property holds qu.e. (quasi everywhere) on a set $S \subset \mathbf{C}$ if it holds on S with possible exceptions on a subset of capacity zero.

Now we are able to define the notion of regularity of the measure: A measure σ with support $S = S(\sigma)$ is called *regular* if

$$\lim_{n\to\infty} \gamma_n(\sigma)^{\frac{1}{n}} = \frac{1}{\mathrm{cap}(S(\sigma))}. \tag{2.7}$$

As it is shown in [ST90], there are several equivalent formulations of regularity, we list only a few of them:

1. $$\limsup_{n\to\infty} |\hat{\pi}_n(x)|^{\frac{1}{n}} = 1 \quad \text{for quasi-every} \quad x \in S(\sigma)$$
   ([ST90, Theorem 3.1(iii)], this is essentially the notion of regularity in [Ull72]).

2. $$\limsup_{n\to\infty} |\hat{\pi}_n(x)|^{\frac{1}{n}} \le 1 \quad \text{for quasi-every} \quad x \in S(\sigma)$$
   ([ST90, Theorem 3.2(iv)]).

3. $$\lim_{n\to\infty} \left(\sup_{x\in S(\sigma)} |\hat{\pi}_n(x)| \right)^{\frac{1}{n}} \le 1$$
   (if $S(\sigma)$is regular with respect to Dirichlet problems in $\mathbf{C} \setminus S(\sigma)$, [ST90, Theorem 3.3(iv)]).

Of course, the definition (2.7) is just what we would like to prove for a measure, i.e. we need some easy to verify criteria for the regularity of a measure. Some of them can be found in [ST90]. For our purposes their "criterion Λ" suffices: The condition

$$\mathrm{cap}\left(\left\{x : \limsup_{r\to 0+} \frac{\log 1/\sigma(B(x,r))}{\log 1/r} < \infty\right\}\right) = \mathrm{cap}(S(\sigma)) \tag{2.8}$$

implies the regularity of σ ([ST90, Theorem 4.8], recall $B(x,r) = \{z : |z-x| \le r\}$). We remark here that the condition (1.10) in the Definition 1.3 is stronger than (2.8): It says that the limit superior in (2.8) is not greater than s for all (not only quasi-every) $x \in S(\sigma)$. The condition (1.10) is satisfied for all examples considered in our investigations.

Since regularity also implies regular limiting zero distribution, we can state our first result:

Theorem 2.3 *Let σ be a regular measure, S its support and ω_S the equilibrium distribution of S. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} \alpha_k = \int x \, d\omega_S(x), \tag{2.9}$$

$$\lim_{n\to\infty} \frac{1}{n} \left(\sum_{k=0}^{n} \alpha_k^2 + 2 \sum_{k=1}^{n} \beta_k \right) = \int x^2 \, d\omega_S(x), \tag{2.10}$$

$$\lim_{n\to\infty} (\beta_1 \dots \beta_n)^{\frac{1}{n}} = \mathrm{cap}(S)^2. \tag{2.11}$$

Proof: For any $n > 0$ let ν_n denote the measure that places mass 1 to every zero of $\hat{\pi}_n$. The regularity of σ implies ([ST90, Theorem 5.8(a)])

$$\frac{1}{n}\nu_n \xrightarrow{*} \omega_S \quad \text{for} \quad n \to \infty,$$

and this means

$$\lim_{n\to\infty} \frac{1}{n} \int f(x)\, d\nu_n(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(\tau_k) = \int f(x)\, d\omega_S(x)$$

for any continuous function f on S. Now for $f(x) = x$ with (2.4) we obtain (2.9), for $f(x) = x^2$ with (2.5) we obtain (2.10). The third limit (2.11) is an immediate consequence of (7) and (2.7).□

Let us illustrate this theorem by two examples.

2.2.1 Bessis' example

In his article [Bes90], D. Bessis investigates orthogonal polynomials with respect to the equilibrium distribution of the Julia set of a polynomial $T(z)$. In general, this is a set in the complex plane. In the special case $T(z) = z^2 - \lambda$ with λ real and $\lambda > 2$, he shows that the Julia set is a Cantor type set on the real line $K \subset [-\xi, \xi]$, where

$$\xi = \frac{1 + \sqrt{1 + 4\lambda}}{2}$$

(one easily shows $2 < \xi < \lambda$). Here $\sigma = \omega_K$ is symmetric and invariant under the transformation $T(x) = x^2 - \lambda$, and we obtain

$$\int x\, d\omega_K(x) = 0 \quad \text{and} \quad \int (x^2 - \lambda)\, d\omega_K(x) = \int x\, d\omega_K(x) = 0.$$

Consequently,

$$\int x^2\, d\omega_K(x) = \int \lambda\, d\omega_K(x) = \lambda.$$

The regularity of σ follows from the explicit bound

$$|\hat{\pi}_n(x)| < \frac{1}{2} \left(\frac{4\sqrt{\lambda}}{\sqrt{1 + 4\lambda} - 1} \right)^k \quad \text{for} \quad 2^{k-1} \le n \le 2^k,$$

when $x \in K$ (this is [Bes90, eq. IV.141]), since this implies

$$\limsup_{n\to\infty} |\hat{\pi}_n(x)|^{\frac{1}{n}} \le 1 \quad \text{for} \quad x \in K,$$

one of the equivalent formulations of regularity. Thus our theorem 2.3 is applicable, but we need the capacity of K. Implicitely, it is calculated in [Bes90, eq.

II.35]: if we compare this with the general formula for the Green function (e.g. [ST90, p. 397]), we see that $\log \operatorname{cap}(K) = 0$, i.e.

$$\operatorname{cap}(K) = 1.$$

For the symmetric measure σ, of course, all α_n are zero, so (2.9) is trivial, and (2.10) becomes

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \beta_k = \frac{\lambda}{2}. \tag{2.12}$$

(2.11) takes the form

$$\lim_{n\to\infty} (\beta_1 \dots \beta_n)^{\frac{1}{n}} = 1. \tag{2.13}$$

We can now compare our results with the recurrences derived by Bessis: his system of equations [Bes90, IV.105] in our notations reads

$$\beta_{2k}\beta_{2k-1} = \beta_k \tag{2.14}$$

$$\beta_{2k} + \beta_{2k+1} = \lambda \tag{2.15}$$

$$\beta_1 = \lambda \tag{2.16}$$

for $k \geq 2$. From (2.14) easily follows (2.13) (if the existence of the limit is known), (2.15) immediately gives (2.12). Observe that for a measure satisfying the conditions of Rakhmanov on the interval $[-\xi, \xi]$ we would have (after an obvious rescaling)

$$\lim_{n\to\infty} \beta_n = \frac{\xi^2}{4},$$

which is strictly between 1 and $\lambda/2$ (since $\xi^2 = \xi + \lambda$ and $2 < \xi < \lambda$).

2.2.2 Cantor measure

In the article [GW93] moments of the Cantor measure (together with interesting asymptotics) are investigated and the first recurrence coefficients β_n are generated. We will not give a detailed construction of the Cantor set and the Cantor measure here, but only recall a few facts sufficient to prove the regularity of the measure: The Cantor set C is defined by

$$C = \bigcap_{n=0}^{\infty} \bigcup_{k=1}^{2^n} I_{nk},$$

where for any $n \geq 0$ the $I_{nk} \subset [0,1]$ are disjoint intervals of length 3^{-n}. The Cantor measure σ is defined by

$$\sigma(I_{nk}) = 2^{-n}.$$

Let $3^{-n} \leq r \leq 3^{-n+1}$ and $x \in C$. This means that there is a $k = k(x)$ with $x \in I_{nk}$. Since the length of I_{nk} is $3^{-n} \leq r$, we have $B(x,r) \supset I_{nk}$ and

$$\sigma(B(x,r)) \geq \sigma(I_{nk}) = 2^{-n} \geq \left(\frac{r}{3}\right)^{\alpha},$$

where

$$\alpha = \frac{\log 2}{\log 3}.$$

Consequently, we have an inequality (1.10), and criterion Λ (2.8) is satisfied (see the discussion after the formulation of this criterion). We have proved

Lemma 2.4 *The Cantor measure σ is regular.*

The measure σ is symmetric about $\frac{1}{2}$, hence $\alpha_n \equiv \frac{1}{2}$.

this means that Equation (2.9) is again trivial. The propositions (2.10) and (2.11) of our theorem 2.3 now give

Corollary 2.5 *The recurrence coefficients β_n for the Cantor measure satisfy*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \beta_k = \frac{1}{2}\left(\int x^2 \, d\omega_C(x) - \frac{1}{4}\right) \tag{2.17}$$

and

$$\lim_{n\to\infty} (\beta_1 \ldots \beta_n)^{\frac{1}{n}} = \operatorname{cap}(C)^2. \tag{2.18}$$

Unfortunately, this is far less explicit than (2.12) and (2.13), since $\operatorname{cap}(C)$ and ω_C are unknown (to the author).

2.3 The case $S = [0,1]$

We could now specify our result to the case $S = [0,1]$, where

$$\operatorname{cap}(S) = \frac{1}{4} \quad \text{and} \quad d\omega_S(x) = \frac{1}{\pi\sqrt{x(1-x)}}\, dx,$$

but instead we shall prove a stronger theorem in a more direct way. First we need some elementary estimates.

Lemma 2.6 *Let $S = [0,1]$. Then for the recurrence coefficients β_n we have the inequality*

$$\sqrt{\beta_1} + \ldots + \sqrt{\beta_n} \leq \frac{n}{4}. \tag{2.19}$$

Proof: It was mentioned already in Section 0.2 that the zeros of $\hat{\pi}_n(x)$ are the eigenvalues of the Jacobi matrix

$$J_n = \begin{bmatrix} \alpha_0 & \sqrt{\beta_1} & & & & 0 \\ \sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \sqrt{\beta_{n-1}} & \alpha_{n-1} & \sqrt{\beta_n} \\ 0 & & & & \sqrt{\beta_n} & \alpha_n \end{bmatrix}$$

(see Theorem 0.6). These are all in $(0,1)$ and consequently

$$0 \le \langle x, J_n x \rangle \le \langle x, x \rangle \quad \text{for any} \quad x \in \mathbf{R}^n.$$

From the first inequality with $x = (1, -1, 1, \ldots)^T$ we obtain

$$\sum_{k=0}^{n} \alpha_k - 2 \sum_{k=1}^{n} \sqrt{\beta_k} \ge 0,$$

from the second with $x = (1, 1, 1, \ldots)^T$

$$\sum_{k=0}^{n} \alpha_k + 2 \sum_{k=1}^{n} \sqrt{\beta_k} \le n.$$

Subtracting these two inequalities and dividing by 4 we arrive at (2.19).□

Lemma 2.7 *Let $S(\sigma) = [0,1]$ and*

$$M_n = \sup_{x \in [0,1]} |\hat{\pi}_n(x)|.$$

Then

$$\frac{M_n}{\gamma_n} \ge 2^{1-2n} \quad \textit{for} \quad n \ge 1. \tag{2.20}$$

Proof: We have

$$\frac{M_n}{\gamma_n} = \sup_{x \in [0,1]} |\pi_n(x)|,$$

(compare the proof of [Saf90, Proposition 3.2]) and thus the proposition is a consequence of the fact that of all monic polynomials the monic shifted Chebyshev polynomial has minimal sup-norm on $[0,1]$. □

Lemma 2.8 *If $S(\sigma) = [0,1]$ and*

$$M_n = \sup_{x\in[0,1]} |\hat{\pi}_n(x)|,$$

then the inequality

$$\sum_{k=1}^{n} \left(4\sqrt{\beta_k} - 1 - \log 4\sqrt{\beta_k}\right) \leq -\log \frac{2}{\sqrt{\beta_0}} + \log M_n$$

holds for $n \geq 1$.

Proof: The inequality follows from (7), (2.20) and (2.19) just by elementary algebra.□

The preceding lemma says nothing about the α_n, so we have to investigate the symmetrization ν of σ, first (see Section 1.1). Of course, we must have a result like

Lemma 2.9 *If $S(\sigma) = [0,1]$ and σ is regular, then ν is regular, too.*

Proof: Evidently, $S(\nu) = [-1,1]$. The regularity of σ by definition (2.7) implies

$$\lim_{n\to\infty} \gamma_n(\sigma)^{\frac{1}{n}} = \frac{1}{\text{cap}(S(\sigma))} = 4.$$

To verify the regularity of ν, we have to find a relation between $\gamma_n(\sigma)$ and $\gamma_n(\nu)$. But from (7) and (1.4) we have

$$\frac{1}{\gamma_{2n}(\nu)} = (\beta_0(\nu)\dots\beta_{2n}(\nu))^{\frac{1}{2}} = (\beta_0(\sigma)\dots\beta_n(\sigma))^{\frac{1}{2}} = \frac{1}{\gamma_n(\sigma)}$$

and (since all $\beta_n(\nu) < 1$ by virtue of Inequality (11))

$$\frac{1}{\gamma_{2n}(\nu)} \geq \frac{1}{\gamma_{2n+1}(\nu)} \geq \frac{1}{\gamma_{2n+2}(\nu)},$$

and consequently

$$\lim_{n\to\infty} \gamma_n(\nu)^{\frac{1}{n}} = \lim_{n\to\infty} \gamma_{2n}(\nu)^{\frac{1}{2n}} = \lim_{n\to\infty} \gamma_n(\sigma)^{\frac{1}{2n}} = \sqrt{4} = 2.$$

This means the regularity of ν, because of $\text{cap}([-1,1]) = \frac{1}{2}$. □

To formulate our next result, we need some definitions:

Let M be a subset of the natural numbers. We call the limit

$$d(M) = \limsup_{n\to\infty} \frac{\#\{k \leq n : k \in M\}}{n}$$

the *density* of M. We say that a sequence (a_n) converges to a *in density* ($a_n \stackrel{d}{\to} a$) if

$$d(\{n : |a_n - a| \geq \delta\}) = 0 \quad \text{for all} \quad \delta > 0.$$

This is weaker than convergence (in the case of convergence, the set $\{n : |a_n - a| \geq \delta\}$ is finite for all $\delta > 0$, and a finite set has density zero). But it is meaningful, as the following lemma shows:

Lemma 2.10 *$a_n \xrightarrow{d} a$ if and only if*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(a_k) = f(a) \tag{2.21}$$

for all bounded continuous functions f.

Proof: Suppose first $a_n \xrightarrow{d} a$. Then we have

$$\left|\frac{1}{n}\sum_{k=1}^{n} f(a_k) - f(a)\right| = \left|\frac{1}{n}\sum_{k=1}^{n}(f(a_k) - f(a))\right| \le \frac{1}{n}\sum_{k=1}^{n}|f(a_k) - f(a)|.$$

Now if f is bounded, $|f(x)| \le K$ for all x. Let ε be any real number > 0, then there exists a $\delta > 0$ with $|f(x) - f(a)| < \varepsilon$ whenever $|x - a| < \delta$. If we write

$$N_\delta = \{n : |a_n - a| \ge \delta\}, \tag{2.22}$$

then we have $|f(a_k) - f(a)| < 2K$ for $k \in N_\delta$ and $|f(a_k) - f(a)| < \varepsilon$ for $k \notin N_\delta$, consequently

$$\frac{1}{n}\sum_{k=1}^{n}|f(a_k) - f(a)| \le \varepsilon + \frac{2K}{n}\#\{k \le n : k \in N_\delta\}.$$

From this follows

$$\limsup_{n\to\infty}\left|\frac{1}{n}\sum_{k=1}^{n} f(a_k) - f(a)\right| \le \limsup_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}|f(a_k) - f(a)| \le \varepsilon,$$

i.e.

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} f(a_k) = f(a),$$

since $\varepsilon > 0$ was arbitrary.

To prove the "if"-part, suppose (2.21) holds. For $\delta > 0$ let

$$f(x) = \min\{\frac{1}{\delta}|x - a|, 1\}.$$

Obviously, f is continuous and bounded, and with the notation (2.22) we have

$$\frac{1}{n}\sum_{k=1}^{n} f(a_k) \ge \frac{1}{n}\#\{k \le n : k \in N_\delta\}.$$

The equation (2.21) for this special function f now immediately gives $d(N_\delta) = 0$. □

In a similar way, one easily shows that sums and products of convergent in density sequences are convergent in density (to the sum or product of the d-limit).

The main result of this section is the following

Theorem 2.11 *Let σ be regular and $S(\sigma) = [0,1]$. Then*

$$\beta_n \xrightarrow{d} \frac{1}{16} \quad \text{and} \quad \alpha_n \xrightarrow{d} \frac{1}{2}.$$

Proof: First we prove the proposition concerning β_n. Let again

$$M_n = \sup_{x\in[0,1]} |\hat{\pi}_n(x)|.$$

The equivalent formulation 3 of regularity (the condition for $S(\sigma)$ is satisfied) gives

$$\lim_{n\to\infty} M_n^{\frac{1}{n}} = 1,$$

i.e.

$$\lim_{n\to\infty} \frac{1}{n} \log M_n = 0.$$

Now we can apply our Lemma 2.8. The elementary inequality

$$\log x \le x - 1 \quad \text{for} \quad x > 0,$$

shows that the sum in this lemma is positive! Moreover, substituting $x^{\frac{1}{2}}$ instead of x, we obtain

$$x - 1 - \log x \ge x - 1 - 2(x^{\frac{1}{2}} - 1) = (x^{\frac{1}{2}} - 1)^2.$$

Thus,

$$\frac{1}{n}\sum_{k=1}^{n}\left(4\sqrt{\beta_k} - 1 - \log 4\sqrt{\beta_k}\right) \ge \frac{1}{n}\sum_{k=1}^{n}(2\beta_k^{\frac{1}{4}} - 1)^2$$
$$\ge \frac{\delta^2}{n}\#\{k \le n : |2\beta_k^{\frac{1}{4}} - 1| \ge \delta\}$$

for any $\delta > 0$. Since the lefthand side converges to zero as $n \to \infty$ by Lemma 2.8, we have $2\beta_n^{\frac{1}{4}} \xrightarrow{d} 1$ and this immediately gives $\beta_n \xrightarrow{d} \frac{1}{16}$.

Now we are going to prove the second proposition concerning α_n. To do this we consider the symmetrisation ν of σ, by Lemma 2.9 it is regular and we have $S(\nu) = [-1,1]$. What we have shown above implies $\beta_n(\nu) \xrightarrow{d} \frac{1}{16}[1-(-1)]^2 = \frac{1}{4}$ by virtue of the scaling properties of recurrence coefficients (compare [Chi78, Chapter I, ex. 4.4(a)]). From this by (1.3) and (1.4) follows $\alpha_n(\sigma) \xrightarrow{d} \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$ and $\beta_n(\sigma) \xrightarrow{d} \frac{1}{4}\cdot\frac{1}{4} = \frac{1}{16}$. □

Remark: As a result of Mercer in his note [Mer91] shows, our Theorem 2.11 is nearly optimal: If we know that $\beta_n \xrightarrow{d} \frac{1}{16}$ and $\alpha_n \xrightarrow{d} \frac{1}{2}$, then his Theorem 3 implies that the asymptotic zero distribution is $\omega_{[0,1]}$. This, of course, is not fully equivalent, but closely related to regularity of the measure (see the discussion in Chapter 3 of [ST92]).

To illustrate our Theorem 2.11, we consider two examples of self-similar distributions σ with $S(\sigma) = [0,1]$.

2.3.1 Riesz-Nagy distributions

Another example for singular measures investigated in [GW93] are the Riesz-Nagy singular distributions (more precisely, a special case of them). We need more information about them than given in the cited paper [GW93], so we recall the construction here:

Suppose (τ_n) is a sequence of real numbers, $0 < \tau_n < 1$. Then we define inductively a sequence of continuous functions $F_n(x)$ on $[0, 1]$ as follows. Let $F_0(x) = x$, for $n \geq 0$ define $F_{n+1}(x)$ by

$$\begin{aligned} F_{n+1}\left(\frac{k}{2^n}\right) &= F_n\left(\frac{k}{2^n}\right), \qquad k = 0, 1, \ldots, 2^n; \\ F_{n+1}\left(\frac{2k+1}{2^{n+1}}\right) &= \frac{1-\tau_{n+1}}{2} F_n\left(\frac{k}{2^n}\right) + \frac{1+\tau_{n+1}}{2} F_n\left(\frac{k+1}{2^n}\right) \end{aligned}$$

for $k = 0, 1, \ldots, 2^n - 1$ and $F_{n+1}(x)$ linear in the intervals $\left[\frac{k}{2^{n+1}}, \frac{k+1}{2^{n+1}}\right]$ for $k = 0, 1, \ldots, 2^{n+1} - 1$. Finally, we define the function $F(x)$ by

$$F(x) = \lim_{n \to \infty} F_n(x).$$

F is continuous and strictly increasing, so there is a unique measure σ with distribution function $\sigma((-\infty, x)) = F(x)$. This measure is called Riesz-Nagy distribution. In order to apply our results, we need

Lemma 2.12 *If $\tau_n \leq 1 - \delta < 1$ for $n \geq 1$, then $S(\sigma) = [0, 1]$ and σ is regular.*

Proof: Let

$$I_{n,k} = \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right], \quad k = 0, 1, \ldots, 2^n - 1$$

for $n \geq 0$. We show inductively $\sigma(I_{n,k}) \geq (\delta/2)^n$ for all $0 \leq k \leq 2^n - 1$: For $n = 0$ this is trivial. Now we observe that the definition of F_{n+1} implies

$$F_m\left(\frac{k}{2^n}\right) = F_n\left(\frac{k}{2^n}\right), \qquad k = 0, 1, \ldots, 2^n$$

for all $m \geq n + 1$ and thus

$$F\left(\frac{k}{2^n}\right) = F_n\left(\frac{k}{2^n}\right).$$

Consequently, we have

$$\sigma(I_{n,k}) = F_n\left(\frac{k+1}{2^n}\right) - F_n\left(\frac{k}{2^n}\right).$$

From the definition now follows

$$\sigma(I_{n+1,2k}) = \frac{1+\tau_{n+1}}{2}\sigma(I_{n,k}) \geq \frac{1-\tau_{n+1}}{2}\sigma(I_{n,k}) \geq \frac{\delta}{2}\sigma(I_{n,k})$$

and

$$\sigma(I_{n+1,2k+1}) = \frac{1-\tau_{n+1}}{2}\sigma(I_{n,k}) \geq \frac{\delta}{2}\sigma(I_{n,k}),$$

and we are done.

If $x \in [0,1]$ and $2^{-n} \leq r \leq 2^{-n+1}$, then the interval $B(x,r)$ contains one of the $I_{n,k}$, and it must be

$$\sigma(B(x,r)) \geq \left(\frac{\delta}{2}\right)^n \geq \left(\frac{r}{2}\right)^\alpha,$$

where

$$\alpha = \frac{\log\frac{2}{\delta}}{\log 2}.$$

This proves both $x \in S(\sigma)$ and regularity of σ. □

Application of our theorem 2.11 gives

Corollary 2.13 *For the Riesz-Nagy distributions with $\tau_n \leq 1-\delta < 1$ for $n \geq 1$ we have*

$$\beta_n \xrightarrow{d} \frac{1}{16} \quad \textit{and} \quad \alpha_n \xrightarrow{d} \frac{1}{2}.$$

The result is stronger than convergence of arithmetical means and it recovers the special case $\tau_n \equiv \tau \in (0,1)$ considered in [GW93]. But nevertheless we feel that it is not strong enough — here we should have convergence in the usual sense.

2.3.2 An atomic measure

Our last example shows that the recurrence coefficients may behave quiet regularly even in the case of an atomic (or discrete) measure:

We define

$$\sigma\left(\left\{\frac{m}{2^n}\right\}\right) = 3^{-n}$$

for $n \geq 1$, m odd and $0 < m < 2^n$, and $\sigma(B) = 0$ for any Borel set B not containing a dyadic rational. The dyadic rationals are dense in $[0,1]$, so we have $S(\sigma) = [0,1]$. Any interval of length 2^{-n} contains a number of the form $m2^{-n}$, and we easily prove $\sigma(B(x,r)) \geq (r/2)^\alpha$ with $\alpha = \log 3/\log 2$, i.e. we have

Lemma 2.14 *The atomic measure σ is regular.*

By theorem 2.11 we obtain (note that we have $\alpha_n \equiv \frac{1}{2}$ by symmetry)

Corollary 2.15 *For the measure defined above the coefficients β_n satisfy*

$$\beta_n \xrightarrow{d} \frac{1}{16}.$$

Again we feel that in fact the coefficients simply converge to $\frac{1}{16}$. This is suggested by numerical experiments: The measure has a simple self-similar structure, and one easily shows

$$\int f(x)\, d\sigma(x) = \frac{1}{3}\left[f\left(\frac{1}{2}\right) + \int f\left(\frac{x}{2}\right)\, d\sigma(x) + \int f\left(\frac{x+1}{2}\right)\, d\sigma(x)\right]$$

for any bounded function f. From this equation we can derive simple recurrences not only for ordinary moments

$$\int x^n\, d\sigma(x),$$

but for modified moments

$$\int T_n^*(x)\, d\sigma(x),$$

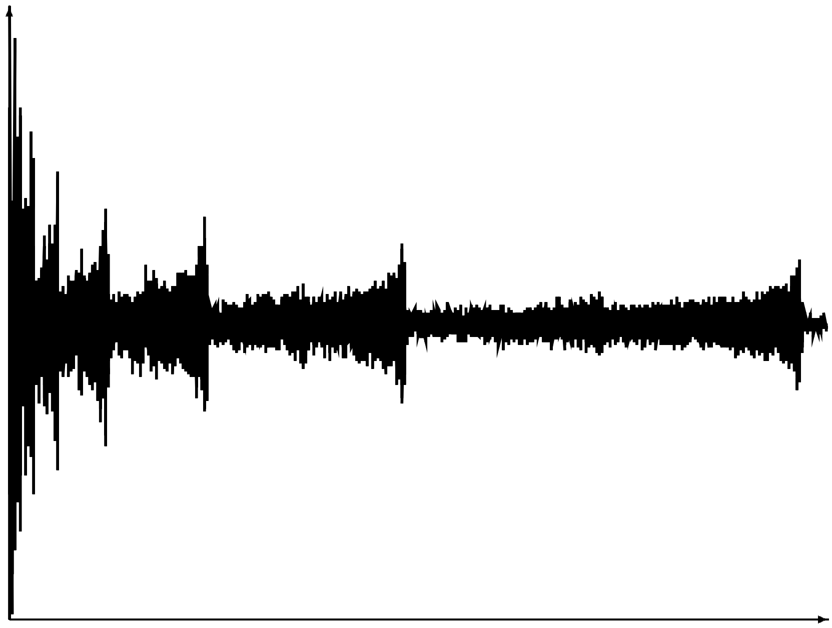
too, where $T_n^*(x) = T_n(2x-1)$ are the shifted Chebyshev polynomials of first kind. With an own implementation of the method of modified moments (see [Gau90]) adapted to Chebyshev moments we can calculate numerically some β_n beyond the $n = 18$ in [GW93] (actually up to $n = 5000$), and the algorithm seems to be stable. The values

$$\varepsilon_n = \max_{2^{n-1} \le k < 2^n} \left| \beta_k - \frac{1}{16} \right|$$

should give some feeling for the long-range behavior of the coefficients β_n:

n	ε_n
4	$3.05E-02$
5	$2.72E-02$
6	$2.04E-02$
7	$1.80E-02$
8	$1.47E-02$
9	$1.14E-02$
10	$9.76E-03$
11	$7.55E-03$
12	$6.23E-03$

Indeed, these figures (and rather similar results for the Riesz-Nagy distribution with $\tau_n \equiv \frac{1}{2}$) suggest convergence. The following plot of the first about 3000 coefficients β_j seems to confirm this expectation:



2.4 Bibliographical references

The class of orthogonal polynomials (respectively of their measures of orthogonality) with converging recurrence coefficients is sometimes called *Nevai's class*, since Nevai investigated them thoroughly in his famous memoir [Nev79]. However, he did not give any sufficient condition for a measure to belong to this class. For measures in Szegö's class, much stronger asymptotics are known, but the assumptions are also much stronger. Especially, these measures are absolutely continuous w.r.t. Lebesgue measure, whereas there exist atomic measures in Nevai's class. Examples were given in Lubinsky's paper [Lub88]. The same results were obtained later by van Assche and Magnus in their paper [vAM89] using other techniques.

Rakhmanov obtained his result (2.1) in the paper [Rak77]. Unfortunately, his proof contained a gap, as Máté, Nevai and Totik pointed out. In [Rak82], Rakhmanov gave a correct proof of his result. In the meantime, other proofs were found, but all are long and technical (and reduce the problem to measures on the unit circle).

The class of regular measures has still weaker so-called nth root asymptotics. Potential-theoretic methods of proof were used already by Ullman in [Ull72]. Since potential theory in the complex plane (dimension 2) is rather special, the relevant results had to be stated in many articles on polynomials. Most of them referred to the book of Tsuji [Tsu58], though Doob considered the case of dimension 2 (and even of dimension 1) in his book [Doo84]. The already mentioned book [NS88] has a chapter on potential theory. The standard reference book now is the book of Saff and Totik [ST97]. In the article of Stahl and Totik [ST90], they formulate regularity criteria, the detailed proofs of which appeared in the book [ST92].

The results of this chapter (which first appeared in our paper [Fis95b]) were inspired by the article [GW93], where Goh and Wimp conjectured the convergence of arithmetic means of recurrence coefficients. The general Theorem 2.3 seems not to be very deep. If the support of the measure is an interval, we can prove much more. In fact, in this case the proof of Theorem 2.11 together with Mercer's Theorem 3 in [Mer91] show that our convergence in density of the recurrence coefficients is *equivalent* to regular zero asymptotics of the orthogonal polynomials — and this is a rather surprising fact (*regular zero asymptotics* is formulated in the proof of Theorem 2.3).

Of course, convergence in density does not exclude the possibility for a subsequence of density zero to diverge (or converge to another limit). As examples show, this is indeed possible even if the stronger condition (1.10) is satisfied. The most striking example can be found in the marvellous paper of Magnus [Mag95]: In Theorem 2 he (explicitly!) constructs polynomials, the measure of orthogonality of which has a countable set of mass points dense in an interval. The recurrence coefficients are almost constant, the exceptional set of indices has indeed density zero (and can be related to the Fibonacci numbers, if Magnus' parameter $\theta/(2\pi)$ is equal to the golden ratio $(1+\sqrt{5})/2$). It can be easily seen that condition (1.10) is satisfied at least for algebraic values of $\theta/(2\pi)$.

This shows that there hardly exists a simple proof for the conjecture that our Example 2.3.2 is in Nevai's class. We remark here that our use of the method of modified moments for this example will be justified in the following chapter: The measure satisfies the conditions of Corollary 3.10.

Chapter 3

From modified moments to recurrence coefficients

It is well known, that the generation of the orthogonal polynomials (or, equivalently, of the coefficients α_j and β_j) from ordinary moments

$$\mu_k = \int x^k \, d\sigma(x)$$

is severely ill-conditioned, see the classical paper of Walter Gautschi [Gau78]. More promising are modified moments

$$m_k = \int p_k(x) \, d\sigma(x), \tag{3.1}$$

with some properly chosen set of polynomials p_k (e.g. Chebyshev polynomials). Of course, we should be able to calculate or estimate the condition of the map K_n from $[m_0, \ldots, m_{2n-1}]^T$ to $[\alpha_0, \ldots, \alpha_{n-1}, \beta_0, \ldots, \beta_{n-1}]^T$. In [Gau78] and [Gau90], the map G_n from modified moments to the vector of nodes and weights of a Gauss quadrature rule is considered, instead, and from this estimates are given for the condition of our original map K_n. Fortunately, a more direct approach is possible.

3.1 Maps and norms

If we denote by $m \in \mathbf{R}^{2n}$ the vector of the first $2n$ modified moments $[m_0, \ldots, m_{2n-1}]^T$ and by ρ the vector of the recursion coefficients

$$[\alpha_0, \ldots, \alpha_{n-1}, \beta_0, \ldots, \beta_{n-1}]^T,$$

then we are interested in the ition of the map

$$K_n : \mathbf{R}^{2n} \to \mathbf{R}^{2n} \quad m \to \rho.$$

In order to compare our results with those of Walter Gautschi [Gau78], [Gau86] and [Gau90], we introduce the vector

$$\gamma = [\tau_1, \dots, \tau_n, \sigma_1, \dots, \sigma_n]^T$$

of nodes and weights of the Gauss-Christoffel quadrature rule

$$\int q(x)\, d\sigma(x) = \sum_{k=1}^{n} \sigma_k\, q(\tau_k)$$

valid for any polynomial q of degree $\leq 2n-1$ (see Section 0.2), and denote by G_n the map

$$G_n : \mathbf{R}^{2n} \to \mathbf{R}^{2n} \quad m \to \gamma,$$

and by H_n the map

$$H_n : \mathbf{R}^{2n} \to \mathbf{R}^{2n} \quad \gamma \to \rho.$$

Then one has the decomposition

$$K_n = H_n \circ G_n,$$

and this was used by Gautschi to reduce the investigation of K_n to that of G_n (the map H_n is usually well-conditioned, see Chapter 4). More precisely, Gautschi considered the map defined by

$$\tilde{G}_n : \mathbf{R}^{2n} \to \mathbf{R}^{2n} \quad \tilde{m} \to \gamma,$$

where

$$\tilde{m}_k = d_k^{-\frac{1}{2}}\, m_k \quad \text{with} \quad d_k = \int p_k^2(x)\, ds(x),$$

and the polynomials p_k are orthogonal w.r.t. the measure s. The map $\tilde{G}_n$ differs from G_n only by a trivial diagonal map. The sensitivity of those nonlinear maps can be measured by the norm of the Fréchet derivatives K_n' and G_n', which are linear maps (the Jacobians of K_n and G_n) from $\mathbf{R}^{2n}$ into $\mathbf{R}^{2n}$. We consider here only two norms of matrices $A = (a_{j,k})_{j,k=0,\dots,2n-1}$: The supremum-norm

$$\|A\|_\infty = \max_{0 \leq j \leq 2n-1} \sum_{k=0}^{2n-1} |a_{jk}|$$

and the Frobenius norm

$$\|A\|_F = (\mathrm{Tr}\, A^T A)^{1/2} = \left(\sum_{j,k=0}^{2n-1} a_{jk}^2 \right)^{1/2}.$$

We mention here that the condition number defined in [Gau78, 2.1] for the map $\gamma = G_n(m)$ in our notations will be

$$\mathrm{cond}(G_n, m) = \frac{\|m\|}{\|G_n(m)\|}\, \|G_n'(m)\| \tag{3.2}$$

(compare [Gau78, 2.2]).

The partial derivatives constituting the Jacobian K_n' will be evaluated in the following section.

3.2 Calculation of some partial derivatives

We assume that $(p_j)_{n=0,\ldots,N}$ is any polynomial base in the space $\mathcal{P}_N$ of all polynomials of degree $\leq N$, i.e. there is a representation

$$q(x) = \sum_{j=0}^{N} c_j(q) p_j(x) \tag{3.3}$$

for all polynomials $q(x) \in \mathcal{P}_N$, with some linear functionals c_j. Note that we do *not* assume $\deg(p_j) = j$, i.e. the case of Bernstein polynomials or Lagrange interpolation polynomials are included! The case of ordinary moments is included as well, here we have $p_j(x) = x^j$ and $c_j(q) = q^{(j)}(0)/j!$ (remember that we are in a finite-dimensional space, where all norms are equivalent and all linear operations are continuous).

Lemma 3.1 *Under our assumptions (3.1) and (3.3), we obtain the equations*

$$\frac{\partial \alpha_j}{\partial m_k} = \gamma_j^2 c_k(\pi_j \pi_{j+1}) - \gamma_{j-1}^2 c_k(\pi_{j-1}\pi_j)$$

and

$$\frac{\partial \beta_j}{\partial m_k} = \beta_j[\gamma_j^2 c_k(\pi_j^2) - \gamma_{j-1}^2 c_k(\pi_{j-1}^2)].$$

Proof: These equations are an almost trivial consequence of our Lemma 1.2. From (3.3) and (3.1) we have immediately

$$l(q) = \int q(x)\, d\sigma(x) = \sum_{k=0}^{N} c_k(q) m_k,$$

and this implies

$$\frac{\partial l}{\partial m_k} = c_k. \square$$

The Jacobian K_j' consists of the partial derivatives

$$\frac{\partial \alpha_j}{\partial m_k} \quad \text{and} \quad \frac{\partial \beta_j}{\partial m_k} \quad \text{with} \quad j = 0, \ldots, n-1,\ k = 0, \ldots, 2n-1,$$

so it is clear that we must have a polynomial base with $N \geq 2n-1$. Now we can introduce the following notations:

$$\psi_{2j}(x) = \beta_j[\gamma_j^2 \pi_j^2(x) - \gamma_{j-1}^2 \pi_{j-1}^2(x)] \tag{3.4}$$

and

$$\psi_{2j+1}(x) = \gamma_j^2 \pi_j(x)\pi_{j+1}(x) - \gamma_{j-1}^2 \pi_{j-1}(x)\pi_j(x) \tag{3.5}$$

for $j = 0, \ldots, n-1$. Then ψ_j is a polynomial of degree j, and obviously $\|K_n'\|_\infty = \|\Psi\|_\infty$ and $\|K_n'\|_F = \|\Psi\|_F$, where $\Psi = (\psi_{jk})_{j,k=0,\ldots,2n-1}$ with $\psi_{jk} = c_k(\psi_j)$. These norms can be estimated (or evaluated exactly in some special cases) for various choices of polynomial bases $(p_j)_{n=0,\ldots,N}$. We first reconsider ordinary moments, investigate the (somewhat exotic) bases of Bernstein polynomials and Lagrange interpolation polynomials, and finally we are concerned with the practically important case, where the p_j are orthonormal polynomials with respect to some other measure s.

To evaluate the norms, we just have to find a way to express

$$a_\infty(q) = \sum_{k=0}^{N} |c_k(q)|$$

and

$$a_2(q) = \sum_{k=0}^{N} c_k^2(q)$$

for all $q \in \mathcal{P}_N$. Then evidently we will have

$$\|K_n'\|_\infty = \|\Psi\|_\infty = \max_{0\le j\le 2n-1} a_\infty(\psi_j) \tag{3.6}$$

and

$$\|K_n'\|_F = \|\Psi\|_F = \left(\sum_{j=0}^{2n-1} a_2(\psi_j)\right)^{1/2}. \tag{3.7}$$

One easily observes that $a_\infty(\cdot)$ and $a_2(\cdot)^{1/2}$ satisfy the norm inequalities

$$a_\infty(q+r) \le a_\infty(q) + a_\infty(r)$$

and

$$a_2(q+r)^{1/2} \le a_2(q)^{1/2} + a_2(r)^{1/2}$$

for $q, r \in \mathcal{P}_N$ (since the coordinate functionals c_k are linear).

3.3 Some examples

In order to illustrate our results, we consider Chebyshev polynomials of first and second kind.

3.3.1 Chebyshev polynomials of first kind

Chebyshev polynomials of first kind are defined by the identity

$$T_j(\cos\theta) = \cos j\theta, \quad j = 0, 1, \ldots$$

(this is [AS84, Equation 22.3.15]). These polynomials are the orthogonal polynomials with respect to the measure

$$d\sigma(x) = \frac{dx}{\pi\sqrt{1-x^2}} \quad \text{for} \quad x \in [-1, 1]$$

(observe that we use the normalization $\sigma([-1,1]) = 1$, i.e. $\beta_0 = 1$). In this case, we have $\alpha_j = 0$ by symmetry, and

$$\beta_j = \begin{cases} 1/2 & \text{for} \quad j = 1 \\ 1/4 & \text{for} \quad j > 1 \end{cases},$$

the orthonormal polynomials will be

$$\hat{\pi}_j = \begin{cases} 1 & \text{for} \quad j = 0 \\ \sqrt{2}\,T_j & \text{for} \quad j > 0 \end{cases}.$$

Then, the polynomials ψ_j defined above can be calculated easily (using the well-known identities for Chebyshev polynomials, see [AS84, 22.7.24]):

$$\psi_{2j} = \begin{cases} 1 & \text{for} \quad j = 0 \\ \frac{1}{2}T_2 & \text{for} \quad j = 1 \\ \frac{1}{4}(T_{2j} - T_{2j-2}) & \text{for} \quad j > 1 \end{cases} \tag{3.8}$$

and

$$\psi_{2j+1} = \begin{cases} T_1 & \text{for} \quad j = 0 \\ \frac{1}{2}(T_{2j+1} - T_{2j-1}) & \text{for} \quad j > 0 \end{cases}. \tag{3.9}$$

To compare our results with those of Gautschi [Gau78], we consider shifted Chebyshev polynomials as well (we will mark any quantity related to the shifted polynomials with an asterisk): The measure here is defined by

$$d\sigma^*(x) = \frac{dx}{\pi\sqrt{x(1-x)}} \quad \text{for} \quad x \in [0, 1],$$

again it has total weight 1, and the orthonormal polynomials are connected with the above via

$$\hat{\pi}_j^*(x) = \hat{\pi}_j(2x-1) = \begin{cases} 1 & \text{for} \quad j = 0 \\ \sqrt{2}\,T_j^*(x) & \text{for} \quad j > 0 \end{cases},$$

where the common notation for shifted Chebyshev polynomials of first kind is used (see [AS84, 22.5.14]):

$$T_n^*(x) = T_n(2x-1).$$

The recurrence coefficients of these shifted polynomials can be readily determined from their scaling properties:

$$\alpha_j^* = \frac{\alpha_j + 1}{2} = \frac{1}{2}$$

and

$$\beta_j^* = \frac{1}{2^2}\beta_j = \begin{cases} 1/8 & \text{for} \quad j = 1 \\ 1/16 & \text{for} \quad j > 1 \end{cases}$$

(this is essentially [Chi78, Chapter I, exercise 4.4(a)]). Of course, we have $\beta_0^* = 1$. From (3.4) we obtain

$$\psi_{2j}^*(x) = \frac{1}{4}\psi_{2j}(2x-1) \quad \text{for} \quad j \geq 1$$

and from (3.5)

$$\psi_{2j+1}^*(x) = \frac{1}{2}\psi_{2j+1}(2x-1) \quad \text{for} \quad j \geq 0,$$

clearly ψ_0^* must be $\equiv 1$. Thus, the explicit formulas will read

$$\psi_{2j}^* = \begin{cases} 1 & \text{for} \quad j = 0 \\ \frac{1}{8}T_2^* & \text{for} \quad j = 1 \\ \frac{1}{16}(T_{2j}^* - T_{2j-2}^*) & \text{for} \quad j > 1 \end{cases} \tag{3.10}$$

and

$$\psi_{2j+1}^* = \begin{cases} \frac{1}{2}T_1^* & \text{for} \quad j = 0 \\ \frac{1}{4}(T_{2j+1}^* - T_{2j-1}^*) & \text{for} \quad j > 0 \end{cases} . \tag{3.11}$$

3.3.2 Chebyshev polynomials of second kind

Chebyshev polynomials of second kind are defined by the identity

$$U_j(\cos\theta) = \frac{\sin(j+1)\theta}{\sin\theta}, \quad j = 0, 1, \ldots$$

(this is [AS84, Equation 22.3.16]). These polynomials are the orthogonal polynomials with respect to the measure

$$d\sigma(x) = \frac{2}{\pi}\sqrt{1-x^2}\,dx \quad \text{for} \quad x \in [-1, 1]$$

(again we use the normalization $\sigma([-1,1]) = 1$, i.e. $\beta_0 = 1$). In this case, we have $\alpha_j = 0$ by symmetry, and

$$\beta_j = 1/4 \quad \text{for} \quad j > 0,$$

the orthonormal polynomials will be

$$\hat{\pi}_j = U_j.$$

The polynomials ψ_j can be calculated easily (using the well-known identities for Chebyshev polynomials of second kind, see [AS84, 22.7.25]):

$$\psi_{2j} = \begin{cases} 1 & \text{for} \quad j = 0 \\ \frac{1}{4} U_{2j} & \text{for} \quad j > 0 \end{cases} \tag{3.12}$$

and

$$\psi_{2j+1} = \frac{1}{2} U_{2j+1}. \tag{3.13}$$

As in the preceeding subsection, we will consider shifted polynomials. The measure is defined by

$$d\sigma^*(x) = \frac{8}{\pi}\sqrt{x(1-x)}\,dx \quad \text{for} \quad x \in [0,1].$$

The explicit formulas for our polynomials ψ_j^* can be obtained analogously as above:

$$\psi_{2j}^* = \begin{cases} 1 & \text{for} \quad j = 0 \\ \frac{1}{16} U_{2j}^* & \text{for} \quad j > 0 \end{cases} \tag{3.14}$$

and

$$\psi_{2j+1}^* = \frac{1}{4} U_{2j+1}^*, \tag{3.15}$$

where U_n^* denotes the shifted Chebyshev polynomials of second kind (see [AS84, 22.5.15]):

$$U_n^*(x) = U_n(2x-1).$$

3.4 Ordinary moments

For our special choice $p_j(x) = x^j$ and $c_j(q) = q^{(j)}(0)/j!$ the following lemma holds:

Lemma 3.2 *Let $q \in \mathcal{P}_N$ a be polynomial with real coefficients. If the coefficients of q have alternating sign, then $a_\infty(q) = |q(-1)|$, if q contains only odd (or only even) powers of x with coefficients of alternating sign, then $a_\infty(q) = |q(i)|$.*

$a_2(q)$ can be expressed as

$$a_2(q) = \frac{1}{2\pi}\int_0^{2\pi} |q(e^{i\phi})|^2\, d\phi.$$

Proof: Let the coefficients of q have alternating sign. Then we have $\varepsilon(-1)^j c_j(q) \geq 0$ with some $\varepsilon = \pm 1$, and from this follows

$$a_\infty(q) = \sum_{j=0}^{N} |c_j(q)| = \sum_{j=0}^{N} \varepsilon(-1)^j c_j(q) = \varepsilon \sum_{j=0}^{N} c_j(q)\, p_j(-1) = \varepsilon q(-1).$$

But, since $a_\infty(q)$ is non-negative and $|\varepsilon| = 1$, we must have $\varepsilon q(-1) = |q(-1)|$.

If only the odd-indexed coefficients of q are different from zero and have alternating sign, then we have $c_{2j}(q) = 0$ for $2j \leq N$, and again $\varepsilon(-1)^j c_{2j+1}(q) \geq 0$ for $2j+1 \leq N$ with some $\varepsilon = \pm 1$. Consequently, from $(-1)^j = (-i)i^{2j+1} = (-i)p_{2j+1}(i)$ we obtain

$$\begin{aligned} a_\infty(q) &= \sum_{j=0}^{[\frac{N-1}{2}]} |c_{2j+1}(q)| = \sum_{j=0}^{[\frac{N-1}{2}]} \varepsilon(-1)^j c_{2j+1}(q) = (-i)\varepsilon \sum_{j=0}^{N} c_j(q)\, p_j(i) \\ &= (-i)\varepsilon q(i). \end{aligned}$$

As above, we have $|(-i)\varepsilon| = 1$, and the equation $(-i)\varepsilon q(i) = |q(i)|$ follows. The case when only the even-indexed coefficients of q are different from zero and have alternating sign can be dealt with analogously.

The formula for $a_2(q)$ is an immediate consequence of the orthogonality of the functions $e^{ij\phi}$. □

We are now ready to give explicit results for the simple Examples 3.3.1 and 3.3.2 above.

Theorem 3.3 *Let the measure σ be defined as in Example 3.3.1 (Chebyshev polynomials of first kind). Then the supremum-norm of K_n' will be*

$$\|K_n'\|_\infty = 4\, U_{n-2}(3) \quad \text{for} \quad n \geq 2. \tag{3.16}$$

In the case of Example 3.3.1 (shifted Chebyshev polynomials of first kind) this norm evaluates as*

$$\|K_n'\|_\infty = 4\, U_{2n-3}(3) \quad \text{for} \quad n \geq 2. \tag{3.17}$$

For Chebyshev polynomials of second kind (Example 3.3.2) we have the result

$$\|K_n'\|_\infty = U_{n-1}(3) \quad \textit{for} \quad n \geq 2. \tag{3.18}$$

Finally, in the case of shifted Chebyshev polynomials of second kind (Example 3.3.2) we obtain the formula*

$$\|K_n'\|_\infty = \frac{1}{4} U_{2n-1}(3) \quad \textit{for} \quad n \geq 2. \tag{3.19}$$

Proof: First, we prove Equation (3.16). According to equation (3.6), we have to investigate the functionals $a_\infty(\psi_j)$. The fact, that the Chebyshev polynomials $T_n(x)$ and $U_n(x)$ contain only odd or only even powers of x with coefficients of alternating sign, is well-known (see the explicit expressions [AS84, 22.3.6] and [AS84, 22.3.7] or [Mei64, 4.15] and [Mei64, 4.16]). For the polynomials $T_{j+1}(x) - T_{j-1}(x)$ this follows from the identity

$$T_{j+1}(x) - T_{j-1}(x) = 2(x^2 - 1)U_{j-1}(x) \tag{3.20}$$

(this is Equation [AS84, 22.7.25] with $n = j$ and $m = 1$). Thus, the polynomials ψ_j satisfy the assumptions of the preceeding lemma, and we have $a_\infty(\psi_j) = |\psi_j(i)|$. From equations (3.6), (3.8), (3.9) and (3.20) we can see, that $\|K_n'\|_\infty$ will be the maximum of

$$\max_{0 \leq j \leq 2} |\psi_j(i)| = \frac{3}{2}, \quad \max_{2 \leq j \leq n-1} |U_{2j-2}(i)| \quad \text{and} \quad 2 \max_{1 \leq j \leq n-1} |U_{2j-1}(i)|.$$

We will need the following formulas for Chebyshev polynomials of first and second kinds

$$T_n(x) = \frac{1}{2}\left[(x + \sqrt{x^2 - 1})^n + (x - \sqrt{x^2 - 1})^n\right] \tag{3.21}$$

(this is [Mei64, 4.13]) and

$$U_n(x) = \frac{1}{2\sqrt{x^2 - 1}}\left[(x + \sqrt{x^2 - 1})^{n+1} - (x - \sqrt{x^2 - 1})^{n+1}\right] \tag{3.22}$$

(see [Mei64, 4.14]). We see easily, that always holds

$$|U_j(i)| \geq |U_{j-1}(i)|.$$

Consequently, the maximal value will be

$$2 \max_{1 \leq j \leq n-1} |U_{2j-1}(i)| = 2\, |U_{2n-3}(i)|,$$

if only this is not less than $\frac{3}{2}$. To convert the result to real expressions, we need another identity:

$$U_{2j-1}(x) = 2xU_{j-1}(2x^2 - 1). \tag{3.23}$$

This equation follows at once from the trigonometric form of the definition

$$U_n(\cos\theta) = \frac{\sin(n+1)\theta}{\sin\theta}$$

(see Subsection 3.3.2), since we have

$$U_{2j-1}(\cos\theta) = \frac{\sin 2j\theta}{\sin\theta} = \frac{\sin 2\theta}{\sin\theta}\cdot\frac{\sin(j\cdot 2\theta)}{\sin 2\theta} = 2\cos\theta\, U_{j-1}(\cos 2\theta).$$

Setting $x = i$, we obtain $U_{2n-3}(i) = 2iU_{n-2}(-3)$ and $\|K_n'\|_\infty = 2\,|2iU_{n-2}(-3)| = 4U_{n-2}(3)$ (this value is greater than $\frac{3}{2}$ for all $n \geq 2$).

The Equation (3.17) can be shown analogously. First, we have to check the assumptions of the lemma. Indeed, the polynomial $U_j^*(x)$ has coefficients of alternating sign: From (3.23) we see

$$U_j^*(x) = U_j(2x-1) = \frac{U_{2j+1}(\sqrt{x})}{2\sqrt{x}}.$$

The same is true for the polynomials $T_{j+1}^*(x) - T_{j-1}^*(x)$, since (3.20) gives (with $2x-1$ instead of x)

$$T_{j+1}^*(x) - T_{j-1}^*(x) = 8(x^2 - x)U_{j-1}^*(x). \tag{3.24}$$

Hence, all polynomials ψ_j^* have coefficients of alternating sign, and $a_\infty(\psi_j^*) = |\psi_j^*(-1)|$. Again, from equations (3.6), (3.10), (3.11) and (3.24) we obtain

$$\|K_n'\|_\infty = |\psi_{2n-1}^*(-1)| = 4U_{2n-3}(3),$$

if only this is not less than $\max\limits_{j=0,1,2} |\psi_j^*(-1)| = \frac{3}{2}$, and this is true for all $n \geq 2$.

The proof of equations (3.18) and (3.19) goes along the same lines.□

The comparison of (3.17) with the result of Gautschi [Gau78, 5.9] is very interesting. Gautschi investigates the map G_n instead of K_n and obtains the estimate

$$\operatorname{cond}_\infty(G_n, m) > \frac{(17+6\sqrt{8})^n}{64n^2}$$

(where $\operatorname{cond}_\infty$ is defined via (3.2) with the norm $\|\cdot\| = \|\cdot\|_\infty$). From our Equation (3.17) together with (3.22) we have the asymptotics

$$\|K_n'\|_\infty \sim 4\,\frac{1}{2\sqrt{8}}(3+\sqrt{8})^{2n-2} = \frac{2(17-6\sqrt{8})}{\sqrt{8}}\,(17+6\sqrt{8})^n.$$

3.5 Bernstein polynomials

The well-known Bernstein polynomials $\mathcal{B}_N = (B_{N,j})_{j=0,\ldots,N}$, defined by

$$B_{N,j}(x) = \binom{N}{j} x^j (1-x)^{N-j},$$

form a base of $\mathcal{P}_N$. Since $\mathcal{B}_N$ is not a subset of $\mathcal{B}_M$ for $M \geq N$, we have to indicate the dependence of coordinate functionals and norms on the degree N explicitly:

$$q(x) = \sum_{j=0}^{N} c_{N,j}(q) B_{N,j}(x) \quad \text{for} \quad q \in \mathcal{P}_N$$

and

$$a_{N,\infty}(q) = \sum_{j=0}^{N} |c_{N,j}(q)|$$

(we consider here the supremum-norm only).

Lemma 3.4 *Let $q \in \mathcal{P}_N \subset \mathcal{P}_M$ (i. e. $N \leq M$). Then we have the inequality*

$$a_{M,\infty}(q) \leq \frac{M+1}{N+1}\, a_{N,\infty}(q). \tag{3.25}$$

Proof: Obviously, we have to prove our proposition only for $M = N+1$, the general case follows via induction. From the identity

$$B_{N,j}(x) = \frac{N+1-j}{N+1}\, B_{N+1,j}(x) + \frac{j+1}{N+1}\, B_{N+1,j+1}(x)$$

for $j = 0, \ldots, N$ we obtain

$$c_{N+1,j}(q) = \frac{N+1-j}{N+1}\, c_{N,j}(q) + \frac{j}{N+1}\, c_{N,j-1}(q)$$

for $q \in \mathcal{P}_N$ and $j = 0, \ldots, N+1$ (the undefined functionals $c_{N,N+1}$ and $c_{N,-1}$ have coefficient zero). This gives

$$\begin{aligned} a_{N+1,\infty}(q) &= \sum_{j=0}^{N+1} |c_{N+1,j}(q)| \leq \sum_{j=0}^{N} \frac{N+1-j}{N+1}\, |c_{N,j}(q)| + \sum_{j=1}^{N+1} \frac{j}{N+1}\, |c_{N,j-1}(q)| \\ &= \sum_{j=0}^{N} \left(\frac{N+1-j}{N+1} + \frac{j+1}{N+1} \right) |c_{N,j}(q)| \\ &= \frac{N+2}{N+1} \sum_{j=0}^{N} |c_{N,j}(q)| = \frac{N+2}{N+1}\, a_{N,\infty}(q). \square \end{aligned}$$

The following result enables us to evaluate $a_{N,\infty}(q)$ explicitly in some special cases:

Lemma 3.5 *Let $q \in \mathcal{P}_N$, and let the coefficients of q in the base $\mathcal{B}_N$ have alternating sign. Then the equality*

$$a_{N,\infty}(q) = (N+1) \left| \int_0^1 (2u-1)^N \, q\left(\frac{u}{2u-1}\right) du \right| \tag{3.26}$$

holds. The coefficients have alternating sign, if all zeros of the polynomial are real and $\in [0,1]$.

Proof: Suppose first, that all coefficients $c_{N,j}(q)$ are positive. From the well-known fact

$$\int_0^1 B_{N,j}(u)\, du = \frac{1}{N+1}$$

we obtain

$$a_{N,\infty}(q) = \sum_{j=0}^{N} c_{N,j}(q) = (N+1) \int_0^1 q(u)\, du.$$

The first proposition now follows easily from the identity

$$(2u-1)^N \, B_{N,j}\left(\frac{u}{2u-1}\right) = (-1)^{N-j} B_{N,j}(u).$$

If all zeros $x_1, ..., x_N$ of $q \in \mathcal{P}_N$ are real and $\in [0,1]$, then we have

$$q(x) = c \prod_{i=1}^{N} (x - x_i) = c \prod_{i=1}^{N} [(1-x_i)x - x_i(1-x)],$$

and this polynomial obviously has coefficients of alternating sign in the base $\mathcal{B}_N$.$\Box$

We choose $N = 2n-1$, and Equation (3.6) takes the form

$$\|K_n'\|_\infty = \max_{0 \le j \le N} a_{N,\infty}(\psi_j).$$

Of course, our Equation (3.26) enables us to evaluate $a_{j,\infty}(\psi_j)$ only. But in some cases we will be able to show

$$\max_{0 \le j \le N} a_{N,\infty}(\psi_j) = a_{N,\infty}(\psi_N)$$

using the inequality

$$a_{N,\infty}(\psi_j) \le \frac{N+1}{j+1} a_{j,\infty}(\psi_j).$$

Since the natural interval for Bernstein polynomials is $[0,1]$, we illustrate our results with shifted Chebyshev polynomials only.

Theorem 3.6 *Let the measure σ be defined as in Example 3.3.1* (shifted Chebyshev polynomials of first kind). Then the supremum-norm of K_n' can be estimated as*

$$\|K_n'\|_\infty = \frac{n}{2}\left[2^{2n-2}\,\frac{\Gamma(\frac{1}{2})\Gamma(2n-\frac{3}{2})}{\Gamma(2n-1)} - 2^{2n}\,\frac{\Gamma(\frac{3}{2})\Gamma(2n-\frac{1}{2})}{\Gamma(2n+1)} + O(1)\right] \qquad (3.27)$$

for $n \to \infty$. In the case of Example 3.3.2 (shifted Chebyshev polynomials of second kind) we have the estimate*

$$\|K_n'\|_\infty = \frac{n}{2}\left[2^{2n-1}\,\frac{\Gamma(\frac{1}{2})\Gamma(2n+\frac{1}{2})}{\Gamma(2n+1)} + O(1)\right] \quad \textit{for} \quad n \to \infty. \qquad (3.28)$$

Proof: First, we consider shifted Chebyshev polynomials of first kind. We define the polynomials

$$q_j(x) = T_j^*(x) - T_{j-2}^*(x) \quad \text{for} \quad j \geq 2.$$

From (3.24) we have $q_j(x) = 8x(x-1)U_{j-2}^*(x)$, this polynomial has all its zeroes in $[0,1]$ and thus fulfills the assumptions of Lemma 3.5. Using Equation (3.22), we get

$$q_j(x) = 2\sqrt{x^2-x}\left[\left(2x-1+2\sqrt{x^2-x}\right)^{j-1} - \left(2x-1-2\sqrt{x^2-x}\right)^{j-1}\right]$$

and from this

$$(2u-1)^j q_j\left(\frac{u}{2u-1}\right) =$$
$$2\sqrt{u-u^2}\left[\left(1+2\sqrt{u-u^2}\right)^{j-1} - \left(1-2\sqrt{u-u^2}\right)^{j-1}\right].$$

Our Lemma 3.5 yields

$$\frac{1}{j+1}\,a_{j,\infty}(q_j) =$$
$$\int_0^1 2\sqrt{u-u^2}\left[\left(1+2\sqrt{u-u^2}\right)^{j-1} - \left(1-2\sqrt{u-u^2}\right)^{j-1}\right]\,du$$

Now we can see that the integrand is monotonically increasing in j, and consequently we obtain

$$\frac{1}{j+1}\,a_{j,\infty}(q_j) \leq \frac{1}{N+1}\,a_{N,\infty}(q_N) \quad \text{for} \quad 2 \leq j \leq N.$$

Together with (3.25) this implies

$$a_{N,\infty}(q_j) \leq \frac{N+1}{j+1}\,a_{j,\infty}(q_j) \leq a_{N,\infty}(q_N) \quad \text{for} \quad 2 \leq j \leq N.$$

Since according to equations (3.10) and (3.11) we have $\psi_j = \frac{1}{16} q_j$ for even $j > 2$ and $\psi_j = \frac{1}{4} q_j$ for odd $j > 2$ and $N = 2n - 1$ is odd, the inequality

$$a_{N,\infty}(\psi_j) \le \frac{1}{4} a_{N,\infty}(q_j) \le \frac{1}{4} a_{N,\infty}(q_N) = a_{N,\infty}(\psi_N) \quad \text{for} \quad 2 < j \le N$$

follows immediately. This yields the equation

$$\max_{2<j\le N} a_{N,\infty}(\psi_j) = a_{N,\infty}(\psi_N),$$

and since we will see that $a_{N,\infty}(\psi_N)$ tends to infinity as $N \to \infty$, we have

$$\|K_n'\|_\infty = \max_{0\le j\le N} a_{N,\infty}(\psi_j) = a_{N,\infty}(\psi_N) = \frac{1}{4} a_{N,\infty}(q_N) \tag{3.29}$$

for N large enough. This quantity can be estimated easily: Our lemma gives the expression

$$\frac{1}{N+1} a_{N,\infty}(q_N) = \\ \int_0^1 2\sqrt{u-u^2}\left[\left(1+2\sqrt{u-u^2}\right)^{N-1} - \left(1-2\sqrt{u-u^2}\right)^{N-1}\right] du.$$

But $\left|1 - 2\sqrt{u-u^2}\right| \le 1$, and consequently the right-hand side equals

$$\int_0^1 2\sqrt{u-u^2}\left[\left(1+2\sqrt{u-u^2}\right)^{N-1} + \left(1-2\sqrt{u-u^2}\right)^{N-1}\right] du + O(1).$$

The last integral can be evaluated: First, in the integral

$$\int_0^1 2\sqrt{u-u^2}\left(1+2\sqrt{u-u^2}\right)^{N-1} du = \\ 2\int_0^{\frac{1}{2}} 2\sqrt{u-u^2}\left(1+2\sqrt{u-u^2}\right)^{N-1} du$$

we substitute $v = \sqrt{u} + \sqrt{1-u}$ to obtain

$$2\int_1^{\sqrt{2}} \frac{(v^2-1)^2}{\sqrt{2-v^2}} v^{2N-2}\, dv.$$

Further, in the integral

$$\int_0^1 2\sqrt{u-u^2}\left(1-2\sqrt{u-u^2}\right)^{N-1} du = \\ 2\int_0^{\frac{1}{2}} 2\sqrt{u-u^2}\left(1-2\sqrt{u-u^2}\right)^{N-1} du$$

we substitute $v = \sqrt{1-u} - \sqrt{u}$ to obtain

$$2\int_0^1 \frac{(v^2-1)^2}{\sqrt{2-v^2}} v^{2N-2}\, dv.$$

Putting these integrals together, we obtain the equation

$$\frac{1}{N+1}\,a_{N,\infty}(q_N) = 2\int_0^{\sqrt{2}} \frac{(v^2-1)^2}{\sqrt{2-v^2}}\,v^{2N-2}\,dv + O(1).$$

Using the simple identity

$$\frac{(v^2-1)^2}{\sqrt{2-v^2}}\,v^{2N-2} = \frac{1}{\sqrt{2-v^2}}\,v^{2N-2} - \sqrt{2-v^2}\,v^{2N},$$

the substitution $v = \sqrt{2t}$ and the well-known formulas for the Beta integrals (see [AS84, 6.2.1 and 6.2.2])

$$B(z,w) = \int_0^1 t^{z-1}(1-t)^{w-1}\,dt = \frac{\Gamma(z)\Gamma(w)}{\Gamma(z+w)},$$

we finally arrive at

$$\frac{1}{N+1}\,a_{N,\infty}(q_N) = 2^{N-1}\,\frac{\Gamma(\frac{1}{2})\Gamma(N-\frac{1}{2})}{\Gamma(N)} - 2^{N+1}\,\frac{\Gamma(\frac{3}{2})\Gamma(N+\frac{1}{2})}{\Gamma(N+2)} + O(1).$$

Inserting this in Equation (3.29) and observing $N = 2n-1$, we obtain our proposition (3.27).

The proof of our second proposition (3.28) is almost identical to the above and will be omitted here.□

3.6 Lagrange interpolation polynomials

For a given set $\{\tau_i\}_{i=1,\dots,n}$ of nodes, the polynomials l_i defined by Equation (1.15) form a base of $\mathcal{P}_N$ with $N = n-1$ (this was explained in Section 1.5). Here, we need a base for $\mathcal{P}_N$ with $N = 2n-1$, i.e. we must have nodes τ_i and polynomials l_i for $i = 1,\dots,2n$. The coordinate functionals have the simple form $c_i(q) = q(\tau_i)$ for all $q \in \mathcal{P}_N$, and clearly the following result holds:

Lemma 3.7 *Let $q \in \mathcal{P}_N$, with $N = 2n-1$, and let $[a,b]$ be any interval containing all nodes $(\tau_1)_{j=1,\dots,2n}$. Then we have the estimate*

$$a_\infty(q) \le (N+1)\,\|q\|_\infty, \tag{3.30}$$

where $\|q\|_\infty$ is the usual supremum-norm:

$$\|q\|_\infty = \sup_{x\in[a,b]} |q(x)|.$$

Now we are able to give simple estimates for our illustrative examples.

Theorem 3.8 *Let the measure σ be defined as in Example 3.3.1 (Chebyshev polynomials of first kind), and let all nodes $\tau_1, \ldots, \tau_{2n}$ be in the interval of orthogonality $[-1,1]$. Then the supremum-norm of K_n' can be estimated as*

$$\|K_n'\|_\infty \leq 2n. \tag{3.31}$$

In the case of Example 3.3.2 (Chebyshev polynomials of second kind) we have the estimate

$$\|K_n'\|_\infty \leq 2n^2. \tag{3.32}$$

Proof: In the case of Chebyshev polynomials of first kind, we can see from our equations (3.8) and (3.9) that $\|\psi_j\|_\infty \leq 1$ for all $j \geq 0$. Consequently, we have the estimate

$$\|K_n'\|_\infty = \max_{0\leq j\leq N} a_\infty(\psi_j) \leq (N+1) \max_{0\leq j\leq N} \|\psi_j\|_\infty \leq N+1 = 2n.$$

In the case of Chebyshev polynomials of second kind, our proposition follows easily from equations (3.12) and (3.13) and from

$$\|\psi_j\|_\infty \leq \frac{1}{2}\|U_j\|_\infty = \frac{1}{2}(j+1) \quad \text{for} \quad j > 0. \Box$$

Surprisingly, we have simple estimates for $\|K_n'\|_\infty$ of only polynomial growth in n! Moreover, the result does not depend on the distribution of the nodes inside the interval of orthogonality. However, this is somewhat misleading. The norm $\|K_n'\|_\infty$ is a measure for sensitivity to small *absolute* perturbations of the modified moments. But if the distribution of the nodes is bad, the basic polynomials may have large peaks, and consequently the modified moments may vary greatly in magnitude (compare the discussion in [Gau86, Sections 2. and 6.]). If the support of our measure σ is an interval, we could, of course, use well-known good nodes of interpolation (e.g. Chebyshev nodes), where the basic polynomials remain uniformly bounded. Unfortunately, this is only theoretically useful, too. The numerical computation of the "Lagrange moments" $\int l_i(x)\, d\sigma(x)$ is hardly ever possible even in the case of very simple measures σ.

3.7 Orthonormal systems

Let now $(p_j)_{n=0,\ldots,N}$ be a system of orthonormal polynomials with respect to some measure s. Then the coordinate functional $c_j(\cdot)$ will be $c_j(q) = \int q(x)p_j(x)\, ds(x)$. We shall concentrate in this case on Frobenius norm, clearly we have

$$a_2(q) = \sum_{j=0}^{N} c_j^2(q) = \int q^2(x)\, ds(x) \quad \text{for} \quad q \in \mathcal{P}_N.$$

This immediately yields

Theorem 3.9 *The Frobenius norm of K_n' is*

$$\|K_n'\|_F = \left\{\int w_n(x)\, ds(x)\right\}^{1/2},$$

where

$$w_n(x) = \sum_{j=0}^{2n-1} \psi_j^2(x) =$$
$$\sum_{j=0}^{n-1} \left\{\beta_j^2[\hat{\pi}_j^2(x) - \hat{\pi}_{j-1}^2(x)]^2 + [\beta_{j+1}^{1/2}\hat{\pi}_{j+1}(x) - \beta_j^{1/2}\hat{\pi}_{j-1}(x)]^2\hat{\pi}_j^2(x)\right\}.$$

In the last equation, we used the representations

$$\psi_{2j}(x) = \beta_j(\hat{\pi}_j^2(x) - \hat{\pi}_{j-1}^2(x))$$

and

$$\psi_{2j+1}(x) = \beta_{j+1}^{1/2}\hat{\pi}_j(x)\hat{\pi}_{j+1}(x) - \beta_j^{1/2}\hat{\pi}_{j-1}(x)\hat{\pi}_j(x)$$

in terms of orthonormal polynomials. In most applications, the measure s will be some classical measure, the support of which is an interval. We can see that this interval should be contained in the support of σ, since otherwise we have to expect bad numerical condition (the orthonormal polynomials $\hat{\pi}_n$ grow exponentially with n outside $S(\sigma)$). The support of σ should not be larger than $S(s)$ for the same reason: The polynomials p_n grow exponentially outside $S(s)$, and we would have large *absolute* errors in our modified moments. However, if $S(\sigma) = S(s)$ is an interval, we know that Markov's inequality (with $\alpha = 2$) holds, and in many cases we can use the inequalities in Section 1.3.

Corollary 3.10 *Let $S(\sigma) = S(s)$ be an interval, and let the measure satisfy condition (1.10). Then there are constants $C > 0$ and γ satisfying*

$$\|K_n'\|_F \leq Cn^{\gamma}.$$

Proof: This is an immediate consequence of the above theorem and Proposition 1.4. □

3.8 An outlook to Sobolev-type polynomials

Let $(\cdot,\cdot)$ be a more general inner product in a function space containing all polynomials given by

$$(p, q) = l_0(pq) + l_1(p'q')$$

with some positive linear functionals l_0 and l_1. Again, there uniquely (up to the sign) exists a family of orthonormal polynomials $\hat{\pi}_j$ with

$$(\hat{\pi}_l, \hat{\pi}_j) = 0 \quad \text{for} \quad l < j \quad \text{and} \quad (\hat{\pi}_j, \hat{\pi}_j) = 1, \quad j = 0, 1, \ldots \tag{3.33}$$

They satisfy a recurrence relation

$$x\hat{\pi}_j(x) = \sum_{l=0}^{j+1} \beta_{jl}\hat{\pi}_l(x) \quad \text{for} \quad j = 0, 1, \ldots. \tag{3.34}$$

If we set

$$\beta_{jl} = 0 \quad \text{for} \quad l > j + 1,$$

then we can write

$$x\hat{\pi} = \mathbf{B}\hat{\pi} \tag{3.35}$$

with the (formally infinite) lower Hessenberg matrix $\mathbf{B} = [\beta_{lj}]$ and the column vector $\hat{\pi} = [\hat{\pi}_0, \hat{\pi}_1, \ldots]^T$. We stress here that $\mathbf{B}$ is not necessarily tridiagonal, since the operator of multiplication by x is no longer symmetric (compare Section 0.1).

In order to calculate the inner product, we must be given the values of the functionals l_0 and l_1 for some polynomials forming a base. The generation of the orthogonal polynomials (or, equivalently, of the coefficients β_{jl}) from ordinary moments

$$\mu_k^0 = l_0(x^k) \quad \text{and} \quad \mu_k^1 = l_1(x^k)$$

is severely ill-conditioned even in the special case $l_1 \equiv 0$, as the above results show clearly (compare Section 3.4). More promising are modified moments

$$m_k^0 = l_0(p_k) \quad \text{and} \quad m_k^1 = l_1(p_k) \tag{3.36}$$

with some properly chosen set of polynomials p_k (e.g. Chebyshev polynomials). Of course, we should be able to estimate the sensitivity of the coefficients β_{jl} to small perturbations of the modified moments. As we already mentioned, in the traditional case $l_1 \equiv 0$, this sensitivity was estimated by Gautschi, where the map from modified moments to the vector of nodes and weights of a Gauss quadrature rule is considered, instead. In the general case, the sensitivity was estimated in [Zha94], where some auxiliary map is considered, too.

Again, a more direct approach is possible. For the special case $l_1 \equiv 0$ results of the author were published in [Fis96], formulated in terms of *orthonormal* polynomials. In this section, we use a similar device. We assume that $(p_j)_{n=0,\ldots,N}$ is any polynomial base in the space $\mathcal{P}_N$ of all polynomials of degree $\leq N$ (as above, **without** assumptions like $\deg(p_j) = j$).

For the following, we need partial derivatives of inner products.

Lemma 3.11 *Let $p, q \in \mathcal{P}_N$ be some polynomials depending on*

$$m_0^0, \ldots, m_N^0, m_0^1, \ldots, m_N^1$$

with continuous partial derivatives in some neighbourhood of the point

$$[m_0^0, \ldots, m_N^0, m_0^1, \ldots, m_N^1]^T,$$

$pq \in \mathcal{P}_N$ and (3.3) and (3.36) hold. Then,

$$\frac{\partial}{\partial m_k^0}(p,q) = c_k(pq) + \left(\frac{\partial p}{\partial m_k^0}, q\right) + \left(p, \frac{\partial q}{\partial m_k^0}\right) \tag{3.37}$$

and

$$\frac{\partial}{\partial m_k^1}(p,q) = c_k(p'q') + \left(\frac{\partial p}{\partial m_k^1}, q\right) + \left(p, \frac{\partial q}{\partial m_k^1}\right). \tag{3.38}$$

Proof: From (3.3) and (3.36) we have immediately

$$(p,q) = \sum_{j=0}^{N} c_j(pq)\, m_j^0 + \sum_{j=0}^{N} c_j(p'q')\, m_j^1,$$

and differentiating

$$\begin{aligned}
\frac{\partial}{\partial m_k^0}(p,q) &= c_k(pq) + \sum_{j=0}^{N} \frac{\partial c_j(pq)}{\partial m_k^0}\, m_j^0 + \sum_{j=0}^{N} \frac{\partial c_j(p'q')}{\partial m_k^0}\, m_j^1 \\
&= c_k(pq) + \sum_{j=0}^{N} c_j \left(\frac{\partial p}{\partial m_k^0} q + p \frac{\partial q}{\partial m_k^0}\right) m_j^0 \\
&\quad + \sum_{j=0}^{N} c_j \left(\frac{\partial p'}{\partial m_k^0} q' + p' \frac{\partial q'}{\partial m_k^0}\right) m_j^1 \\
&= c_k(pq) + \left(\frac{\partial p}{\partial m_k^0}, q\right) + \left(p, \frac{\partial q}{\partial m_k^0}\right).
\end{aligned}$$

The exchanche of derivative, partial derivatives and the functionals c_j is justified. Equation (3.38) can be shown analogously. □

The $\hat{\pi}_j$ are differentiable functions of $m_0^0, \ldots, m_N^0, m_0^1, \ldots, m_N^1$ in some sufficiently small neighbourhood of $[m_0^0, \ldots, m_N^0, m_0^1, \ldots, m_N^1]^T$. Differentiating the relations (3.33) with the help of (3.37) we obtain first

$$\begin{aligned}
0 &= \frac{\partial}{\partial m_k^0}(\hat{\pi}_l, \hat{\pi}_j) = c_k(\hat{\pi}_l \hat{\pi}_j) + \left(\frac{\partial \hat{\pi}_l}{\partial m_k^0}, \hat{\pi}_j\right) + \left(\hat{\pi}_l, \frac{\partial \hat{\pi}_j}{\partial m_k^0}\right) \\
&= c_k(\hat{\pi}_l \hat{\pi}_j) + \left(\hat{\pi}_l, \frac{\partial \hat{\pi}_j}{\partial m_k^0}\right)
\end{aligned}$$

for $l < j$, i.e.

$$\left(\hat{\pi}_l, \frac{\partial \hat{\pi}_j}{\partial m_k^0}\right) = -c_k(\hat{\pi}_l \hat{\pi}_j).$$

The second relation gives

$$0 = \frac{\partial}{\partial m_k^0}\,(\hat{\pi}_j, \hat{\pi}_j) = c_k(\hat{\pi}_j^2) + 2\left(\hat{\pi}_j, \frac{\partial \hat{\pi}_j}{\partial m_k^0}\right),$$

and this means

$$\left(\hat{\pi}_j, \frac{\partial \hat{\pi}_j}{\partial m_k^0}\right) = -\frac{1}{2}\,c_k(\hat{\pi}_j^2).$$

In the same way we can obtain the equations

$$\left(\hat{\pi}_l, \frac{\partial \hat{\pi}_j}{\partial m_k^1}\right) = -c_k(\hat{\pi}_l' \hat{\pi}_j')$$

and

$$\left(\hat{\pi}_j, \frac{\partial \hat{\pi}_j}{\partial m_k^1}\right) = -\frac{1}{2}\,c_k(\hat{\pi}_j'^2).$$

Thus we have

Corollary 3.12 *Under our assumptions (3.3), (3.36) and (3.33) the following formulas hold:*

$$\frac{\partial \hat{\pi}_j}{\partial m_k^0} = -\sum_{l<j} c_k(\hat{\pi}_l \hat{\pi}_j)\hat{\pi}_l - \frac{1}{2}\,c_k(\hat{\pi}_j^2)\hat{\pi}_j$$

and

$$\frac{\partial \hat{\pi}_j}{\partial m_k^1} = -\sum_{l<j} c_k(\hat{\pi}_l' \hat{\pi}_j')\hat{\pi}_l - \frac{1}{2}\,c_k(\hat{\pi}_j'^2)\hat{\pi}_j.$$

Again, we can write down these equations in a more compact form using matrix notation. Let

$$\gamma_{k;jl} = \begin{cases} 0 & \text{for} \quad l > j \\ \frac{1}{2}\,c_k(\hat{\pi}_j^2) & \text{for} \quad l = j \\ c_k(\hat{\pi}_l \hat{\pi}_j) & \text{for} \quad l < j \end{cases}$$

and

$$\delta_{k;jl} = \begin{cases} 0 & \text{for} \quad l > j \\ \frac{1}{2}\,c_k(\hat{\pi}_j'^2) & \text{for} \quad l = j \\ c_k(\hat{\pi}_l' \hat{\pi}_j') & \text{for} \quad l < j \end{cases}$$

and the lower triangular matrices $\mathbf{C}_k$ and $\mathbf{D}_k$ be defined as $\mathbf{C}_k = [\gamma_{k;jl}]_{j,l=0,1,\ldots}$ and $\mathbf{D}_k = [\delta_{k;jl}]_{j,l=0,1,\ldots}$. Then the proposition of our corollary is just

$$\frac{\partial \hat{\pi}}{\partial m_k^0} = -\mathbf{C}_k \hat{\pi} \quad \text{and} \quad \frac{\partial \hat{\pi}}{\partial m_k^1} = -\mathbf{D}_k \hat{\pi}. \tag{3.39}$$

Now we are ready to prove our main result.

Theorem 3.13 *The partial derivatives of* $\mathbf{B}$ *with respect to the modified moments can be calculated from the equations*

$$\frac{\partial \mathbf{B}}{\partial m_k^0} = \mathbf{B}\mathbf{C}_k - \mathbf{C}_k \mathbf{B} \tag{3.40}$$

and

$$\frac{\partial \mathbf{B}}{\partial m_k^1} = \mathbf{B}\mathbf{D}_k - \mathbf{D}_k \mathbf{B}. \tag{3.41}$$

Proof: From (3.33) and (3.34) we have

$$\beta_{ij} = (x\hat{\pi}_i, \hat{\pi}_j),$$

and with (3.37) we obtain

$$\frac{\partial \beta_{ij}}{\partial m_k^0} = c_k(x\hat{\pi}_i \hat{\pi}_j) + \left(x\frac{\partial \hat{\pi}_i}{\partial m_k^0}, \hat{\pi}_j \right) + \left(x\hat{\pi}_i, \frac{\partial \hat{\pi}_j}{\partial m_k^0} \right).$$

Expanding $x\hat{\pi}_i$ using (3.34) and observing Corollary 3.12 we can write the right-hand side as

$$\begin{aligned} &\sum_l \left[\beta_{il} c_k(\hat{\pi}_l \hat{\pi}_j) - \gamma_{k;il}(x\hat{\pi}_l, \hat{\pi}_j) - (x\hat{\pi}_i, \hat{\pi}_l)\gamma_{k;jl}\right] \\ &= \sum_l \left[\beta_{il} c_k(\hat{\pi}_l \hat{\pi}_j) - \gamma_{k;il}\beta_{lj} - \beta_{il}\gamma_{k;jl}\right]. \end{aligned}$$

Now the full matrix $[c_k(\hat{\pi}_l \hat{\pi}_j)]_{l,j=0,1,\ldots}$ is just $\mathbf{C}_k + \mathbf{C}_k^T$, and we obtain in matrix notation

$$\frac{\partial \mathbf{B}}{\partial m_k^0} = \mathbf{B}(\mathbf{C}_k + \mathbf{C}_k^T) - \mathbf{C}_k \mathbf{B} - \mathbf{B}\mathbf{C}_k^T = \mathbf{B}\mathbf{C}_k - \mathbf{C}_k \mathbf{B}.$$

Thus Equation (3.40) is true, and the Equation (3.41) can be shown in the same way. □

We remark here that there is, however, a significant difference between the matrices $\mathbf{C}_k$ and $\mathbf{D}_k$: The matrix formed by the elements $c_k(x\hat{\pi}_i\hat{\pi}_j)$ is symmetric, i.e. we have the identity

$$\mathbf{B}(\mathbf{C}_k + \mathbf{C}_k^T) = (\mathbf{C}_k + \mathbf{C}_k^T)\mathbf{B}^T.$$

There is no analogous identity for $\mathbf{D}_k$.

3.9 Bibliographical references

In this chapter, we concentrated on the numerical condition of the map K_n. The numerical implementation of the map K_n, numerical results as well as an alternative approach to sensitivity (via the maps G_n and H_n) can be found in papers of W. Gautschi (especially in the survey papers [Gau90] and [Gau96] and in the description of algorithms [Gau94]). Of course, all of these results consider the classical case of absolutely continuous measures, where we have a positive weight function (the case of indefinite weights was considered thoroughly in a paper of Golub and Gutknecht [GG90]). However, we showed that the method of modified moments is justified sometimes for singular measures, if the support is an interval (see Corollary 3.10). In the case of a fractal support, it is difficult to find an appropriate polynomial base for modified moments: They can't be orthogonal, since only orthogonal polynomials with the same support can give reasonable numerical condition. Beckermann and Bourreau came to the same conclusion in their paper [BB97], using somewhat different techniques.

A possible choice of a base would be Newton interpolation polynomials in the Leja points of the support. Formally, they satisfy a recurrence of type (1.13) with $\beta_j \equiv 0$. This gives an even simpler algorithm of modified moments. The functionals a_∞ and a_2 can be analyzed with Reichel's method (see his paper [Rei90]), and we would have subexponential growth of $\|K_n'\|$. Indeed, we calculated the recurrence coefficients of Cantor measure with this method, first. We did not report any results in this direction here, since the methods of Chapter 6 are more natural. The method could be competetive for unbounded supports, if we use weighted Leja-points (compare Chapter V in the book of Saff and Totik [ST97]), but this is far beyond the scope of the present work.

The results of the first sections were published in our paper [Fis96] (in terms of orthonormal polynomials). The last section is a mere illustration of the power of our method and was inspired by the paper of Zhang [Zha94]. To apply these results, we would need estimates of Sobolev-type polynomials generalizing those of Section 1.3. This should be possible, but we did not elaborate on it.

Chapter 4

From modified moments to quadrature formulas

If we are given modified moments $m_k = \int p_k(x)\, d\sigma(x)$, we can try to calculate directly the nodes $(\tau_i)_{i=1,\ldots,n}$ and the weights $(\sigma_i)_{i=1,\ldots,n}$ of a Gaussian quadrature formula for the measure σ. Since the quadrature formula is exact for polynomials of degree $2n-1$, we obtain the nonlinear system of equations

$$\sum_{i=1}^{n} \sigma_i\, p_j(\tau_i) = m_j, \qquad j = 0, \ldots, 2n-1.$$

This can be written as

$$F(\sigma, \tau) = m$$

with $\sigma = [\sigma_1, \ldots, \sigma_n]^T$, $\tau = [\tau_1, \ldots, \tau_n]^T$, $m = [m_1, \ldots, m_{2n-1}]^T$ and the function

$$F(\sigma, \tau) = \sum_{i=1}^{n} \sigma_i\, p_j(\tau_i).$$

The method of Newton leads to an iteration

$$\mathbf{P}_c\mathbf{D} \begin{bmatrix} \sigma^{(k+1)} - \sigma^{(k)} \\ \tau^{(k+1)} - \tau^{(k)} \end{bmatrix} = m - F(\sigma^{(k)}, \tau^{(k)}),$$

where $\mathbf{P}_c = \mathbf{P}_c(\tau^{(k)})$ will be defined in the next section and

$$\mathbf{D} = \mathbf{D}(\sigma^{(k)}) = \operatorname{diag}(1, \ldots, 1, \sigma_1^{(k)}, \ldots, \sigma_n^{(k)}).$$

The calculation of the residue vector can be avoided. We have

$$F(\sigma, \tau) = \mathbf{P}_c(\tau) \begin{bmatrix} \sigma \\ \mathbf{0} \end{bmatrix} = \mathbf{P}_c(\tau)\mathbf{D}(\sigma) \begin{bmatrix} \sigma \\ \mathbf{0} \end{bmatrix},$$

and this implies that we have to solve the system

$$\mathbf{P}_c\mathbf{D} \begin{bmatrix} \sigma^{(k+1)} \\ \tau^{(k+1)} - \tau^{(k)} \end{bmatrix} = m$$

in each iteration.

The solution of linear systems with a general $n \times n$ matrix, of course, requires $O(n^3)$ arithmetic operations. Fortunately, there are known fast algorithms for this special structure of matrices with only $O(n^2)$ operations, if the polynomials p_j satisfy a three-term recurrence relation, see the bibliography in [Lu94]. In this 1994 paper of Hao Lu an even faster algorithm is proposed for confluent Vandermonde matrices, which is based on fast polynomial arithmetic and requires only $O(n \log^2 n)$ arithmetic operations. However, the resulting iterative procedure is slower than traditional methods. It is only interesting from a theoretical point of view, since it shows that time complexity can be reduced to $O(n \log^2 n)$.

In this chapter, we rederive some known $O(n^2)$-algorithms within the framework of polynomial arithmetic and extend the ideas of Hao Lu to the case of (confluent) Chebyshev-Vandermonde matrices.

4.1 Vandermonde-like matrices and their inversion

Many problems in interpolation or construction of quadrature formulas require the solution of linear systems with (possibly confluent) Vandermonde-like matrices or their transposes. If we are given a set of polynomials $\{p_j\}_{j=0,\dots,n-1}$, where p_j has degree j, and a set of distinct nodes $\{\tau_i\}_{i=1,\dots,n}$, then the $n \times n$ Vandermonde-like matrix is defined by

$$\mathbf{P} = \begin{bmatrix} p_0(\tau_1) & p_0(\tau_2) & \cdots & p_0(\tau_n) \\ p_1(\tau_1) & p_1(\tau_2) & \cdots & p_1(\tau_n) \\ \vdots & \vdots & & \vdots \\ p_{n-1}(\tau_1) & p_{n-1}(\tau_2) & \cdots & p_{n-1}(\tau_n) \end{bmatrix}. \tag{4.1}$$

The following lemma shows that the Lagrange interpolation base is important for our problem.

Lemma 4.1 *Let*

$$l_i(x) = \sum_{j=1}^{n} s_{ij}\, p_{j-1}(x) \quad \textit{for} \quad i = 1, 2, \dots, n.$$

Then we have

$$\mathbf{P}^{-1} = [s_{ij}]\,.$$

Proof: This follows immediately from $l_i(\tau_j) = \delta_{ij}$. $\square$

We consider here only the simplest case of a $2n \times 2n$ confluent Vandermonde-like matrix related to the set of polynomials $\{p_j\}_{j=0,\ldots,2n-1}$ and the set of nodes $\{\tau_i\}_{i=1,\ldots,n}$ defined by

$$\mathbf{P}_c = \begin{bmatrix} p_0(\tau_1) & \cdots & p_0(\tau_n) & p_0'(\tau_1) & \cdots & p_0'(\tau_n) \\ p_1(\tau_1) & \cdots & p_1(\tau_n) & p_1'(\tau_1) & \cdots & p_1'(\tau_n) \\ \vdots & & \vdots & \vdots & & \vdots \\ p_{2n-1}(\tau_1) & \cdots & p_{2n-1}(\tau_n) & p_{2n-1}'(\tau_1) & \cdots & p_{2n-1}'(\tau_n) \end{bmatrix}. \tag{4.2}$$

For the inversion of confluent Vandermonde-like matrices we need the fundamental polynomials of Hermite interpolation. The problem of the inversion of $\mathbf{P}_c$ is now equivalent to the transition from the base $\{h_i, k_i\}_{i=1,\ldots,n}$ of $\mathcal{P}_{2n-1}$ to $\{p_j\}_{j=0,\ldots,2n-1}$:

Lemma 4.2 *Let*

$$h_i(x) = \sum_{j=1}^{2n} a_{ij}\, p_{j-1}(x) \quad \textit{for} \quad i = 1, 2, \ldots, n.$$

and

$$k_i(x) = \sum_{j=1}^{2n} b_{ij}\, p_{j-1}(x) \quad \textit{for} \quad i = 1, 2, \ldots, n.$$

Then we have

$$\mathbf{P}_c^{-1} = \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix},$$

where

$$\mathbf{A} = [a_{ij}] \quad \textit{and} \quad \mathbf{B} = [b_{ij}]\,.$$

4.2 Auxiliary bases

The connection to interpolation suggests the use of interpolational polynomials in Newton form. Thus, we consider the base $\{\omega_j\}_{j=0,\ldots,n-1}$ in $\mathcal{P}_{n-1}$ defined by

$$\omega_j(x) = \prod_{i=1}^{j} (x - \tau_i) \quad \text{for} \quad j = 0, \ldots, n-1$$

(where empty products as usually are assumed to be one). Obviously, the polynomials satisfy a simple recurrence relation, since we have

$$\omega_j(x) = \omega_{j-1}(x)(x - \tau_j) \quad \text{for} \quad j \geq 1 \tag{4.3}$$

(this is a special case of (1.13)). It's elementary to express the polynomials ω_j by l_i:

Lemma 4.3 *We can write*

$$\omega_{j-1} = \sum_{i=1}^{n} u_{ji} l_i,$$

where the coefficients are $u_{ji} = \omega_{j-1}(\tau_i)$, *and*

$$\mathbf{U} = [u_{ji}]$$

is an upper triangular matrix.

Proof: The equation follows immediately from (1.16), and $\omega_{j-1}(\tau_i) = 0$ for $i \leq j-1$ is trivial.□

We stress here that the matrix $\mathbf{U}$ can be computed in $O(n^2)$ operations, since $u_{j+1,i} = \omega_j(\tau_i) = \omega_{j-1}(\tau_i)(\tau_i - \tau_j) = u_{ji}\,(\tau_i - \tau_j)$. This, of course, means that for any vector $\mathbf{v}$ the vector $\mathbf{U}^{-1}\mathbf{v}$ can be computed in $O(n^2)$ arithmetic operations, too.

On the other hand, we can express the polynomials ω_{j-1} in terms of the p_{k-1}, and this gives finally

Theorem 4.4 *We assume that the base* $\{p_0, \ldots, p_{n-1}\}$ *has Property X. Then we have the decomposition*

$$\mathbf{P} = \mathbf{V}^{-1}\mathbf{U} \quad \textit{and thus} \quad \mathbf{P}^{-1} = \mathbf{U}^{-1}\mathbf{V}, \tag{4.4}$$

where the matrix $\mathbf{V} = [v_{jk}]$ *is defined via*

$$\omega_{j-1} = \sum_{k=1}^{n} v_{jk} p_{k-1}$$

and can be computed in $O(n^2)$ *arithmetic operations. If the condition (1.11) is satisfied, then* $\mathbf{V}$ *is lower triangular.*

Proof: From the defining equation

$$\omega_{j-1}(x) = \sum_{k=1}^{n} v_{jk} p_{k-1}(x)$$

with $x = \tau_i$ we obtain

$$\omega_{j-1}(\tau_i) = \sum_{k=1}^{n} v_{jk} p_{k-1}(\tau_i),$$

and this in matrix notation reads $\mathbf{U} = \mathbf{VP}$. As a consequence of Equation (4.3) and Property X, the $(j+1)$-th row of $\mathbf{V}$ can be generated from the j-th row in $O(n)$ arithmetic operations. For the whole matrix $\mathbf{V}$ this sums up to $O(n^2)$ arithmetic operations. And, of course, if (1.11) is fulfilled, then we must have $v_{jk} = 0$ for $k > j$, since the degree of ω_{j-1} is $j-1$. □

```
for k = 1 to n do
  v_k ← p*_{k-1}(1){coordinates of polynomial ≡ 1}
end for
for j = 1 to n do
  y_j ← Σ_{k=1}^{n} v_k b_{k-1}
  if j < n then
    [v_1, ..., v_n] ← Ψ([v_1, ..., v_n]) − τ_j[v_1, ..., v_n] {using Property X}
  end if
end for
for i = n downto 1 do
  u_1 ← 1
  for j = 1 to i − 1 do
    u_{j+1} ← (τ_i − τ_j) × u_j
  end for
  y_i ← y_i/u_i
  for j = 1 to i − 1 do
    y_j ← y_j − u_j × y_i
  end for
end for
```

Algorithm 4.1: Solve Vandermonde-like system

Our result gives a simple algorithm for the solution $\mathbf{y} = [y_1, \ldots, y_n]^T$ of the system $\mathbf{Py} = \mathbf{b}$ with $\mathbf{b} = [b_0, \ldots, b_{n-1}]^T$ (note that we have to store only the actual row $[v_1, \ldots, v_n]$ of $\mathbf{V}$ and the actual column $[u_1, \ldots, u_n]$ of $\mathbf{U}$) (see Algorithm 4.1).

Of course, the system $\mathbf{P}^T\mathbf{y} = \mathbf{b}$ can be solved by the same method, based on the decomposition $\left(\mathbf{P}^T\right)^{-1} = \mathbf{V}^T\left(\mathbf{U}^T\right)^{-1}$. This is the original idea of [Hig88] — $\left(\mathbf{U}^T\right)^{-1}\mathbf{b}$ is the process of generating divided differences from function values at the interpolation nodes.

It's not very hard to see how we can extend this method to confluent systems. If we had in our matrix instead of the derivative $p_k'(\tau_i)$ the difference quotient

$$\frac{p_k(\tau_i + h) - p_k(\tau_i)}{h},$$

then our task obviously would reduce to the inversion of a Vandermonde-like $2n \times 2n$-matrix with the set of nodes $\{\tau_i, \tau_i + h\}_{i=1,\ldots,n}$. The limit $h \to 0$ then will give the following auxiliary base $\nu_0, \ldots, \nu_{2n-1}$ in $\mathcal{P}_{2n-1}$:

$$\nu_j(x) = \begin{cases} \omega_i^2(x) & \text{for} \quad j = 2i \\ \omega_i^2(x)(x - \tau_{i+1}) & \text{for} \quad j = 2i + 1 \end{cases}, \quad i = 0, \ldots n - 1.$$

This implies the recursive form of definition

$$\nu_{j+1}(x) = \nu_j(x)\left(x - \tau_{\lceil j/2 \rceil}\right) \quad \text{for} \quad j = 0, \ldots, 2n - 2 \tag{4.5}$$

and with the initial value $\nu_0 \equiv 1$. Differentiating we obtain a recursion for ν_j',

$$\nu_{j+1}'(x) = \nu_j'(x)\left(x - \tau_{\lceil j/2 \rceil}\right) + \nu_j(x) \quad \text{for} \quad j = 0, \ldots, 2n-2 \tag{4.6}$$

with the initial value $\nu_0' \equiv 0$. We can express the polynomials ν_{j-1} in the base $\{h_i, k_i\}_{i=1,\ldots,n}$.

Lemma 4.5 *We can write*

$$\nu_{j-1} = \sum_{i=1}^{n} \left(u_{j,2i-1}\, h_i + u_{j,2i}\, k_i\right),$$

where the coefficients are $u_{j,2i-1} = \nu_{j-1}(\tau_i)$ *and* $u_{j,2i} = \nu_{j-1}'(\tau_i)$, *and*

$$\mathbf{U} = [u_{ji}]$$

is an upper triangular matrix.

Proof: The equation follows immediately from (1.18). The fact that $\mathbf{U}$ is upper triangular is again trivial.□

We can prove the following decomposition of the modified matrix $\mathbf{P}_c$ (the proof is almost literally the same as the proof of (4.4) and can be omitted).

Theorem 4.6 *Let the matrix* $\tilde{\mathbf{P}}_c$ *be defined as*

$$\tilde{\mathbf{P}}_c = \begin{bmatrix} p_0(\tau_1) & p_0'(\tau_1) & \cdots & p_0(\tau_n) & p_0'(\tau_n) \\ p_1(\tau_1) & p_1'(\tau_1) & \cdots & p_1(\tau_n) & p_1'(\tau_n) \\ \vdots & \vdots & & \vdots & \vdots \\ p_{2n-1}(\tau_1) & p_{2n-1}'(\tau_1) & \cdots & p_{2n-1}(\tau_n) & p_{2n-1}'(\tau_n) \end{bmatrix}. \tag{4.7}$$

We assume that the base $\{p_0, \ldots, p_{2n-1}\}$ *has Property X. Then we have the decomposition*

$$\tilde{\mathbf{P}}_c = \mathbf{V}^{-1}\mathbf{U} \quad \textit{and thus} \quad \tilde{\mathbf{P}}_c^{-1} = \mathbf{U}^{-1}\mathbf{V}, \tag{4.8}$$

where the matrix $\mathbf{V} = [v_{jk}]$ *is defined via*

$$\nu_{j-1} = \sum_{k=1}^{2n} v_{jk} p_{k-1}$$

and can be computed in $O(n^2)$ *arithmetic operations. If the condition (1.11) is satisfied, then* $\mathbf{V}$ *is lower triangular.*

This theorem gives a surprisingly simple algorithm for the solution $\mathbf{y} = [y_1, \ldots, y_{2n}]^T$ of the system $\tilde{\mathbf{P}}_c\mathbf{y} = \mathbf{b}$ with $\mathbf{b} = [b_0, \ldots, b_{2n-1}]^T$ (here we have to store only the actual row $[v_1, \ldots, v_{2n}]$ of $\mathbf{V}$ and two columns $[u_{1,1}, \ldots, u_{1,2n}]$ and $[u_{2,1}, \ldots, u_{2,2n}]$ of $\mathbf{U}$, see Algorithm 4.2).

for $k = 1$ to $2n$ **do**
  $v_k \leftarrow p^*_{k-1}(1)\{$coordinates of polynomial $\equiv 1\}$
end for
for $j = 1$ to $2n$ **do**
  $y_j \leftarrow \sum_{k=1}^{2n} v_k b_{k-1}$
  if $j < 2n$ **then**
    $[v_1, \ldots, v_{2n}] \leftarrow \Psi([v_1, \ldots, v_{2n}]) - \tau_{\lceil j/2 \rceil}[v_1, \ldots, v_{2n}]$ {using Property X}
  end if
end for
for $i = n$ downto 1 **do**
  $u_{11} \leftarrow 1$
  $u_{21} \leftarrow 0$
  for $j = 1$ to $2i - 1$ **do**
    $u_{1,j+1} \leftarrow (\tau_i - \tau_{\lceil j/2 \rceil}) \times u_{1,j}$ {see Equation (4.5)}
    $u_{2,j+1} \leftarrow (\tau_i - \tau_{\lceil j/2 \rceil}) \times u_{2,j} + u_{1,j}$ {see Equation (4.6)}
  end for
  $y_{2i} \leftarrow y_{2i}/u_{2,2i}$
  for $j = 1$ to $2i - 1$ **do**
    $y_j \leftarrow y_j - u_{2,j} \times y_{2i}$
  end for
  $y_{2i-1} \leftarrow y_{2i-1}/u_{1,2i-1}$
  for $j = 1$ to $2i - 2$ **do**
    $y_j \leftarrow y_j - u_{1,j} \times y_{2i-1}$
  end for
end for

Algorithm 4.2: Solve confluent Vandermonde-like system

4.3 The method of Hao Lu

In this section, we assume that the polynomials p_j are orthogonal with respect to some measure σ. There are several normalizations due to historical reasons, so we do not expect them to be monic or orthonormal. Instead, we assume the leading coefficient of p_j to be K_j. The polynomials satisfy a three-term recurrence (1.13), and their orthogonality relation is

$$\int p_j(x)\,p_k(x)\,d\sigma(x) = \frac{K_j^2}{\gamma_j^2}\,\delta_{kj} = H_j\,\delta_{kj}, \tag{4.9}$$

where the leading coefficients γ_j of *orthonormal* polynomials can be found from

$$\gamma_j^{-2} = \gamma_0^{-2}\prod_{k=1}^{j}\beta_k$$

($\gamma_0^{-2} = \int 1\,d\sigma(x)$ is the full weight of the measure). Since our polynomials are not necessarily monic, we define the *numerator polynomials* by

$$q_j(x) = \int \frac{p_j(x) - p_j(t)}{x - t}\,d\sigma(t) \tag{4.10}$$

(note the slight difference from Equation (12)).

It's well known that the polynomials q_j for $j > 0$ satisfy the same recurrence (1.13) with the initial values $q_0 \equiv 0$ and $q_1 \equiv \gamma_0^{-2}K_1$ (and this means that the degree of q_j is $j-1$). We can use them to formulate the following identity:

Lemma 4.7 *Let p_j be the orthogonal polynomials with respect to some measure σ satisfying Equation (1.13) and the orthogonality conditions (4.9), and let the numerator polynomials be defined via (4.10). Then we have*

$$\frac{p_n(x) - p_n(\tau)}{x - \tau} = \sum_{j=0}^{n-1}\frac{1}{H_j}\left[p_j(x)\,q_n(x) - p_n(x)\,q_j(x)\right]p_j(\tau). \tag{4.11}$$

Proof: We follow closely the proof of [NS88, Chapter 2, Lemma 7.4]. The expression

$$\frac{p_n(x) - p_n(\tau)}{x - \tau}$$

is a polynomial of degree $n-1$ in τ and thus it can be written as

$$\frac{p_n(x) - p_n(\tau)}{x - \tau} = \sum_{j=0}^{n-1} c_j\,p_j(\tau).$$

Multiplying both sides by $p_j(\tau)$ and integrating we obtain from (4.9)

$$\int \frac{p_n(x) - p_n(\tau)}{x - \tau}\,p_j(\tau)\,d\sigma(\tau) = H_j\,c_j.$$

The proposition now follows from the identity

$$\begin{aligned}\frac{p_n(x)-p_n(\tau)}{x-\tau}\,p_j(\tau) \quad &= \quad \frac{p_n(x)-p_n(\tau)}{x-\tau}\,p_j(x)-\frac{p_j(x)-p_j(\tau)}{x-\tau}\,p_n(x)+\\ &\quad +\frac{p_j(x)-p_j(\tau)}{x-\tau}\,p_n(\tau),\end{aligned}$$

since

$$\int\frac{p_n(x)-p_n(\tau)}{x-\tau}\,p_j(x)\,d\sigma(\tau)=q_n(x)\,p_j(x),$$

$$\int\frac{p_j(x)-p_j(\tau)}{x-\tau}\,p_n(x)\,d\sigma(\tau)=q_j(x)\,p_n(x)$$

and

$$\int\frac{p_j(x)-p_j(\tau)}{x-\tau}\,p_n(\tau)\,d\sigma(\tau)=0$$

due to orthogonality (observe that $\frac{p_j(x)-p_j(\tau)}{x-\tau}$ is a polynomial of degree $j-1<n$ in τ). $\square$

This auxiliary result enables us to extend the method of Hao Lu for the inversion of confluent Vandermonde matrices to arbitrary orthogonal polynomials. Now we are in a situation to formulate a variant of Theorem 1.1 of [Lu94].

Theorem 4.8 *The solution $\mathbf{y}=[y_1,\dots,y_n]^T$ of the Vandermonde-like system*

$$\mathbf{Py}=\mathbf{b}$$

with $\mathbf{b}=[b_0,\dots,b_{n-1}]^T$ is given by

$$y_i=\frac{g(\tau_i)}{\omega_n'(\tau_i)},$$

where

$$\omega_n(x)=(x-\tau_1)\dots(x-\tau_n),\quad g(x)=\omega_n(x)b(x)\div p_n(x)$$

and

$$b(x)=q_n(x)\sum_{j=0}^{n-1}\frac{b_j}{H_j}\,p_j(x)\bmod p_n(x).$$

Proof: We multiply the equation

$$\sum_{i=1}^{n} y_i p_j(\tau_i) = b_j$$

by $\dfrac{1}{H_j}\,[p_j(x)\,q_n(x) - p_n(x)\,q_j(x)]$ and sum up, this (in view of the identity (4.11)) gives the equation

$$\begin{aligned}\sum_{i=1}^{n} y_i\,\frac{p_n(x)-p_n(\tau_i)}{x-\tau_i} &= \sum_{j=0}^{n-1}\frac{b_j}{H_j}\,[p_j(x)\,q_n(x)-p_n(x)\,q_j(x)]\\ &= q_n(x)\sum_{j=0}^{n-1}\frac{b_j}{H_j}\,p_j(x) - p_n(x)\sum_{j=0}^{n-1}\frac{b_j}{H_j}\,q_j(x).\end{aligned}$$

Since the left-hand side (and thus the right-hand side) is a polynomial of degree $\leq n-1$, it coincides with $b(x)$. If we multiply this equation by $\omega_n(x)$, we have

$$p_n(x)\sum_{i=1}^{n} y_i\,\frac{\omega_n(x)}{x-\tau_i} - \sum_{i=1}^{n} y_i\,p_n(\tau_i)\frac{\omega_n(x)}{x-\tau_i} = \omega_n(x)b(x).$$

Now the polynomial $\displaystyle\sum_{i=1}^{n} y_i\,p_n(\tau_i)\frac{\omega_n(x)}{x-\tau_i}$ has degree $\leq n-1$, consequently

$$\sum_{i=1}^{n} y_i\,\frac{\omega_n(x)}{x-\tau_i} = \omega_n(x)b(x) \div p_n(x) = g(x).$$

Substituting $x=\tau_i$, we obtain $y_i\,\omega_n'(\tau_i) = g(\tau_i)$, and this proves our theorem. □

Now we formulate the result for the confluent case.

Theorem 4.9 *The solution*

$$\begin{bmatrix}\mathbf{y}\\ \mathbf{z}\end{bmatrix} = [y_1,\ldots,y_n,z_1,\ldots,z_n]^T$$

of the confluent Vandermonde-like system

$$\mathbf{P}_c\begin{bmatrix}\mathbf{y}\\ \mathbf{z}\end{bmatrix} = \mathbf{b}$$

with $\mathbf{b} = [b_0,\ldots,b_{2n-1}]^T$ *is given by*

$$y_i = \frac{g'(\tau_i)}{\omega_n'^2(\tau_i)} - 2\,\frac{\omega_n''(\tau_i)\,g(\tau_i)}{\omega_n'^3(\tau_i)} \quad \textit{and} \quad z_i = \frac{g(\tau_i)}{\omega_n'^2(\tau_i)},$$

where

$$\omega_n(x) = (x-\tau_1)\ldots(x-\tau_n), \quad g(x) = \omega_n^2(x)b(x) \div p_{2n}(x)$$

and

$$b(x) = q_{2n}(x)\sum_{j=0}^{2n-1}\frac{b_j}{H_j}\,p_j(x) \bmod p_{2n}(x).$$

Proof: From the identity (4.11) with $2n$ instead of n we obtain by differentiating with respect to τ

$$\frac{p_{2n}(x)-p_{2n}(\tau)}{(x-\tau)^2}-\frac{p_{2n}'(\tau)}{x-\tau}=\sum_{j=0}^{2n-1}\frac{1}{H_j}\left[p_j(x)\,q_{2n}(x)-p_{2n}(x)\,q_j(x)\right]p_j'(\tau).$$

Together with the original identity we obtain

$$\sum_{i=1}^{n}\left[y_i\,\frac{p_{2n}(x)-p_{2n}(\tau_i)}{x-\tau_i}+z_i\left\{\frac{p_{2n}(x)-p_{2n}(\tau_i)}{(x-\tau_i)^2}-\frac{p_{2n}'(\tau_i)}{x-\tau_i}\right\}\right]=b(x)$$

as above. Multiplying by $\omega_n^2(x)$ we have

$$p_{2n}(x)\sum_{i=1}^{n}\left[y_i\,\frac{\omega_n^2(x)}{x-\tau_i}+z_i\,\frac{\omega_n^2(x)}{(x-\tau_i)^2}\right]+r_{2n-1}(x)=b(x)\,\omega_n^2(x),$$

where

$$r_{2n-1}(x)=-\sum_{i=1}^{n}\left[(y_i\,p_{2n}(\tau_i)+z_i\,p_{2n}'(\tau_i))\,\frac{\omega_n^2(x)}{x-\tau_i}+z_i\,p_{2n}(\tau_i)\,\frac{\omega_n^2(x)}{(x-\tau_i)^2}\right]$$

is a polynomial of degree $\leq 2n-1$. This gives the equation

$$\sum_{i=1}^{n}\left[y_i\,\frac{\omega_n^2(x)}{x-\tau_i}+z_i\,\frac{\omega_n^2(x)}{(x-\tau_i)^2}\right]=b(x)\,\omega_n^2(x)\div p_{2n}(x)=g(x).$$

Now substituting $x=\tau_i$ we obtain

$$z_i\,\omega_n'^2(\tau_i)=g(\tau_i),$$

and differentiating with respect to x and substituting $x=\tau_i$ we have

$$y_i\,\omega_n'^2(\tau_i)+2z_i\,\omega_n''(\tau_i)\,\omega_n'(\tau_i)=g'(\tau_i).$$

Our proposition follows immediately from these two equations. □

4.4 Some numerical experiments

In this section, we report the results of some numerical experiments with the iterative method mentioned in the beginning of this chapter. More precisely, we considered a variant, namely the system

$$\mathbf{P}_c\mathbf{D}(\sigma^{(k+1)})\begin{bmatrix}\sigma^{(k+1)}\\ \tau^{(k+1)}-\tau^{(k)}\end{bmatrix}=m.$$

We assume that we are given Chebyshev moments, i.e. $p_k=T_k$. This makes our algorithm very efficient: Up to a numerical factor the U_n are the second kind polynomials for T_n. Division by T_n is trivial in view of the identities

$$2T_n(x)T_m(x)=T_{n+m}(x)+T_{|n-m|}(x)$$

Table 4.1: Defekt after k iterations, 1. case, $n = 500$

k	defekt
1	4.22E-04
2	1.44E-04
3	2.87E-05
4	4.67E-06
5	5.94E-08
6	4.95E-12
7	1.65E-16

and

$$2U_n(x)T_m(x) = \begin{cases} U_{n+m}(x) + U_{n-m}(x) & , n \geq m \\ U_{2m-1} & , n = m-1 \\ U_{n+m}(x) + U_{m-n-2}(x), & n \leq m-2 \end{cases}$$

and the product of two polynomials of degree n can be calculated by fast DCT in $O(n \log n)$ arithmetical operations. This is interpolation at the zeros of Chebyshev polynomials, and it was also used for the computation of the polynomial ω_n in Theorem 4.9.

As starting values $\tau^{(0)}$ we used the zeros of T_n, $\sigma^{(0)}$ is not needed. We didn't investigate conditions for convergence. The measure σ in our experiments was the measure of orthogonality of Jacobi-polynomials in the following three cases:

1. $\alpha = \beta = 0$ — Legendre polynomials,

2. $\alpha = \beta = \frac{1}{2}$ — Chebyshev polynomials of second kind,

3. $\alpha = -\beta = -\frac{1}{2}$ — Chebyshev polynomials of third kind.

As a defekt criterion, we used the largest (in absolute value) component of the computed correction $\tau^{(k+1)} - \tau^{(k)}$. The results are displayed in the tables 4.1–4.3.

Surprisingly, in the first case our iteration converged for all values of n we tried, up to $n = 500$. However, to obtain a defekt of the order of the machine eps, we needed 7 iterations. The CPU-time required was 7.09 seconds, whereas the traditional method (solving the eigenvalue problem for the Jacobi matrix) required only 0.88 seconds. The zeros of Chebyshev polynomials of second and third kinds, of course, are not so near to the starting values. The values of n displayed in these cases were the largest for which the iteration converged at all.

Table 4.2: Defekt after k iterations, 2. case, $n = 51$

k	defekt
1	8.65E-03
2	6.03E-03
3	5.06E-03
4	2.73E-03
5	9.90E-04
6	1.29E-04
7	2.32E-06
8	5.62E-10
9	1.76E-15

Table 4.3: Defekt after k iterations, 3. case, $n = 25$

k	defekt
1	3.33E-02
2	2.29E-02
3	1.78E-02
4	6.55E-03
5	1.61E-03
6	9.11E-05
7	2.40E-07
8	1.14E-12
9	1.17E-15

4.5 The numerical condition in special cases

Our algorithm implements the map G_n from modified moments m to the quadrature formula γ (see the definitions in Section 3.1). We will consider here the map $\tilde{G}_n$, instead (it differs from G_n by a trivial diagonal map). In the paper [Gau86], Gautschi proved the equation

$$\|\tilde{G}_n'\|_F = \left\{\int g_n(x)\, ds(x)\right\}^{\frac{1}{2}}$$

with some polynomial g_n of degree $4n-2$ determined by γ, which are defined as follows (this is essentially Equation (3.2) of [Gau86], though the result is implicit already in the proof of Theorem 3.1 of the article [Gau82]):

$$g_n(x) = \sum_{j=1}^{n} h_j^2(x) + \sum_{j=1}^{n} \sigma_j^{-2} k_j^2(x), \tag{4.12}$$

where σ_j are the weights of the Gauss-Christoffel quadrature rule and h_j and k_j are the fundamental polynomials of Hermite interpolation in the nodes of the quadrature rule. In our case, the nodes τ_i are the zeros of some orthogonal polynomial p_n. We remark here that in Equations (1.15) and (1.17), we can use the (generally not monic) polynomial p_n instead of ω_n, since the leading coefficient will simply cancel out.

4.5.1 Another interpolation identity

In the following, we will need an elementary interpolation identity of higher order.

Lemma 4.10 *Let $\{\tau_i\}_{i=1,\dots,n}$ be any set of distinct nodes, and let $b_i \in \mathcal{P}_{4n-1}$ be polynomials with the property*

$$b_i(\tau_j) = \delta_{ij}, \quad b_i'(\tau_j) = b_i''(\tau_j) = b_i'''(\tau_j) = 0 \quad \textit{for} \quad i,j = 1,\dots,n.$$

Then the identity

$$\sum_{i=1}^{n} b_i(x) \equiv 1$$

is satisfied.

Proof: Under our assumptions the polynomial

$$r(x) = \sum_{i=1}^{n} b_i(x) - 1 \in \mathcal{P}_{4n-1}$$

vanishs at all nodes τ_i together with all derivatives up to the third, and thus it is divisible by the polynomial $(x-\tau_1)^4 \cdots (x-\tau_n)^4$ having degree $4n$. Consequently, the polynomial r must be identically zero. □

In fact, the polynomials b_i are well defined by the above properties and can be written down explicitly in terms of the polynomials l_i.

Lemma 4.11 *Let l_i be the fundamental polynomials of Lagrange interpolation for the nodes $\{\tau_i\}_{i=1,\dots,n}$ and let the first few coefficients of the Taylor expansion of $l_i(x)$ around $x = \tau_i$ be*

$$l_i(x) = 1 + c_1(\tau_i)(x-\tau_i) + c_2(\tau_i)(x-\tau_i)^2 + c_3(\tau_i)(x-\tau_i)^3 + \cdots .$$

Then we have

$$b_i(x) = [1 + d_1(\tau_i)(x-\tau_i) + d_2(\tau_i)(x-\tau_i)^2 + d_3(\tau_i)(x-\tau_i)^3]\, l_i^4(x),$$

where the coefficients d_i can be obtained from

$$\begin{aligned} d_1 &= -4c_1, \\ d_2 &= 10c_1^2 - 4c_2, \\ d_3 &= -20c_1^3 + 20c_1c_2 - 4c_3. \end{aligned}$$

Proof: If $1 + d_1(\tau_i)(x-\tau_i) + d_2(\tau_i)(x-\tau_i)^2 + d_3(\tau_i)(x-\tau_i)^3$ coincides with the first few terms of the Taylor series of $l_i(x)^{-4}$, then the $b_i(x)$ as defined in our proposition satisfies $b_i(x) = 1 + O((x-\tau_i)^4)$. Since this b_i clearly has degree $\leq 4n-1$ and vanishes at each $\tau_j \neq \tau_i$, it follows that it is identical with the b_i defined in Lemma 4.10. The Taylor series can be obtained easily: From

$$\sum_{j=0}^{\infty} d_j(x-\tau_i)^j = \left(\sum_{j=0}^{\infty} c_j(x-\tau_i)^j\right)^{-4}$$

by differentiating and multiplying by $\sum_{j=0}^{\infty} c_j(x-\tau_i)^j$ we have the equation

$$\sum_{j=1}^{\infty} j\, d_j(x-\tau_i)^{j-1} \sum_{j=0}^{\infty} c_j(x-\tau_i)^j = -4 \sum_{j=1}^{\infty} j\, c_j(x-\tau_i)^{j-1} \sum_{j=0}^{\infty} d_j(x-\tau_i)^j .$$

A comparison of the coefficients gives the linear system

$$\begin{aligned} d_1 &= -4c_1, \\ c_1d_1 + 2d_2 &= -4(c_1d_1 + 2c_2), \\ c_2d_1 + 2c_1d_2 + 3d_3 &= -4(c_1d_2 + 2c_2d_1 + 3c_3), \end{aligned}$$

and solving for d_j we arrive at our proposition. □

4.5.2 Explicit calculation of $\sum_{i=1}^{n} \sigma_i^{-2} k_i^2(x)$ for Jacobi polynomials

Of course, we will be able to compute explicit formulas only in some special cases. The Jacobi polynomials $P_n^{(\alpha,\beta)}$ are orthogonal with respect to the measure defined by

$$d\sigma(x) = (1-x)^\alpha (1+x)^\beta\, dx,$$

their leading coefficient (in their traditional standardization as in [AS84, 22.2.1]) is

$$K_n = \frac{1}{2^n}\binom{2n+\alpha+\beta}{n}$$

and the constant in their orthogonality relation is

$$H_n = \frac{2^{\alpha+\beta+1}}{2n+\alpha+\beta+1}\,\frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{n!\Gamma(n+\alpha+\beta+1)}.$$

We will write simply p_n instead of $P_n^{(\alpha,\beta)}$. The three-term recurrence for p_n has the explicit form

$$\begin{aligned}2n(n+\alpha+\beta)(2n+\alpha+\beta-2)p_n(x) = {} & [(2n+\alpha+\beta-1)(\alpha^2-\beta^2)+ \\ & +(2n+\alpha+\beta-2)(2n+\alpha+\beta-1)(2n+\alpha+\beta)x]p_{n-1}(x)- \\ & -2(n+\alpha-1)(n+\beta-1)(2n+\alpha+\beta)p_{n-2}(x). \qquad (4.13)\end{aligned}$$

For these special polynomials we have some more useful relations. Especially, they satisfy the differential equation

$$(1-x^2)p_n''(x)+[\beta-\alpha-(\alpha+\beta+2)x]p_n'(x)+n(n+\alpha+\beta+1)p_n(x) = 0 \qquad (4.14)$$

and their derivative can be expressed as

$$\begin{aligned}(2n+\alpha+\beta)(1-x^2)p_n'(x) &= n[\alpha-\beta-(2n+\alpha+\beta)x]p_n(x)+ \\ &\quad +2(n+\alpha)(n+\beta)p_{n-1}(x). \qquad (4.15)\end{aligned}$$

Now we can prove the following

Theorem 4.12 *Let τ_i be the zeros of the Jacobi polynomial p_n. Then the polynomials k_i defined in Equations (1.17) satisfy the identity*

$$\sum_{i=1}^{n}\sigma_i^{-2}k_i^2(x) = C_n^2p_n^2(x)\left[p_{n-1}^2(x)-B_np_{n-2}(x)p_n(x)\right]$$

with the constants

$$C_n = \frac{K_{n-1}}{K_nH_{n-1}} \quad \textit{and} \quad B_n = \frac{(2n+\alpha+\beta)(n+\alpha-1)(n+\beta-1)}{(2n+\alpha+\beta-1)(n+\alpha)(n+\beta)}.$$

Proof: For any set of orthogonal polynomials the weights of Gauss-Christoffel quadrature formulas can be written as

$$\sigma_i^{-1} = C_np_n'(\tau_i)p_{n-1}(\tau_i).$$

From the definition (1.17) of the polynomials k_i and $l_i(x) = \dfrac{p_n(x)}{(x-\tau_i)p_n'(\tau_i)}$ we obtain easily

$$\sum_{i=1}^{n}\sigma_i^{-2}k_i^2(x) = C_n^2p_n^2(x)\sum_{i=1}^{n}p_{n-1}^2(\tau_i)\,l_i^2(x).$$

The sum on the right-hand side can be evaluated from the interpolation identity (1.18). First we observe, that from the differential equation the explicit formula

$$h_i(x) = \left[1 - \frac{(\alpha+\beta+2)\,\tau_i + \alpha - \beta}{1-\tau_i^2}\,(x-\tau_i)\right]\, l_i^2(x)$$

can be derived. Thus Equation (1.18) can be written as

$$\begin{aligned} &\sum_{i=1}^{n} q(\tau_i)\, l_i^2(x) = q(x) - \\ &\qquad \sum_{i=1}^{n}\left[q'(\tau_i) - \frac{(\alpha+\beta+2)\,\tau_i + \alpha - \beta}{1-\tau_i^2} q(\tau_i)\right](x-\tau_i)\, l_i^2(x) \end{aligned} \tag{4.16}$$

for any polynomial $q \in \mathcal{P}_{2n-1}$. Now, we set $q(x) = p_{n-1}^2(x)$. Since we have

$$(x-\tau_i)\, l_i(x) = \frac{1}{p_n'(\tau_i)}\, p_n(x)$$

and from Equation (4.15)

$$p_n'(\tau_i) = \frac{2(n+\alpha)(n+\beta)}{(2n+\alpha+\beta)(1-\tau_i^2)}\, p_{n-1}(\tau_i),$$

we obtain

$$\begin{aligned} &\left[q'(\tau_i) - \frac{(\alpha+\beta+2)\,\tau_i + \alpha - \beta}{1-\tau_i^2} q(\tau_i)\right]\frac{1}{p_n'(\tau_i)} = \\ &\frac{2n+\alpha+\beta}{(n+\alpha)(n+\beta)}\left\{(1-\tau_i^2)\, p_{n-1}'(\tau_i) - \frac{1}{2}[(\alpha+\beta+2)\,\tau_i + \alpha - \beta]\, p_{n-1}(\tau_i)\right\}. \end{aligned}$$

If we express $(1-\tau_i^2)\, p_{n-1}'(\tau_i)$ using (4.15) and eliminate the terms $\tau_i p_{n-1}(\tau_i)$ using (4.13), after an easy (but tedious) calculation we arrive at

$$\begin{aligned} &(1-\tau_i^2)\, p_{n-1}'(\tau_i) - \frac{1}{2}[(\alpha+\beta+2)\,\tau_i + \alpha - \beta]\, p_{n-1}(\tau_i) = \\ &\quad \frac{(n+\alpha-1)(n+\beta-1)}{2n+\alpha+\beta-1}\, p_{n-2}(\tau_i). \end{aligned}$$

This implies

$$\begin{aligned} &\sum_{i=1}^{n}\left[q'(\tau_i) - \frac{(\alpha+\beta+2)\,\tau_i + \alpha - \beta}{1-\tau_i^2} q(\tau_i)\right](x-\tau_i)\, l_i^2(x) \\ &= B_n p_n(x) \sum_{i=1}^{n} p_{n-2}(\tau_i)\, l_i(x), \end{aligned}$$

and the sum on the right clearly is equal to $p_{n-2}(x)$ due to the Lagrange interpolation formula (1.16). $\Box$

Since we are able to obtain further results in the special case of Chebyshev polynomials, where $\alpha, \beta = \pm\frac{1}{2}$, and the traditional standardization of Chebyshev polynomials is different from that of Jacobi polynomials, we reformulate our result in this case. While Chebyshev polynomials of first and second kinds are well known, the nomenclature for the mixed cases ($\alpha = -\beta$) was introduced by Gautschi, e.g. in [Gau92], and adopted by other authors (in the papers [ME93] and [Mas93] of Mason and Elliott we can find some properties of Chebyshev polynomials of third and fourth kinds).

First we recall the connection between these special Jacobi polynomials and Chebyshev polynomials (see Equations 22.5.30 – 22.5.32 in [AS84]):

$$T_n(x) = \frac{n!\sqrt{\pi}}{\Gamma\left(n+\frac{1}{2}\right)}\, P^{(-\frac{1}{2},-\frac{1}{2})}(x),$$

$$U_n(x) = \frac{(n+1)!\sqrt{\pi}}{2\Gamma\left(n+\frac{3}{2}\right)}\, P^{(\frac{1}{2},\frac{1}{2})}(x)$$

and

$$U_{2n}(x) = \frac{n!\sqrt{\pi}}{\Gamma\left(n+\frac{1}{2}\right)}\, P^{(\frac{1}{2},-\frac{1}{2})}(2x^2-1).$$

We will rewrite the last equation in a more convenient way. The trigonometric representation of Chebyshev polynomials implies

$$U_{2n}(x) = U_n(2x^2-1) + U_{n-1}(2x^2-1),$$

and this equation gives immediately

$$U_n(x) + U_{n-1}(x) = \frac{n!\sqrt{\pi}}{\Gamma\left(n+\frac{1}{2}\right)}\, P^{(\frac{1}{2},-\frac{1}{2})}(x).$$

Since we have $P^{(-\frac{1}{2},\frac{1}{2})}(x) = (-1)^n P^{(\frac{1}{2},-\frac{1}{2})}(-x)$, we obtain easily the equation

$$U_n(x) - U_{n-1}(x) = \frac{n!\sqrt{\pi}}{\Gamma\left(n+\frac{1}{2}\right)}\, P^{(-\frac{1}{2},\frac{1}{2})}(x).$$

That's why we can call the polynomials $V_n(x) = U_n(x) - U_{n-1}(x)$ *Chebyshev polynomials of third kind* and the polynomials $W_n(x) = U_n(x) + U_{n-1}(x)$ *Chebyshev polynomials of fourth kind.* Their trigonometric representation is

$$V_n(\cos\theta) = \frac{\cos\left(n+\frac{1}{2}\right)\theta}{\cos\frac{1}{2}\theta} \quad \text{and} \quad W_n(\cos\theta) = \frac{\sin\left(n+\frac{1}{2}\right)\theta}{\sin\frac{1}{2}\theta}, \tag{4.17}$$

see Equations (1.3) and (1.4) in the paper [ME93]. In view of the obvious property $W_n(x) = (-1)^n V_n(-x)$, there will be no need to consider the polynomials of fourth kind in the following.

Our Theorem 4.12 can be specialized for Chebyshev polynomials.

Corollary 4.13 *For Chebyshev polynomials of first kind, we have*

$$\sum_{i=1}^{n} \sigma_i^{-2} k_i^2(x) = \frac{1}{\pi^2} \left[1 - x^2 + \frac{3}{2n} T_{n-2}(x) T_n(x)\right] T_n^2(x),$$

whereas for Chebyshev polynomials of second kind,

$$\sum_{i=1}^{n} \sigma_i^{-2} k_i^2(x) = \frac{1}{\pi^2} \left[1 + \frac{3}{2(n+1)} U_{n-2}(x) U_n(x)\right] U_n^2(x),$$

and for Chebyshev polynomials of third kind,

$$\sum_{i=1}^{n} \sigma_i^{-2} k_i^2(x) = \frac{1}{4\pi^2} \left[2(1-x) + \frac{3}{2n+1} V_{n-2}(x) V_n(x)\right] V_n^2(x).$$

Proof: By simple substitution for the left-hand sum we have

$$\frac{1}{\pi^2} \left[T_{n-1}^2(x) - \frac{n - \frac{3}{2}}{n} T_{n-2}(x) T_n(x)\right] T_n^2(x)$$

in the first case,

$$\frac{1}{\pi^2} \left[U_{n-1}^2(x) - \frac{n - \frac{1}{2}}{n+1} U_{n-2}(x) U_n(x)\right] U_n^2(x)$$

in the second case, and

$$\frac{1}{4\pi^2} \left[V_{n-1}^2(x) - \frac{n-1}{n + \frac{1}{2}} V_{n-2}(x) V_n(x)\right] V_n^2(x)$$

in the third case. Now our proposition follows at once from the elementary identities

$$T_{n-1}^2(x) - T_{n-2}(x) T_n(x) = 1 - x^2,$$

$$U_{n-1}^2(x) - U_{n-2}(x) U_n(x) = 1$$

and

$$V_{n-1}^2(x) - V_{n-2}(x) V_n(x) = 2(1-x).\Box$$

4.5.3 An estimate of $\sum_{i=1}^{n} h_i^2(x)$ for $\alpha, \beta \in (-1, 0]$

Interestingly, the sum $\sum_{i=1}^{n} h_i^2(x)$ in this case is uniformly bounded.

Theorem 4.14 *Let τ_i be the zeros of $P_n^{(\alpha,\beta)}$ and $\alpha, \beta \in (-1, 0]$. Then the polynomials h_i defined in (1.17) satisfy the inequality*

$$\sum_{i=1}^{n} h_i^2(x) \leq 1 \quad \textit{for} \quad x \in [-1, 1].$$

Proof: All we have to prove is the nonnegativity of $h_i(x)$ in $[-1, 1]$: From (1.18) with $q \equiv 1$ we obtain

$$\sum_{i=1}^{n} h_i(x) \equiv 1,$$

consequently $0 \leq h_i(x) \leq 1$ and

$$\sum_{i=1}^{n} h_i^2(x) \leq \sum_{i=1}^{n} h_i(x) \equiv 1.$$

Since

$$h_i(x) = \left[1 - \frac{(\alpha + \beta + 2)\,\tau_i + \alpha - \beta}{1 - \tau_i^2}\,(x - \tau_i)\right]\, l_i^2(x)$$

and all $\tau_i \in (-1, 1)$, it is sufficient to show

$$1 - \frac{(\alpha + \beta + 2)\,\tau + \alpha - \beta}{1 - \tau^2}\,(x - \tau) \geq 0 \quad \text{for} \quad x \in [-1, 1], \tau \in (-1, 1).$$

The left-hand side is linear in x, so we have to show the inequality only for $x = -1$ and $x = 1$. For $x = -1$ we have

$$1 - \frac{(\alpha + \beta + 2)\,\tau + \alpha - \beta}{1 - \tau^2}\,(-1 - \tau) = \frac{(\alpha + \beta + 1)\,\tau + 1 + \alpha - \beta}{1 - \tau}.$$

Now the numerator is again a linear function in τ, and it takes the values $-2\beta \geq 0$ for $\tau = -1$ and $2 + 2\alpha > 0$ for $\tau = 1$, hence it is ≥ 0 for all $\tau \in (-1, 1)$. The case $x = 1$ can be handled in the same way: We have

$$1 - \frac{(\alpha + \beta + 2)\,\tau + \alpha - \beta}{1 - \tau^2}\,(1 - \tau) = \frac{1 - \alpha + \beta - (\alpha + \beta + 1)\,\tau}{1 + \tau},$$

and the numerator is $2 + 2\beta > 0$ for $\tau = -1$ and $-2\alpha \geq 0$ for $\tau = 1$. $\square$

Naturally, this bound is sharp — the sum is equal to 1 for $x = \tau_i$.

4.5.4 Explicit calculation of $\sum_{i=1}^{n} h_i^2(x)$ for Chebyshev polynomials

The explicit calculation of this sum is possible for Chebyshev polynomials of first, second and third (or mixed) kinds. Our starting point is the interpolation identity from Lemma 4.10. The coefficients c_j defined in Lemma 4.11 can be calculated from the differential equation (4.14) for any values of α and β. But we will see that for the subsequent simplifications we need explicit (and simple) formulas for $p_n'(\tau_i)$. This is possible only for Chebyshev polynomials.

Lemma 4.15 *For any zero τ of T_n we have*

$$(1-\tau^2)\,T_n'(\tau) = \frac{n}{U_{n-1}(\tau)} \quad \textit{and} \quad T_n'^2(\tau) = \frac{n^2}{1-\tau^2},$$

whereas for any zero τ of U_n,

$$(1-\tau^2)\,U_n'(\tau) = \frac{n+1}{U_{n-1}(\tau)} \quad \textit{and} \quad U_n'^2(\tau) = \left(\frac{n+1}{1-\tau^2}\right)^2,$$

and for any zero τ of the third-kind Chebyshev polynomial $V_n(x)$,

$$(1-\tau^2)\,V_n'(\tau) = \frac{n+\frac{1}{2}}{U_{n-1}(\tau)} \quad \textit{and} \quad V_n'^2(\tau) = \frac{(2n+1)^2}{2(1-\tau^2)(1+\tau)}.$$

Proof: This follows immediately from the well-known trigonometric representation of Chebyshev polynomials. Differentiating $T_n(\cos\theta) = \cos n\theta$ we obtain

$$\sin\theta\, T'(\cos\theta) = n\,\sin n\theta. \tag{4.18}$$

If $\cos\theta$ is a zero of T_n, then we have $\cos n\theta = 0$ and consequently $\sin^2 n\theta = 1$. This implies

$$\begin{aligned}&(1-\cos^2\theta)\,T'(\cos\theta) = \sin^2\theta\, T'(\cos\theta) = n\,\sin\theta\sin n\theta\\ &= \frac{n\,\sin\theta}{\sin n\theta} = \frac{n}{U_{n-1}(\cos\theta)}.\end{aligned}$$

From the above equation (4.18), by taking squares, we obtain

$$(1-\cos^2\theta)\,T'^2(\cos\theta) = \sin^2\theta\, T'^2(\cos\theta) = n^2\sin^2 n\theta = n^2,$$

if $\cos n\theta = 0$. The derivation of the other equations is very similar. If $\cos\theta$ is a zero of U_n, then $\sin(n+1)\theta = 0$ and consequently $\cos^2(n+1)\theta = 1$. From $\sin\theta\, U_n(\cos\theta) = \sin(n+1)\theta$ by differentiating we have the identity

$$-\sin^2\theta\, U_n'(\cos\theta) + \cos\theta\, U_n(\cos\theta) = (n+1)\,\cos(n+1)\theta, \tag{4.19}$$

and by squaring both sides (using $U_n(\cos\theta) = 0$ and $\cos^2(n+1)\theta = 1$) we obtain

$$(1-\cos^2\theta)^2\, U_n'^2(\cos\theta) = (n+1)^2.$$

From the equation (4.19) we conclude

$$\begin{aligned} -\sin^2\theta\, U_n'(\cos\theta) &= (n+1)\,\cos(n+1)\theta = \frac{n+1}{\cos(n+1)\theta} \\ &= \frac{(n+1)\,\sin\theta}{\sin\theta\cos(n+1)\theta - \cos\theta\sin(n+1)\theta} = \frac{(n+1)\,\sin\theta}{-\sin n\theta} = -\frac{n+1}{U_{n-1}(\cos\theta)}, \end{aligned}$$

where we used again $\cos^2(n+1)\theta = 1$ and $\sin(n+1)\theta = 0$. Observing $\sin^2\theta = 1-\cos^2\theta$ we arrive at our proposition.

The proof of the last equations relies on the trigonometric representation (4.17) and the fact that $\cos\left(n+\frac{1}{2}\right)\theta = 0$ implies $\sin^2\left(n+\frac{1}{2}\right)\theta = 1$ and

$$\sin\left(n+\frac{1}{2}\right)\theta = \frac{1}{\sin\left(n+\frac{1}{2}\right)\theta}$$

(we omit details here). □

Now we are able to formulate the main result of this section.

Theorem 4.16 *In the case of Chebyshev polynomials of first kind ($\alpha = \beta = -\frac{1}{2}$) the following formula holds:*

$$\sum_{i=1}^{n} h_i^2(x) = 1 - \left[\frac{2}{3} - \frac{1}{6n^2} - \frac{1}{2n^2}\cdot\frac{1-\dfrac{1}{2n}\,x\,U_{2n-1}(x)}{1-x^2}\right] T_n^2(x).$$

For Chebyshev polynomials of second kind ($\alpha = \beta = \frac{1}{2}$) we have

$$\begin{aligned} \sum_{i=1}^{n} h_i^2(x) &= 1 - \frac{1}{(n+1)^2}\left[\frac{7}{2}x^2 + \frac{2}{3}(n^2+n-3)(1-x^2)\right] U_n^2(x) \\ &\quad + \frac{1}{(n+1)^3}\left[\frac{9}{4}U_{n-2}(x) - \frac{5}{4}U_{n-4}(x)\right] U_n^3(x). \end{aligned}$$

Finally, for Chebyshev polynomials of third kind ($\alpha = -\frac{1}{2}$, $\beta = \frac{1}{2}$) we find

$$\begin{aligned} \sum_{i=1}^{n} h_i^2(x) &= 1 - \frac{1}{(2n+1)^2}\left[-4x + \frac{4}{3}(n^2+n)(x+1) - s_n(x)\right] V_n^2(x) \\ &\quad + \frac{2}{(2n+1)^3}\left[U_{n-2}(x) + U_{n-3}(x)\right] V_n^3(x), \end{aligned}$$

where

$$s_n(x) = \frac{1 + \dfrac{1}{2n+1}\left[U_{n-2}(x) - 3\,U_{n-1}(x)\right] V_n(x)}{1-x},$$

and

$$V_n(x) = U_n(x) - U_{n-1}(x).$$

Proof: As we can see easily, we have $h_i(x) = \left[1 + \frac{1}{2}\, d_1(\tau_i)(x-\tau_i)\right]\, l_i^2(x)$ and thus, using Lemma 3,

$$\begin{aligned} h_i^2(x) &= b_i(x) + \left[\left(\frac{1}{4}\, d_1^2(\tau_i) - d_2(\tau_i)\right)\,(x-\tau_i)^2 - d_3(\tau_i)\,(x-\tau_i)^3\right]\, l_i^4(x) \\ &= b_i(x) + p_n^2(x)\, \left[\left(\frac{1}{4}\, d_1^2(\tau_i) - d_2(\tau_i)\right) - d_3(\tau_i)\,(x-\tau_i)\right]\, \frac{1}{p_n'^2(\tau_i)}\, l_i^2(x) \end{aligned}$$

for any set of orthogonal polynomials. Now we use our explicit formulas for $p_n'^2(\tau_i)$ in the three special cases we consider.

Chebyshev polynomials of first kind: Computing the $c_j(\tau_i)$ from the differential equation (4.14), $d_j(\tau_i)$ from Lemma 4.11 and observing

$$h_i(x) = \left[1 - \tau_i\,(x-\tau_i)/(1-\tau_i^2)\right]\, l_i^2(x)$$

we obtain after an easy (but tedious) calculation

$$h_i^2(x) = b_i(x) + T_n^2(x)\, \left[\frac{1}{2n^2}\cdot\frac{1}{1-\tau_i^2}\, l_i^2(x) - \left(\frac{2}{3} - \frac{1}{6n^2}\right)\, h_i(x)\right].$$

From the above interpolation identities we know that

$$\sum_{i=1}^n h_i(x) = \sum_{i=1}^n b_i(x) \equiv 1,$$

but we have to evaluate the polynomial

$$r_n(x) = \sum_{i=1}^n \frac{1}{1-\tau_i^2}\, l_i^2(x).$$

The trick here is to calculate $(1-x^2)\, r_n(x)$, since we have

$$\begin{aligned} &\frac{1-x^2}{1-\tau_i^2}\, l_i^2(x) = \frac{1-\tau_i x}{1-\tau_i^2}\, l_i^2(x) - x\,\frac{x-\tau_i}{1-\tau_i^2}\, l_i^2(x) \\ &= h_i(x) - \frac{x\, T_n(x)}{n}\, U_{n-1}(\tau_i)\, l_i(x) \end{aligned}$$

(where we used $(x-\tau_i)\, l_i(x) = T_n(x)/T_n'(\tau_i)$ and Lemma 4.15), and this gives immediately

$$(1-x^2)\, r_n(x) = 1 - \frac{x\, T_n(x)}{n}\, U_{n-1}(x) = 1 - \frac{1}{2n}\, x\, U_{2n-1}(x).$$

Substituting this into the above equation we obtain our first proposition.

Chebyshev polynomials of second kind: As above we compute the representation

$$\begin{aligned} h_i^2(x) &= b_i(x) + \frac{1}{(n+1)^2}\, U_n^2(x)\left\{\left[-\frac{7}{2}\,\tau_i^2 - \frac{2}{3}\,(n^2+2n-3)(1-\tau_i^2)\right]\right. \\ &\qquad \left. + \left[\frac{10\,\tau_i^3}{1-\tau_i^2} + \left(\frac{10}{3}\, n^2 + \frac{20}{3}\, n - \frac{15}{2}\right)\, \tau_i\right]\,(x-\tau_i)\right\}\, l_i^2(x). \end{aligned}$$

The sum

$$\sum_{i=1}^{n}\left[-\frac{7}{2}\,\tau_i^2-\frac{2}{3}\,(n^2+2n-3)(1-\tau_i^2)\right]\,l_i^2(x)$$

can be evaluated by Equation (4.16)

$$\text{(with}\quad \alpha=\beta=\frac{1}{2}\quad\text{and}\quad q(x)=-\frac{7}{2}\,x^2-\frac{2}{3}\,(n^2+2n-3)(1-x^2)),$$

and this gives

$$\sum_{i=1}^{n}h_i^2(x)=1+\frac{1}{(n+1)^2}\,U_n^2(x)\left\{q(x)+\sum_{i=1}^{n}\frac{\frac{19}{2}\,\tau_i-10\tau_i^3}{1-\tau_i^2}\,(x-\tau_i)\,l_i^2(x)\right\}.$$

Now we can use $(x-\tau_i)\,l_i(x)=U_n(x)/U_n'(\tau_i)$ and the formula for $U_n'(\tau_i)$ from Lemma 4.15 to obtain

$$\sum_{i=1}^{n}\frac{\frac{19}{2}\,\tau_i-10\tau_i^3}{1-\tau_i^2}\,(x-\tau_i)\,l_i^2(x)=$$

$$\frac{1}{n+1}\,U_n(x)\sum_{i=1}^{n}\left(\frac{19}{2}\,\tau_i-10\tau_i^3\right)\,U_{n-1}(\tau_i)\,l_i(x).$$

Repeatedly applying the recurrence relation for the polynomials U_j and observing $U_n(\tau_i)=0$, we can calculate

$$\left(\frac{19}{2}\,\tau_i-10\tau_i^3\right)\,U_{n-1}(\tau_i)=\frac{9}{4}\,U_{n-2}(\tau_i)-\frac{5}{4}\,U_{n-4}(\tau_i),$$

and by the Lagrange interpolation formula we have

$$\sum_{i=1}^{n}\left[\frac{9}{4}\,U_{n-2}(\tau_i)-\frac{5}{4}\,U_{n-4}(\tau_i)\right]\,l_i(x)=\frac{9}{4}\,U_{n-2}(x)-\frac{5}{4}\,U_{n-4}(x).$$

Chebyshev polynomials of third kind: Our starting point is again a representation

$$h_i^2(x)=b_i(x)-\frac{V_n^2(x)}{(2n+1)^2}\,[C_0(\tau_i)+C_1(\tau_i)(x-\tau_i)]\,l_i^2(x),$$

where

$$C_0(\tau)=-4\tau+\frac{4}{3}(n^2+n)(\tau+1)+\frac{1}{\tau-1}$$

and

$$C_1(\tau)=\frac{2}{3}\,\frac{6(n^2+n-2)\tau^2+2(n^2+n+3)\tau-4n^2-4n+3}{\tau^2-1}$$

are rational functions of τ. The sum

$$\sum_{i=1}^{n}q(\tau_i)\,l_i^2(x)$$

with $q(\tau) = -4\tau + \frac{4}{3}(n^2+n)(\tau+1)$ can be calculated from formula (4.16) again. The sum on the right-hand side of (4.16) combines with $\sum_{i=1}^n C_1(\tau_i)(x-\tau_i)\, l_i^2(x)$ to give

$$\sum_{i=1}^n \frac{-4\tau_i^2+2}{1-\tau_i^2}(x-\tau_i)\, l_i^2(x) = \frac{V_n(x)}{n+\frac{1}{2}}\sum_{i=1}^n(-4\tau_i^2+2)\, U_{n-1}(\tau_i)\, l_i(x) =$$
$$\frac{V_n(x)}{n+\frac{1}{2}}\sum_{i=1}^n[-U_{n-3}(\tau_i)-U_{n-2}(\tau_i)]\, l_i(x) = \frac{V_n(x)}{n+\frac{1}{2}}[-U_{n-3}(x)-U_{n-2}(x)]$$

(where we used Lemma 4.15, the recurrence relation for U_j and the fact that $U_n(\tau_i) = U_{n-1}(\tau_i)$). However, we have to calculate the sum

$$s_n(x) = \sum_{i=1}^n \frac{1}{1-\tau_i}\, l_i^2(x).$$

We use the above trick to calculate $(1-x)\, s_n(x)$, since we have the identity

$$\frac{1-x}{1-\tau_i}\, l_i^2(x) = h_i(x) + \frac{\tau_i - 2}{1-\tau_i^2}(x-\tau_i)\, l_i^2(x).$$

With the same procedure as before we easily obtain the proposition.

This finally proves our theorem. We remark here that the simple but tedious algebraic manipulations were performed partially with MAPLE. □

Our last result will be the proof of an inequality conjectured by W. Gautschi already in his paper [Gau86].

Corollary 4.17 *In the case $d\sigma(x) = dx/\sqrt{1-x^2}$, the inequality $g_n(x) \le 1$ holds for all $x \in [-1,1]$ and all $n \ge 2$.*

Proof: As an immediate consequence of Theorem 4.16 and Corollary 4.13, we obtain the equation

$$g_n(x) = 1 - \left[\frac{2}{3} - \frac{1}{6n^2} - \frac{1}{2n^2}\, r_n(x) - \frac{1}{\pi^2}\left\{1 - x^2 + \frac{3}{2n}\, T_{n-2}(x)T_n(x)\right\}\right] T_n^2(x), \tag{4.20}$$

where as in the proof of Theorem 4.16, we use the abbreviation

$$r_n(x) = \sum_{i=1}^n \frac{1}{1-\tau_i^2}\, l_i^2(x).$$

In order to estimate $r_n(x)$, we use another way to evaluate this sum. From the Equation (1.15) and Lemma 4.15 we have

$$\frac{1}{1-\tau_i^2}\, l_i^2(x) = \frac{1}{n^2}\frac{T_n^2(x)}{(x-\tau_i)^2} = \frac{1}{n^2}\left[\frac{T_n(x)-T_n(\tau_i)}{x-\tau_i}\right]^2.$$

Now the right-most expression is a polynomial in τ_i of degree $2n-2$, and thus by virtue of the quadrature formula

$$r_n(x) = \frac{1}{\pi n}\sum_{i=1}^{n}\sigma_i\left[\frac{T_n(x)-T_n(\tau_i)}{x-\tau_i}\right]^2 = \frac{1}{\pi n}\int\left[\frac{T_n(x)-T_n(y)}{x-y}\right]^2 d\sigma(y).$$

Using the identity

$$\frac{T_n(x)-T_n(y)}{x-y} = U_{n-1}(x) + 2\sum_{k=1}^{n-1}U_{k-1}(x)T_{n-k}(y)$$

(which can be derived easily from the generating functions of Chebyshev polynomials) and the orthogonality of T_{n-k} with respect to σ, we obtain

$$r_n(x) = \frac{1}{n}\left[U_{n-1}^2(x) + 2\sum_{k=1}^{n-1}U_{k-1}^2(x)\right].$$

This equation implies the inequality

$$r_n(x) \le \frac{1}{n}\left(n^2 + 2\sum_{k=1}^{n-1}k^2\right) = \frac{2}{3}n^2 + \frac{1}{3}.$$

Observing $T_{n-2}(x)T_n(x) = x^2 - 1 + T_{n-1}^2(x) \le x^2$ and thus

$$1 - x^2 + \frac{3}{2n}T_{n-2}(x)T_n(x) \le 1 - x^2 + \frac{3}{2n}x^2 \le 1,$$

we see that the expression in square brackets in Equation (4.20) is not less than

$$\frac{1}{3} - \frac{1}{3n^2} - \frac{1}{\pi^2} \ge \frac{1}{4} - \frac{1}{\pi^2} > 0$$

for $x \in [-1,1]$ and $n \ge 2$. □

4.6 Another application of the method of Hao Lu

Some problems in weighted L^2-approximation require the solution of linear systems with Gram matrices. The $n\times n$ Gram matrix is defined by

$$\mathbf{G} = [h_{ij}]_{i,j=1}^{n} \text{ with } h_{ij} = \int p_{i-1}(x)p_{j-1}(x)\,d\sigma(x),$$

where σ is some measure. Of course, all integrals have to be finite. Furthermore, we assume that the support $S(\sigma)$ of our measure contains at least n points. The monic orthogonal polynomials π_j with respect to σ satisfy the three-term recurrence relation (2).

We recall here the Gauss-Christoffel quadrature (see Section 0.2): There is an n-point quadrature formula with nodes $\{\tau_i\}_{i=1,\dots,n}$ and weights $\{\sigma_i\}_{i=1,\dots,n}$ integrating exactly all polynomials of degree $\leq 2n-1$.

$\{\tau_i\}_{i=1,\dots,n}$ are the zeros of π_n and the weights σ_i are defined by

$$\sigma_i^{-1} = \gamma_{n-1}^2 \pi_{n-1}(\tau_i)\pi_n'(\tau_i) \quad \text{for} \quad i = 1, \dots, n.$$

This immediately gives a factorization of $\mathbf{G}$:

Lemma 4.18 *With the notations defined before and* $\mathbf{D} = \mathrm{diag}(\sigma_1, \dots, \sigma_n)$ *we have the representation*

$$\mathbf{G} = \mathbf{PDP}^T,$$

where the Vandermonde-type matrix $\mathbf{P}$ *is defined by (4.1).*

The proof of this lemma is trivial.

We assume in this section that the polynomials are monic and orthogonal with respect to some measure s. To fix notations, we rewrite the relevant equations of Section 0.1 for polynomials orthogonal w.r.t. the measure s:

The p_j satisfy a three-term recurrence relation of the form

$$xp_j(x) = p_{j+1}(x) + a_j p_i(x) + b_j p_{j-1}(x)$$

for $j = 0, 1, \dots$, where $p_{-1} \equiv 0$ and $p_0 \equiv 1$ (this includes all sets of orthogonal polynomials).

The polynomials p_j have the orthogonality relations

$$\int p_j(x)\, p_k(x)\, ds(x) = c_j^{-2}\, \delta_{kj},$$

where the leading coefficients c_j of the *orthonormal* polynomials can be found from

$$c_j^{-2} = c_0^{-2} \prod_{k=1}^{j} b_k$$

($c_0^{-2} = \int 1\, ds(x)$ is the full weight of the measure). Let

$$q_j(x) = c_0^2 \int \frac{p_j(x) - p_j(t)}{x-t}\, ds(t)$$

be the monic polynomials of second kind.

Now we can formulate our result for Gram systems.

Theorem 4.19 *Let* $\mathbf{y} = [y_0, \ldots, y_{n-1}]^T$ *be the solution of the linear system*

$$\mathbf{G}\mathbf{y} = \mathbf{z},$$

where $\mathbf{G}$ *is defined as above and* $\mathbf{z} = [z_0, \ldots, z_{n-1}]^T$*. Then the polynomial*

$$y(x) = \sum_{j=0}^{n-1} y_j p_j(x)$$

can be computed from

$$y(x) = \gamma_{n-1}^2 g(x)\, \pi_{n-1}(x) \bmod \pi_n(x),$$

with

$$g(x) = z(x)\, \pi_n(x) \div p_n(x)$$

and

$$z(x) = c_0^{-2}\, q_n(x) \sum_{j=0}^{n-1} c_j^2\, z_j\, p_j(x) \bmod p_n(x).$$

Proof: The solution $\mathbf{g} = [g_1, \ldots, g_n]^T$ of $\mathbf{P}\mathbf{g} = \mathbf{z}$ is given by

$$g_i = \frac{g(\tau_i)}{\pi_n'(\tau_i)}$$

(since we have $\omega_n(x) = \pi_n(x)$), and multiplying by $\mathbf{D}^{-1}$ we obtain

$$\mathbf{P}^T\mathbf{y} = [\gamma_{n-1}^2\pi_{n-1}(\tau_1)g(\tau_1), \ldots, \gamma_{n-1}^2\pi_{n-1}(\tau_n)g(\tau_n)]^T.$$

But we have

$$\mathbf{P}^T\mathbf{y} = [y(\tau_1), \ldots, y(\tau_n)]^T$$

by the definition of the polynomial y, and if the polynomial $y \in \mathcal{P}_{n-1}$ coincides with $\gamma_{n-1}^2\pi_{n-1}(x)g(x)$ in all zeros of π_n, then this implies

$$y(x) = \gamma_{n-1}^2\pi_{n-1}(x)g(x) \bmod \pi_n(x).\square$$

4.7 Bibliographical references

The map G_n was introduced by Gautschi mainly to study the numerical condition of K_n, see [Gau78], [Gau82] and [Gau86]. In this analysis, he used known results on inverses of confluent Vandermonde-like matrices. In practice, the map G_n is rather implemented as $H_n^{-1} \circ K_n$, since H_n^{-1} is an eigenvalue problem for the tridiagonal Jacobi matrix. These problems can be solved very efficiently, see [Gau90], [Gau96] and [Gau94]. As the numerical experiments showed, our direct method is much slower even for very large n. However, the explicit formulas for the solution of (confluent) Vandermonde-like systems are interesting in its own, as well as the application to Gram systems. The latter result was inspired by an article of Gemignani [Gem97], where he re-invents a known method (see [HHR89]).

The results on the structure of the inverses of (confluent) Vandermonde-like matrices are known for long. Algorithm 4.1 and Algorithm 4.2 are due to Higham, see his papers [Hig88] and [Hig90]. Kailath and Olshevsky use another approach in their paper [KO97], where many bibliographical references are given.

Our results, Theorem 4.8 and Theorem 4.9, were inspired by the article of Hao Lu [Lu94]. In the meantime, he generalized his results in the article [Lu96]. In fact, our result is just the identification of what Hao Lu calls "J-match" in terms of numerator polynomials. The implementation in the case of Chebyshev moments uses fast polynomial arithmetic (the algorithms for fast discrete cosine transforms are taken from the paper of Steidl and Tasche [ST91]). This makes our algorithm faster than that of Gohberg and Olshevsky in [GO94], though their method is based essentially on the same identity (4.11).

The idea of using nested polynomial division (and FFT) to speed up the evaluation of a polynomial in n given points (this gives the operation count of $O(n \log^2 n)$ in [Lu94]) was used already in the paper of Finck, Heinig and Rost [FHR93] in an algorithm for Cauchy-Vandermonde systems (with the same operation count). We add the remark that the evaluation of a polynomial in the Chebyshev base at n given points can be speeded up (asymptotically) by another method based on interpolation and FFT, see the article of Dutt and Rokhlin [DR93].

The explicit calculation of some polynomials ocurring in Gautschi's analysis is possible for special cases of Jacobi weights. The results are published in the article [Fis98a]. We were able to prove a conjecture in his paper [Gau86] concerning the uniform boundedness of $g_n(x)$ for the Chebyshev weight. A weaker result was shown already in our paper [Fis96].

Chapter 5

From quadrature data to recurrence coefficients

5.1 Stieltjes procedure and its stability

The coefficients α_j and β_j can be expressed by simple formulas (see (8) and (9)). These formulas together with the recurrence relation (2) can be used to calculate recursively as many coefficients α_j and β_j (and polynomials π_j) as desired (compare Gautschi [Gau82, p. 292] or [Gau90, p. 198]): Starting from $\pi_{-1} = 0$ and $\pi_0 = 1$ we can compute α_0 from (8) with $j = 0$. Now (2) gives the new polynomial π_1 allowing to compute α_1 from (8) and β_1 from (9).

At step k of this algorithm we have calculated the recurrence coefficients up to α_{k-1} and β_{k-1}. Then we use the recurrence relation (2) with $j = k-1$ to compute π_k and from this (via (8) and (9) with $j = k$) we obtain the new coefficients α_k and β_k.

As was mentioned by Gautschi [Gau90, p. 198], in order to implement this algorithm we have to find an appropriate codification of the polynomials π_j allowing the evaluation of the integrals in (8) and (9). If the given measure is discrete or has a discretization

$$\sigma^{(N)} = \sum_{i=1}^{N} \sigma_i \, \delta_{\tau_i}$$

of sufficiently high degree of exactness, then we can use equation (14) to compute the integrals. In this case, the polynomial π_j will be represented by the vector $[\pi_j(\tau_1), \ldots, \pi_j(\tau_N)]^T$. In this form the algorithm is called *(discretized) Stieltjes procedure.*

The stability of the algorithm depends on the stability of the recurrence relation (2). Though we are not able to analyze this difficult question completely, we can say that the recurrence (2) is unstable in isolated mass points x of the measure σ. We will give rough estimates for the order of magnitude of the resulting error magnification (compare the similar analysis in the paper of Gautschi [Gau93]).

We start from the well-known fact that instability of the recurrence relation (2) is caused by the introduction (via rounding errors) of another independent solution of the recursive equation. Thus we compute not $\pi_j(x)$, but $\pi_j(x)+\varepsilon\kappa_j(x)$, where ε is of the order of the machine eps and $\kappa_j(x)$ are the numerator polynomials (see Section 1.3). Consequently, instead of the quantities

$$\int \pi_j^2(x)\, d\sigma(x) \quad \text{and} \quad \int x\pi_j^2(x)\, d\sigma(x)$$

occuring in equations (8) and (9) we calculate

$$\int (\pi_j(x) + \varepsilon\kappa_j(x))^2\, d\sigma(x) = \int \pi_j^2(x)\, d\sigma(x) + \varepsilon^2 \int \kappa_j^2(x)\, d\sigma(x)$$

and

$$\int x(\pi_j(x) + \varepsilon\kappa_j(x))^2\, d\sigma(x) = \int x\pi_j^2(x)\, d\sigma(x) + 2\varepsilon\gamma_j^{-2} + \varepsilon^2 \int x\kappa_j^2(x)\, d\sigma(x)$$

(observe that $\kappa_j(x) \in \mathcal{P}_{j-1}$ and that $x\kappa_j(x)$ is a monic polynomial $\in \mathcal{P}_j$). Since we consider measures σ with bounded support $S(\sigma)$, the error in the recurrence coefficients α_j and β_j is characterized by

$$E_j = \varepsilon^2 \frac{\int \kappa_j^2(x)\, d\sigma(x)}{\int \pi_j^2(x)\, d\sigma(x)} = \varepsilon^2\gamma_j^2 \int \kappa_j^2(x)\, d\sigma(x).$$

In the Section 5.3 we show that the numerator polynomials may grow exponentially with their degree in isolated mass points of the measure σ (see Lemma 5.10 below).

Proposition 5.1 *If τ is a mass point with weight λ, we can expect an error magnification $M_n = E_n/\varepsilon$ in the coefficients ρ_n that satisfies*

$$M_n \geq \varepsilon\lambda\gamma_n^2\kappa_n^2(\tau).$$

In contrast to this, if the support of the measure has some regularity property, then the error magnification grows only polynomially with the degree n (see Lemma 5.6 in the Section 5.3).

5.2 Adding point masses

Nevai [Nev79] formulated his results in terms of the *orthonormal* polynomials $\hat{\pi}_j$:

Lemma 5.2 *Let σ be a measure, $\hat{\pi}_j$ the orthonormal polynomials with respect to this measure satisfying the recurrence relation (6) and $\nu = \sigma + \lambda\,\delta_\tau$. Then the recurrence coefficients for the orthogonal polynomials with respect to ν can be computed from*

$$\alpha_j(\nu) = \alpha_j + \lambda\,\frac{\beta_{j+1}^{1/2}\hat{\pi}_j(\tau)\hat{\pi}_{j+1}(\tau)}{1+\lambda\sum\limits_{i=0}^{j}\hat{\pi}_i^2(\tau)} - \lambda\,\frac{\beta_j^{1/2}\hat{\pi}_j(\tau)\hat{\pi}_{j-1}(\tau)}{1+\lambda\sum\limits_{i=0}^{j-1}\hat{\pi}_i^2(\tau)} \tag{5.1}$$

and

$$\beta_j(\nu) = \beta_j\,\frac{\left[1+\lambda\sum\limits_{i=0}^{j-2}\hat{\pi}_i^2(\tau)\right]\left[1+\lambda\sum\limits_{i=0}^{j}\hat{\pi}_i^2(\tau)\right]}{\left[1+\lambda\sum\limits_{i=0}^{j-1}\hat{\pi}_i^2(\tau)\right]^2}.$$

For computational purposes, it is better to work with the monic polynomials in order to avoid the square roots in (6) and (5.1). Moreover, if we apply the lemma to a discrete measure $\sigma = \sum_{i=1}^N \sigma_i\,\delta_{\tau_i}$, there is a difficulty: if all τ_i and τ are different and all σ_i and λ are positive (i.e. we have really $N+1$ mass points) then $\alpha_N(\nu)$ is defined, but we can't compute it from (5.1), since $\alpha_N(\sigma)$ is not defined.

For these reasons we rederive the relevant equations here. First, we need some auxiliary results. Using Lemma 1.1, we are able to prove a simple relation for all existing recurrence coefficients α_j of a discrete measure.

Lemma 5.3 *Let $\sigma = \sum_{i=1}^N \sigma_i\,\delta_{\tau_i}$ be a discrete measure, where $\sigma_i > 0$ for $i = 1,\dots,N$ and all τ_i are different. Then we have*

$$\sum_{j=0}^{N-1}\alpha_j = \sum_{i=1}^{N}\tau_i.$$

Proof: We prove first that

$$\pi_N(x) = (x-\tau_1)\dots(x-\tau_N). \tag{5.2}$$

Indeed, both polynomials of the same degree N are monic, and so its difference $r(x)$ is in $\mathcal{P}_{N-1}$. Consequently, r is orthogonal to π_N. But r is also orthogonal to the polynomial $(x-\tau_1)\dots(x-\tau_N)$. Thus r must be orthogonal to itself, i.e.

$$\int r^2(x)\,d\sigma(x) = \sum_{i=1}^{N}\sigma_i r^2(\tau_i) = 0.$$

From this we see that $r(\tau_i) = 0$ for $i = 1,\dots,N$. Now $r \in \mathcal{P}_{N-1}$ implies $r \equiv 0$. Comparing the leading coefficients of both sides in equation (5.2), we immediately obtain our proposition: The coefficient of x^{N-1} in $\pi_N(x)$ is $-\sum_{j=0}^{N-1}\alpha_j$ (see Lemma 1.1), and the same coefficient in the right-hand side is $-\sum_{i=1}^{N}\tau_i$, thus our proposition is true. $\square$

After these preparations we can formulate the main result.

Theorem 5.4 *Let $\sigma = \sum_{i=1}^{N} \sigma_i \, \delta_{\tau_i}$ be a discrete measure, where $\sigma_i > 0$ for $i = 1, \ldots, N$ and all τ_i are different. Further, let $\lambda > 0$, τ be different from all τ_i, and $\nu = \sigma + \lambda \delta_\tau$. Then we have*

$$\alpha_j(\nu) = \alpha_j + \lambda \, \frac{\gamma_j^2 \pi_{j+1}(\tau) \pi_j(\tau)}{1 + \lambda \sum\limits_{i=0}^{j} \gamma_i^2 \pi_i^2(\tau)} - \lambda \, \frac{\gamma_{j-1}^2 \pi_j(\tau) \pi_{j-1}(\tau)}{1 + \lambda \sum\limits_{i=0}^{j-1} \gamma_i^2 \pi_i^2(\tau)} \tag{5.3}$$

and

$$\beta_j(\nu) = \beta_j \, \frac{\left[1 + \lambda \sum\limits_{i=0}^{j-2} \gamma_i^2 \pi_i^2(\tau)\right] \left[1 + \lambda \sum\limits_{i=0}^{j} \gamma_i^2 \pi_i^2(\tau)\right]}{\left[1 + \lambda \sum\limits_{i=0}^{j-1} \gamma_i^2 \pi_i^2(\tau)\right]^2} \tag{5.4}$$

for $j < N$, and furthermore

$$\alpha_N(\nu) = \tau - \lambda \, \frac{\gamma_{N-1}^2 \pi_N(\tau) \pi_{N-1}(\tau)}{1 + \lambda \sum\limits_{i=0}^{N-1} \gamma_i^2 \pi_i^2(\tau)} \tag{5.5}$$

and

$$\beta_N(\nu) = \frac{\lambda \gamma_{N-1}^2 \pi_N^2(\tau) \left[1 + \lambda \sum\limits_{i=0}^{N-2} \gamma_i^2 \pi_i^2(\tau)\right]}{\left[1 + \lambda \sum\limits_{i=0}^{N-1} \gamma_i^2 \pi_i^2(\tau)\right]^2}. \tag{5.6}$$

Proof: We denote the monic orthogonal polynomials with respect to ν by p_j. Then we must have

$$p_j(x) = \pi_j(x) + \sum_{l<j} c_{jl} \pi_l(x) \tag{5.7}$$

with some unknown coefficients c_{jl}. Multiplying this by $\pi_l(x)$ und integrating we obtain

$$\begin{aligned} 0 &= \int p_j(x) \pi_l(x) \, d\nu(x) = \int p_j(x) \pi_l(x) \, d\sigma(x) + \lambda p_j(\tau) \pi_l(\tau) \\ &= c_{jl} \gamma_l^{-2} + \lambda p_j(\tau) \pi_l(\tau) \end{aligned}$$

and from this

$$c_{jl} = -\lambda \gamma_l^2 p_j(\tau) \pi_l(\tau).$$

Putting the last result into Eq. (5.7) for $x = \tau$ and then solving for $p_j(\tau)$, we finally obtain

$$p_j(\tau) = \frac{\pi_j(\tau)}{1 + \lambda \sum\limits_{i=0}^{j-1} \gamma_i^2 \pi_i^2(\tau)} \tag{5.8}$$

and

$$c_{jl} = -\frac{\lambda \gamma_l^2 \pi_j(\tau)\pi_l(\tau)}{1 + \lambda \sum\limits_{i=0}^{j-1} \gamma_i^2 \pi_i^2(\tau)}. \tag{5.9}$$

Comparing the coefficients of x^{j-1} in both sides of Eq. (5.7) we have by virtue of Lemma 1.1

$$-\sum_{i=0}^{j-1} \alpha_i(\nu) = -\sum_{i=0}^{j-1} \alpha_i + c_{jj-1} = -\sum_{i=0}^{j-1} \alpha_i - \frac{\lambda \gamma_{j-1}^2 \pi_j(\tau)\pi_{j-1}(\tau)}{1 + \lambda \sum\limits_{i=0}^{j-1} \gamma_i^2 \pi_i^2(\tau)} \tag{5.10}$$

for $j \leq N$. Taking differences we immediately arrive at Eq. (5.3) for $j+1 \leq N$. The last coefficient $\alpha_N(\nu)$ can be calculated from Lemma 5.3. Applying the lemma to the measures σ and ν we have

$$\sum_{i=0}^{N-1} \alpha_i = \sum_{i=1}^{N} \tau_i$$

and

$$\sum_{i=0}^{N} \alpha_i(\nu) = \sum_{i=1}^{N} \tau_i + \tau$$

and consequently

$$\sum_{i=0}^{N} \alpha_i(\nu) = \sum_{i=0}^{N-1} \alpha_i + \tau.$$

Solving for $\alpha_N(\nu)$ and using Eq. (5.10) for $j = N$ we obtain (5.5).

In order to calculate the coefficients $\beta_j(\nu)$ we use Eq. (9). From Eq. (5.7) we have

$$\begin{aligned} \int p_j^2(x)\, d\nu(x) &= \int p_j^2(x)\, d\sigma(x) + \lambda p_j^2(\tau) \\ &= \int \pi_j^2(x)\, d\sigma(x) + \sum_{l=0}^{j-1} c_{jl}^2 \int \pi_l^2(x)\, d\sigma(x) + \lambda p_j^2(\tau). \end{aligned}$$

Since

$$\int \pi_j^2(x)\, d\sigma(x) = \gamma_j^{-2}$$

for $j < N$ and

$$\int \pi_N^2(x)\, d\sigma(x) = 0,$$

using (5.8) and (5.9) we get

$$\gamma_j(\nu)^{-2} = \int p_j^2(x)\, d\nu(x) = \gamma_j^{-2} + \frac{\lambda \pi_j^2(\tau)}{1 + \lambda \sum\limits_{i=0}^{j-1} \gamma_i^2 \pi_i^2(\tau)}$$

$$= \gamma_j^{-2} \frac{1 + \lambda \sum\limits_{i=0}^{j} \gamma_i^2 \pi_i^2(\tau)}{1 + \lambda \sum\limits_{i=0}^{j-1} \gamma_i^2 \pi_i^2(\tau)} \tag{5.11}$$

for $j < N$ and

$$\int p_N^2(x)\, d\nu(x) = \frac{\lambda \pi_N^2(\tau)}{1 + \lambda \sum\limits_{i=0}^{N-1} \gamma_i^2 \pi_i^2(\tau)}.$$

Now from (9) (for the measure ν) and (5) we obtain (5.4) and (5.6). $\square$

Using equations (5.3) – (5.6) together with the recurrence relations (2) and (5) it is possible to design a new algorithm which adds point masses one by one and computes the recurrence coefficients of a discrete measure starting from nothing.

Numerical stability of this algorithm obviously depends on the stability of the recurrence relations (2), as well. If we add points in ascending order, then any new point will be outside the interval of orthogonality of the measure constructed so far, and Lemma 5.7 will be applicable.

5.3 Inequalities II

If the measure σ and and its support S fulfil the conditions in Definition 1.3, then the numerator polynomials grow not essentially faster than the orthogonal polynomials on S. In order to estimate numerator polynomials, we need the division inequality [Jon94, Proposition 2].

Lemma 5.5 *Let $1 \leq p \leq \infty$, let $S \in \mathcal{M}^p(M, \alpha)$ and let q be a polynomial $\in \mathcal{P}_n$ of the form $q(x) = (x - x_0)\, q_1(x)$, where $x_0 \in \mathbf{R}$. Then it holds that*

$$\|q_1\|_p \leq 6Mn^\alpha \|q\|_p.$$

Lemma 5.6 *Let σ be a measure with $S(\sigma) \in \mathcal{M}^2(M, \alpha)$, and let π_n and κ_n be the monic orthogonal and numerator polynomials, respectively. Then we have*

$$\frac{\|\kappa_n\|_2^2}{\|\pi_n\|_2^2} \leq 72M^2 n^{2\alpha}.$$

Proof: For fixed x, we set $q(t) = \pi_n(x) - \pi_n(t) = (x - t)\, q_1(t)$. Then we have from (12)

$$|\kappa_n(x)| \leq \gamma_0^2 \|q_1\|_1 \leq \gamma_0 \|q_1\|_2 \leq 6Mn^\alpha \gamma_0 \|q\|_2$$

and (if $n > 0$)

$$\|q\|_2^2 = \gamma_0^{-2} \pi_n^2(x) + \gamma_n^{-2}.$$

Thus, we obtain

$$\kappa_n^2(x) \leq 36M^2 n^{2\alpha} [\pi_n^2(x) + \gamma_0^2 \gamma_n^{-2}]$$

and consequently

$$\|\kappa_n^2\|_2^2 \leq 72M^2 n^{2\alpha} \gamma_n^{-2} = 72M^2 n^{2\alpha} \|\pi_n^2\|_2^2. \square$$

The result, of course, is applicable only to discretizations of continuous measures (since finite sets trivially do not preserve Markov's property).

Outside the interval of orthogonality the numerator polynomials grow not faster than the orthogonal polynomials.

Lemma 5.7 *Let $\sigma = \sum_{i=1}^N \sigma_i\, \delta_{\tau_i}$ be a discrete measure with support $S(\sigma) = \{\tau_1, \ldots, \tau_N\} \subset [a, b]$ and let τ be a point with $\tau > b$. Then we have the estimate*

$$\frac{\kappa_n(\tau)}{\pi_n(\tau)} \leq \gamma_0^2 \sum_{i=1}^N \frac{\sigma_i}{\tau - \tau_i} \quad \textit{for} \quad n \leq N.$$

Proof: Since all zeros of π_n and κ_n are in the interval $[a, b]$ and the polynomials are monic, they are positive for $\tau > b$. Now from (13) the inequality

$$\frac{\kappa_n(\tau)}{\pi_n(\tau)} < \frac{\kappa_{n+1}(\tau)}{\pi_{n+1}(\tau)} \quad \text{for} \quad n + 1 \leq N$$

follows immediately. Thus we have

$$\frac{\kappa_n(\tau)}{\pi_n(\tau)} \leq \frac{\kappa_N(\tau)}{\pi_N(\tau)} = \gamma_0^2 \sum_{i=1}^N \frac{\sigma_i}{\tau - \tau_i}$$

by the definition of numerator polynomials, since $\pi_N(\tau_i) = 0$. $\square$

However, in isolated points of the support the behavior of orthogonal polynomials and numerator polynomials changes – the first show exponential decay, the latter exponential growth.

Lemma 5.8 *Let σ be a measure with bounded support $S(\sigma) \subset [a, b]$ and let τ be a point with dist$(\tau, S(\sigma)) \geq \delta$. Then there exists a number $c(\tau) > 1$ with*

$$\sum_{i=0}^n \gamma_i^2 \pi_i^2(\tau) \geq \frac{\gamma_0^2}{4}\, c(\tau)^{n-1}.$$

Proof: In terms of the orthonormal polynomials $\hat{\pi}_i$ the left-hand side of our inequality is

$$\sum_{i=0}^{n} \hat{\pi}_i^2(\tau).$$

But its reciprocal value is the Christoffel function

$$\left(\sum_{i=0}^{n} \hat{\pi}_i^2(\tau)\right)^{-1} = \min_{q\in\mathcal{P}_n} \frac{\int q^2(x)\,d\sigma(x)}{q^2(\tau)} \tag{5.12}$$

(see [Chi78, I-Th. 7.3] or [Nev79, p. 4]). Now we construct a polynomial $q \in \mathcal{P}_n$ with $|q| \leq 1$ on $S(\sigma)$. If we set

$$M(\tau) = \max\{(a-\tau)^2, (b-\tau)^2\},$$

then we have

$$\delta^2 \leq (x-\tau)^2 \leq M(\tau) \quad \text{for any} \quad x \in S(\sigma),$$

and from this follows

$$-1 \leq 2\,\frac{(x-\tau)^2 - \delta^2}{M(\tau) - \delta^2} - 1 \leq 1$$

on $S(\sigma)$. Consequently, the polynomial

$$q(x) = T_{[\frac{n}{2}]}\left(2\,\frac{(x-\tau)^2 - \delta^2}{M(\tau) - \delta^2} - 1\right)$$

has the desired properties. We see easily

$$\int q^2(x)\,d\sigma(x) \leq \int 1\,d\sigma(x) = \gamma_0^{-2} \tag{5.13}$$

and

$$q^2(\tau) = T^2_{[\frac{n}{2}]}\left(\frac{2\delta^2}{M(\tau) - \delta^2} + 1\right).$$

The explicit expression

$$T_m(x) = \frac{1}{2}\left[\left(x + \sqrt{x^2-1}\right)^m + \left(x - \sqrt{x^2-1}\right)^m\right]$$

(this is [Mei64, 4.13]) shows that

$$q^2(\tau) \geq \frac{1}{4}\,\left(x(\tau) + \sqrt{x(\tau)^2 - 1}\right)^{2[\frac{n}{2}]} \geq \frac{1}{4}\,\left(x(\tau) + \sqrt{x(\tau)^2 - 1}\right)^{n-1} \tag{5.14}$$

with

$$x(\tau) = \frac{2\delta^2}{M(\tau) - \delta^2} + 1 > 1.$$

From (5.12), (5.13) and (5.14) our proposition with

$$c(\tau) = x(\tau) + \sqrt{x(\tau)^2 - 1}$$

follows immediately. □

Lemma 5.9 *Let ν be a measure with bounded support $S(\nu) \subset [a,b]$ and let τ be a point with $\nu(\{\tau\}) = \lambda > 0$ and dist$(\tau, S(\nu) \setminus \{\tau\}) \geq \delta$. Let p_j be the monic orthogonal polynomials with respect to ν. Then there exists a number $c(\tau) > 1$ with*

$$\gamma_n(\nu)^2\, p_n^2(\tau) \leq \frac{1}{\lambda\left(1 + \frac{1}{4}\lambda\gamma_0(\nu)^2 c(\tau)^{n-2}\right)}.$$

Proof: Under our assumptions $\sigma = \nu - \lambda\delta_\tau$ is a measure with support $S(\sigma) = S(\nu)\setminus\{\tau\}$ and we have $\operatorname{dist}(\tau, S(\sigma)) \geq \delta$. Thus, the assumptions of the preceding lemma are fulfilled, and the monic orthogonal polynomials π_j with respect to σ satisfy the inequality in Lemma 5.8. Measure ν is obtained from measure σ by adding a point mass at τ of weight λ. Now from equations (5.8) and (5.11) we obtain easily

$$\begin{aligned}
\gamma_n(\nu)^2\, p_n^2(\tau) &= \gamma_n^2\, \frac{1 + \lambda \sum\limits_{i=0}^{n-1} \gamma_i^2\pi_i^2(\tau)}{1 + \lambda \sum\limits_{i=0}^{n} \gamma_i^2\pi_i^2(\tau)} \left[\frac{\pi_n(\tau)}{1 + \lambda \sum\limits_{i=0}^{n-1} \gamma_i^2\pi_i^2(\tau)}\right]^2 = \\
&= \frac{1}{\lambda}\left(\frac{1}{1 + \lambda \sum\limits_{i=0}^{n-1} \gamma_i^2\pi_i^2(\tau)} - \frac{1}{1 + \lambda \sum\limits_{i=0}^{n} \gamma_i^2\pi_i^2(\tau)}\right) \\
&\leq \frac{1}{\lambda\left(1 + \lambda \sum\limits_{i=0}^{n-1} \gamma_i^2\pi_i^2(\tau)\right)}.
\end{aligned}$$

Our proposition follows from Lemma 5.8, since

$$\gamma_0(\nu)^2 = \frac{\gamma_0^2}{1 + \lambda\gamma_0^2} \leq \gamma_0^2\,.\square$$

Lemma 5.10 *Let σ be a measure with bounded support $S(\sigma) \subset [a,b]$ and let τ be a point with $\sigma(\{\tau\}) = \lambda > 0$ and dist$(\tau, S(\sigma) \setminus \{\tau\}) \geq \delta$. Let κ_j be the monic numerator polynomials with respect to σ. Then there exists a number $c(\tau) > 1$ with*

$$\max\{\gamma_n^2\, \kappa_n^2(\tau), \gamma_{n+1}^2\, \kappa_{n+1}^2(\tau)\} \geq \frac{\lambda\gamma_0^4}{(b-a)^2}\left(1 + \frac{1}{4}\lambda\gamma_0^2 c(\tau)^{n-2}\right).$$

Proof: From equation (13) we have

$$\gamma_n\gamma_{n+1}[\pi_n(x)\kappa_{n+1}(x) - \pi_{n+1}(x)\kappa_n(x)] = \gamma_0^2\frac{\gamma_{n+1}}{\gamma_n} = \gamma_0^2\beta_{n+1}^{-1/2} > \frac{2\gamma_0^2}{b-a},$$

where we used the well-known inequality (11). Consequently, the inequality

$$\max\{|\gamma_n\gamma_{n+1}\pi_n(x)\kappa_{n+1}(x)|, |\gamma_n\gamma_{n+1}\pi_{n+1}(x)\kappa_n(x)|\} > \frac{\gamma_0^2}{b-a}$$

holds. Squaring both sides and using inequality (5.9) for the measure σ we obtain our proposition. $\Box$

The estimate of the condition of our problem given in Theorem 5.13 requires information about the size of orthogonal polynomials and their derivative in the nodes τ_i. But even if our discrete measure is the discretization of some continuous measure with support S preserving Markov's inequality, this may be not very helpful, since the nodes are not necessarily contained in S. For this reason we give some rough estimates based on the discretization itself.

Lemma 5.11 *Let $\sigma = \sum_{i=1}^N \sigma_i \, \delta_{\tau_i}$ be a discrete measure with $S(\sigma) \subset [a, b]$ and π_n the monic orthogonal polynomials with respect to σ. Then the following estimates hold:*

$$|\sigma_j \gamma_n^2 \pi_n^2(\tau_j)| \le 1,$$

$$|\sigma_j \gamma_n^2 \pi_n(\tau_j) \pi'_{n+1}(\tau_j)| \le 1 + \frac{b-a}{2} \sum_{i \neq j} \frac{1 + \sigma_i^{1/2} \sigma_j^{-1/2}}{|\tau_i - \tau_j|},$$

$$|\sigma_j \gamma_n^2 \pi_{n+1}(\tau_j) \pi'_n(\tau_j)| \le \frac{b-a}{2} \sum_{i \neq j} \frac{1 + \sigma_i^{1/2} \sigma_j^{-1/2}}{|\tau_i - \tau_j|},$$

$$|\sigma_j \gamma_n^2 \pi_n(\tau_j) \pi'_n(\tau_j)| \le \sum_{i \neq j} \frac{1 + \sigma_i^{1/2} \sigma_j^{-1/2}}{|\tau_i - \tau_j|}.$$

Proof: The first estimate follows easily from the obvious equation

$$\gamma_n^{-2} = \int \pi_n^2(x) \, d\sigma(x) = \sum_{i=1}^N \sigma_i \pi_n^2(\tau_i) \ge \sigma_j \pi_n^2(\tau_j).$$

Now we prove the second estimate: The polynomial

$$\frac{\pi_{n+1}(x) - \pi_{n+1}(\tau_j)}{x - \tau_j} \in \mathcal{P}_n$$

is monic and thus we have

$$\int \frac{\pi_{n+1}(x) - \pi_{n+1}(\tau_j)}{x - \tau_j} \, \pi_n(x) \, d\sigma(x) = \gamma_n^{-2}.$$

This equation can be written as

$$\sigma_j \pi'_{n+1}(\tau_j) \pi_n(\tau_j) + \sum_{i \neq j} \sigma_i \frac{\pi_{n+1}(\tau_i) - \pi_{n+1}(\tau_j)}{\tau_i - \tau_j} \, \pi_n(\tau_i) = \gamma_n^{-2}. \tag{5.15}$$

From our first estimate we obtain $|\gamma_n \pi_n(\tau_j)| \le \sigma_j^{-1/2}$ and

$$|\gamma_n \pi_{n+1}(\tau_j)| \le \frac{\gamma_n}{\gamma_{n+1}} \sigma_j^{-1/2} = \beta_n^{1/2} \sigma_j^{-1/2} \le \frac{b-a}{2} \sigma_j^{-1/2}$$

using (11). Substituting these inequalities into (5.15) we arrive at our proposition.

The last two inequalities can be shown analogously starting from

$$\int \frac{\pi_n(x) - \pi_n(\tau_j)}{x - \tau_j} \pi_n(x)\, d\sigma(x) = 0$$

and

$$\int \frac{\pi_n(x) - \pi_n(\tau_j)}{x - \tau_j} \pi_{n+1}(x)\, d\sigma(x) = 0.$$

□

5.4 Questions of numerical condition

In this section we investigate the sensitivity of the map $H_n : \gamma \to \rho$ from the nodes and weights of a discrete measure $\sigma = \sum_{i=1}^N \sigma_i\, \delta_{\tau_i}$ to the recurrence coefficients $\rho = \rho_n(\sigma)$.

Our analysis is based on Lemma 1.2 concerning dependence on a parameter. With its help, we are able to calculate the partial derivatives of the coefficients α_j and β_j with respect to a single node or weight.

Lemma 5.12 *Let σ be a measure not depending on λ and τ and let $\nu = \sigma + \lambda\delta_\tau$. We denote the monic orthogonal polynomials with respect to ν by p_j (as in the proof of Theorem 5.4). Then the following formulas hold:*

$$\frac{\partial \alpha_j(\nu)}{\partial \lambda} = \gamma_j^2(\nu) p_j(\tau) p_{j+1}(\tau) - \gamma_{j-1}^2(\nu) p_{j-1}(\tau) p_j(\tau), \tag{5.16}$$

$$\frac{\partial \beta_j(\nu)}{\partial \lambda} = \beta_j(\nu)[\gamma_j^2(\nu) p_j^2(\tau) - \gamma_{j-1}^2(\nu) p_{j-1}^2(\tau)], \tag{5.17}$$

$$\begin{aligned} \frac{\partial \alpha_j(\nu)}{\partial \tau} =& \lambda\gamma_j^2(\nu)[p_j'(\tau)p_{j+1}(\tau) + p_j(\tau)p_{j+1}'(\tau)] \\ &-\lambda\gamma_{j-1}^2(\nu)[p_{j-1}'(\tau)p_j(\tau) + p_{j-1}(\tau)p_j'(\tau)] \end{aligned} \tag{5.18}$$

and

$$\frac{\partial \beta_j(\nu)}{\partial \tau} = 2\lambda\beta_j(\nu)[\gamma_j^2(\nu)p_j'(\tau)p_j(\tau) - \gamma_{j-1}^2(\nu)p_{j-1}'(\tau)p_{j-1}(\tau)]. \tag{5.19}$$

Proof: The proposition is an almost trivial consequence of our Lemma 1.2. The linear functional l is defined via

$$l(q) = \int q(x)\, d\nu(x) = \int q(x)\, d\sigma(x) + \lambda\, q(\tau)$$

for any polynomial q, and obviously we have

$$\frac{\partial l}{\partial \lambda}(q) = q(\tau)$$

and

$$\frac{\partial l}{\partial \tau}(q) = \lambda\, q'(\tau).$$

□

Now we are going to define an appropriate condition number condH_n. As in [Gau82, Section 3.1] we consider each α_j and β_j individually as functions of one particular σ_i or τ_i. However, we feel that relative error is an adequate measure for β_j or σ_i only. In contrast to this, some node τ_i and even all coefficients α_j can be zero. Thus we base our condition number on relative error of β_j and σ_i and on absolute error of α_j and τ_i, i.e.

$$(\mathrm{cond}\,\alpha_j)(\sigma_i) = \left|\sigma_i \frac{\partial \alpha_j}{\partial \sigma_i}\right| \quad \text{for} \quad j = 0, \ldots, n-1, i = 1, \ldots, N,$$

$$(\mathrm{cond}\,\alpha_j)(\tau_i) = \left|\frac{\partial \alpha_j}{\partial \tau_i}\right| \quad \text{for} \quad j = 0, \ldots, n-1, i = 1, \ldots, N,$$

$$(\mathrm{cond}\,\beta_j)(\sigma_i) = \left|\beta_j^{-1} \sigma_i \frac{\partial \beta_j}{\partial \sigma_i}\right| \quad \text{for} \quad j = 1, \ldots, n-1, i = 1, \ldots, N$$

and

$$(\mathrm{cond}\,\beta_j)(\tau_i) = \left|\beta_j^{-1} \frac{\partial \beta_j}{\partial \tau_i}\right| \quad \text{for} \quad j = 1, \ldots, n-1, i = 1, \ldots, N.$$

The sensitivity of a single coefficient α_j or β_j to errors in the initial data can now be measured by

$$\mathrm{cond}\,\alpha_j = \sum_{i=1}^{N} [(\mathrm{cond}\,\alpha_j)(\sigma_i) + (\mathrm{cond}\,\alpha_j)(\tau_i)] \quad \text{for} \quad j = 0, \ldots, n-1$$

and

$$\mathrm{cond}\,\beta_j = \sum_{i=1}^{N} [(\mathrm{cond}\,\beta_j)(\sigma_i) + (\mathrm{cond}\,\beta_j)(\tau_i)] \quad \text{for} \quad j = 1, \ldots, n-1.$$

A suitable condition number is the maximum over all this quantities.

$$\mathrm{cond}\, H_n = \max\left\{\max_{j=0,\ldots,n-1} \mathrm{cond}\,\alpha_j, \max_{j=1,\ldots,n-1} \mathrm{cond}\,\beta_j\right\}.$$

The main4 result of this section follows immediately from the equations (5.16) – (5.19) above.

Theorem 5.13 *Let $\sigma = \sum_{i=1}^N \sigma_i\, \delta_{\tau_i}$ be a discrete measure, ρ_n the recurrence coefficients, γ_j^2 the normalization constants and π_j the monic orthogonal polynomials with respect to σ. Then we have*

$$\begin{aligned}\operatorname{cond} H_n = \max\Big\{ & \max_{j=0,\dots,n-1} \sum_{i=1}^N \sigma_i\gamma_j^2(|\pi_j(\tau_i)\pi_{j+1}(\tau_i) - \beta_j\pi_{j-1}(\tau_i)\pi_j(\tau_i)|+ \\ & |[\pi_j'(\tau_i)\pi_{j+1}(\tau_i) + \pi_j(\tau_i)\pi_{j+1}'(\tau_i)] - \beta_j[\pi_{j-1}'(\tau_i)\pi_j(\tau_i) + \pi_{j-1}(\tau_i)\pi_j'(\tau_i)]|), \\ & \max_{j=1,\dots,n-1} \sum_{i=1}^N \sigma_i\gamma_j^2(|\pi_j^2(\tau_i) - \beta_j\pi_{j-1}^2(\tau_i)| + 2|\pi_j'(\tau_i)\pi_j(\tau_i) - \beta_j\pi_{j-1}'(\tau_i)\pi_{j-1}(\tau_i)|)\Big\}.\end{aligned}$$

Proof: The proof is just by insertion of equations (5.16) – (5.19) into the above definitions (recall $\gamma_{j-1}^2 = \beta_j\gamma_j^2$).$\square$

However, these formulas are applicable only as a-priori bounds. If instead we use the results of some unstable algorithm, the obtained figures will show this instability, but not the conditioning of our problem. Some relevant a-priori estimates of the size of orthogonal polynomials and their derivatives are summarized in the Section 1.3. We can use them to obtain the following results.

Corollary 5.14 *Let $\sigma^{(N)} = \sum_{i=1}^N \sigma_i\delta_{\tau_i}$ be a discretization of degree of exactness M of some measure σ satisfying (1.10) with support $S(\sigma) \in \mathcal{M}^2(C,\alpha)$ (see Definition 1.3). Then the condition number of our problem can be estimated by*

$$\operatorname{cond} H_n \le C_1(C,\alpha)\, n^\alpha \quad \textit{for} \quad 2n \le M,$$

where $C_1(C,\alpha)$ denotes a constant depending on C and α.

Proof: We simply estimate the sums in Theorem 5.13 using Cauchy-Schwarz inequality. The sums not including derivatives are uniformly bounded, if we assume $S(\sigma) \subset [a,b]$:

$$\begin{aligned}&\sum_{i=1}^N \sigma_i\gamma_j^2(|\pi_j(\tau_i)\pi_{j+1}(\tau_i) - \beta_j\pi_{j-1}(\tau_i)\pi_j(\tau_i)| \le \sum_{i=1}^N \sigma_i\gamma_j^2|\pi_j(\tau_i)\pi_{j+1}(\tau_i)| + \\ &\sum_{i=1}^N \sigma_i\gamma_j^2\beta_j|\pi_{j-1}(\tau_i)\pi_j(\tau_i)| \le \gamma_j^2\left(\sqrt{\sum_{i=1}^N \sigma_i\pi_j^2(\tau_i)\sum_{i=1}^N \sigma_i\pi_{j+1}^2(\tau_i)} + \right. \\ &\left.\beta_j\sqrt{\sum_{i=1}^N \sigma_i\pi_{j-1}^2(\tau_i)\sum_{i=1}^N \sigma_i\pi_j^2(\tau_i)}\right) = \gamma_j^2(\|\pi_j\|_2\|\pi_{j+1}\|_2 + \beta_j\|\pi_{j-1}\|_2\|\pi_j\|_2) \\ &= \beta_{j+1}^{1/2} + \beta_j^{1/2} \le b-a,\end{aligned}$$

(for the last inequality we used the estimate (11), again) and

$$\sum_{i=1}^N \sigma_i\gamma_j^2(|\pi_j^2(\tau_i) - \beta_j\pi_{j-1}^2(\tau_i)| \le \gamma_j^2(\|\pi_j\|_2^2 + \beta_j\|\pi_{j-1}\|_2^2) = 2.$$

The sums containing derivatives can be similarly estimated using

$$\|\pi_j'\|_2 \leq Cj^\alpha \|\pi_j\|_2 = Cj^\alpha \gamma_j^{-1},$$

and we arrive at our proposition. □

Corollary 5.15 *Let $\sigma^{(N)} = \sum_{i=1}^N \sigma_i \delta_{\tau_i}$ be a discrete measure satisfying*

$$\sigma_i^{-1} \leq C_1 N^\alpha \quad \textit{and} \quad \frac{1}{|\tau_i - \tau_j|} \leq C_2 N^\beta \quad \textit{for} \quad i \neq j \in \{1, \ldots, N\}.$$

Then we have the estimate

$$\operatorname{cond} H_n \leq C_3(C_1, C_2, \alpha, \beta)\, N^{\alpha/2+\beta+1} \quad \textit{for} \quad n \leq N.$$

Proof: This follows immediately from Theorem 5.13 and the inequalities in Lemma 5.11, since under our assumptions we have

$$\frac{1 + \sigma_i^{1/2}\sigma_j^{-1/2}}{|\tau_i - \tau_j|} \leq (1 + \gamma_0^{-1} C_1^{1/2} N^{\alpha/2}) C_2 N^\beta$$

for $i \neq j$. □

5.5 Examples

In this section we compare the performance of the Stieltjes procedure (`stieltj`), the stabilized Lanczos process (`lancz1` — for technical reasons we used not the original procedure `lancz` from Gautschi's ORTHPOL-package (see [Gau94]), but an own implementation in Pascal based on the Algorithm RKPW in [GH84, p. 328]) and of the new algorithm (`points`) based on Equations (5.3) – (5.6) in a number of examples. In any of these examples we have more or less explicit formulas for the recurrence coefficients and thus we can compute the actual error magnification.

5.5.1 Chebyshev measure

We consider the measure

$$d\sigma(x) = \frac{dx}{\pi\sqrt{1-x^2}} \quad \text{on} \quad [-1, 1].$$

This, of course, is a continuous measure, but the Gauss-Christoffel quadrature formula is well-known in this case. If we set

$$\tau_i = \cos\frac{(2i-1)\pi}{2N} \quad \text{and} \quad \sigma_i = \frac{1}{N}$$

for $i = 1, \ldots, N$, then the discrete measure

$$\sigma^{(N)} = \sum_{i=1}^{N} \sigma_i \delta_{\tau_i}$$

is a discretization of degree of exactness $2N-1$ and consequently $\rho_n(\sigma) = \rho_n(\sigma^{(N)})$ for $n \leq N$. The coefficients ρ_n are known explicitly in this example. We have

$$\alpha_n \equiv 0$$

due to symmetry of the measure, and

$$\beta_n = \begin{cases} 1/2 & \text{for} \quad n = 1 \\ 1/4 & \text{for} \quad n \geq 2 \end{cases} \quad \text{and} \quad \gamma_n^2 = \begin{cases} 1 & \text{for} \quad n = 0 \\ 2^{2n-1} & \text{for} \quad n \geq 1 \end{cases}.$$

The orthogonal polynomials are well-known, too. The monic polynomials are

$$\pi_n = 2^{-n+1} T_n \quad \text{and} \quad \kappa_n = 2^{-n+1} U_{n-1},$$

where T_n and U_n denote Chebyshev polynomials of first and second kind. From this we can calculate easily

$$\operatorname{cond} H_n = O(n)$$

uniformly in N. Thus our problem is well-conditioned, but we have to estimate the quantities characterizing the stability of the algorithms `stieltj` and `points`. Since

$$E_n = \varepsilon^2 \gamma_n^2 \int \kappa_n^2(x)\, d\sigma(x) = 2\varepsilon^2 n,$$

the Stieltjes procedure is expected to be stable. The same is true for the new algorithm `points`, if points are added in ascending order: The stability is characterized by the bounds in Lemma 5.7, which grow only polynomially in N.

The results of numerical computations are shown in Table 5.1. The error magnification tabulated there is defined as the maximum of the absolute errors in the α_j and the relative errors in the β_j, divided by the machine eps (of order 1E-19). The figures confirm our expectations concerning the qualitative behaviour of the algorithms.

5.5.2 Bessis' example

Now, we return to our Example 2.2.1. Again, it is not a discrete measure, since the support $S(\sigma) = K$ is not finite. But a reasonable discretization is readily found.

Table 5.1: Error magnification in ρ_n for Example 5.5.1, $N = 50$

n	`points`	`lancz1`	`stieltj`
10	9.00e+00	6.00e+00	4.00e+00
20	1.40e+01	1.00e+01	1.00e+01
30	3.40e+01	3.20e+01	1.50e+01
40	3.40e+01	3.20e+01	2.80e+01
50	3.40e+01	4.00e+01	2.80e+01

Definition 5.16 *Let the polynomial T be defined by $T(x) = x^2 - \lambda$. For any set $S \subset \mathbf{R}$ we define*

$$T^{-1}(S) = \{x \in \mathbf{R} : T(x) \in S\}$$

and $T^{-k}(S)$ recursively via

$$T^{-k-1}(S) = T^{-1}(T^{-k}(S)).$$

The measure is characterized by the invariance property

$$\int f(x)\, d\sigma(x) = \int \frac{f(-\sqrt{x+\lambda}) + f(\sqrt{x+\lambda})}{2}\, d\sigma(x). \tag{5.20}$$

This property together with the normalization $\sigma([-\infty, \infty]) = 1$ uniquely defines σ. Indeed, for a constant function $f \equiv c$ we have

$$\int f(x)\, d\sigma(x) = \int c\, d\sigma(x) = c = f(0)$$

due to the normalization. If $f \in \mathcal{P}_n$ then

$$g(x) = \frac{f(-\sqrt{x+\lambda}) + f(\sqrt{x+\lambda})}{2}$$

defines a polynomial $g \in \mathcal{P}_{[n/2]}$. Consequently, equation (5.20) recursively defines integrals with respect to measure σ for polynomials f of any degree and thus for all continuous functions by continuity. Moreover, we see by induction that for $N = 2^k$ the set

$$S_N = \{\tau_1, \ldots, \tau_N\} = T^{-k}(\{0\})$$

of nodes together with the weights $\sigma_i = 1/N$ for $i = 1, \ldots, N$ give a discretization of degree of exactness $2N - 1$: If we set

$$\sigma^{(N)} = \frac{1}{N} \sum_{\tau \in S_N} \delta_\tau$$

and this is a discretization of degree of exactness M, then σ_{2N} has degree of exactness $\geq 2M+1$ by Equation (5.20). But $\sigma^{(1)} = \delta_0$ has degree of exactness 1, since

$$\int x \, d\sigma(x) = 0$$

due to symmetry of the measure.

The recurrence coefficients are not known explicitly, but can be calculated recursively from the equations

$$\beta_1 = \lambda, \quad \beta_{2k}\beta_{2k-1} = \beta_k \quad \text{and} \quad \beta_{2k} + \beta_{2k+1} = \lambda \quad \text{for} \quad k \geq 1 \tag{5.21}$$

mentioned already in Subsection 2.2.1. Of course, we have $\alpha_n \equiv 0$ due to symmetry. Thus we can compute the actual error magnification in our algorithms. To see what we have to expect we need some estimates of orthogonal polynomials, of their derivatives and of numerator polynomials on S_N.

On the support of the measure we have explicit bounds for the orthonormal polynomials $\hat{\pi}_n$. Bessis' estimate [Bes90, IV.141] is

$$|\hat{\pi}_n(x)| < \frac{1}{2}\left(\frac{4\sqrt{\lambda}}{\sqrt{1+4\lambda}-1}\right)^k \quad \text{for} \quad 2^{k-1} \leq n \leq 2^k,$$

where $x \in K$. This seems to be of little use, since obviously $S_N \cap K = \emptyset$. Fortunately, a careful analysis of the proof shows that estimates of the form

$$|\hat{\pi}_n(x)| \leq C(\lambda)^k \quad \text{for} \quad 2^{k-1} \leq n \leq 2^k$$

are valid not only for $x \in K$, but also for $x \in T^{-k}([-\xi, \xi])$, and this set contains $T^{-k}(\{0\}) = S_N$ for $N = 2^k$. The same is true for the polynomials $\hat{\pi}_n'(x)$ and $\kappa_n(x)$. We will make this more precise now.

Proposition 5.17 *Let $\hat{\pi}_j(x)$ be the orthonormal polynomials with respect to Bessis measure σ. Let $\xi = (1+\sqrt{1+4\lambda})/2$ and $\Delta = [-\xi, \xi]$. Then we have*

$$|\hat{\pi}_j(x)| \leq \frac{\xi}{\sqrt{\lambda}} C(\lambda)^k \quad \textit{for} \quad j \leq 2^k + 1, \; x \in T^{-k}(\Delta),$$

where the constant $C(\lambda)$ can be chosen as

$$C(\lambda) = \frac{1}{\sqrt{\lambda - \xi}}\left(1 + \frac{\lambda}{\sqrt{\lambda} - 1}\right).$$

The derivative of the orthonormal polynomials satisfies the inequality

$$|\hat{\pi}_j'(x)| \leq \left[\frac{2(\xi+\lambda)}{\sqrt{\lambda}} + \frac{\xi}{\sqrt{\lambda(\lambda-\xi)}}\right] C_1(\lambda)^k \quad \textit{for} \quad j \leq 2^k + 1, \; x \in T^{-k}(\Delta),$$

where

$$C_1(\lambda) = \max\{2C(\lambda), 2\xi\}.$$

Proof: The invariance property of our measure can be expressed in the terms of shift and symmetrization: σ coincides with the symmetrization of the measure ν obtained by a shift of σ (by λ). Again, we denote by $\pi_j(x)$ the monic orthogonal polynomials with respect to σ and by $p_j(x)$ the monic orthogonal polynomials with respect to ν. Then we have $p_j(x) = \pi_j(x-\lambda)$, $\pi_{2j}(x) = p_j(x^2) = \pi_j(x^2-\lambda)$ and $\pi_{2j+1}(x) = xp_j^*(0; x^2)$, where

$$\begin{aligned} p_j(0;x) &= x^{-1}\left[p_{j+1}(x) - \frac{p_{j+1}(0)}{p_j(0)}p_j(x)\right] \\ &= x^{-1}\left[\pi_{j+1}(x-\lambda) - \frac{\pi_{j+1}(-\lambda)}{\pi_j(-\lambda)}\pi_j(x-\lambda)\right]. \end{aligned}$$

Thus, the monic orthogonal polynomials satisfy the equations

$$\begin{aligned} \pi_{2j}(x) &= \pi_j(x^2-\lambda) \\ \pi_{2j+1}(x) &= x^{-1}\left[\pi_{j+1}(x^2-\lambda) - \frac{\pi_{j+1}(-\lambda)}{\pi_j(-\lambda)}\pi_j(x^2-\lambda)\right]. \end{aligned}$$

The orthonormal polynomials can be obtained from $\hat{\pi}_j(x) = \gamma_j\pi_j(x)$. The leading coefficient γ_j satisfies the recursion

$$\frac{\gamma_{j-1}^2}{\gamma_j^2} = \beta_j,$$

and this (taking into account Equations (5.21)) implies the equations

$$\gamma_{2j}^2 = \gamma_j^2 \quad \text{and} \quad \gamma_{2j+1}^2 = \beta_{2j+1}^{-1}\gamma_j^2 = \beta_{2j+2}\gamma_{j+1}^2.$$

Thus for the orthonormal polynomials we have

$$\begin{aligned} \hat{\pi}_{2j}(x) &= \hat{\pi}_j(x^2-\lambda) \\ \hat{\pi}_{2j+1}(x) &= x^{-1}\left[\beta_{2j+2}^{1/2}\hat{\pi}_{j+1}(x^2-\lambda) - \frac{\pi_{j+1}(-\lambda)}{\pi_j(-\lambda)}\beta_{2j+1}^{-1/2}\hat{\pi}_j(x^2-\lambda)\right]. \end{aligned}$$

Now due to symmetry we have

$$-\frac{\pi_{j+1}(-\lambda)}{\pi_j(-\lambda)} = \frac{\pi_{j+1}(\lambda)}{\pi_j(\lambda)},$$

and since all zeros of $\pi_j(x)$ are in Δ, the monic polynomials $\pi_j(x)$ must be > 0 for $x \geq \lambda > \xi$. Consequently, the recurrence (2) implies

$$0 < \frac{\pi_{j+1}(\lambda)}{\pi_j(\lambda)} < \lambda.$$

Since $\beta_{2j+2} \leq 1$ and $\beta_{2j+1} \geq \lambda - 1$ (see [Bes90, (IV.128) and (IV.129)]) and $|x| \geq \sqrt{\lambda - \xi}$ for $x \in T^{-k}(\Delta) \subset T^{-1}(\Delta)$, from the above equation follows

$$|\hat{\pi}_{2j+1}(x)| \leq \frac{1}{\sqrt{\lambda-\xi}} \left[|\hat{\pi}_{j+1}(x^2-\lambda)| + \frac{\lambda}{\sqrt{\lambda-1}} |\hat{\pi}_j(x^2-\lambda)| \right].$$

But $x \in T^{-k-1}(\Delta)$ is equivalent to $x^2 - \lambda \in T^{-k}(\Delta)$, and $2j+1 \leq 2^{k+1}+1$ is equivalent to $j+1 \leq 2^k+1$. If we set

$$M_k = \max_{j \leq 2^k+1, x \in T^{-k}(\Delta)} |\hat{\pi}_j(x)|,$$

we obtain $M_{k+1} \leq C(\lambda) M_k$. The inequality $M_k \leq C(\lambda)^k M_0$ follows immediately from this estimate, and $M_0 = \frac{\xi}{\sqrt{\lambda}}$ is obvious.

The proof of the second bound is similar. The derivatives $\hat{\pi}'_j(x)$ fulfill the equations

$$\begin{aligned} \hat{\pi}'_{2j}(x) &= 2x \hat{\pi}'_j(x^2-\lambda) \\ x\hat{\pi}'_{2j+1}(x) + \hat{\pi}'_{2j+1}(x) &= 2x \left[\beta_{2j+2}^{1/2} \hat{\pi}'_{j+1}(x^2-\lambda) - \frac{\pi_{j+1}(-\lambda)}{\pi_j(-\lambda)} \beta_{2j+1}^{-1/2} \hat{\pi}'_j(x^2-\lambda) \right]. \end{aligned}$$

With

$$M'_k = \max_{j \leq 2^k+1, x \in T^{-k}(\Delta)} |\hat{\pi}'_j(x)|,$$

this gives the estimate

$$M'_{k+1} \leq C_1(\lambda) M'_k + \frac{1}{\sqrt{\lambda-\xi}} M_k,$$

and this implies our proposition. □

From the bounds for the orthogonal polynomials and their derivatives we can obtain bounds for the (normalized) polynomials of second kind:

Corollary 5.18 *Let*

$$\hat{\kappa}_j(x) = \gamma_0^2 \int \frac{\hat{\pi}_j(x) - \hat{\pi}_j(t)}{x-t}\, d\sigma(t).$$

There exists a constant $C_2(\lambda)$ *with*

$$|\hat{\kappa}_j(x)| \leq C_2(\lambda)^k \quad \textit{for} \quad j \leq 2^k+1,\ x \in T^{-k}(\Delta).$$

Proof: The set $T^{-k}(\Delta)$ consists of 2^k components with σ-measure 2^{-k}. If x and t are in different components, we have $|x-t| \geq 2\sqrt{\lambda-\xi}\ (2\xi)^{-k}$, otherwise

$$\left| \frac{\hat{\pi}_j(x) - \hat{\pi}_j(t)}{x-t} \right| \leq \left[\frac{2(\xi+\lambda)}{\sqrt{\lambda}} + \frac{\xi}{\sqrt{\lambda(\lambda-\xi)}} \right] C_1(\lambda)^k.$$

Table 5.2: Error magnification in ρ_n for Example 2.2.1, $N = 128$

n	`points`	`lancz1`	`stieltj`
10	1.00e+01	1.70e+01	6.00e+00
20	1.40e+01	1.70e+01	8.00e+00
30	2.40e+01	1.70e+01	1.60e+01
40	2.60e+01	1.80e+01	3.30e+01
50	3.80e+01	4.70e+01	3.40e+01
128	9.70e+01	1.09e+02	8.20e+01

Since $\gamma_0 = 1$, we have

$$|\hat{\kappa}_j(x)| \leq 2^{-k}\left[\frac{2(\xi+\lambda)}{\sqrt{\lambda}} + \frac{\xi}{\sqrt{\lambda(\lambda-\xi)}}\right] C_1(\lambda)^k + \frac{1}{2\sqrt{\lambda-\xi}}(2\xi)^k \cdot 2M_k$$

for $j \leq 2^k + 1,\ x \in T^{-k}(\Delta)$, and this proves our proposition. □ Consequently, the condition of our problem can be bounded by a polynomial in n and N. This fact we could derive from Corollary 5.15, too: For $N = 2^k$ and $S_N = \{\tau_1, \ldots, \tau_N\}$ we have $\sigma_i^{-1} = N$ and

$$|\tau_i - \tau_j| \geq C \cdot (2\xi)^{-k} \quad \text{for} \quad i \neq j. \tag{5.22}$$

This follows from the definition of S_N and the elementary inequality

$$|\sqrt{\lambda + x} - \sqrt{\lambda + y}| \geq \frac{|x-y|}{2\sqrt{\lambda+\xi}} = \frac{|x-y|}{2\xi} \quad \text{for} \quad x, y \in [-\xi, \xi].$$

The estimate (5.22) together with Lemma 5.7 shows that procedure `points` should be stable. The same is true for Stieltjes procedure `stieltj`. The numerical results confirm the expected behaviour of the algorithms.

5.5.3 An atomic measure with a dense set of mass points

In the preceding example we assumed $\lambda > 2$ and obtained a singular measure concentrated on a Cantor-type set K of Lebesgue measure 0. For the "degenerate" case $\lambda = 2$ we would obtain an absolutely continuous measure on $[-2, 2]$ — in fact this is simply a Chebyshev measure scaled to this interval. If we modify the defining invariance property (5.20), then we have a more interesting example: Let the measure σ with support $S(\sigma) = [-2, 2]$ be defined by

$$\int f(x)\, d\sigma(x) = \int \frac{f(-\sqrt{x+2}) + f(0) + f(\sqrt{x+2})}{3}\, d\sigma(x). \tag{5.23}$$

and the normalization $\int 1\, d\sigma(x) = \sigma([-2,2]) = \gamma_0^{-2} = 1$. Clearly, this measure puts mass 3^{-k-1} on the points $2\cos(2m-1)2^{-k-1}\pi$ for $k \geq 0$ and $m = 1, 2, \ldots, 2^k$. This is an infinite set of points dense in $[-2,2]$. Again, it is easy to find a suitable discretization of the measure. The measure $\sigma^{(1)} = \delta_0$ is a discretization of degree of exactness 1. If we are given a discretization $\sigma^{(N)}$ of degree of exactness M, then we can construct a new discretization

$$\sigma^{(2N+1)} = \frac{1}{3}\,\delta_0 + \sum_{\tau^2-2\in S_N} \frac{1}{3}\,\sigma^{(N)}(\{\tau^2-2\})\,\delta_\tau$$

with degree of exactness $2M+1$ using the invariance (5.23). Thus by induction we obtain discretizations $\sigma^{(N)}$ of degree of exactness N with $N = 2^k - 1$.

The recurrence coefficients ρ_n can be calculated from a system of equations similar to (though more complicated than) the equations in the preceding example:

Proposition 5.19 *Let ν be measure σ without the point mass at 0, i.e. $\nu(B) = \sigma(B \setminus \{0\})$ for any measurable set B on the real line. We have $\alpha_n(\sigma) \equiv 0$ due to symmetry of the measure. The coefficients $\beta_n(\sigma)$ can be calculated from*

$$\beta_1(\nu) = 2, \quad \beta_{2n}(\nu) + \beta_{2n+1}(\nu) = 2$$

and

$$\beta_{2n}(\nu)\beta_{2n-1}(\nu) = q_n\beta_n(\nu) = \beta_n(\sigma)$$

for $n \geq 1$, where

$$q_n = \begin{cases} \frac{2}{3} & \text{for} \quad n = 1 \\ \frac{1+\frac{1}{2}(1+s_k)}{1+\frac{1}{2}(1+s_{k-1})} & \text{for} \quad n = 2k, \ k \geq 1 \\ \frac{1}{q_{n-1}} & \text{for} \quad n = 2k+1, \ k \geq 1 \end{cases}$$

and

$$s_k = \sum_{l=1}^{k} \frac{\beta_1(\nu)\ldots\beta_{2l-1}(\nu)}{\beta_2(\nu)\ldots\beta_{2l}(\nu)}$$

for $k \geq 1$.

Proof: Let σ be our measure, σ_2 the shifted by 2 measure σ, and ν the measure σ restricted to $\mathbf{R} \setminus \{0\}$. The symmetrization of σ_2 (up to a factor) coincides with ν, and from the above formulas we have

$$\beta_{2j}(\nu) + \beta_{2j+1}(\nu) = \alpha_j(\sigma_2) \equiv 2$$

and

$$\beta_{2j-1}(\nu)\beta_{2j}(\nu) = \beta_j(\sigma_2) = \beta_j(\sigma).$$

Since $\nu + \frac{1}{3}\delta_0 = \sigma$, using the formulas for adding a point mass to a measure we obtain our proposition (see Theorem 5.4, which is just Nevai's Lemma 7.15 in [Nev79] in terms of monic polynomials): If we denote the monic orthogonal polynomials with respect to measure ν by p_j, then we have

$$\gamma_j(\sigma)^{-2} = \gamma_j(\nu)^{-2} \frac{1 + \frac{1}{3}\sum\limits_{i=0}^{j} \gamma_i(\nu)^2 p_i^2(0)}{1 + \frac{1}{3}\sum\limits_{i=0}^{j-1} \gamma_i(\nu)^2 p_i^2(0)} = c_j \gamma_j(\nu)^{-2}.$$

Since $\alpha_j(\nu) \equiv 0$ by symmetry, the recurrence relation gives

$$p_{j+1}(0) = -\beta_j(\nu) p_{j-1}(0),$$

thus $p_j(0) = 0$ for odd j, and

$$p_j(0) = (-1)^k \prod_{l=1}^{k} \beta_{2l-1}(\nu) \quad \text{for} \quad j = 2k.$$

But we have

$$\gamma_i(\nu)^2 = \frac{\gamma_0(\nu)^2}{\prod\limits_{l=1}^{i} \beta_l(\nu)} \quad \text{and} \quad \gamma_0(\nu)^2 = \frac{3}{2},$$

from this we obtain (using the notations of Example 5.5.3)

$$c_j = \begin{cases} \dfrac{3}{2} & \text{for} \quad j = 0 \\ \dfrac{1 + \frac{1}{2}(1 + s_k)}{1 + \frac{1}{2}(1 + s_{k-1})} & \text{for} \quad j = 2k, \ k \geq 1 \\ 1 & \text{for} \quad j = 2k+1, \ k \geq 0 \end{cases}$$

Now we have to observe

$$\beta_j(\sigma) = \frac{\gamma_j(\sigma)^{-2}}{\gamma_{j-1}(\sigma)^{-2}} = \frac{c_j}{c_{j-1}} \cdot \frac{\gamma_j(\nu)^2}{\gamma_{j-1}(\nu)^{-2}} = \frac{c_j}{c_{j-1}} \beta_j(\nu),$$

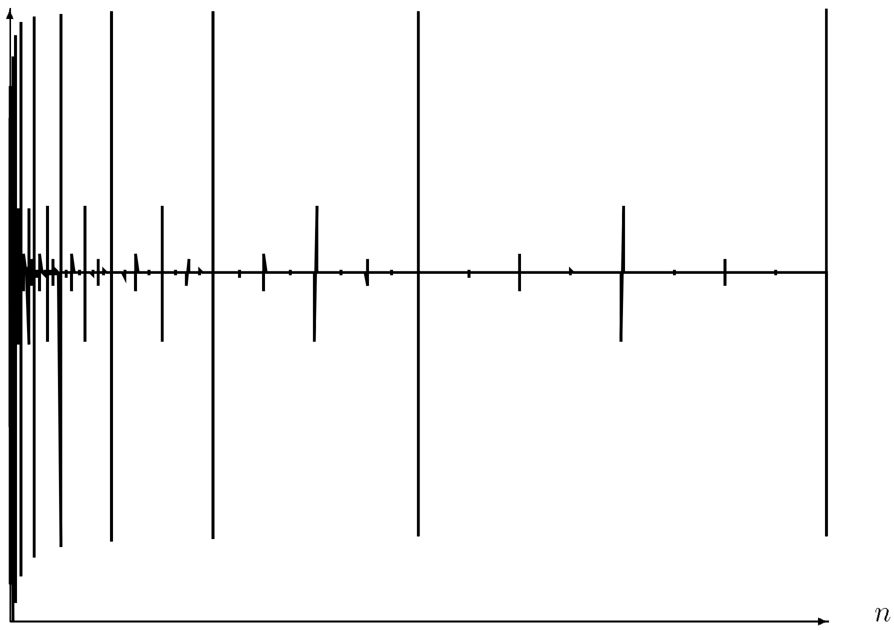
i.e. $q_j = c_j / c_{j-1}$, and we finally arrive at our proposition.□

Estimates for derivatives of orthogonal polynomials are readily available in this example, since $S = [-2, 2]$ preserves Markov's inequality: In the notations of the Section 1.3 we have $S \in \mathcal{M}^\infty(\frac{1}{2}, 2)$, and obviously our measure σ satisfies condition (1.10) with $s = \ln 3 / \ln 2$. Consequently, S preserves Markov's inequality in $L^2(\sigma)$, i.e. $S \in \mathcal{M}^2(M, 2+s)$, due to [Jon94, Proposition 1]. Thus, our problem is well-conditioned due to Corollary 5.14, and procedure `stieltj` is stable in virtue of Lemma 5.6. The assumptions of Corollary 5.15 are satisfied as well (with $\alpha = \ln 3 / \ln 2$ and $\beta = 2$), and procedure `points` is stable due to Lemma 5.7. The numerical results in Table 5.3 agree with these expectations (observe

Table 5.3: Error magnification in ρ_n for Example 5.5.3, $N = 127$

n	`points`	`lancz1`	`stieltj`
10	2.80e+01	2.20e+01	3.00e+01
20	4.80e+01	4.60e+01	3.00e+01
30	4.80e+01	4.60e+01	3.00e+01
40	5.50e+01	4.60e+01	3.00e+01
50	5.50e+01	4.60e+01	3.00e+01

that the degree of exactness of our discretization $\sigma^{(N)}$ is only N for $N = 2^k - 1$ and thus we can calculate only recursion coefficients ρ_n with $2n - 1 \leq N$). As the following display of the first 2048 coefficients suggests, we have only convergence in density in this example, but no convergence in the usual sense:



5.5.4 Chebyshev measure plus point mass

In this example we use the measure $\sigma = \sigma_c + \delta_2$, where measure σ_c is defined by

$$d\sigma_c(x) = \frac{dx}{\pi\sqrt{1 - x^2}} \quad \text{on} \quad [-1, 1],$$

i.e. we add a point mass of weight 1 at $t = 2$ to the measure from Example 1. The recurrence coefficients can be calculated easily from the formulas of Section 5.2. We obtain $\gamma_0^2 = 1/2$ and

$$\gamma_j^2 = 2^{2j-1} \frac{U_{2j-2}(2) + 2j + 1}{U_{2j}(2) + 2j + 3} \quad \text{for} \quad j \geq 1.$$

The coefficients β_j can be computed from Equation (5), and the coefficients α_j from the equation

$$\sum_{i=0}^{j} \alpha_i = \frac{2T_j(2)T_{j+1}(2)}{U_{2j}(2) + 2j + 3}.$$

A discretization of the measure can be obtained trivially from the discretization $\sigma^{(N-1)}$ in Example 1 by adding $\sigma_N = 1$ and $\tau_N = 2$ (and consequently the degree of exactness of $\sigma^{(N)}$ is only $2N - 3$, i.e. we can calculate ρ_n for $n \le N - 1$). The assumptions of Corollary 5.15 are satisfied, and thus our problem is well-conditioned. The procedure `points` should be stable by Lemma 5.7. In the Stieltjes procedure `stieltj`, however, we can expect instability: The inequalities in Lemma 5.9 and 5.10 show that the ratio $\kappa_j(2)/\pi_j(2)$ grows exponentially (in fact like $(2 + \sqrt{3})^{2j} = (7 + 4\sqrt{3})^j$). The data in Table 5.4 indeed show this behaviour. We mention here that instability of Stieltjes procedure in this situation was reported by Gautschi (compare [Gau82, Example 4.8]).

5.5.5 Discrete Chebyshev polynomials

These polynomials are orthogonal with respect to the discrete measure

$$\sigma^{(N)} = \frac{1}{N} \sum_{i=1}^{N} \delta_{i-1}.$$

The recurrence coefficients are known explicitly: We have

$$\alpha_k = \frac{N-1}{2} \quad \text{for} \quad k = 0, 1, \ldots, N-1$$

and

$$\beta_k = \frac{1 - (k/N)^2}{4(1 - 1/k^2)} \quad \text{for} \quad k = 1, \ldots, N-1.$$

Table 5.4: Error magnification in ρ_n for Example 5.5.4, $N = 80$

n	`points`	`lancz1`	`stieltj`
10	5.20e+01	4.80e+01	5.53e+01
20	1.02e+02	1.04e+02	1.08e+02
30	1.56e+02	1.58e+02	1.05e+13
40	2.32e+02	2.28e+02	4.45e+19
50	2.32e+02	2.28e+02	4.45e+19
79	4.66e+02	4.76e+02	5.13e+19

Table 5.5: Error magnification in ρ_n for Example 5.5.5, $N = 80$

n	`points`	`lancz1`	`stieltj`
60	2.80e+01	4.70e+01	4.00e+01
65	2.90e+01	6.00e+01	9.44e+05
70	4.70e+01	6.00e+01	1.49e+11
75	8.00e+01	6.00e+01	3.36e+17
80	1.24e+02	6.00e+01	5.05e+20

The values of the *orthonormal* polynomials $\hat{\pi}_n(x)$ at $x = 0$ can be evaluated easily using the explicit formulas in [AS84, 22.19]. After some calculation we obtain

$$\hat{\pi}_n(0) = (-1)^n (N-1)! \sqrt{\frac{(2n+1)N}{(N+n)!(N-n-1)!}} \quad \text{for} \quad n = 0, \ldots, N-1.$$

Now from Equation (13) follows

$$\frac{\kappa_{n+1}(0)}{\pi_{n+1}(0)} - \frac{\kappa_n(0)}{\pi_n(0)} = \frac{1}{\gamma_n^2 \pi_n(0) \pi_{n+1}(0)} = \frac{1}{\beta_{n+1}^{1/2} \hat{\pi}_n(0) \hat{\pi}_{n+1}(0)},$$

since we have $\gamma_0^2 = 1$. This gives

$$\frac{\kappa_n(0)}{\pi_n(0)} = \sum_{l=1}^{n} \frac{1}{\beta_l^{1/2} \hat{\pi}_{l-1}(0) \hat{\pi}_l(0)} = -2 \sum_{l=1}^{n} \frac{(N+l-1)!(N-l-1)!}{(N-1)!^2},$$

and this quantity grows very rapidly as n approaches N (for $n = N-1$ its absolute value is of order 2^{2N}). Consequently, we can expect instable behaviour of the Stieltjes procedure. The problem is well-conditioned, however, and the new algorithm `points` should be stable, as the estimates in Lemma 5.7 and Corollary 5.15 show.

The numerical results reported in Table 5.5 confirm our expectations (the instability of the Stieltjes procedure in this case was observed already by Gautschi, see [Gau82, Example 4.1]).

5.6 Bibliographical references

The problem of numerical condition has an interesting history: In his paper [Gau78] Gautschi explained that the inverse map H_n^{-1} is essentially an eigenvalue problem for the symmetric positive definite Jacobi matrix and thus should be well-conditioned. However, no estimates for condition numbers were given. In his article [Gau82] Gautschi gave estimates for the condition number which (if used as a-posteriori bounds!) seemed to show bad conditioning in some cases.

Unfortunately, his formulas were partially erroneous, and Gautschi corrected his statement in [Gau86, Note added in proof]. But for all we know no correct estimates of the sensitivity of the map H_n appeared in the literature.

In fact, a-posteriori bounds can be completely wrong if the algorithm is unstable. As Gautschi observed in his paper [Gau93], the behavior of second kind polynomials is crucial for the stability analysis of the recurrence relation (2). In this chapter, we gave some a-priori estimates of the growth of numerator polynomials. Their exponential growth in isolated mass points (versus expontial decay of the orthogonal polynomials) is the origin if instability of all Stieltjes-type algorithms in this case (compare also Example 6.2.5 in Chapter 6). Since not always we have a discretization of a continuous measure, we derived some (rough) bounds in terms of the nodes and weights of the quadrature data. An alternative approach to such bounds can be found in the preprint of Beckermann and Bourreau [BB97].

However, stable techniques are known, which are usually called Lanczos algorithm. As Gragg and Harrod explain in their article [GH84], this is not correct: The original Lanczos process is manifestly numerically unstable. The stabilization of the Lanczos process in the Hermitian case is due to Rutishauser (for the indefinite case, see the article of Reichel [Rei93]). For comparison, we implemented Algorithm RKPW of [GH84]. We did not use Gautschi's algorithm `lancz` for technical reasons (lack of a Fortran compiler). Moreover, we designed an own algorithm `points` using Nevai's formulas for adding point masses (generalized to the case of finite support). The algorithm works only with quantities related to orthogonal polynomials and is stable, if points are added in ascending order.

The results of this chapter were published in the recent paper [Fis98b].

Chapter 6

Self-similar measures

However, our measure σ can be defined by some invariance property of the measure as well. This means that an equation like

$$\int h(x)\,d\sigma(x) = \int Ah(x)\,d\sigma(x) \quad \text{for} \quad h \in \mathbf{H} \tag{6.1}$$

with some linear function space $\mathbf{H}$ and some linear operator $A : \mathbf{H} \to \mathbf{H}$ is satisfied. This kind of equation can be obtained from integration by parts, if the weight function of the measure satisfies some linear differential equation with polynomial coefficients. But it can be derived from a self-similarity of the measure, too. Independently of its origin, from the invariance (6.1) sometimes we can obtain recurrence relations between the integrals in (8) and (9). This will be made more precise in the next section.

6.1 Definition of self-similarity

Now we will specify some types of operators A, for which the invariance (6.1) together with results like (10) lead to recurrence relations for the integrals $\int \pi_j^2(x)\,d\sigma(x)$ and $\int x\pi_j^2(x)\,d\sigma(x)$. For classical measures, the weights of which satisfy simple differential equations, the operator A will be some differential operator. Indeed, this leads to recursion formulas for α_j and β_j, as was shown in the literature long ago.

We will consider here two other types of operators closely related to *fractal interpolation functions* (see Barnsley and Harrington [BH89]).

6.1.1 Fractal interpolation functions

In their paper [BH89], Barnsley and Harrington give the following definition of fractal interpolation functions:

Definition 6.1 *Let $x_0 < x_1 < \ldots < x_N$. Let L_i be the affine map satisfying $L_i(x_0) = x_{i-1}$, $L_i(x_N) = x_i$, $i = 1, 2, \ldots, N$. Let $y_0, y_1, \ldots, y_N \in \mathbf{R}$. Let $-1 < \lambda_i < 1$, $i = 1, 2, \ldots, N$. Let $X = [x_0, x_N]$. Let $F_i : X \times \mathbf{R} \to \mathbf{R}$ satisfy*

$$|F_i(x, y_1) - F_i(x, y_2)| \leq |\lambda_i||y_1 - y_2|, \quad x \in X, y_1, y_2 \in \mathbf{R}. \tag{6.2}$$

$$F_i(x_0, y_0) = y_{i-1}, \quad F_i(x_N, y_N) = y_i, \quad i = 1, 2, \ldots, N. \tag{6.3}$$

The fractal interpolation function, abbreviated FIF, associated with

$$\{L_i(x), F_i(x, y)\}_{i=1}^N$$

is the unique function $f : X \to \mathbf{R}$ satisfying

$$f(L_i(x)) = F_i(x, f(x)), \quad i = 1, 2, \ldots, N, x \in X \tag{6.4}$$

As in [BH89], we deal with FIFs where

$$F_i(x, y) = \lambda_i y + q_i(x). \tag{6.5}$$

Thus, we assume that the cumulative distribution function of our measure σ (which for simplicity we will denote by $\sigma(x)$) is a FIF.

Lemma 6.2 *Let the cumulative distribution function of our measure σ be a FIF defined in Definition 6.1 and let F_i be defined by (6.5). Then σ satisfies the invariance property (6.1), where the operator A is defined via*

$$Ah(x) = \sum_{i=1}^N \Big(\lambda_i h(L_i(x)) + \gamma_0^2 \int h(L_i(y))\, dq_i(y) \Big) \quad \textit{for} \quad h \in \mathbf{H} \tag{6.6}$$

and

$$\mathbf{H} = \{h \in L^1(\sigma) : h(L_i(\cdot)) \in L^1(q_i) \quad \textit{for} \quad i = 1, \ldots, N\}.$$

Proof: From the definitions (with σ instead of f) we obtain

$$\int h(x)\, d\sigma(x) = \sum_{i=1}^N \int_{[x_{i-1}, x_i]} h(x)\, d\sigma(x) = \sum_{i=1}^N \int h(L_i(x))\, d\sigma(L_i(x))$$
$$= \sum_{i=1}^N \Big(\lambda_i \int h(L_i(x))\, d\sigma(x) + \int h(L_i(x))\, dq_i(x) \Big),$$

and this can be written as

$$\int h(x)\, d\sigma(x) = \int \sum_{i=1}^N \Big(\lambda_i h(L_i(x)) + \gamma_0^2 \int h(L_i(y))\, dq_i(y) \Big)\, d\sigma(x)$$

for $h \in \mathbf{H}$. This, however, is an equation of type (6.1) with an operator A defined by (6.6).□

Remark: From Definition 6.1 follows that the linear functionals L_i are of the form

$$L_i(x) = a_i x + b_i \tag{6.7}$$

with $a_i > 0$. This guarantees the continuity of the FIF, i.e. the absence of atomic parts in our measure. If we drop the assumption $a_i > 0$, our equations still make sense. Simply our distribution function will no longer be continuous, and our measure will be atomic. In the degenerate case, where $a_i = 0$ and $q_i \equiv 0$ for $i = 1, \ldots, N$, we obviously get a discrete measure with weight $\gamma_0^{-2}\lambda_i$ on $\{b_i\}$.

Obviously, the operator A converts polynomials to polynomials of the same degree, i.e. if $h \in \mathcal{P}_D$ then $Ah \in \mathcal{P}_D$, too.

Lemma 6.3 *There must be a representation*

$$\pi_j(L_i(x)) = \sum_{k=0}^{j} c_{jk}^{(i)} \pi_k(x) \tag{6.8}$$

with appropriate real coefficients $c_{jk}^{(i)}$. They can be determined recursively from their initial values $c_{00}^{(i)} = 1$ via the relation

$$c_{j+1,k}^{(i)} = a_i c_{j,k-1}^{(i)} + (a_i\alpha_k + b_i - \alpha_j) c_{jk}^{(i)} + a_i \beta_{k+1} c_{j,k+1}^{(i)} - \beta_j c_{j-1,k}^{(i)} \tag{6.9}$$

valid for $0 \le k \le j+1$ (we assume $c_{jk}^{(i)} = 0$ for $k < 0$ or $k > j$).

Proof: This can be obtained immediately from the relation (2): If we substitute the affine functions

$$L_i(x) = a_i x + b_i$$

instead of x and compare the coefficients of $\pi_k(x)$, then after an easy calculation we obtain the equation (6.9).□

One easily observes that we can compute all coefficients $c_{jk}^{(i)}$, knowing only α_k, β_k up to $\alpha_{j-1}, \beta_{j-1}$. Now we can use the invariance property (6.6) with $h(x) = \pi_j^2(x)$ and $h(x) = x\pi_j^2(x)$ to derive recurrences for α_j and β_j.

Theorem 6.4 *Let the measure σ satisfy equation (6.1) with an operator A defined by (6.6), where the functionals L_i have the explicit form (6.7), let the coefficients $c_{jk}^{(i)}$ be defined via Equation (6.8). Then γ_j^{-2} and α_j can be calculated from the equations*

$$\gamma_j^{-2}\left(1 - \sum_{i=1}^{N} \lambda_i c_{jj}^{(i)2}\right) = \sum_{i=1}^{N}\sum_{k=0}^{j-1} \lambda_i c_{jk}^{(i)2} \gamma_k^{-2} + r_j \tag{6.10}$$

and

$$\alpha_j\gamma_j^{-2}\left(1-\sum_{i=1}^{N}\lambda_i a_i c_{jj}^{(i)2}\right)=\sum_{i=1}^{N}\lambda_i\left\{\sum_{k=0}^{j-1}[b_i c_{jk}^{(i)}+a_i(\alpha_k c_{jk}^{(i)}+2c_{j,k-1}^{(i)})]c_{jk}^{(i)}\gamma_k^{-2}+\right.$$
$$\left.[b_i c_{jj}^{(i)}+2a_i c_{j,j-1}^{(i)}]c_{jj}^{(i)}\gamma_j^{-2}\right\}+s_j, \tag{6.11}$$

where

$$r_j=\sum_{i=1}^{N}\int \pi_j^2(L_i(x))\,dq_i(x)$$

and

$$s_j=\sum_{i=1}^{N}\int L_i(x)\pi_j^2(L_i(x))\,dq_i(x).$$

Proof: In the first case $h(x)=\pi_j^2(x)$ we have

$$\int h(L_i(x))\,d\sigma(x)=\int\left(\sum_{k=0}^{j}c_{jk}^{(i)}\pi_k(x)\right)^2\,d\sigma(x)=\sum_{k=0}^{j}c_{jk}^{(i)2}\gamma_k^{-2},$$

using the relations (10). Thus, (6.6) may be written as

$$\gamma_j^{-2}=\sum_{i=1}^{N}\sum_{k=0}^{j}\lambda_i c_{jk}^{(i)2}\gamma_k^{-2}+r_j.$$

If we solve the last equation for γ_j^{-2}, we have Equation (6.10).

Analogously, in the second case $h(x)=x\pi_j^2(x)$ using the formula

$$\int x\left(\sum_{k=0}^{j}c_{jk}^{(i)}\pi_k(x)\right)^2\,d\sigma(x)=\sum_{k=0}^{j}c_{jk}^{(i)}(\alpha_k c_{jk}^{(i)}+2c_{j,k-1}^{(i)})\gamma_k^{-2}$$

we obtain

$$\alpha_j\gamma_j^{-2}=\sum_{i=1}^{N}\lambda_i\sum_{k=0}^{j}[b_i c_{jk}^{(i)}+a_i(\alpha_k c_{jk}^{(i)}+2c_{j,k-1}^{(i)})]c_{jk}^{(i)}\gamma_k^{-2}+s_j.$$

This gives finally Equation (6.11). □

Now it is clear how to design an algorithm for recursive computation of the coefficients α_j and β_j:

- The functions q_i (if non-constant) are by assumption polynomials, so we can generate with standard methods appropriate discretizations of the integrals in the quantities r_j and s_j in the preceding theorem.

- If we have already computed all α_l and β_l with $l < j$ (i.e. none, if $j = 0$), we can calculate the new coefficients $c_{jk}^{(i)}$ from Equation (6.9) or take their initial values for $j = 0$.
- The polynomials π_j are determined by the coefficients already computed, so we are able to calculate r_j and s_j.
- We compute γ_j^{-2} from (6.10) and $\alpha_j\gamma_j^{-2}$ from (6.11),
- and finally we obtain $\alpha_j = \alpha_j\gamma_j^{-2}/\gamma_j^{-2}$ and (if $j > 0$) $\beta_j = \gamma_j^{-2}/\gamma_{j-1}^{-2}$.

6.1.2 A special case of nonlinear self-similarity

In the article [Bes90], Bessis considers another type of self-similarity based on the operators B_1 and B_2 defined by

$$B_1 h(x) = \frac{1}{2}\left[h(\sqrt{x+\lambda}) + h(-\sqrt{x+\lambda})\right] \tag{6.12}$$

and

$$B_2 h(x) = \frac{1}{2\sqrt{x+\lambda}}\left[h(\sqrt{x+\lambda}) - h(-\sqrt{x+\lambda})\right]. \tag{6.13}$$

Obviously, if $h \in \mathcal{P}_d$, then $B_1 h, B_2 h \in \mathcal{P}_{[d/2]}$. Consequently, there must be representations

$$B_1\pi_j(x) = \sum_l c_{jl}\pi_l(x), \qquad B_2\pi_j(x) = \sum_l d_{jl}\pi_l(x), \tag{6.14}$$

where the sums extend over all indices l with $2l \leq j$. We can derive recursive relations for the coefficients analogous to (6.9):

Lemma 6.5 *Let the coefficients c_{jk} and d_{jk} be defined by Equation (6.14). Then we have the initial values $c_{0l} = \delta_{0l}$ and $d_{0l} \equiv 0$ and the recurrences*

$$c_{j+1,l} = d_{j,l-1} + (\alpha_l + \lambda)d_{jl} + \beta_{l+1}d_{j,l+1} - \alpha_j c_{jl} - \beta_j c_{j-1,l} \tag{6.15}$$

and

$$d_{j+1,l} = c_{jl} - \alpha_j d_{jl} - \beta_j d_{j-1,l} \tag{6.16}$$

for $j \geq 0$ and $0 \leq 2l \leq j+1$.

Proof: The initial values are consequences from the trivial identities $B_1 1 = 1$ and $B_2 1 = 0$. From the recurrence formulas (2) and the definition of the operators B_1 and B_2 we obtain the identities

$$B_1\pi_{j+1}(x) = (x+\lambda)B_2\pi_j(x) - \alpha_j B_1\pi_j(x) - \beta_j B_1\pi_{j-1}(x)$$

and

$$B_2\pi_{j+1}(x) = B_1\pi_j(x) - \alpha_j B_2\pi_j(x) - \beta_j B_2\pi_{j-1}(x).$$

Substituting (6.14) and using (2) again, we arrive at our proposition.□

If we want to derive results analogous to Theorem 6.4 for measures satisfying some invariance (6.1) with an operator A containing B_1 and B_2, we need some auxiliary results, first.

Lemma 6.6 *Let the coefficients c_{jk} and d_{jk} be defined via Equations (6.14), and let*

$$\phi_j(x) = \pi_j^2(x), \qquad \psi_j(x) = x\pi_j^2(x).$$

Then we have

$$\int B_1\phi_j(x)\, d\sigma(x) = \int B_2\psi_j(x)\, d\sigma(x) = \sum_l \left[c_{jl}^2 + 2d_{j,l-1}d_{jl} + (\lambda+\alpha_l)d_{jl}^2\right]\gamma_l^{-2},$$

$$\int B_2\phi_j(x)\, d\sigma(x) = 2\sum_l c_{jl}d_{jl}\gamma_l^{-2}$$

and

$$\int B_1\psi_j(x)\, d\sigma(x) = 2\sum_l \left[c_{j,l-1}d_{jl} + c_{jl}d_{j,l-1} + (\lambda+\alpha_l)c_{jl}d_{jl}\right]\gamma_l^{-2}.$$

Proof: From the definition of the operators B_1 and B_2 we readily obtain the identities

$$B_1\phi_j(x) = [B_1\pi_j(x)]^2 + (x+\lambda)[B_2\pi_j(x)]^2,$$

$$B_2\phi_j(x) = 2B_1\pi_j(x)\cdot B_2\pi_j(x),$$

$$B_1\psi_j(x) = (x+\lambda)B_2\phi_j(x)$$

and

$$B_2\psi_j(x) = B_1\phi_j(x).$$

If we substitute (6.14) and use (10), we obtain our propositions.□

After this preparations we can build up an algorithm for the recursive computation of the coefficients α_j and β_j. For simplicity, we will consider some special case only.

Theorem 6.7 *Let σ be the measure defined by the invariance*

$$\int h(x)\,d\sigma(x) = \lambda_0 h(0) + \lambda_1 \int B_1 h(x)\,d\sigma(x) + \lambda_2 \int B_2 h(x)\,d\sigma(x)$$

for all polynomials h, where B_1 and B_2 are defined via Equations (6.12) and (6.13) and $\lambda \geq 2$ is some real parameter. Let the coefficients c_{jl} and d_{jl} be defined via Equation (6.14). Then we have

$$\gamma_0^{-2} = \frac{\lambda_0}{1-\lambda_1}, \quad \alpha_0 = \lambda_2$$

and for $j \geq 1$

$$\gamma_j^{-2} = \lambda_0 \pi_j^2(0) + \lambda_1 \sum_{2l \leq j} \left[c_{jl}^2 + 2d_{j,l-1}d_{jl} + (\lambda + \alpha_l) d_{jl}^2\right] \gamma_l^{-2} + 2\lambda_2 \sum_{2l \leq j} c_{jl} d_{jl} \gamma_l^{-2}$$

and

$$\begin{aligned} \alpha_j \gamma_j^{-2} &= 2\lambda_1 \sum_{2l \leq j} \left[c_{j,l-1}d_{jl} + c_{jl}d_{j,l-1} + (\lambda + \alpha_l) c_{jl} d_{jl}\right] \gamma_l^{-2} + \\ &\quad + \lambda_2 \sum_{2l \leq j} \left[c_{jl}^2 + 2d_{j,l-1}d_{jl} + (\lambda + \alpha_l) d_{jl}^2\right] \gamma_l^{-2}. \end{aligned}$$

Proof: We just have to insert $h(x) = \phi_j(x) = \pi_j^2(x)$ and $h(x) = \psi_j^2(x) = x\pi_j^2(x)$ into the defining invariance property (using the formulas of the preceding lemma). For $j = 0$ this gives

$$\gamma_0^{-2} = \lambda_0 + \lambda_1 \gamma_0^{-2}$$

and

$$\alpha_0 \gamma_0^{-2} = \lambda_2 \gamma_0^{-2},$$

for $j > 0$ we obtain the above equations immediately. □

This result can be transformed into an algorithm in the same way as above.

6.2 Examples

In the following we give some examples illustrating the actual behaviour of the algorithms developed in this paper. The error magnification tabulated there is defined as the maximum of the absolute errors in the α_j and the relative errors in the β_j, divided by the machine eps. Of course, the errors can not always be calculated exactly — we have explicit formulas for the coefficients only in some cases. In the other examples, we estimate the errors by comparison with the results of the same algorithm computed with higher precision.

One remark is in place here: If we want to calculate *many* of the coefficients α_j and β_j with this algorithm (e.g. we could compute some thousands of them to get some feeling for the asymptotic behaviour of these numbers), then we must ensure that our computer can represent floating point numbers in a large range of magnitude. It's well-known that γ_j^{-2} in sufficiently regular cases behaves like a geometric sequence with ratio $\mathrm{cap} S(\sigma)$, and thus for really large j numerical overflow or underflow occurs. We can avoid this to some extent by an appropriate scaling of the support of our measure to make its capacity near to 1. Of course, this is impossible if the capacity is zero, and it's difficult if we don't know the capacity of the support.

6.2.1 Special Riesz-Nagy measures

We consider here only the special case of Lemma 4.1 in the paper of Goh and Wimp [GW93]: The distribution function σ of the measure (with support $[0,1]$ and total weight $\gamma_0^{-2} = 1$) satisfies the equations

$$\frac{1-\tau}{1+\tau}\sigma(x) + \sigma\left(\frac{1}{2}\right) = \sigma\left(x + \frac{1}{2}\right) \quad \text{for} \quad x \in \left[0, \frac{1}{2}\right] \tag{6.17}$$

and

$$\sigma\left(\frac{x}{2}\right) = \frac{1+\tau}{2}\sigma(x) \quad \text{for} \quad x \in [0,1] \tag{6.18}$$

(we wrote down Equations 4.4 and 4.5 of the cited paper in our notations). If we substitute $\frac{x}{2}$ instead of x in (6.17) and use (6.18), we obtain the equation

$$\sigma\left(\frac{x+1}{2}\right) = \frac{1-\tau}{2}\sigma(x) + \sigma\left(\frac{1}{2}\right) \quad \text{for} \quad x \in [0,1]. \tag{6.19}$$

If we now compare the last two equations with Definition 6.1, we see that σ is indeed a FIF, where $N = 2$, $X = [0,1]$, $x_0 = 0$, $x_1 = \frac{1}{2}$, $x_2 = 1$, $L_1(x) = \frac{x}{2}$ and $L_2(x) = \frac{x+1}{2}$. Moreover, we have $\lambda_1 = \frac{1+\tau}{2}$ and $\lambda_2 = \frac{1-\tau}{2}$, whereas q_1 and q_2 are constant and thus can be ignored in our algorithm (the quantities r_j and s_j in Equations (6.10) and (6.11) are all equal to 0).

We remark here that the real parameter τ in [GW93] is assumed to be $\in (0,1)$. The above equations, however, suggest that the natural interval for τ should be $(-1,1)$. Of course, for $\tau = 0$ the distribution function is not a fractal, and the measure σ is not singular — in fact, this measure is simply Lebesgue measure restricted to $[0,1]$. For $\tau \neq 0$, of course, no explicit results are known, so we have to estimate errors. The numerical results show that our algorithm seems to be stable.

Interestingly, there is some numerical evidence that the measure belongs to Nevai's class $M(\frac{1}{2}, \frac{1}{8})$ not only for $\tau = 0$, but for all $\tau \in (-1,1)$. However, the author is not able to give a proof for this conjecture.

Table 6.1: Error magnification for Riesz-Nagy measure with $\tau = 0.5$

n	error magnification
10	3.00E+00
20	1.25E+01
50	4.00E+01
100	4.30E+01
200	6.35E+01

6.2.2 Cantor measure

Let's consider Example 2.2.2, again. Cantor measure is defined by its cumulative distribution function — the well-known Cantor function. Standard properties of this function are mentioned in the paper [GW93]: The function is continuous, non-decreasing and satisfies the equations

$$\sigma\left(\frac{x}{3}\right) = \frac{1}{2}\sigma(x) \quad \text{for} \quad x \in [0,1]$$

and

$$\sigma\left(\frac{x+2}{3}\right) = \frac{1}{2} + \frac{1}{2}\sigma(x) \quad \text{for} \quad x \in [0,1].$$

Again, we see easily that this is a FIF: Here $N = 3$, $X = [0,1]$, $L_1(x) = \frac{x}{3}$, $L_3(x) = \frac{x+2}{3}$ (the missing $L_2(x)$ would be $\frac{x+1}{3}$, but can be omitted, since $\lambda_2 = 0$ and q_2 is constant), $\lambda_1 = \lambda_3 = \frac{1}{2}$ and q_1 and q_3 constant. Again, only estimates of the errors in the algorithm are available. There are no explicit formulas for the recurrence coefficients, and for all we know this is the first attempt to calculate numerically *many* of these coefficients. Recursive equations for the moments of Cantor measure based on self-similarity, of course, are well- known (see [GW93]). But numerical methods using moments are severely ill-conditioned. The same is true in this case for computations using modified moments — the orthogonal polynomials with respect to Cantor measure are very large in the gaps of the support, and the theory predicts bad condition of our problem (see the thorough discussion in the articles of Gautschi [Gau86] or [Gau90] and the alternative approach in our paper [Fis96]). In their paper [GW93], Goh and Wimp tried to compute recurrence coefficients from ordinary moments using exact (rational) arithmetic. This avoids rounding errors, but one can expect rapid growth of the length of numerators and denominators of the rationals in the algorithm, making this method impractical.

The algorithm proposed here is reasonably fast, and the (estimated) error magnification reported below indicates stability. Of course, the recurrence coefficients β_j show a very irregular behaviour due to the gaps in the support of the measure (but we have $\alpha_j \equiv \frac{1}{2}$ by symmetry of the measure).

Table 6.2: Error magnification for Cantor measure

n	error magnification
10	1.97E+00
20	4.87E+00
50	2.25E+01
100	3.15E+01
200	7.98E+01

6.2.3 An atomic measure with application to an exotic infinite sum

In our paper [Fis93], we considered the infinite sum of the reciprocals of all positive natural numbers, which (in their decimal representation) don't contain the digit 9. If we denote the set of all such numbers by M, then it has some self-similar structure: M satisfies

$$M = \{1,2,3,4,5,6,7,8\} \cup \bigcup_{z=0}^{8} \{10m + z : m \in M\}$$

(see Equation (1) in [Fis93]). We proved that our sum can be calculated from

$$\sum_{m \in M} \frac{1}{m} = l(r),$$

where $r \in C[0,1]$ is the function defined by

$$r(x) = \frac{1}{1+x} + \frac{1}{2+x} + \ldots + \frac{1}{8+x},$$

the linear functional l on $C[0,1]$ is defined via the equation

$$l(h - Ah) = h(0) \quad \text{for} \quad h \in C[0,1]$$

and the linear operator $A : C[0,1] \to C[0,1]$ satisfies

$$Ah(x) = \frac{1}{10}\left[h\left(\frac{x}{10}\right) + h\left(\frac{x+1}{10}\right) + \ldots + h\left(\frac{x+8}{10}\right)\right] \quad \text{for} \quad h \in C[0,1]$$

(see Equation (3) and Section 2 in [Fis93]). As was shown in the cited paper, the functional l is positive and $l(1) = 10$. We add here that, consequently, l represents some measure σ with support $\subset [0,1]$ and total weight $\beta_0 = \gamma_0^{-2} = 10$. The above equations fit in the general framework of Definition 6.1: We have here $N = 10$, $\lambda_k = \frac{1}{10}$ and $L_k(x) = \frac{x+k-1}{10}$ for $k = 1, \ldots, 9$, the additional term is $\lambda_{10} = \frac{1}{10}$ and $L_{10} \equiv 0$ (making the measure atomic). Thus, we can compute the recurrence coefficients of the orthogonal polynomials. Now we mentioned already in Section 0.3 that the Jacobi fraction

$$\cfrac{\beta_0}{x-\alpha_0-\cfrac{\beta_1}{x-\alpha_1-\cfrac{\beta_2}{x-\alpha_2-\ddots}}}$$

converges (uniformly on any compact not intersecting $[0,1]$) to the function

$$F(x)=\int\frac{d\sigma(t)}{x-t}.$$

Obviously, our sum can be written as

$$l(r)=-F(-1)-F(-2)-\ldots-F(-8).$$

It's not very difficult to estimate the truncation error of the continued fraction: Since in fact the support of our measure is contained in $[0,\frac{8}{9}]$, Lemma 1.6 of Chapter 1 gives the bound

$$|F(-1)-F_n(-1)|\le 90e^2n(n+1)\left(\frac{13+3\sqrt{17}}{4}\right)^{-2n}$$

This is $\approx 4.809333\cdot 10^{-18}$ for $n=14$, and the errors of the other terms are much smaller. In reality, this bound is very conservative: Computing the recurrence coefficients up to α_{10} and β_{10} with the help of our algorithm, we obtain the approximation

$$\sum_{m\in M}\frac{1}{m}\approx 22.9206766192641503.$$

We see that all of its digits are correct, since in [Fis93] with another method we computed the value

$$\begin{aligned}\sum_{m\in M}\frac{1}{m} &= 22.920676619264150348163657094375931914944762436998481\\ &\quad 568541998356572156338189911129445626037448 2\ldots\end{aligned}$$

6.2.4 An absolutely continuous measure with an infinite number of gaps

In the paper [Fis95a], we gave an example of an absolutely continuous measure having oscillating terms in the asymptotics of the moments. This strange behavior is caused by an infinite number of gaps in the support of the measure. We should expect that the recurrence coeffients for the same reason behave very irregularly, too. As for the Cantor measure, traditional algorithms are not able to compute more than a few of these coefficients. However, the defining Equation (10) in [Fis95a] for the distribution function in our notations reads

$$\sigma(x) = \begin{cases} 2x & \text{for} \quad x \in [0, \frac{1}{4}], \\ \frac{1}{2} & \text{for} \quad x \in [\frac{1}{4}, \frac{1}{2}], \\ \frac{1}{2} + \frac{1}{2}\sigma(2x-1) & \text{for} \quad x \in [\frac{1}{2}, 1], \end{cases}$$

and can be rewritten in the form

$$\begin{aligned} \sigma\left(\frac{x}{4}\right) &= \frac{x}{2} \\ \sigma\left(\frac{x+1}{4}\right) &= \frac{1}{2} \\ \sigma\left(\frac{x+1}{2}\right) &= \frac{1}{2} + \frac{1}{2}\sigma(x) \end{aligned}$$

(for $x \in [0,1]$) and this clearly defines a FIF. Note, however, that in this example we have a nonconstant function $q_1(x) = \frac{x}{2}$ and consequently we have to calculate the quantities r_j and s_j in the algorithm by some appropriate discretization. This is, of course, a trivial task. Since the measure is essentially the Lebesgue measure,

Table 6.3: Error magnification for the absolutely continuous measure

n	error magnification
10	4.80E+01
20	9.40E+01
50	4.20E+02
100	4.83E+02
200	4.83E+02
300	8.76E+02

we used a Gauss quadrature formula which integrates exactly polynomials of all degrees occuring in our algorithm (i.e. up to degree $2n-1$).

Remark: The actual numerical computations were performed not for this measure, but for the measure with support scaled by a factor of 4 (the smallest enclosing interval is not $[0,1]$, but $[0,4]$). This was done for the reasons explained in the beginning of this section.

6.2.5 A special case of little q-Jacobi polynomials

In this example, our algorithm will turn out to be numerically not stable. We consider a measure concentrated in a countable set of isolated mass points:

$$\sigma(\{q^n\}) = (1-q)q^n \quad \text{for} \quad n = 0, 1, \ldots$$

with some real parameter $q \in (0,1)$. We can identify this measure as the measure of orthogonality of some very special little q-Jacobi polynomials: These polynomials in general are defined by

$$p_n(x;a,b;q) = {}_2\Phi_1\left[\begin{matrix} q^{-n}, abq^{n+1} \\ aq \end{matrix}; q, qx\right]$$

and their orthogonality relations read

$$\begin{aligned} &\frac{(q^{\alpha+1};q)_\infty (q^{\beta+1};q)_\infty}{(1-q)(q;q)_\infty (q^{\alpha+\beta+2};q)_\infty} \int_0^1 p_n(x)\, p_m(x)\, x^\alpha \frac{(qx;q)_\infty}{(q^{\beta+1}x;q)_\infty}\, d_q x \\ &= \frac{q^{n(\alpha+1)}(1-q^{\alpha+\beta+1})(q^{\beta+1};q)_n (q;q)_n}{(1-q^{2n+\alpha+\beta+1})(q^{\alpha+1};q)_n (q^{\alpha+\beta+1};q)_n}\, \delta_{nm} \quad \text{for} \quad \alpha,\beta > -1. \end{aligned}$$

These are Equations (5.7) and (5.8) of [Koo90], and we add the missing relations $a = q^\alpha$ and $b = q^\beta$. For the notations used we refer to [Rah90]). The polynomials are not monic; their recurrence relations are

$$xp_n(x) = A_n p_{n+1}(x) - (A_n + B_n)p_n(x) + B_n p_{n-1}(x),$$

where

$$A_n = \frac{(1-aq^{n+1})(1-abq^{n+1})}{(1-abq^{2n+1})(1-abq^{2n+2})}\,(-q^n)$$

and

$$B_n = \frac{(1-q^n)(1-bq^n)}{(1-abq^{2n})(1-abq^{2n+1})}\,(-aq^n)$$

(see [Rah90, p. 328]). The transformation to monic polynomials is elementary. After an easy calculation we obtain

$$\alpha_n = -A_n - B_n \quad \text{and} \quad \beta_n = B_n A_{n-1}.$$

As an examination of the orthogonality relations shows, our measure corresponds to the case $\alpha = \beta = 0$, i.e. $a = b = 1$. The polynomials $p_n(x;1,1;q)$ in [Koo90] are called *little q-Legendre polynomials.*

The self-similar structure of our measure is obvious. We have

$$\begin{aligned} &\int h(x)\, d\sigma(x) = \int_0^1 h(x)\, d_q x = \sum_{n=0}^\infty h(q^n)(1-q)q^n \\ &= (1-q)h(1) + q\int_0^1 h(qx)\, d_q x = (1-q)h(1) + q\int h(qx)\, d\sigma(x) \end{aligned}$$

for any polynomial h. Here we have the degenerate case $L_1 \equiv 1$ (and $\lambda_1 = (1-q)\gamma_0^2$, $L_2(x) = qx$, $\lambda_2 = q$). Thus, only $c_{j0}^{(1)} \neq 0$, and Equation (6.9) reduces simply to the well-known three- term recurrence (2) for $\pi_j(1)$. However, this point is an isolated mass-point of the measure, and the recurrence is unstable (see the discussion in our paper [Fis97]). The numerical results confirm our expectations.

Table 6.4: Error magnification for little q-Legendre polynomials with q=0.5

n	error magnification
9	2.00E+00
10	2.54E+02
11	1.34E+08
12	2.80E+14
13	2.35E+21

6.2.6 Bessis' example

We return to our Example 2.2.1. Its defining invariance property (5.20) reads

$$\int h(x)\, d\sigma(x) = \int B_1 h(x)\, d\sigma(x) \tag{6.20}$$

with the operator B_1 defined via Equation (6.12). As we mentioned, the recurrence coefficients β_j are not known explicitly, but can be calculated recursively from the equations (5.21). Of course, we have $\alpha_j \equiv 0$ due to symmetry.

In this example we can use the algorithm derived from our Theorem 6.7 (the invariance doesn't determine γ_0^{-2}, since $\lambda_0 = 1 - \lambda_1 = 0$, so we assume $\gamma_0 = 1$). The numerical results indicate stability.

Table 6.5: Error magnification for Bessis measure with $\lambda = 2.1$

n	error magnification
10	2.00E+00
20	2.00E+00
50	5.00E+00
100	8.00E+00
200	1.40E+01

6.2.7 Another atomic measure

Now, we return to our Example 5.5.3. With the notation introduced above, its invariance property can be written as

$$\int h(x)\, d\sigma(x) = \frac{1}{3}h(0) + \frac{2}{3}\int B_1 h(x)\, d\sigma(x). \tag{6.21}$$

Since the mass points are not isolated, we can expect stability of the algorithm derived from Theorem 6.7. The recurrence coefficients β_j can be calculated from a system of equations similar to (though more complicated than) equations (5.21). This was shown already in Example 5.5.3. Thus we can exactly calculate error magnification in our algorithm. The results confirm the expected stability. We should remark, however, that for $\lambda > 2$ we would have isolated mass points and nevertheless a stable recursion: The errors stay zero by symmetry.

Table 6.6: Error magnification for another atomic measure

n	error magnification
10	2.00E+00
20	3.00E+00
50	8.00E+00
100	1.60E+01
200	4.60E+01

6.3 Bibliographical references

In the last years, measures on fractals became very interesting for physicists, see the papers [BBM82], [Bes90], [BM91], [GM94] and [Man96] and the bibliographical references therein. These measures and their supports can be defined by some invariance properties, which are usually called self-similarity. The mathematical description in the simplest case is due to Barnsley, e.g. in the article [BH89]. Since self-similarity implies homogeneity of the measure in the sense (1.10), the regularity of those measures is trivial. This is not true for Markov's inequality on their supports. The results of Wagner in [Wag91] imply Hölder-continuity of the Green function of some Cantor-type sets (not including the classical middle-third Cantor set), but he did not state this result explicitly. This Hölder-continuity property (HCP) of the Cantor set and some generalizations were shown by Białas and Volberg [BV93] and by Totik [Tot95]. Their methods obviously extend to supports of IFS-measures. The fact that HCP implies Markov's inequality is well known, see the article of Jonsson [Jon94]. Almost nothing seems to be known about numerical methods for the computation of the equilibrium measures of fractals (except the extremal point method, see Chapter V of [ST97]). It is hard to decide wether the ideas of an article of Jones, Ma and Rokhlin [JMR94] can be useful for computations (due to the level of mathematical rigour).

It should be mentioned that an algorithm for the computation of Jacobi matrices was developed independently by Mantica in [Man96]. His method is very similar to our algorithm, though not identical. Mantica works with orthonormal polynomials, this gives more complicated formulas. He gives a proof for the stability of his algorithm. This proof, however, seems to be wrong: In a personal communication, Mantica recently reported the same numerical difficulties for our Example 6.2.5. The algorithm for the nonlinear case seems to be new.

In the presence of gaps in the support of the measure, we can expect an irregular behavior of the recurrence coefficients. The existence of infinitely many limit points of the coefficients for measures supported on the union of two disjoint intervals was shown by Peherstorfer in his paper [Peh95] (except in special cases, e.g. for two subintervals of equal length).

We add the remark that orthogonal polynomials of the type of our Example 6.2.5 were considered already in the paper of Hahn [Hah49].

Bibliography

[AS84] Milton Abramowitz and Irene Stegun, editors. *Pocketbook of Mathematical Functions* . Verlag Harri Deutsch, Thun — Frankfurt am Main, 1984.

[BB97] Bernhard Beckermann and Emmanuel Bourreau. How to choose modified moments. preprint, 1997.

[BBM82] J. Bellisard, D. Bessis, and P. Moussa. Chaotic states of almost periodic Schrödinger operators. *Physical Review Letters*, 49(10):701–704, 1982.

[Bes90] Daniel Bessis. Orthogonal polynomials, Padé approximations and Julia sets. In Paul Nevai, editor, *Orthogonal Polynomials, Theory and Practice*, volume 294 of *NATO ASI Series C*, pages 55–97. Kluwer Academic Publishers, Dordrecht, Boston, London, 1990.

[BH89] Barnsley and Harrington. The calculus of fractal interpolation functions. *Journal of Approximation Theory*, 57:14–34, 1989.

[BM91] Daniel Bessis and Giorgio Mantica. Construction of multifractal measures in dynamical systems from their invariance properties. *Physical Review Letters*, 66(23):2939–2942, 1991.

[BV93] Leokadia Białas and Alexander Volberg. Markov's property of the Cantor ternary set. *Studia Mathematica*, 104(3):259–268, 1993.

[Chi78] T. S. Chihara. *An Introduction to Orthogonal Polynomials*. Gordon & Breach, New York, 1978.

[Doo84] J. L. Doob. *Classical Potential Theory and Its Probabilistic Counterpart*, volume 262 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin et al., 1984.

[DR93] A. Dutt and V. Rokhlin. Fast fourier transforms for nonequispaced data. *SIAM Journal of Scientific Computations*, 14:1368–1393, 1993.

[FHR93] Tilo Finck, Georg Heinig, and Karla Rost. An inversion formula and fast algorithms for Cauchy-Vandermonde matrices. *Linear Algebra and its Applications*, 183:179–191, 1993.

[Fis93] Hans-Jürgen Fischer. Die Summe der Reziproken der natürlichen Zahlen ohne Ziffer 9. *Elemente der Mathematik*, 48:100–106, 1993.

[Fis95a] Hans-Jürgen Fischer. On the paper "Asymptotics for the moments of singular distributions". *Journal of Approximation Theory*, 82:362–374, 1995.

[Fis95b] Hans-Jürgen Fischer. Recurrence coefficients of orthogonal polynomials with respect to some self- similar singular distributions. *Zeitschrift für Analysis und ihre Anwendungen*, 14(1):141–155, 1995.

[Fis96] Hans-Jürgen Fischer. On the condition of orthogonal polynomials via modified moments. *Zeitschrift für Analysis und ihre Anwendungen*, 15(1):223–244, 1996.

[Fis97] Hans-Jürgen Fischer. On generating orthogonal polynomials for self-similar measures. *talk at the conference CMFT'97 in October 1997 in Nicosia*, 1997. to appear in the Proceedings.

[Fis98a] Hans-Jürgen Fischer. Explicit calculation of some polynomials introduced by W. Gautschi. *Zeitschrift für Analysis und ihre Anwendungen*, 17(4):963–977, 1998.

[Fis98b] Hans-Jürgen Fischer. On generating orthogonal polynomials with respect to discrete measures. *Zeitschrift für Analysis und ihre Anwendungen*, 17(1):183–205, 1998.

[Fre71] G. Freud. *Orthogonal Polynomials*. Pergamon Press, Oxford, 1971.

[Gau78] Walter Gautschi. Questions of numerical condition related to polynomials. In C. de Boor and G. H. Golub, editors, *Recent Advances in Numerical Analysis*, pages 45–72. Academic Press, New York, 1978.

[Gau82] Walter Gautschi. On generating orthogonal polynomials. *SIAM Journal of Scientific and Statistical Computations*, 3(3):289 – 317, 1982.

[Gau86] Walter Gautschi. On the sensitivity of orthogonal polynomials to perturbations in the moments. *Numerische Mathematik*, 48:369–382, 1986.

[Gau90] Walter Gautschi. Computational aspects of orthogonal polynomials. In Paul Nevai, editor, *Orthogonal Polynomials, Theory and Practice*, volume 294 of *NATO ASI Series C*, pages 181–216. Kluwer Academic Publishers, Dordrecht, Boston, London, 1990.

[Gau92] Walter Gautschi. On mean convergence of extended Lagrange interpolation. *J. Comput. Appl. Math.*, 43:19–35, 1992.

[Gau93] Walter Gautschi. Is the recurrence relation for orthogonal polynomials always stable? *BIT*, 33:277–284, 1993.

[Gau94] Walter Gautschi. Algorithm 726: ORTHPOL — a package of routines for generating orthogonal polynomials and Gauss-type quadrature rules. *ACM Trans. Math. Software*, 20:21–62, 1994.

[Gau96] Walter Gautschi. Orthogonal polynomials: applications and computations. In *Acta Numerica 1996*, pages 45–119. Cambridge University Press, Cambridge, 1996.

[Gem97] L. Gemignani. A fast algorithm for generalized Hankel matrices arising in finite-moment problems. *Linear Algebra and its Applications*, 267:41–52, 1997.

[GG90] Gene Golub and Martin Gutknecht. Modified moments for indefinite weight functions. *Numerische Mathematik*, 57:607–624, 1990.

[GH84] W.B. Gragg and W.J. Harrod. The numerically stable reconstruction of jacobi matrices from spectral data. *Numerische Mathematik*, 44:317–335, 1984.

[GM94] Italo Guarneri and Giorgio Mantica. Multifractal energy spectra and their dynamical implications. *Physical Review Letters*, 73(25):3379–3382, 1994.

[GO94] I. Gohberg and V. Olshevsky. Fast inversion of Chebyshev-Vandermonde matrices. *Numerische Mathematik*, 67:71–92, 1994.

[GW93] William Goh and Jet Wimp. Asymptotics for the moments of singular distributions. *Journal of Approximation Theory*, 74:301–334, 1993.

[Hah49] W. Hahn. Über Orthogonalpolynome die q-Differenzengleichungen genügen. *Math. Nachr.*, 2:4–34, 1949.

[HHR89] Georg Heinig, Wolfram Hoppe, and Karla Rost. Structured matrices in interpolation and approximation problems. *Wissenschaftliche Zeitschrift der TU Karl-Marx-Stadt*, 31:196–202, 1989.

[Hig88] N. J. Higham. Fast solution of Vandermonde-like systems involving orthogonal polynomials. *IMA Journal of Numerical Analysis*, 8:473–486, 1988.

[Hig90] N. J. Higham. Stability analysis of algorithms for solving confluent Vandermonde-like systems. *SIAM Journal of Matrix Analysis and Applications*, 11:23–41, 1990.

[JMR94] P. Jones, J. Ma, and V. Rokhlin. A fast direct algorithm for the solution of the Laplace equation on regions with fractal boundaries. *Journal of Computational Physics*, 113:35–51, 1994.

[Jon94] Alf Jonsson. Markov's inequality and zeros of orthogonal polynomials on fractal sets. *Journal of Approximation Theory*, 78:87–97, 1994.

[KO97] Thomas Kailath and Vadim Olshevsky. Displacement-structure approach to polynomial Vandermonde and related matrices. *Linear Algebra and its Applications*, 216:49–90, 1997.

[Koo90] T. H. Koornwinder. Orthogonal polynomials in connection with quantum groups. In Paul Nevai, editor, *Orthogonal Polynomials, Theory and Practice*, volume 294 of *NATO ASI Series C*, pages 257–292. Kluwer Academic Publishers, Dordrecht, Boston, London, 1990.

[Lu94] Hao Lu. Fast solution of confluent Vandermonde linear systems. *SIAM Journal of Matrix Analysis and Applications*, 15:1277–1289, 1994.

[Lu96] Hao Lu. Solution of Vandermonde-like systems and confluent Vandermonde-like systems. *SIAM Journal of Matrix Analysis and Applications*, 17:127–138, 1996.

[Lub88] D. S. Lubinsky. Jump distributions on $[-1, 1]$ whose orthogonal polynomials have leading coefficients with given asymptotics. *Proc. Amer. Math. Soc.*, 104:516–524, 1988.

[Mag95] Alphonse Magnus. Special nonuniform lattice (snul) orthogonal polynomials on discrete dense sets of points. *Journal of Computational and Applied Mathematics*, 65:253–265, 1995.

[Man96] Giorgio Mantica. A stable Stieltjes technique for computing orthogonal polynomials and Jacobi matrices associated with a class of singular measures. *Constructive Approximation*, 12:509–530, 1996.

[Mas93] J. C. Mason. Chebyshev polynomials of second, third and fourth kinds. *J. Comput. Appl. Math.*, 49:169–178, 1993.

[ME93] J. C. Mason and G.H. Elliott. Near-minimax complex approximation by four kinds of Chebyshev polynomial expansions. *J. Comput. Appl. Math.*, 46:291–300, 1993.

[Mei64] Günter Meinardus. *Approximation von Funktionen und ihre numerische Behandlung.* Springer-Verlag, Berlin et al., 1964.

[Mer91] A. McD. Mercer. A note on the zero distribution of orthogonal polynomials. *Journal of Approximation Theory*, 64:238–242, 1991.

[Nev79] Paul Nevai. Orthogonal polynomials. *Memoir of the American Mathematical Society*, 18(213), 1979.

[NS88] Никишин, Е. М. and В. Н. Сорокин. *Рациональные аппроксимации и ортогональность.* Наука, Москва, 1988.

[Peh95] Franz Peherstorfer. Elliptic orthogonal and extremal polynomials. *Proceedings of the London Mathematical Society*, pages 605–624, 1995.

[Rah90] Mizan Rahman. Some extensions of the beta integral and the hypergeometric function. In Paul Nevai, editor, *Orthogonal Polynomials, Theory and Practice*, volume 294 of *NATO ASI Series C*, pages 319–344. Kluwer Academic Publishers, Dordrecht, Boston, London, 1990.

[Rak77] E. A. Rakhmanov. On the asymptotics of the ratio of orthogonal polynomials. *Mat. Sb.*, 103(145):237–252, 1977.

[Rak82] E. A. Rakhmanov. On the asymptotics of the ratio of orthogonal polynomials II. *Mat. Sb.*, 118(160):104–117, 1982.

[Rei90] Lothar Reichel. Newton interpolation at Leja-points. *BIT*, 30:332–346, 1990.

[Rei93] Lothar Reichel. Construction of polynomials that are orthogonal with respect to a discrete bilinear form. *Adv. Comput. Math.*, 1:241–258, 1993.

[Saf90] Edward B. Saff. Orthogonal polynomials from a complex perspective. In Paul Nevai, editor, *Orthogonal Polynomials, Theory and Practice*, volume 294 of *NATO ASI Series C*, pages 363–393. Kluwer Academic Publishers, Dordrecht, Boston, London, 1990.

[ST90] H. Stahl and V. Totik. Nth root asymptotic behavior of orthonormal polynomials. In Paul Nevai, editor, *Orthogonal Polynomials, Theory and Practice*, volume 294 of *NATO ASI Series C*, pages 395–417. Kluwer Academic Publishers, Dordrecht, Boston, London, 1990.

[ST91] G. Steidl and M. Tasche. A polynomial approach to fast algorithms for discrete Fourier-cosine and Fourier-sine transforms. *Mathematics of Computation*, 56(193):281–296, 1991.

[ST92] Herbert Stahl and Vilmos Totik. *General Orthogonal Polynomials*, volume 43 of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press, Cambridge, 1992.

[ST97] Edward B. Saff and Vilmos Totik. *Logarithmic Potentials with External Fields*, volume 316 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin et al., 1997.

[Sze75] G. Szegö. *Orthogonal Polynomials*, volume 23 of *Colloquium Publications*. Amer. Math. Soc., Providence, 1975.

[Tot95] Vilmos Totik. Markoff constants for Cantor sets. *Acta Sci. Math.*, 60:715–734, 1995.

[Tsu58] M. Tsuji. *Potential Theory in Modern Function Theory*. Chelsea, New York, 1958.

[Ull72] J. L. Ullman. On the regular behavior of orthogonal polynomials. *Proceedings of the London Mathematical Society*, 24:119–148, 1972.

[vAM89] W. van Assche and A. P. Magnus. Sieved orthogonal polynomials and discrete measures with jumps dense in an interval. *Proc. Amer. Math. Soc.*, 106:163–173, 1989.

[Wag91] Gerold Wagner. On the maximal modulus of polynomials on Cantor sets. *Journal of Approximation Theory*, 67:1–18, 1991.

[Zha94] Minda Zhang. Sensitivity analysis for computing orthogonal polynomials of Sobolev type. In *Approximation and Computation*, volume 119 of *ISNM*, pages 563–576. Birkhäuser, Boston, 1994.

List of Symbols

$\mathbf{R}$	real line, p. 1
$S(\sigma)$	support (or spectrum) of the measure σ, p. 1
π_j	orthogonal polynomials, p. 1
α_j, β_j	recurrence coefficients, p. 1
(p, q)	inner product of two polynomials p and q, p. 2
Ψ	operator of multiplication by the independent variable, p. 2
$\hat{\pi}_j$	orthonormal polynomials, p. 2
γ_j	leading coefficient of $\hat{\pi}_j$, p. 2
κ_j	monic numerator polynomials, p. 4
$\mathcal{P}_d$	set of all polynomials of degree $\leq d$, p. 5
$\deg(q)$	degree of a polynomial q, p. 5
J_n	Jacobi matrix, p. 6
δ_t	point mass (weight 1) at t, p. 7
$\sigma^{(N)}$	discretization of the measure σ, p. 7
ρ_n	vector of the first recurrence coefficients, p. 7
$[n-1, n]_F$	$n-1, n$-Padé approximation to F, p. 9
$B(x, r)$	the closed ball with center x and radius r, p. 14
$S \in \mathcal{M}^p(M, \alpha)$	S has Markov's inequality in L^p, p. 14
$\Vert \cdot \Vert_p$	norm in L^p, p. 14
$\mathrm{dist}(A, B)$	distance between two sets A and B, p. 15
$\mathcal{P}_N^*$	dual space of $\mathcal{P}_N$, p. 17
p_j^*	coordinate functional in the base $\{p_j\}_{j=0,\ldots,N}$, p. 17
l_i	fundamental polynomials of Lagrange interpolation, p. 18
h_i, k_i	fundamental polynomials of Hermite interpolation, p. 18
$q \div r$, $q \bmod r$	partial quotient and remainder in polynomial division, p. 20
$\mathbf{C}$	complex plane, p. 24
$\mathrm{cap}(B)$	logarithmic capacity of B, p. 24
ω_K	equilibrium distribution of K, p. 24
$d(M)$	density of a set M of natural numbers, p. 30
$\xrightarrow{d}$	convergence in density, p. 30
cond	condition number, p. 39
T_j	Chebyshev polynomials of first kind, p. 42

T_j^*	shifted Chebyshev polynomials of first kind, p. 43
U_j	Chebyshev polynomials of second kind, p. 43
U_j^*	shifted Chebyshev polynomials of second kind, p. 44
V_j	Chebyshev polynomials of third kind, p. 77
W_j	Chebyshev polynomials of fourth kind, p. 77

Thesen

zur Habilitationsschrift

Analytic, numeric and algorithmic aspects of orthogonal polynomials

zur Erlangung des akademischen Grades eines "Dr. rer. nat. habil."

an der Technischen Universität Chemnitz
Fakultät für Mathematik

vorgelegt von Dr. rer. nat. Hans-Jürgen Fischer

1. Es ist seit langem bekannt, daß zu einem gegebenen Maß mit endlichen Momenten eine eindeutig bestimmte Folge monischer orthogonaler Polynome $(\pi_j)_{j=0,1,\ldots,}$ gehört. Dies Folge genügt einer dreigliedrigen Rekursionsbeziehung der Gestalt

$$\pi_{j+1}(x) = (x - \alpha_j)\pi_j(x) - \beta_j \pi_{j-1}(x) \quad \text{für} \quad j = 0, 1, \ldots,$$

den Vektor der Rekursionskoeffizienten $\rho = [\alpha_j, \beta_j]^T_{j=0,\ldots}$ kann man als Kodierung der Folge der orthogonalen Polynome selbst betrachten. Als Kodierung für das Maß wurden bisher entweder eine Folge von (modifizierten) Momenten oder eine Diskretisierung im Sinne einer Quadraturformel betrachtet. Das setzte voraus, daß das Maß absolutstetig oder diskret war. Auch in diesen Fällen ist freilich die numerische Kondition der Abbildung von der Kodierung des Maßes auf ρ (bzw. einen endlichen Abschnitt ρ_n) entscheidend. Da außerdem in letzter Zeit singuläre Maße in physikalischen Anwendungen eine wachsende Rolle spielen, ist eine Berechnung der orthogonalen Polynome auch in diesem Fall wünschenswert. Als geeignete Kodierung des Maßes erweisen sich hier Selbstähnlichkeitseigenschaften.

In dieser Arbeit werden nur Maße mit kompaktem Support behandelt.

2. Neben lange bekannten algebraischen Eigenschaften der orthogonalen Polynome benötigt man Abschätzungen ihres Betrags auf dem Support des Orthogonalitätsmaßes. Diese kann man aus einer Markov-Ungleichung auf dem Support und einer gewissen Homogenität des Maßes leicht erhalten, erstaunlicherweise wurde dieses Resultat bisher anscheinend nirgends explizit formuliert. Es ist nützlich, da die Voraussetzungen bei selbstähnlichen Maßen erfüllt sind.

Ein weiteres zwar elementares, aber sehr mächtiges Werkzeug ist ein allgemeines Lemma über die Abhängigkeit der Rekursionskoeffizienten von einem Parameter des Maßes. Die numerische Kondition sowohl von momenten- als auch von diskretisierungsbasierten Verfahren läßt sich damit effektiv behandeln.

3. Über das asymptotische Verhalten der Rekursionskoeffizienten für singuläre Maße war bisher nicht viel bekannt. Für sogenannte reguläre Maße wird hier die Konvergenz von Mittelwerten der Koeffizienten bewiesen. Im Falle eines zusammenhängenden Supports (also eines Intervalls) kann man reguläre Maße noch genauer charakterisieren: Hier hat man auch reguläre Asymptotik der Nullstellenverteilung der orthogonalen Polynome und eine Fastkonvergenz der Rekursionskoeffizienten (bis auf eine Indexmenge der Dichte 0).

4. Mit diesen Hilfsmitteln läßt sich die numerische Kondition von momentenbasierten Verfahren zur Berechnung orthogonaler Polynome erstmals direkt untersuchen — bisher war das nur auf Umwegen möglich (etwa in den Arbeiten von W. Gautschi). Für spezielle Polynombasen und spezielle Orthogonalitätsmaße lassen sich sogar geschlossene Ausdrücke für die Norm des Jacobians $\|K_n'\|$ der Abbildung K_n von den (modifizierten) Momenten auf ρ_n angeben. Eine Verallgemeinerung auf orthogonale Polynome vom Sobolev-Typ ist möglich.

5. Die Abbildung G_n von modifizierten Momenten auf Quadraturformeln wurde bisher nur als theoretisches Hilfsmittel betrachtet. Die numerische Implementation führt auf ein nichtlineares Gleichungssystem, dessen iterative Lösung mittels Newton-Verfahren versucht wird. Für die auftretenden linearen Gleichungssysteme mit konfluenten Vandermonde-Typ-Matrizen wird ein schnelles Lösungsverfahren hergeleitet. Dazu werden Algorithmen für die Multiplikation und Division von Polynomen entwickelt, die als Linearkombination orthogonaler Polynome gegeben sind. Obwohl das Iterationsverfahren (zumindest für Chebyshev-Momente und im Falle der Konvergenz) asymptotisch schneller als traditionelle Verfahren sein sollte, ist es ihnen für praktisch auftretende Werte von n unterlegen.

Es wurde gezeigt, daß sich die von Gautschi definierten Polynome g_n, die die Frobenius-Norm $\|G_n'\|_F$ bestimmen, für Jacobi-Gewichte abschätzen oder (bei Chebyshev-Polynomen 1.–3. Art) sogar geschlossen darstellen lassen. Die von Gautschi vermutete gleichmäßige Beschränktheit dieser Norm für Chebyshev-Polynome 1. Art konnte damit bewiesen werden.

6. Für die Abbildung H_n von Quadraturformeln auf ρ_n waren bisher im wesentlichen zwei numerische Implementationen bekannt: Der klassische Stieltjes-Algorithmus und die neuere Lanczos-Prozedur. Die bereits bekannte numerische Instabilität des ersteren Verfahrens im Falle isolierter Massepunkte des Orthogonalitätsmaßes wird hier nochmals theoretisch begründet und ein neues numerisches Verfahren vorgestellt. Die numerische Kondition von H_n kann mit dem allgemeinen Lemma aus Kapitel 1 ebenfalls behandelt werden.

7. Schließlich wird ein numerisches Verfahren zur Berechnung der Rekursionskoeffizienten von orthogonalen Polynomen bezüglich singulärer Maße hergeleitet. Diese Maße müssen beschreibbar sein durch eine Selbstähnlichkeitseigenschaft: Entweder muß ihre Verteilungsfunktion eine fraktale Interpolationsfunktion (FIF) im Sinne von Barnsley und Harrington sein, oder eine spezielle nichtlineare Verallgemeinerung. Da es sich letztlich um eine Stieltjes-Prozedur handelt, ist das Verfahren numerisch instabil im Falle der Existenz isolierter Massepunkte des Orthogonalitätsmaßes.

8. Alle theoretischen Aussagen und Algorithmen werden durch konkrete Beispiele illustriert. Diese sind zwar meist nicht unmittelbar praktisch relevant, erfassen aber alle praktisch auftretenden Problemfälle, zumindest für Maße mit kompaktem Support.

Dr. rer. nat. Hans-Jürgen Fischer
Fakultät für Mathematik
Technische Universität Chemnitz
09107 Chemnitz

Erklärung

Hiermit erkläre ich, daß die vorliegende Arbeit von mir selbst und ohne andere als die darin angegebenen Hilfsmittel angefertigt wurde.

Chemnitz, den 7.10.1998