Contents

1 The Hamburger moment problem 7
  1.1 Introduction .......................................................... 7
  1.1.1 Moment sequences and positive definiteness .................. 7
  1.1.2 Determinacy of the moment problem ............................ 11
  1.1.3 Example due to Stieltjes .............................. 13
  1.2 Orthonormal polynomials ............................................ 16
    1.2.1 Construction and properties of orthonormal polynomials ... 16
    1.2.2 The kernel polynomial .................................. 18
  1.3 Proof of Theorem 1.18 ........................................... 21

2 The matrix moment problem 27
  2.1 Matrix measures .................................................. 27
  2.2 Formulation of the matrix moment problem ...................... 30
  2.3 Matrix inner products and orthonormal matrix polynomials .... 33
    2.3.1 Matrix polynomials .................................. 33
    2.3.2 Properties of orthonormal matrix polynomials .............. 39
    2.3.3 Matrix polynomials of the first and second kind ....... 42
    2.3.4 The kernel polynomial .................................. 46
  2.4 Some difficulties that arise while generalizing Theorem 1.18 .. 48

3 The operator approach to the moment problem 53
  3.1 The Jacobi operator .............................................. 53
  3.2 The indices of deficiency ....................................... 56
    3.2.1 Relation between the deficiency indices and determinacy of the matrix
    moment problem ........................................... 57
    3.2.2 The deficiency indices in terms of orthonormal matrix polynomials . 58
  3.3 Examples of the matrix moment problem ........................ 61
    3.3.1 Examples with diagonalizable weights .................... 61
    3.3.2 Example arising from a doubly infinite Jacobi operator .. 64

4 Criteria for (in)determinacy 71
  4.1 Conditions for completely indeterminate moment problems ...... 71
  4.2 Generalizations from Akhiezer .................................. 71

A Complex measures 77
  A.1 Complex and positive measures ................................. 77
  A.2 Radon-Nikodym Theorem ........................................... 78
B Spectral theory

B.1 Bounded operators
   B.1.1 Hilbert spaces and operators
   B.1.2 The spectral theorem for bounded self-adjoint operators

B.2 Unbounded operators
   B.2.1 The spectral theorem for unbounded self-adjoint operators

Bibliography

4
Introduction

The term moment problem originates from 1894, when Stieltjes introduced it as the following problem: It is required to find the distribution of positive mass on the interval $[0, \infty)$, given the moments of order $n$ ($n = 0, 1, 2, \ldots$) of the distribution. Thus, in Stieltjes’ problem, a certain sequence of numbers $(s_n)_{n \geq 0}$ is given and a non-decreasing function $\sigma(x)$ ($x \geq 0$) is sought such that

$$\int_0^\infty x^n d\sigma(x) = s_n \text{ for all } n \geq 0.$$ 

Note that we will use the terminology of measures instead of distributions in this thesis. In the statement of Stieltjes’ problem, the carrier of the mass is the semi-infinite interval $[0, \infty)$, but instead one can look at the moment problem on any given interval or any other point set that is contained in $[0, \infty)$, by requiring that a certain part of the semi-axis $[0, \infty)$ must be free of mass. Here however, we study the ‘extended’ moment problem on $(-\infty, \infty)$, which is known as the Hamburger moment problem. The moment problem has been researched extensively in works like [1], and both orthogonal polynomials and spectral theory have turned out to be major tools in discussing this problem. Since the papers [20] and [21] of Krein, there is a general theory on matrix-valued orthogonal polynomials. By using this theory amongst others, the matrix-valued analogue of the moment problem, called the matrix moment problem, has been studied.

The original purpose of this thesis was to generalize the main result of [6], which states that the scalar moment problem is determinate (i.e. has a unique solution) if and only if the smallest eigenvalues of the Hankel matrices $H_N$ tend to zero whenever $N \to \infty$. In Chapter 1 the classical moment problem is introduced, and the determinacy of the measure(s) involved in the moment problem is discussed. Moreover the corresponding orthonormal polynomials are defined and some properties of these polynomials are studied. These concepts are necessary in order to understand the proof given in [6], which is the content of Section 1.3.

Since we wanted to generalize [6], Chapter 2 is dedicated to the matrix moment problem. A brief overview of matrix measures is given, before discussing the matrix moment problem itself and the associated orthonormal matrix polynomials, following a similar structure as the first chapter. Some attempts have been made in order to give a generalization, but to no avail. A few of these attempts are collected in Section 2.4, along with the problems one encounters while trying to find such a generalization.

Afterwards the matrix moment problem is treated from a functional analytic perspective in Chapter 3. Herefore a summary of spectral theory is given. In this chapter a connection between the orthonormal polynomials and the so-called Jacobi operator (which forms the starting point of the operator approach to the moment problem) is established. Chapter 3 concludes with some examples that for instance illustrate the difficulty in finding explicit expressions for the orthonormal polynomials in non-trivial cases, or the difficulty to even generate an example from a given operator with explicit expressions for the measure.
In Chapter 4 some conditions for the (in)determinacy of the matrix moment problem are collected, either results taken from the literature or generalizations from [1].
Chapter 1

The Hamburger moment problem

This chapter is mostly based on [5], [6] and [19]. See Appendix A for a brief overview of complex measures.

1.1 Introduction

The (Hamburger) moment problem consists of the following two questions:

1. Given a sequence \((s_n)_{n \geq 0}\) in \(\mathbb{R}\), does there exist a positive Borel measure \(\mu\) on \(\mathbb{R}\) such that \(s_n = \int_{\mathbb{R}} x^n d\mu(x)\) for every \(n \geq 0\)?

2. If such a measure exists, is it unique?

Assumption 1.1. In this chapter we assume that \(\text{supp}(\mu)\) is infinite, and that \(\mu\) is not a finite discrete measure (see also Section 1.2).

Assumption 1.2. Without loss of generality we will always assume that \(s_0 = 1\). Note that this can be achieved by normalizing the involved measures to be probability measures.

If \(\mu\) is a positive measure on \(\mathbb{R}\) such that \(s_n = \int_{\mathbb{R}} x^n d\mu(x)\) for every \(n \geq 0\), we say that \(\mu\) is a solution to the moment problem of \((s_n)_{n \geq 0}\). If \(\mu\) is unique, we speak of a determinate moment problem. Otherwise the moment problem is called indeterminate. Observe that in the indeterminate case, there exists some convex set of probability measures on \(\mathbb{R}\) solving the moment problem.

The first question formulated above will be answered in Theorem 1.14, while Theorem 1.18 'solves' the second one.

1.1.1 Moment sequences and positive definiteness

Now we will describe the moment problem in some more detail by introducing some notation and definitions.

Definition 1.3. Let \(\mu\) be a positive Borel measure on \(\mathbb{R}\) with infinite support and finite moments of any order

\[ s_n := s_n(\mu) = \int_{\mathbb{R}} x^n d\mu(x). \tag{1.1} \]

We call \((s_n)_{n \geq 0}\) a moment sequence.
**Definition 1.4.** For a real sequence \((s_n)_{n \geq 0}\) we form for \(N \geq 0\) the so-called **Hankel matrices**

\[
H_N = (s_{i+j})_{0 \leq i,j \leq N}
\]

which are matrices of size \((N + 1) \times (N + 1)\). Written out we get

\[
\begin{pmatrix}
  s_0 & s_1 & s_2 & \cdots & s_N \\
  s_1 & s_2 & s_3 & \cdots & s_{N+1} \\
  s_2 & s_3 & s_4 & \cdots & s_{N+2} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  s_N & s_{N+1} & s_{N+2} & \cdots & s_{2N}
\end{pmatrix}
\]

Moreover we define the infinite Hankel matrix to be

\[
H_\infty = (s_{i+j})_{i,j \geq 0}.
\]

**Notation 1.5.** We denote by \(v \in \mathbb{C}^k\) a column vector with entries \(v_i\), where \(0 \leq i \leq k - 1\). Moreover \(v^*\) is the row vector we obtain by conjugating and transposing \(v\). With this notation, the inner product in \(\mathbb{C}^k\) is then defined to be

\[
\langle w, v \rangle = \sum_{i=0}^{k-1} \overline{w_i}v_i = v^*w
\]

for \(v, w \in \mathbb{C}^k\). Observe that the inner product defined above is indeed linear in the first slot, and antilinear in the second, which agrees with Definition B.1. The associated norm \(\| \cdot \|\) is defined by

\[
\|v\| = \sqrt{\langle v, v \rangle} = \sqrt{v^*v}.
\]

Similar notation is used for matrices: given \(A \in \mathbb{C}^{k \times k}\) we denote by \(A^*\) the matrix obtained by conjugating and transposing \(A\). Moreover we denote the zero matrix by \(\theta\) and the identity matrix by \(\mathbb{1}_k\).

For future purposes, we now define a norm on \(\mathbb{C}^{k \times k}\), given the vector norm on \(\mathbb{C}^k\).

**Definition 1.6.** Given the vector norm from Notation 1.5, we define

\[
\|A\| = \max_{v \neq 0} \frac{\|Av\|}{\|v\|} = \max_{\|v\|=1} \|Av\|
\]

for \(A \in \mathbb{C}^{k \times k}\). This **matrix norm** is said to be **induced** by the vector norm, and is alternatively called the **operator norm**.

This matrix norm satisfies \(\|Av\| \leq \|A\|\|v\|\) for all \(A \in \mathbb{C}^{K \times K}\) and \(v \in \mathbb{C}^k\), and \(\|\mathbb{1}_k\| = 1\) (see [16], Section 5.6, for an extensive treatment of matrix norms, and Theorem 5.6.2 for the proof of these statements in particular).

Before being able to give a characterization of moment sequences, we need the concept of positive hermitian matrices.
Definition 1.7. A matrix \( A \in \mathbb{C}^{k \times k} \), say \( A = (a_{ij}) \), is called positive hermitian if it is hermitian (i.e. \( A^* = A \)) and positive definite, i.e. if it is hermitian and
\[
\langle Av, v \rangle = v^* A v > 0, \text{ i.e. } \sum_{i,j=0}^{k-1} a_{ij} \overline{v_i} v_j > 0 \text{ for all } 0 \neq v \in \mathbb{C}^k.
\] (1.6)

In fact (1.6) implies that \( A \) is hermitian (see the proof below), so the former condition is actually not necessary in the above definition.

Proof. Assume that \( A \) is positive definite. Write \( A \) as
\[
B = \frac{A + A^*}{2} \quad \text{and} \quad C = \frac{A - A^*}{2i}.
\]
Then \( B \) and \( C \) are clearly hermitian. Now for all \( 0 \neq v \in \mathbb{C}^k \) we have
\[
0 < \langle Av, v \rangle = \langle Bh, v \rangle + i\langle Ch, v \rangle,
\] (1.7)
where both inner products on the right hand side are real. Therefore \( \langle Ch, v \rangle \) must be equal to 0, in order for the right hand side to be real. Since \( v \in \mathbb{C}^k \setminus \{0\} \) was chosen arbitrarily, \( C = 0 \). We conclude that \( A = B \) is hermitian.

On the set of hermitian matrices we can define a strict partial ordering (i.e. a relation that is irreflexive, transitive and antisymmetric) as follows. For two hermitian \( K \times K \) matrices \( A \) and \( B \) we write \( B \prec A \) if \( A - B \) is a positive hermitian matrix, i.e. if
\[
\langle Bv, v \rangle < \langle Av, v \rangle \text{ for all } 0 \neq v \in \mathbb{C}^K,
\] (1.8)
according to (1.6). Then the relation \( \prec \) is a strict partial ordering on the set of hermitian matrices.

Remark 1.8. The positive hermitian matrices are the matrices \( A \) for which \( \theta \prec A \). ■

In the following lemmas some properties of positive hermitian matrices are collected.

Lemma 1.9. The eigenvalues of a positive hermitian matrix are positive.

Proof. Let \( A \in \mathbb{C}^{k \times k} \) be a positive hermitian matrix and suppose that \( \lambda \) is an eigenvalue of \( A \) with corresponding eigenvector \( v \in \mathbb{C}^k \). Then by (1.6),
\[
0 < \langle Av, v \rangle = \lambda \langle v, v \rangle = \lambda |v|^2
\]
so that
\[
\lambda = \frac{\langle Av, v \rangle}{|v|^2} > 0
\]
as it is the ratio of two positive numbers. ■

9
Lemma 1.10. Let $A \in \mathbb{C}^{k \times k}$ be a positive hermitian matrix. Then $A \prec \text{tr}(A)1_k$, where $1_k$ denotes the $k \times k$ identity matrix.

Proof. Since $A$ is hermitian, it is diagonalizable\(^1\), which means that there exists an invertible matrix $B$ such that $A = B \text{ diag}(\lambda_1, \ldots, \lambda_k) B^{-1}$, where the $\lambda_i$ are the eigenvalues of $A$, which are all positive by Lemma 1.9. Denote the corresponding eigenvectors by $v_i$ and observe that $\{v_i : 1 \leq i \leq k\}$ is a basis for $\mathbb{C}^k$.

From positivity of the eigenvalues it is clear that

$$\langle Av_i, v \rangle = \langle \lambda_i v_i, v \rangle = \lambda_i \langle v_i, v \rangle < \sum_{j=1}^{k} \lambda_j \langle v_i, v \rangle = \text{tr}(A) \langle v_i, v \rangle = \langle \text{tr}(A) v_i, v \rangle$$

for all $0 \neq v \in \mathbb{C}^k$. Now write $v$ in terms of the basis, say $v = \sum_{i=1}^{k} c_i v_i$. Then

$$\langle Av, v \rangle = \sum_{i=1}^{k} c_i \langle Av_i, v \rangle < \sum_{i=1}^{k} c_i \langle \text{tr}(A) v_i, v \rangle = \langle \text{tr}(A)v, v \rangle,$$

so that $A \prec \text{tr}(A)1_k$. \(\square\)

Lemma 1.11. Let $A \in \mathbb{C}^{k \times k}$ be a regular and positive hermitian matrix. Then $A^{-1}$ is also positive hermitian.

Proof. Given $0 \neq v \in \mathbb{C}^k$, define $w = A^{-1}v \neq 0$. Then

$$\langle A^{-1}v, v \rangle = v^* A^{-1}v = (Aw)^* A^{-1} (Aw) = w^* A^* A^{-1} Aw = w^*Aw > 0,$$

hence $A^{-1}$ is positive hermitian. \(\square\)

Lemma 1.12. Let $A \in \mathbb{C}^{k \times k}$ be a positive hermitian matrix. Then for all $v \in \mathbb{C}^k$:

$$Av = 0 \iff v^* Av = 0.$$

Proof. If $Av = 0$, then obviously $v^* Av = 0$.

For the other implication, define $\langle v, w \rangle_A := \langle Av, w \rangle = w^* Av$. Then by the Cauchy-Schwarz inequality (see Appendix B.1.1) we have

$$|\langle v, w \rangle_A|^2 \leq \langle v, v \rangle_A \langle w, w \rangle_A.$$

If $0 = v^* Av = \langle v, v \rangle_A$, then $\langle v, w \rangle_A = 0$ for all $v \in \mathbb{C}^k$, hence $Av = 0$. \(\square\)

We will now give an answer of the first question formulated at the beginning of this chapter. For this reason we introduce the following definition:

Definition 1.13. A real sequence $(s_n)_{n \geq 0}$ is called positive definite if its Hankel matrix $H_N$ is positive hermitian for every $N$. Since $H_N$ is a real and symmetric matrix, $(s_n)_{n \geq 0}$ is positive definite if

$$\langle H_N v, v \rangle = \sum_{i,j=0}^{N} s_{i+j} v_i v_j > 0 \text{ for all } N \geq 0 \text{ and } 0 \neq v \in \mathbb{R}^{N+1}. \quad (1.9)$$

---

\(^1\)See [16], Theorem 4.1.5.
Theorem 1.14 (Hamburger). A sequence \((s_n)_{n \geq 0}\) is a moment sequence if and only if it is a positive definite sequence.

Proof. Suppose \((s_n)_{n \geq 0}\) is a moment sequence, i.e.

\[
s_n = \int_{\mathbb{R}} x^n d\mu(x) \tag{1.10}
\]

for some positive measure \(\mu\) on \(\mathbb{R}\). Then for any \(N \geq 0\) and \(v \in \mathbb{R}^{N+1} \setminus \{0\},
\[
\sum_{i,j=0}^{N} s_{i+j} v_i v_j = \sum_{i,j=0}^{N} \left( \int_{\mathbb{R}} x^{i+j} d\mu(x) \right) v_i v_j \\
= \int_{\mathbb{R}} \left( \sum_{i=0}^{N} v_i x^i \right) \left( \sum_{j=0}^{N} v_j x^j \right) d\mu(x) \\
= \int_{\mathbb{R}} \left( \sum_{i=0}^{N} v_i x^i \right)^2 d\mu(x) > 0. \tag{1.11}
\]

Hence \((s_n)_{n \geq 0}\) is a positive definite sequence.

The other implication is more complicated and can be found in [1], Theorem 2.1.1. Here we only sketch the main idea of the proof. Given the positive definite sequence \((s_n)_{n \geq 0}\), the so-called truncated moment problem of order \(2k-1\) is considered for every \(k \geq 0\). The problem is to find a positive measure \(\mu\) such that (1.10) holds for every \(0 \leq n \leq 2k - 1\). Then a sequence of positive measures \((\mu_k)_{k}\) is constructed, each of which is a solution of the truncated moment problem, in other words

\[
s_n = \int_{\mathbb{R}} x^n d\mu_k(x)
\]

for \(0 \leq n \leq 2k - 1\). Then, according to Helly’s theorem\(^2\), there exists a subsequence \((\mu_{k_i})_i\) of \((\mu_k)_k\) that converges to a measure \(\mu\) that in turn is a solution of the complete moment problem, i.e. \(\mu\) solves (1.10) for all \(n \geq 0\). It is thus shown that \((s_n)_{n \geq 0}\) is a moment sequence. \(\Box\)

1.1.2 Determinacy of the moment problem

In order to give an answer to the second question formulated at the beginning of this chapter, we consider the eigenvalues of the Hankel matrices. Given a positive measure \(\mu\), the associated Hankel matrices \(H_N\) are positive hermitian. Hence all eigenvalues of \(H_N\) are positive (see Lemma 1.9). Denote the smallest eigenvalue of \(H_N\) by \(\lambda_N\). Note that it can be obtained by the classical Rayleigh quotient, of which the definition is given below.

Definition 1.15. For a given hermitian matrix \(A \in \mathbb{C}^{k \times k}\) and a non-zero vector \(v \in \mathbb{C}^k\), the Rayleigh quotient \(R(A, v)\) is defined as

\[
R(A, v) = \langle Av, v \rangle = \frac{v^* A v}{v^* v}. \tag{1.12}
\]

\(^2\)This theorem is formulated in [11], Chapter II, Theorem 2.2 as follows: Let \((\varphi_n)\) be a uniformly bounded sequence of non-decreasing functions defined on \((-\infty, \infty)\). Then \((\varphi_n)\) has a subsequence which converges on \((-\infty, \infty)\) to a bounded, non-decreasing function.
Lemma 1.16. The Rayleigh quotient reaches its minimum value \( \lambda_{\text{min}} \) (i.e. the smallest eigenvalue of \( A \)) when \( v = v_{\text{min}} \) is the corresponding eigenvector.

Proof. Since \( A \) is hermitian, there is a unitary \( U \in \mathbb{C}^{k \times k} \) such that \( A = U \Lambda U^* \) where \( \Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k) \) and \( \lambda_{\text{min}} = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_k = \lambda_{\text{max}} \) are the eigenvalues of \( A \) (note that these are all real, hence such an ordering is possible). Then for all \( v \in \mathbb{C}^k \),

\[
v^* A v = v^* U \Lambda U^* v = (U^* v)^* \Lambda (U^* v) = \sum_{i=1}^{k} \lambda_i |(U^* v)_i|^2
\]

and thus (since \( U \) is unitary),

\[
\lambda_{\text{min}} v^* v = \lambda_{\text{min}} \sum_{i=1}^{k} |v_i|^2 = \lambda_{\text{min}} \sum_{i=1}^{k} |(U^* v)_i|^2 \leq \sum_{i=1}^{k} \lambda_i |(U^* v)_i|^2 = v^* A v.
\]

It follows that

\[
\lambda_{\text{min}} \leq \frac{v^* A v}{v^* v}
\]

for all \( v \in \mathbb{C}^k \setminus \{0\} \). Equality holds when \( v = v_{\text{min}} \), since in that case \( v^* A v = v^* \lambda_{\text{min}} v = \lambda_{\text{min}} v^* v \). We conclude that

\[
\lambda_{\text{min}} = \min_{0 \neq v \in \mathbb{C}^k} \frac{v^* A v}{v^* v}.
\]

Likewise it can be shown that the Rayleigh quotient attains its maximum value \( \lambda_{\text{max}} \) for \( v = v_{\text{max}} \).

We thus obtain the following expression\(^3\) for \( \lambda_N \), namely

\[
\lambda_N = \min_{0 \neq v \in \mathbb{C}^{N+1}} \frac{v^* H_N v}{v^* v} = \min_{v \in \mathbb{C}^{N+1}, \|v\| = 1} v^* H_N v = \min \left\{ \sum_{i=0}^{N} \sum_{j=0}^{N} s_{i+j \ell} v_i \mid \sum_{i=0}^{N} |v_i|^2 = 1, v_i \in \mathbb{C} \text{ for } 0 \leq i \leq N \right\}.
\]

(1.13)

From (1.13) it follows that \( \lambda_N \) is a decreasing function of \( N \). Indeed, suppose that \( v \in \mathbb{C}^N \), with \( \sum_{i=0}^{N} |v_i|^2 = 1 \), minimizes \( \sum_{i=0}^{N} \sum_{j=0}^{N} s_{i+j \ell} v_i \). Then construct \( w \in \mathbb{C}^{N+1} \) by putting \( w_i = v_i \) for \( 1 \leq i \leq N \), and \( w_{N+1} = 0 \). Then clearly \( \sum_{i=0}^{N+1} |w_i|^2 = \sum_{i=0}^{N} |v_i|^2 = 1 \), so that \( w \) lies in the set over which the minimum is taken while computing \( \lambda_{N+1} \). Then \( w \) could either minimize \( \sum_{i=0}^{N+1} \sum_{j=0}^{N+1} s_{i+j \ell} v_i \), or there could be another vector in \( \mathbb{C}^{N+1} \) for which this quantity is smaller. We conclude that \( \lambda_{N+1} \leq \lambda_N \).

\(^3\)As \( H_N \) is a real and symmetric matrix, it is actually sufficient to take the minimum over real vectors, i.e.

\[
\lambda_N = \min \left\{ \sum_{i=0}^{N} \sum_{j=0}^{N} s_{i+j \ell} v_i \mid \sum_{i=0}^{N} v_i^2 = 1, v_i \in \mathbb{R} \text{ for } 0 \leq i \leq N \right\}.
\]

In order to give the proof of Theorem 1.18 however, it is convenient to the more general expression given in (1.13).
Remark 1.17. If $\lambda_N = 0$ for some $N \geq 0$, then also $\lambda_n = 0$ for all $n \geq N$, and in that case $\mu$ is a finite sum of point masses. This case implies determinacy since any measure with compact support is automatically determinate. We, however, exclude this case. Whenever $\mu$ has infinite support, it holds that $\lambda_N > 0$ for all $N \geq 0$, which agrees with the remark made earlier that all eigenvalues of $H_N$ are positive. ■

The following theorem is the main result of [6], and gives a condition for the determinacy of the moment problem.

**Theorem 1.18.** The moment problem associated with the moments given by (1.1) is determinate if and only if $\lim_{N \to \infty} \lambda_N = 0$.

We will postpone the proof of this theorem until we have developed some more machinery (see Section 1.3). In the proof of this theorem, the reciprocal of $\lambda_N$ will turn out to play a role.

Note that it can be written as

$$\frac{1}{\lambda_N} = \max_{v \in \mathbb{C}^{N+1}} \frac{v^* v}{v^* H_N v} = \max_{v \in \mathbb{C}^{N+1}, (H_N v, v) = 1} v^* v = \max \left\{ \sum_{i=0}^{N} |v_i|^2 : \sum_{i=0}^{N} \sum_{j=0}^{N} s_{i+j} v_i v_j = 1, v_i \in \mathbb{C} \text{ for } 0 \leq i \leq N \right\}. \quad (1.14)$$

### 1.1.3 Example due to Stieltjes

Here we work out the moment problem for an explicit measure, originally studied by Stieltjes. Consider the measure $\mu$ on $\mathbb{R}$ given by

$$d\mu(x) = C_{\alpha, \gamma} e^{-\gamma |x|^\alpha} dx$$

where it is assumed that $\alpha, \gamma > 0$ (note that Stieltjes considered the case $\gamma = 1$), and $C_{\alpha, \gamma}$ is a constant depending on these two parameters. The corresponding moments are thus given by

$$s_n = \int_{\mathbb{R}} x^n d\mu(x) = C_{\alpha, \gamma} \int_{\mathbb{R}} x^n e^{-\gamma |x|^\alpha} dx. \quad (1.16)$$

Both the moments and the value of $C_{\alpha, \gamma}$ can be determined by exploiting the equality

$$\int_0^\infty x^{c-1} e^{-b x} dx = b^{-c} \Gamma(c) \quad (1.17)$$

which holds true\(^4\) for $\text{Re}(b) > 0$ and $c > 0$. Here $\Gamma$ is the well-known Gamma function, defined for all $z \in \mathbb{C}$ with $\text{Re}(z) > 0$ by

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

Then

$$1 = s_0 = C_{\alpha, \gamma} \int_{\mathbb{R}} e^{-\gamma |x|^\alpha} dx = 2 C_{\alpha, \gamma} \int_0^\infty e^{-\gamma x^\alpha} dx$$

Indeed, making use of the substitution $bx = y$, we obtain

$$\int_0^\infty x^{c-1} e^{-b x} dx = \int_0^\infty \left( \frac{y}{b} \right)^{c-1} e^{-y} \frac{dy}{b} = b^{-c} \int_0^\infty y^{c-1} e^{-y} dy = b^{-c} \Gamma(c).$$

13
where, substituting $y = x^\alpha$, so that $dy = \alpha x^{\alpha - 1} dx$, and applying (1.17) for $c = \frac{1}{\alpha}$ and $b = \gamma$,

$$
\int_0^\infty e^{-y x^\alpha} dx = \int_0^\infty e^{-y} \frac{1}{\alpha y} dy = \frac{1}{\alpha} \int_0^\infty y^{\frac{1}{\alpha}-1} e^{-y} dy = \frac{1}{\alpha} \gamma^{-\frac{1}{\alpha}} \Gamma \left( \frac{1}{\alpha} \right).
$$

We conclude that $C_{\alpha, \gamma} = \frac{\alpha}{2} \gamma^{-\frac{1}{\alpha}} \left( \Gamma \left( \frac{1}{\alpha} \right) \right)^{-1}$. In a similar way the moments $s_n$ can be calculated. First note that all odd moments vanish, since the integral

$$
\int_0^\infty e^{-y} \frac{1}{\alpha y} dy = \frac{1}{\alpha} \int_0^\infty y^{\frac{1}{\alpha}-1} e^{-y} dy
$$

has an odd integrand for every $n \in \mathbb{N}$. The even moments can be found by applying (1.17) with $c = \frac{2n+1}{\alpha}$ and $b = \gamma$, since

$$
s_{2n+1} = C_{\alpha, \gamma} \int_\mathbb{R} x^{2n+1} e^{-\gamma |x|^\alpha} dx
$$

has a odd integrand for every $n \in \mathbb{N}$. The even moments can be found by applying (1.17) with $c = \frac{2n+1}{\alpha}$ and $b = \gamma$.

$$
s_{2n} = C_{\alpha, \gamma} \int_\mathbb{R} x^{2n} e^{-\gamma |x|^\alpha} dx = 2C_{\alpha, \gamma} \int_0^\infty x^{2n} e^{-\gamma x^n} dx
$$

First we prove the indeterminacy whenever $0 < \alpha < 1$. Rewriting (1.17) with $x = y^\alpha$ and $c = \frac{2n+1}{\alpha}$ we obtain

$$
b^{-\frac{2n+1}{\alpha}} \Gamma \left( \frac{2n+1}{\alpha} \right) = b^{-c} \Gamma (c) = \int_0^\infty x^{c-1} e^{-bx} dx = \frac{1}{2} \int_\mathbb{R} |x|^{c-1} e^{-b|x|} dx = \frac{1}{2} \int_\mathbb{R} |y|^{2n+1-\alpha} e^{-b|y|^\alpha} \alpha y^{\alpha-1} dy = \frac{\alpha}{2} \int_\mathbb{R} y^{2n} e^{-b|y|^\alpha} dy
$$

As

$$
\text{Re} \left( b^{-\frac{2n+1}{\alpha}} \right) = \text{Re} \left( \left| b \right| e^{i \arg(b)}^{-\frac{2n+1}{\alpha}} \right) = \left| b \right|^{-\frac{2n+1}{\alpha}} \cos \left( - \arg(b) \frac{2n+1}{\alpha} \right)
$$

and

$$
\text{Re} \left( e^{-b|y|^\alpha} \right) = \text{Re} \left( e^{-\left( \text{Re}(b) + i \text{Im}(b) \right)|y|^\alpha} \right) = e^{-\text{Re}(b)|y|^\alpha} \cos \left( -\text{Im}(b) |y|^\alpha \right)
$$

it follows by taking the real parts on both sides of (1.18) that

$$
\frac{\alpha}{2} \int_\mathbb{R} y^{2n} e^{-\text{Re}(b)|y|^\alpha} \cos \left( -\text{Im}(b) |y|^\alpha \right) dy = \left| b \right|^{-\frac{2n+1}{\alpha}} \cos \left( - \arg(b) \frac{2n+1}{\alpha} \right) \Gamma \left( \frac{2n+1}{\alpha} \right).
$$

Thus the right-hand side is equal to 0 for all $n \in \mathbb{N}$ whenever $\arg(b) = \frac{1}{2} \alpha \pi$. Since $\text{Re}(b) > 0$, this is the case for $|\alpha| < 1$, and because we assumed $\alpha > 0$, it in fact holds for $0 < \alpha < 1$. But every such value of $\alpha$ gives rise to a measure $\mu_{\alpha,b}$ different from $\mu$ that has the same moments, namely

$$
d\mu_{\alpha,b}(x) = C_{\alpha, \gamma} \left( e^{-\gamma |x|^\alpha} + e^{-\text{Re}(b) |x|^\alpha} \cos \left( -\text{Im}(b) |x|^\alpha \right) \right)
$$

where $b \in \mathbb{C}$ is chosen such that $\arg(b) = \frac{1}{2} \alpha \pi$. We thus conclude that the moment problem is indeterminate whenever $0 < \alpha < 1$. 


Now we consider the case in which \( \alpha \geq 1 \). It is known that moment problems for which the even moments satisfy
\[
\sum_{n=1}^{\infty} s_{2n}^{-\frac{1}{2n}} = \infty,
\]
are determinate (this result is due to Carleman, see [1], Chapter 2, Problem 11). Here this is also the case, since
\[
\sum_{n=1}^{\infty} s_{2n}^{-\frac{1}{2n}} = \sum_{n=1}^{\infty} \gamma_n^{-\frac{1}{\alpha}} \left( \frac{\Gamma \left( \frac{2n+1}{\alpha} \right)}{\Gamma \left( \frac{1}{\alpha} \right)} \right)^{-\frac{1}{2n}} = \infty. \tag{1.19}
\]
This follows by the following observations. According to [28], Section 12.33, the logarithm of the Gamma function can be approximated as
\[
\log \Gamma(z) \sim z \log z - z + \frac{1}{2} \log \frac{2\pi}{z}
\]
for \(|z| \to \infty\) and \(|\arg(z)| < \pi\). Then under these conditions we also have
\[
\Gamma(z) \sim \frac{2\pi}{z} \left( \frac{z}{e} \right)^z,
\]
which is known as Stirling’s formula. Then, for large enough \( n \), we have
\[
\Gamma \left( \frac{2n+1}{\alpha} \right)^{-\frac{1}{2n}} \sim \left( \sqrt{\frac{2\pi \alpha}{2n+1}} \left( \frac{2n+1}{\alpha e} \right)^{\frac{2n+1}{2n+1}} \right)^{-\frac{1}{2n}} = \left( \frac{2n+1}{2n} \right)^{\frac{1}{2n}} \left( \frac{\alpha e}{2n+1} \right)^{\frac{2n+1}{2n}}.
\tag{1.20}
\]
Since \( \alpha \geq 1 \),
\[
\frac{2n+1}{2n} \leq \frac{2n+1}{2n} = 1 + \frac{1}{2n}
\]
and thus
\[
\frac{1}{4n} - \frac{2n+1}{2n} \geq \frac{1}{4n} - \left( 1 + \frac{1}{2n} \right) = -\frac{1}{4n} - 1.
\]
Thus (again for large enough \( n \)), (1.20) can be further approximated as
\[
\Gamma \left( \frac{2n+1}{\alpha} \right)^{-\frac{1}{2n}} \geq \alpha e(2n+1)^{-\frac{1}{4n}}.
\]
As \( \Gamma \left( \frac{1}{\alpha} \right) \) is constant, (1.19) now follows, since
\[
\sum_{n=1}^{\infty} n^{-\frac{1}{2n}} = \infty.
\]
Thus the determinacy of the moment problem with \( \alpha \geq 1 \) has been proved.
1.2 Orthonormal polynomials

As said before, we need some more machinery in order to give a proof of Theorem 1.18. In this section we therefore introduce orthonormal polynomials, which will play a central role in the treatment of the moment problem.

1.2.1 Construction and properties of orthonormal polynomials

Consider $L^2(\mu)$, the space of square integrable functions on $\mathbb{R}$ with respect to $\mu$, i.e. $f \in L^2(\mu)$ if and only if

$$\int_{\mathbb{R}} |f(x)|^2 d\mu(x) < \infty.$$ (1.21)

Note that $L^2(\mu)$ is a Hilbert space (see Definition B.2) after identifying two functions $f$ and $g$ for which $\int_{\mathbb{R}} |f(x) - g(x)|^2 d\mu(x) = 0$, with respect to the inner product

$$\langle f, g \rangle_{L^2(\mu)} := \int_{\mathbb{R}} f(x)\overline{g(x)} d\mu(x).$$ (1.22)

Assume that all moments of $\mu$ exist, so that all polynomials are integrable. In applying the Gram-Schmidt orthogonalisation process to the sequence $\{1, x, x^2, x^3, \ldots\}$ we may end up in one of the following two situations:

1. The polynomials are linearly dependent in $L^2(\mu)$. Then there is a non-zero polynomial $p$ such that $\int_{\mathbb{R}} |p(x)|^2 d\mu(x) = 0$. This implies that $\mu$ is a finite sum of Dirac measures at the zeros of $p$. We will exclude this case.

2. The polynomials are linearly independent. We then end up with a set of orthonormal polynomials as in the definition that follows below. Note that these polynomials form a basis of the vector space $\mathbb{C}[x]$.

Observe that the polynomials $p_n$ are real-valued for $x \in \mathbb{R}$, so that their coefficients are real. Moreover it follows from the Gram-Schmidt process that the leading coefficients are positive.

**Definition 1.19.** A sequence of polynomials $(p_n)_{n=0}^{\infty}$ with $\deg(p_n) = n$ for every $n \in \mathbb{N}$ is a set of orthonormal polynomials with respect to $\mu$ if

$$\langle p_n, p_m \rangle = \int_{\mathbb{R}} p_n(x)\overline{p_m(x)} d\mu(x) = \int_{\mathbb{R}} p_n(x)p_m(x) d\mu(x) = \delta_{n,m}. $$ (1.23)

From (2.27) it follows that

$$\|p_n\|_{L^2(\mu)} = \langle p_n, p_n \rangle^{\frac{1}{2}} = \left(\int_{\mathbb{R}} |p_n(x)|^2 d\mu(x)\right)^{\frac{1}{2}} = 1.$$ (1.24)

**Remark 1.20.** The orthonormal polynomials $(p_n)_{n=0}^{\infty}$ for $\mu$ are uniquely determined if we require the polynomials to satisfy (2.27) and $p_n$ is a polynomial of degree $n$ with positive leading coefficient.

---

5To be more precise, we let $L^2(\mu)$ be the space of square integrable functions, i.e.

$$L^2(\mu) = \left\{ f : \int_{\mathbb{R}} |f(x)|^2 d\mu(x) < \infty \right\}$$

and then define $L^2(\mu) = L^2(\mu) / \sim$, where $f \sim g$ if and only if $\int_{\mathbb{R}} |f(x) - g(x)|^2 d\mu(x) = 0$.  

16
The following theorem describes a fundamental property of the orthonormal polynomials.

**Theorem 1.21** (Three-term recurrence relation). Let \((p_n)_{n=0}^{\infty}\) be a set of orthonormal polynomials in \(L^2(\mu)\). Then there exist sequences \((a_n)_{n=0}^{\infty}\), \((b_n)_{n=0}^{\infty}\) with \(a_n > 0\) and \(b_n \in \mathbb{R}\) for every \(n \in \mathbb{N}\), such that

\[
 xp_n(x) = a_{n+1}p_{n+1}(x) + b_np_n(x) + a_np_{n-1}(x), \quad \text{for } n \geq 1,
\]

(1.25)

\[
 xp_0(x) = a_1p_1(x) + b_0p_0(x).
\]

(1.26)

Moreover, if \(\mu\) has compact support, then the coefficients \(a_n\) and \(b_n\) are bounded.

**Proof.** Since the polynomial \(xp_n(x)\) has degree \(n+1\), it can be written as

\[
 xp_n(x) = \sum_{i=0}^{n+1} c_ip_i(x)
\]

(1.27)

for certain constants \(c_i\). By the orthonormality relation (2.27),

\[
 \int_{\mathbb{R}} p_i(x)xp_n(x)d\mu(x) = \int_{\mathbb{R}} p_i(x)\left(\sum_{j=0}^{n+1} c jp_j(x)\right) d\mu(x)
\]

\[
 = \sum_{j=0}^{n+1} c_j \int_{\mathbb{R}} p_i(x)p_j(x)d\mu(x)
\]

\[
 = \sum_{j=0}^{n+1} c_j \delta_{i,j} = c_i,
\]

(1.28)

hence

\[
 c_i = \int_{\mathbb{R}} xp_i(x)p_n(x)d\mu(x).
\]

(1.29)

The polynomial \(xp_i(x)\) has degree \(i + 1\), so that the above equation implies that \(c_i = 0\) for \(i + 1 < k\), i.e. for \(i < k - 1\). Thus (1.27) can be rewritten as

\[
 xp_n(x) = c_{n+1}p_{n+1}(x) + c_np_n(x) + c_{n-1}p_{n-1}.
\]

(1.30)

Now define

\[
 a_n = \int_{\mathbb{R}} p_{n-1}(x)xp_n(x)d\mu(x),
\]

(1.31)

\[
 b_n = \int_{\mathbb{R}} x(p_n(x))^2d\mu(x).
\]

(1.32)

Then \(a_n = c_{n-1}\),

\[
 a_{n+1} = \int_{\mathbb{R}} p_n(x)xp_{n+1}(x)d\mu(x) = \int_{\mathbb{R}} xp_{n+1}(x)p_n(x)d\mu(x) = c_{n+1}
\]

and \(b_n = c_n\). Moreover it is clear that \(b_n \in \mathbb{R}\). Denote the leading coefficient of \(p_n(x)\) by \(l_n\), which is positive for all \(n\). Then comparing the terms of order \(n+1\) in (1.25) yields \(l_n = a_{n+1}/b_{n+1}\), hence \(a_{n+1} = (l_n/b_{n+1})\) is positive.

Now suppose that \(\text{supp}(\mu)\) is compact. Then

\[
 |a_n| = \left|\int_{\mathbb{R}} xp_{n-1}(x)p_n(x)d\mu(x)\right| \leq \sup_{x \in \text{supp}(\mu)} |x| \int_{\mathbb{R}} |p_{n-1}(x)||p_n(x)|d\mu(x)
\]

\[
 \leq \sup_{x \in \text{supp}(\mu)} |x| \|p_{n-1}\|_{L^2(\mu)} \|p_n\|_{L^2(\mu)} = \sup_{x \in \text{supp}(\mu)} |x| < \infty,
\]

(1.33)
where we made use of the fact that $\|p_n\|_{L^2(\mu)} = 1$ by orthonormality. Moreover in the second inequality we applied the Cauchy-Schwarz inequality (see Appendix B.1). Likewise

$$|b_n| = \left| \int_\mathbb{R} x(p_n(x))^2 d\mu(x) \right| \leq \sup_{x \in \text{supp}(\mu)} |x| \int_\mathbb{R} |p_n(x)|^2 d\mu(x)$$

$$\leq \sup_{x \in \text{supp}(\mu)} |x| \cdot \|p_n\|_{L^2(\mu)}^2 = \sup_{x \in \text{supp}(\mu)} |x| < \infty.$$  \quad (1.34)

Thus, if $\mu$ is compactly supported, then the coefficients $a_n$ and $b_n$ are bounded.

Note that (1.25) and (1.26) together with the initial condition $p_0(x) = 1$ completely determine the polynomials $p_n$ for all $n \in \mathbb{N}$.

The orthonormal polynomials are also useful to characterize the determinacy of the associated moment problem, as the following remark illustrates.

**Remark 1.22.** The moment problem is indeterminate if and only if there exists a non-real number $z_0$ such that

$$\sum_{n=0}^{\infty} |p_n(z_0)|^2 < \infty.$$  \quad (1.35)

In the indeterminate case the series (1.35) actually converges for all $z_0 \in \mathbb{C}$, uniformly on compact sets. In the determinate case the series in (1.35) diverges for all non-real $z_0$ and also for all real numbers except the at most countably many points where $\mu$ has a positive mass. The proof of these statements will be omitted. In Chapter 3 we will look at the general (i.e. matrix-valued) case, from which the above statement can be derived (see Remark 3.15).

### 1.2.2 The kernel polynomial

We will now introduce the kernel polynomial, which is defined in terms of the orthonormal polynomials discussed in the previous subsection. Moreover we will see the connection between the kernel polynomial and the Hankel matrices. We won’t make explicit use of the results stated in this section; in Section 2.3.4 however we will generalize the concept of the kernel polynomial to the matrix-valued case, and this turns out to be a handy tool in discussing the matrix moment problem.

**Definition 1.23.** The reproducing kernel for the polynomials of degree \( \leq N \) is defined as

$$K_N(x, y) = \sum_{k=0}^{N} p_k(x)p_k(y)$$  \quad (1.36)

and is called the *kernel polynomial*.

**Remark 1.24.** Note that

$$\int_\mathbb{R} p(x)K_N(x, y)d\mu(x) = p(y)$$  \quad (1.37)
for any polynomial $p$ of degree $\leq N$. This is an immediate consequence of the fact that

$$\int_{\mathbb{R}} p_k(x) K_N(x,y) d\mu(x) = \int_{\mathbb{R}} p_k(x) \left( \sum_{l=0}^{N} p_l(x)p_l(y) \right) d\mu(x)$$

$$= \sum_{l=0}^{N} \left( \int_{\mathbb{R}} p_k(x)p_l(x) d\mu(x) \right) p_l(y)$$

$$= \sum_{l=0}^{N} \delta_{k,l} p_l(y) = p_k(y)$$

(1.38)

for $0 \leq k \leq N$.

Observe that the kernel polynomial can alternatively be written as

$$K_N(x,y) = \sum_{i=0}^{N} \sum_{j=0}^{N} a_{ij}^{(N)} x^i y^j$$

(1.39)

where the numbers $a_{ij}^{(N)}$ are uniquely determined and satisfy $a_{ij}^{(N)} = a_{ji}^{(N)}$. This can be shown by writing $p_k(x) = \sum_{i=0}^{k} b_{i}^{(k)} x^i$ for certain $b_{i}^{(k)} \in \mathbb{C}$. Then

$$K_N(x,y) = \sum_{k=0}^{N} \left( \sum_{i=0}^{k} b_{i}^{(k)} x^i \sum_{j=0}^{k} b_{j}^{(k)} y^j \right)$$

$$= \sum_{k=0}^{N} \sum_{i=0}^{k} \sum_{j=0}^{k} b_{i}^{(k)} b_{j}^{(k)} x^i y^j$$

$$= \sum_{i=0}^{N} \sum_{j=0}^{N} a_{ij}^{(N)} x^i y^j$$

(1.40)

where $a_{ij}^{(N)}$ satisfies

$$a_{ij}^{(N)} = \sum_{k=\max(i,j)}^{N} b_{i}^{(k)} b_{j}^{(k)}$$

(1.41)

so that $a_{ij}^{(N)} = a_{ji}^{(N)}$ clearly holds.

Now define the matrix $A_N \in \mathbb{C}^{(N+1)\times(N+1)}$ by $A_N = \left( a_{ij}^{(N)} \right)_{0 \leq i,j \leq N}$. Then $A_N$ is the inverse of the Hankel matrix $H_N$ as the following theorem shows (see also [4], Theorem 2.1).

**Theorem 1.25.** The matrix $A_N$ is the inverse of $H_N$, i.e.

$$A_N H_N = \mathbb{1}_{N+1} = H_N A_N$$

(1.42)

where $\mathbb{1}_{N+1}$ is the $(N+1) \times (N+1)$ unit matrix.

**Proof.** For $0 \leq k \leq N$ it follows that

$$\int_{\mathbb{R}} x^k K_N(x,y) d\mu(x) = y^k$$

(1.43)

19
by Remark 1.24. On the other hand we have

$$\int_{\mathbb{R}} x^k K_N(x,y) d\mu(x) = \int_{\mathbb{R}} x^k \left( \sum_{i=0}^{N} \sum_{j=0}^{N} a_{ij}^{(N)} x^i y^j \right) d\mu(x)$$

$$= \sum_{i=0}^{N} \sum_{j=0}^{N} a_{ij}^{(N)} \left( \int_{\mathbb{R}} x^{k+i} d\mu(x) \right) y^j$$

$$= \sum_{j=0}^{N} \left( \sum_{i=0}^{N} s_{k+i} a_{ij}^{(N)} \right) y^j. \quad (1.44)$$

Combining (1.43) and (1.44), we see that

$$y^k = \sum_{j=0}^{N} \left( \sum_{i=0}^{N} s_{k+i} a_{ij}^{(N)} \right) y^j \quad (1.45)$$

so that

$$\sum_{i=0}^{N} s_{k+i} a_{ij}^{(N)} = \delta_{k,j} \quad (1.46)$$

We thus have shown that $H_N A_N = I_{N+1}$. By uniqueness of the inverse, the claim follows.

The kernel polynomial can also be directly written in terms of the moments and the Hankel matrix, as the following lemma shows.

**Lemma 1.26.** The kernel polynomial is equal to

$$K_N(x,y) = - \det(H_N)^{-1} \det \begin{pmatrix} s_0 & s_1 & \cdots & s_N & 1 \\ s_1 & s_2 & \cdots & s_{N+1} & x \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ s_N & s_{N+1} & \cdots & s_{2N} & x^N \\ 1 & y & \cdots & y^N & 0 \end{pmatrix}. \quad (1.47)$$

Note that $\det(H_N) > 0$ as $H_N$ is positive hermitian, so the above expression is well-defined.

**Proof.** We know that the kernel polynomial satisfies (1.43). Hence the claim follows whenever the same holds for the right-hand side of (1.47).
Note that

\[
\int_{\mathbb{R}} x^k \det \begin{pmatrix} s_0 & s_1 & \ldots & s_N & 1 \\ s_1 & s_2 & \ldots & s_{N+1} & x \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ s_N & s_{N+1} & \ldots & s_{2N} & x^N \\ 1 & y & \ldots & y^N & 0 \end{pmatrix} \mathrm{d}\mu(x) = \int_{\mathbb{R}} \det \begin{pmatrix} s_0 & s_1 & \ldots & s_N & x^k \\ s_1 & s_2 & \ldots & s_{N+1} & x^{k+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ s_N & s_{N+1} & \ldots & s_{2N} & x^{k+N} \\ 1 & y & \ldots & y^N & 0 \end{pmatrix} \mathrm{d}\mu(x)
\]

\[
= \det \begin{pmatrix} s_0 & s_1 & \ldots & s_k & \ldots & s_N & s_N \\ s_1 & s_2 & \ldots & s_{k+1} & \ldots & s_{N+1} & s_{k+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ s_N & s_{N+1} & \ldots & s_{k+N} & \ldots & s_{2N} & s_{k+N} \\ 1 & y & \ldots & y^k & \ldots & y^N & 0 \end{pmatrix}
\]

\[
= \det \begin{pmatrix} s_0 & s_1 & \ldots & 0 & \ldots & s_N & s_k \\ s_1 & s_2 & \ldots & 0 & \ldots & s_{N+1} & s_{k+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ s_N & s_{N+1} & \ldots & 0 & \ldots & s_{2N} & s_{k+N} \\ 1 & y & \ldots & y^k & \ldots & y^N & 0 \end{pmatrix}
\]

\[
= -\det \begin{pmatrix} s_0 & s_1 & \ldots & s_N & 0 \\ s_1 & s_2 & \ldots & s_{N+1} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ s_N & s_{N+1} & \ldots & s_{2N} & 0 \\ 1 & y & \ldots & y^N & y^k \end{pmatrix}
\]

\[
= -(-1)^{N+1}(-1)^{N+1} y^k \det \begin{pmatrix} s_0 & s_1 & \ldots & s_N \\ s_1 & s_2 & \ldots & s_{N+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_N & s_{N+1} & \ldots & s_{2N} \end{pmatrix} = -y^k \det(H_N).
\]

Here we subtracted column \(N+2\) from column \(k+1\) at the third equality, and then interchanged them at the fourth equality. Dividing by \(-\det(H_N)\) on both sides yields the desired result.

\[\square\]

### 1.3 Proof of Theorem 1.18

By using the machinery of Section 1.2, we are able to give a proof of Theorem 1.18 (see also Section 1 and 2 from [6]), which states that the scalar moment problem is determinate if and only if the smallest eigenvalues of the Hankel matrices tend to zero whenever \(N \to \infty\).

**Proof.** Define

\[
\pi_N(x) = \sum_{j=0}^{N} v_j x^j
\]

(1.48)
where \( v_j \in \mathbb{C} \). Then

\[
\int_{\mathbb{R}} |\pi_N(x)|^2 d\mu(x) = \int_{\mathbb{R}} \pi_N(x) \overline{\pi_N(x)} d\mu(x) \\
= \int_{\mathbb{R}} \left( \sum_{j=0}^{N} v_j x^j \right) \left( \sum_{k=0}^{N} v_k x^k \right) d\mu(x) \\
= \int_{\mathbb{R}} \sum_{j=0}^{N} \sum_{k=0}^{N} x^{j+k} v_j v_k d\mu(x) \\
= \sum_{j=0}^{N} \sum_{k=0}^{N} \left( \int_{\mathbb{R}} x^{j+k} d\mu(x) \right) v_j v_k \\
= \sum_{j=0}^{N} \sum_{k=0}^{N} s_{j+k} v_j v_k 
\]

and

\[
\int_{0}^{2\pi} \left| \pi_N \left( e^{i\theta} \right) \right|^2 \frac{d\theta}{2\pi} = \int_{0}^{2\pi} \frac{\pi_N \left( e^{i\theta} \right) \overline{\pi_N \left( e^{i\theta} \right)} d\theta}{2\pi} \\
= \int_{0}^{2\pi} \sum_{j=0}^{N} \sum_{k=0}^{N} \overline{v_j} v_k e^{-i(j-k)} d\theta \\
= \sum_{j=0}^{N} \sum_{k=0}^{N} \overline{v_j} v_k \int_{0}^{2\pi} e^{-i(j-k)} \frac{d\theta}{2\pi} \\
= \sum_{j=0}^{N} \sum_{k=0}^{N} s_{j+k} v_j v_k = \sum_{k=0}^{N} |v_k|^2. 
\]

By (1.13), (1.49) and (1.50) it follows that the smallest eigenvalue \( \lambda_N \) of \( H_N \) is determined by

\[
\lambda_N = \min_{\pi_N} \left\{ \int_{\mathbb{R}} |\pi_N(x)|^2 d\mu(x) : \int_{0}^{2\pi} \left| \pi_N \left( e^{i\theta} \right) \right|^2 \frac{d\theta}{2\pi} = 1 \right\}. 
\]

The reciprocal of \( \lambda_N \) then is equal to

\[
\frac{1}{\lambda_N} = \max_{\pi_N} \left\{ \int_{0}^{2\pi} \left| \pi_N \left( e^{i\theta} \right) \right|^2 \frac{d\theta}{2\pi} : \int_{\mathbb{R}} |\pi_N(x)|^2 d\mu(x) = 1 \right\}. 
\]

Let \( (p_n)_{n=0}^{\infty} \) denote the orthonormal polynomials with respect to \( \mu \), so that (2.27) is satisfied. Moreover \( p_n \) has a positive leading coefficient. As \( (p_n)_{n=0}^{\infty} \) forms a basis, we can write \( \pi_N(x) \) as a linear combination of the orthonormal polynomials \( p_k \), say

\[
\pi_N(x) = \sum_{j=0}^{N} c_j p_j(x), 
\]

where \( c_j \in \mathbb{C} \). We will now rewrite the integrals that appear in both (1.51) and (1.52) by using
We obtain
\[
\int_{0}^{2\pi} \left| \pi_N(e^{i\theta}) \right|^2 \frac{d\theta}{2\pi} = \int_{0}^{2\pi} \left( \sum_{j=0}^{N} c_j p_j(e^{-i\theta}) \right) \left( \sum_{k=0}^{N} c_k p_k(e^{i\theta}) \right) \frac{d\theta}{2\pi}
\]
\[
= \sum_{j=0}^{N} \sum_{k=0}^{N} c_j c_k \int_{0}^{2\pi} p_j(e^{-i\theta}) p_k(e^{i\theta}) \frac{d\theta}{2\pi}
\]
\[
= \sum_{j=0}^{N} \sum_{k=0}^{N} K_{jk} c_j c_k \tag{1.54}
\]
where we have defined
\[
K_{jk} = \int_{0}^{2\pi} p_j(e^{-i\theta}) p_k(e^{i\theta}) \frac{d\theta}{2\pi}. \tag{1.55}
\]
Similarly
\[
\int_{\mathbb{R}} |\pi_N(x)|^2 d\mu(x) = \int_{\mathbb{R}} \left( \sum_{j=0}^{N} c_j p_j(x) \right) \left( \sum_{k=0}^{N} c_k p_k(x) \right) d\mu(x)
\]
\[
= \sum_{j=0}^{N} \sum_{k=0}^{N} c_j c_k \int_{\mathbb{R}} p_j(x) p_k(x) d\mu(x)
\]
\[
= \sum_{j=0}^{N} \sum_{k=0}^{N} c_j c_k \delta_{jk} = \sum_{j=0}^{N} |c_j|^2. \tag{1.56}
\]
Thus (1.52) can be rewritten as
\[
\frac{1}{\lambda_N} = \max_{c_j} \left\{ \sum_{j=0}^{N} \sum_{k=0}^{N} K_{jk} c_j c_k : \sum_{j=0}^{N} |c_j|^2 = 1 \right\}. \tag{1.57}
\]
Since the matrix $K_N := (K_{jk})_{0 \leq j, k \leq N}$ is positive definite\(^6\), all its eigenvalues are positive, and the sum of these eigenvalues is equal to the trace of the matrix. Note that (1.57) implies that $\frac{1}{\lambda_N}$ is the largest eigenvalue of $K_N$. Hence we obtain the inequality
\[
\frac{1}{\lambda_N} \leq \text{tr}(K_N) = \sum_{k=0}^{N} K_{kk} = \sum_{k=0}^{N} \int_{0}^{2\pi} p_k(e^{-i\theta}) p_k(e^{i\theta}) \frac{d\theta}{2\pi} = \int_{0}^{2\pi} \sum_{k=0}^{N} |p_k(e^{i\theta})|^2 \frac{d\theta}{2\pi}. \tag{1.58}
\]
According to Remark 1.22 it holds that
\[
\sum_{k=0}^{\infty} |p_k(e^{i\theta})|^2 < \infty. \tag{1.59}
\]
whenever the moment problem is indeterminate. Thus in the case of indeterminacy it follows from (1.58) and (1.59) that
\[
\frac{1}{\lambda_N} \leq \int_{0}^{2\pi} \sum_{k=0}^{N} |p_k(e^{i\theta})|^2 \frac{d\theta}{2\pi} \leq \int_{0}^{2\pi} \sum_{k=0}^{\infty} |p_k(e^{i\theta})|^2 \frac{d\theta}{2\pi} < \infty. \tag{1.60}
\]
\(^6\)This immediately follows from (1.54), since the left hand side is $\geq 0$.  

23
This equality shows that
\[
\lambda_N \geq \left( \int_0^{2\pi} \sum_{k=0}^{\infty} \left| p_k \left( e^{i\theta} \right) \right|^2 \frac{d\theta}{2\pi} \right)^{-1} > 0. \tag{1.61}
\]

We thus have established that in the indeterminate case, the smallest eigenvalue \( \lambda_N \) is bounded from below. Put differently, if \( \lim_{N \to \infty} \lambda_N = 0 \), then the moment problem is determinate.

Conversely, assume that \( \lambda_N \geq \gamma \) for all \( N \), where \( \gamma > 0 \). We will show that in this case the moment problem is indeterminate. Since \( \frac{1}{\lambda_N} \leq \frac{1}{\gamma} \) for all \( N \), and \( \frac{1}{\lambda_N} \) is the largest eigenvalue of the positive definite matrix \( K_N \), it follows from the Rayleigh quotient that for all \( c = (c_0, c_1, \ldots, c_N) \in \mathbb{C}^{N+1} \),
\[
\frac{\sum_{j=0}^{N} \sum_{k=0}^{N} K_{jk} c_j c_k}{\sum_{j=0}^{N} |c_j|^2} \leq \frac{1}{\lambda_N} \leq \frac{1}{\gamma},
\]

hence
\[
\sum_{j=0}^{N} \sum_{k=0}^{N} K_{jk} c_j c_k \leq \frac{1}{\gamma} \sum_{j=0}^{N} |c_j|^2. \tag{1.62}
\]

Now let \( p \) be an arbitrary complex polynomial of degree \( \leq N \), say
\[
p(x) = \sum_{k=0}^{N} c_k p_k(x). \tag{1.63}
\]

Then (1.62) can be reformulated as
\[
\int_0^{2\pi} \left| p \left( e^{i\theta} \right) \right|^2 \frac{d\theta}{2\pi} = \int_0^{2\pi} \left| \sum_{k=0}^{N} c_k p_k \left( e^{i\theta} \right) \right|^2 \frac{d\theta}{2\pi} = \sum_{j=0}^{N} \sum_{k=0}^{N} K_{jk} c_j c_k
\]
\[
\leq \frac{1}{\gamma} \sum_{j=0}^{N} |c_j|^2 = \frac{1}{\gamma} \sum_{j=0}^{N} \sum_{k=0}^{N} c_j c_k \delta_{j,k}
\]
\[
= \frac{1}{\gamma} \sum_{j=0}^{N} \sum_{k=0}^{N} c_j c_k \int_{\mathbb{R}} p_j(x) p_k(x) d\mu(x)
\]
\[
= \frac{1}{\gamma} \int_{\mathbb{R}} \left( \sum_{j=0}^{N} c_j p_j(x) \right) \left( \sum_{k=0}^{N} c_k p_k(x) \right) d\mu(x)
\]
\[
= \frac{1}{\gamma} \int_{\mathbb{R}} \sum_{k=0}^{N} c_k p_k(x) \left| d\mu(x) = \frac{1}{\gamma} \int_{\mathbb{R}} \left| p(x) \right|^2 d\mu(x). \tag{1.64}
\]

Let \( z_0 \) be an arbitrary non-real number in the open unit disc, i.e. \( |z_0| < 1 \). Then it follows from the Cauchy integral formula that
\[
p(z_0) = \frac{1}{2\pi} \int_0^{2\pi} \frac{p \left( e^{i\theta} \right)}{e^{i\theta} - z_0} e^{i\theta} d\theta, \tag{1.65}
\]
For this particular polynomial we have

\[ |p(z_0)|^2 = \frac{1}{4\pi^2} \left| \int_0^{2\pi} p(e^{i\theta}) e^{i\theta} d\theta \right|^2 \leq \frac{1}{4\pi^2} \left( \int_0^{2\pi} \frac{|p(e^{i\theta})|}{e^{i\theta} - z_0} d\theta \right)^2 \]

\[ \leq \left( \int_0^{2\pi} \frac{|p(e^{i\theta})|^2}{2\pi} d\theta \right) \left( \int_0^{2\pi} \frac{1}{|e^{i\theta} - z_0|^2} d\theta \right) \]

\[ \leq \left( \int_0^{2\pi} \frac{|p(e^{i\theta})|^2}{2\pi} d\theta \right) \left( \int_0^{2\pi} \frac{1}{(1 - |z_0|)^2} d\theta \right) \]

\[ = \frac{1}{(1 - |z_0|)^2} \left( \int_0^{2\pi} |p(e^{i\theta})|^2 d\theta \right). \tag{1.66} \]

At the second inequality we made use of the Cauchy-Schwarz inequality for integrals\(^7\), while the third inequality follows from

\[ |e^{i\theta} - z_0| = |1 - z_0 e^{-i\theta}| \geq 1 - |z_0|. \]

We define

\[ \kappa = \frac{1}{\gamma(1 - |z_0|)^2}. \]

Then combining (1.64) and (1.66) yields the inequality

\[ |p(z_0)|^2 \leq \kappa \int_{\mathbb{R}} |p(x)|^2 d\mu(x) \tag{1.67} \]

which holds for every complex polynomial \( p \) of degree \( \leq N \). Lastly we define the complex polynomial

\[ p(x) = \sum_{k=0}^{N} p_k(z_0) p_k(x). \]

For this particular polynomial we have

\[ |p(z_0)|^2 = \overline{p(z_0)} p(z_0) = \left( \sum_{k=0}^{N} p_k(z_0) p_k(z_0) \right) \left( \sum_{l=0}^{N} p_l(z_0) p_l(z_0) \right) \]

\[ = \left( \sum_{k=0}^{N} |p_k(z_0)|^2 \right) \left( \sum_{l=0}^{N} |p_l(z_0)|^2 \right) = \left( \sum_{k=0}^{N} |p_k(z_0)|^2 \right)^2 \tag{1.68} \]

and

\[ \int_{\mathbb{R}} |p(x)|^2 d\mu(x) = \int_{\mathbb{R}} \left( \sum_{k=0}^{N} p_k(z_0) p_k(x) \sum_{l=0}^{N} p_l(z_0) p_l(x) \right) d\mu(x) \]

\[ = \sum_{k=0}^{N} \sum_{l=0}^{N} p_k(z_0) p_l(z_0) \left( \int_{\mathbb{R}} p_k(x) p_l(x) d\mu(x) \right) \]

\[ = \sum_{k=0}^{N} \sum_{l=0}^{N} p_k(z_0) p_l(z_0) \delta_{k,l} = \sum_{k=0}^{N} |p_k(z_0)|^2. \tag{1.69} \]

\(^7\)This inequality comprises the following statement: for integrable \( f, g : [a, b] \to \mathbb{R} \),

\[ \left( \int_a^b f(x) g(x) dx \right)^2 \leq \int_a^b f(x)^2 dx \int_a^b g(x)^2 dx. \]
From subsequently (1.68), (1.69) and (1.67) we thus obtain

\[ |p(z_0)|^2 = \left( \sum_{k=0}^{N} |p_k(z_0)|^2 \right)^2 = \left( \int_{\mathbb{R}} |p(x)|^2 d\mu(x) \right)^2 \leq \kappa \int_{\mathbb{R}} |p(x)|^2 d\mu(x). \]

Dividing on both sides by \( \int_{\mathbb{R}} |p(x)|^2 d\mu(x) \) yields

\[ \sum_{k=0}^{N} |p_k(z_0)|^2 \leq \kappa. \tag{1.70} \]

Since \( N \) is arbitrary, it follows that

\[ \sum_{k=0}^{\infty} |p_k(z_0)|^2 \leq \kappa < \infty, \tag{1.71} \]

in other words the moment problem is indeterminate by Remark 1.22.

As stated before, the original purpose of this thesis was to generalize the above proof to the matrix-valued case. Therefore we will revert our attention to the matrix-valued moment problem in the next chapter, and discuss some difficulties that arise while attempting to give this generalization.
Chapter 2

The matrix moment problem

In this chapter we generalize the moment problem introduced in Chapter 1 to the matrix-valued setting. Before formulating the matrix moment problem, we will take a closer look at matrix measures. Afterwards orthonormal matrix polynomials are treated, and their connection to the moment problem will be discussed (as briefly done in Chapter 1 for the classical moment problem). This chapter is concluded by describing some of the difficulties in generalizing Theorem 1.18.

The treatment of the matrix moment problem and orthonormal polynomials in this thesis is mostly based on [5].

2.1 Matrix measures

We will generalize the notion of complex measures to matrix measures. See Appendix A for a brief overview of complex measures.

Definition 2.1. A matrix measure \( \mu \) on a measurable space \((X, \mathcal{E})\) is a function \( \mu : \mathcal{E} \to \mathbb{C}^{K \times K} \) such that

\[
\mu \left( \bigcup_{n=1}^{\infty} E_n \right) = \sum_{n=1}^{\infty} \mu(E_n)
\]

(2.1)

for any sequence \((E_n)_{n \geq 1}\) of pairwise disjoint sets from \( \mathcal{E} \).

\( \blacksquare \)

Note that a matrix measure \( \mu = (\mu_{ij})_{0 \leq i,j \leq K-1} \) can be considered as a matrix of \( K^2 \) complex measures.

Definition 2.2. A matrix measure \( \mu \) is positive if all values of the measure are positive hermitian matrices, i.e. \( \mu(E) \) is a positive hermitian matrix for all \( E \in \mathcal{E} \). We denote the set of positive matrix measures on \((X, \mathcal{E})\) with \( M_K(X) \).

\( \blacksquare \)

In what follows, we will always assume that \( \mu \) is a positive matrix measure (unless stated otherwise).

Lemma 2.3. A positive matrix measure \( \mu \) is

(i) increasing, which means that \( E \subseteq F \) implies that \( \mu(E) \prec \mu(F) \) for all \( E, F \in \mathcal{E} \).

(ii) countably subadditive, i.e. for every sequence \((E_n)_{n \geq 1}\) in \( \mathcal{E} \) we have

\[
\mu \left( \bigcup_{n=1}^{\infty} E_n \right) \prec \sum_{n=1}^{\infty} \mu(E_n)
\]

(2.2)
Proof.

(i) Let $E \subseteq F$. Then $\mu(F) = \mu(E) + \mu(F \setminus E)$, and since $\mu(F) - \mu(E) = \mu(F \setminus E)$ is positive hermitian, it follows that $\mu(E) \prec \mu(F)$.

(ii) Define the sequence $(F_n)_{n \geq 1}$ by $F_1 := E_1$ and $F_n := E_n \setminus (E_1 \cup \ldots \cup E_{n-1})$ for $n > 1$. Note that the $F_n$ are pairwise disjoint sets by construction. Moreover $\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} F_n$, from which it follows that

$$\mu \left( \bigcup_{n=1}^{\infty} E_n \right) = \mu \left( \bigcup_{n=1}^{\infty} F_n \right) = \sum_{n=1}^{\infty} \mu(F_n) = \lim_{k \to \infty} \sum_{n=1}^{k} \mu(F_n). \tag{2.3}$$

As $F_n \subseteq E_n$ for every $n \geq 1$, invoking (i) yields $\mu(F_n) \prec \mu(E_n)$, and thus

$$\sum_{n=1}^{k} \mu(F_n) \prec \sum_{n=1}^{k} \mu(E_n). \tag{2.4}$$

We conclude that

$$\mu \left( \bigcup_{n=1}^{\infty} E_n \right) = \lim_{k \to \infty} \sum_{n=1}^{k} \mu(F_n) \prec \lim_{k \to \infty} \sum_{n=1}^{k} \mu(E_n) = \sum_{n=1}^{\infty} \mu(E_n), \tag{2.5}$$

as was to be shown.

In the remainder of this section (see Theorem 2.6) we will show that every positive matrix measure $\mu$ can be written as

$$\mu(dx) = W(x)d\tau_{\mu}(x) \tag{2.6}$$

where $W(x)$ is a positive hermitian matrix and $\tau_{\mu}$ is the so-called trace measure (which is defined below).

**Definition 2.4.** Let $\mu$ be a positive matrix measure. Then the diagonal measures $\mu_{ii}$ are positive finite measures and so is

$$\tau_{\mu} := \text{tr}(\mu) = \mu_{00} + \ldots + \mu_{K-1,K-1}. \tag{2.7}$$

**Lemma 2.5.** For any $\mu \in \mathcal{M}_K(X)$ we have

$$|\mu_{ij}(E)| \leq |\mu_{ij}|(E) \leq \tau_{\mu}(E) \tag{2.8}$$

for all $E \in \mathcal{E}$.

In this lemma $|\mu_{ij}|$ denotes the variation of $\mu_{ij}$ (see Definition A.5).

**Proof.** Let $E \in \mathcal{E}$. The first inequality is obvious by construction, see Appendix A. In order to prove the second one, we introduce the notation $a_{ij} = \mu_{ij}(E)$ and $A = (a_{ij})$. For $i \neq j$, the $2 \times 2$ matrix

$$\begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix}$$

is positive definite, so $a_{ij} \neq 0$. For $i = j$, the matrix

$$\begin{pmatrix} a_{ii} & 0 \\ 0 & a_{jj} \end{pmatrix}$$

is positive definite, so $a_{ii} > 0$ for $i = 0, 1, \ldots, K-1$. Therefore

$$|\mu_{ij}(E)| \leq |\mu_{ij}|(E) \leq \tau_{\mu}(E),$$

as was to be shown.

\[\Box\]
is positive hermitian. In particular it has a non-negative determinant, so that \(a_{ii}a_{jj} - a_{ij}a_{ji} \geq 0\), and thus \(|a_{ij}|^2 \leq a_{ii}a_{jj}\). Hence

\[
|a_{ij}| \leq \sqrt{a_{ii}a_{jj}} \leq 2\sqrt{a_{ii}a_{jj}} \leq a_{ii} + a_{jj} \leq \text{tr}(A)
\]  

where the third inequality follows by rewriting \((\sqrt{a_{ii}} - \sqrt{a_{jj}})^2 \geq 0.1\) We conclude from (2.9) that \(|\mu_{ij}(E)| \leq \tau_\mu(E)\). Since \(|\mu_{ij}|\) is the smallest positive measure satisfying these inequalities (see Appendix A), we get \(|\mu_{ij}(E)| \leq \tau_\mu(E)\) for all \(E \in \mathcal{E}\). □

Before we are able to give a proof of (2.6), we need the Radon-Nikodym Theorem, which is formulated in Theorem A.8. Observe that \(\mu_{ij} \ll \tau_\mu\) for all \(0 \leq i, j \leq K - 1\). Indeed, suppose that \(\tau_\mu(E) = 0\). Then clearly \(\mu_{ii}(E) = 0\) for every \(i\), and by (2.9), \(\mu_{ij}(E) = 0\) for all \(0 \leq i, j \leq K - 1\).

Thus by Theorem A.8 there exist measurable functions \(f_{ij}\) such that

\[
\mu_{ij}(E) = \int_E f_{ij} d\tau_\mu \quad (2.10)
\]

for all \(E \in \mathcal{E}\). Hence \(|\mu_{ij}(E)| = \int_E |f_{ij}| d\tau_\mu\) by Proposition A.9. Thus, according to Lemma 2.5, \(|f_{ij}(x)| \leq 1\) for \(\tau_\mu\)-almost all \(x \in X\), since

\[
\int_E |f_{ij}| d\tau_\mu = |\mu_{ij}(E)| \leq \tau_\mu(E) = \int_E d\tau_\mu.
\]

It follows that \(\int_E |f_{ij}| d\tau_\mu < \infty\) (as \(\tau_\mu\) is a finite measure), in other words \(f_{ij} \in L^1(\tau_\mu)\). Now we can finally state the desired result.

**Theorem 2.6.** Let \(\mu\) be a positive matrix measure with Radon-Nikodym derivatives \(f_{ij} \in L^1(\tau_\mu)\) such that

\[
\mu_{ij}(E) = \int_E f_{ij}(x) d\tau_\mu(x) \quad (2.11)
\]

for \(E \in \mathcal{E}\). Then the matrix \(W(x) := (f_{ij}(x))_{0 \leq i, j \leq K - 1}\) is positive hermitian for \(\tau_\mu\)-almost all \(x \in X\).

**Proof.** Since \(\mu\) is a positive matrix measure, it follows that \(\mu(E)\) is positive hermitian for any \(E \in \mathcal{E}\). Let \(v \in \mathbb{C}^K \setminus \{0\}\). Then

\[
\int_E F(x, v) d\tau_\mu(x) := \int_E \langle W(x)v, v \rangle d\tau_\mu(x) = \left\langle \int_E W(x) d\tau_\mu(x)v, v \right\rangle = \langle \mu(E)v, v \rangle > 0 \quad (2.12)
\]

where we have defined \(F(x, v) = \langle W(x)v, v \rangle\) and made use of the fact that (2.11) implies

\[
\mu(E) = \int_E W(x) d\tau_\mu(x).
\]

Thus the function \(x \mapsto F(x, v)\) has a positive integral over all sets \(E \in \mathcal{E}\) w.r.t. the positive measure \(\tau_\mu\). Hence \(F(x, v)\) is real-valued and positive for \(\tau_\mu\)-almost all \(x \in \mathcal{E}\). It follows that there exists a set \(\Omega_v\) with \(\tau_\mu(\Omega_v) = 0\) such that \(F(x, v) > 0\) for all \(x \in X \setminus \Omega_v\).

Now define \(\Omega = \bigcup_{v \in D} \Omega_v\) for some countable dense subset \(D\) of \(\mathbb{C}^K\). Then

\[
\tau_\mu(\Omega) = \tau_\mu\left(\bigcup_{v \in D} \Omega_v\right) \leq \sum_{v \in D} \tau_\mu(\Omega_v) = 0
\]

\[1\text{Indeed, } 0 \leq (\sqrt{a_{ii}} - \sqrt{a_{jj}})^2 = a_{ii} - 2\sqrt{a_{ii}a_{jj}} + a_{jj}, \text{ so that } a_{ii} + a_{jj} \geq 2\sqrt{a_{ii}a_{jj}}.\]
so that $\tau_\mu(\Omega) = 0$. We also have $F(x, v) > 0$ for all $x \in X \setminus \Omega$ and $v \in D$. For fixed $x \in X$ the function $v \mapsto F(x, v)$ is continuous. Moreover it is positive on the dense set $D$ whenever $x \in X \setminus \Omega$, and hence it is positive on $\mathbb{C}^K$. But then $W(x)$ is positive hermitian for $\tau_\mu$-almost all $x \in X$, namely for all $x \in X \setminus \Omega$.

**Remark 2.7.** Moreover it can be shown that $W(x) \prec \mathbb{1}_K$ for $\tau_\mu$-almost all $x \in X$. Indeed, invoking Lemma 1.10 we see that

\begin{align*}
\int_E \langle (\mathbb{1}_K - W(x))v, v \rangle d\tau_\mu(x) &= \int_E \langle v, v \rangle d\tau_\mu(x) - \int_E \langle W(x)v, v \rangle d\tau_\mu(x) \\
&= \langle \tau_\mu(E)v, v \rangle - \langle \mu(E)v, v \rangle > 0
\end{align*}

for all $v \in \mathbb{C}^K \setminus \{0\}$. Following the reasoning in the above proof, we see that $\mathbb{1}_K - W(x)$ is positive hermitian for $\tau_\mu$-almost all $x \in X$, and from this the claim follows.

In the following we always assume that the functions $f_{ij}$ are chosen such that $W(x)$ is positive hermitian and $|f_{ij}(x)| \leq 1$ for all $x \in X$.

### 2.2 Formulation of the matrix moment problem

In this section we generalize the scalar moment problem to the matrix-valued case.

Suppose $K \geq 1$. Let $\mu = W \tau_\mu$ be a positive matrix measure, supported on the real line, and with moments of any order. Here $W(x) \in \mathbb{C}^{K \times K}$ is positive hermitian, i.e. $W(x) = (f_{ij}(x))_{0 \leq i,j \leq K-1}$ where the functions $f_{ij}$ are chosen as in Theorem 2.6.

**Definition 2.8.** Denote the $n$th moment of the measure $\mu$ with

\[ S_n := S_n(\mu) = \int_\mathbb{R} x^n d\mu(x) = \int_\mathbb{R} x^n W(x) d\tau_\mu(x) \quad (2.13) \]

for $n \in \mathbb{N}$. We call $(S_n)_{n \geq 0}$ a matrix moment sequence.

The integration in (2.13) has to be taken entrywise, which means that the $(i,j)$ entry of the matrix $S_n \in \mathbb{C}^{K \times K}$ is given by

\[ (S_n)_{ij} = \int_\mathbb{R} x^n f_{ij}(x) d\tau_\mu(x). \quad (2.14) \]

Note that $S_n^* = S_n$ as $W(x)$ is hermitian. Under certain conditions it is possible to assume $S_0 = \mathbb{1}_K$ without loss of generality (see Remark 2.29). This normalization however, is not always convenient in explicit examples.

**Notation 2.9.** We denote the set of positive matrix measures on $\mathbb{R}$ with moments of any order as $\mathcal{M}_K^* = \mathcal{M}_K^*(\mathbb{R})$. For $\mu \in \mathcal{M}_K^*$ we denote the set of all $\nu \in \mathcal{M}_K^*$ with the same moments as $\mu$ by $[\mu]$, i.e.

\[ [\mu] = \{\nu \in \mathcal{M}_K^* : S_n(\nu) = S_n(\mu) \text{ for all } n \geq 0\}. \]

The **matrix moment problem** consists of the following two questions:

1. Which sequences $(S_n)_{n \geq 0}$ are matrix moment sequences?
2. To which extent is \( \mu \in M_K^* \) determined by its moment sequence?

The answer to the latter question is given in terms of the determinacy of the measure \( \mu \):

**Definition 2.10.** Let \( \mu \in M_K^* \). Then \( \mu \) or the corresponding moment sequence \((S_n)_{n \geq 0}\) is called **determinate** if \([\mu] = \{\mu\}\), and **indeterminate** otherwise.

If \( \mu \) is indeterminate, then \([\mu]\) is a convex set with at least two elements, and thus infinite.

In the previous section we have seen that the trace measure \( \tau_\mu \) is useful to give another characterization of positive matrix measures, namely via (2.6) with \( W(x) = (f_{ij}(x))_{0 \leq i,j \leq K-1} \). As we shall show next, this particular measure also gives a sufficient condition for determinacy of its associated matrix measure \( \mu \). To show this, we first need to state a result regarding the determinacy of matrix measures and its components.

**Theorem 2.11.** Let \( \mu, \nu \) be positive matrix measures with moments of any order and assume that they have the same moments.

(i) If \( \mu_{ii} \) is determinate for some \( i \in \{0, 1, \ldots, K-1\} \), then \( \mu_{ij} = \nu_{ij} \) for \( j \in \{0, 1, \ldots, K-1\} \).

(ii) If \( \mu_{ii} \) is determinate for all \( i \in \{0, 1, \ldots, K-1\} \), then \( \mu = \nu \), so that \( \mu \) is determinate.

The proof is omitted and can be found in [5], Theorem 3.6. Note that (ii) immediately follows from (i).

**Corollary 2.12.** Let \( \mu \) be a positive matrix measure with moments of any order. If \( \tau_\mu \) is determinate, then \( \mu \) is determinate.

**Proof.** Fix \( i \) and note that \( \mu_{ii} \leq \tau_\mu \). Then it follows from Lemma A.4 that \( \mu_{ii} \) is determinate, since \( \tau_\mu \) is assumed to be determinate. But then \( \mu \) is determinate according to Theorem 2.11. \( \Box \)

The above corollary thus relates determinacy of the matrix moment problem and the classical moment problem.

Completely similar to Section 1.1, we now form the **Hankel block matrices** corresponding to a sequence \((S_n)_{n \geq 0}\) of hermitian \( K \times K \)-matrices, namely

\[
H_N = (S_{i+j})_{0 \leq i,j \leq N}
\]  
(2.15)

for \( N \geq 0 \). Observe that the Hankel matrices are of size \( K(N+1) \times K(N+1) \). Written out we thus get

\[
H_N = \begin{pmatrix}
S_0 & S_1 & S_2 & \cdots & S_N \\
S_1 & S_2 & S_3 & \cdots & S_{N+1} \\
S_2 & S_3 & S_4 & \cdots & S_{N+2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
S_N & S_{N+1} & S_{N+2} & \cdots & S_{2N}
\end{pmatrix}
\]  
(2.16)

Moreover we define the infinite Hankel matrix to be

\[
H_\infty = (S_{i+j})_{i,j \geq 0}
\]  
(2.17)
Theorem 2.14 (Krein). The generalized form of Hamburger’s theorem (i.e. Theorem 1.14), is the following:

\[ \langle H_N v, v \rangle = \sum_{i,j=0}^{N} v_i^* S_{i+j} v_j > 0 \text{ for all } N \geq 0 \text{ and } 0 \neq v \in (\mathbb{C}^K)^{N+1}. \] (2.18)

The generalized form of Hamburger’s theorem (i.e. Theorem 1.14), is the following:

Definition 2.13. A matrix sequence \((S_n)_{n \geq 0}\) is called positive definite if its Hankel matrix \(H_N\) is positive hermitian for every \(N\), which is equivalent to

\[ \langle H_N v, v \rangle = \sum_{i,j=0}^{N} v_i^* S_{i+j} v_j > 0 \text{ for all } N \geq 0 \text{ and } 0 \neq v \in (\mathbb{C}^K)^{N+1}. \]

The other implication is more involved (as was the case in Theorem 1.14), and can be given in a similar fashion by using (a generalized form of) Helly’s theorem. Another proof can be found in [5], Theorem 3.2, and is based on [20]. Here we only present a sketch of the proof.

The large part of that proof relies on spectral theory, which will also be discussed in Chapter 3 (and Appendix B).

Given a positive definite sequence \((S_n)_{n \geq 0}\), define the positive hermitian form \(\langle \cdot, \cdot \rangle\) on the set of vector polynomials \(L = \{ g(x) = \sum_i c_i x^i : c_i \in \mathbb{C}^K \}\) by

\[ \left\langle \sum d_j x^j, \sum c_i x^i \right\rangle = \sum_{i,j} c_i^* S_{i+j} d_j \]

with associated seminorm \( \| \sum_i c_i x^i \|^2 = \sum_{i,j} c_i^* S_{i+j} c_j \). Consider the multiplication operator in \(L\), denoted by \(A_0\), i.e. \((A_0 g)(x) = xg(x)\). It can then be shown that \(A_0\) induces an operator \(A\) in the quotient space \(L/L_0\), where \(L_0 = \{ g \in L : \|g\| = 0 \}\), and that the form defined above defines an inner product on \(L/L_0\). Now take \(\mathcal{H}\) to be the Hilbert space completion of \(L/L_0\) with respect to that inner product, and let \(\tilde{H}\) be a Hilbert space which contains \(\mathcal{H}\) as a closed subspace. It is known that one can find a self-adjoint extension of \(A\) in \(\tilde{H}\), say \(T\). Applying the Spectral Theorem, \(T\) can be written in the form \(T = \int_{\mathbb{R}} x dW(x)\) for some spectral measure \(W(x)\) on the Borel sets of \(\mathbb{R}\). Finally a particular positive measure \(\mu\) is defined in terms of \(W(x)\) (similar as in (3.9)), and again invoking the Spectral Theorem one sees that \((S_n)_{ij} = \int_{\mathbb{R}} x^n d\mu_{ij}(x)\) where \(\mu_{ij}\) is a component of the matrix measure \(\mu\). Hence \((S_n)_{n \geq 0}\) is a matrix moment sequence. □
2.3 Matrix inner products and orthonormal matrix polynomials

2.3.1 Matrix polynomials

The orthonormal polynomials with respect to some positive measure turned out to have a direct connection with the associated (classical) moment problem; see Section 1.2. Here we proceed completely similar by introducing orthonormal matrix polynomials. First we revert our attention to matrix polynomials.

**Definition 2.15.** A matrix polynomial $P$ is a polynomial in one complex variable $x$ which has $K \times K$-matrices as coefficients, i.e.

$$P(x) = \sum_{k=0}^{n} A_k x^k = \sum_{k=0}^{n} x^k A_k \quad \text{with} \quad A_k \in \mathbb{C}^{K \times K} \quad \text{for every} \quad 0 \leq k \leq K. \quad (2.19)$$

The degree of $P$ is the highest power $k$ of $x$ for which $A_k \neq 0$. The set of matrix polynomials with coefficients in $\mathbb{C}^{K \times K}$ is denoted by $\mathbb{C}^{K \times K}[x]$.

If a matrix polynomials of degree $\leq n$ equals the zero matrix for more than $n$ values of the variable $x$, then all matrix coefficients are equal to the zero matrix.

**Notation 2.16.** If $P$ is a polynomial of degree $n$, then we denote its leading coefficient by $\text{lc}(P) = A_n$.

**Remark 2.17.** We consider the set $\mathbb{C}^{K \times K}[x]$ of matrix polynomials as a module over the matrix ring $\mathbb{C}^{K \times K}$. This is possible since $\mathbb{C}^{K \times K}$ is a complex vector space in which left and right multiplication by matrices is possible.

**Notation 2.18.** Let $P(x) = \sum_{k=0}^{n} A_k x^k$ be a matrix polynomial. Then we denote by $P^*$ the matrix polynomial

$$P^*(x) = \sum_{k=0}^{n} A_k^* x^k. \quad (2.20)$$

Note that $P(x)^* = P^*(x)$ for $x \in \mathbb{C}$.

In the scalar case the orthonormal polynomials $(p_n)_{n \geq 0}$ satisfy $\text{deg}(p_n) = n$ and $\text{lc}(p_n) > 0$. For the orthonormal polynomials in the matrix-valued case, similar properties hold. For this reason we introduce the following terminology:

**Definition 2.19.** A sequence of matrix polynomials $(P_n)_{n \geq 0}$ is called simple if

(i) $P_n$ has degree $n$;

(ii) The leading coefficient of $P_n$ is regular.

**Proposition 2.20.** Let $(P_n)_{n \geq 0}$ be a simple sequence of matrix polynomials. Then every matrix polynomial $P$ of degree $n$ can be uniquely expressed as

$$P(x) = \sum_{k=0}^{n} A_k P_k(x) \quad \text{where} \quad A_k \in \mathbb{C}^{K \times K} \quad \text{for every} \quad 0 \leq k \leq n. \quad (2.21)$$
Put differently, we thus see that a simple sequence of matrix polynomials forms a basis of $\mathbb{C}^{K \times K}[x]$, as a left module over $\mathbb{C}^{K \times K}$.

**Proof.** Let $P$ be a polynomial of degree $n$, and let $(P_n)_{n \geq 0}$ be a simple sequence of matrix polynomials wherein $P_n$ has leading coefficient $L_n$. Then it can obviously be written as a linear combination of the polynomials $P_n$. Assume that

$$P(x) = \sum_{k=0}^{n} A_k P_k(x) = \sum_{k=0}^{n} B_k P_k(x)$$

are two different ways to write $P$. Then the leading coefficient of $P$ equals $\text{lc}(P) = A_n \text{lc}(P_n) = A_n L_n$ and likewise $\text{lc}(P) = B_n L_n$. Regularity of $L_n$ then implies that $A_n = B_n$. It is then inductively clear that also $A_{n-1} = B_{n-1}, \ldots, A_0 = B_0$. Hence $P$ can be uniquely expressed as a linear combination of the polynomials $P_n$. 

Before we are able to give a definition of orthonormal matrix polynomials, we need to consider matrix inner products.

**Definition 2.21.** A matrix inner product on $\mathbb{C}^{K \times K}[x]$ is a mapping

$$\langle \cdot, \cdot \rangle : \mathbb{C}^{K \times K}[x] \times \mathbb{C}^{K \times K}[x] \to \mathbb{C}^{K \times K}$$

such that

(i) $\langle P, Q \rangle = \langle Q, P \rangle^*$;

(ii) $\langle A_1 P_1 + A_2 P_2, Q \rangle = A_1 \langle P_1, Q \rangle + A_2 \langle P_2, Q \rangle$, where $A_i \in \mathbb{C}^{K \times K}$ for $i = 1, 2$;

(iii) $\theta < \langle P, P \rangle$.

The matrix inner product is called non-degenerate if it also satisfies

(iv) for all $P \in \mathbb{C}^{K \times K}[x]$, if $\langle P, P \rangle = \theta$, then $P = \theta$.

The matrix inner product is called degenerate if there exists some non-zero matrix polynomial $P$ for which $\langle P, P \rangle = \theta$.

**Remark 2.22.** From (i) it follows that $\langle P, P \rangle$ is always hermitian, while (iii) implies that it is positive. Moreover it follows from (i) and (ii) that

$$\langle P, A_1 Q_1 + A_2 Q_2 \rangle = \langle A_1 Q_1 + A_2 Q_2, P \rangle^* = (A_1 \langle Q_1, P \rangle + A_2 \langle Q_2, P \rangle)^* = \langle Q_1, P \rangle^* A_1^* + \langle Q_2, P \rangle^* A_2^* = \langle P, Q_1 \rangle A_1^* + \langle P, Q_2 \rangle A_2^*.$$

**Lemma 2.23.** A matrix inner product $\langle \cdot, \cdot \rangle$ is non-degenerate if and only if for all $n \geq 0$ and for all $v_0, \ldots, v_n \in \mathbb{C}^{K}$ the following condition holds:

$$\sum_{i,j=0}^{n} v_i^* \langle x^i 1_K, x^j 1_K \rangle v_j = 0 \iff v_0 = \ldots = v_n = 0. \quad (2.24)$$

Note that the implication $\Leftarrow$ in (2.24) is trivial.
Proof. Let $P$ be a matrix polynomial, say $P(x) = \sum_{i=0}^{n} A_i x^i$. Then

$$\langle P, P \rangle = \sum_{i,j=0}^{n} A_i \langle x^i 1_K, x^j 1_K \rangle A_j^*.$$ 

Suppose that $\langle \cdot, \cdot \rangle$ is non-degenerate and let $v_0, \ldots, v_n \in \mathbb{C}^K$ satisfy

$$\sum_{i,j=0}^{n} v_i^* \langle x^i 1_K, x^j 1_K \rangle v_j = 0.$$ 

Now let $A_i$ be the matrix for which the zero'th row is equal to $v_i^*$, while all other rows are zero. Then the $kl$'th entry of $A_i \langle x^i 1_K, x^j 1_K \rangle A_j^* v$ equals zero, unless $k = l = 0$, in which case the entry equals $v_i^* \langle x^i 1_K, x^j 1_K \rangle v_j$. Taking the sum over all $0 \leq i, j \leq n$, it follows that $\langle P, P \rangle = 0$, and thus $P = \theta$ by non-degeneracy of the inner product. But then $A_i = \theta$ for every $i$, from which it follows that $v_i = 0$. This proves the first implication.

Conversely assume that (2.24) holds. Moreover let $P$ be a matrix polynomial that satisfies $\langle P, P \rangle = 0$, say $P(x) = \sum_{i=0}^{n} A_i x^i$. Note that

$$0 = v^* \langle P, P \rangle v = \sum_{i,j=0}^{n} v_i^* A_i \langle x^i 1_K, x^j 1_K \rangle A_j^* v$$

for any $v \in \mathbb{C}^K$, so that $A_i^* v = 0$ for every $0 \leq i \leq n$ by assumption. Since $v$ was chosen arbitrarily, we have $A_i = \theta$ for all $i$. But then $P = \theta$, hence the inner product is non-degenerate.

Having introduced the concept of matrix inner products, we are now able to give a specific inner product with respect to a given matrix measure $\mu \in \mathcal{M}_K^\times$.

**Definition 2.24.** Let $\mu \in \mathcal{M}_K^\times$. Then we define the matrix inner product with respect to $\mu$ by

$$\langle P, Q \rangle_{\mu} := \int_{\mathbb{R}} P(x) d\mu(x) Q^*(x) = \int_{\mathbb{R}} P(x) W(x) Q^*(x) d\tau_\mu(x).$$

(2.25)

We will often just write $\langle \cdot, \cdot \rangle$ for the above inner product if it is clear with respect to which measure $\mu$ the inner product has to be taken.

**Definition 2.25.** A matrix measure $\mu$ is called non-degenerate if $\langle \cdot, \cdot \rangle_{\mu}$ is non-degenerate. □

**Corollary 2.26.** A positive matrix measure $\mu$ is non-degenerate if and only if for all $n \geq 0$ and for all $v_0, \ldots, v_n \in \mathbb{C}^K$ the following condition holds:

$$\sum_{i,j=0}^{n} v_i^* S_{i+j} v_j = 0 \iff v_0 = \ldots = v_n = 0.$$  

(2.26)

Proof. This is an immediate corollary to Lemma 2.23, since

$$\langle x^i 1_K, x^j 1_K \rangle_{\mu} = \int_{\mathbb{R}} x^{i+j} d\mu(x) = S_{i+j}.$$ 

□
Example 2.27. Let \( \mu_0, \ldots, \mu_{K-1} \) be \( K \) positive Borel measures on \( \mathbb{R} \) with moments of any order, and let \( \mu \) be the matrix of measures defined by

\[
\mu_{ij} = \begin{cases} 
\mu_i & \text{if } i = j; \\
0 & \text{if } i \neq j.
\end{cases}
\]

We assume that \( \text{supp}(\mu_i) \) contains infinitely many points for each \( i = 0, 1, \ldots, K - 1 \), so that \( \text{supp}(\mu) \) is also infinite. Let \( v \in \mathbb{C}^K \) with components \( v_i \) (\( 0 \leq i \leq K - 1 \)). Then

\[
\langle \mu(E)v, v \rangle = v^*\mu(E)v = (v_0, \overline{v_1}, \ldots, \overline{v_{K-1}})^t \begin{pmatrix} \mu_0(E) & \mu_1(E) & \cdots & \mu_{K-1}(E) \end{pmatrix} \begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_{K-1} \end{pmatrix}
\]

\[
= \sum_{i=0}^{K-1} v_i \mu_i(E)v_i = \sum_{i=0}^{K-1} \mu_i(E)|v_i|^2 > 0
\]

for every Borel set \( E \), so that \( \mu \) is a positive matrix measure.

Now denote the \( n \)th moment of \( \mu_i \) by

\[
s^i_n = \int_{\mathbb{R}} x^n d\mu_i(x).
\]

Then the \( n \)th moment of \( \mu \), which we will denote by \( S_n \), equals

\[
S_n = \int_{\mathbb{R}} x^n d\mu(x) = \int_{\mathbb{R}} x^n \left( d\mu_0(x) d\mu_1(x) \cdots d\mu_{K-1}(x) \right) = \begin{pmatrix} s^0_n \\ s^1_n \\ \vdots \\ s^K_n \end{pmatrix}.
\]

We will show that \( \mu \) is non-degenerate. Let \( v_0, \ldots, v_n \in \mathbb{C}^K \), where we write \( v_i = (v^{(i)}_0, \ldots, v^{(i)}_{K-1})^t \) for every \( i \). Assume that \( v_i \neq 0 \) for at least one \( i \). Then

\[
\sum_{i,j=0}^{n} v_i^* S_{i+j} v_j = \sum_{i,j=0}^{n} \left( v^{(i)}_0, \overline{v^{(i)}_1}, \ldots, \overline{v^{(i)}_{K-1}} \right)^t \begin{pmatrix} s^0_{i+j} \\ s^1_{i+j} \\ \vdots \\ s^K_{i+j} \end{pmatrix} \begin{pmatrix} v^{(j)}_0 \\ v^{(j)}_1 \\ \vdots \\ v^{(j)}_{K-1} \end{pmatrix}
\]

\[
= \sum_{i,j=0}^{n} \sum_{k=0}^{K-1} v^{(i)}_k \overline{v^{(j)}_{i+j} v^{(j)}_k} = \sum_{i,j=0}^{n} \sum_{k=0}^{K-1} v^{(i)}_k x^{i+j} v^{(j)}_k d\mu_k(x)
\]

\[
= \sum_{k=0}^{K-1} \left( \sum_{j=0}^{n} x^j v^{(j)}_k \right)^2 d\mu_k(x) > 0.
\]

According to Corollary 2.26, \( \mu \) is non-degenerate. ■

Definition 2.28. A sequence \( (P_n)_{n \geq 0} \) of matrix polynomials is called orthonormal with respect to \( \langle \cdot, \cdot \rangle \) if it is a simple sequence and moreover

\[
\langle P_n, P_m \rangle = \delta_{n,m} \mathbb{1}_K \tag{2.27}
\]

for any \( n, m \geq 0 \). If the inner product is given by \( \mu \) (i.e. \( \langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_\mu \)), then the polynomials are also said to be orthonormal with respect to \( \mu \). ■
Remark 2.29. Without loss of generality we can always assume that $S_0 = I_K$ whenever $\mu$ is non-degenerate. To see this, note that $\mu_M := M\mu M^*$ is a positive matrix measure for any $M \in \mathbb{C}^{K \times K}$, defined by $\mu_M(B) = M\mu(B)M^*$ for Borel sets $B$. The moments $S^{(M)}_n$ are given by

$$S^{(M)}_n(x) = \int_{\mathbb{R}} x^n d\mu_M(x) = \int_{\mathbb{R}} x^n M d\mu(x) M^* = M \left( \int_{\mathbb{R}} x^n d\mu(x) \right) M^* = MS_n M^*.$$ 

Moreover the inner product with respect to $\mu_M$ is given by

$$\langle P, Q \rangle_{\mu_M} = \int_{\mathbb{R}} P(x) d\mu_M(x) Q^*(x) = \int_{\mathbb{R}} P(x) M d\mu(x) M^* Q^*(x) = \langle PM, QM \rangle_\mu.$$ 

Now suppose that $\mu$ is non-degenerate and $M$ is regular. Moreover assume that $\langle P, P \rangle_{\mu_M} = \theta$, which implies that $\langle PM, PM \rangle_\mu = \theta$. By non-degeneracy of $\mu$ we have $PM = \theta$, so that $P = \theta$ as $M$ is regular. It follows that $\mu_M$ is non-degenerate.

Now $S_0$ can be normalized by considering the matrix measure $\mu_{P_0}$. Indeed, we then have

$$S^{(P_0)}_0 = P_0 S_0 P_0^* = \int_{\mathbb{R}} P_0 d\mu(x) P_0^* = \langle P_0, P_0 \rangle_\mu = I_K.$$ 

In explicit examples however, this normalization is not always the most convenient choice. Alternatively one can put $P_0 = S_0^{-\frac{1}{2}}$ (so that (2.27) holds for $n = m = 0$), which can be done since $S_0$ is a positive definite matrix, hence having a square root and an inverse (that also has a square root). \[\blacksquare\]

Remark 2.30. If $(P_n)_{n \geq 0}$ is a sequence of orthonormal matrix polynomials with respect to $\mu$, we can construct another set of orthonormal polynomials (also with respect to $\mu$) by defining $Q_n = U_n P_n$, where $U_n$ is unitary for every $n \geq 0$. Indeed, $Q_n$ obviously has degree $n$ and $\text{lc}(Q_n) = U_n \text{lc}(P_n)$ is regular. Moreover,

$$\langle Q_n, Q_m \rangle = \int_{\mathbb{R}} Q_n(x) W(x) Q_m^*(x) d\tau_\mu(x) = \int_{\mathbb{R}} U_n P_n(x) W(x) P_m^*(x) U_n^* d\tau_\mu(x) = U_n \left( \int_{\mathbb{R}} P_n(x) W(x) P_m^*(x) d\tau_\mu(x) \right) U_n^* = U_n \delta_{n,m} I_K U_m^* = \delta_{n,m} U_n U_m^* = \delta_{n,m} I_K.$$ 

Of course the same also holds if $(P_n)_{n \geq 0}$ is an orthonormal sequence with respect to an arbitrary inner product $\langle \cdot, \cdot \rangle$, since in that case, $\langle Q_n, Q_m \rangle = \langle U_n P_n, U_m P_m \rangle = U_n \langle P_n, P_m \rangle U_m^* = U_n \delta_{n,m} U_m^* = \delta_{n,m} I_K$. \[\blacksquare\]

It turns out that all orthonormal sequences with respect to the same non-degenerate inner product are related via a sequence $(U_n)_{n \geq 0}$ of unitary matrices. Before we are able to prove this statement, we need the following theorem:

Theorem 2.31 (Matrix Gram-Schmidt). For a matrix inner product $\langle \cdot, \cdot \rangle$ the following statements are equivalent:

(i) $\langle \cdot, \cdot \rangle$ is non-degenerate.

(ii) For every matrix polynomial $P$ with regular leading coefficient, $\langle P, P \rangle$ is regular.

(iii) There exists an orthonormal sequence $(P_n)_{n \geq 0}$ with respect to $\langle \cdot, \cdot \rangle$.

If these statements hold, any matrix polynomial $P$ of degree $n$ can be uniquely expressed as

$$P = \sum_{k=0}^{n} A_k P_k \quad (2.28)$$

where $A_k \in \mathbb{C}^{K \times K}$ is given by $A_k = \langle P, P_k \rangle$. 

37
Proof. \((i) \iff (ii)\): Suppose \(\langle \cdot, \cdot \rangle\) is non-degenerate. For a matrix \(A \in \mathbb{C}^{K \times K}\) we introduce the notation
\[
\ker(A) = \{v \in \mathbb{C}^K : Av = 0\}.
\]
Then for a matrix polynomial \(P\) of degree \(n\), say \(P(x) = \sum_{k=0}^n A_k x^k\), we have
\[
v \in \ker(P, P) \iff \langle P, P \rangle v = 0 \iff \sum_{i,j=0}^n v^* A_i (x^i 1_K, x^j 1_K) A_j^* v = v^* \langle P, P \rangle v = 0
\]
\[
\iff A_i^* v = 0 \text{ for all } 0 \leq i \leq n \iff v \in \bigcap_{i=0}^n \ker(A_i^*).
\]
Here the second equivalence follows from Lemma 1.12, while the third equivalence is a consequence of Lemma 2.23. Now we assume that \(A_n\) is regular, so that \(P\) has regular leading coefficient. Then also \(A_n^*\) is regular, so that \(\ker(A_n^*) = \{0\}\). By the above equivalences, we conclude that \(\ker(P, P) = \{0\}\). Put differently, \(\langle P, P \rangle\) is regular.

\((ii) \iff (iii)\): Consider the sequence \((x^n 1_K)_{n \geq 0}\). We will apply a procedure to this sequence similar to classical Gram-Schmidt orthogonalization, in order to construct an orthonormal sequence with respect to \(\langle \cdot, \cdot \rangle\) (assuming that \(\langle P, P \rangle\) is regular for every \(P\) with regular leading coefficient).

Put \(R_0 = 1_K\). By assumption, \(\langle R_0, R_0 \rangle\) is regular.

Next assume that we already have found matrix polynomials \(R_k\) for \(0 \leq k \leq n - 1\), such that the degree of \(R_k\) equals \(k\), the leading coefficient of \(R_k\) (and hence also \(\langle R_k, R_k \rangle\)) is regular, and the \(R_k\) are orthogonal, i.e. \(\langle R_k, R_l \rangle = \theta\) for \(0 \leq k, l \leq n - 1\), \(k \neq l\). Then define
\[
R_n(x) = x^n 1_K - \sum_{i=0}^{n-1} (x^n 1_K, R_i) \langle R_i, R_i \rangle^{-1} R_i(x).
\]

Note that \(R_n\) has degree \(n\), and moreover has regular leading coefficient, so that \(\langle R_n, R_n \rangle\) is regular. For every \(0 \leq k \leq n - 1\) we now have
\[
\langle R_n, R_k \rangle = \langle x^n 1_K, R_k \rangle - \sum_{i=0}^{n-1} \langle x^n 1_K, R_i \rangle \langle R_i, R_i \rangle^{-1} \langle R_i, R_k \rangle
\]
\[
= \langle x^n 1_K, R_k \rangle - \langle x^n 1_K, R_k \rangle \langle R_k, R_k \rangle^{-1} \langle R_k, R_k \rangle = \theta,
\]
in other words, \(R_n\) and \(R_k\) are orthogonal. As \(\langle R_n, R_n \rangle\) is regular and positive hermitian, so is its inverse by Lemma 1.11. In particular we can define its square root \(\langle R_n, R_n \rangle^{-\frac{1}{2}}\). Finally we define
\[
P_n = \langle R_n, R_n \rangle^{-\frac{1}{2}} R_n
\]
so that \((P_n)_{n \geq 0}\) is an orthonormal sequence with respect to \(\langle \cdot, \cdot \rangle\). Indeed,
\[
\langle P_n, P_m \rangle = \langle R_n, R_n \rangle^{-\frac{1}{2}} \langle R_n, R_m \rangle \langle R_m, R_m \rangle^{-\frac{1}{2}} = \theta
\]
whenever \(n \neq m\), and
\[
\langle P_n, P_n \rangle = \langle R_n, R_n \rangle^{-\frac{1}{2}} \langle R_n, R_n \rangle \langle R_n, R_n \rangle^{-\frac{1}{2}} = \langle R_n, R_n \rangle^{-\frac{1}{2}} \langle R_n, R_n \rangle \langle R_n, R_n \rangle^{-\frac{1}{2}} = 1_K.
\]

\((iii) \implies (i)\): Let \((P_n)_{n \geq 0}\) be an orthonormal sequence with respect to \(\langle \cdot, \cdot \rangle\). Assume that \(P\) is a matrix polynomial of degree \(m\) satisfying \(\langle P, P \rangle = \theta\). According to Proposition 2.20, \(P\) can be written as
\[
P(x) = \sum_{k=0}^m A_k P_k(x)
\]

38
for certain $A_k \in \mathbb{C}^{K \times K}$. Hence
\[
\theta = \langle P, P \rangle = \sum_{k,l=0}^{m} A_k \langle P_k, P_l \rangle A^*_l = \sum_{k,l=0}^{m} A_k \delta_{k,l} \mathbb{1}_K A^*_l = \sum_{k=0}^{m} A_k A^*_k.
\]
Thus, for every $v \in \mathbb{C}^{K}$ it holds that
\[
0 = v^* \left( \sum_{k=0}^{m} A_k A^*_k \right) v = \sum_{k=0}^{m} v^* A_k A^*_k v = \sum_{k=0}^{m} \| A_k v \|^2,
\]
so that $A^*_k v = 0$ for all $0 \leq k \leq m$ and $v \in \mathbb{C}^{K}$. It follows that $A^*_k = \theta$, and subsequently $A_k = \theta$, for all $0 \leq k \leq m$, which in turn implies that $P = \theta$. We conclude that $\langle \cdot, \cdot \rangle$ is non-degenerate.

Now suppose the equivalent conditions (i), (ii) and (iii) hold. Let $P$ be a matrix polynomial of degree $n$. Then, by Proposition 2.20, it can be written as $P = \sum_{k=0}^{n} A_k P_k$ for certain $A_k \in \mathbb{C}^{K \times K}$. As $(P_n)_{n \geq 0}$ is an orthonormal sequence, we obtain
\[
\langle P, P_k \rangle = \left( \sum_{l=0}^{n} A_l P_l, P_k \right) = \sum_{k=0}^{n} A_l \langle P_l, P_k \rangle = \sum_{l=0}^{n} A_l \delta_{k,l} = A_k.
\]

Lemma 2.32. Let $(P_n)_{n \geq 0}$ and $(Q_n)_{n \geq 0}$ be orthonormal sequences with respect to some non-degenerate inner product $\langle \cdot, \cdot \rangle$. Then there exists a sequence $(U_n)_{n \geq 0}$ of unitary matrices such that $Q_n = U_n P_n$.

Proof. By Theorem 2.31 every $Q_n$ can be written as
\[
Q_n = \sum_{k=0}^{n} A_k P_k \text{ where } A_k = \langle Q_n, P_k \rangle.
\]
Since $Q_n$ is orthogonal to all polynomials of degree $< n$, it follows that $A_k = 0$ for $0 \leq k \leq n - 1$, hence $Q_n = A_n P_n$. Then by orthonormality,
\[
\mathbb{1}_K = \langle Q_n, Q_n \rangle = \langle A_n P_n, A_n P_n \rangle = A_n \langle P_n, P_n \rangle A^*_n = A_n A^*_n
\]
which shows that $A_n$ is regular with $A_{n-1}^{-1} = A^*_n$, in other words $A_n$ is unitary. Defining $U_n = A_n$ for every $n \geq 0$ thus gives the desired result.

2.3.2 Properties of orthonormal matrix polynomials

In this and the following subsection some properties of orthonormal matrix polynomials with respect to a matrix measure $\mu$ are listed. Most of these results will be used while discussing the kernel polynomial in Section 2.4, or while treating the matrix moment problem from a functional analytic point of view (see Chapter 3).

Theorem 2.33 (Three-term recurrence relation). Let $(P_n)_{n \geq 0}$ be an orthonormal sequence of matrix polynomials with respect to a non-degenerate measure $\mu \in \mathcal{M}_K^{\times}$. Then there exist two sequences of matrices $(A_n)_{n \geq 0}$, $(B_n)_{n \geq 0}$ such that $A_n$ is regular for $n \geq 1$, $B_n$ is hermitian for $n \geq 0$ and moreover
\[
x P_n (x) = A_{n+1} P_{n+1} (x) + B_n P_n (x) + A^*_n P_{n-1} (x).
\]
In the equation for $n = 0$ we put $P_{-1} = A_0 = \theta$. 

39
Proof. By Theorem 2.31, the polynomial $xP_n(x)$ (which has degree $n + 1$) can be written as

$$xP_n(x) = \sum_{k=0}^{n+1} C_k P_k(x) \quad (2.30)$$

where $C_k \in \mathbb{C}^{K \times K}$ is given by $C_k = \langle xP_n, P_k \rangle$. Since $P_n$ is orthogonal to all matrix polynomials of degree less than $n$, we see that

$$C_k = \langle xP_n, P_k \rangle = \langle P_n, xP_k \rangle = 0 \quad (2.31)$$

for $k < n - 1$. Thus (2.30) can be simplified to

$$xP_n(x) = C_{n-1} P_{n-1}(x) + C_n P_n(x) + C_{n+1} P_{n+1}(x). \quad (2.32)$$

Now define

$$A_{n+1} = \langle xP_n, P_{n+1} \rangle, \quad (2.33)$$

$$B_n = \langle xP_n, P_n \rangle. \quad (2.34)$$

Then $A_{n+1} = C_{n+1}$, and $B_n = C_n$. Note that $B_n$ is hermitian, since

$$C_n^* = \langle xP_n, P_n \rangle^* = \langle P_n, xP_n \rangle = \langle xP_n, P_n \rangle = C_n. \quad (2.35)$$

Now suppose that the leading coefficient of $P_n$ is $L_n$, which is regular as $(P_n)_{n \geq 0}$ is a simple sequence. Comparing the term of order $n + 1$ in (2.29), we see that $L_n = A_{n+1} L_{n+1}$. Hence $A_{n+1} = L_n L_{n+1}^{-1}$ which is clearly regular. Thus $A_n$ is regular for all $n \geq 1$. 

Remark 2.34. Define the polynomials $(Q_n)_{n \geq 0}$ as in Remark 2.30. These polynomials also satisfy a three-term recurrence relation, with $A_n$ and $B_n$ replaced by respectively $\bar{A}_n = U_{n-1} A_n U_n^*$ and $\bar{B}_n = U_n B_n U_n^*$. Indeed, since $P_n = U_n^* Q_n$, it follows from (2.29) that

$$xU_n^* Q_n(x) = A_{n+1} U_{n+1}^* Q_{n+1}(x) + B_n U_n^* Q_n(x) + A_n^* U_{n-1}^* Q_{n-1}(x).$$

Multiplying with $U_n$ from the left gives the result.

In Chapter 3 the moment problem will be studied via the so-called Jacobi operator, which is defined in terms of the matrices $A_n$ and $B_n$ that appear in the three-term recurrence relation.

The following theorem comprises the statement given in Theorem 2.33 in the opposite direction. Thus, combining these theorems, a characterization of orthonormal matrix polynomials via the three-term recurrence relation is established.

Theorem 2.35 (Favard). Let $(P_n)_{n \geq 0}$ be a sequence of matrix polynomials satisfying a three-term recurrence relation (2.29) where $P_0 = 1_K$, $P_{-1} = \theta$, $A_n$ is regular for $n \geq 1$ and $B_n$ is hermitian for $n \geq 0$. Then there exists a non-degenerate $\mu \in \mathcal{M}_K^*$ such that $(P_n)_{n \geq 0}$ is an orthonormal sequence with respect to $\mu$. 

40
Proof. From the initial conditions $P_0 = \mathbb{1}_K$, $P_{-1} = \theta$, all other polynomials $P_n$ are determined recursively. For instance $P_1$ follows by plugging in $n = 0$, and is given by $P_1(x) = A_{1}^{-1}(1_Kx - B_0)$. Note that $P_0$ and $P_1$ have degree 0 and 1, respectively.

Next assume that $P_n$ has degree $n$ with regular leading coefficient $L_n$. Then $P_{n+1}$ has degree $n + 1$ with leading coefficient $L_{n+1} = A_{n+1}^{-1}L_n$ (see the proof of Theorem 2.33), which is also regular. Thus $(P_n)_{n \geq 0}$ is a simple sequence by induction. Therefore we can define a matrix inner product by

$$\langle \sum_k C_k P_k, \sum_l D_l P_l \rangle = \sum_k C_k D_k^*.$$ 

Note that $(P_n)_{n \geq 0}$ is orthonormal with respect to this inner product, since

$$\langle \sum_k C_k P_k, \sum_l D_l P_l \rangle = \sum_{k,l} C_k \langle P_k, P_l \rangle D_l^*$$

and thus $\langle P_k, P_l \rangle = \delta_{k,l} \mathbb{1}_K$. It then follows from Theorem 2.31 that the above defined inner product is non-degenerate.

Next we invoke the three-term recurrence relation and the fact that all $B_n$ are hermitian to show that

$$\langle xP_n, P_m \rangle = \langle A_{n+1}P_{n+1} + B_n P_n + A_n^* P_{n-1}, P_m \rangle$$

$$= A_{n+1} \langle P_{n+1}, P_m \rangle + B_n \langle P_n, P_m \rangle + A_n^* \langle P_{n-1}, P_m \rangle$$

$$= A_{n+1} \delta_{n+1,m} + B_n \delta_{n,m} + A_n^* \delta_{n-1,m}$$

$$= A_{n+1} \delta_{n,m+1} + B_m \delta_{n,m} + A_m \delta_{n,m-1}$$

$$= (P_n, P_{m+1}) A_{m+1}^* + (P_n, P_m) B_m + (P_n, P_{m-1}) A_m$$

$$= (P_n, A_{m+1} P_{m+1} + B_m P_m + A_m^* P_{m-1}) = \langle P_n, xP_m \rangle$$

which in turn can be extended to

$$\langle xP, Q \rangle = \langle P, xQ \rangle$$

for all $P, Q \in \mathbb{C}^{K \times K}[x]$. Defining $S_n = \langle x^n \mathbb{1}_K, \mathbb{1}_K \rangle$ for $n \geq 0$, we see that by repeatedly applying (2.36) it holds that

$$S_n = \langle x^n \mathbb{1}_K, x^n \mathbb{1}_K \rangle = \langle \mathbb{1}_K, x^n \mathbb{1}_K \rangle = \langle x^n \mathbb{1}_K, \mathbb{1}_K \rangle^* = S_n^*$$

where $k + l = n$. Now consider an arbitrary matrix polynomial $P$, say $P(x) = \sum_{k=0}^N C_k x^k$. Then by the definition of the matrix inner product,

$$\theta \prec (P, P) = \sum_{k,l=0}^N C_k \langle x^k \mathbb{1}_K, x^l \mathbb{1}_K \rangle C_l^* = \sum_{k,l=0}^N C_k S_{k+l} C_l^*.$$

(2.37)

In particular it holds that

$$\text{tr} \left( \sum_{k,l=0}^N C_k S_{k+l} C_l^* \right) > 0$$

(2.38)

as the eigenvalues of a positive hermitian matrix are positive (see Lemma 1.9).

Let $v_0, v_1, \ldots, v_N \in \mathbb{C}^K$ be given. Then choose $C_k$ in such a way that the first row is equal to $v_k^*$, and let all other entries be equal to zero, i.e.

$$C_k = \begin{pmatrix} v_k^* \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$
Denoting the standard basis of $\mathbb{C}^K$ by $\{e_i : 0 \leq i \leq K - 1\}$, we see that $v^*_k = e^*_0 C_k$ and thus $v_k = C^*_k e_0$. Hence
\[
\sum_{k,l=0}^{N} v^*_k S_{k+l} v_l = e^*_0 \left( \sum_{k,l=0}^{N} C_k S_{k+l} C_l^* \right) e_0 \tag{2.39}
\]
while
\[
e^*_i \left( \sum_{k,l=0}^{N} C_k S_{k+l} C_l^* \right) e_i = 0 \quad (i \neq 0). \tag{2.40}
\]
Combining (2.38), (2.39) and (2.40) it follows that
\[
\sum_{k,l=0}^{N} v^*_k S_{k+l} v_l = \text{tr} \left( \sum_{k,l=0}^{N} C_k S_{k+l} C_l^* \right) > 0
\]
hence $(S_n)_{n \geq 0}$ is a sequence of positive definite matrices, and thus a moment sequence according to Theorem 2.14. Thus the inner product defined above is given by some $\mu \in \mathcal{M}^*_K$. \qed

2.3.3 Matrix polynomials of the first and second kind

In this section, let $(P_n)_{n \geq 0}$ and $(Q_n)_{n \geq 0}$ be sequences of matrix polynomials satisfying the three-term recurrence relation (2.29) for $n \geq 1$ with initial conditions
\[
P_0 = 1, P_1(x) = A_1^{-1} (1_K x - B_0) \quad \text{and} \quad Q_0 = \theta, Q_1 = A_1^{-1}. \tag{2.41}
\]
Here $A_n$ is regular for $n \geq 1$ and $B_n$ is hermitian for $n \geq 0$, as before. The polynomials $P_n$ and $Q_n$ are often referred to as the matrix polynomials of the first kind and second kind, respectively.

**Lemma 2.36.** The polynomials $P_n$ and $Q_n$ are related by
\[
Q_n(x) = \int_{\mathbb{R}} \frac{P_n(x) - P_n(t)}{x - t} d\mu(t). \tag{2.42}
\]

**Proof.** First we check the initial conditions. Since $P_0 = 1$ and $P_1(x) = A_1^{-1} (1_K x - B_0)$, we see that
\[
\int_{\mathbb{R}} \frac{P_0(x) - P_0(t)}{x - t} d\mu(t) = \int_{\mathbb{R}} \frac{1_K - 1_K}{x - t} d\mu(t) = \theta = Q_0
\]
and
\[
\int_{\mathbb{R}} \frac{P_1(x) - P_1(t)}{x - t} d\mu(t) = \int_{\mathbb{R}} \frac{A_1^{-1} (1_K x - B_0) - A_1^{-1} (1_K t - B_0)}{x - t} d\mu(t) = \int_{\mathbb{R}} A_1^{-1} d\mu(t) = A_1^{-1} = Q_1.
\]
We thus only need to show that the polynomials (2.42) also satisfy the three-term recurrence relation (2.29). Using that (2.29) holds for the polynomials $P_n$, we obtain
\[
x (P_n(x) - P_n(t)) + (x - t) P_n(t) = x P_n(x) - t P_n(t) = A_{n+1} (P_{n+1}(x) - P_{n+1}(t)) + B_n (P_n(x) - P_n(t)) + A^*_n (P_{n-1}(x) - P_{n-1}(t)).
\]
Then dividing by $x - t$ and integrating with respect to $t$ gives
\[
x \int_{\mathbb{R}} \frac{P_n(x) - P_n(t)}{x - t} d\mu(t) + \int_{\mathbb{R}} P_n(t) d\mu(t) = A_{n+1} \int_{\mathbb{R}} \frac{P_{n+1}(x) - P_{n+1}(t)}{x - t} d\mu(t)
\]
\[
+ B_n \int_{\mathbb{R}} \frac{P_n(x) - P_n(t)}{x - t} d\mu(t) + A^*_n \int_{\mathbb{R}} \frac{P_{n-1}(x) - P_{n-1}(t)}{x - t} d\mu(t).
\]
By remarking that $\int_{\mathbb{R}} P_n(t) d\mu(t) = \langle P_n, P_0 \rangle = \theta$ for $n \geq 1$, the claim follows. \qed
We conclude this section by stating two results regarding these particular polynomials.

**Theorem 2.37** (Liouville-Ostrogradsky). Let \((P_n)_{n \geq 0}, (Q_n)_{n \geq 0}\) satisfy (2.29) with initial conditions (2.41). Then for \(x \in \mathbb{C}\) we have

\[
Q_k(x)P_k^*(x) = P_k(x)Q_k^*(x) \quad \text{for} \quad k \geq 0
\]  

(2.43) and

\[
Q_k(x)P_{k-1}^*(x) - P_k(x)Q_{k-1}^*(x) = A_k^{-1} \quad \text{for} \quad k \geq 1.
\]  

(2.44)

**Proof.** Note that

\[
Q_0(x)P_0^*(x) = \theta = P_0(x)Q_0^*(x)
\]

and

\[
Q_1(x)P_1^*(x) = A_1^{-1}(\mathbb{1}_K x - B_0) (A_1^{-1})^* = P_1(x)Q_1^*(x),
\]

so that (2.43) holds for \(k = 0\) and \(k = 1\). Furthermore (2.44) holds for \(k = 1\), as

\[
Q_1(x)P_0^*(x) - P_1(x)Q_0^*(x) = A_1^{-1}\mathbb{1}_K - A_1^{-1}(\mathbb{1}_K x - B_0) \theta = A_1^{-1}.
\]

In the following we assume that both formulas hold for all \(k \leq n\).

We now multiply (2.29) for \(P_n\) by \(Q_n^*\) from the right, and the corresponding relation for \(Q_n\) by \(P_n^*\) from the right. Subtracting the latter equation from the former gives

\[
x(P_n(x)Q_n^*(x) - Q_n(x)P_n^*(x)) = A_{n+1}(P_{n+1}(x)Q_n^*(x) - Q_{n+1}(x)P_n^*(x)) \\
+ B_n(P_n(x)Q_n^*(x) - Q_n(x)P_n^*(x)) \\
+ A_n^*(P_{n-1}(x)Q_n^*(x) - Q_{n-1}(x)P_n^*(x)).
\]  

(2.45)

Invoking (2.43) for \(k = n\) then yields

\[
A_{n+1}(P_{n+1}(x)Q_n^*(x) - Q_{n+1}(x)P_n^*(x)) = -A_n^*(P_{n-1}(x)Q_n^*(x) - Q_{n-1}(x)P_n^*(x)).
\]  

(2.46)

Taking the adjoint of (2.44) for \(k = n\) (and replacing \(x\) by \(\overline{x}\)) gives

\[
P_{n-1}(x)Q_n^*(x) - Q_{n-1}(x)P_n^*(x) = (A_n^{-1})^*.
\]  

(2.47)

Combining (2.46) and (2.47) now gives

\[
A_{n+1}(P_{n+1}(x)Q_n^*(x) - Q_{n+1}(x)P_n^*(x)) = -A_n^*(A_n^{-1})^* = -\mathbb{1}_K.
\]  

(2.48)

This equation can be rewritten as

\[
Q_{n+1}(x)P_n^*(x) - P_{n+1}(x)Q_n^*(x) = A_{n+1}^{-1}
\]  

(2.49)

so that (2.44) holds for \(k = n + 1\).

Taking the adjoint of (2.29) for \(Q_n\) gives

\[
Q_{n+1}^*(x) = (Q_n^*(x)(\mathbb{1}_K x - B_n) - Q_{n-1}(x)A_n) (A_{n+1}^*)^{-1}.
\]  

(2.50)

Inserting this result in the formula

\[
xP_n(x)Q_{n+1}^*(x) = A_{n+1}P_{n+1}(x)Q_n^*(x) + B_nP_n(x)Q_{n+1}(x) + A_n^*P_{n-1}(x)Q_{n+1}^*(x),
\]  

(2.51)
which is obtained by multiplying the three term recurrence relation (2.29) for $P_n$ by $Q_{n+1}^{*}$ from the right, then gives

$$A_{n+1}P_{n+1}(x)Q_{n+1}^{*}(x) = (1_Kx - B_n) P_n(x)Q_{n+1}^{*}(x) - A_{n}^{*} P_{n-1}(x) (Q_{n}^{*}(x)(1_Kx - B_n) - Q_{n-1}^{*}(x)A_n) (A_{n+1}^{*})^{-1}.$$

(2.52)

Similarly it holds that

$$A_{n+1}Q_{n+1}(x)P_{n+1}^{*}(x) = (1_Kx - B_n) Q_n(x)P_{n+1}^{*}(x) - A_{n}^{*} Q_{n-1}(x) (P_{n}^{*}(x)(1_Kx - B_n) - P_{n-1}^{*}(x)A_n) (A_{n+1}^{*})^{-1}$$

(2.53)

by interchanging the roles of $P_n$ and $Q_n$. Now subtracting (2.53) from (2.52) yields

$$A_{n+1} (P_{n+1}(x)Q_{n+1}^{*}(x) - Q_{n+1}(x)P_{n+1}^{*}(x)) = (1_Kx - B_n) (P_n(x)Q_{n+1}^{*}(x) - Q_n(x)P_{n+1}^{*}(x)) - A_{n}^{*} (P_{n-1}(x)Q_{n}^{*}(x) - Q_{n-1}(x)P_{n}^{*}(x)) (1_Kx - B_n) (A_{n+1}^{*})^{-1} + A_{n}^{*} (P_{n-1}(x)Q_{n-1}^{*}(x) - Q_{n-1}(x)P_{n-1}^{*}(x)) A_n (A_{n+1}^{*})^{-1}.$$

(2.54)

The third term on the right-hand side is equal to $\theta$, since we assumed (2.43) to hold for $k \leq n$, in particular for $k = n - 1$. Invoking (2.50) moreover simplifies the first two terms on the right-hand side (note that (2.50) also holds if we replace $n$ by $n + 1$, since we have already shown that (2.44) holds for $k = n + 1$). We thus obtain

$$A_{n+1} (P_{n+1}(x)Q_{n+1}^{*}(x) - Q_{n+1}(x)P_{n+1}^{*}(x)) = (1_Kx - B_n) (A_{n}^{*}(1_Kx - B_n)) (A_{n+1}^{*})^{-1} + A_{n}^{*} (P_{n-1}(x)Q_{n-1}^{*}(x) - Q_{n-1}(x)P_{n-1}^{*}(x)) A_n (A_{n+1}^{*})^{-1} = \theta,$$

(2.55)

and by regularity of $A_{n+1}$ also $P_{n+1}(x)Q_{n+1}^{*}(x) - Q_{n+1}(x)P_{n+1}^{*}(x) = \theta$. Therefore (2.43) holds for $k = n + 1$. By simultaneous induction it is thus shown that (2.43) holds for every $k \geq 0$, and (2.44) for every $k \geq 1$.

**Theorem 2.38 (Christoffel-Darboux).** Let $(P_n)_{n \geq 0}$ satisfy (2.29) with initial conditions (2.41). Then

$$P_{n-1}^{*}(x)A_n P_n(y) - P_n^{*}(x)A_n^{*} P_{n-1}(y) = (y - x) \sum_{k=0}^{n-1} P_k^{*}(x) P_k(y)$$

for $x, y \in \mathbb{C}$

(2.56)

with the special cases

$$P_{n-1}^{*}(z)A_n P_n(z) - P_n^{*}(z)A_n^{*} P_{n-1}(z) = \theta$$

(2.57)

and

$$P_{n-1}^{*}(z)A_n P_n'(z) - P_n^{*}(z)A_n^{*} P_{n-1}'(z) = \sum_{k=0}^{n-1} P_k^{*}(z) P_k(z)$$

(2.58)

for $z \in \mathbb{C}$.

**Proof.** Note that

$$y P_n^{*}(x) P_n(y) = P_n^{*}(x) (y P_n(y)) = P_n^{*}(x) A_{n+1} P_{n+1}(y) + P_n^{*}(x) B_n P_n(y) + P_n^{*}(x) A_n^{*} P_{n-1}(y)$$

(2.59)
and similarly
\[ xP_n^*(x)P_n(y) = (\tau P_n(\tau))^* P_n(y) = P_{n+1}^*(x)A_{n+1}^* P_n(y) + P_n^*(x)B_n P_n(y) + P_{n-1}^*(x)A_n P_n(y) \]  
(2.60)

by exploiting (2.29). Hence, defining
\[ \omega_n(x, y) = P_n^*(x)A_{n+1} P_{n+1}(y) - P_{n+1}^*(x)A_{n+1}^* P_n(y) \]  
(2.61)

for \( n \geq 0 \), we see that
\[ (y - x)P_n^*(x)P_n(y) = \omega_n(x, y) - \omega_{n-1}(x, y) \]  
(2.62)

for all \( n \geq 1 \). Thus by (2.41),
\[
(y - x) \sum_{k=0}^{n-1} P_k^*(x)P_k(y) = (y - x) \left( \sum_{k=0}^{n-1} P_k^*(x)P_k(y) \right)
\]
\[
= (y - x)1_K + \omega_{n-1}(x, y) - \omega_0(x, y)
\]
\[
= \omega_{n-1}(x, y)
\]
\[
= P_{n-1}^*(x)A_n P_n(y) - P_n^*(x)A_n^* P_{n-1}(y),
\]  
(2.63)

hence (2.56) holds. Moreover (2.57) follows by letting \( x = y = z \) in (2.56), while (2.58) follows from (2.56) by differentiation with respect to \( y \) and then substituting \( x = y = z \).

The following result relates convergence of the matrix polynomials of the first and second kind. In Chapter 3 and 4 it will be linked to determinacy of the matrix moment problem.

**Theorem 2.39.** If for fixed \( z_0 \in \mathbb{C} \setminus \mathbb{R} \) the series
\[ \sum_{k=0}^{\infty} P_k(z_0)^* P_k(z_0) \]
is convergent, then the following statements hold.

(i) The series
\[ \sum_{k=0}^{\infty} Q_k(z_0)^* Q_k(z_0) \]
converges.

(ii) The series
\[ \sum_{k=0}^{\infty} P_k(z)^* P_k(z) \text{ and } \sum_{k=0}^{\infty} Q_k(z)^* Q_k(z) \]
are both uniformly convergent in any compact subset of the complex plane.

This result is stated in [22], Theorem 3.2.
2.3.4 The kernel polynomial

As in the scalar case (see Section 1.3), we here introduce the so-called kernel polynomial.

**Definition 2.40.** The reproducing kernel for the polynomials of degree \( \leq N \) is defined as

\[
K_N(x, y) = \sum_{k=0}^{N} P_k^*(x)P_k(y).
\]  

(2.64)

and is called the **kernel polynomial**. ■

Now suppose that \( P_k(x) = \sum_{i=0}^{k} B_i^{(k)} x^i \) where \( B_i^{(k)} \in \mathbb{C}^{K \times K}. \) Then

\[
K_N(x, y) = \sum_{k=0}^{N} \left( \sum_{i=0}^{k} B_i^{(k)} x^i \sum_{j=0}^{k} B_j^{(k)} y^j \right)
\]

\[= \sum_{k=0}^{N} \sum_{i=0}^{k} \sum_{j=0}^{k} \left( B_i^{(k)} \right)^* B_j^{(k)} x^i y^j
\]

\[= \sum_{i=0}^{N} \sum_{j=0}^{N} A_{ij}^{(N)} x^i y^j
\]  

(2.65)

where \( A_{ij}^{(N)} \in \mathbb{C}^{K \times K} \) satisfies

\[
A_{ij}^{(N)} = \sum_{k=\max(i,j)}^{N} \left( B_i^{(k)} \right)^* B_j^{(k)}.
\]  

(2.66)

Observe that the matrix \( A_N \), defined by \( A_N = \left( A_{ij}^{(N)} \right)_{0 \leq i,j \leq N} \), is hermitian, as

\[
\left( A_{ij}^{(N)} \right)^* = \sum_{k=\max(i,j)}^{N} \left( B_j^{(k)} \right)^* B_i^{(k)} = A_{ji}^{(N)}.
\]  

(2.67)

We will prove that \( A_N \) is the inverse of \( H_N \) (this is the matrix-valued generalization of Theorem 1.25).

**Theorem 2.41.** The matrix \( A_N \) is the inverse of \( H_N \), i.e.

\[
A_N H_N = \mathbb{1}_{K(K+1)} = H_N A_N
\]

where \( \mathbb{1}_{K(K+1)} \) is the \( K(K+1) \times K(K+1) \) unit matrix.

**Proof.** Recall that

\[
\int_{\mathbb{R}} P_l(x)W(x)P_k^*(x)d\tau_\mu(x) = \delta_{k,l}\mathbb{1}_K,
\]  

(2.69)

so that

\[
\int_{\mathbb{R}} K_N(y,x)W(x)P_k^*(x)d\tau_\mu(x) = \int_{\mathbb{R}} \left( \sum_{l=0}^{N} P_l^*(y)P_l(x) \right) W(x)P_k^*(x)d\tau_\mu(x)
\]

\[= \sum_{l=0}^{N} P_l^*(y) \int_{\mathbb{R}} P_l(x)W(x)P_k^*(x)d\tau_\mu(x)
\]

\[= \sum_{l=0}^{N} P_l^*(y)\delta_{k,l}\mathbb{1}_K = P_k^*(y)
\]  

(2.70)
for $0 \leq k \leq N$, and thus in particular for these values of $k$ it holds that
\[ \int_{\mathbb{R}} K_N(y, x) W(x) x^k \mathbb{1}_K d\tau_\mu(x) = y^k \mathbb{1}_K. \] (2.71)

We thus obtain
\[
y^k \mathbb{1}_K = \int_{\mathbb{R}} \left( \sum_{i=0}^{N} \sum_{j=0}^{N} A_{ij}^{(N)} y^j x^i \right) W(x) x^k \mathbb{1}_K d\tau_\mu(x)
= \sum_{i=0}^{N} \sum_{j=0}^{N} A_{ij}^{(N)} y^j \int_{\mathbb{R}} x^j W(x) x^k d\tau_\mu(x)
= \sum_{i=0}^{N} \sum_{j=0}^{N} A_{ij}^{(N)} S_{j+k} y^j.
\] (2.72)

Hence
\[
\sum_{j=0}^{N} A_{ij}^{(N)} (H_N)_{jk} = \sum_{j=0}^{N} A_{ij}^{(N)} S_{j+k} = \delta_{i,k} \mathbb{1}_K,
\] (2.73)

for $0 \leq k \leq N$, i.e. $A_N H_N = \mathbb{1}_{K(N+1)}$. Moreover
\[
\delta_{i,k} \mathbb{1}_K = \sum_{j=0}^{N} S_{j+k}^{*} \left(A_{ij}^{(N)}\right)^{*} = \sum_{j=0}^{N} S_{j+k} A_{ji}^{(N)} = \sum_{j=0}^{N} (H_N)_{jk} A_{ji}^{(N)} = \sum_{j=0}^{N} (H_N)_{kj} A_{ji}^{(N)},
\] (2.74)

so that $H_N A_N = \mathbb{1}_{K(N+1)}$. We conclude that $A_N$ is the inverse of $H_N$. □

By using the Christoffel-Darboux formula (2.56), the kernel polynomial can be rewritten as
\[
K_N(x, y) = \frac{1}{y - x} \left(P_N^{*}(x) A_{N+1} P_{N+1}(y) - P_{N+1}^{*}(x) A_{N+1}^{*} P_N(y)\right). \] (2.75)

As we will see in Chapter 3, the boundedness of $v^{*}K_N(\overline{z}, z)v$ (where $v$ ranges over $\mathbb{C}^K$) correlates with the degree of determinacy of the associated moment problem. Therefore we introduce the following notation.

**Notation 2.42.** We define
\[
R_N(z) = \left( \sum_{k=0}^{N} P_k^{*}(\overline{z}) P_k(z) \right)^{-1} = K_N(\overline{z}, z)^{-1}.
\] (2.76)

Note that, since $\sum_{k=1}^{N} P_k(z)^{*} P_k(z)$ is positive definite, we have
\[
\sum_{k=0}^{N} P_k^{*}(\overline{z}) P_k(z) = \mathbb{1}_K + \sum_{k=1}^{N} P_k(z)^{*} P_k(z) \succ \mathbb{1}_K,
\]
so that $\sum_{k=0}^{N} P_k^{*}(\overline{z}) P_k(z)$ is invertible, and $R_N(z)$ is thus well-defined.

For any $v \in \mathbb{C}^K$, we have the equality
\[
v^{*}K_N(\overline{z}, z)v = v^{*}R_N(z)^{-1}v = \sum_{k=0}^{N} v^{*} P_k^{*}(\overline{z}) P_k(z) v = \sum_{k=0}^{N} \|P_k(z)v\|^2,
\] (2.77)

which will play an important role in Chapter 3.
Remark 2.43. One can also consider the limit matrix $R(z)$, defined by $R(z) = \lim_{N \to \infty} R_N(z)$ whenever it exists, where the convergence has to be interpreted with respect to the matrix norm. However, by taking into consideration that the boundedness of $v^*K_N(\tau, z)v$ (with $v \in \mathbb{C}^K$) will be of importance in the following chapter, we define $V = \{v \in \mathbb{C}^K : \sup_N \langle K_N(\tau, z)v, v \rangle < \infty\}$, and then consider $R(z) : V \to \mathbb{C}^{K \times K}$ defined by $R(z)v = \lim_{N \to \infty} R_N(z)v$ (thus we restrict the domain of $R(z)$). In fact, the case $V = \mathbb{C}^K$ corresponds to the previous definition of $R(z)$, i.e. where the convergence is meant with respect to the matrix norm. The dimension of $V$ is related to the deficiency indices; see Section 3.2.

The sequence $(R_N(z))_{N \geq 0}$ has the following property.

Proposition 2.44. For $z \in \mathbb{C} \setminus \mathbb{R}$ the following two statements are equivalent:

(i) The sequence $R_N(z)^{-1}$ converges for $N \to \infty$.

(ii) The matrix $R(z)$ is invertible.

The proof can be found in [29], Proposition 3.1.8.

2.4 Some difficulties that arise while generalizing Theorem 1.18

The intended purpose of this thesis was to give a generalization of Theorem 1.18 to the matrix-valued case. Theorem 4.2 is a partial generalization, in the sense that it contains a statement regarding matrix moment problems with the ‘highest degree of determinacy’ (such a moment problem is called completely indeterminate; this terminology will be introduced formally in Section 3.2). The purpose in this thesis however, was to formulate a result also for the intermediate cases, i.e. moment problems with a smaller degree of indeterminacy. In Section 3.2 this degree is measured by the so-called deficiency indices. It is in fact related to the convergence behaviour of $(R_N(z))_{N \geq 0}$, via (2.77).

From the proof given in Section 1.3 it is clear that a lot of the computations can easily be extended. The main problem however, lies in using the Rayleigh quotient to determine the smallest eigenvalues $\lambda_N$ of the Hankel matrices. In order to generalize the proof of Theorem 1.18, while altering as less as possible, we need an extension of the concept of eigenvalues, i.e. some kind of eigenmatrix.

We first tried to generalize the notion of eigenvalues as either $H_NV = V\Lambda$ or $H_NV = \Lambda V$. Here the ‘eigenvector’ $V$ is of the form

$$V = \begin{pmatrix} V_0 \\ V_1 \\ \vdots \\ V_N \end{pmatrix},$$

where $V_i \in \mathbb{C}^{K \times K}$. Thus in the first case, $\Lambda \in \mathbb{C}^{K \times K}$, while in the second case it must hold that $\Lambda \in \mathbb{C}^{K(N+1) \times K(N+1)}$. In the latter case we interpret $\Lambda$ as a diagonal matrix, i.e. $\Lambda \equiv \text{diag}(\Lambda, \ldots, \Lambda)$. Now the above two eigenequations are written out as respectively

$$H_NV = V\Lambda = \begin{pmatrix} V_0\Lambda \\ V_1\Lambda \\ \vdots \\ V_N\Lambda \end{pmatrix} \quad \text{(2.78)}$$

48
and
\[ H_N V = \Lambda V \equiv \left( \begin{array}{ccc} \Lambda & \cdots & \Lambda \\ \vdots & \ddots & \vdots \\ \Lambda & \cdots & \Lambda \end{array} \right) V = \left( \begin{array}{c} \Lambda V_0 \\ \Lambda V_1 \\ \vdots \\ \Lambda V_N \end{array} \right). \] (2.79)

**Remark 2.45.** Suppose we consider the eigenequation (2.78). Mimicking the expression appearing in (1.13), we try
\[ \Lambda = \sum_{i,j=0}^{N} V^*_i S_{i+j} V_j. \] (2.80)

Writing out the eigenequation yields
\[
\begin{pmatrix}
S_0 & S_1 & \cdots & S_N \\
S_1 & S_2 & \cdots & S_{N+1} \\
\vdots & \vdots & \ddots & \vdots \\
S_N & S_{N+1} & \cdots & S_{2N}
\end{pmatrix}
\begin{pmatrix} V_0 \\ V_1 \\ \vdots \\ V_N \end{pmatrix} =
\begin{pmatrix} V_0 \Lambda \\ V_1 \Lambda \\ \vdots \\ V_N \Lambda \end{pmatrix}.
\]

which is equivalent to the following system of equations:
\[ \sum_{j=0}^{N} S_{k+j} V_j = V_k \Lambda \quad (0 \leq k \leq N). \] (2.81)

Filling in our assumption (2.80) into (2.81) gives
\[
\sum_{j=0}^{N} S_{k+j} V_j = \sum_{i,j=0}^{N} V_k V^*_i S_{i+j} V_j
\]
so that \( V_k V^*_i = \delta_{i,k} \mathbb{I}_K \). \[ \blacksquare \]

We considered some simple examples, i.e. matrix measures with diagonalizable weights.

**Example 2.46.** Consider the matrix moment problem with weight function
\[
W(x) = \frac{1}{\sum_{i=1}^{K} w_i(x)} M \begin{pmatrix}
w_1(x) \\
w_2(x) \\
\vdots \\
w_K(x)
\end{pmatrix} M^* =: \frac{1}{\sum_{i=1}^{K} w_i(x)} MD(x)M^* \] (2.82)

where we assume \( M \) to be unitary and all weights \( w_i \) to be absolutely continuous with respect to the Lebesgue measure. By a result of [25], page 294, it follows that for any \( E \in \mathbb{E} \), we have
\[
\mu(E) = \int_E d\mu(x) = \int_E W(x) d\tau_\mu(x) = \int_E D(x) dx,
\]
where \( d\tau_\mu(x) = \sum_{i=1}^{K} w_i(x) dx \).

Now assume that \( M = \mathbb{I}_K \). The moments of weight \( w_i \) are denoted by \( s_n^i \), i.e.
\[
s_n^i = \int_{\mathbb{R}} x^n w_i(x) dx
\]
for \( i = 1, \ldots, K \). Likewise the corresponding Hankel matrices are denoted by \( H_N^i \), hence

\[
H_N^i = \begin{pmatrix}
  s_0^i & s_1^i & \cdots & s_N^i \\
  s_1^i & s_2^i & \cdots & s_{N+1}^i \\
  \vdots & \vdots & \ddots & \vdots \\
  s_N^i & s_{N+1}^i & \cdots & s_{2N}^i
\end{pmatrix}.
\]

Denoting the smallest eigenvalue of \( H_N^i \) by \( \lambda_N^i \), we consider the eigenequations \( H_N^i v_i = \lambda_N^i v_i \) for \( 1 \leq i \leq K \), where the \( v_i \) are the corresponding eigenvectors with components \( v_i^0, v_i^1, \ldots, v_i^N \).

Writing out these equations yields

\[
\sum_{k=0}^{N} s_{k+l}^i v_k = \lambda_N^i v_l \quad (0 \leq l \leq N).
\] (2.83)

The moments \( S_n \) are computed to be

\[
S_n = \int_{\mathbb{R}} x^n W(x) d\tau(x) = \int_{\mathbb{R}} x^n \begin{pmatrix}
  w_1(x) \\
  w_2(x) \\
  \vdots \\
  w_K(x)
\end{pmatrix} dx = \begin{pmatrix}
  \int_{\mathbb{R}} x^n w_1(x) dx \\
  \int_{\mathbb{R}} x^n w_2(x) dx \\
  \vdots \\
  \int_{\mathbb{R}} x^n w_K(x) dx
\end{pmatrix} = \begin{pmatrix}
  s_1^n \\
  s_2^n \\
  \vdots \\
  s_K^n
\end{pmatrix}.
\]

Now we form the matrix \( \Lambda \) in terms of the smallest eigenvalues \( \lambda_N^i \) by \( \Lambda = \text{diag} (\lambda_N^1, \lambda_N^2, \ldots, \lambda_N^K) \). The corresponding eigenvector \( V \) is then given by

\[
V = \begin{pmatrix}
  V_0 \\
  V_1 \\
  \vdots \\
  V_N
\end{pmatrix} \quad \text{where} \quad V_k = \begin{pmatrix}
  v_k^1 \\
  v_k^2 \\
  \vdots \\
  v_k^K
\end{pmatrix}.
\]
Indeed, using (2.83) we obtain

\[
H_N V = \begin{pmatrix}
S_0 & S_1 & \cdots & S_N \\
S_1 & S_2 & \cdots & S_{N+1} \\
\vdots & \vdots & \ddots & \vdots \\
S_N & S_{N+1} & \cdots & S_{2N}
\end{pmatrix}
\begin{pmatrix}
V_0 \\
V_1 \\
\vdots \\
V_N
\end{pmatrix} =
\begin{pmatrix}
\sum_{k=0}^N S_k V_k \\
\sum_{k=0}^N S_{k+1} V_k \\
\vdots \\
\sum_{k=0}^N S_{k+N} V_k
\end{pmatrix}
\]

\[
= \begin{pmatrix}
\lambda_0^1 v_0^1 \\
\lambda_0^2 v_0^2 \\
\vdots \\
\lambda_N v_0^N
\end{pmatrix}
\]

Invoking Theorem 1.18, we see that the matrix weight \( W(x) \) given by (2.82) is determinate if and only if \( \text{tr}(\Lambda N) \to 0 \).

The above result can be extended to arbitrary unitary \( M \). Denote the Hankel matrices, moments and ‘eigenmatrix’ in this case by respectively \( H'_N, S'_N \) and \( \Lambda' \). It is easy to see that \( S'_N = MS_N M^* \) and \( H'_N = MH_N M^* \). Consider the eigenequation \( H'_N U = U \Lambda' \). Defining \( V = M^* U \), so that \( U = (M^*)^{-1} V \), we have

\[
MH_N V = MH_N M^* U = H'_N U = U \Lambda' = (M^*)^{-1} V \Lambda'.
\]

Invoking unitarity of \( M \) yields \( H_N V = M^* M H_N V = V \Lambda' \), so that the eigenequation is reduced to the form discussed in case \( M = 1_K \).

Several attempts have been made to say something about the (in)determinacy of the matrix moment problem in the same sense as Theorem 1.18, such as letting \( \Lambda \) be a diagonal matrix with the eigenvalues of \( H_N \) as entries on the diagonal, and then looking at the convergence behaviour of ‘blocks of eigenvalues’. Unfortunately, none of these attempts have given a satisfactory result.

In the next chapter we therefore consider the moment problem from another point of view, namely the functional analytic one. This approach is based on the so-called Jacobi operator \( J \), which will be defined in terms of the coefficients appearing in the three-term recurrence relation of the orthonormal matrix polynomials.
Chapter 3

The operator approach to the moment problem

In the previous chapter the matrix moment problem has been introduced and studied mostly by considering the associated orthonormal polynomials. Here we will apply spectral theory in order to get a better understanding of the moment problem. See Appendix B for a summary of spectral theory.

This chapter is largely based on [22] and [17].

3.1 The Jacobi operator

It turns out that the (in)determinacy of the matrix moment problem is related to the indices of deficiency of the so-called Jacobi operator, which we will denote by $J$. Before being able to give a definition of this operator, we will need to introduce the space $\ell^2(\mathbb{C}^K)$.

Construction 3.1. We construct the Hilbert space $\ell^2(\mathbb{C}^K) := \ell^2(\mathbb{N}) \otimes \mathbb{C}^K$, as described in Construction B.4 (see Example B.3 for the definition of $\ell^2(\mathbb{N})$). The orthonormal bases of $\ell^2(\mathbb{N})$ and $\mathbb{C}^K$ are denoted by respectively $\{e_n\}_{n \in \mathbb{N}}$ and $\{\tilde{e}_n\}_{0 \leq n \leq K-1}$. An element $V \in \ell^2(\mathbb{C}^K)$ is of the form

$$V = \sum_{n=0}^{\infty} v_n \otimes e_n \text{ or alternatively } V = \begin{pmatrix} v_0 \\ v_1 \\ \vdots \end{pmatrix} \quad (3.1)$$

where the components $v_n \in \mathbb{C}^K$ satisfy $\sum_{n=0}^{\infty} \|v_n\|^2 < \infty$. Indeed, for $V$ as in (3.1) and $W = \sum_{n=0}^{\infty} e_n \otimes w_n$, the inner product in $\ell^2(\mathbb{C}^K)$ is defined as

$$\langle V, W \rangle := \langle V, W \rangle_{\ell^2(\mathbb{C}^K)} = \sum_{n=0}^{\infty} \langle v_n, w_n \rangle$$

where the latter inner product is just the inner product in $\mathbb{C}^K$, as in (1.5). Thus the associated norm is

$$\|V\| := \|V\|_{\ell^2(\mathbb{C}^K)} = \langle V, V \rangle^{\frac{1}{2}} = \left( \sum_{n=0}^{\infty} \langle v, v \rangle \right)^{\frac{1}{2}} = \left( \sum_{n=0}^{\infty} \|v\|^2 \right)^{\frac{1}{2}},$$

so that $V \in \ell^2(\mathbb{C}^K) \iff \|V\| < \infty \iff \sum_{n=0}^{\infty} \|v_n\|^2 < \infty$. In particular, for $K = 1$ we retrieve the Hilbert space $\ell^2(\mathbb{N})$. ■
In the following we always assume that $V \in \ell^2(\mathbb{C}^K)$ is written in the form (3.1).

Let $(P_n)_{n \geq 0}$ be a sequence of orthonormal matrix polynomials with respect to $\mu$. Furthermore let $A_n, B_n \in \mathbb{C}^{K \times K}$ be the coefficients appearing in the three-term recurrence relation for $(P_n)_{n \geq 0}$, i.e. defined as in Theorem 2.33.

**Definition 3.2.** The Jacobi operator $J : \mathcal{D}(J) \to \ell^2(\mathbb{C}^K)$ is defined by

$$J V = e_0 \otimes (A_1 v_1 + B_0 v_0) + \sum_{n=1}^{\infty} e_n \otimes (A_{n+1} v_{n+1} + B_n v_n + A_n^* v_{n-1}),$$  \hspace{1cm} (3.2)

$$\mathcal{D}(J) = \left\{ V = \sum_{n=0}^{N} e_n \otimes v_n : N \in \mathbb{N} \right\} \subset \ell^2(\mathbb{C}^K).$$ \hspace{1cm} (3.3)

Thus $J$ is defined on the set of finite linear combinations of the basis vectors $\{e_n\}_{n \in \mathbb{N}}$. Furthermore let

$$(J V, W) = (V, J W)$$

for all $V, W \in \mathcal{D}(J)$.

**Proof.** Let $V, W \in \mathcal{D}(J)$. Then

$$\langle J V, W \rangle = \sum_{n=0}^{\infty} \langle (J V)_n, w_n \rangle = \langle (J V)_0, w_0 \rangle + \sum_{n=1}^{\infty} \langle (J V)_n, w_n \rangle$$

$$= \langle A_1 v_1 + B_0 v_0, w_0 \rangle + \sum_{n=1}^{\infty} \langle A_{n+1} v_{n+1} + B_n v_n + A_n^* v_{n-1}, w_n \rangle$$

$$= A_1 \langle v_1, w_0 \rangle + B_0 \langle v_0, w_0 \rangle + \sum_{n=1}^{\infty} \langle A_{n+1} v_{n+1} + B_n v_n + A_n^* v_{n-1}, w_n \rangle$$

$$= \sum_{n=0}^{\infty} A_{n+1} \langle v_{n+1}, w_n \rangle + \sum_{n=0}^{\infty} B_n \langle v_n, w_n \rangle + \sum_{n=1}^{\infty} A_n^* \langle v_{n-1}, w_n \rangle$$

$$= \sum_{n=1}^{\infty} A_n \langle v_n, w_{n-1} \rangle + \sum_{n=0}^{\infty} B_n \langle v_n, w_n \rangle + \sum_{n=0}^{\infty} A_{n+1}^* \langle v_n, w_{n+1} \rangle$$

$$= A_1^* \langle v_0, w_1 \rangle + B_0^* \langle v_0, w_0 \rangle + \sum_{n=1}^{\infty} \langle A_{n+1}^* v_n, w_{n+1} \rangle + B_n \langle v_n, w_n \rangle + A_n \langle v_n, w_{n-1} \rangle$$

$$= \langle v_0, A_1^* v_1 + B_0^* v_0 \rangle + \sum_{n=1}^{\infty} \langle v_n, A_{n+1}^* v_{n+1} + B_n v_n + A_n^* v_{n-1} \rangle$$

$$= \langle v_0, (J W)_0 \rangle + \sum_{n=1}^{\infty} \langle v_n, (J W)_n \rangle = \sum_{n=0}^{\infty} \langle v_n, (J W)_n \rangle = \langle V, J W \rangle.$$  \hspace{1cm} \square

Here we exploited the fact that all $B_n$ are hermitian. Note that all above sums are actually finite, as $V$ and $W$ are finite linear combinations of the $e_n$ ($n \in \mathbb{N}$).
Remark 3.4. Note that $J$ is a bounded operator whenever $\{\|A_n\|\}_{n \geq 1}$ and $\{\|B_n\|\}_{n \geq 0}$ are bounded sequences. In that case $J$ extends to a bounded self-adjoint operator on $\ell^2(\mathbb{C}^K)$. ■

The adjoint of the Jacobi operator will be computed in the following proposition (see Definition B.17 for the definition of the adjoint of an unbounded operator).

**Proposition 3.5.** Let $W = \sum_{n=0}^{\infty} e_n \otimes w_n$. The adjoint of $(J, D(J))$ is given by $(J^*, D(J^*))$ where

$$J^*W = e_0 \otimes (A_1 w_1 + B_0 w_0) + \sum_{n=1}^{\infty} e_n \otimes (A_{n+1} w_{n+1} + B_n w_n + A_n^* w_{n-1}),$$

(3.5)

$$D(J^*) = \left\{ W \in \ell^2(\mathbb{C}^K) : \|A_1 w_1 + B_0 w_0\|^2 + \sum_{n=1}^{\infty} \|A_{n+1} w_{n+1} + B_n w_n + A_n^* w_{n-1}\|^2 < \infty \right\}.$$  

(3.6)

Thus the adjoint of $J$ is the natural extension to its maximal domain.

**Proof.** Let $V \in D(J)$ and $W \in \ell^2(\mathbb{C}^K)$. Then

$$\langle JV, W \rangle = \langle A_1 v_1 + B_0 v_0, w_0 \rangle + \sum_{n=1}^{\infty} \langle A_{n+1} v_{n+1} + B_n v_n + A_n^* v_{n-1}, w_n \rangle$$

$$= \sum_{n=1}^{\infty} \langle v_{n+1}, A_{n+1}^* w_n \rangle + \langle v_n, B_n w_n \rangle + \langle v_{n-1}, A_n w_n \rangle + \langle v_1, A_1^* w_0 \rangle + \langle v_0, B_0 w_0 \rangle$$

$$= \sum_{n=1}^{\infty} \langle v_n, A_n^* w_{n-1} \rangle + \langle v_n, B_n w_n \rangle + \langle v_{n+1}, A_{n+1} w_n \rangle + \langle v_0, A_1 w_1 \rangle + \langle v_0, B_0 w_0 \rangle$$

$$= \sum_{n=1}^{\infty} \langle v_n, A_{n+1} w_{n+1} + B_n w_n + A_n^* w_{n-1} \rangle + \langle v_0, A_1 w_1 + B_0 w_0 \rangle,$$

(3.7)

where we have made use of the fact that all $B_n$ are hermitian. Note that all the above sums are actually finite, as $V \in D(J)$.

Write $D^*$ for the set defined in (3.6). We will show that $D(J^*) = D^*$ indeed holds, where $D(J^*)$ is the set

$$D(J^*) = \left\{ W \in \ell^2(\mathbb{C}^K) : V \mapsto \langle JV, W \rangle \text{ is continuous on } D(J) \right\}.$$  

Take $W \in D^*$. Then $\|A_1 w_1 + B_0 w_0\|^2 + \sum_{n=1}^{\infty} \|A_{n+1} w_{n+1} + B_n w_n + A_n^* w_{n-1}\|^2 \leq C$ for some constant $C$, so that

$$|\langle JV, W \rangle| \leq \|V\| \|J^*W\| \leq C \|V\|$$

for all $V \in D(J)$, by (3.7). We conclude that $D^* \subseteq D(J^*)$.

Conversely take $W \in D(J^*)$. Then $V \mapsto \langle JV, W \rangle$ is continuous on $D(J)$, so that there exists a constant $C$ for which

$$|\langle JV, W \rangle| \leq C \|V\|$$

(3.8)

for all $V \in D(J)$. Now we let

$$V = e_0 \otimes (A_1 w_1 + B_0 w_0) + \sum_{n=1}^{N} e_n \otimes (A_{n+1} w_{n+1} + B_n w_n + A_n^* w_{n-1})$$
for some $N \geq 0$, so that according to (3.7) and (3.8),
\[
\left( \left\| A_1 w_1 + B_0 w_0 \right\|^2 + \sum_{n=1}^{N} \left\| A_{n+1} w_{n+1} + B_n w_n + A_n^* w_{n-1} \right\|^2 \right)^{\frac{1}{2}} \leq C.
\]
This inequality holds for every $N$, and since $C$ does not depend on $N$, it holds also if we take the limit $N \to \infty$. From this it immediately follows that $W \in D^*$, so that $D(J^*) \subseteq D^*$.

It follows that $D(J^*)$ is indeed given by (3.6). The expression for the operator $J^*$ follows from (3.7), as $\langle JV, W \rangle = \langle V, J^* W \rangle$ for all $V \in D(J)$ and $W \in D(J^*)$.

Construction 3.6. Given the Jacobi operator $J$, let $(P_n)_{n \geq 0}$ be the matrix polynomials satisfying (2.29) where the coefficients $A_n$ and $B_n$ are those appearing in the definition of $J$. Assume that $J$ has a self-adjoint extension $J'$. Then, one can construct a measure $\mu$ with respect to which the polynomials $(P_n)_{n \geq 0}$ are orthonormal, as is shown in [3], Section 1.

Let $\{\tilde{e}_n\}_{0 \leq n \leq K-1}$ denote the standard orthonormal basis for $\mathbb{C}^K$. Then for every $n$ we define $\hat{e}_n = (\tilde{e}_n, 0, 0, \ldots) \in \ell^2(\mathbb{C}^K)$. Let $E$ denote the (unique) spectral resolution of the identity for the operator $J'$, which exists by Theorem B.26. In particular $E(B) : \ell^2(\mathbb{C}^K) \to \ell^2(\mathbb{C}^K)$ for all Borel sets $B \subseteq \mathbb{R}$. The matrix measure $\mu$ is then defined by
\[
\mu(B)_{ij} = \langle E(B)\hat{e}_i, \hat{e}_j \rangle \tag{3.9}
\]
for $0 \leq i, j \leq K - 1$.

This construction will be used in Section 3.3.2 to generate an explicit example.

Remark 3.7. From the above construction and Remark B.27 (with $f(t) = t^n$) it follows that the moment problem has a solution whenever $J$ has a self-adjoint extension $J'$. Indeed, one has
\[
\langle (J')^n \hat{e}_i, \hat{e}_j \rangle = \int_{\mathbb{R}} x^n d\mu_{ij}(x) = (S_n)_{ij}
\]
for the measure $\mu$ defined by 3.9.

3.2 The indices of deficiency

In the previous section we have computed the adjoint of the Jacobi operator, which was necessary in order to give the definition of the deficiency indices of $J$.

Definition 3.8. Let $z \in \mathbb{C}$ with $\text{Im}(z) > 0$. Then the deficiency index $n_+$ of the Jacobi operator $J$ is defined as
\[
n_+ = \dim \left( \ker \left( J^* - z \mathbb{1} \right) \right). \tag{3.10}
\]
Likewise, for $z \in \mathbb{C}$ with $\text{Im}(z) < 0$, we define
\[
n_- = \dim \left( \ker \left( J^* - z \mathbb{1} \right) \right). \tag{3.11}
\]

The indices of deficiency can be any natural number from 0 to $K$ (see also Theorem 3.21). It is well-known that $n_+$ and $n_-$ are independent of the choice of $z$ in the upper respectively lower half-plane; see for instance [13], Theorem XII.4.19. In particular, $n_+ = \dim \left( \ker \left( J^* - i \mathbb{1} \right) \right)$ and $n_- = \dim \left( \ker \left( J^* + i \mathbb{1} \right) \right)$. 

56
3.2.1 Relation between the deficiency indices and determinacy of the matrix moment problem

In general the deficiency indices do not have to be equal. Below some cases in which they are equal, are listed.

**Lemma 3.9.** The deficiency indices are equal if one of the following statements holds.

(i) $J^*$ commutes with complex conjugation in $l^2(\mathbb{C}^K)$.

(ii) $A_n, B_n \in \mathbb{R}^{K \times K}$ for all $n$.

(iii) There exists a unitary sequence $(U_n)_{n \geq 0}$ such that $U_{n-1}A_nU_n^*, U_nB_nU_n^* \in \mathbb{R}^{K \times K}$ for all $n$.

**Proof.** The first implication is proven in Lemma B.23. Since (ii) implies (i), the second implication holds. Finally the third implication follows from the second by Remark 2.34. □

Krein has shown in [20] (see also [21], Theorem 3 and 4) that the following connection between the matrix moment problem and the deficiency indices of $J$ holds.

**Theorem 3.10.** The matrix moment problem is determinate if and only if one of the deficiency indices of the associated Jacobi operator is equal to 0.

For all other values of $n_+$ and $n_-$, the corresponding moment problem thus has some degree of indeterminacy. If $n_\pm = K$, the moment problem (or $\mu$) is said to be completely indeterminate. Note that if one of the deficiency indices is equal to $K$, then both of them equal $K$, as is shown in Lemma 3.16. In the not completely indeterminate cases (i.e. $1 \leq n_\pm \leq K - 1$) almost nothing is known.

**Notation 3.11.** A matrix moment problem with associated indices of deficiency $n_+$ and $n_-$, is sometimes referred to as $(n_+, n_-)$. In particular the completely indeterminate case is denoted as $(K, K)$.

Now the following result holds.

**Corollary 3.12.** If $J$ is self-adjoint, then the matrix moment problem is determinate.

**Proof.** Since $J$ is self-adjoint, we have $n_\pm = 0$ by Lemma B.24. Thus by Theorem 3.10, the matrix moment problem is determinate. □

**Remark 3.13.** Consider the case $K = 1$. Then the Jacobi operator is given by

$$J = \begin{pmatrix} b_0 & a_1 & & \\ a_1 & b_1 & a_2 & \\ & a_2 & b_2 & a_3 \\ & & \ddots & \ddots \ddots \end{pmatrix},$$

where $a_n > 0$ and $b_n \in \mathbb{R}$; see Theorem 1.21. It then follows from Lemma 3.9(ii) that the deficiency indices are equal. Hence the only possibilities for the pair $(n_+, n_-)$ are $(0, 0)$ and $(1, 1)$, corresponding with the determinate and indeterminate moment problem, respectively. □
3.2.2 The deficiency indices in terms of orthonormal matrix polynomials

In Definition 3.8 the deficiency indices were given in terms of the Jacobi operator $J$. Since $J$ is related to the orthonormal matrix polynomials $P_n$ via the coefficients of the three-term recurrence relation, it is evident that they can also be written in terms of these polynomials.

**Lemma 3.14.** Consider the equation $(J^* - zI)V = \theta$ where $V \in D(J^*)$ and $z$ is a complex number with $\text{Im}(z) > 0$. Then $V$ is of the form

$$V = \begin{pmatrix} P_0(z)v \\ P_1(z)v \\ P_2(z)v \\ \vdots \end{pmatrix} = \begin{pmatrix} v \\ P_1(z)v \\ P_2(z)v \\ \vdots \end{pmatrix}$$

(3.12)

for a certain $v \in \mathbb{C}^K$.

**Proof.** Assume that the infinite vector $V$ satisfies $(J^* - zI)V = \theta$, i.e. $J^*V = zV$. We write $V$ as

$$V = \begin{pmatrix} v_0 \\ v_1 \\ v_2 \\ \vdots \end{pmatrix}.$$  

(3.13)

Then

$$\begin{pmatrix} B_0v_0 + A_1v_1 \\ A_1^*v_0 + B_1v_1 + A_2v_2 \\ A_2^*v_1 + B_2v_2 + A_3v_3 \\ \vdots \end{pmatrix} = J^*V = zV = \begin{pmatrix} zv_0 \\ zv_1 \\ zv_2 \\ \vdots \end{pmatrix}.  

(3.14)

Comparing the respective entries in both vectors yields the equation

$$A_{n+1}v_{n+1} + B_nv_n + A_n^*v_{n-1} = zv_n$$

(3.15)

for every $n \geq 0$, with the convention $v_{-1} = 0$. As $v_n \in \mathbb{C}^K$ for every $n \geq 0$, we necessarily have $v_n = P_n(z)v$ for some $v \in \mathbb{C}^K$, by (2.29). Thus $V$ is of the form (3.12). \qed

Note that $V \in \ell^2(\mathbb{C}^K)$ if and only if

$$\sum_{k=0}^{\infty} \|P_k(z)v\|^2 < \infty.$$  

(3.16)

Hence $n_+$ is given by

$$n_+ = \dim(\ker(J^* - zI))$$

$$= \dim\{V \in D(J^*) : (J^* - zI)V = \theta\}$$

$$= \dim \left\{ v \in \mathbb{C}^K : \sum_{k=0}^{\infty} \|P_k(z)v\|^2 < \infty \right\}$$

$$= \dim \left\{ v \in \mathbb{C}^K : v^*R_N(z)^{-1}v \text{ is bounded for every } N \right\}$$  

(3.17)

where $\text{Im}(z) > 0$. Note that this is a consequence of (2.77). A similar equation holds for $n_-$ if we take $z \in \mathbb{C}$ with $\text{Im}(z) < 0$. Of course, we may specify (3.17) to $z = i$ (respectively $z = -i$ for $n_-$).
Remark 3.15. Now suppose that \( K = 1 \). Then

\[
n_+ = \dim \left\{ v \in \mathbb{C} : \sum_{k=0}^{\infty} |p_k(z)v|^2 < \infty \right\} = \dim \left\{ v \in \mathbb{C} : |v|^2 \sum_{k=0}^{\infty} |p_k(z)|^2 < \infty \right\}
\]

(3.18)

where \( \text{Im}(z) > 0 \). Then \( n_+ = 1 \) if and only if \( \sum_{k=0}^{\infty} |p_k(z)|^2 < \infty \). In Remark 3.13 it is already shown that \( n_+ = n_- \). Therefore the above result coincides with the statement of Remark 1.22, as the scalar moment problem is indeterminate whenever \((n_+, n_-) = (1, 1)\).

Lemma 3.16. If \( n_+ = K \) or \( n_- = K \), then \((n_+, n_-) = (K, K)\), in other words the moment problem is completely indeterminate.

Proof. Suppose without loss of generality that \( n_+ = K \), and fix \( z \in \mathbb{C} \) with \( \text{Im}(z) > 0 \). Then

\[
v^* \left( \sum_{k=0}^{\infty} P_k(z)^* P_k(z) \right) v = \sum_{k=0}^{\infty} \|P_k(z)v\|^2 < \infty
\]

for all \( v \in \mathbb{C}^K \), according to (3.17). But invoking Theorem 2.39 the same holds if we replace \( z \) by \( \overline{z} \), so that \( n_- = K \).

Notation 3.17. For further convenience we introduce the following notation:

\[
\mathcal{N}_z = \left\{ v \in \mathbb{C}^K : \sum_{k=0}^{\infty} \|P_k(z)v\|^2 < \infty \right\} = \left\{ v \in \mathbb{C}^K : v^* R_N(z)^{-1} v \text{ is bounded for every } N \right\};
\]

\[
\mathcal{N}_z = \ker(J^* - zI) = \left\{ \sum_{k=0}^{\infty} e_k \otimes P_k(z)v : \sum_{k=0}^{\infty} \|P_k(z)v\|^2 < \infty \right\}.
\]

Observe that \( \mathcal{N}_z \cong \mathcal{N}_z \). Indeed, introducing the notation

\[
V_z := \sum_{k=0}^{\infty} e_k \otimes P_k(z)v = \begin{pmatrix} v \\ P_1(z)v \\ P_2(z)v \\ \vdots \end{pmatrix} \in \mathcal{N}_z,
\]

(3.19)

we see that

\[
\varphi_z : \mathcal{N}_z \rightarrow \mathcal{N}_z, \ v \mapsto V_z
\]

is an isomorphism (with inverse \( \varphi_z^{-1} : \mathcal{N}_z \rightarrow \mathcal{N}_z, \ V_z \mapsto P_0(z)v = v \)).

Remark 3.18. Note that \( \mathcal{N}_z \) is a linear space. Indeed, for fixed \( z \) with \( \text{Im}(z) > 0 \), define \( \langle w, v \rangle_{z,N} := v^* R_N(z)^{-1} w = v^* K_N(z, z) w \) for \( N \geq 0 \) and \( v, w \in \mathbb{C}^K \). As \( K_N(z, z) \) is positive hermitian, \( \langle \cdot, \cdot \rangle_{z,N} \) is an inner product. Now observe that \( v \in \mathcal{N}_z \) if and only if \( \langle v, v \rangle_{z,N} < \infty \) for all \( N \geq 0 \). Let \( \lambda \in \mathbb{C} \), and \( v, w \in \mathcal{N}_z \). Then \( \langle \lambda v, \lambda w \rangle_{z,N} = |\lambda|^2 \langle v, w \rangle_{z,N} < \infty \) and furthermore \( \langle v+w, v+w \rangle_{z,N} = \langle v, v \rangle_{z,N} + \langle w, w \rangle_{z,N} < \infty \), since the Cauchy-Schwarz inequality implies that \( |\langle v, w \rangle| \leq \|v\| \|w\| = \langle v, v \rangle^{1/2} \langle w, w \rangle^{1/2} < \infty \).

Some sufficient conditions for determinacy of the matrix moment problem are summarized in the following proposition.

Proposition 3.19. The following statements are equivalent:

(i) \( n_+ = n_- = 0 \).
(ii) For all \( v \in \mathbb{C}^{K}, \sum_{k=0}^{\infty} \| P_k(z)v \|^2 = \infty \) for \( z \in \mathbb{C} \setminus \mathbb{R} \).

(iii) \( N_z = \{0\} \).

(iv) \( J \) has a unique self-adjoint extension, i.e. is essentially self-adjoint.

If one of these statements holds true, then the matrix moment problem is determinate.

All of these equivalencies, with the exception of (iv), are already shown to be true. Statement (iv) is due to [2], Section 79. Determinacy of the moment problem follows from these statements by invoking Theorem 3.10.

**Remark 3.20.** Note that \( N_z \), and also the deficiency indices \( n_{\pm} \) are independent of the choice of the orthonormal matrix polynomials. Indeed, define \( Q_n = U_n P_n \) for \( n \geq 0 \) with \( U_n \) unitary and \( (P_n)_{n \geq 0} \) a sequence of orthonormal matrix polynomials. Then \( (Q_n)_{n \geq 0} \) are also orthonormal, according to Remark 2.30. We have

\[
\sum_{k=0}^{N} v^* Q_k^*(z) Q_k(z)v = \sum_{k=0}^{N} v^* P_k^*(z) U_k^* U_k P_k(z)v = \sum_{k=0}^{N} v^* P_k^*(z) P_k(z)v
\]

from which the claim follows.

The following theorem relates the limit function \( R(z) \) defined in Remark 2.43 and the deficiency indices of \( J \).

**Theorem 3.21.** The rank of \( R(z) \) is constant in the half-planes \( \text{Im}(z) > 0 \) and \( \text{Im}(z) < 0 \), and it coincides with the index of deficiency of the operator \( J \). Put differently,

\[
\dim \text{Im}(R(z)) = \dim N_z.
\]

See [22], Theorem 3.1. In this proof it is shown that \( \text{Im}(R(z)) = N_z \).

**Remark 3.22.** In Chapter 2 and 3 we have stressed several times the connection between \( R(z) \) and the determinacy of the matrix moment problem. Therefore it seems likely that there exists an explicit correspondence between \( R(z) \) and the Hankel matrices \( H_N \). However, we didn’t succeed in finding such an explicit relation.

**Remark 3.23.** There exists a one-to-one map between \( N_{z_0} \) and \( N_{z_1} \) whenever \( \text{Im}(z_0), \text{Im}(z_1) > 0 \). Let \( J_1 \) be the restriction of \( J^* \) to \( D(J_1) = N_{z_1} \oplus D(J) \). We then define \( K(z_0, z_1) = (J_1 - z_0 1) A(z_1) \) with \( A(z_1) = (J_1 - z_1 1)^{-1} \) (note that \( A(z_1) \) is an everywhere defined bounded operator; see [13], Theorem XII.4.19). The restriction \( J_1 \) in the definition of \( K(z_0, z_1) \) is necessary to establish boundedness of \( A(z_1) \). It can be shown that\(^1\)

- \( K(z_0, z_0) = 1 \);
- \( K(z_0, z_1)K(z_1, z_2) = K(z_0, z_2) \).

\(^1\)Indeed,

\[
K(z_0, z_0) = (J_1 - z_0 1)(J_1 - z_0 1)^{-1} = 1;
\]

\[
K(z_0, z_1)K(z_1, z_2) = (J_1 - z_0 1)(J_1 - z_1 1)^{-1}(J_1 - z_1 1)(J_1 - z_2 1)^{-1} = (J_1 - z_0 1)(J_1 - z_1 1)^{-1} = K(z_0, z_2);
\]

\[
K(z_0, z_1) = (J_1 - z_0 1)(J_1 - z_1 1)^{-1} = ((J_1 - z_1 1) + (z_1 - z_0 1)(J_1 - z_1 1)^{-1} = 1 + (z_1 - z_0) A(z_1).
\]

60
Therefore, whenever $x \kappa$ to a map $K$ in other words $K$ is an everywhere defined bounded operator with bounded inverse.

We now claim that $\hat{K}$ is an operator with bounded inverse.

Note that $\hat{K}$ is continuous with respect to the Lebesgue measure. We assume

\[
\text{Remark 3.24. Note that the set } N_{z_0}, \text{ so that } \hat{K}(z_0, z_1) = 0, \text{ and independent of } \theta.
\]

\[\text{Example 3.25. Consider the matrix measure } \mu = W\tau_\mu \text{ where } W \text{ is given by}
\]

\[
W(x) = \frac{1}{\sum_{i=1}^{K} w_i(x)} M \begin{pmatrix}
  w_1(x) \\
  w_2(x) \\
  \vdots \\
  w_K(x)
\end{pmatrix} M^* (3.21)
\]

and $d\tau_\mu(x) = \sum_{i=1}^{K} w_i(x) dx$. As in Example 2.46, the weights $w_i$ are taken to be absolutely continuous with respect to the Lebesgue measure. We assume $M \in \mathbb{C}^{K \times K}$ to be regular and independent of $x$. Moreover we suppose that $w_1(x), \ldots, w_l(x)$ are determinate, and that

\[
\text{Remark 3.24. Note that the set } V \text{ defined in Remark 2.43 is actually } N_{z_0}, \text{ so it seems plausible that there exists an explicit connection between } R(z) \text{ and } \kappa(z_0, z_1). \]

3.3 Examples of the matrix moment problem

3.3.1 Examples with diagonalizable weights

\textbf{Example 3.25.} Consider the matrix measure $\mu = W\tau_\mu$ where $W$ is given by
$w_{l+1}(x), \ldots, w_K(x)$ are indeterminate.²

The $n$th moment of $\mu$ is given by

$$S_n = \int_{\mathbb{R}} x^n W(x) d\tau_{\mu}(x) = M \begin{pmatrix} s_1^n & s_2^n & \cdots & s_K^n \end{pmatrix} M^* \quad (3.22)$$

where

$$s_i^n = \int_{\mathbb{R}} x^n w_i(x) dx \quad (3.23)$$

We denote the orthonormal polynomials associated with $w_i(x)$ by $(p^i_n)_{n=0}^\infty$, i.e.

$$\langle p^i_n, p^j_m \rangle = \int_{\mathbb{R}} p^i_n(x) w_i(x)(p^j_m(x))^* dx = \delta_{n,m}. \quad (3.24)$$

Hence the orthonormal polynomials with respect to $\mu$ are given by

$$P_n(x) = \begin{pmatrix} p^1_n(x) \\ p^2_n(x) \\ \vdots \\ p^K_n(x) \end{pmatrix} M^{-1}. \quad (3.25)$$

Indeed, by [25], page 294, we have

$$\mu(E) = \int_E W(x) d\tau_{\mu}(x) = \int_E MD(x) M^* dx$$

where $D(x) = \text{diag}(w_1(x), \ldots, w_K(x))$, so that

$$\langle P_n, P_m \rangle = \int_{\mathbb{R}} P_n(x) W(x) P^*_m(x) d\tau_{\mu}(x)$$

$$= \int_{\mathbb{R}} P_n(x) M \begin{pmatrix} w_1(x) \\ w_2(x) \\ \vdots \\ w_K(x) \end{pmatrix} M^* P^*_m(x) dx$$

$$= \int_{\mathbb{R}} \begin{pmatrix} p^1_n(x) w_1(x) p^1_m(x)^* \\ p^2_n(x) w_2(x) p^2_m(x)^* \\ \vdots \\ p^K_n(x) w_K(x) p^K_m(x)^* \end{pmatrix} dx$$

$$= \begin{pmatrix} \delta_{n,m} & \delta_{n,m} & \cdots & \delta_{n,m} \\ \delta_{n,m} & \delta_{n,m} & \cdots & \delta_{n,m} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{n,m} & \delta_{n,m} & \cdots & \delta_{n,m} \end{pmatrix} = \delta_{n,m} \mathbb{I}_K. \quad (3.26)$$

²We say that the weight $w_i(x)$ is (in)determinate if the corresponding scalar measure $\mu_{ii}$, defined by

$$\mu_{ii}(E) = \int_E w_i(x) dx$$

for all $E \in \mathcal{E}$, is (in)determinate.
Using the notation $u := M^{-1}v \in \mathbb{C}^K$, we see that for any $v \in \mathbb{C}^K$,

$$
\sum_{n=0}^{\infty} \|P_n(z)v\|^2 = \sum_{n=0}^{\infty} \left\| \begin{pmatrix} p_n^1(z) \\ p_n^2(z) \\ \vdots \\ p_n^K(z) \end{pmatrix} \right\|^2 = \sum_{n=0}^{\infty} \left\| \begin{pmatrix} p_n^1(z)u_1 \\ p_n^2(z)u_2 \\ \vdots \\ p_n^K(z)u_K \end{pmatrix} \right\|^2
$$

$$
= \sum_{n=0}^{\infty} \left( \sum_{i=1}^{K} |p_n^i(z)u_i|^2 \right) = \sum_{i=1}^{l} \left( \sum_{n=0}^{\infty} |p_n^i(z)u_i|^2 \right) + \sum_{i=l+1}^{K} \left( \sum_{n=0}^{\infty} |p_n^i(z)u_i|^2 \right)
$$

$$
= \sum_{i=1}^{l} |u_i|^2 \left( \sum_{n=0}^{\infty} |p_n^i(z)|^2 \right) + \sum_{i=l+1}^{K} |u_i|^2 \left( \sum_{n=0}^{\infty} |p_n^i(z)|^2 \right)
$$

(3.27)

Invoking Remark 1.22 we note that $\sum_{n=0}^{\infty} |p_n^i(z)|^2$ diverges for $1 \leq i \leq l$, and converges for $l + 1 \leq i \leq K$. Therefore $\sum_{n=0}^{\infty} \|P_n(z)v\|^2 < \infty$ whenever $u_i = 0$ for $1 \leq i \leq l$. Thus $v \in \mathcal{N}_z$ if and only if

$$
v = M \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda_{l+1} \\ \vdots \\ \lambda_K \end{pmatrix}
$$

for certain $\lambda_i \in \mathbb{C}$, $l + 1 \leq i \leq K$. Put differently,

$$
\mathcal{N}_z = \bigoplus_{i=l+1}^{K} \mathbb{C}M e_i,
$$

(3.28)

which is independent of $z$. In particular it follows that $n_+ = n_- = K - l$.

This example also illustrates the general facts known about the completely (in)determinate case. Indeed, if all weights $w_i(x)$ are determinate (so that $l = K$), the deficiency indices are equal to $n_+ = n_- = 0$. On the other hand we have $l = 0$ if all the $w_i(x)$ are indeterminate, so that $n_+ = n_- = K$ in this case. $\blacksquare$

In the previous example we assumed $M$ to be independent of $x$. The next example illustrates the difficulties that may arise when $M$ does in fact depend on $x$.

**Example 3.26.** In this example we let $K = 2$. Consider the matrix measure $\mu = W\tau_{\mu}$ where $W$ is given by

$$
W(x) = \frac{1}{\sum_{i=1}^{K} w_i(x)} M \begin{pmatrix} w_1(x) & 0 \\ 0 & w_2(x) \end{pmatrix} M^* \text{ with } M = \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}
$$

(3.29)

and $d\tau_{\mu}(x) = \sum_{i=1}^{K} w_i(x)dx$. We assume $w_1(x)$ and $w_2(x)$ to be respectively determinate and indeterminate. Note that

$$
W(x) = \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} \begin{pmatrix} w_1(x) & 0 \\ 0 & w_2(x) \end{pmatrix} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} w_1(x) & xw_1(x) \\ xw_1(x) & x^2w_1(x) + w_2(x) \end{pmatrix}.
$$

(3.30)

According to Theorem 2.11(i), $xw_1(x)$ is also determinate.

The $n$th moment of $\mu$ is given by

$$
S_n = \int_{\mathbb{R}} x^n W(x) d\tau_{\mu}(x) = \int_{\mathbb{R}} x^n \begin{pmatrix} w_1(x) & xw_1(x) \\ xw_1(x) & w_2(x) \end{pmatrix} dx = \begin{pmatrix} s_n^1 & s_n^1 \\ s_{n+1}^1 & s_{n+1}^1 + s_n^2 \end{pmatrix},
$$

(3.31)
Our first guess for the orthonormal polynomials with respect to $\mu$ is similar to that of the previous example, namely
\[ P_n(x) = \begin{pmatrix} p_n^1(x) & 0 \\ 0 & p_{n-1}^2(x) \end{pmatrix} M^{-1} = \begin{pmatrix} p_n^1(x) & 0 \\ 0 & p_{n-1}^2(x) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -x & 1 \end{pmatrix} = \begin{pmatrix} p_n^1(x) & 0 \\ -xp_{n-1}^2(x) & p_{n-1}^2(x) \end{pmatrix}. \] (3.32)

The problem now is that the leading coefficient of the above polynomial is equal to
\[ \begin{pmatrix} l_n^1 \\ -l_{n-1}^2 \end{pmatrix} \]
(where we have used $l_n^i$ to denote the leading coefficient of $p_n^i$), and this is certainly not a regular matrix (and this is a requirement for $(P_n)_{n \geq 0}$ to be a simple sequence)!

Thus, in general it is not easy to find the orthonormal polynomials associated to some moment problem.

### 3.3.2 Example arising from a doubly infinite Jacobi operator

In this section an explicit example for $K = 2$ is worked out. All results stated without proof are taken from either [18], or [17] Section 5.3.

First we consider a so-called doubly infinite Jacobi operator, i.e. an operator $L$ on the Hilbert space $\ell^2(\mathbb{Z})$. Let $\{\tilde{e}_n\}_{n \in \mathbb{Z}}$ be the standard orthonormal basis of $\ell^2(\mathbb{Z})$. Then $L$ is defined by
\[ L\tilde{e}_n = a_n \tilde{e}_{n+1} + b_n \tilde{e}_n + a_{n-1} \tilde{e}_{n-1} \] (3.33)
with $a_n > 0$ and $b_n \in \mathbb{R}$. This operator is defined on the dense domain
\[ \mathcal{D}(L) = \left\{ v = \sum_{n=0}^{N} v_n \tilde{e}_n : N \in \mathbb{N} \right\} \subset \ell^2(\mathbb{Z}), \] (3.34)
i.e. the set of finite linear combinations of the basis vectors $\tilde{e}_n$. It might not seem clear why such an operator would give rise to a 2-dimensional matrix moment problem, but three-term recurrence operators on $\ell^2(\mathbb{Z})$ can in fact be related to a $2 \times 2$-matrix recurrence on $\mathbb{N}$. More explicitly, we identify $\ell^2(\mathbb{Z})$ with $\ell^2(\mathbb{C}^2) = \ell^2(\mathbb{N}) \otimes \mathbb{C}^2$ (see Construction B.4). The standard orthonormal basis of $\ell^2(\mathbb{N})$ is denoted by $\{e_n\}_{n \in \mathbb{N}}$. The aforementioned identification is then given by $U : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{C}^2)$,
\[ \tilde{e}_n \mapsto e_n \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \tilde{e}_{n-1} \mapsto e_n \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad n \in \mathbb{N}. \] (3.35)

Note that $UU^* = 1_{\ell^2(\mathbb{C}^2)}$ and $U^*U = 1_{\ell^2(\mathbb{Z})}$. Considering the diagram
\[ \ell^2(\mathbb{Z}) \supset \mathcal{D}(L) \xrightarrow{L} \ell^2(\mathbb{Z}) \]
\[ U \downarrow \quad \quad \downarrow U \]
\[ \ell^2(\mathbb{C}^2) \supset \mathcal{D}(J) \xrightarrow{J} \ell^2(\mathbb{C}^2) \]
we see that the corresponding operator $J$, acting on $\mathcal{D}(J) \subset \ell^2(\mathbb{C}^2)$, should satisfy $JU = UL$, in other words $J = ULU^*$. Here $\mathcal{D}(J)$ denotes the set of finite linear combinations of the basis vectors of $\ell^2(\mathbb{C}^2)$, i.e.
\[ \mathcal{D}(J) = \left\{ w = \sum_{n=0}^{N} e_n \otimes \begin{pmatrix} x_n \\ y_n \end{pmatrix} : N \in \mathbb{N} \right\} \subset \ell^2(\mathbb{C}^2) \] (3.36)
Lemma 3.27. The operator \( J : \ell^2(\mathbb{C}^2) \ni D(J) \rightarrow \ell^2(\mathbb{C}^2) \) is given by

\[
\sum_{n=0}^{\infty} e_n \otimes v_n \mapsto e_0 \otimes (A_1 v_1 + B_0 v_0) + \sum_{n=1}^{\infty} e_n \otimes (A_{n+1} v_{n+1} + B_n v_n + A_n^* v_{n-1}),
\]

where

\[
A_n = \begin{pmatrix} a_{n-1} & 0 \\ 0 & a_{-n-1} \end{pmatrix} \quad (n \in \mathbb{N}), \quad B_n = \begin{pmatrix} b_n & 0 \\ 0 & b_{-n-1} \end{pmatrix} \quad (n \geq 1), \quad B_0 = \begin{pmatrix} b_0 & a_{-1} \\ a_{-1} & b_{-1} \end{pmatrix}.
\]

Proof. Let \( v = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{C}^2 \). Then we compute

\[
J(e_n \otimes v) = U L U^* \begin{pmatrix} a \\ b \end{pmatrix} = U L \begin{pmatrix} a \tilde{e}_n + b \tilde{e}_{-n-1} \\ a a_{n+1} + b a_n e_0 + a a_{-1} e_{-1} + b b_{-n-1} \tilde{e}_{-1} + b b_{-n-2} \tilde{e}_{-2} \end{pmatrix}.
\]

Note that no basis vectors inside the parentheses coincide if \( n - 1 > -n \), i.e. if \( n > \frac{1}{2} \). Therefore we distinguish between two cases, namely \( n \geq 1 \) and \( n = 0 \).

For \( n \geq 1 \) we obtain

\[
J(e_n \otimes v) = e_{n+1} \begin{pmatrix} a a_n \\ b a_{n-2} \end{pmatrix} + e_n \begin{pmatrix} a b_n \\ b b_{n-1} \end{pmatrix} + e_{n-1} \begin{pmatrix} a a_{n-1} \\ b a_{n-1} \end{pmatrix} + e_n \begin{pmatrix} b_n \\ 0 \end{pmatrix} + e_{n-1} \begin{pmatrix} 0 \\ a_{n-1} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix},
\]

while in case \( n = 0 \) it holds that

\[
J(e_n \otimes v) = e_1 \begin{pmatrix} a a_0 \\ b a_2 \end{pmatrix} + e_0 \begin{pmatrix} a b_0 + b a_1 \\ a a_1 + b b_1 \end{pmatrix} + e_0 \begin{pmatrix} b_0 \\ a_{-1} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}.
\]

Therefore it follows that

\[
J \left( \sum_{n=0}^{\infty} e_n \otimes v_n \right) = \sum_{n=0}^{\infty} J(e_n \otimes v_n) = J(e_0 \otimes v_0) + \sum_{n=1}^{\infty} J(e_n \otimes v_n)
\]

which coincides with (3.37). \( \Box \)
Example 3.28. In [18] the operator

$$2L\bar{e}_n = a_n\bar{e}_{n+1} + a_{n-1}\bar{e}_{n-1}$$  \hspace{1cm} (3.38)

on $\mathcal{D}(L) \subset \ell^2(\mathbb{Z})$ (see (3.34)) is considered, where $a_n > 0$. This is a special case of (3.33), namely with $b_n = 0$ for all $n \in \mathbb{Z}$ (and a factor 2 for convenience). This operator is related to the so-called hypergeometric difference equation, namely

$$2zf_n(z) = \frac{1 + a^2tq^{n-1}}{atq^{n-1}}f_{n+1}(z) - \frac{1 - q^{-n-1}t}{atq^{n-1}}f_{n-1}(z) \quad (n \in \mathbb{Z}),$$  \hspace{1cm} (3.39)

where it is assumed that $0 < q < 1$, $a \neq 0$ and $t \neq 0$. Moreover we assume that the coefficients do not vanish, i.e. $t \notin q\mathbb{Z}$ and $-ta^2 \notin q\mathbb{Z}$. In [18] three branches of solutions of (3.39) are given, namely

$$u_n(z) = 2\varphi_1\left(\frac{ay, a/y}{-q}; q, tq^n\right)$$

$$v_n(z) = (-1)^n2\varphi_1\left(-ay, -a/y; -q; q, tq^n\right)$$

$$F_n(y) = (ay)^{-n}2\varphi_1\left(\frac{a, -ay}{qy^2}; q, -\frac{q^{2-n}}{a^2t}\right)$$  \hspace{1cm} (3.40)

where $z = \frac{1}{2}(y + y^{-1})$ and $y^2 \notin q^{-N}$. Moreover $F_n(y^{-1})$ is a solution. Here $2\varphi_1$ denotes the series

$$2\varphi_1\left(\frac{a, b}{c}; q, z\right) = \sum_{k=0}^{\infty} \frac{(a, b; q)_k}{(c, q; q)_k}z^k$$  \hspace{1cm} (3.41)

where $(a, b; q)_k = (a; q)_k(b; q)_k$ and $(a; q)_k = \prod_{r=0}^{k-1}(1 - aq^r)$. Note that this last expression also holds for $k = \infty$, as the product is convergent.

The solutions given in (3.40) are related by

$$u_n(z) = c(y; a, t)F_n(y) + c(y^{-1}; a, t)F_n(y^{-1})$$

$$v_n(z) = c(y; -a, t)F_n(y) + c(y^{-1}; -a, t)F_n(y^{-1})$$

$$c(y; a, t) = \frac{(a/y, -q/ay, at, q/aty; q)_\infty}{(-q, y^{-2}, t, q/t; q)_\infty}.$$  \hspace{1cm} (3.42)

Here the notation $(a_1, \ldots, a_r; q)_k = \prod_{i=1}^k(a_i; q)_k$ is used.

For the purpose of [18] it is necessary to rewrite (3.39) in a symmetric form; this is also the reason that the operator under consideration is symmetric (see (3.38)). It can be shown that this can be achieved whenever $a = \sqrt{q}e^{i\psi}$ and $t = ire^{-i\psi}$ for certain $r \in \mathbb{R} \setminus \{0\}$. The values of the coefficients $a_n$ are fixed as

$$a_n = |r|^{-n}q^{-n}\sqrt{1 - 2rq^n \sin \psi + r^2q^{2n}} = \left|\frac{1 + ire^{i\psi}q^n}{irq^n}\right|. \hspace{1cm} (3.43)$$

Lastly it is assumed that $t \notin \mathbb{R}_{>0}$.

Note that the operator $(L, \mathcal{D}(L))$ is a symmetric operator that commutes with conjugation (since $a_n \in \mathbb{R}$). It follows that the deficiency indices are equal (see Lemma B.23). Since the solution space of $Lv = zv$ is two-dimensional, the only possibilities for the deficiency indices are
(0, 0), (1, 1) and (2, 2). The adjoint operator \((L^*, \mathcal{D}(L^*))\) is given by

\[
2L^* \left( \sum_{n \in \mathbb{Z}} v_n \tilde{e}_n \right) = \sum_{n \in \mathbb{Z}} (a_n v_{n+1} + a_{n-1} v_{n-1}) e_n \quad (3.44)
\]

and \(\mathcal{D}(L^*) = \{ v \in l^2(\mathbb{Z}) : L^* v \in l^2(\mathbb{Z}) \}\). This can be shown similarly as Proposition 3.6.

We will consider the case in which the deficiency indices are (1, 1). Observe that the coefficients \(a_n\) are bounded for \(n \to -\infty\), due to the fact that \(0 < q < 1\). From this last fact it follows that the space

\[
S^-(z) = \left\{ v = \sum_{n \in \mathbb{Z}} v_n \tilde{e}_n : L^* v = z v, \sum_{n=-\infty}^0 |v_n|^2 < \infty \right\}
\]

(3.45)
is one-dimensional whenever \(z \in \mathbb{C} \setminus \mathbb{R}\). We assume \(S^-(z)\) to be spanned by \(\Psi(z) = \sum_{n \in \mathbb{Z}} \Psi_n(z) \tilde{e}_n\) satisfying \(\Psi_n(z) = \overline{\Psi_n(\overline{z})}\). This condition can be imposed since \(L\) commutes with complex conjugation.

Likewise we define the space

\[
S^+(z) = \left\{ v = \sum_{n \in \mathbb{Z}} v_n \tilde{e}_n : L^* v = z v, \sum_{n=0}^\infty |v_n|^2 < \infty \right\}
\]

(3.46)
which is at least one-dimensional for \(z \in \mathbb{C} \setminus \mathbb{R}\), and at most two-dimensional. It can be shown that \(S^+(z)\) is two-dimensional. From this it follows that

\[
\dim \ker(L^* + i) = \dim \left( S^+(\mp i) \cap S^-(\mp i) \right) = 1
\]

(3.47)
so that the deficiency indices indeed are equal to (1, 1), as stated above. Observe that \(\Psi(\pm i) \in \ker(L^* \mp i)\).

Now the self-adjoint extensions of \((L, \mathcal{D}(L))\) are given by \((L^*, \mathcal{D}_\theta)\) where

\[
\mathcal{D}_\theta = \left\{ v \in \mathcal{D}(L^*) : \lim_{N \to \infty} \left[ v, e^{-i\theta} \Psi(i) + e^{i\theta} \Psi(-i) \right]_N = 0 \right\}
\]

(3.48)
for \(\theta \in [0, 2\pi)\). Here \([, ,]_n\) denotes the Wronskian or Casorati determinant defined by

\[
[u, v]_n = \frac{1}{2} a_n (u_{n+1} v_n - u_n v_{n+1})
\]

(3.49)
for \(u = \sum_{n \in \mathbb{Z}} u_n \tilde{e}_n\) and \(v = \sum_{n \in \mathbb{Z}} v_n \tilde{e}_n\).

For \(\overline{\Psi(z)} \in S^+(z) \cap \mathcal{D}_\theta\), we define the Green kernel

\[
G_{k,l}(z) = \frac{1}{[\Psi(z), \overline{\Psi(z)}]} \begin{cases} 
    \Psi_k(z) \overline{\Psi_l(z)} & \text{if } k \leq l, \\
    \Psi_l(z) \overline{\Psi_k(z)} & \text{if } l \leq k.
\end{cases}
\]

(3.50)
The resolvent \(G(z) = (z - L)^{-1}\) is then given by

\[
G(z) v = \sum_{k=-\infty}^\infty (G(z) v)_k \tilde{e}_k \quad \text{where} \quad (G(z) v)_k = \sum_{l=-\infty}^\infty v_l G_{k,l}(z).
\]

(3.51)
The corresponding spectral measure $E$ of the self-adjoint operator $(L, \mathcal{D}_B)$ can be obtained from the resolvent by

$$E_{\xi,\eta}(x_1, x_2) = \lim_{\delta \downarrow 0} \lim_{\varepsilon \downarrow 0} \frac{1}{2\pi i} \int_{x_1 + \delta}^{x_2 - \delta} (G(x - i\varepsilon)\xi, \eta) - (G(x + i\varepsilon)\xi, \eta) \, dx$$

for $\xi, \eta \in \ell^2(\mathbb{Z})$. This formula is called the Stieltjes-Perron inversion formula (see also Theorem B.14).

Now we are finally able to state the spectral measure of the self-adjoint extensions of $(L, \mathcal{D})$ as explicitly as possible. We introduce the notation $\xi, \eta$ for $x \in \ell^2(\mathbb{Z})$. This formula is called the Stieltjes-Perron inversion formula (see also Theorem B.14).

The following two results are taken directly from [18], Proposition 5.5 and Proposition 5.6. Lemma 5.4, but these aren’t of particular interest for the example we are considering here. The constants $\alpha_n$ appear in the expression $g_n(z) = \alpha_n f_n(z)$, where $(f_n(z))_{n \in \mathbb{Z}}$ satisfies (3.39) and $(g_n(z))_{n \in \mathbb{Z}}$ satisfies the symmetric form derived from (3.39), i.e.

$$2zg_k(z) = a_k g_{k+1}(z) + a_{k-1} g_{k-1}(z).$$

Thus the coefficients $a_n$ and $\alpha_n$ are related to each other; see [18], Section 3. The $\alpha_n$ are explicitly given by

$$\alpha_n = e^{i\phi_n} q^{\frac{1}{2}n}, \quad \phi_{n+1} - \phi_n \equiv \arg \left(1 + i e^{i\psi} q^n\right) - \frac{1}{2} \pi \text{ sgn}(r) \mod 2\pi. \quad (3.53)$$

In fact, one can show that $\mathcal{T} = B$. Explicit expressions for $\psi_n(z)$ and $A$ can be found in [18], Lemma 5.4, but these aren’t of particular interest for the example we are considering here. The following two results are taken directly from [18], Proposition 5.5 and Proposition 5.6.

**Proposition 3.29.** $[-1, 1]$ is contained in the continuous spectrum of $(L, \mathcal{D}_B)$ and for $0 \leq \chi_1 < \chi_2 \leq \pi$ the spectral measure is determined by

$$\langle E \left([\cos \chi_2, \cos \chi_1]\right) \xi, \eta \rangle = \frac{1}{2\pi} \int_{\chi_1}^{\chi_2} \frac{\langle \xi, \psi(\cos \chi) \rangle \langle \eta, \psi(\cos \chi) \rangle}{|A c(e^{i\chi}; a, t) + A c(e^{i\chi}; -a, t)|^2} \, d\chi. \quad (3.54)$$

Here $c(y; a, t)$ is defined as in (3.42).

**Proposition 3.30.** There is no continuous spectrum of $(L^*, \mathcal{D}_B)$ in $(-\infty, -1) \cup (1, \infty)$. The point spectrum of $(L^*, \mathcal{D}_B)$ occurs at the set

$$S = \left\{ x_0 = \frac{1}{2} (y_0 + y_0^{-1}) : |y_0| > 1, A c(y_0; -a, t) + A c(y_0; a, t) = 0 \right\}. \quad (3.55)$$

and the spectral projection is determined by

$$\langle E(\{x_0\}) \xi, \eta \rangle = \text{Res}_{y=y_0} \frac{\langle \xi, \psi(x_0) \rangle \langle \eta, \psi(x_0) \rangle}{y \left(A c(y^{-1}; a, t) + A c(y^{-1}; -a, t)\right) \left(A c(y; -a, t) + A c(y; a, t)\right)}.$$

Combining the above two propositions gives the following result.
Theorem 3.31. The spectral decomposition of the self-adjoint extension \((L^*, D_\theta)\) is given by
\[
\langle L^* \xi, \eta \rangle = \frac{1}{2\pi} \int_0^\pi \frac{\cos \chi \langle \xi, \psi(\cos \chi) \rangle \langle \eta, \psi(\cos \chi) \rangle}{|\operatorname{Ac}(e^{i\chi}; a, t) + \overline{\operatorname{Ac}}(e^{i\chi}; -a, t)|^2} \, d\chi \\
+ \sum_{x_0 \in S} \operatorname{Res}_{y=y_0} \frac{x_0(\xi, \psi(x_0)) \langle \eta, \psi(x_0) \rangle}{y (\operatorname{Ac}(y^{-1}; a, t) + \overline{\operatorname{Ac}}(y^{-1}; -a, t)) (\overline{\operatorname{Ac}}(y; -a, t) + \operatorname{Ac}(y; a, t))}
\]  
\[(3.57)\]
for \(\xi \in D_\theta\) and \(\eta \in \ell^2(\mathbb{Z})\).

Proof. According to the Spectral Theorem,
\[
\langle L^* \xi, \eta \rangle = \int \theta \, dE_{\xi,\eta}(t)
\]
for \(\xi \in D(L^*)\) and \(\eta \in \ell^2(\mathbb{Z})\). As the continuous spectrum equals \([-1, 1]\), while \(S\) constitutes the point spectrum, we have
\[
\langle L^* \xi, \eta \rangle = \int_{-1}^1 t \, d\langle E([-1, 1]) \xi, \eta \rangle + \sum_{x_0 \in S} x_0 \langle E(\{x_0\}) \xi, \eta \rangle
\]
\[
= \frac{1}{2\pi} \int_0^\pi \frac{\cos \chi \langle \xi, \psi(\cos \chi) \rangle \langle \eta, \psi(\cos \chi) \rangle}{|\operatorname{Ac}(e^{i\chi}; a, t) + \overline{\operatorname{Ac}}(e^{i\chi}; -a, t)|^2} \, d\chi + \\
\sum_{x_0 \in S} x_0 \operatorname{Res}_{y=y_0} \frac{\langle \xi, \psi(x_0) \rangle \langle \eta, \psi(x_0) \rangle}{y (\operatorname{Ac}(y^{-1}; a, t) + \overline{\operatorname{Ac}}(y^{-1}; -a, t)) (\overline{\operatorname{Ac}}(y; -a, t) + \operatorname{Ac}(y; a, t))}.
\]

With the explicit expressions for the spectral measure, we are able to compute the corresponding measure by using Construction 3.6. Let \(B\) be a Borel set of \(\mathbb{R}\) in \([-1, 1]\), so that \(B = (\cos \chi_2, \cos \chi_1)\) for certain \(0 \leq \chi_1 < \chi_2 \leq \pi\). It then follows from (3.54) that
\[
\langle E(B) \xi, \eta \rangle = \frac{1}{2\pi} \int_{\chi_1}^{\chi_2} \frac{\langle \xi, \psi(\cos \chi) \rangle \langle \eta, \psi(\cos \chi) \rangle}{|\operatorname{Ac}(e^{i\chi}; a, t) + \overline{\operatorname{Ac}}(e^{i\chi}; -a, t)|^2} \, d\chi
\]  
\[(3.58)\]
for \(\xi, \eta \in \ell^2(\mathbb{Z})\). In the following, the operator on \(\ell^2(\mathbb{C}^2)\) corresponding to \(E(B)\) is denoted by \(\mathcal{E}(B)\). Using the identification \(U\), defined in (3.35), we obtain the diagram
\[
\begin{array}{ccc}
\ell^2(\mathbb{Z}) & \xrightarrow{E(B)} & \ell^2(\mathbb{Z}) \\
U & \downarrow U & \\
\ell^2(\mathbb{C}^2) & \xrightarrow{\mathcal{E}(B)} & \ell^2(\mathbb{C}^2)
\end{array}
\]
from which it follows that \(\mathcal{E}(B) = U E(B) U^*\). Hence
\[
\langle \mathcal{E}(B) \xi, \eta \rangle = \langle U E(B) U^* \xi, \eta \rangle = \langle E(B) U^* \xi, U^* \eta \rangle = \frac{1}{2\pi} \int_{\chi_1}^{\chi_2} \frac{\langle U^* \xi, \psi(\cos \chi) \rangle \langle U^* \eta, \psi(\cos \chi) \rangle}{|\operatorname{Ac}(e^{i\chi}; a, t) + \overline{\operatorname{Ac}}(e^{i\chi}; -a, t)|^2} \, d\chi
\]  
\[(3.59)\]
for \(\xi, \eta \in \ell^2(\mathbb{C}^2)\). Defining
\[
\hat{u}_0 = e_0 \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \hat{u}_1 = e_0 \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix},
\]
69
the measure $\mu(B)$ has components

$$
\mu(B)_{ij} = \langle \mathcal{E}(B)\hat{u}_i, \hat{u}_j \rangle;
$$

(3.60)

see (3.9). Observe that $U^*\hat{u}_0 = \hat{\pi}_0$ and $U^*\hat{u}_1 = \hat{\pi}_{-1}$. Therefore we obtain

$$
\mu(B)_{00} = \langle \mathcal{E}(B)\hat{u}_0, \hat{u}_0 \rangle = \frac{1}{2\pi} \int_{\chi_1}^{\chi_2} \langle U^*\hat{u}_0, \psi(\cos \chi) \rangle \overline{\langle U^*\hat{u}_0, \psi(\cos \chi) \rangle} d\chi
$$

$$
= \frac{1}{2\pi} \int_{\chi_1}^{\chi_2} \frac{\langle \bar{\pi}_0, \psi(\cos \chi) \rangle \langle \bar{\pi}_0, \psi(\cos \chi) \rangle}{|\mathcal{A}(e^{i\chi}; a, t) + \overline{\mathcal{A}(e^{i\chi}; -a, t)}|^2} d\chi
$$

$$
= \frac{1}{2\pi} \int_{\chi_1}^{\chi_2} \frac{\langle \bar{\pi}_0, \psi(\cos \chi) \rangle \langle \bar{\pi}_0, \psi(\cos \chi) \rangle}{|\mathcal{A}(e^{i\chi}; a, t) + \overline{\mathcal{A}(e^{i\chi}; -a, t)}|^2} d\chi.
$$

In a completely similar manner we obtain the expressions for the other components of $\mu(B)$, namely

$$
\mu(B)_{11} = \frac{1}{2\pi} \int_{\chi_1}^{\chi_2} \frac{\langle \bar{\pi}_1, \psi(\cos \chi) \rangle \langle \bar{\pi}_1, \psi(\cos \chi) \rangle}{|\mathcal{A}(e^{i\chi}; a, t) + \overline{\mathcal{A}(e^{i\chi}; -a, t)}|^2} d\chi;
$$

$$
\mu(B)_{01} = \frac{1}{2\pi} \int_{\chi_1}^{\chi_2} \frac{\langle \bar{\pi}_0, \psi(\cos \chi) \rangle \langle \bar{\pi}_1, \psi(\cos \chi) \rangle}{|\mathcal{A}(e^{i\chi}; a, t) + \overline{\mathcal{A}(e^{i\chi}; -a, t)}|^2} d\chi;
$$

$$
\mu(B)_{10} = \frac{1}{2\pi} \int_{\chi_1}^{\chi_2} \frac{\langle \bar{\pi}_1, \psi(\cos \chi) \rangle \langle \bar{\pi}_0, \psi(\cos \chi) \rangle}{|\mathcal{A}(e^{i\chi}; a, t) + \overline{\mathcal{A}(e^{i\chi}; -a, t)}|^2} d\chi.
$$

Combining the above results we can write down the following expression for $\mu(B)$:

$$
\mu(B) = \frac{1}{2\pi} \int_{\chi_1}^{\chi_2} \left( \begin{array}{cc}
|\psi_0(\cos \chi)|^2 & \psi_0(\cos \chi)\overline{\psi_1(\cos \chi)} \\
\psi_1(\cos \chi)\overline{\psi_0(\cos \chi)} & |\psi_1(\cos \chi)|^2
\end{array} \right) \frac{d\chi}{|\mathcal{A}(e^{i\chi}; a, t) + \overline{\mathcal{A}(e^{i\chi}; -a, t)}|^2}.
$$

(3.61)

Note that

\[
\det \left( \begin{array}{cc}
|\psi_0(\cos \chi)|^2 & \psi_0(\cos \chi)\overline{\psi_1(\cos \chi)} \\
\psi_1(\cos \chi)\overline{\psi_0(\cos \chi)} & |\psi_1(\cos \chi)|^2
\end{array} \right) = 0,
\]

so that $\det \mu(B) = 0$. The above matrix has eigenvalues equal to 0 and its trace, namely $|\psi_0(\cos \chi)|^2 + |\psi_1(\cos \chi)|^2 > 0$. The operator $J$ on $l^2(\mathbb{C}^2)$ corresponding to $L$ (i.e. that is obtained from $L$ via Lemma 3.27) has multiplicity 1.

Note that we only considered the case in which $B \subseteq [-1, 1]$. The same result however also holds for general Borel sets $\mathcal{B}$, since the point spectrum gives rise to similar expressions for the components of $\mu(B)$. More explicitly we have (note that $x_0 = \frac{1}{2} (y_0 + y_0^{-1})$ as in Proposition 3.30),

$$
\mu(B) = \frac{1}{2\pi} \int_{\chi_1}^{\chi_2} \left( \begin{array}{cc}
|\psi_0(\cos \chi)|^2 & \psi_0(\cos \chi)\overline{\psi_1(\cos \chi)} \\
\psi_1(\cos \chi)\overline{\psi_0(\cos \chi)} & |\psi_1(\cos \chi)|^2
\end{array} \right) \frac{d\chi}{|\mathcal{A}(e^{i\chi}; a, t) + \overline{\mathcal{A}(e^{i\chi}; -a, t)}|^2}
$$

$$
+ \sum_{x_0 \in S^* \cap \mathcal{B}} \left( \begin{array}{cc}
|\psi_0(x_0)|^2 & \psi_0(x_0)\overline{\psi_1(x_0)} \\
\psi_1(x_0)\overline{\psi_0(x_0)} & |\psi_1(x_0)|^2
\end{array} \right) \times
$$

$$
\frac{1}{\text{Res}_{y = y_0} \frac{(\mathcal{A}(y^{-1}; a, t) + \overline{\mathcal{A}(y^{-1}; -a, t)}) (\mathcal{A}(y; -a, t) + \mathcal{A}(y; a, t))}{\mathcal{A}(e^{i\chi}; a, t) + \overline{\mathcal{A}(e^{i\chi}; -a, t)}}. \tag{3.62}
$$

for an arbitrary Borel set $\mathcal{B}$, and from this expression it is clear that $\det \mu(B) = 0$. But then $\mu(B)$ is certainly not positive definite, which is in contradiction with the positivity of $\mu$ (this was assumed to hold true in Section 2.1).
Chapter 4

Criteria for (in)determinacy

4.1 Conditions for completely indeterminate moment problems

Here some criteria for completely indeterminate matrix moment problems are listed.

**Theorem 4.1.** A matrix measure $\mu$ is completely indeterminate if and only if for some $z_0 \in \mathbb{C} \setminus \mathbb{R}$ the series

$$\sum_{k=0}^{\infty} P_k(z)^* P_k(z) = \sum_{k=0}^{\infty} P_k^*(z) P_k(z)$$

converges. If this is the case, the series actually converges for all $z \in \mathbb{C}$, even uniformly on compact subsets of $\mathbb{C}$.

This result is taken from the proof of Corollary 3.1 in [8], and Section 2 and 3 from [22].

From Theorem 1.18 it follows that the scalar moment problem is indeterminate if and only if $\lim_{N \to \infty} \lambda_N > 0$. The following theorem is the analogue to this statement in the matrix-valued case. As in the scalar case, $\lambda_N$ denotes the smallest eigenvalue of the Hankel matrix $H_N$.

**Theorem 4.2.** The matrix moment problem associated with the moments (2.13) is completely indeterminate if and only if the sequence $(\lambda_N)_{N \geq 0}$ is bounded below, i.e. $\lambda_N \geq c$ for some constant $c > 0$.

The proof can be found in [8], Corollary 3.1.

4.2 Generalizations from Akhiezer

In this section we generalize some criteria given in [1] for (in)determinacy of the scalar Hamburger moment problem to the matrix-valued case.

**Theorem 4.3** (Carleman). If

$$\sum_{n=0}^{\infty} \frac{1}{\|A_{n+1}\|} = \infty,$$

then the associated matrix moment problem is determinate.

**Proof.** Let $z \in \mathbb{C} \setminus \mathbb{R}$. Then, since

$$\sum_{k=0}^{n} P_k^*(z) P_k(z) \succ P_0^*(z) P_0(z) = 1_K$$

(4.3)
it follows that
\[
\|v\|^2 = v^*v \leq v^* \left( \sum_{k=0}^{n} P^*_k(\bar{z}) P_k(z) \right) v = v^* K_n(\bar{z}, z)v
\]
\[
= \frac{1}{|z - \bar{z}|} \left| v^* P^*_n(\bar{z}) A_{n+1} P_{n+1}(z)v - v^* P^*_n(\bar{z}) A^*_n P_n(z)v \right|
\]
\[
\leq \frac{2}{|z - \bar{z}|} \|P_n(z)v\| \|A_{n+1}\| \|P_n(z)v\|.
\] (4.4)

Here we applied (2.75) with \( x = \bar{z} \) and \( y = z \). Rewriting this inequality we obtain
\[
\frac{1}{\|A_{n+1}\|} \leq \frac{2}{|z - \bar{z}|} \frac{\|P_n(z)v\| \|P_{n+1}(z)v\|}{\|v\|^2}.
\] (4.5)

It follows by the Cauchy-Schwarz inequality that
\[
\sum_{n=0}^{\infty} \frac{1}{\|A_{n+1}\|} \leq \frac{2}{|z - \bar{z}|} \frac{\sum_{n=0}^{\infty} \|P_n(z)v\| \|P_{n+1}(z)v\|}{\|v\|^2}
\]
\[
\leq \frac{2}{|z - \bar{z}|} \frac{1}{\|v\|^2} \left( \sum_{n=0}^{\infty} \|P_n(z)v\|^2 \right)^{\frac{1}{2}} \left( \sum_{n=1}^{\infty} \|P_n(z)v\|^2 \right)^{\frac{1}{2}}
\]
\[
\leq \frac{2}{|z - \bar{z}|} \frac{1}{\|v\|^2} \sum_{n=0}^{\infty} \|P_n(z)v\|^2.
\] (4.6)

Now suppose that \( \sum_{n=0}^{\infty} \frac{1}{\|A_{n+1}\|} = \infty \). Then \( \sum_{n=0}^{\infty} \|P_n(z)v\|^2 = \infty \) for all \( v \in \mathbb{C}^K \setminus \{0\} \). Therefore \( \mathcal{N}_z = \{0\} \) for any \( z \in \mathbb{C} \setminus \mathbb{R} \), from which we conclude that \( n_+ = n_- = 0 \). Thus the associated moment problem is determinate. \( \square \)

The following theorem can be seen as a partial converse of Theorem 4.3.

**Theorem 4.4** (Berezanskii). If
\[
\|B_n\| \leq M < \infty \text{ for all } n \geq 1,
\] (4.7)
\[
\sum_{n=0}^{\infty} \frac{1}{\|A_{n+1}\|} < \infty
\] (4.8)

and if there exists \( N \in \mathbb{N} \) such that for all \( n \geq N \) we have
\[
\|A_{n+2}\| \|A_{n+1}^{-1}\| \|A_{n}\|^\frac{1}{2} \leq 1,
\] (4.9)

then the associated matrix moment problem is completely indeterminate.

**Proof.** Rewriting the three-term recurrence relation (2.29) we obtain
\[
P_{n+1}(x) = A_{n+1}^{-1}(xI_K - B_n)P_n(x) - A_{n+1}^{-1}A_n^*P_{n-1}(x).
\] (4.10)

Now let \( v \in \mathbb{C}^K \) and put \( x = i \). Then
\[
\|P_{n+1}(i)v\| \leq \|A_{n+1}^{-1}\| \|iI_K - B_n\| \|P_n(i)v\| + \|A_{n+1}^{-1}\|\|A_n^*\|\|P_{n-1}(i)v\|
\]
\[
\leq \|A_{n+1}^{-1}\| (1 + \|B_n\|) \|P_n(i)v\| + \|A_{n+1}^{-1}\| \|A_n\| \|P_{n-1}(i)v\|
\] (4.11)
Multiplying both sides with $\|A_{n+2}\|^{\frac{1}{2}}$ we obtain
\[
\|A_{n+2}\|^{\frac{1}{2}}\|P_{n+1}(i)v\| \leq \frac{\|A_{n+2}\|^{\frac{1}{2}}\|A_{n+1}^{-1}\|}{\|A_{n+1}\|^{\frac{1}{2}}} (1 + \|B_n\|) \|A_{n+1}\|^{\frac{1}{2}}\|P_n(i)v\|
+ \|A_{n+2}\|^{\frac{1}{2}}\|A_{n+1}^{-1}\| \|A_n\|^{\frac{1}{2}}\|A_{n-1}(i)v\|.
\] (4.12)

Invoking (4.7) and (4.9) it follows that for all $n \geq N$,
\[
\|A_{n+2}\|^{\frac{1}{2}}\|P_{n+1}(i)v\| \leq \frac{1 + M}{\|A_n\|^{\frac{1}{2}}\|A_{n+1}\|^{\frac{1}{2}}} (\|A_{n+1}\|^{\frac{1}{2}}\|P_n(i)v\| + \|A_n\|^{\frac{1}{2}}\|P_{n-1}(i)v\|).
\] (4.13)

If we define $\alpha_n = \max_{N \leq k \leq n} \|A_{k+1}\|^{\frac{1}{2}}\|P_k(i)v\|$ for $n \geq N$, we get
\[
\alpha_{n+1} \leq \alpha_n \left( 1 + \frac{1 + M}{\|A_n\|^{\frac{1}{2}}\|A_{n+1}\|^{\frac{1}{2}}} \right)
\] (4.14)

for all $n \geq N$. Hence for these values of $n$ it holds that
\[
\|A_{n+1}\|^{\frac{1}{2}}\|P_n(i)v\| \leq \alpha_n \leq \alpha_N \prod_{k=N}^{\infty} \left( 1 + \frac{1 + M}{\|A_k\|^{\frac{1}{2}}\|A_{k+1}\|^{\frac{1}{2}}} \right) =: C < \infty.
\] (4.15)

The fact that $C < \infty$ is a consequence of (4.8) and (4.9). Indeed, by the ratio test and the convergence of $\sum_{n=0}^{\infty} \frac{1}{\|A_{n+1}\|}$ we know that
\[
\lim_{n \to \infty} \frac{\|A_{n+1}\|}{\|A_{n+2}\|} = \lim_{n \to \infty} \frac{\frac{1}{\|A_{n+2}\|}}{\frac{1}{\|A_{n+1}\|}} < 1.
\]

But then from (4.9) we obtain
\[
\lim_{n \to \infty} \frac{\|A_{n+1}\|^{\frac{1}{2}}\|A_{n+2}\|^{\frac{1}{2}}}{\|A_n\|^{\frac{1}{2}}\|A_{n+1}\|^{\frac{1}{2}}} = \lim_{n \to \infty} \frac{\|A_n\|^{\frac{1}{2}} \|A_{n+1}\|^{\frac{1}{2}}}{\|A_{n+2}\|} = \lim_{n \to \infty} \frac{\|A_n\|^{\frac{1}{2}} \|A_{n+1}\|^{\frac{1}{2}}}{\|A_{n+2}\|}
\leq \lim_{n \to \infty} \frac{\|A_{n+1}\|^{-1}}{\|A_{n+2}\|} \leq \lim_{n \to \infty} \frac{\|A_{n+1}\|}{\|A_{n+2}\|} < 1,
\]

since $\|A^{-1}\| \geq \|A\|^{-1}$ for any $A \in \mathbb{C}^{K \times K}$. Thus
\[
\sum_{n=N}^{\infty} \gamma_n := \sum_{n=N}^{\infty} \frac{1 + M}{\|A_n\|^{\frac{1}{2}}\|A_{n+1}\|^{\frac{1}{2}}}
\]
converges by the ratio test. Observe that $1 + \gamma_n \leq e^{\gamma_n}$ since $\gamma_n > 0$, so that
\[
\prod_{n=N}^{\infty} (1 + \gamma_n) \leq \prod_{n=N}^{\infty} e^{\gamma_n} = \exp \left( \sum_{n=N}^{\infty} \gamma_n \right) < \infty.
\]

Thus $C\infty$. We then conclude from (4.8) that
\[
\sum_{n=0}^{\infty} \|P_n(i)v\|^2 = \sum_{n=0}^{N-1} \|P_n(i)v\|^2 + \sum_{n=N}^{\infty} \|P_n(i)v\|^2 \leq \sum_{n=0}^{N-1} \|P_n(i)v\|^2 + \sum_{n=N}^{\infty} \|C^2\| \|A_{n+1}\| < \infty.
\] (4.16)
Since $\sum_{n=0}^{\infty} \|P_n(x)\|^2 < \infty$ for all $x \in \mathbb{C}^K$, we conclude that $N_1 = \mathbb{C}^K$, and thus $n_+ = K$. In a completely similar way (by putting $x = -i$ in the above computations) it follows that $n_- = K$. Thus the associated moment problem is completely indeterminate. \hfill \square

**Theorem 4.5** (Dennis/Wall). If

$$\sum_{n=1}^{\infty} |v^* A_{n+1}^{-1} B_n A_n^{-1} v| = \infty \quad (4.17)$$

and

$$\sum_{n=1}^{\infty} \| (A_{n+1}^*)^{-1} v \| \| A_n^{-1} v \| < \infty \quad (4.18)$$

for all $v \in \mathbb{C}^K \setminus \{0\}$, and moreover $P_n^* = P_n$ for any $n \in \mathbb{N}$, then the associated matrix moment problem is determinate.

Note that for $K = 1$, the condition $P_n^* = P_n$ is always satisfied, as the scalar orthonormal polynomials are real-valued.

**Proof.** Recall from (2.43) that $Q_k(x) P_k^*(x) = P_k(x) Q_k^*(x)$ for all $n \geq 0$, and from (2.44) that $Q_n(x) P_{n-1}^*(x) - P_n(x) Q_{n-1}^*(x) = A_n^{-1}$ for any $n \geq 1$. Hence it follows from (2.29) that

$$Q_{n+1}(x) P_{n+1}^*(x) - P_{n+1}(x) Q_{n+1}^*(x) = A_{n+1}^{-1} (x Q_n(x) - B_n Q_n(x) - A_n^* Q_{n-1}(x)) P_{n-1}^*(x)$$

$$- A_{n+1}^{-1} (x P_n(x) - B_n P_n(x) - A_n^* P_{n-1}(x)) Q_{n-1}^*(x)$$

$$= A_{n+1}^{-1} (x 1_K - B_n) (Q_n(x) P_{n-1}^*(x) - P_n(x) Q_{n-1}^*(x))$$

$$+ A_{n+1}^{-1} A_n^* (P_{n-1}(x) Q_{n-1}^*(x) - Q_{n-1}(x) P_{n-1}^*(x))$$

$$= A_{n+1}^{-1} (x 1_K - B_n) A_n^{-1} = x A_{n+1}^{-1} A_n^{-1} - A_{n+1}^{-1} B_n A_n^{-1}$$

for any $n \geq 1$. Now note that since $P_n^* = P_n$ for any $n \in \mathbb{N}$, the same holds for the polynomials $Q_n$, by (2.42). Thus

$$\sum_{n=1}^{\infty} |v^* A_{n+1}^{-1} B_n A_n^{-1} v| \leq |x| \sum_{n=1}^{\infty} |v^* A_{n+1}^{-1} A_n^{-1} v| + \sum_{n=1}^{\infty} |v^* Q_{n+1}(x) P_{n+1}^*(x) v| + \sum_{n=1}^{\infty} |v^* P_{n+1}(x) Q_{n+1}^*(x) v|$$

$$\leq |x| \sum_{n=1}^{\infty} \| (A_{n+1}^*)^{-1} v \| \| A_n^{-1} v \| + \sum_{n=1}^{\infty} \| Q_{n+1}(x) v \| \| P_{n+1}^*(x) v \|$$

$$+ \sum_{n=1}^{\infty} \| P_{n+1}(x) v \| \| Q_{n+1}^*(x) v \|$$

$$\leq |x| \sum_{n=1}^{\infty} \| (A_{n+1}^*)^{-1} v \| \| A_n^{-1} v \| + \left( \sum_{n=1}^{\infty} \| Q_{n+1}(x) v \|^2 \right)^{\frac{1}{2}} \left( \sum_{n=1}^{\infty} \| P_{n+1}(x) v \|^2 \right)^{\frac{1}{2}}$$

$$+ \left( \sum_{n=1}^{\infty} \| P_{n+1}(x) v \|^2 \right)^{\frac{1}{2}} \left( \sum_{n=1}^{\infty} \| Q_{n+1}^*(x) v \|^2 \right)^{\frac{1}{2}}$$

$$\leq |x| \sum_{n=1}^{\infty} \| (A_{n+1}^*)^{-1} v \| \| A_n^{-1} v \| + 2 \left( \sum_{n=1}^{\infty} \| P_n(x) v \|^2 \right)^{\frac{1}{2}} \left( \sum_{n=0}^{\infty} \| Q_n(x) v \|^2 \right)^{\frac{1}{2}}.$$
the former by Theorem 2.39, so that we may conclude that \( \sum_{n=0}^{\infty} \| P_n(x)v \|^2 = \infty \). It follows that \( n_+ = 0 \), which in turn implies determinacy of the associated matrix moment problem by Theorem 3.10. \( \square \)
Appendix A

Complex measures

For an extensive treatment of the theory of complex measures, we refer to [12].

A.1 Complex and positive measures

Let \((X, \mathcal{E})\) be a measurable space, i.e. a set \(X\) equipped with a \(\sigma\)-algebra \(\mathcal{E}\) of subsets of \(X\).

**Definition A.1.** A complex measure \(\mu\) on \((X, \mathcal{E})\) is a function \(\mu : \mathcal{E} \to \mathbb{C}\) such that

\[
\mu \left( \bigcup_{n=1}^{\infty} E_n \right) = \sum_{n=1}^{\infty} \mu(E_n) \tag{A.1}
\]

for any sequence \((E_n)_{n \geq 1}\) in \(\mathcal{E}\) of pairwise disjoint sets. A positive measure \(\mu : \mathcal{E} \to [0, \infty)\) is defined likewise.

Clearly \(\mu(\emptyset) = 0\), by letting \(E_n = \emptyset\) for every \(n \geq 1\).

**Definition A.2.** For a (complex) measure \(\mu\) on \((X, \mathcal{E})\), the support \(\text{supp}(\mu)\) is defined as the complement of the union of all open sets which are \(\mu\)-null sets, i.e. \(\text{supp}(\mu)\) is the smallest closed set \(C\) such that \(\mu(X \setminus C) = 0\).

**Notation A.3.** For positive measures \(\mu\) and \(\nu\), we say that \(\mu \leq \nu\) if \(\mu(E) \leq \nu(E)\) for all \(E \in \mathcal{E}\).

We now state a result involving the determinacy of positive measures (recall that such a measure is determinate if it is the only measure generating its moment sequence; see Section 1.1).

**Lemma A.4.** Let \(\mu\) and \(\nu\) be positive measures satisfying \(\mu \leq \nu\). If \(\nu\) is determinate, then so is \(\mu\).

**Proof.** Let \(\nu\) be determinate, but assume \(\mu\) to be indeterminate. Then there exists another positive measure with the same moments as \(\mu\), say \(\sigma\). As \(\mu \leq \nu\), we see that \(\sigma + (\nu - \mu)\) is a positive measure with the same moments as \(\nu\). Then, by determinacy of \(\nu\), we obtain \(\nu = \sigma + (\nu - \mu)\), from which it follows that \(\mu = \sigma\). This contradicts the indeterminacy of \(\mu\), so that it must in fact be determinate. \(\square\)
A.2 Radon-Nikodym Theorem

Definition A.5. Given a complex measure $\mu$, we define

$$|\mu|(E) = \sup \left\{ \sum_{n=1}^{\infty} |\mu(E_n)| : \bigcup_{n=1}^{\infty} E_n = E, (E_n)_{n=1}^{\infty} \text{ pairwise disjoint} \right\} \quad (A.2)$$

for every $E \in \mathcal{E}$. It is called the variation or absolute value of $\mu$. ■

The function $|\mu|$ turns out to be a positive measure with $|\mu(E)| \leq |\mu|(E)$ for any $E \in \mathcal{E}$ (see [5], Theorem 1.1). Moreover it can be shown that $|\mu|$ is the smallest positive measure $\sigma$ satisfying $|\mu(E)| \leq \sigma(E)$ for all $E \in \mathcal{E}$ (see [5], Theorem 1.5).

The following concepts and results are needed in order to prove (2.6).

Definition A.6. Let $\mu$ be a positive measure and let $\nu$ be an arbitrary measure on $(X, \mathcal{E})$, i.e. $\nu$ is either a complex or positive measure. Then $\nu$ is said to be absolute continuous with respect to $\mu$ if, for every $E \in \mathcal{E}, \mu(E) = 0$ implies that $\nu(E) = 0$. This is denoted as $\nu \ll \mu$. ■

Definition A.7. A positive measure $\mu$ is called $\sigma$-finite if there exists a partition $(E_n)_{n \geq 1}$ of $X$ (i.e. if there exist $E_n \in \mathcal{E}$ with $X = \bigcup_{n \geq 1} E_n$) such that $\mu(E_n) < \infty$ for $n \geq 1$. ■

Theorem A.8 (Radon-Nikodym). Let $\nu$ be a complex measure on $(X, \mathcal{E})$ and let $\mu$ be a $\sigma$-finite positive measure. If $\nu \ll \mu$, then there exists an $\mathcal{E}$-measurable function $f : X \to \mathbb{C}$ such that for all $E \in \mathcal{E}$,

$$\nu(E) = \int_{E} f \, d\mu. \quad (A.3)$$

The function $f$ is unique $\mu$-almost everywhere.

See [12], Theorem 4.2.3. The function $f$ is called the Radon-Nikodym derivative and denoted by $\frac{d\nu}{d\mu}$.

We conclude this appendix by stating a relation of a complex measure to its variation.

Proposition A.9. Let $(X, \mathcal{E})$ be a measurable space, $f$ an integrable function on $X$ and $\nu$ the complex measure defined by (A.3). Then

$$|\nu|(E) = \int_{E} |f| \, d\mu \quad (A.4)$$

holds for each $E \in \mathcal{E}$.

The proof can be found in [12], Proposition 4.2.4.
Appendix B

Spectral theory

In this appendix some facts from functional analysis are collected, which are needed in order to
discuss the operator approach to the moment problem in Chapter 3. It is largely based on [17].

B.1 Bounded operators

B.1.1 Hilbert spaces and operators

Definition B.1. A vector space $\mathcal{H}$ over $\mathbb{C}$ is called an inner product space if there exists a map
$\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ which satisfies

(i) $\langle av + bw, u \rangle = a \langle v, u \rangle + b \langle w, u \rangle$ for all $u, v, w \in \mathcal{H}$ and $a, b \in \mathbb{C}$;

(ii) $\langle v, u \rangle = \overline{\langle u, v \rangle}$ for all $u, v \in \mathcal{H}$;

(iii) $\langle v, v \rangle \geq 0$ for all $v \in \mathcal{H}$;

(iv) $\langle v, v \rangle = 0$ if and only if $v = 0$.

To the inner product $\langle \cdot, \cdot \rangle$ we associate the norm $\|v\| = \sqrt{\langle v, v \rangle}$ (sometimes the notation $\|v\|_\mathcal{H}$ is used to specify on which Hilbert space the norm is defined), and the topology arising from the corresponding metric $d(u, v) = \|u - v\|$. For this norm the so-called Cauchy-Schwarz inequality holds, i.e. $|\langle u, v \rangle| \leq \|u\| \|v\|$ for all $u, v \in \mathcal{H}$.

Definition B.2. A Hilbert space $\mathcal{H}$ is a complete inner product space, i.e. for every Cauchy
sequence $^1 (x_n)$ in $\mathcal{H}$ there exists $x \in \mathcal{H}$ such that $x_n \to x$ whenever $n \to \infty$.

Example B.3 (Square summable sequences). We define

$$\ell^2(\mathbb{N}) = \left\{ x = (x_n)_{n \in \mathbb{N}} : x_n \in \mathbb{C}, \sum_{k \in \mathbb{N}} |x_k|^2 < \infty \right\} \quad (B.1)$$

and likewise for $\ell^2(\mathbb{Z})$. Both are Hilbert spaces, where the inner product is given by $\langle x, y \rangle = \sum_{n \in \mathbb{N}} x_n \overline{y}_n$, respectively with $\mathbb{N}$ substituted by $\mathbb{Z}$.

$^1$Recall that $(x_n)$ is a Cauchy sequence if for every $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that for all $n, m \geq N$ we have $\|x_n - x_m\| < \varepsilon$. 79
Construction B.4. Given two Hilbert space $\mathcal{H}_1$ and $\mathcal{H}_2$, we construct its algebraic tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$ by equipping it with an inner product defined by $\langle v_1 \otimes v_2, w_1 \otimes w_2 \rangle = \langle v_1, w_1 \rangle_{\mathcal{H}_1} \langle v_2, w_2 \rangle_{\mathcal{H}_2}$ where $\langle \cdot, \cdot \rangle_{\mathcal{H}_i}$ is the inner product on $\mathcal{H}_i$. Taking the completion of $\mathcal{H}_1 \otimes \mathcal{H}_2$ gives a Hilbert space, which is denoted by $\mathcal{H}_1 \hat{\otimes} \mathcal{H}_2$. ■

Definition B.5. Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces, and $T : \mathcal{H} \to \mathcal{K}$ an operator. Then $T$ is called \textit{linear} if $T(au + bv) = aT(u) + bT(v)$ for all $a, b \in \mathbb{C}$ and $u, v \in \mathcal{H}$. Moreover it is \textit{bounded} if there exists a constant $M$ such that $\|Tu\|_\mathcal{K} \leq M\|u\|_\mathcal{H}$ for all $u \in \mathcal{H}$. The smallest $M$ for which this inequality holds is called the \textit{norm} of $T$, and is denoted by $\|T\|$. ■

Note that bounded linear operators are continuous.

Definition B.6. Let $T : \mathcal{H} \to \mathcal{K}$ be a bounded linear operator. The \textit{adjoint} of $T$ is defined to be the map $T^* : \mathcal{K} \to \mathcal{H}$ which satisfies $\langle Tu, v \rangle_\mathcal{K} = \langle u, T^*v \rangle_\mathcal{H}$. We say that $T^*$ is unitary if $T^*T = \mathbb{1}_\mathcal{H}$ and $TT^* = \mathbb{1}_\mathcal{K}$. ■

Definition B.7. An operator $T : \mathcal{H} \to \mathcal{H}$ is called \textit{self-adjoint} if $T^* = T$. If it moreover satisfies $T^2 = T$, then it is said to be a \textit{projection}. ■

B.1.2 The spectral theorem for bounded self-adjoint operators

Before defining the \textit{spectrum} of an operator, we introduce the following notation.

Notation B.8. Let $T$ be a bounded operator on $\mathcal{H}$. For each $z \in \mathbb{C}$ we form the \textit{resolvent operator} $G(z) := (T - z\mathbb{1}_\mathcal{H})^{-1}$ (if it exists).\footnote{Whenever several different operators are under consideration, we use the notation $G(z, T)$ in order to specify with respect to which operator the resolvent is defined.}

The spectrum can be defined conveniently in terms of the resolvent operator.

Definition B.9. The \textit{resolvent set} of a bounded operator $T$, denoted by $\rho(T)$, is defined as $\rho(T) = \{z \in \mathbb{C} : G(z) \text{ exists and is a bounded linear operator}\}$. The \textit{spectrum} of $T$, written as $\sigma(T)$, is then given by $\sigma(T) = \mathbb{C} \setminus \rho(T)$. ■

Definition B.10. We define the \textit{point spectrum}, \textit{continuous spectrum} and \textit{residual spectrum} as subsequently

$\sigma_p(T) = \{z \in \sigma(T) : T - z\mathbb{1}_\mathcal{H} \text{ is not one-to-one}\}$;
$\sigma_c(T) = \{z \in \sigma(T) : T - z\mathbb{1}_\mathcal{H} \text{ is one-to-one, and } (T - z\mathbb{1}_\mathcal{H})\mathcal{H} \subseteq \mathcal{H} \text{ is dense}\}$;
$\sigma_r(T) = \sigma(T) \setminus (\sigma_p(T) \cup \sigma_c(T))$. ■

Remark B.11.

- Note that if $z \in \sigma_p(T)$, then there exists $v \in \mathcal{H}$ such that $(T - z\mathbb{1}_\mathcal{H})v = 0$, in other words $z$ is an eigenvalue.

- For self-adjoint operators, the spectrum only consists of the point and continuous spectrum.
For a bounded operator $T$, the spectrum $\sigma(T)$ is a compact subset of the disk of radius $\|T\|$. Moreover, if $T$ is also self-adjoint, then $\sigma(T) \subset \mathbb{R}$, so that $\sigma(T) \subset [-\|T\|,\|T\|]$ and the spectrum consists of the point spectrum and the continuous spectrum.

Definition B.12. A resolution of the identity $E$ of a Hilbert space $\mathcal{H}$ is a projection-valued Borel measure on $\mathbb{R}$ such that for all Borel sets $A, B \subseteq \mathbb{R}$,

(i) $E(A)$ is a self-adjoint projection, i.e. $E(A)^* = E(A)$ and $E(A)^2 = E(A)$;

(ii) $E(A \cap B) = E(A)E(B)$;

(iii) $E(\emptyset) = 0$ and $E(\mathbb{R}) = 1_{\mathcal{H}}$;

(iv) $E(A \cup B) = E(A) + E(B)$ whenever $A \cap B = \emptyset$;

(v) for all $u, v \in \mathcal{H}$, the map $A \mapsto E_{u,v}(A) := \langle E(A)u, v \rangle$ is a complex Borel measure.

The following theorem can be found in [13], Section X.2.

Theorem B.13 (Spectral theorem for bounded self-adjoint operators). Let $T : \mathcal{H} \to \mathcal{H}$ be a bounded self-adjoint linear map. Then there exists a unique resolution of the identity $E$ such that $T = \int_{\mathbb{R}} t dE(t)$, i.e. $\langle Tu, v \rangle = \int_{\mathbb{R}} t dE_{u,v}(t)$. Moreover, $E$ is supported on the spectrum $\sigma(T)$, which is contained in the interval $[-\|T\|,\|T\|]$. Furthermore, any of the spectral projections $E(A)$ (with $A \subseteq \mathbb{R}$ a Borel set) commutes with $T$.

Note that a more general theorem of this kind holds for normal operators, i.e. operators $T$ that satisfy $T^*T = TT^*$.

The spectral measure can be obtained from the resolvent operators by the Stieltjes-Perron inversion formula (see [13], Theorem X.6.1).

Theorem B.14. The spectral measure of the open interval $(a, b) \subset \mathbb{R}$ is given by

$$E_{u,v}((a, b)) = \lim_{\delta \downarrow 0} \lim_{\varepsilon \downarrow 0} \frac{1}{2\pi i} \int_{a+\delta}^{b-\delta} (G(x+i\varepsilon)u, v) - (G(x-i\varepsilon)u, v) dx.$$ 

This limit holds in the strong operator topology, which means that $T_n x \to Tx$ for all $x \in \mathcal{H}$.

### B.2 Unbounded operators

In this section $T$ denotes an unbounded operator.

Notation B.15. An unbounded operator is usually denoted as $(T, \mathcal{D}(T))$, where $\mathcal{D}(T)$ is the domain of $T$, i.e. $\mathcal{D}(T)$ is a linear subspace of $\mathcal{H}$ and $T : \mathcal{D}(T) \to \mathcal{H}$.

Definition B.16. An unbounded linear operator $T$ is said to be densely defined if the closure of $\mathcal{D}(T)$ equals $\mathcal{H}$.

From now on we always assume that $T$ is densely defined. Then we define the adjoint operator $(T^*, \mathcal{D}(T^*))$ as follows.
**Definition B.17.** The domain $\mathcal{D}(T^*)$ is defined as

$$\mathcal{D}(T^*) = \{ v \in \mathcal{H} : u \mapsto \langle Tu, v \rangle \text{ is continuous on } \mathcal{D}(T) \}.$$ 

Since $T$ is densely defined, we can extend the map $u \mapsto \langle Tu, v \rangle$ for $v \in \mathcal{D}(T^*)$ to a continuous linear functional $\omega : \mathcal{H} \to \mathbb{C}$. By the Riesz representation theorem there exists a unique $w \in \mathcal{H}$ such that $\omega(u) = \langle u, w \rangle$ for all $u \in \mathcal{H}$. The adjoint $T^*$ is then defined by $T^*v = w$, so that

$$\langle Tu, v \rangle = \langle u, T^*v \rangle$$

for all $u \in \mathcal{D}(T)$ and $v \in \mathcal{D}(T^*)$. ■

**Definition B.18.** Let $S$ and $T$ be unbounded operators on $\mathcal{H}$. We say that $T$ is an extension of $S$ if $\mathcal{D}(S) \subset \mathcal{D}(T)$ and $Sv = Tv$ for all $v \in \mathcal{D}(S)$. This is denoted as $S \subset T$. If both $S \subset T$ and $T \subset S$, we say that $S$ and $T$ are equal, written as $S = T$. ■

In terms of the graph $G(T) = \{(u, Tu) : u \in \mathcal{D}(T)\} \subset \mathcal{H} \times \mathcal{H}$ we see that $S \subset T$ if and only if $G(S) \subset G(T)$. An operator $T$ is closed if its graph is closed in the product topology of $\mathcal{H} \times \mathcal{H}$. The adjoint of a densely defined operator is a closed operator, since the graph of the adjoint is given by

$$G(T^*) = \{((-Tu, u) : u \in \mathcal{D}(T))\}^\perp$$

for the inner product $\langle (u, v), (x, y) \rangle = \langle u, x \rangle + \langle v, y \rangle$ on $\mathcal{H} \times \mathcal{H}$.

**Proof.** This follows from the following equivalent statements:

$$(v, w) \in G(T^*) \iff T^*v = w$$

$$\iff \langle Tu, v \rangle = \langle u, T^*v \rangle = \langle u, w \rangle \text{ for all } u \in \mathcal{D}(T)$$

$$\iff \langle (-Tu, u), (v, w) \rangle = -\langle Tu, v \rangle + \langle u, w \rangle = 0 \text{ for all } u \in \mathcal{D}(T)$$

$$\iff (v, w) \in \{((-Tu, u) : u \in \mathcal{D}(T))\}^\perp.$$

□

**Definition B.19.** Let $T$ be a densely defined unbounded operator. Then $T$ is symmetric if $T \subset T^*$, or

$$\langle Tu, v \rangle = \langle u, Tv \rangle$$

for all $u, v \in \mathcal{D}(T)$. Moreover $T$ is self-adjoint if $T = T^*$ (so that a self-adjoint operator is closed). ■

**Notation B.8, Definition B.9, Definition B.10 and the first two remarks made in Remark B.11 remain valid in this case, with the following exception:**

$$\rho(T) = \{ z \in \mathbb{C} : G(z) \text{ exists, and is densely defined and bounded} \}.$$ 

Note that if $G(z)$ is indeed densely defined and bounded, it can be extended to a bounded linear operator on $\mathcal{H}$.

Moreover we remark that the spectrum of an unbounded self-adjoint operator is contained in $\mathbb{R}$. 82
Remark B.20. Clearly $\mathcal{D}(T) \subset \mathcal{D}(T^*)$, hence $\mathcal{D}(T^*)$ is a dense subspace and taking the adjoint once more gives $(T^*, \mathcal{D}(T^*))$ as the minimal closed extension of $(T, \mathcal{D}(T))$, i.e. any densely defined symmetric operator has a closed extension. Thus $T \subset T^{**} \subset T^*$.

Definition B.21. A densely defined symmetric operator $T$ is essentially self-adjoint if its closure is self-adjoint, i.e. $T \subset T^{**} = T^*$.

In general, a densely defined symmetric operator $T$ might not have self-adjoint extensions. This can be measured by the so-called deficiency indices. For $z \in \mathbb{C} \setminus \mathbb{R}$ we define

$$N_z = \{ v \in \mathcal{D}(T^*) : T^* v = zv \}. \tag{B.2}$$

According to [13], Theorem XII.4.19, $\dim N_z$ is constant on both the upper and lower half-plane. For the remainder of this section, let $T$ be a densely defined symmetric operator.

Definition B.22. The deficiency indices of $T$ are defined as $n_{\pm} = \dim N_{\pm i}$.

Lemma B.23. If $T^*$ commutes with complex conjugation of $\mathcal{H}$, then $n_+ = n_- \neq 0$.

Proof. Let $v \in N_i$, i.e. $T^* v = i v$. Then\(^3\)

$$T^* \overline{v} = \overline{T^* v} = \overline{iv} = -i \overline{v}$$

so that $\overline{v} \in N_{-i}$. Similarly it follows that $\overline{v} \in N_i$ whenever $v \in N_{-i}$. We conclude that $\dim N_i = \dim N_{-i}$, in other words $n_+ = n_-$. $\square$

Lemma B.24. If $T$ is self-adjoint, then $n_+ = n_- = 0$.

Proof. Let $\lambda$ be an eigenvalue of $T$ with corresponding eigenvector $v$. Then, since $T$ is self-adjoint, it follows that

$$\lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle Tv, v \rangle = \langle v, Tv \rangle = \langle v, \lambda v \rangle = \lambda^* \langle v, v \rangle$$

hence $\lambda = \lambda^*$, in other words $\lambda \in \mathbb{R}$. Thus all eigenvalues of $T$ are real, from which we conclude $n_{\pm} = 0$. $\square$

For the proof of the following theorem, we refer to [13], Section XII.4.

Theorem B.25. Let $(T, \mathcal{D}(T))$ be a densely defined symmetric operator. Then the following statements hold.

(i) $\mathcal{D}(T^*) = \mathcal{D}(T^{**}) \oplus N_i \oplus N_{-i}$ as an orthogonal direct sum with respect to the graph norm for $T^*$ coming from the inner product $\langle u, v \rangle_{T^*} = \langle u, v \rangle + \langle T^* u, T^* v \rangle$. As a direct sum, $\mathcal{D}(T^*) = \mathcal{D}(T^{**}) + N_+ + N_-$ for general $z \in \mathbb{C} \setminus \mathbb{R}$.

(ii) Let $U$ be an isometric bijection $U : N_i \rightarrow N_{-i}$ and define $(S, \mathcal{D}(S))$ by

$$\mathcal{D}(S) = \{ u + v + Uv : u \in \mathcal{D}(T^{**}), v \in N_i \}, \quad Sw = T^* w.$$  

Then $(S, \mathcal{D}(S))$ is a self-adjoint extension of $(T, \mathcal{D}(T))$, and every self-adjoint extension of $T$ arises in this way.

From (ii) it follows that $T$ has self-adjoint extensions if and only if $n_+ = n_-$. Moreover $T$ has a unique self-adjoint extension (i.e. is self-adjoint) if and only if $n_+ = n_- = 0$; see [2], Section 79.

\(^3\)Here $\overline{v}$ denotes the vector obtained from $v$ by applying the complex conjugation of $\mathcal{H}$.

83
B.2.1 The spectral theorem for unbounded self-adjoint operators

Here we extend Theorem B.13 to the case in which $T$ is unbounded.

**Theorem B.26** (Spectral theorem for unbounded self-adjoint operators). Let $T : \mathcal{D}(T) \to \mathcal{H}$ be an unbounded self-adjoint linear map. Then there exists a unique resolution of the identity $E$ such that $T = \int_{\mathbb{R}} t \, dE(t)$, i.e. $\langle Tu, v \rangle = \int_{\mathbb{R}} t \, dE_{u,v}(t)$ for $u \in \mathcal{D}(T)$ and $v \in \mathcal{H}$. Moreover, $E$ is supported on the spectrum $\sigma(T)$, which is contained in $\mathbb{R}$. For any bounded operator $S$ that satisfies $ST \subset TS$ it holds that $E(A)S = SE(A)$, where $A \subset \mathbb{R}$ is a Borel set.

This result is taken from [13], Section XII.4. Note that the Stieltjes-Perron inversion formula, stated in Theorem B.14, remains valid in the unbounded case.

**Remark B.27.** For any measurable function $f$ we define $f(T)$ by

$$\langle f(T)u, v \rangle = \int_{\mathbb{R}} f(t) \, dE_{u,v}(t)$$

for $u \in \mathcal{D}(f(T)) = \{ u \in \mathcal{H} : \int_{\mathbb{R}} |f(t)|^2 \, dE_{u,u}(t) < \infty \}$ and $v \in \mathcal{H}$. ■
Bibliography


