Philipp J. di Dio

An Introduction to T-Systems

with a special Emphasis on Sparse Moment Problems, Sparse Positivstellensätze, and Sparse Nichtnegativstellensätze

April 4, 2024

Lecture Notes

Samuel Karlin *(June 8, 1924 – December 18, 2007)*

He solved almost unnoticed an important algebraic question.

Preface

These are the lecture notes based on [dD23] for the (upcoming) lecture *T-systems with a special emphasis on sparse moment problems and sparse Positivstellensätze* in the summer semester 2024 at the University of Konstanz.

The main purpose of this lecture is to prove the sparse Positiv- and Nichtnegativstellensätze of Samuel Karlin (1963) and to apply them to the algebraic setting. That means given finitely many monomials, e.g.

$$1, x^2, x^3, x^6, x^7, x^9,$$

how do all linear combinations of these look like which are strictly positive or non-negative on some interval $[a, b]$ or $[0, \infty)$, e.g. describe and even write down all

$$f(x) = a_0 + a_1 x^2 + a_2 x^3 + a_3 x^6 + a_4 x^7 + a_5 x^9$$

with $f(x) > 0$ or $f(x) \geq 0$ on $[a, b]$ or $[0, \infty)$, respectively.

To do this we introduce the theoretical framework in which this question can be answered: T-systems. We study these T-systems to arrive at Karlin's Positiv- and Nichtnegativstellensatz but we also do not hide the limitations of the T-systems approach.

The main limitation is the Curtis–Mairhuber–Sieklucki Theorem which essentially states that every T-system is only one-dimensional and hence we can only apply these results to the *univariate* polynomial case. This can also be understood as a lesson or even a warning that this approach has been investigated and found to fail, i.e., learning about these results and limitations shall save students and researchers from following old footpaths which lead to a dead end.

We took great care finding the correct historical references where the results appeared first but are perfectly aware that like people before we not always succeed.

Konstanz,
February 2024 *Philipp J. di Dio*

Acknowledgements

The author was supported by the Deutsche Forschungsgemeinschaft DFG with the grant DI-2780/2-1 and his research fellowship at the Zukunfskolleg of the University of Konstanz, funded as part of the Excellence Strategy of the German Federal and State Government.

The author thanks Konrad Schmüdgen for reading large parts of the manuscript and providing valuable remarks.

Contents

Chapter 0
Preliminaries

Pure mathematics is, in its way, the poetry of logical ideas.

Albert Einstein [Ein35]

The purpose of this preliminary chapter is not to establish and prove results but to clarify notation and to give the reader a survey of what will be assumed as known.

For the representation theorems of linear functionals of Daniell (Signed Daniell's Representation Theorem 0.17) and Riesz (Signed Riesz' Representation Theorem 0.18) more care is invested since these are the essential representation theorems in the theory of moments in the following chapters, i.e., we include the proofs.

0.1 Sets, Relations, and Orders

We let $\mathbb{N} := \{1, 2, 3, \dots\}$ be the natural numbers, $\mathbb{N}_0 := \{0, 1, 2, \dots, \}$ be the natural numbers including zero, and as usual $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$. The unit circle is denoted by $\mathbb{T} := \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 = 1\}$.

For inclusions we use $\subseteq$ and $\subsetneq$. To avoid any confusion we avoid the use of $\subset$ since $\subset$ is used in the literature by different authors either as $\subseteq$ or $\subsetneq$.

For a set $\mathcal{X}$ we denote by $\mathcal{P}(\mathcal{X})$ the set of all subsets of $\mathcal{X}$.

A *partial order* on a set $\mathcal{X}$ is a relation $R \subseteq \mathcal{X} \times \mathcal{X}$, usually denoted by $\leq$, such that

(i) $x = y \quad \Leftrightarrow \quad x \leq y$ and $y \leq x$,
(ii) $x \leq y$ and $y \leq z \quad \Rightarrow \quad x \leq z$.

A relation $\leq$ is a *total order* if for all $x, y \in \mathcal{X}$ we have either $x \leq y$ or $y \leq x$. A vector space E with a partial order $\leq$ such that

(i) $x \leq y$ and $z \in \mathcal{X} \quad \Rightarrow \quad x + z \leq y + z$,
(ii) $x \leq y$ and $a \in [0, \infty) \quad \Rightarrow \quad ax \leq ay$

is called an *ordered vector space*. If E is an ordered vector space then $E_+ := \{x \in E \,|\, 0 \leq x\}$ denotes the *positive cone* and $E_- := \{x \in E \,|\, x \leq 0\}$ denoted the *negative cone*. Let $C \subseteq E$ be a cone in a vector space E. Then E with $x \leq y$ if and only if $y - x \in C$ is a (partially) ordered vector space.

For a vector space E a (linear) function $f : E \to \mathbb{R}$ is called (*linear*) *functional*. For a vector space E the (*algebraic*) *dual* E^* is the set of all linear functionals $f : E \to \mathbb{R}$. A functional $f : E \to \mathbb{R}$ is called *sublinear* if $f(\rho x) \leq \rho f(x)$ and $f(x + y) \leq f(x) + f(y)$ hold for all $\rho \geq 0$ and $x, y \in E$. It is called *superlinear* if $-f$ is sublinear.

Hahn–Banach Theorem 0.1. *Let $\mathcal{X}$ be a real vector space, let $p : \mathcal{X} \to \mathbb{R}$ be a sublinear function, $\mathcal{V} \subseteq \mathcal{X}$ be a subspace, and $f : \mathcal{V} \to \mathbb{R}$ be a linear functional such that $f(x) \leq p(x)$ for all $x \in \mathcal{V}$. Then there exists a linear functional $F : \mathcal{X} \to \mathbb{R}$ such that*

(i) $f(x) = F(x)$ for all $x \in \mathcal{V}$, and
(ii) $F(x) \leq p(x)$ for all $x \in \mathcal{X}$.

The Hahn–Banach Theorem 0.1 was proved by H. Hahn [Hah27] and S. Banach [Ban29a, Ban29b]. A previous version is due to E. Helly [Hel12]. For more see e.g. [Pie07] or standard functional analysis textbooks like [Yos68, Wer07].

0.2 Topology

A *topology* $\mathcal{T}$ on a set $\mathcal{X}$ is a set $\mathcal{T} \subseteq \mathcal{P}(\mathcal{X})$ of subsets of $\mathcal{X}$ which is closed under finite intersections and arbitrary unions, i.e., especially $\emptyset, \mathcal{X} \in \mathcal{T}$. $(\mathcal{X}, \mathcal{T})$ is called a *topological space* and sets $A \in \mathcal{T}$ are called *open*. A set $A \subseteq \mathcal{X}$ is called *closed* if $\mathcal{X} \setminus A$ is open. The *interior* $\operatorname{int} A$ of a set $A \subseteq \mathcal{X}$ is the union of all open sets $O \subseteq A$. A subset U of a topological space $(\mathcal{X}, \mathcal{T})$ is called a *neighborhood* of x if $x \in \operatorname{int} U$.

A function $f : \mathcal{X} \to \mathcal{Y}$ between two topological spaces $\mathcal{X}$ and $\mathcal{Y}$ is called *continuous* at $x \in \mathcal{X}$ if for each neighborhood V of $y = f(x)$ the set $f^{-1}(V)$ is a neighborhood of x. The function f is called *continuous* if it is continuous at every $x \in \mathcal{X}$. The set of continuous functions $f : \mathcal{X} \to \mathcal{Y}$ is denoted by $C(\mathcal{X}, \mathcal{Y})$. A set $K \subseteq \mathcal{X}$ is called *compact* if every open cover $K \subseteq \bigcup_{i \in I} U_i$, $U_i \in \mathcal{T}$, has a finite subcover $K \subseteq \bigcup_{k=1}^n U_{i_k}$. For a function $f : \mathcal{X} \to \mathbb{R}$ we have the support $\operatorname{supp} f := \{x \in \mathcal{X} \mid f(x) \neq 0\}$. The set of all continuous functions with compact support are denoted by $C_c(\mathcal{X}, \mathbb{R})$.

A topological space $\mathcal{X}$ is called *Hausdorff space* if each pair of distinct points $x, y \in \mathcal{X}$ have disjoint neighborhoods. A Hausdorff space $\mathcal{X}$ is called *locally compact* if every point $x \in \mathcal{X}$ has a compact neighborhood. On Hausdorff spaces we have the following important topological result.

Urysohn's Lemma 0.2 (see [Ury25])**.** *Let $\mathcal{X}$ be a Hausdorff space. The following are equivalent:*

(i) For every pair of disjoint closed sets $A, B \subseteq \mathcal{X}$ there exist a neighborhood U of A and a neighborhood V of B such that $U \cap V = \emptyset$.
(ii) For each pair $A, B \subseteq \mathcal{X}$ of disjoint closed sets there exists a continuous function $f : \mathcal{X} \to [0, 1]$ such that $f(x) = 1$ for all $x \in A$ and $f(y) = 0$ for all $y \in B$.

0.3 Stone–Weierstrass Theorem

Stone–Weierstrass Theorem 0.3 ([Wei85] and [Sto37, pp. 467–468] or e.g. [Yos68, p. 9])**.** *Let $\mathcal{X}$ be a compact set and let $B \subseteq C(\mathcal{X}, \mathbb{R})$ be such that*

(i) $fg, \alpha f + \beta g \in B$ for all $f, g \in B$ and $\alpha, \beta \in \mathbb{R}$,
(ii) there exists a $f \in B$ with $f > 0$ on $\mathcal{X}$, and
(iii) for all $x, y \in \mathcal{X}$ with $x \neq y$ there is a $f \in B$ such that $f(x) \neq f(y)$

then for any $f \in C(\mathcal{X}, \mathbb{R})$ there exists $\{f_n\}_{n\in\mathbb{N}_0} \subseteq B$ such that

$$\|f - f_n\|_\infty \xrightarrow{n\to\infty} 0.$$

Especially $\mathbb{R}[x_1, \dots, x_n]$ on any compact $K \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, is dense in $C(K, \mathbb{R})$ in the sup-norm.

For more on the history of the Stone–Weierstrass Theorem 0.3 see e.g. [Pie07, §4.5.6–§4.5.8].

0.4 Convex Geometry

A set $\mathcal{X}$ is *convex* if $\lambda x + (1-\lambda)y \in \mathcal{X}$ for all $x, y \in \mathcal{X}$ and $\lambda \in [0, 1]$. A set $\mathcal{X}$ is a *cone* if $\lambda x \in \mathcal{X}$ for all $x \in \mathcal{X}$ and $\lambda \in [0, \infty)$. For a set $A \subseteq \mathbb{R}^n$ we denote by $\operatorname{conv} A$ the *convex hull* of A.

Carathéodory's Theorem 0.4 (see [Car11])**.** *Let $n \in \mathbb{N}$ and let $S \subseteq \mathbb{R}^n$ be a set. If $x \in \operatorname{conv} A$ then there is a $k \leq n+1$, points $x_1, \dots, x_k \in A$, and $\lambda_1, \dots, \lambda_k > 0$ with*

$$x = \lambda_1 x_1 + \dots + \lambda_k x_k \qquad \text{and} \qquad \lambda_1 + \dots + \lambda_k = 1.$$

For more on convex geometry we recommend [Roc72] and [Sch14].

0.5 Linear Algebra

A matrix $M = (a_{i,j})_{i,j=1}^n$ with $a_{i,j} = a_{k,l}$ if $i + j = k + l$ is called *Hankel matrix*. For a sequence $s = (s_\alpha)_{\alpha\in\mathbb{N}_0 : |\alpha|\leq 2n}$ with $n \in \mathbb{N}_0$ we denote by

$$\mathcal{H}(s) := (s_{\alpha+\beta})_{\alpha,\beta\in\mathbb{N}_0 : |\alpha|,|\beta|\leq n} \tag{0.1}$$

the *Hankel matrix of s*.

0.6 Measures

For a set $\mathcal{X}$ an *algebra* $\mathfrak{A}$ is a set $\mathfrak{A} \subseteq \mathcal{P}(\mathcal{X})$ such that $\emptyset, \mathcal{X} \in \mathfrak{A}$ and for all $A, B \in \mathfrak{A}$ we have $A \cap B, A \cup B, A \setminus B \in \mathfrak{A}$. If additionally $\bigcup_{n=1}^{\infty} A_n \in \mathfrak{A}$ for all $A_n \in \mathfrak{A}$ then $\mathfrak{A}$ called a *σ-algebra* and $(\mathcal{X}, \mathfrak{A})$ is called a *measurable space*. By $\mathfrak{B}(\mathbb{R}^n)$ we denote the *Borel σ-algebra*. For $A \subseteq \mathcal{P}(\mathcal{X})$ we denote by $\sigma(A)$ the smallest σ-algebra containing A. A function $f : (\mathcal{X}, \mathfrak{A}) \to (\mathcal{Y}, \mathfrak{B})$ between two measurable spaces is called *measurable* if $f^{-1}(B) \in \mathfrak{A}$ for all $B \in \mathfrak{B}$.

A *measure*[1] μ is a function $\mu : \mathfrak{A} \to [0, \infty]$ on an algebra $\mathfrak{A}$ such that μ is *countably additive*, i.e.,

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

for all pairwise disjoint sets $A_n \in \mathfrak{A}$. A measure μ on $\mathfrak{B}(\mathbb{R}^n)$ is called *Borel measure*. A Borel measure μ is called a *Radon measure* if for every $A \in \mathfrak{B}(\mathbb{R}^n)$ and $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subseteq A$ such that $\mu(A \setminus K_\varepsilon) < \varepsilon$. We denote by $\mathcal{M}(\mathcal{X})_+$ the set of all Borel measures on $(\mathcal{X}, \mathfrak{A})$. By $(\mathcal{X}, \mathfrak{A}, \mu)$ we denote a *measure space*. A measurable function $f : (\mathcal{X}, \mathfrak{A}) \to \mathbb{R}$ is called *μ-integrable* if

$$\int_{\mathcal{X}} |f(x)| \, \mathrm{d}\mu(x) < \infty.$$

For any $p \geq 1$ we denote by $\mathcal{L}^p(\mathcal{X}, \mu)$ all μ-integrable functions on $\mathcal{X}$. For $p = \infty$, i.e., $\mathcal{L}^\infty(\mathcal{X}, \mu)$, the essential supremum is bounded.

Since we are proving the (signed) Daniell's Theorem and the (signed) Riesz' Representation Theorem we will give a more detailed background on measures. For more on measure theory we recommend [Bog07] and [Fed69].

Definition 0.5. Let $\mathcal{X}$ be a set. A function $\mu : \mathcal{P}(\mathcal{X}) \to [0, \infty]$ with

(i) $\mu(\emptyset) = 0$,
(ii) $\mu(A) \leq \mu(B)$ for all $A \subseteq B \subseteq \mathcal{X}$, and
(iii) $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i)$ for all $A_i \in \mathcal{X}$

is called a (*Carathéodory) outer measure*.

Definition 0.6. For an outer measure μ on $\mathcal{X}$ a set $A \subseteq \mathcal{X}$ is called (*Carathéodory*) *μ-measurable* if for every $E \subseteq \mathcal{X}$ we have $\mu(E) = \mu(E \cap A) + \mu(E \setminus A)$.

Remark 0.7. Since by Definition 0.5 (iii) we always have

$$\mu(E) = \mu((E \cap A) \cup (E \setminus A)) \leq \mu(E \cap A) + \mu(E \setminus A)$$

it is sufficient for μ-measurability to test

$$\mu(E) \geq \mu(E \cap A) + \mu(E \setminus A). \tag{0.2}$$

[1] For us all measures are non-negative unless stated otherwise. In [Bog07] the theory is developed in greater generality.

An outer measure is in fact a measure on all its measurable sets.

Theorem 0.8. *Let μ be an outer measure on a set $\mathcal{X}$ and $\mathcal{A}_\mu \subseteq \mathcal{P}(\mathcal{X})$ be the set of all μ-measurable sets. Then $\mathcal{A}_\mu$ is a σ-algebra of $\mathcal{X}$ and μ is a measure on $(\mathcal{X}, \mathcal{A}_\mu)$.*

Proof. See e.g. [Bog07, Thm. 1.11.4 (iii)]. □

Outer measures give another characterization of measurable functions.

Lemma 0.9. *Let μ be an outer measure on $\mathcal{X}$ and $f : \mathcal{X} \to [-\infty, \infty]$ be a function. Then f is μ-measurable if and only if*

$$\mu(A) \geq \mu(\{x \in A \mid f(x) \leq a\}) + \mu(\{x \in A \mid f(x) \geq b\})$$

for all $A \subseteq \mathcal{X}$ and $-\infty < a < b < \infty$.

Proof. See e.g. [Fed69, §2.3.2(7), pp. 74–75]. □

Definition 0.10. An outer measure μ is called *regular* if for each set $A \subseteq \mathcal{X}$ there exists a μ-measurable set $B \subseteq \mathcal{X}$ with $A \subseteq B$ and $\mu(A) = \mu(B)$.

Definition 0.11. Let $f, g : (\mathcal{X}, \mathcal{A}) \to \mathbb{R}$ be two functions. Then we define $\inf(f, g)$ by

$$\inf(f, g)(x) := \inf(f(x), g(x))$$

for all $x \in \mathcal{X}$ and similarly $\sup(f, g)$. Additionally, $f \leq g$ iff $f(x) \leq g(x)$ for all $x \in \mathcal{X}$. We have $f_+ := \sup(f, 0)$, $f_- := f - f_+$, and $|f| = f_+ - f_-$.

Definition 0.12. Let $\mathcal{X}$ be a set. We call a set $\mathcal{F}$ of functions $f : \mathcal{X} \to \mathbb{R}$ a *lattice* (*of functions*) if the following holds:

(i) $c \cdot f \in \mathcal{F}$ for all $c \geq 0$ and $f \in \mathcal{F}$,
(ii) $f + g \in \mathcal{F}$ for all $f, g \in \mathcal{F}$,
(iii) $\inf(f, g) \in \mathcal{F}$ for all $f, g \in \mathcal{F}$,
(iv) $\inf(f, c) \in \mathcal{F}$ for all $c \geq 0$ and $f \in \mathcal{F}$, and
(v) $g - f \in \mathcal{F}$ for all $f, g \in \mathcal{F}$ with $f \leq g$.

Some authors require that a lattice of functions is a vector space (*lattice space*). But for proving Daniell's Representation Theorem 0.15 it is only necessary that a lattice is a convex cone as in Definition 0.12.

Example 0.13. Let $\mathcal{X}$ be a locally compact Hausdorff space. Then $C_c(\mathcal{X}, \mathbb{R})$ is a lattice of functions and even a lattice space. ○

Given a lattice $\mathcal{F}$ we get another lattice $\mathcal{F}_+$ by taking only the non-negative functions.

Lemma 0.14 (see e.g. [Fed69, §2.5.1, p. 91])**.** *Let $\mathcal{F}$ be a non-empty lattice on a set $\mathcal{X}$ and define*

$$\mathcal{F}_+ := \{f \in \mathcal{F} \mid f \geq 0\}.$$

Then

(i) $f_+, f_-, |f| \in \mathcal{F}_+$ *for all* $f \in \mathcal{F}$ *and*
(ii) $\mathcal{F}_+$ *is a non-empty lattice on* $\mathcal{X}$.

Proof. (i): Since $\inf(f,0) \in \mathcal{F}$ and $\inf(f,0) \leq f$ we have $f_+ = \sup(f,0) = f - \inf(f,0) \in \mathcal{F}_+$ for all $f \in \mathcal{F}$. Since $f \leq f_+ = \sup(f,0) \in \mathcal{F}$ we have $f_- = f_+ - f \in \mathcal{F}_+$ for all $f \in \mathcal{F}$. It follows that $|f| = f_+ + f_- \in \mathcal{F}_+$ for all $f \in \mathcal{F}$.

(ii): Since $\mathcal{F}$ is non-empty there is a $f \in \mathcal{F}$ and by (ii) we have $|f| \in \mathcal{F}$ and hence $|f| \in \mathcal{F}_+$. $\mathcal{F}_+$ is a lattice by directly checking the Definition 0.12. □

0.7* Daniell's Representation Theorem

The question when a linear functional acting on (not necessarily measurable) functions is represented by a measure was already fully answered by P. J. Daniell in 1918 [Dan18], see also [Dan20].

Nowadays only the Riesz' Representation Theorem 0.20 is given in standard texts for the moment problem. We therefore take the time to present also Daniell's approach which is more general and has some interesting features the standard Riesz' Representation Theorem 0.20 does not have.

Note, that $h_n \nearrow g$ denotes a sequence $(h_n)_{n\in\mathbb{N}}$ with $h_1 \leq h_2 \leq \dots \leq g$, i.e., point-wise non-decreasing, with $\lim_{n\to\infty} h_n(x) = g(x)$ for all $x \in \mathcal{X}$. Equivalently, $h_n \searrow 0$ denotes a point-wise non-increasing sequence with $\lim_{n\to\infty} h_n(x) = 0$ for all $x \in \mathcal{X}$.

Daniell's Representation Theorem 0.15 ([Dan18], see also [Dan20] or [Fed69, Thm. 2.5.2])**.** *Let* $\mathcal{F}$ *be a lattice of functions on a set* $\mathcal{X}$ *and let* $L : \mathcal{F} \to \mathbb{R}$ *be such that*

(i) $L(f+g) = L(f) + L(g)$ *for all* $f, g \in \mathcal{F}$,
(ii) $L(c \cdot f) = c \cdot L(f)$ *for all* $c \geq 0$ *and* $f \in \mathcal{F}$,
(iii) $L(f) \leq L(g)$ *for all* $f, g \in \mathcal{F}$ *with* $f \leq g$,
(iv) $L(f_n) \nearrow L(g)$ *as* $n \to \infty$ *for all* $g \in \mathcal{F}$ *and* $f_n \in \mathcal{F}$ *with* $f_n \nearrow g$.

Then there exists a measure μ *on* $(\mathcal{X}, \mathcal{A})$ *with*

$$\mathcal{A} := \sigma(\{f^{-1}((-\infty, a]) \,|\, a \in \mathbb{R},\ f \in \mathcal{F}\}) \tag{0.3}$$

such that

$$L(f) = \int_{\mathcal{X}} f(x)\, \mathrm{d}\mu(x)$$

for all $f \in \mathcal{F}$.

We follow the proof in [Fed69, Thm. 2.5.2, pp. 92–93].

Proof. By assumption (iii) we have $L(f) \geq L(0 \cdot f) = 0$ for all $f \in \mathcal{F}_+$.

For any $A \subseteq \mathcal{X}$ we say a sequence $(f_n)_{n\in\mathbb{N}}$ *suits* A if and only if $f_n \in \mathcal{F}_+$ and $f_n \leq f_{n+1}$ for all $n \in \mathbb{N}$ and

$$\lim_{n\to\infty} f_n(x) \geq 1 \qquad \text{for all } x \in A.$$

Note, that we can even assume equality by replacing the f_n's by $\tilde{f}_n := \inf(f_n, 1) \in \mathcal{F}_+$. Then we define

$$\mu(A) := \inf\left\{\lim_{n\to\infty} L(f_n) \;\middle|\; (f_n)_{n\in\mathbb{N}} \text{ suits } A\right\} \tag{0.4}$$

which is ∞ if there is no sequence $(f_n)_{n\in\mathbb{N}}$ that suits A.

We prove that μ is an outer measure, see Definition 0.5. By assumption (iii) $L(f_n)$ is a non-negative increasing sequence and therefore $\lim_{n\to\infty} L(f_n)$ exists and is in $[0,\infty]$. Hence, $\mu : \mathcal{P}(\mathcal{X}) \to [0,\infty]$. For $A = \emptyset$ the zero sequence $f_n = 0 \in \mathcal{F}_+$ is suited and therefore $\mu(\emptyset) = 0$. Let $A \subseteq B \subseteq \mathcal{X}$, then a suited sequence $(f_n)_{n\in\mathbb{N}}$ of B is also a suited sequence for A and therefore $\mu(A) \leq \mu(B)$. Let $A_i \subseteq \mathcal{X}$, $i \in \mathbb{N}$, and set $A := \bigcup_{i=1}^\infty A_i$. Any suited sequence for A is a suited sequences for all A_i. Assume there is an A_i which has no suited sequence, then A has no suited sequence and $\mu(A) = \infty \leq \sum_{i=1}^\infty \mu(A_i) = \infty$. So assume all A_i have suited sequences, say $(f_{i,n})_{n\in\mathbb{N}}$ suits A_i, $i \in \mathbb{N}$. Then $f_n := \sum_{i=1}^n f_{i,n}$ suits A and

$$\mu(A) \leq \lim_{n\to\infty} L(f_n) = \lim_{n\to\infty} \sum_{i=1}^n L(f_{i,n}) \leq \sum_{i=1}^\infty \lim_{m\to\infty} L(f_{i,m}).$$

Taking the infimum on the right side for all A_i's retains the inequality and gives

$$\mu\left(\bigcup_{i=1}^\infty A_i\right) = \mu(A) \leq \sum_{i=1}^\infty \mu(A_i).$$

Hence, all conditions in Definition 0.5 are fulfilled and μ is an outer measure.

Since μ is an outer measure on $\mathcal{X}$ by Theorem 0.8 the set $\tilde{\mathcal{A}}$ of all μ-measurable sets of $\mathcal{X}$ is a σ-algebra and μ is a measure on $(\mathcal{X}, \tilde{\mathcal{A}})$.

It remains to show that all $f \in \mathcal{F}$ are μ-measurable, μ is a measure on $(\mathcal{X}, \mathcal{A})$ with $\mathcal{A} = \sigma(\{f^{-1}((-\infty,a]) \mid a \in \mathbb{R},\ f \in \mathcal{F}\})$, and $L(f) = \int_{\mathcal{X}} f(x)\,\mathrm{d}\mu(x)$ for all $f \in \mathcal{F}$.

Since $f = f_+ - f_-$ with $f_+, f_- \in \mathcal{F}_+$ it is sufficient to show that every function in $\mathcal{F}_+$ is μ-measurable. So let $f \in \mathcal{F}_+$. To show that f is μ-measurable it is sufficient to show that $A := f^{-1}((-\infty,a]) = \{x \in \mathcal{X} \mid f(x) \leq a\} \in \mathcal{A}$ for all $a \in \mathbb{R}$, i.e., A is μ-measurable by Definition 0.6 resp. Remark 0.7 if (0.2) holds for all $E \subseteq \mathcal{X}$. From $E \setminus A = E \cap (\mathcal{X} \setminus A) = E \cap \{x \in \mathcal{X} \mid f(x) > a\}$ we have to verify

$$\mu(E) \geq \mu\big(\{x \in E \mid f(x) \leq a\}\big) + \mu\big(\{x \in E \mid f(x) > a\}\big)$$

and by Lemma 0.9 this is equivalent to

$$\mu(E) \geq \mu\big(\underbrace{\{x \in E \mid f(x) \leq a\}}_{=:E_a}\big) + \mu\big(\underbrace{\{x \in E \mid f(x) \geq b\}}_{=:E_b}\big) \tag{0.5}$$

for all $a < b$. For $a < 0$ or $\mu(E) = \infty$ (0.5) is trivial, so assume $a \geq 0$ and $\mu(E) < \infty$.
Let $(g_n)_{n\in\mathbb{N}}$ be a sequence that suits E and set

$$h := (b-a)^{-1} \cdot [\inf(f,b) - \inf(f,a)] \in \mathcal{F}_+ \qquad \text{and} \qquad k_n := \inf(g_n, h) \in \mathcal{F}_+.$$

Then we have $0 \leq k_{n+1} - k_n \leq g_{n+1} - g_n$,

$$h(x) = 1 \qquad \text{for all } x \in \mathcal{X} \text{ with } f(x) \geq b,$$

and

$$h(x) = 0 \qquad \text{for all } x \in \mathcal{X} \text{ with } f(x) \leq a.$$

It follows that $(k_n)_{n\in\mathbb{N}}$ suits E_b and $(g_n - k_n)_{n\in\mathbb{N}}$ suits E_a. Therefore,

$$\lim_{n\to\infty} L(g_n) = \lim_{n\to\infty} [L(g_n - k_n) + L(k_n)] \geq \mu(E_a) + \mu(E_b)$$

and taking the infimum on the left side retains the inequality and proves (0.5). Hence, all $f \in \mathcal{F}_+$ and therefore all $f \in \mathcal{F}$ are μ-measurable.

Let us show that μ remains a measure on $(\mathcal{X}, \mathcal{A})$. Since all $f \in \mathcal{F}$ are μ- and $\mathcal{A}$-measurable we have

$$f^{-1}((-\infty, a]) \in \tilde{\mathcal{A}}$$

for all $a \in \mathbb{R}$ and $f \in \mathcal{F}$. Therefore,

$$\mathcal{A} = \sigma(\{f^{-1}((-\infty, a]) \mid a \in \mathbb{R},\ f \in \mathcal{F}\}) \subseteq \tilde{\mathcal{A}}$$

is a σ-algebra and we can restrict μ resp. $\tilde{\mathcal{A}}$ to $\mathcal{A}$. μ is a measure on $(\mathcal{X}, \mathcal{A})$.

We show that $L(f) = \int_{\mathcal{X}} f(x)\, \mathrm{d}\mu(x)$ holds for all $f \in \mathcal{F}_+$. Let $f \in \mathcal{F}_+$ and set

$$f_t := \inf(f, t)$$

for $t \geq 0$. If $\varepsilon > 0$ and $k \in \mathbb{N}$ then

$$\begin{aligned} 0 \leq f_{k\varepsilon}(x) - f_{(k-1)\varepsilon}(x) &\leq \varepsilon \quad \text{for all } x \in \mathcal{X}, \\ f_{k\varepsilon}(x) - f_{(k-1)\varepsilon}(x) &= \varepsilon \quad \text{for all } x \in \mathcal{X} \text{ with } f(x) \geq k\varepsilon, \end{aligned}$$

and

$$f_{k\varepsilon}(x) - f_{(k-1)\varepsilon}(x) = 0 \quad \text{for all } x \in \mathcal{X} \text{ with } f(x) \leq (k-1)\varepsilon.$$

The constant sequence $(\varepsilon^{-1} \cdot (f_{k\varepsilon} - f_{(k-1)\varepsilon}))_{n\in\mathbb{N}}$ suits $\{x \in \mathcal{X} \mid f(x) \geq k\varepsilon\}$ and consequently

$$\begin{aligned} L(f_{k\varepsilon} - f_{(k-1)\varepsilon}) &\geq \varepsilon \cdot \mu(\{x \in \mathcal{X} \mid f(x) \geq k\varepsilon\}) \\ &\geq \int_{\mathcal{X}} f_{(k+1)\varepsilon}(x) - f_{k\varepsilon}(x)\, \mathrm{d}\mu(x) \end{aligned}$$

$$\geq \varepsilon \cdot \mu(\{x \in \mathcal{X} \mid f(x) \geq (k+1)\varepsilon\}) \geq L(f_{(k+2)\varepsilon} - f_{(k+1)\varepsilon}).$$

Summing with respect to k from 1 to n we find

$$L(f_{n\varepsilon}) \geq \int_{\mathcal{X}} f_{(n+1)\varepsilon}(x) - f_{\varepsilon}(x)\, \mathrm{d}\mu(x) \quad \geq L(f_{(n+2)\varepsilon} - f_{2\varepsilon})$$

and since $f_{n\varepsilon} \nearrow f$ as $n \to \infty$ we get from assumption (iv) for $n \to \infty$

$$L(f) \geq \int_{\mathcal{X}} f(x) - f_{\varepsilon}(x)\, \mathrm{d}\mu(x) \quad \geq L(f - f_{2\varepsilon})$$

which gives again from assumption (iv) for $\varepsilon \searrow 0$

$$L(f) \geq \int_{\mathcal{X}} f(x)\, \mathrm{d}\mu(x) \quad \geq L(f).$$

Hence, $L(f) = \int_{\mathcal{X}} f(x)\, \mathrm{d}\mu(x)$ for all $f \in \mathcal{F}_+$.

Finally, for all $f \in \mathcal{F}$ we have $f = f_+ - f_-$ with $f_+, f_- \in \mathcal{F}_+$ which implies

$$\int_{\mathcal{X}} f(x)\, \mathrm{d}\mu(x) = \int_{\mathcal{X}} f_+(x)\, \mathrm{d}\mu(x) - \int_{\mathcal{X}} f_-(x)\, \mathrm{d}\mu(x) = L(f_+) - L(f_-) = L(f)$$

where the last equality follows from $f_+ = f + f_-$ and assumption (i). □

The most impressive part is that the functional $L : \mathcal{F} \to \mathbb{R}$ lives only on a lattice $\mathcal{F}$ of functions $f : \mathcal{X} \to \mathbb{R}$ where $\mathcal{X}$ is a set without any structure. Daniell's Representation Theorem 0.15 provides a representing measure μ by (0.4) including the σ-algebra $\mathcal{A}$ of the measurable space $(\mathcal{X}, \mathcal{A})$ by (0.3).

Remark 0.16. In Daniell's Representation Theorem 0.15 the assumption (iv) is equivalent to

(iv') $L(h_n) \searrow 0$ as $n \to \infty$ for all $h_n \in \mathcal{F}$ with $h_n \searrow 0$ as $n \to \infty$

since $f_n \nearrow g$ implies $f_n \leq g$ and $0 \leq h_n = g - f_n \in \mathcal{F}$:

$$L(g) = L(g - f_n + f_n) = L(g - f_n) + \underbrace{L(f_n)}_{\nearrow L(g)} = \underbrace{L(h_n)}_{\searrow 0} + L(f_n). \qquad \circ$$

The representing measure μ in Daniell's Representation Theorem 0.15 is not unique. But the representing measure μ constructed in (0.4) has further properties, see e.g. [Fed69, §2.5.3].

Daniell's Representation Theorem 0.15 also has a signed version.

Signed Daniell's Representation Theorem 0.17 ([Dan18], see also [Fed69, Thm. 2.5.5])**.** *Let $\mathcal{F}$ be a lattice of functions on some set $\mathcal{X}$ and let $L : \mathcal{F} \to \mathbb{R}$ be such that for all $f, g, h_1, h_2, h_3, \ldots \in \mathcal{F}$ we have*

(a) $L(f+g) = L(f) + L(g)$,
(b) $L(c \cdot f) = c \cdot L(f)$ *for all* $c \geq 0$,
(c) $\sup L\big(\{k \in \mathcal{F} \mid 0 \leq k \leq f\}\big) < \infty$,
(d) $h_n \nearrow g$ *as* $n \to \infty$ *implies* $L(h_n) \to L(g)$ *as* $n \to \infty$.

Let L_+ *and* L_- *be the functionals on* $\mathcal{F}_+$ *defined by*

$$L_+(f) := \sup L\big(\{k \in \mathcal{F} \mid 0 \leq k \leq f\}\big)$$

and

$$L_-(f) := -\inf L\big(\{k \in \mathcal{F} \mid 0 \leq k \leq f\}\big)$$

for all $f \in \mathcal{F}_+$. *Then there exist* $\mathcal{F}_+$ *regular measures* μ_+ *and* μ_- *on* $\mathcal{X}$ *such that*

(i) $L_+(f) = \int_{\mathcal{X}} f(x)\, \mathrm{d}\mu_+(x)$ *for all* $f \in \mathcal{F}_+$,
(ii) $L_-(f) = \int_{\mathcal{X}} f(x)\, \mathrm{d}\mu_-(x)$ *for all* $f \in \mathcal{F}_+$, *and*
(iii) $L(f) = L_+(f) - L_-(f)$ *for all* $f \in \mathcal{F}$.

The proof is taken from [Fed69, pp. 96–97] and uses Daniell's Representation Theorem 0.15.

Proof. Let $f_+ \in \mathcal{F}_+$. Then $f \geq g \in \mathcal{F}_+$ implies $f \geq f - g \in \mathcal{F}_+$ and

$$L(g) - L_-(f) \leq L(g) + L(f-g) \leq L(g) + L_+(f).$$

Hence,

$$L_+(f) - L_-(f) \leq L(f) \leq -L_-(f) + L_+(f)$$

so that

$$L(f) = L_+(f) - L_-(f).$$

Now let $f, g \in \mathcal{F}_+$. If $f + g \geq h \in \mathcal{F}_+$ then

$$f \geq k := \inf(f, h) \in \mathcal{F}_+ \quad \text{and} \quad g \geq h - k \in \mathcal{F}_+$$

and hence

$$L_+(f) + L_+(g) \geq L(k) + L(h-k) = L(h).$$

Therefore, $L_+(f) + L_+(g) \geq L_+(f+g)$. Since the opposite inequality is clear, we have that L_+ is additive on $\mathcal{F}_+$. Additionally, L_+ is positively homogeneous and monotone.

We now show that L_+ preserves increasing convergence. Suppose $h_n \nearrow g$ as $n \nearrow \infty$ with $g, h_n \in \mathcal{F}_+$. If $g \geq k \in \mathcal{F}_+$ then $f_n := \inf(h_n, k) \nearrow k$ as $n \nearrow \infty$, i.e.,

$$L(k) = \lim_{n\to\infty} L(f_n) \leq \lim_{n\to\infty} L_+(h_n).$$

Hence, $L_+(h_n) \nearrow L_+(g)$ as $n \nearrow \infty$. By Daniell's Representation Theorem 0.15 we have that there is a $\mathcal{F}_+$ regular measure μ_+ on $\mathcal{X}$ such that $L_+(f) = \int f(x)\, \mathrm{d}\mu_+(x)$ for all $f \in \mathcal{F}_+$.

Similarly, we have $L_-(f) = \int f(x)\, \mathrm{d}\mu_-(x)$ for some measure μ_- on $\mathcal{X}$. $\square$

0.8 Riesz' Representation Theorem

The Riesz' Representation Theorem 0.20 was developed in several stages. A first version for continuous functions on the unit interval $[0, 1]$ is due to F. Riesz [Rie09]. It was extended by Markov to some non-compact spaces [Mar38] and then by Kakutani to locally compact Hausdorff spaces [Kak41]. It is therefore sometimes also called the *Riesz–Markov–Kakutani Representation Theorem*.

However, we will see now that the general version already follows from the Signed Daniell's Representation Theorem 0.17 and Daniell's Representation Theorem 0.15 from 1918 [Dan18] combined with Urysohn's Lemma 0.2 from 1925 [Ury25], see also [Fed69, Sect. 2.5]. Urysohn's Lemma 0.2 is used to ensure that $C_c(\mathcal{X}, \mathbb{R})$ is large enough.

At first let us give the signed version.

Signed Riesz' Representation Theorem 0.18 (see e.g. [Fed69, Thm. 2.5.13])**.** *Let $\mathcal{X}$ be a locally compact Hausdorff space. If $L : C_c(\mathcal{X}, \mathbb{R}) \to \mathbb{R}$ is a linear functional such that*

$$\sup L(\{g \in C_c(\mathcal{X}, \mathbb{R}) \mid 0 \leq g \leq f\}) < \infty \tag{0.6}$$

for all $f \in C_c(\mathcal{X}, \mathbb{R})_+$ then there exist $C_c(\mathcal{X}, \mathbb{R})$ regular measures μ_+ and μ_- such that

$$L(f) = \int_{\mathcal{X}} f(x)\, \mathrm{d}\mu_+(x) - \int_{\mathcal{X}} f(x)\, \mathrm{d}\mu_-(x)$$

for all $f \in C_c(\mathcal{X}, \mathbb{R})$.

The following proof is taken from [Fed69, Thm. 2.5.13, pp. 106–107].

Proof. It is sufficient to verify condition (d) in the Signed Daniell's Representation Theorem 0.17.

Let $g, h_1, h_2, \ldots \in C_c(\mathcal{X}, \mathbb{R})_+$ be such that $h_n \nearrow g$ as $n \to \infty$. By Urysohn's Lemma 0.2 there exists a $f \in C_c(\mathcal{X}, \mathbb{R})_+$ such that $f(x) = 1$ for all $x \in \operatorname{supp} g$. Then

$$c := \sup\{|L(k)| \mid k \in C_c(\mathcal{X}, \mathbb{R}) \text{ and } 0 \leq k \leq f\} < \infty.$$

For each $\varepsilon > 0$ the intersection of all compact sets

$$S_n := \{x \in \mathcal{X} \mid g(x) \geq h_n(x) + \varepsilon\}$$

is empty. Since $S_{n+1} \subset S_n$ for all $n \in \mathbb{N}$ it follows that $S_n = \emptyset$ when n is sufficiently large. But $S_n = \emptyset$ implies $0 \leq g - h_n \leq \varepsilon f$ and $|L(g - h_n)| \leq \varepsilon c$ which proves condition (d). □

Corollary 0.19 (see e.g. [Fed69, §2.5.14])**.** *If in the Signed Riesz' Representation Theorem 0.18 we additionally have that the topology of $\mathcal{X}$ has a countable base then μ_+ and μ_- are Radon measures.*

Since positivity of L on $C_c(\mathcal{X}, \mathbb{R})_+$ implies (0.6) by

$$0 \leq g \leq f \quad \Rightarrow \quad 0 \leq f - g \quad \Rightarrow \quad 0 \leq L(f-g) \quad \Rightarrow \quad 0 \leq L(g) \leq L(f) < \infty$$

we have as an immediate consequence of the Signed Riesz' Representation Theorem 0.18 the non-negative version.

Riesz' Representation Theorem 0.20. *Let $\mathcal{X}$ be a locally compact Hausdorff space and $L : C_c(\mathcal{X}, \mathbb{R}) \to \mathbb{R}$ be a non-negative linear functional on $C_c(\mathcal{X}, \mathbb{R})_+$. Then there exists a measure μ on $\mathcal{X}$ such that*

$$L(f) = \int_{\mathcal{X}} f(x)\, \mathrm{d}\mu(x)$$

for all $f \in C_c(\mathcal{X}, \mathbb{R})$.

If additionally $\mathcal{X}$ as a topological space has a countable base then μ can be chosen to be a Radon measure.

From a topological point of view measures can also be introduced abstractly as linear functionals over certain spaces, see e.g. [Trè67, p. 216]. The Riesz representation theorem is then used to show the equivalence of the measure theoretic approach and the topological approach.

0.9* Riesz Decomposition

The results in this section about the Riesz decomposition will be used only in Theorem 2.13 (ii) about adapted cones and extensions of linear functionals on these. Theorem 2.13 is not used for the T-systems and can be omitted on first reading.

In Definition 0.12 we introduced lattices. Lattice spaces fulfill the following.

Riesz Decomposition Lemma 0.21 (see e.g. [Cho69, Lem. 10.5])**.** *Let $\mathcal{F}$ be a lattice space and $x, y_1, y_2 \geq 0$ with $x \leq y_1 + y_2$. Then there exist $x_1, x_2 \geq 0$ such that*

$$x = x_1 + x_2, \quad x_1 \leq y_1, \quad \textit{and} \quad x_2 \leq y_2$$

hold.

While the previous results holds for lattice spaces, also other spaces have this property.

Definition 0.22. Let F be an ordered vector space. We say F has the *Riesz decomposition property* if

$$x, y_1, y_2 \in F_+ : x \leq y_1 + y_2 \quad \Rightarrow \quad \exists x_1, x_2 \in F_+ : x = x_1 + x_2,\ x_1 \leq y_1,\ x_2 \leq y_2. \tag{0.7}$$

We have the following corollary.

Corollary 0.23 (see e.g. [Cho69, Cor. 10.6])**.** *Let F be an ordered vector space with the Riesz decomposition property, let $x_1, \dots, x_n \in F_+$, and let $y_1, \dots, y_m \in F_+$ with*

$$\sum_{i=1}^{n} x_i = \sum_{j=1}^{m} y_j.$$

Then for all $i = 1, \ldots, n$ *and* $j = 1, \ldots, m$ *there exist* $z_{i,j} \in F_+$ *such that*

$$x_i = \sum_{j=1}^{m} z_{i,j} \qquad \text{and} \qquad y_j = \sum_{i=1}^{n} z_{i,j}.$$

Part I
Introduction to Moments

Chapter 1
Moments and Moment Functionals

Extremes in nature equal ends produce;
In man they join to some mysterious use.

Alexander Pope: Essay on Man, Epistle II

In this chapter we deal with the basics of moments and moment functionals. More on moments and moment functionals can be found e.g. in [Sch17, Lau09, Mar08] and the classical literature [ST43, AK62, KN77].

1.1 Moments and Moment Functionals

Definition 1.1. Let $(\mathcal{X}, \mathfrak{A}, \mu)$ be a measure space and let $f : \mathcal{X} \to \mathbb{R}$ be a μ-integrable function. The real number

$$\int_{\mathcal{X}} f(x)\ \mathrm{d}\mu(x)$$

is called the *f-moment of μ*.

The name *moment* comes from the most famous example of moments: $\mathcal{X} = \mathbb{R}^3$ and $f(x,y,z) = f_\alpha(x,y,z) = x^{\alpha_1} \cdot y^{\alpha_2} \cdot z^{\alpha_3}$. Then

$$\int_{\mathbb{R}^3} (x^2 + y^2) \cdot \rho(x,y,z)\ \mathrm{d}x\ \mathrm{d}y\ \mathrm{d}z$$

is the z-rotational moment of a body with mass distribution ρ in $\mathbb{R}^3$.

In the modern theory of moments the investigation is about moment *functionals*.

Definition 1.2. Let $(\mathcal{X}, \mathfrak{A})$ be a measurable space and let $\mathcal{V}$ be a vector space of real-valued measurable functions on $(\mathcal{X}, \mathfrak{A})$. A linear functional $L : \mathcal{V} \to \mathbb{R}$ is called a *moment functional* if there exists a measure μ such that

$$L(f) = \int_{\mathcal{X}} f(x)\ \mathrm{d}\mu(x) \tag{1.1}$$

for all $f \in \mathcal{V}$. Any measure μ such that (1.1) holds is called a *representing measure of L*. We denote by $\mathcal{M}(L)$ the set of all representing measures of L.

Corollary 1.3. *Let $(\mathcal{X}, \mathfrak{A})$ be a measurable space, $\mathcal{V}$ be a space of measurable functions $f : \mathcal{X} \to \mathbb{R}$, and let $L : \mathcal{V} \to \mathbb{R}$ be a moment functional. Then $\mathcal{M}(L)$ is convex.*

Proof. See Problem 1.2. □

While a moment functional comes from a measure, conversely a measure μ gives a moment functional on μ-integrable functions.

Definition 1.4. Let $(\mathcal{X}, \mathfrak{A})$ be a measurable space and let $\mathcal{V}$ be a vector space of measurable functions on $(\mathcal{X}, \mathfrak{A})$. Given a measure μ such that all $f \in \mathcal{V}$ are μ-integrable then

$$L_\mu : \mathcal{V} \to \mathbb{R}, \quad f \mapsto L_\mu(f) := \int_{\mathcal{X}} f(x)\, \mathrm{d}\mu(x)$$

is the *moment functional generated by* μ.

We did not give any restrictions to the possible representing measures μ of a moment functional L. In practice and hence also in theory restrictions can and even must be made, e.g., $\operatorname{supp}\mu \subseteq K$ for some $K \in \mathfrak{A}$.

Definition 1.5. Let $(\mathcal{X}, \mathfrak{A})$ be a measurable space, $K \in \mathfrak{A}$ be a measurable set, let $\mathcal{V}$ be a vector space of measurable functions $f : \mathcal{X} \to \mathbb{R}$, and let $L : \mathcal{V} \to \mathbb{R}$ be a linear functional. We call L to be a *K-moment functional* if there exists a measure μ on $\mathcal{X}$ such that

$$L(f) = \int_{\mathcal{X}} f(x)\, \mathrm{d}\mu(x)$$

for all $f \in \mathcal{V}$ and $\operatorname{supp}\mu \subseteq K$.

A linear functional $L : \mathcal{V} \to \mathbb{R}$ can also be described by the numbers $s_i := L(f_i)$ for a basis $\{f_i\}_{i\in I}$ of $\mathcal{V}$.

Definition 1.6. Let $(\mathcal{X}, \mathfrak{A})$ be a measurable space, let $\mathcal{V}$ be a space of measurable functions $f : \mathcal{X} \to \mathbb{R}$ with basis $\{f_i\}_{i\in I}$ for some index set I. Given any real sequence $s = (s_i)_{i\in I}$ the linear functional $L_s : \mathcal{V} \to \mathbb{R}$ defined by

$$L_s(f_i) := s_i$$

for all $i \in I$ is called the *Riesz functional of s*. The sequence s is called a *moment sequence* if $L_s : \mathcal{V} \to \mathbb{R}$ is a moment functional.

Example 1.7. Let $n \in \mathbb{N}$, $\mathcal{X} = \mathbb{R}^n$ with $\mathfrak{A} = \mathfrak{B}(\mathbb{R}^n)$ the Borel σ-algebra, and let $\mathcal{V} = \mathbb{R}[x_1, \dots, x_n]$ be the ring of polynomials. Then a real sequence $s = (s_\alpha)_{\alpha\in\mathbb{N}_0^n}$ gives a linear functional $L_s : \mathbb{R}[x_1, \dots, x_n] \to \mathbb{R}$ by $L_s(x^\alpha) := s_\alpha$ for all $\alpha \in \mathbb{N}_0^n$. The matrix $\mathcal{H}(s) = (s_{\alpha+\beta})_{\alpha,\beta\in\mathbb{N}_0^n}$ is the *Hankel matrix* of the sequence s (resp. the linear functional L_s). ∘

In practice and hence also in theory we have the special case that $\mathcal{V}$ is finite dimensional.

Definition 1.8. Let $(\mathcal{X}, \mathfrak{A})$ be a measurable space, let $\mathcal{V}$ be a vector space of measurable functions $f : \mathcal{X} \to \mathbb{R}$, and $L : \mathcal{V} \to \mathbb{R}$ be a moment functional. Then L is called a *truncated* moment functional if $\mathcal{V}$ is finite dimensional.

1.2 Determinacy and Indeterminacy

We introduced the set of all representing measures $\mathcal{M}(L)$ of a moment functional in Definition 1.2. We have the special and important case when $\mathcal{M}(L)$ is a singleton, i.e., the moment functional L has a unique representing measure.

Definition 1.9. Let $(\mathcal{X}, \mathfrak{A})$ be a measurable space, $\mathcal{V}$ a real vector space of measurable functions $f : \mathcal{X} \to \mathbb{R}$, and let $L : \mathcal{V} \to \mathbb{R}$ be a moment functional. If $\mathcal{M}(L)$ is a singleton, i.e., L has a unique representing measure, then L is called *determinate*. Otherwise it is call *indeterminate*.

Corollary 1.10. *Let $(\mathcal{X}, \mathfrak{A})$ be a measurable space, $\mathcal{V}$ a real vector space of measurable functions $f : \mathcal{X} \to \mathbb{R}$, and let $L : \mathcal{V} \to \mathbb{R}$ be an indeterminate moment functional. Then L has infinitely many representing measures.*

Proof. See Problem 1.3. □

The first example of an indeterminate moment functional/sequence was given by T. J. Stieltjes [Sti94]. In [Sti94, p. J.105, §55] he states that all

$$s_k = \int_0^\infty x^k \cdot \left(1 + c \cdot \sin(\sqrt[4]{x})\right) \cdot e^{-\sqrt[4]{x}}\, \mathrm{d}x \qquad (k \in \mathbb{N}_0)$$

are independent on $c \in [-1, 1]$.

The first explicit example then follows in [Sti94, pp. J.106–J.107, §56].

Example 1.11 (see [Sti94, pp. J.106–J.107, §56])**.** Let $c \in [-1, 1]$ and

$$f(x) = \frac{1}{\sqrt{\pi}} \cdot \exp\left(-\frac{1}{2}(\ln x)^2\right)$$

for all $x \in [0, \infty)$. Then the measure $\mu_c \in \mathcal{M}(\mathbb{R})$ defined by

$$\mathrm{d}\mu_c(x) := [1 + c \cdot \sin(2\pi \ln x)] \cdot f(x)\, \mathrm{d}x$$

has the moments

$$s_k = \int_0^\infty x^k\, \mathrm{d}\mu_c(x) = e^{\frac{1}{4}(k+1)^2}$$

for all $k \in \mathbb{N}_0$, i.e., independent on $c \in [-1, 1]$. ∘

Criteria for determinacy and indeterminacy are well-studied, see e.g. [Sch17] and reference therein.

Problems

1.1 Let $n \in \mathbb{N}$ and let $L : \mathbb{R}[x_1, \ldots, x_n] \to \mathbb{R}$ be a moment functional with a representing measure μ such that $\operatorname{supp} \mu \subseteq K$ for some compact $K \subset \mathbb{R}^n$. Show that L is determinate, i.e., show that μ is the unique representing measure of L.

Hint: Use the Stone–Weierstrass Theorem 0.3.

1.2 Prove Corollary 1.3.

1.3 Prove Corollary 1.10.

Chapter 2
Choquet's Theory and Adapted Spaces

Progress imposes not only new possibilities for the future but new restrictions.

Norbert Wiener [Wie88, p. 46]

This chapter is devoted to the theory of Choquet and the concept of adapted spaces. The results can also be found in e.g. [Cho69, Phe01, Sch17].

2.1 Extensions of Linear Functionals preserving Positivity

We remind the reader that a convex cone $C \subseteq F$ in a real vector space F induces an order $\leq$ on F, i.e., for any $x, y \in F$ we have $x \leq y$ iff $y - x \in C$, see Section 0.1.

Lemma 2.1 (see e.g. [Cho69, Prop. 34.1])**.** *Let F be a real vector space, $E \subseteq F$ be a linear subspace, and let $C \subseteq F$ be a convex cone which induces the order $\leq$ on F. Then the following are equivalent:*

(i) $F + C$ *is a vector space.*
(ii) $F + C = F - C$.
(iii) *Any* $x \in (F + C) \cup (F - C)$ *is majorized by some* $z \in F$*, i.e.,* $x \leq z$*, and is minorized by some* $y \in F$*, i.e.,* $y \leq x$.

Proof. See Problem 2.1. □

Definition 2.2. Let F be a real vector space and $C \subseteq F$ be a convex cone. A linear functional $L : F \to \mathbb{R}$ is called *C-positive* if $L(f) \geq 0$ holds for all $f \in C$. L is called *strictly* C-positive if $L(f) > 0$ holds for all $f \in C \setminus \{0\}$.

Theorem 2.3 (see e.g. [Cho69, Thm. 34.2])**.** *Let F be a real vector space, $E \subseteq F$ be a linear subspace, and $C \subseteq F$ be a convex cone with $F = E + C$. Then any $(C \cap E)$-positive linear functional $L : E \to \mathbb{R}$ can be extended to a C-positive linear functional $\tilde{L} : F \to \mathbb{R}$.*

The extension $\tilde{L}$ is unique if and only if for all $x \in E$ we have

$$\sup\{L(y) \mid y \leq x,\ y \in F\} = \inf\{L(y) \mid x \leq y,\ y \in F\}. \tag{2.1}$$

The proof is taken from [Cho69, vol. 2, p. 270–271]. It adapts the idea behind the proof of the Hahn–Banach Theorem 0.1.

Proof. Let $\mathcal{H} := \{(H, h)\}_{H \text{ subspace: } E \subseteq H \subseteq F}$ where $h : H \to \mathbb{R}$ extends L. The family $\mathcal{H}$ has a natural order by the extension property, i.e., we have $(H_1, h_1) \leq (H_2, h_2)$ if $h_2 : H_2 \to \mathbb{R}$ is an extension of $h_1 : H_1 \to \mathbb{R}$. By Zorn's Lemma $\mathcal{H}$ has a maximal element (G, g). We have to show $G = F$. For that it is sufficient that E is a hyperplane in F and L can be extended to F.

Let $x_0 \in F \setminus E$. By Lemma 2.1 (iii) there exist $y, z \in E$ with $y \leq x_0 \leq z$. We define

$$\alpha := \sup\{L(y) \,|\, y \leq x_0 \text{ and } y \in E\}$$

and

$$\beta := \inf\{L(z) \,|\, x_0 \leq z \text{ and } z \in E\}.$$

Since L is C-positive we have $\alpha \leq \beta$ and any extension $\tilde{L}$ must satisfy $\alpha \leq \tilde{L}(x_0) \leq \beta$.

We show that for each $\gamma \in [\alpha, \beta]$ there exists an extension $\tilde{L}$ with $\tilde{L}(x_0) = \gamma$. Each point $u \in F$ can be uniquely written as $u = y - \lambda x_0$ with $y \in E$ and $\lambda \in \mathbb{R}$. Define $\tilde{L}(u) := L(y) - \lambda\gamma$. Then $\tilde{L}$ is a linear extension of L and we have to show that $\tilde{L}$ is C-positive. Let $u \in C$, i.e., $y \geq \lambda x_0$. If $\lambda > 0$ then $x_0 \leq y/\lambda$ and $\beta \leq L(y/\lambda)$. Hence, $L(y) \geq \lambda\beta \geq \lambda\gamma$ and so $\tilde{L}(u) \geq 0$. If on the other hand $\lambda < 0$ then $x_0 \geq y/\lambda$ and $\alpha \geq L(y/\lambda)$ which implies $L(y) \geq \lambda\alpha \geq \lambda\gamma$ and $\tilde{L}(u) \geq 0$. At last, if $\lambda = 0$ then $\tilde{L}(u) = \tilde{L}(y) \geq 0$. In summary, we proved that $\tilde{L}$ is C-positive.

For the uniqueness it is sufficient to note that if (2.1) holds for all $x \in E$ then $\tilde{L}$ is uniquely determined since every extension $\tilde{L}$ arises from this construction. If on the other hand $\alpha < \beta$, i.e., (2.1) does not hold, then some extension $(H, h) \in \mathcal{H}$ is not unique for H and consequently $\tilde{L}$ is not a unique extension of L. □

From the previous proof we see that by redoing the proof of the Hahn–Banach Theorem the uniqueness criteria (2.1) can be incorporated. A second proof using the Hahn–Banach Theorem is much shorter but loses the uniqueness condition (2.1), see e.g. [Sch17, Prop. 1.7].

A third proof of Theorem 2.3 follows from the following lemma.

Lemma 2.4 (see e.g. [Cho69, Prop. 34.3])**.** *Let E be a real vector space, let $g : E \to \mathbb{R}$ be superlinear and let $h : E \to \mathbb{R}$ be sublinear. Then there exists a linear map $f : E \to \mathbb{R}$ such that $g \leq f \leq h$.*

Proof. Equip E with the topology of all semi-norms. Then $p(x) := \sup\{h(x), h(-x)\}$ is a semi-norm and $h \leq p$. Since p is continuous and h is convex we have that h is continuous. Thus g and h can be separated by a closed hyperplane. □

Lemma 2.4 not only gives a third proof of Theorem 2.3 but also has a generalization which is known as *Strassen's Theorem* [Str65].

Strassen's Theorem states that if $(\mathcal{Y}, \mu)$ is a measure space, $\{h_y : E \to \mathbb{R}\}_{y \in \mathcal{Y}}$ is a family of sublinear maps, and let $l : E \to \mathbb{R}$ be a linear map with

$$l \leq \int_{\mathcal{Y}} h_y \,\mathrm{d}\mu(y).$$

Then there exists a family $\{l_y : E \to \mathbb{R}\}_{y\in\mathcal{Y}}$ of linear maps l_y with $l_y \leq h_y$ such that

$$l = \int_{\mathcal{Y}} l_y \, \mathrm{d}\mu(y).$$

For more on Strassen's Theorem see e.g. [Edw78, Ska93, Lin99] and references therein.

2.2 Adapted Spaces of Continuous Functions

We now come to the adapted spaces. To define them we need the following.

Definition 2.5. Let $\mathcal{X}$ be a locally compact Hausdorff space and $f, g \in C(\mathcal{X}, \mathbb{R})_+$. We say f *dominates* g if for any $\varepsilon > 0$ there is an $h_\varepsilon \in C_c(\mathcal{X}, \mathbb{R})$ such that $g \leq \varepsilon f + h_\varepsilon$.

Equivalent expressions are the following.

Lemma 2.6 (see e.g. [Sch17, Lem. 1.4])**.** *Let $\mathcal{X}$ be a locally compact Hausdorff space and let $f, g \in C(\mathcal{X}, \mathbb{R})_+$. Then the following are equivalent:*

(i) f dominates g.
(ii) For every $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subseteq \mathcal{X}$ such that $g(x) \leq \varepsilon \cdot f(x)$ holds for all $x \in \mathcal{X} \setminus K_\varepsilon$.
(iii) For every $\varepsilon > 0$ there exists an $\eta_\varepsilon \in C_c(\mathcal{X}, \mathbb{R})$ with $0 \leq \eta_\varepsilon \leq 1$ such that $g \leq \varepsilon \cdot f + \eta_\varepsilon \cdot g$.

Proof. See Problem 2.2. □

The main definition of this chapter is the following.

Definition 2.7. Let $\mathcal{X}$ be a locally compact Hausdorff space and let $E \subseteq C(\mathcal{X}, \mathbb{R})$ be a vector space. Then E is called an *adapted space* if the following conditions hold:

(i) $E = E_+ - E_+$,
(ii) for all $x \in \mathcal{X}$ there is a $f \in E_+$ such that $f(x) > 0$, and
(iii) every $g \in E_+$ is dominated by some $f \in E_+$.

The space $C_c(\mathcal{X}, \mathbb{R})_+$ is of special interest because of the Riesz' Representation Theorem 0.20. The following result shows that any $g \in C_c(\mathcal{X}, \mathbb{R})_+$ is dominated (and even bounded) by some $f \in E_+$ for any given adapted space $E \subseteq C(\mathcal{X}, \mathbb{R})$.

Lemma 2.8. *Let $\mathcal{X}$ be a locally compact Hausdorff space, $g \in C_c(\mathcal{X}, \mathbb{R})_+$, and let $E \subseteq C(\mathcal{X}, \mathbb{R})$ be an adapted space. Then there exists a $f \in E_+$ such that $f > g$.*

Proof. See Problem 2.6. □

2.3 Existence of Integral Representations

One important reason adapted spaces have been introduced is to get the following representation theorem. It is a general version of Haviland's Theorem 3.4 and will be used to solve most moment problems in an efficient way.

Basic Representation Theorem 2.9 (see e.g. [Cho69, Thm. 34.6])**.** *Let $\mathcal{X}$ be a locally compact Hausdorff space, $E \subseteq C(\mathcal{X}, \mathbb{R})$ be an adapted subspace, and let $L : E \to \mathbb{R}$ be a linear functional. The following are equivalent:*

(i) The functional L is E_+-positive.
(ii) L is a moment functional, i.e., there exists a (Radon) measure μ on $\mathcal{X}$ such that

(a) all $f \in E$ are μ-integrable and
(b) $L(f) = \int_{\mathcal{X}} f(x)\, \mathrm{d}\mu(x)$ holds for all $f \in E$.

The following proof is adapted from [Cho69, vol. 2, p. 276–277].

Proof. The direction (ii) $\Rightarrow$ (i) is clear. It is therefore sufficient to prove (i) $\Rightarrow$ (ii).

Define

$$F := \{f \in C(\mathcal{X}, \mathbb{R}) \mid |f| \le g \text{ for some } g \in E_+\}. \tag{2.2}$$

Then F_+ is a convex cone. We have $F = E + F_+$. To see this let $f \in F$ and write $f = -g + (f + g)$ where $|f| \le g$ for some $g \in E_+$, i.e., $f \in E + F_+$ and hence $F \subseteq E + F_+$. The inclusion $E + F_+ \subseteq F$ is clear and we therefore have $F = E + F_+$.

By Theorem 2.3 we can extend L to a F_+-positive linear functional $\tilde{L} : F \to \mathbb{R}$. By Lemma 2.8 we have $C_c(\mathcal{X}, \mathbb{R}) \subseteq F$ and hence by the Riesz' Representation Theorem 0.20 there exists a representing Radon measure μ on $\mathcal{X}$ of $\tilde{L}|_{C_c(\mathcal{X},\mathbb{R})}$.

We need to show that μ is also a representing measure of L. Let $f \in E_+$. Since μ is Radon we have

$$\int_{\mathcal{X}} f(x)\, \mathrm{d}\mu(x) = \sup \left\{ \int_{\mathcal{X}} \varphi(x)\, \mathrm{d}\mu(x) \,\middle|\, \varphi \in C_c(\mathcal{X}, \mathbb{R}),\ \varphi \le f \right\} \le \tilde{L}(f) = L(f) \tag{2.3}$$

and hence f is μ-integrable. Since $E = E_+ - E_+$ we have that all $f \in E$ are μ-integrable.

Then

$$K(f) := \tilde{L}(f) - \int_{\mathcal{X}} f(x)\, \mathrm{d}\mu(x) \tag{2.4}$$

for all $f \in F$ defines a F_+-positive linear functional on F which vanishes on $C_c(\mathcal{X}, \mathbb{R})$. For every $g \in E_+$ there is an $f \in E_+$ dominating g. Let $\varepsilon > 0$ and $h_\varepsilon \in C_c(\mathcal{X}, \mathbb{R})$ be such that $g \le \varepsilon f + h_\varepsilon$. Then $0 \le K(g) \le \varepsilon \cdot K(f) \xrightarrow{\varepsilon \to 0} 0$, i.e., $K = 0$ on E_+ and hence on E which proves that μ is a representing measure of L. □

We actually proved that L can be extended to $\tilde{L}$ on F in (2.2) and that μ is a representing measure for $\tilde{L}$. This is included in (ii-b).

For the uniqueness of the representing measure μ of L we have the following.

Corollary 2.10 (see e.g. [Cho69, Cor. 34.7]). *Let $\mathcal{X}$ be a locally compact Hausdorff space, $E \subseteq C(\mathcal{X}, \mathbb{R})$ be an adapted space, and let $L : E \to \mathbb{R}$ be a E_+-positive linear functional. Then the following are equivalent:*

(i) The representing measure μ of L from the Basic Representation Theorem 2.9 is unique.
(ii) For any $f \in C_c(\mathcal{X}, \mathbb{R})$ and $\varepsilon > 0$ there are $f_1, f_2 \in E$ with $f_1 \leq f \leq f_2$ and $0 \leq T(f_2 - f_1) \leq \varepsilon$.

Proof. Reformulating (i) we get that the measure μ must be uniquely defined by the extension of $L : E \to \mathbb{R}$ to $\tilde{L} : E + C_c(\mathcal{X}, \mathbb{R}) \to \mathbb{R}$. By Theorem 2.3 eq. (2.1) this is equivalent to

$$\sup\{L(\varphi) \mid \varphi \leq f,\ \varphi \in E\} = \inf\{L(\varphi) \mid f \leq \varphi,\ \varphi \in E\}.$$

But this is equivalent to our condition (ii), i.e., we showed (i) $\Leftrightarrow$ (ii). □

2.4* Adapted Cones

A generalization of adapted spaces is to go to adapted cones, i.e., dropping the vector space property. This is presented in [Cho69] but not included in [Sch17] and we want to show it to the reader for the sake (or at least a glimpse) of completeness. It is not used in the T-systems and can be omitted on first reading.

Definition 2.11. Let F be an ordered vector space and let $E \subseteq F$ be a convex cone. For $x, y \in F$ with $x, y \geq 0$ we say that y *dominates* x (*relative to* E) if for any $\varepsilon > 0$ there exists a $z_\varepsilon \in E$ such that $x \leq \varepsilon y + z_\varepsilon$.

For two convex cones $C, E \subseteq F_+$ we say that (C, E) are *adapted* (*cones*) if every $x \in C$ is dominated by some $x' \in C$ (relative to E) and for each $g \in E$ there is an $f \in C$ so that $g \leq f$.

The previous definition is a generalization of Definition 2.5. The convex cone C has the role of $C_c(\mathcal{X}, \mathbb{R})_+$, F has the role of $C(\mathcal{X}, \mathbb{R})$, and E is the adapted space.

Lemma 2.12 (see e.g. [Cho69, Prop. 35.3]). *Let F be an ordered vector space, let (C, E) be adapted cones, and let $L : E \to \mathbb{R}$ be a positive linear functional. Then*

$$L|_E = 0 \quad \Rightarrow \quad L|_C = 0.$$

Proof. Let $x \in C$. Since (C, E) are adapted cones there exists a $x' \in C$ such that for any $\varepsilon > 0$ there is a $z_\varepsilon \in E$ with

$$0 \leq x \leq \varepsilon x' + z_\varepsilon.$$

Since $L \geq 0$ on E we have

$$0 \leq L(x) \leq \varepsilon L(x') \xrightarrow{\varepsilon \to 0} 0$$

which proves $L|_C = 0$. □

Theorem 2.13 (see e.g. [Cho69, Thm. 35.4])**.** *Let F be an ordered vector space.*

(i) Let $C \subseteq F_+$ be a convex cone and let $L : C \to [0,\infty)$ be a positive linear functional. Define

$$\hat{C} := \{g \in F_+ \,|\, g \leq f \text{ for some } x \in C\}.$$

Then L has an extension to a positive linear functional $\hat{L} : \hat{C} \to [0,\infty)$.

(ii) Let (C, E) be adapted cones such that $E \subseteq \hat{C}$ and $\hat{C}$ has the Riesz decomposition property (0.7). Then for each $f \in \hat{C}$ we have

$$\hat{L}(f) = \sup\left\{\hat{L}(g) \,\middle|\, g \in E \text{ with } g \leq f\right\}.$$

Proof. (i): First, extend L by linearity to the vector space $C - C$. Let $F_0 := \hat{C} - \hat{C}$. Then $F_0 = C - C + \hat{C} = -C + \hat{C}$. By Theorem 2.3 L extends to a $\hat{C}$-positive linear functional on F_0.

(ii): Define $L_0 : \hat{C} \to \mathbb{R}$ by

$$L_0(f) := \sup\{L(g) \,|\, g \in E \text{ with } g \leq f\}.$$

Hence, $0 \leq L_0(f) \leq \hat{L}(f)$ for all $f \in \hat{C}$. Clearly, $L_0(\lambda f) = \lambda L_0(f)$ holds for all $\lambda \geq 0$ and $f \in \hat{C}$. Additionally,

$$L_0(f_1 + f_2) = \sup\left\{\hat{L}(g) \,\middle|\, g \in E,\ g \leq f_1 + f_2\right\}$$

which is by the Riesz decomposition property (0.7)

$$\begin{aligned} &= \sup\left\{\hat{L}(g_1 + g_2) \,\middle|\, g_1, g_2 \in E,\ g_1 \leq f_1,\ g_2 \leq f_2\right\} \\ &= L_0(f_1) + L_0(f_2) \end{aligned}$$

for all $f_1, f_2 \in \hat{C}$ and hence by linearity extension L_0 is linear on F_0.

We now show at last that $L - L_0 = 0$ on $\hat{C}$. Since (C, E) are adapted cones we have that $(\hat{C}, E)$ are adapted cones. We have $L(f) - L_0(f) = 0$ for all $f \in E$ and hence by Lemma 2.12 we have $L = L_0$ on $\hat{C}$ which proves (ii). □

Theorem 2.13 (ii) is the analogue of extending a Radon measure on $C_c(\mathcal{X}, \mathbb{R})$ to continuous integrable functions.

Example 2.14 (see e.g. [Cho69, Exm. 35.5])**.** Let $(\mathcal{X}, \mathfrak{A}, \mu)$ be a measure space. Let $C = (\mathcal{L}^1(\mathcal{X}, \mu))_+$ and $E = \mathcal{L}^\infty(\mathcal{X}, \mu) \cap (\mathcal{L}^1(\mathcal{X}, \mu))_+$. Then (C, E) are adapted cones. Hence, every positive linear functional is uniquely determined by its values on $\mathcal{L}^\infty \cap \mathcal{L}^1$. ∘

2.5* Continuity of Positive Linear Functionals

At the end of this chapter we want to point out some continuity results. But we will leave out the proofs since these results will not be used for our T-system treatment.

Theorem 2.15 (see e.g. [Cho69, Thm. 36.1])**.** *Let E be an ordered Hausdorff topological vector space such that $E = E_+ - E_+$ and let either*

(i) $\operatorname{int} E_+ \neq \emptyset$

or

(ii) E is complete, metrizable, and E_+ is closed.

Then any positive linear functional $L : E \to \mathbb{R}$ is continuous.

The previous results holds for general convex pointed cones in E.

Corollary 2.16 (see e.g. [Cho69, Cor. 36.1])**.** *Let E be a Hausdorff topological vector space and $P \subset E$ be a convex pointed cone. The following hold:*

(i) If $\operatorname{int} P \neq \emptyset$ *then any linear P-positive functional $T : E \to \mathbb{R}$ is continuous.*
(ii) If E is complete, metrizable, P is closed, and $E = P - P$, then any linear P-positive functional $T : E \to \mathbb{R}$ is continuous.

Further conditions for continuity can be found e.g. in [Cho69, Ch. 36] or [SW99]. [Cho69, Ch. 36] also gives results for positive linear functionals on C*-algebras, the Schwartz space $\mathcal{S}(\mathbb{R}^n, \mathbb{R})$, Lipschitz functions, and on general vector lattices.

Another direction is more operator theoretic and deals with linear functionals over algebras. An *algebra* $\mathcal{A}$ is a (complex) vector space with a multiplication $\cdot : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$, $(a, b) \mapsto ab$ such that

(i) $a(bc) = (ab)c$,
(ii) $(a + b)c = ac + bc$, and
(iii) $\alpha(ab) = (\alpha a)b = a(\alpha b)$

for all $a, b, c \in \mathcal{A}$ and $\alpha \in \mathbb{C}$. An element $1 \in \mathcal{A}$ is called *unit element* if $1a = a = a1$ for all $a \in \mathcal{A}$. A $*$-*algebra* is an algebra with an involution $^* : \mathcal{A} \to \mathcal{A}$, $a \mapsto a^*$ that also satisfies $(ab)^* = b^* a^*$ and $(\alpha a)^* = \overline{\alpha} a^*$. A *linear functional* $L : \mathcal{A} \to \mathbb{C}$ is called *non-negative* if $L(a^* a) \geq 0$ for all $a \in \mathcal{A}$. A *topological $*$-algebra* is a $*$-algebra with a topology $\mathcal{T}$ such that the multiplication and involution are continuous. A *Fréchet topological $*$-algebra* is a topological algebra which is a Fréchet space, i.e., a complete metrizable locally convex space. An example is $\mathbb{C}[x_1, \dots, x_n]$.

We have the following.

Theorem 2.17 ([Xia59] and [NW72]; or e.g. [Sch90, Thm. 3.6.1])**.** *Let $\mathcal{A}$ be a Fréchet topological $*$-algebra with unit element and let $L : \mathcal{A} \to \mathbb{C}$ be a linear functional. If L is non-negative then it is continuous.*

A more general statement is [NW72, Thm. 1]. For more see e.g. [Sch90, Ch. 3.6] and references therein.

Problems

2.1 Prove Lemma 2.1.

2.2 Prove Lemma 2.6.

2.3 Let $\mathcal{X}$ be a compact topological Hausdorff space and let $E \subseteq C(\mathcal{X}, \mathbb{R})$ be a subspace such that there exists an $e \in E$ such that $e(x) > 0$ for all $x \in \mathcal{X}$. Show that E is an adapted space.

2.4 Let $n \in \mathbb{N}$ and $\mathcal{X} \subseteq \mathbb{R}^n$ be closed. Show that $\mathbb{R}[x_1, \ldots, x_n]$ on $\mathcal{X}$ is an adapted space.

2.5 Let $n \in \mathbb{N}$, $\mathcal{X} \subseteq \mathbb{R}^n$ be closed, and let $E \subseteq \mathbb{R}[x_1, \ldots, x_n]$ be an adapted space. Show that if E is finite dimensional then $\mathcal{X}$ is compact.

2.6 Prove Lemma 2.8.

Chapter 3
The Classical Moment Problems

Those who cannot remember the past are condemned to repeat it.

George Santayana [San05]

In this chapter we give several classical solutions of moment problems: the Stieltjes, Hamburger, and Hausdorff moment problem. Additionally, we collect other classical results such as Haviland's Theorem, Richter's Theorem on the existence of finitely atomic representing measures for truncated moment functionals, and Boas' Theorem on the existence of signed representing measures for any linear functional $L : \mathbb{R}[x_1, \ldots, x_n] \to \mathbb{R}$.

3.1 Classical Results

In this section we give a chronological list of the early moment problems which have been solved. We will explicitly discuss the historical (first) proofs of these results. Our modern proofs here will be based on the Choquet's theory from Chapter 2 and for a modern operator theoretic approach see e.g. [Sch17].

The first moment problem was solved by T. J. Stieltjes [Sti94]. He was the first who fully stated the moment problem, solved the first one, and by doing that also introduced the integral theory named after him: the Stieltjes integral.

Stieltjes' Theorem 3.1. *Let $s = (s_i)_{i\in\mathbb{N}_0}$ be a real sequence. The following are equivalent:*

(i) *s is a $[0,\infty)$-moment sequence (Stieltjes moment sequence).*
(ii) $L_s(p) \geq 0$ *for all* $p \in \operatorname{Pos}([0,\infty))$.
(iii) $L_s(p^2) \geq 0$ *and* $L_{Xs}(p^2) = L_s(x\cdot p^2) \geq 0$ *for all* $p \in \mathbb{R}[x]$.
(iv) *s and $Xs = (s_{i+1})_{i\in\mathbb{N}_0}$ are positive semidefinite.*
(v) $\mathcal{H}(s) \geq 0$ *and* $\mathcal{H}(Xs) \geq 0$ *for all* $d \in \mathbb{N}_0$.

Proof. See Problem 3.1. □

In the original proof of Stieltjes' Theorem 3.1 Stieltjes [Sti94] does not use non-negative polynomials. Instead he uses continued fractions and introduces new sequences which we (nowadays) denote by s and Xs.

Stieltjes only proves (i) ⇔ (iv). The implication (i) ⇔ (ii) is Haviland's Theorem 3.4, (ii) ⇔ (iii) is the description of $\mathrm{Pos}([0,\infty))$, and (iv) ⇔ (v) is a reformulation of s and Xs being positive semi-definite.

The next moment problem was solved by H. L. Hamburger [Ham20, Satz X and Existenztheorem (§8, p. 289)].

Hamburger's Theorem 3.2. *Let $s = (s_i)_{i\in\mathbb{N}_0}$ be a real sequence. The following are equivalent:*

(i) s is a $\mathbb{R}$-moment sequence (Hamburger moment sequence or short moment sequence).
(ii) $L_s(p) \geq 0$ for all $p \in \mathrm{Pos}(\mathbb{R})$.
(iii) $L_s(p^2) \geq 0$ for all $p \in \mathbb{R}[x]$.
(iv) s is positive semidefinite.
(v) $\mathcal{H}(s) \succeq 0$.

Proof. See Problem 3.2. □

Hamburger proved similar to Stieltjes the equivalence (i) ⇔ (iv) via continued fractions. In [Ham20, Satz XIII] Hamburger solves the full moment problem by approximation with truncated moment problems. This was later in a slightly more general framework proved in [Sto01], see also Section 3.5. Hamburger needed to assume that the sequence of measures μ_k (which he called "Belegungen" and denoted by $\mathrm{d}\Phi^{(k)}(u)$) to converge to some measure μ (condition 2 of [Ham20, Satz XIII]). Hamburgers additional condition 2 is nowadays replaced by the vague convergence and the fact that the solution set of representing measures is vaguely compact [Sch17, Thm. 1.19], i.e., it assures the existence of a μ as required by Hamburger in the additional condition 2.

Shortly after Hamburger the moment problem on $[0,1]$ was solved by F. Hausdorff [Hau21a, Satz II and III].

Hausdorff's Theorem 3.3. *Let $s = (s_i)_{i\in\mathbb{N}_0}$ be a real sequence. The following are equivalent:*

(i) s is a $[0,1]$-moment sequence (Hausdorff moment sequence).
(ii) $L_s(p) \geq 0$ for all $p \in \mathrm{Pos}([0,1])$.
(iii) $L_s(p^2) \geq 0$, $L_{Xs}(p^2) \geq 0$, and $L_{(1-X)s}(p^2) \geq 0$ for all $p \in \mathbb{R}[x]$.
(iv) s, Xs, and $(1-X)s$ are positive semidefinite.
(v) $\mathcal{H}(s) \succeq 0$, $\mathcal{H}(Xs) \succeq 0$, and $\mathcal{H}((1-X)s) \succeq 0$.

Proof. See Problem 3.3. □

Hausdorff proved the equivalence (i) ⇔ (iii) via so called C-sequences. In [Toe11] Toeplitz treats general linear averaging methods. In [Hau21a] Hausdorff uses these. Let the infinite dimensional matrix $\lambda = (\lambda_{i,j})_{i,j\in\mathbb{N}_0}$ be row-finite, i.e., for every row i only finitely many $\lambda_{i,j}$ are non-zero. Then the averaging method

$$A_i = \sum_{j\in\mathbb{N}_0} \lambda_{i,j} a_j$$

shall be consistent: If $a_j \to \alpha$ converges then $A_i \to \alpha$ converges to the same limit. Toeplitz proved a necessary and sufficient condition on λ for this property. Hausdorff uses only part of this property. He calls a matrix $(\lambda_{i,j})_{i,j\in\mathbb{N}_0}$ with the property that a convergent sequence $(a_j)_{j\in\mathbb{N}_0}$ is mapped to a convergent sequence $(A_j)_{j\in\mathbb{N}_0}$ (the limit does not need to be preserved) a C-matrix (convergence preserving matrix). Hausdorff gives the characterization of C-matrices [Hau21a, p. 75, conditions (A) – (C)]. Additionally, if λ is a C-matrix and a diagonal matrix with diagonal entries $\lambda_{i,i} = s_i$ then $s = (s_i)_{i\in\mathbb{N}_0}$ is called a C-sequence. The equivalence (i) $\Leftrightarrow$ (iii) is then shown by Hausdorff in the result that a sequence is a $[0, 1]$-moment sequence if and only if it is a C-sequence [Hau21a, p. 102].

A much simpler approach to solve the K-moment problem for any closed $K \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, was presented by E. K. Haviland in [Hav36, Theorem], see also [Hav35, Theorem] for the earlier case $K = \mathbb{R}^n$. He no longer used continued fractions but employed the Riesz' Representation Theorem 0.20, i.e., representing a linear functional by integration, and connected the existence of a representing measure to the non-negativity of the linear functional on

$$\operatorname{Pos}(K) := \{f \in \mathbb{R}[x_1, \dots, x_n] \mid f \geq 0 \text{ on } K\}. \tag{3.1}$$

Haviland's Theorem 3.4. *Let $n \in \mathbb{N}$, $K \subseteq \mathbb{R}^n$ be closed, and $s = (s_\alpha)_{\alpha\in\mathbb{N}_0^n}$ be a real sequence. The following are equivalent:*

(i) s is a K-moment sequence.
(ii) $L_s(p) \geq 0$ for all $p \in \operatorname{Pos}(K)$.

Proof. See Problem 3.4. □

As noted before, in [Hav35, Theorem] Haviland proves "only" the case $K = \mathbb{R}^n$ with the extension method by M. Riesz. In [Hav36, Theorem] this is extended to any closed $K \subseteq \mathbb{R}^n$. The idea to do so is attributed by Haviland to A. Wintner [Hav36, p. 164]:

> A. Wintner has subsequently suggested that it should be possible to extend this result [[Hav35, Theorem]] by requiring that the distribution function [measure] solving the problem have a spectrum [support] contained in a preassigned set, a result which would show the well-known criteria for the various standard special momentum problems (Stieltjes, Herglotz [trigonometric], Hamburger, Hausdorff in one or more dimensions) to be put particular cases of the general n-dimensional momentum problem mentioned above. The purpose of this note [[Hav36]] is to carry out this extension.

In [Hav36] after the general Theorem 3.4 Haviland then goes through all the classical results (Theorems 3.1 to 3.3, and the Herglotz (trigonometric) moment problem on the unit circle $\mathbb{T}$ which we did not included here) and shows how all these results (i.e., conditions on the sequences) are recovered from the at this point known representations of non-negative polynomials.

For the *Hamburger moment problem* (Hamburger's Theorem 3.2) Haviland uses

$$\operatorname{Pos}(\mathbb{R}) = \left\{f^2 + g^2 \,\middle|\, f, g \in \mathbb{R}[x]\right\} \tag{3.2}$$

which was already known to D. Hilbert [Hil88]. We prove a stronger version of (3.2) in Theorem 10.7. For the *Stieltjes moment problem* (Stieltjes' Theorem 3.1) he uses

$$\operatorname{Pos}([0,\infty)) = \left\{ f_1^2 + f_2^2 + x \cdot (g_1^2 + g_2^2) \,\middle|\, f_1, f_2, g_1, g_2 \in \mathbb{R}[x] \right\} \tag{3.3}$$

with the reference to G. Pólya and G. Szegö (previous editions of [PS64, PS70]). In [PS64, p. 82, ex. 45] the representation (3.3) is still included while it was already known before, see [ST43, p. 6, footnote], that

$$\operatorname{Pos}([0,\infty)) = \left\{ f^2 + x \cdot g^2 \,\middle|\, f, g \in \mathbb{R}[x] \right\} \tag{3.4}$$

is sufficient. Also in [Sch17, Prop. 3.2] the representation (3.3) is used, not the simpler representation (3.4). We prove a stronger version of (3.4) in Corollary 10.2.

For the $[-1,1]$-moment problem Haviland uses

$$\operatorname{Pos}([-1,1]) = \left\{ f^2 + (1-x^2) \cdot g^2 \,\middle|\, f, g \in \mathbb{R}[x] \right\}. \tag{3.5}$$

For the *Hausdorff moment problem* (Hausdorff's Theorem 3.3) he uses that any strictly positive polynomial on $[0,1]$ is a linear combination of

$$x^m \cdot (1-x)^p \tag{3.6}$$

with $m, p \in \mathbb{N}_0$, $p \geq m$, and with non-negative coefficients.

Haviland gives this with the references to a previous edition of [PS70]. This result is actually due to S. N. Bernstein [Ber12, Ber15].

Bernstein's Theorem 3.5 ([Ber12] for (i), [Ber15] for (ii); or see e.g. [Ach56, p. 30] or [Sch17, Prop. 3.4])**.** *Let* $f \in C([0,1],\mathbb{R})$ *and let*

$$B_{f,d}(x) := \sum_{k=0}^{d} \binom{d}{k} \cdot x^k \cdot (1-x)^{d-k} \cdot f\left(\frac{k}{d}\right) \tag{3.7}$$

be the Bernstein polynomials *of* f *with* $d \in \mathbb{N}$. *Then the following hold:*

(i) The polynomials $B_{f,d}$ *converge uniformly on* $[0,1]$ *to* f, *i.e.,*

$$\|f - B_{f,d}\|_\infty \xrightarrow{d\to\infty} 0.$$

(ii) If additionally $f \in \mathbb{R}[x]$ *with* $f > 0$ *on* $[0,1]$ *then there exist a constant* $D = D(f) \in \mathbb{N}$ *and constants* $c_{k,l} \geq 0$ *for all* $k, l = 0, \dots, D$ *such that*

$$f(x) = \sum_{k,l=0}^{D} c_{k,l} \cdot x^k \cdot (1-x)^l.$$

(iii) The statements (i) and (ii) also hold on $[0,1]^n$ *for any* $n \in \mathbb{N}$. *Especially every* $f \in \mathbb{R}[x_1,\dots,x_n]$ *with* $f > 0$ *on* $[0,1]^n$ *is of the form*

$$f(x) = \sum_{\alpha_1,\dots,\beta_n=0}^{D} c_{\alpha_1,\dots,\beta_n} \cdot x_1^{\alpha_1} \cdots x_n^{\alpha_n} \cdot (1-x_1)^{\beta_1} \cdots (1-x_n)^{\beta_n}$$

for some $D \in \mathbb{N}$ *and* $c_{\alpha_1,\dots,\beta_n} \geq 0$.

The multidimensional statement (iii) follows from the classical one-dimensional cases (i) and (ii). For this and more on Bernstein polynomials see e.g. [Lor86].

Bernstein's Theorem 3.5 only holds for $f > 0$. Allowing zeros at the interval end points is possible and gives the following "if and only if"-statement.

Corollary 3.6. *Let* $f \in \mathbb{R}[x] \setminus \{0\}$. *The following are equivalent:*

(i) $f > 0$ *on* $(0,1)$.

(ii) $f(x) = \sum_{i=0}^{D} c_{k,l} \cdot x^l \cdot (1-x)^k$ *for some* $D \in \mathbb{N}$, $c_{k,l} \geq 0$ *for all* $k, l = 0, \dots, D$, *and* $c_{k',l'} > 0$ *at least once.*

Proof. See Problem 3.5. □

On $[-1,1]$ a strengthened version of Bernstein's Theorem 3.5 (ii) is attributed to F. Lukács [Luk18] (*Lukács Theorem*). Note that Lukács in [Luk18] reproves several results/formulas which already appeared in a work by M. R. Radau [Rad80], as pointed out by L. Brickman [Bri59, p. 196]. Additionally, in [KN77, p. 61, footnote 4] M. G. Krein and A. A. Nudel'man state that A. A. Markov proved a more precise version of Lukács Theorem already in 1906 [Mar06],[1] see also [Mar95]. Krein and Nudel'man call it *Markov's Theorem*. It is the following.

Lukács–Markov Theorem 3.7 ([Mar06] or e.g. [Luk18], [KN77, p. 61, Thm. 2.2])**.** *Let* $-\infty < a < b < \infty$ *and let* $p \in \mathbb{R}[x]$ *be with* $\deg p = n$ *and* $p \geq 0$ *on* $[a,b]$. *The following hold:*

(i) If $\deg p = 2m$ *for some* $m \in \mathbb{N}_0$ *then* p *is of the form*

$$p(x) = f(x)^2 + (x-a)(b-x) \cdot g(x)^2$$

for some $f, g \in \mathbb{R}[x]$ *with* $\deg f = m$ *and* $\deg g = m-1$.

(ii) If $\deg p = 2m+1$ *for some* $m \in \mathbb{N}_0$ *then* p *is of the form*

$$p(x) = (x-a) \cdot f(x)^2 + (b-x) \cdot g(x)^2$$

for some $f, g \in \mathbb{R}[x]$ *with* $\deg f = \deg g = m$.

For case (i) note that the relation

$$(x-a)(b-x) = \frac{1}{b-a}\left[(x-a)^2(b-x) + (x-a)(b-x)^2\right] \tag{3.8}$$

implies

[1] We do not have access to [Mar06] and can therefore neither confirm nor decline this statement.

$$\operatorname{Pos}([a,b]) = \left\{ f(x)^2 + (x-a)\cdot g(x)^2 + (b-x)\cdot h(x)^2 \,\middle|\, f,g,h \in \mathbb{R}[x] \right\}. \quad (3.9)$$

The special part about the Lukács–Markov Theorem 3.7 are the degree bounds on the polynomials f and g. Equation (3.8) destroyes these degree bounds since we have to go one degree higher.

In the Lukács–Markov Theorem 9.5 we will see how from Karlin's Positivstellensatz 7.3 an even stronger version follows which describes the polynomials f and g more precisely and up to a certain point uniquely. In [KN77, p. 61 Thm. 2.2 and p. 373 Thm. 6.4] the Lukács–Markov Theorem 3.7 is called *Markov–Lukács Theorem* since Markov gave the more precise version much earlier than Lukács. In [Hav36] Haviland uses this result without any reference or attribution to either Lukács or Markov.

For the two-dimensional Hausdorff moment problem Haviland uses with a reference to [HS33] that any polynomial $f \in \mathbb{R}[x,y]$ which is strictly positive on $[0,1]^2$ is a linear combination of $x^m \cdot y^n \cdot (1-x)^p \cdot (1-y)^q$, $n,m,q,p \in \mathbb{N}_0$, with non-negative coefficients. This is actually Bernstein's Theorem 3.5 (iii).

T. H. Hildebrandt and I. J. Schoenberg [HS33] already solved the moment problem on $[0,1]^n$ for all $n \in \mathbb{N}$ getting the same result as Haviland. The idea of using $\operatorname{Pos}(K)$-descriptions to solve the moment problem was therefore already used by Hildebrandt and Schoenberg in 1933 [HS33] before Haviland uses this in [Hav35] and generalized this in [Hav36] as suggested to him by Wintner.

With these broader historical remarks we see that of course more people are connected to Theorem 3.4. It might also be appropriate to call Theorem 3.4 the *Haviland–Wintner* or *Haviland–Hildebrandt–Schoenberg–Wintner Theorem*. But as so often, the list of contributors is long (and maybe even longer) and hence the main contribution (the general proof) is rewarded by calling it just Haviland's Theorem.

The last classical moment problem which we want to mention on the long list was solved by K. I. Švenco [Šve39].

Švenco's Theorem 3.8. *Let $s = (s_i)_{i\in\mathbb{N}_0}$ be a real sequence. The following are equivalent:*

(i) s is a $(-\infty,0] \cup [1,\infty)$-moment sequence.
(ii) $L_s(p) \geq 0$ for all $p \in \operatorname{Pos}((-\infty,0]\cup[1,\infty))$.
(iii) $L_s(p^2) \geq 0$, $L_{(X^2-X)s}(p^2) \geq 0$ for all $p \in \mathbb{R}[x]$.
(iv) s and $(X^2-X)s$ are positive semi-definite.
(v) $\mathcal{H}(s) \succeq 0$ and $\mathcal{H}((X^2-X)s) \succeq 0$.

The general case of Švenco's Theorem 3.8 on

$$\mathbb{R} \setminus \bigcup_{i=1}^{n} (a_i, b_i) \quad (3.10)$$

for any $n \in \mathbb{N}$ and $a_1 < b_1 < \dots < a_n < b_n$ was proved by V. A. Fil'štinskiĭ [Fil64]. All non-negative polynomials on (3.10) can be explicitly written down. More precisely, all moment problems on closed and semi-algebraic sets $K \subseteq \mathbb{R}$ follow nowadays easily from Haviland's Theorem 3.4 resp. the Basic Representation

Theorem 2.9 and some well established results from real algebraic geometry, see e.g. [Mar08, Prop. 2.7.3].

Haviland's Theorem 3.4 was important to give the solutions of the classical moment problem, i.e., mostly one-dimensional cases. After that is was no longer used and only became important again when descriptions of strictly positive and non-negative polynomials on $K \subseteq \mathbb{R}^n$ with $n \geq 2$ be came available. This process was started with [Sch91] and real algebraic geometry was revived by it.

3.2 Early Results with Gaps

The early history of moment problems with gaps is very thin. We discuss only [Hau21b] and [Boa39a].

Hausdorff just solved Hausdorff's Theorem 3.3 in [Hau21a][2] and in [Hau21b][3] he treats

$$s_n = \int_0^1 x^{k_n}\, \mathrm{d}\mu(x)$$

for all $n \in \mathbb{N}_0$ with

$$k_0 = 0 < k_1 < k_2 < \cdots < k_n < \ldots$$

for a sequence of real numbers k_i, i.e., not necessarily in $\mathbb{N}_0$. See also [ST43, p. 104]. Since Hausdorff in [Hau21b] did not have access to Haviland's Theorem 3.4 [Hav36] or the description of all non-negative linear combinations of $1, x^{k_1}, \ldots, x^{k_n}, \ldots$ the results in [Hau21b] need complicated formulations and are not very strong. Only with the description of non-negative linear combinations by Karlin [Kar63] an easy formulation of the result is possible. We will therefore postpone the exact formulation to Theorem 9.6 and Theorem 9.8 where we present easy proofs using also the theory of adapted spaces from Chapter 2, especially the Basic Representation Theorem 2.9.

In [Boa39a] Boas investigates the Stieltjes moment problem ($K = [0, \infty)$) with gaps. Similar to [Hau21b] the results are difficult to read and they are unfortunately incomplete since Boas (like Hausdorff) did not have access to the description of all non-negative or strictly positive polynomials with gaps (or more general exponents). We will give the complete solution of the $[0, \infty)$-moment problem with gaps and more general exponents in Theorem 10.4.

3.3 Finitely Atomic Representing Measures: Richter's Theorem

When working with a truncated moment sequence resp. functionals it is often useful in theory and applications to find a representing measure with finitely many atoms.

[2] Submitted: February 11, 1920.

[3] Submitted: September 8, 1920.

That this is always possible for truncated moment functionals was first proved in full generality by H. Richter [Ric57, Satz 4].

Its proof proceeds by induction via the dimension of the moment cone. To do that we need to look at the boundary of the moment cone. We need that when part of the boundary of the moment cone is cut out by a supporting hyperplane then this intersection is again a moment cone of strictly smaller dimension. That is the content of the following lemma.

Lemma 3.9. *Let $n \in \mathbb{N}$, $(\mathcal{X}, \mathfrak{A})$ be a measurable space, $\mathcal{F} = \{f_i\}_{i=1}^n$ be a family of measurable functions $f_i : \mathcal{X} \to \mathbb{R}$, $\mathcal{S}_\mathcal{F}$ be the moment cone spanned by $\mathcal{F}$, and let H be a supporting hyperplane of $\mathcal{S}_\mathcal{F}$. Then $\mathcal{S}_\mathcal{F} \cap H$ is a moment cone of dimension $m = \dim(\mathcal{S}_\mathcal{F} \cap H) < n$ spanned by a family $\mathcal{G} \subset \operatorname{lin} \mathcal{F}$ on a measurable space $(\mathcal{Y}, \mathfrak{A}|_\mathcal{Y})$ with $\mathcal{Y} \subseteq \mathcal{X}$.*

Proof. See Problem 3.6. □

With the previous lemma we can now prove Richter's Theorem.

Richter's Theorem 3.10 ([Ric57, Satz 4]; or see e.g. [Kem68, Thm. 1], [FP01, p. 198, Thm. 1])**.** *Let $n \in \mathbb{N}$, let $(\mathcal{X}, \mathfrak{A})$ be a measurable space, and let $\{f_i\}_{i=1}^n$ be a family of real linearly independent measurable functions $f_i : \mathcal{X} \to \mathbb{R}$. Then for every measure μ on $\mathcal{X}$ such that all f_i are μ-integrable, i.e.,*

$$s_i := \int_\mathcal{X} f_i(x)\, \mathrm{d}\mu(x) \quad \in \mathbb{R}$$

for all $i = 1, \ldots, n$, there exist a $k \in \mathbb{N}_0$ with $k \leq n$, points $x_1, \ldots, x_k \in \mathcal{X}$ pairwise different, and $c_1, \ldots, c_k \in (0, \infty)$ such that

$$s_i = \sum_{j=1}^k c_j \cdot f_i(x_j) = \int_\mathcal{X} f_i(x)\, \mathrm{d}\nu(x) \quad \text{with} \quad \nu = \sum_{j=1}^k c_j \cdot \delta_{x_j}$$

holds for all $i = 1, \ldots, n$.

Proof. We show that every truncated moment sequence $s = (s_1, \ldots, s_n)$ has a finitely atomic representing measure with at most n atoms in $\mathcal{X}$. We prove this statement by induction on n.

$n = 1$: We have

$$s_1 = \int_\mathcal{X} f_1(x)\, \mathrm{d}\mu(x).$$

If $s_1 = 0$ then take $\nu = 0$ which proves the statement. Let us assume $s_1 \neq 0$. Since $\mu \geq 0$ on $\mathcal{X}$ there exists a point $x_1 \in \mathcal{X}$ such that $\operatorname{sgn} f_1(x_1) = \operatorname{sgn} s_1$. Hence, we have $\frac{s_1}{f_1(x_1)} =: c_1 > 0$ and

$$s_1 = \frac{s_1}{f_1(x_1)} \cdot f_1(x_1) = \int_\mathcal{X} f_1(x)\, \mathrm{d}(c_1 \cdot \delta_{x_1})(x)$$

which proves the statement.

$n \geq 2$: Let $\mathcal{S}_\mathcal{F} \subseteq \mathbb{R}^n$ be the moment cone generated from $\mathcal{F}$. We make the distinction of the two cases

(a) $s = (s_1, \dots, s_n) \in \operatorname{int} \mathcal{S}_\mathcal{F}$ and
(b) $s \in \partial\mathcal{S}_\mathcal{F} \cap \mathcal{S}_\mathcal{F}$.

For (a) let $\mathcal{S} := \operatorname{cone}\{(f_1(x), \dots, f_n(x))^T \mid x \in \mathcal{X}\}$ be the cone generated by all point evaluations $(f_1(x), \dots, f_n(x))^T$. By Carathéodory's Theorem 0.4 every $s \in \mathcal{S}$ is a moment sequences with a k-atomic representing measure with $k \leq n$. Additionally, we have that $\operatorname{int} \mathcal{S}$ is non-empty since $\mathcal{S}$ is full dimensional.

Assume $\operatorname{int} \mathcal{S} \neq \operatorname{int} \mathcal{S}_\mathcal{F}$ then $\operatorname{int}(\mathcal{S}_\mathcal{F} \setminus \mathcal{S}) \neq \emptyset$. Let $s \in \operatorname{int}(\mathcal{S}_\mathcal{F} \setminus \mathcal{S})$ with a representing measure μ. Then there exists a separating linear functional l, i.e., $l(s) < 0$ and $l(t) > 0$ for all $t \in \mathcal{S}$. Since $(f_1(x), \dots, f_n(x))^T \in \mathcal{S}$ we have that $f(x) := l((f_1(x), \dots, f_n(x)) > 0$ for all $x \in \mathcal{X}$ but

$$\int_\mathcal{X} f(x)\, \mathrm{d}\mu(x) = l(s) < 0$$

with is a contradiction to $\mu \geq 0$. Hence, $\operatorname{int} \mathcal{S} = \operatorname{int} \mathcal{S}_\mathcal{F}$ and every $s \in \operatorname{int} \mathcal{S}_\mathcal{F}$ has a k-atomic representing measure with $k \leq n$.

For (b) assume $s \in \partial\mathcal{S}_\mathcal{F} \cap \mathcal{S}_\mathcal{F}$. Since $\mathcal{S}_\mathcal{F}$ is a convex cone there exists a supporting hyperplane H of $\mathcal{S}_\mathcal{F}$ at s. But then $\mathcal{S}_\mathcal{F} \cap H$ is by Lemma 3.9 a moment cone of dimension at most $n - 1$ and here the theorem holds by induction. □

The previous proof is the original proof by Richter and only the mathematical language is updated. The following historical overview about Richter's Theorem 3.10 first appeared in [dDS22].

Replacing integration by finitely many point evaluations was already used and investigated by C. F. Gauß [Gau15]. The k-atomic representing measures from Richter's Theorem 3.10 are therefore also called (*Gaussian*) *cubature formulas*.

The history of Richter's Theorem 3.10 is confusing and the literature is often misleading. We therefore list in chronological order previous versions or versions which appeared almost at the same time. The conditions of these versions (including Richter) are the following:

(A) A. Wald 1939[4] [Wal39, Prop. 13]: $\mathcal{X} = \mathbb{R}$ and $f_i(x) = |x - x_0|^{d_i}$ with $d_i \in \mathbb{N}_0$, $0 \leq d_1 < d_2 < \dots < d_n$, and $x_0 \in \mathcal{X}$.

(B) P. C. Rosenbloom 1952 [Ros52, Cor. 38e]: $(\mathcal{X}, \mathfrak{A})$ a measurable space and f_i bounded measurable functions.

(C) H. Richter 1957[5] [Ric57, Satz 4]: $(\mathcal{X}, \mathfrak{A})$ a measurable space and f_i measurable functions.

(D) M. V. Tchakaloff 1957[6] [Tch57, Thm. II]: $\mathcal{X} \subset \mathbb{R}^n$ compact and f_i monomials of degree at most d.

---

[4] Received: February 25, 1939. Published: September 1939.
[5] Received: December 27, 1956. Published: April, 1957.
[6] Published: July-September, 1957

(E) W. W. Rogosinski 1958[7] [Rog58, Thm. 1]: $(\mathcal{X}, \mathfrak{A})$ measurable space and f_i measurable functions.

From this list we see that Tchakaloff's result (D) from 1957 is a special case of Rosenbloom's result (E) from 1952 and that the general case was proved by Richter and Rogosinski almost about at the same time, see the exact dates in the footnotes. If one reads Richter's paper, one might think at first glance that he treats only the one-dimensional case, but a closer look reveals that his Proposition (Satz) 4 covers actually the general case of measurable functions. Rogosinski treats the one-dimensional case, but states at the end of the introduction of [Rog58]:

> Lastly, the restrictions in this paper to moment problems of dimension one is hardly essential. Much of our geometrical arguments carries through, with obvious modifications, to any finite number of dimensions, and even to certain more general measure spaces.

The above proof of Richter's Theorem 3.10, and likewise the one in [Sch17, Theorem 1.24], are nothing but modern formulations of the proofs of Richter and Rogosinski without additional arguments. Note that Rogosinki's paper [Rog58] was submitted about a half year after the appearance of Richter's [Ric57].

It might be of interest that the general results of Richter and Rogosinski from 1957/58 can be derived from Rosenbloom's Theorem from 1952, see Problem 3.7. With that wider historical perspective in mind it might be justified to call Richter's Theorem 3.10 also the *Richter–Rogosinski–Rosenbloom Theorem*.

Richter's Theorem 3.10 was overlooked in the modern literature on truncated polynomial moment problems. The problem probably arose around 1997/98 when it was stated as an open problem in a published paper.[8] The paper [Ric57] and numerous works of J. H. B. Kemperman were not included back then. Especially [Kem68, Thm. 1] where Kemperman fully states the general theorem (Richter's Theorem 3.10) and attributed it therein to Richter and Rogosinski is missing. Later on, this missing piece was not added in several other works. The error continued in the literature for several years and Richter's Theorem 3.10 was reproved in several papers in weaker forms. Even nowadays papers appear not aware of Richter's Theorem 3.10 or of the content of [Ric57].

3.4 Signed Representing Measures: Boas' Theorem

In the theory of moments almost exclusively the representation by non-negative measures is treated. The reason is the following result due to R. P. Boas [Boa39b].

Boas' Theorem 3.11 ([Boa39b] or e.g. [ST43, p. 103, Thm. 3.11]). *Let $s = (s_i)_{i\in\mathbb{N}_0}$ be a real sequence. Then there exist infinitely many signed measures μ on $\mathbb{R}$ and infinitely many signed measures ν on $[0,\infty)$ such that*

[7] Received: August 22, 1957. Published: May 6, 1958.

[8] We do not give the references for this and subsequent papers who reproved Richter's Theorem 3.10.

$$s_i = \int_{\mathbb{R}} x^i \, \mathrm{d}\mu(x) = \int_0^\infty x^i \, \mathrm{d}\nu(x)$$

holds for all $i \in \mathbb{N}_0$.

The proof follows the arguments in [ST43, pp. 103–104].

Proof. We prove the case on $[0,\infty)$. The case on $\mathbb{R}$ is then only a special case.

By induction we write $s = v - w$ such that v and w are positive definite sequences where we can apply the Basic Representation Theorem 2.9.

$i = 0$: We can chose $v_0, w_0 \gg 1$ with $s_0 = v_0 - w_0$, i.e., $L_v(p), L_w(p) \geq 0$ for all $p \in \mathrm{Pos}([0,\infty))_{\leq 0} = [0,\infty)$.

$i \to i{+}1$: Assume we found $(v_j)_{j=0}^i$ and $(w_j)_{j=0}^i$ such that $L_v(p), L_w(p) \geq 0$ for all $p \in \mathrm{Pos}([0,\infty))_{\leq i}$. Since for $i{+}1$ the term x^{i+1} appears additionally to $1, x, x^2, \dots, x^i$, the convex cone $\mathrm{Pos}([0,\infty))_{\leq i+1}$ has compact base, and L is continuous on $\mathbb{R}[x]_{\leq i+1}$ we find $v_{i+1}, w_{i+1} \gg 1$ with $s_{i+1} = v_{i+1} - w_{i+1}$ such that $L_v(p), L_w(p) \geq 0$ for all $p \in \mathrm{Pos}([0,\infty))_{\leq i+1}$.

Hence, we found sequences v, w with $s = v - w$ and $L_v(p), L_w(p) \geq 0$ for all $p \in \mathrm{Pos}([0,\infty))$. By the Basic Representation Theorem 2.9 L_v is represented by some non-negative μ_+ and L_w is represented by some non-negative μ_- both with support in $[0,\infty)$, i.e., L_s is represented by $\mu = \mu_+ - \mu_-$ supported on $[0,\infty)$. □

T. Sherman showed that Boas' Theorem 3.11 (even when L is a complex linear functional) also holds in the n-dimensional case on $\mathbb{R}^n$ and $[0,\infty)^n$ for any $n \in \mathbb{N}$ [She64, Thm. 1]. Similar results are proved for linear functionals on the universal enveloping algebra $\mathcal{E}(G)$ of a Lie group G by K. Schmüdgen [Sch78]. If the Lie group G is $\mathbb{R}^n$ then this again gives Sherman's result. G. Pólya [Pól38] (see also [ST43, p. 104]) showed an extension that special kinds of measures can be chosen. Essentially, it already appeared in [Bor95], as pointed out by Pólya himself [Pól38], see [ST43, p. 104].

Pólya's Signed Representation Theorem 3.12 (see [Bor95] or [Pól38])**.** *Let* $(x_i)_{i\in\mathbb{N}_0} \subseteq \mathbb{R}$ *be such that*

(a) $x_i < x_{i+1}$ *and* $x_i \xrightarrow{i\to\infty} \infty$

or

(b) $x_i > x_{i+1}$ *and* $x_i \xrightarrow{i\to\infty} -\infty$.

Then for every sequence $(s_i)_{i\in\mathbb{N}_0} \subseteq \mathbb{R}$ *there exists a sequence* $(c_i)_{i\in\mathbb{N}_0} \subseteq \mathbb{R}$ *such that*

$$s_i = \int_{\mathbb{R}} x^i \cdot f_c(x) \, \mathrm{d}x$$

holds for all $i \in \mathbb{N}_0$ *with*

$$f_c = \sum_{i\in\mathbb{N}_0} c_i \cdot \chi_{[x_i, x_{i+1})} \qquad \textit{resp.} \qquad f_c = \sum_{i\in\mathbb{N}_0} c_i \cdot \chi_{(x_{i+1}, x_i]}$$

where χ_A is the characteristic function of a set A, i.e., every sequence has a representing measure absolutely continuous with respect to the Lebesgue measure and the density function is a step function f_c where the positions x_i of the steps can be chosen as any strictly increasing or strictly decreasing divergent sequence.

From the theory of distributions, see e.g. [Gru09], we have that the derivative of $\chi_{[a,b)}$ in the distributional sense is a signed measure, i.e.,

$$\int f(x)\cdot \chi'_{[a,b)}\,\mathrm{d}x = -\int_a^b f'(x)\,\mathrm{d}x = f(a) - f(b) = \int f(x)\,\mathrm{d}(\delta_a - \delta_b)(x) \tag{3.11}$$

for all $f \in C^1(\mathbb{R},\mathbb{R})$. Eq. (3.11) can be shortly written down by abusing notation as "$\chi_{[a,b)} = \delta_a - \delta_b$". From (3.11) and Pólya's Signed Representation Theorem 3.12 we get the following immediate consequence.

Signed Atomic Representation Theorem 3.13 (see [Blo53, Thm. 3.1])**.** *Let $(x_i)_{i\in\mathbb{N}_0} \subseteq \mathbb{R}$ be such that*

(a) $x_i < x_{i+1}$ *and* $x_i \xrightarrow{i\to\infty} \infty$

or

(b) $x_i > x_{i+1}$ *and* $x_i \xrightarrow{i\to\infty} -\infty$.

Then for any sequence $(s_i)_{i\in\mathbb{N}_0} \subseteq \mathbb{R}$ there exists a sequence $(c_i)_{i\in\mathbb{N}_0} \subseteq \mathbb{R}$ such that

$$s_i = \int_{\mathbb{R}} x^i\,\mathrm{d}\mu_c(x)$$

holds for all $i \in \mathbb{N}_0$ with

$$\mu_c = \sum_{i\in\mathbb{N}_0} c_i \cdot \delta_{x_i}.$$

Proof. Let f_d be the step function representing measure of the sequence $(t_i)_{i\in\mathbb{N}_0}$ from Pólya's Signed Representation Theorem 3.12 with $t_i := -\frac{1}{i+1} s_i$ for all $i \in \mathbb{N}_0$. Then

$$\begin{aligned}
s_i &= -(i+1)\cdot t_i \\
&= -(i+1)\cdot \int x^i \cdot f_d(x)\,\mathrm{d}x \\
&= -\int (x^{i+1})' \cdot f_d(x)\,\mathrm{d}x \\
&= \int x^{i+1}\cdot f_d(x)'\,\mathrm{d}x \\
&= \int x^{i+1}\,\mathrm{d}\mu_{\tilde{c}}(x) \\
&= \int x^i\,\mathrm{d}\mu_c(x)
\end{aligned}$$

holds for all $i \in \mathbb{N}_0$ with some $\tilde{c} = (\tilde{c})_{i\in\mathbb{N}_0} \subseteq \mathbb{R}$ and $c = (c_i)_{i\in\mathbb{N}_0} \subseteq \mathbb{R}$. The last equality (going from $\tilde{c}$ to c) holds since

$$\int f(x) \cdot x \, \mathrm{d}\delta_y(x) = f(y) \cdot y = \int f(x) \, \mathrm{d}(y \cdot \delta_y)(x),$$

i.e., $c_i = x_i \cdot \tilde{c}_i$ for all $i \in \mathbb{N}_0$. □

We get Pólya's Signed Representation Theorem 3.12 from the Signed Atomic Representation Theorem 3.13 by reversing the previous proof.

Corollary 3.14. *Let $(s_i)_{i\in\mathbb{N}_0}$ and let $U \subseteq \mathbb{R}$ be unbounded. Then there exists a signed measure μ on $\mathbb{R}$ with $\operatorname{supp}\mu = U$ such that*

$$s_i = \int x^i \, \mathrm{d}\mu(x)$$

holds for all $i \in \mathbb{N}_0$.

On $\mathbb{R}^n$ it is even possible to find a Schwartz function $f \in \mathcal{S}(\mathbb{R}^n)$ such that

$$s_\alpha = \int_{\mathbb{R}^n} x^\alpha \cdot f(x) \, \mathrm{d}x$$

for all $\alpha \in \mathbb{N}_0^n$. Use e.g. [CdD22].

Boas' Theorem 3.11, Pólya's Signed Representation Theorem 3.12, and Signed Atomic Representation Theorem 3.13 also cover the cases with gaps. If any gaps in the real sequence s are present then fill them with any real number you like.

Note, neither Boas' Theorem 3.11, Pólya's Signed Representation Theorem 3.12, nor Signed Atomic Representation Theorem 3.13 hold with the restriction that the representing signed measure shall have a bounded and therefore compact support. That is seen from the following result due to Hausdorff [Hau23, p. 232, II.].

Signed Hausdorff's Theorem 3.15 (see [Hau23, p. 232, II.] or e.g. [Lor86, Thm. 3.3.1])**.** *Let $(s_i)_{i\in\mathbb{N}_0} \subseteq \mathbb{R}$ be a real sequence. The following are equivalent:*

(i) There exist positive ($C([0,1],\mathbb{R})$-regular) measures μ_1 and μ_2, i.e., a signed ($C([0,1],\mathbb{R})$-regular) measure $\mu = \mu_1 - \mu_2$, such that

$$s_i = \int_0^1 x^i \, \mathrm{d}\mu_1(x) - \int_0^1 x^i \, \mathrm{d}\mu_2(x) = \int_0^1 x^i \, \mathrm{d}\mu(x)$$

holds for all $i \in \mathbb{N}_0$.

(ii) There exists a $C > 0$ such that

$$\sum_{k=0}^{d} \binom{d}{k} \cdot \left| L_s(x^k \cdot (1-x)^{d-k}) \right| < C$$

holds for all $d \in \mathbb{N}_0$.

The following proof is due to T. H. Hildebrandt [Hil32], see also [Lor86, pp. 58–59]. We employ the Signed Riesz' Representation Theorem 0.18 together with the Bernstein polynomials $B_{f,d}$ in (3.7) and Bernstein's Theorem 3.5.

Proof. (i) $\Rightarrow$ (ii): Since s is represented by $\mu = \mu_1 - \mu_2$ and since by Bernstein's Theorem 3.5 (i) we have $\|1 - B_{1,d}\|_\infty \to 0$ as $d \to \infty$ we have

$$\sum_{k=0}^{d} \binom{d}{k} \cdot \left|L_s(x^k \cdot (1-x)^{d-k})\right| \leq \sum_{k=0}^{d} \left(\int_0^1 B_{1,d}(x)\, \mathrm{d}\mu_1(x) + \int_0^1 B_{1,d}(x)\, \mathrm{d}\mu_2(x) \right)$$
$$\xrightarrow{d\to\infty} \mu_1([0,1]) + \mu_2([0,1]) < \infty$$

which proves (ii).

(ii) $\Rightarrow$ (i): Let $f \in C([0,1],\mathbb{R})$. Then

$$\begin{aligned} |L_s(B_{f,d})| &= \left| \sum_{k=0}^{d} f\left(\frac{k}{n}\right) \cdot \binom{d}{k} \cdot L_s(x^k \cdot (1-x)^{d-k}) \right| \\ &\leq \|f\|_\infty \cdot \sum_{k=0}^{d} \binom{d}{k} \cdot \left|L_s(x^k \cdot (1-x)^{d-k})\right| \\ &\leq C \cdot \|f\|_\infty \end{aligned}$$

which proves that L_s can be continuously extended from $\mathbb{R}[x]$ to $C([0,1],\mathbb{R})$ and the extension fulfills (0.6). To see this let $f_d \in \mathbb{R}[x]$ with $\|f - f_d\|_\infty \to 0$ as $d \to \infty$. Then $|L_s(f_d - f_{d'})| \leq c \cdot \|f_d - f_{d'}\|_\infty \to 0$ as $d, d' \to \infty$. Therefore, $(L_s(f_d))_{d\in\mathbb{N}_0}$ is a Cauchy sequence with a unique limit: $L_s(f) := \lim_{d\to\infty} L_s(f_d)$. Then (0.6) holds since $C_c([0,1],\mathbb{R}) = C([0,1],\mathbb{R})$ and $|L_s(f)| \leq C \cdot \|f\|_\infty$.

Hence, by the Signed Riesz' Representation Theorem 0.18 we have (i). □

With Bernstein's Theorem 3.5 (iii) the previous result also holds on $[0,1]^n$ for any $n \in \mathbb{N}_0$.

More on signed or complex representing measures can be found e.g. in [Blo53, Hor77, BCJ79, Kow84, Dur89, Hoi92] and references therein.

3.5 Solving all Truncated Moment Problems solves the Moment Problem

The following result was already indicated by Hamburger in [Ham20] and formalized by J. Stochel in [Sto01]. We have the following.

Theorem 3.16. *Let $n \in \mathbb{N}$, $K \subseteq \mathbb{R}^n$ be closed, $\mathcal{V} \subseteq \mathbb{R}[x_1,\dots,x_n]$ be an adapted space on K, and let $L : \mathcal{V} \to \mathbb{R}$ be a linear functional on $\mathcal{V}$. The following are equivalent:*

(i) $L : \mathcal{V} \to \mathbb{R}$ *is a K-moment functional.*
(ii) $L_k := L|_{\mathcal{V}\cap\mathbb{R}[x_1,\dots,x_n]_{\leq k}}$ *are truncated K-moment functionals for all $k \in \mathbb{N}_0$.*

Proof. While "(i) $\Rightarrow$ (ii)" is clear it is sufficient to prove the reverse direction.

Let L_k be a truncated K-moment functionals for all $k \in \mathbb{N}_0$. Since $\mathcal{V} \subseteq \mathbb{R}[x_1,\dots,x_n]$ for any $p \in \mathcal{V}$ we have that $L : \mathcal{V} \to \mathbb{R}$ is well-defined by $L(p) := L_{\deg p}(p)$. Let $p \in \mathcal{V}$ with $p \geq 0$ on K then $L(p) = L_{\deg p}(p) \geq 0$, i.e., by the Basic Representation Theorem 2.9 we have that L is a K-moment functional. □

Note, $\mathcal{V}$ can also be finite dimensional when K is compact. Then the result is trivial. For unbounded K the adapted space $\mathcal{V}$ is always infinite dimensional.

A more general version of Theorem 3.16 can e.g. be found in [Sch17, Thm. 1.20].

Problems

3.1 Prove Stieltjes' Theorem 3.1 with the Basic Representation Theorem 2.9 and the representation (3.4).

3.2 Prove Hamburger's Theorem 3.2 with the Basic Representation Theorem 2.9 and the representation (3.2).

3.3 Prove Hausdorff's Theorem 3.3 with the Basic Representation Theorem 2.9 and the Lukács–Markov Theorem 3.7, resp. the representation of $\mathrm{Pos}([a,b])$ in (3.9).

3.4 Prove Haviland's Theorem 3.4 with the Basic Representation Theorem 2.9.

3.5 Use Bernstein's Theorem 3.5 (ii) to prove Corollary 3.6.

3.6 Prove Lemma 3.9.

3.7 Show that Richter's Theorem 3.10 follows from Rosenbloom's Theorem, i.e., show that the additional assumption that all f_i are bounded on the measurable space $(\mathcal{X}, \mathfrak{A})$ can be removed.

Part II
Tchebycheff Systems

Chapter 4
T-Systems

There is nothing more practical than a good theory.

Kurt Lewin [Lew43]

In this chapter we introduce the Tchebycheff systems or short T-systems. We give basic examples and properties.

4.1 The Early History of T-Systems

In our presentation we mostly limit ourselves to the works [Kre51, Kar63, KS66, KN77]. However, the concept of T-system was introduces much earlier. It goes back to its name giver: P. L. Tchebycheff [Tch74]. See especially [Kre51] for a good overview of the history of the development of T-systems and also [Gon00].

In [Tch74] Tchebycheff states the following open problem:

> Let
> $$a < \xi < \eta < b$$
> be real numbers and let the numbers
> $$s_k = \int_a^b x^k f(x)\, \mathrm{d}x \tag{4.1}$$
> for $k = 0, 1, \ldots, n-1$ for some $n \in \mathbb{N}_0$ be given. Find the bounds on the integral
> $$\int_\xi^\eta f(x)\, \mathrm{d}x \tag{4.2}$$
> under the conditions that $f \geq 0$ on $[a, b]$ and (4.1) holds.

From this investigation Tchebycheff arrives at the method of continued fractions, which was used in the early results in the moment problems, see Section 3.1. Tchebycheff gives without proof the inequalities (upper and lower bounds) of (4.2). The proof was independently found by others, see [Kre51, pp. 3–4]. The key here is to work over a finitely dimensional space spanned by $f_0, \ldots, f_n$.

A well-known and guiding example are the functions $1, x, \ldots, x^n$.

Example 4.1. Let $n \in \mathbb{N}$ and $\mathcal{X} \subseteq \mathbb{R}$ with $|\mathcal{X}| \geq n+1$. Then the family $\mathcal{F} = \{x^i\}_{i=0}^n$ is a T-system, see Definition 4.2 below. This follows immediately from the Vandermonde determinant

$$\det \left(x_i^j\right)_{i,j=0}^n = \prod_{0\leq i<j\leq n} (x_j - x_i)$$

for any $x_0, \ldots, x_n \in \mathcal{X}$. $\circ$

Krein states that he developed "the connection between ideas of Markov and functional-geometric ideas" which made it possible to remove the Wronskian approach (Definition 5.6) and replacing it with continuity and the condition

> The curve Γ of the $(n+1)$-dimensional space $\mathbb{R}^{n+1}$:
>
> $$y_0 = f_0(x), \quad y_1 = f_1(x), \quad \ldots, y_n = f_n(x)$$
>
> does not intersect itself and no hyperplane through the origin intersects it in more than n points.

which is equivalent to

> No linear combination
>
> $$\sum_{i=0}^n a_i f_i \quad \text{with} \quad \sum_{i=0}^n a_i^2 > 0$$
>
> vanishes more than n times in the closed interval $[a, b]$.

see [Kre51, pp. 19–20]. The later is then generalized to leave out continuity and replacing $[a, b]$ with any set $\mathcal{X}$, see Definition 4.2. For a family $\{f_i\}_{i=0}^n$ with this property S. N. Bernstein [Ber37] introduced the name *Tchebycheff system* and Krein [Kre51, p. 20] and Archieser [Ach56, p. 73, §47] continued using this terminology.

For more on the history see e.g. [Kre51]. We especially recommend the very nice survey article [Gon00] with the references therein for more on the works, the contributions, and the impact of Tchebycheff's work.

4.2 Definition and Basic Properties

Definition 4.2. Let $n \in \mathbb{N}_0$, $\mathcal{X}$ be a set with $|\mathcal{X}| \geq n+1$, and $\mathcal{F} = \{f_i\}_{i=0}^n$ be a family of real functions $f_i : \mathcal{X} \to \mathbb{R}$. We call a linear combination

$$f = \sum_{i=0}^n a_i \cdot f_i \quad \in \operatorname{lin} \mathcal{F} := \{a_0 f_0 + \cdots + a_n f_n \mid a_0, \ldots, a_n \in \mathbb{R}\} \tag{4.3}$$

a *polynomial*. The family $\mathcal{F}$ on $\mathcal{X}$ is called a *Tchebycheff system* (or short *T-system*) *of order n on* $\mathcal{X}$ if every polynomial $f \in \operatorname{lin} \mathcal{F}$ with $\sum_{i=0}^n a_i^2 > 0$ has at most n zeros in $\mathcal{X}$.

If additionally $\mathcal{X}$ is a topological space and $\mathcal{F}$ is a family of continuous functions we call $\mathcal{F}$ a *continuous T-system*. If additionally $\mathcal{X}$ is the unit circle $\mathbb{T}$ then we call $\mathcal{F}$ a *periodic T-system*.

The following immediate consequence shows that we can restrict the domain $\mathcal{X}$ of the T-system $\mathcal{F}$ to some $\mathcal{Y} \subseteq \mathcal{X}$ and as long as $|\mathcal{Y}| \geq n+1$ the restricted T-system remains a T-system. In applications and examples we therefore only need to prove the T-system property on some larger set $\mathcal{X}$.

Corollary 4.3. *Let $n \in \mathbb{N}_0$ and let $\mathcal{F} = \{f_i\}_{i=0}^n$ be a T-system of order n on some set $\mathcal{X}$ with $|\mathcal{X}| \geq n+1$. Let $\mathcal{Y} \subseteq \mathcal{X}$ with $|\mathcal{Y}| \geq n+1$. Then $\mathcal{G} := \{f_i|_{\mathcal{Y}}\}_{i=0}^n$ is a T-system of order n on $\mathcal{Y}$.*

Proof. See Problem 4.1. □

The set $\mathcal{X}$ does not require any structure or property except $|\mathcal{X}| \geq n+1$.

In the theory of T-systems we often deal with one special matrix. We use the following abbreviation.

Definition 4.4. Let $n \in \mathbb{N}_0$, $\mathcal{F} = \{f_i\}_{i=0}^n$ be a family of real functions on a set $\mathcal{X}$ with $|\mathcal{X}| \geq n+1$. We define the matrix

$$\begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x_0 & x_1 & \dots & x_n \end{pmatrix} := \begin{pmatrix} f_0(x_0) & f_1(x_0) & \dots & f_n(x_0) \\ f_0(x_1) & f_1(x_1) & \dots & f_n(x_1) \\ \vdots & \vdots & & \vdots \\ f_0(x_n) & f_1(x_n) & \dots & f_n(x_n) \end{pmatrix} = (f_i(x_j))_{i,j=0}^n \tag{4.4}$$

for any $x_0, \dots, x_n \in \mathcal{X}$.

Lemma 4.5 (see e.g. [KN77, p. 31])**.** *Let $n \in \mathbb{N}_0$, $\mathcal{X}$ be a set with $|\mathcal{X}| \geq n+1$, and $\mathcal{F} = \{f_i\}_{i=0}^n$ be a family of real functions $f_i : \mathcal{X} \to \mathbb{R}$. The following are equivalent:*

(i) $\mathcal{F}$ is a T-system of order n on $\mathcal{X}$.
(ii) The determinant

$$\det \begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x_0 & x_1 & \dots & x_n \end{pmatrix}$$

does not vanish for any pairwise distinct points $x_0, \dots, x_n \in \mathcal{X}$.

Proof. (i) ⇒ (ii): Let $x_0, \dots, x_n \in \mathcal{X}$ be pairwise distinct. Since $\mathcal{F}$ is a T-system we have that any non-trivial polynomial f has at most n zeros, i.e., the matrix

$$\begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x_0 & x_1 & \dots & x_n \end{pmatrix}$$

has trivial kernel and hence its determinant is non-zero. Since $x_0, \dots, x_n \in \mathcal{X}$ are arbitrary pairwise distinct we have (ii).

(ii) ⇒ (i): Assume there is a polynomial f with $\sum_{i=0}^n a_i^2 > 0$ which has the $n+1$ pairwise distinct zeros $z_0, \dots, z_n \in \mathcal{X}$. Then the matrix

$$Z = \begin{pmatrix} f_0 & f_1 & \dots & f_n \\ z_0 & z_1 & \dots & z_n \end{pmatrix}$$

has non-trivial kernel since $0 \neq (a_0, a_1, \dots, a_n)^T \in \ker Z$ and hence $\det Z = 0$ in contradiction to (ii). □

Lemma 4.5 is used in [KS66, p. 3, Dfn. 2.1] as the definition of a continuous T-system where it is called a weak T-system. In [KS66, p. 22, Thm. 4.1] then the equivalence to Definition 4.2 is shown.

Remark 4.6. Lemma 4.5 implies that for any $x \in \mathcal{X}$ there is a $f \in \operatorname{lin} \mathcal{F}$ such that $f(x) \neq 0$, i.e., the $f_0, \dots, f_n$ do not have common zeros. ∘

Remark 4.7. After adjusting the sign of f_n in a continuous T-system $\mathcal{F} = \{f_i\}_{i=0}^n$ on $[a, b]$ we can assume that

$$\det \begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x_0 & x_1 & \dots & x_n \end{pmatrix} > 0$$

holds for all $a \leq x_1 < x_2 < \dots < x_n \leq b$. ∘

The previous lemma implies the following.

Corollary 4.8 (see e.g. [KN77, p. 33]). *Let $n \in \mathbb{N}_0$, and $\mathcal{F} = \{f_i\}_{i=0}^n$ be a T-system of order n on some set $\mathcal{X}$ with $|\mathcal{X}| \geq n+1$. Let $\mathcal{W}$ be a set with $n+1 \leq |\mathcal{W}| \leq |\mathcal{X}|$ and let $g : \mathcal{W} \to \mathcal{X}$ be injective. Then $\mathcal{G} = \{g_i\}_{i=0}^n$ with $g_i := f_i \circ g$ is a T-system of order n on $\mathcal{W}$.*

Proof. See Problem 4.2. □

Corollary 4.9 (see e.g. [KS66, p. 10] or [KN77, p. 33]). *Let $n \in \mathbb{N}_0$, and $\mathcal{F} = \{f_i\}_{i=0}^n$ be a T-system of order n on some set $\mathcal{X}$ with $|\mathcal{X}| \geq n+1$. Let $g : \mathcal{X} \to \mathbb{R}$ be such that $g > 0$ on $\mathcal{X}$. Then $\mathcal{G} = \{g_i\}_{i=0}^n$ with $g_i := g \cdot f_i$ is a T-system of order n on $\mathcal{X}$.*

Proof. See Problem 4.3. □

Corollary 4.10 (see e.g. [KN77, p. 33]). *Let $n \in \mathbb{N}_0$, and $\mathcal{F} = \{f_i\}_{i=0}^n$ be a T-system of order n on some set $\mathcal{X}$ with $|\mathcal{X}| \geq n+1$. The following hold:*

(i) The functions $f_0, \dots, f_n$ are linearly independent over $\mathcal{X}$.
(ii) For any $f = \sum_{i=0}^n a_i \cdot f_i \in \operatorname{lin} \mathcal{F}$ the coefficients $a_0, \dots, a_n \in \mathbb{R}$ are unique.

Proof. See Problem 4.4. □

The previous corollary extends to the following result.

Theorem 4.11 (see e.g. [KN77, p. 33]). *Let $n \in \mathbb{N}_0$, $\mathcal{F}$ be a T-system on some set $\mathcal{X}$ with $|\mathcal{X}| \geq n+1$, and let $x_0, \dots, x_n \in \mathcal{X}$ be $n+1$ pairwise different points. The following hold:*

(i) Every $f \in \operatorname{lin} \mathcal{F}$ is uniquely determined by its values $f(x_0), \dots, f(x_n)$.
(ii) For any $y_0, \dots, y_n \in \mathbb{R}$ there exists a unique $f \in \operatorname{lin} \mathcal{F}$ such that $f(x_i) = y_i$ holds for all $i = 0, \dots, n$.

Proof. (i): Since $f \in \operatorname{lin} \mathcal{F}$ we have $f = \sum_{i=0}^n a_i \cdot f_i$. Let $x_1, \dots, x_n \in \mathcal{X}$ be pairwise distinct points. Then by Lemma 4.5 (i) $\Rightarrow$ (ii) we have that

$$\begin{pmatrix} f(x_0) \\ \vdots \\ f(x_n) \end{pmatrix} = \begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x_0 & x_1 & \dots & x_n \end{pmatrix} \cdot \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix}$$

has the unique solution $\alpha_0 = a_0, \dots, \alpha_n = a_n$.

(ii): By the same argument as in (i) the system

$$\begin{pmatrix} y_0 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x_0 & x_1 & \dots & x_n \end{pmatrix} \cdot \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix}$$

has the unique solution $\alpha_0 = a_0, \dots, \alpha_n = a_n$. □

4.3 The Curtis–Mairhuber–Sieklucki Theorem

So far we imposed no structure on the set $\mathcal{X}$. We now get a structure of $\mathcal{X}$. The following structural result was proved in [Mai56, Thm. 2] for compact subsets $\mathcal{X}$ of $\mathbb{R}^n$ and for arbitrary compact sets $\mathcal{X}$ in [Sie58] and [Cur59, Thm. 8 and Cor.].

Curtis–Mairhuber–Sieklucki Theorem 4.12. *Let $n \in \mathbb{N}_0$ and $\mathcal{F}$ be a continuous T-system of order n on a topological space $\mathcal{X}$. If $\mathcal{X}$ is a compact metrizable space then $\mathcal{X}$ can be homeomorphically embedded in the unit circle $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$.*

The proof is not difficult but technical and too lengthy for our purposes. We therefore refer the reader to [Cur59, Thm. 8].

An immediate consequence of the Curtis–Mairhuber–Sieklucki Theorem 4.12 is that every T-system is up to homomorphisms one-dimensional, i.e., in algebraic applications of the theory of T-systems we can only deal with the univariate case. Additionally, we have the following result.

Corollary 4.13 (see e.g. [Cur59, Cor. after Thm. 8])**.** *The order n of a periodic T-system is even.*

Proof. Let $\varphi : [0, 2\pi] \to S = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2\}$ with $\varphi(\alpha) = (\sin \alpha, \cos \alpha)$ and $\mathcal{F} = \{f_i\}_{i=0}^n$ be a periodic T-system. Then the f_i are continuous and hence also

$$\det \begin{pmatrix} f_0 & f_1 & \dots & f_n \\ t_0 & t_1 & \dots & t_n \end{pmatrix}$$

is continuous in $t_0, \dots, t_n \in S$. If $\mathcal{F}$ is a T-system we have that

$$d(\alpha) := \det \begin{pmatrix} f_0 & f_1 & \dots & f_n \\ \varphi(\alpha) & \varphi(\alpha + 2\pi/(n+1)) & \dots & \varphi(\alpha + 2n\pi/(n+1)) \end{pmatrix}$$

in non-zero for all $\alpha \in [0, 2\pi]$ and never changes singes. If n is odd then $d(0) = -d(2\pi/(n+1))$ which is a contradiction. Hence, n must be even. □

4.4 Examples of T-Systems

Example 4.14 (Example 4.1 continued)**.** Let $n \in \mathbb{N}_0$ and $\mathcal{X} = \mathbb{R}$. Then the family $\mathcal{F} = \{x^i\}_{i=0}^n$ of monomials is a T-system. To see this let $x_0 < x_1 < \cdots < x_n$ be $n+1$ points in $\mathbb{R}$. We then have by the Vandermonde determinant

$$\det \begin{pmatrix} 1 & x & \dots & x^n \\ x_0 & x_1 & \dots & x_n \end{pmatrix} = \prod_{0 \le i < j \le n} (x_j - x_i) \tag{4.5}$$

which is always non-zero and hence $\mathcal{F}$ is a T-system of order n on $\mathbb{R}$ by Lemma 4.5. Additionally, by Corollary 4.3 we have that $\mathcal{F}$ is a T-system of order n on any $\mathcal{Y} \subseteq \mathbb{R}$ with $|\mathcal{Y}| \geq n+1$. ∘

Note, that in (4.5) the functions f_i should be written more precisely as

$$f_i : \mathbb{R} \to \mathbb{R},\ x \mapsto x^i$$

and not just as x^i. However, we then would have the notation

$$\begin{pmatrix} \cdot^0 & \cdot^1 & \dots & \cdot^n \\ x_0 & x_1 & \dots & x_n \end{pmatrix} \quad \text{or more general} \quad \begin{pmatrix} \cdot^{\alpha_0} & \cdot^{\alpha_1} & \dots & \cdot^{\alpha_n} \\ x_0 & x_1 & \dots & x_n \end{pmatrix}$$

for α_i with $-\infty < \alpha_0 < \alpha_1 < \cdots < \alpha_n < \infty$ which seems to be hard to read. We will therefore abuse the notation and use x^i, x^{α_i}, and (4.5).

Example 4.14 can be generalized to non-negative real exponents.

Example 4.15 (see e.g. [KS66, p. 9, Exm. 1] or [KN77, p. 38, §2(d)])**.** Let $n \in \mathbb{N}_0$ and let $0 = \alpha_0 < \alpha_1 < \cdots < \alpha_n$ be non-negative reals. Then

$$\mathcal{F} = \{x^{\alpha_0}, x^{\alpha_1}, \dots, x^{\alpha_n}\}$$

is a T-system of order n on any $\mathcal{X} \subseteq [0, \infty)$ with $|\mathcal{X}| \geq n+1$. ∘

If we restrict $\mathcal{X}$ to $\mathcal{X} \subseteq (0, \infty)$ then we can allow arbitrary real exponents α_i.

Example 4.16. Let $n \in \mathbb{N}$ and $\alpha_0 < \alpha_1 < \cdots < \alpha_n$ be reals. Then

$$\mathcal{F} = \{x^{\alpha_0}, x^{\alpha_1}, \dots, x^{\alpha_n}\}$$

is a T-system on any $\mathcal{X} \subseteq (0, \infty)$ with $|\mathcal{X}| \geq n+1$. ∘

By using $\exp : \mathbb{R} \to (0, \infty)$ we find that the previous example is by Corollary 4.8 equivalent to the following.

Example 4.17 (see e.g. [KN77, p. 38]). Let $n \in \mathbb{N}$ and $\alpha_0 < \alpha_1 < \cdots < \alpha_n$ be reals. Then

$$\mathcal{G} = \{e^{\alpha_0 x}, e^{\alpha_1 x}, \dots, e^{\alpha_n x}\}$$

is a T-system on any $\mathcal{Y} \subseteq \mathbb{R}$ with $|\mathcal{Y}| \geq n + 1$. $\circ$

That the equivalent Examples 4.16 and 4.17 are T-systems will be postponed to Examples 5.18. The reason is that with the introduction of ET-systems in Chapter 5 and especially Theorem 5.14 we generate plenty of examples of ET- and T-systems.

Example 4.18 (see e.g. [PS64, p. 41, no. 26] or [KN77, p. 37-38]). Let $n \in \mathbb{N}$ and $\alpha_0 < \alpha_1 < \cdots < \alpha_n$ be reals. Then

$$\mathcal{F} = \left\{\frac{1}{x + \alpha_0}, \frac{1}{x + \alpha_1}, \dots, \frac{1}{x + \alpha_n}\right\}$$

is a continuous T-system on any $[a, b]$ or $[a, \infty)$ with $-\alpha_0 < a < b$. $\circ$

Proof. See Problem 4.5. $\square$

Example 4.19 (see e.g. [KN77, p. 38]). Let $n \in \mathbb{N}$ and let $f \in C^n(\mathcal{X}, \mathbb{R})$ with $\mathcal{X} = [a, b]$, $a < b$, and $f^{(n)} > 0$ on $\mathcal{X}$. Then

$$\mathcal{F} = \{1, x, x^2, \dots, x^{n-1}, f\}$$

is a continuous T-system of order n on $\mathcal{X} = [a, b]$. We can also allow $\mathcal{X} = (a, b)$, $[a, \infty)$, $(-\infty, b)$, ... $\circ$

With the techniques developed in Chapter 5 it will be easy to show that Example 4.19 is not only a T-system but in fact also an ET- and ECT-system. We will therefore postpone its proof to Problem 5.5.

4.5 Representation as a Determinant, Zeros, and Non-Negativity

The following result shows that when enough zeros of a polynomial $f \in \operatorname{lin} \mathcal{F}$ are known then f has the following representation as a determinant.

Theorem 4.20 (see e.g. [KN77, p. 33]). *Let $n \in \mathbb{N}$, $\mathcal{F} = \{f_i\}_{i=0}^n$ be a T-system on some set $\mathcal{X}$ with $|\mathcal{X}| \geq n + 1$, $x_1, \dots, x_n \in \mathcal{X}$ be n pairwise distinct points, and let $f \in \operatorname{lin} \mathcal{F}$. The following are equivalent:*

(i) $f(x_i) = 0$ holds for all $i = 1, \dots, n$.
(ii) There exists a constant $c \in \mathbb{R}$ such that

$$f(x) = c \cdot \det \begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x & x_1 & \dots & x_n \end{pmatrix}. \tag{4.6}$$

Proof. (ii) $\Rightarrow$ (i): Clear.

(i) $\Rightarrow$ (ii): If $f = 0$ then $c = 0$ so the assertion holds. If $f \neq 0$ then there exists a point $x_0 \in \mathcal{X} \setminus \{x_1, \dots, x_n\}$ such that $f(x_0) \neq 0$ since $\mathcal{F}$ is a T-system. Then also the determinant in (ii) is non-zero and we can choose c such that both f and the scaled determinant coincide also in x_0. By Corollary 4.10 a polynomial f is uniquely determined by its values $f(x_i)$ at $x_0, \dots, x_n$. This shows that (4.6) is the only polynomial which fulfills (i). □

So far we treated general T-systems. For further properties we go to continuous T-systems. By the Curtis–Mairhuber–Sieklucki Theorem 4.12 we can assume $\mathcal{X} \subseteq \mathbb{R}$.

Definition 4.21. Let $n \in \mathbb{N}_0$, $\mathcal{F}$ be a continuous T-system on $\mathcal{X} \subseteq \mathbb{R}$ an interval, $f \in \operatorname{lin} \mathcal{F}$, and let x_0 be a zero of f. Then $x_0 \in \operatorname{int} \mathcal{X}$ is called a *non-nodal* zero if f does not change sign at x_0. Otherwise the zero x_0 is called *nodal*, i.e., either f changes signs at x_0 or x_0 is a boundary point of $\mathcal{X}$.

The following result bounds the number of nodal and non-nodal zeros.

Theorem 4.22 (see [Kre51, Lem. 3.1] or e.g. [KN77, p. 34, Thm. 1.1])**.** *Let $n \in \mathbb{N}_0$, $\mathcal{F}$ be a continuous T-system of order n on $\mathcal{X} = [a, b]$ with $-\infty < a < b < \infty$. If $f \in \operatorname{lin} \mathcal{F}$ has $k \in \mathbb{N}_0$ non-nodal zeros and $l \in \mathbb{N}_0$ nodal zeros in $\mathcal{X}$ then $2k + l \leq n$.*

The proof is adapted from [KN77, pp. 34, Thm. 1.1].

Proof. We make two case distinctions, one for $k = 0$ and one for $k \geq 1$.

$k = 0$: If $f \in \operatorname{lin} \mathcal{F}$ has l zeros then $l \leq n$ by Definition 4.2.

$k \geq 1$: Let $x_1, \dots, x_p \in \operatorname{int} \mathcal{X}$ with $p \leq k + l$ be the zeros of f in $\operatorname{int} \mathcal{X}$. Set

$$M_i := \max_{x_{i-1} \leq x \leq x_i} |f(x)|$$

for all $i = 1, \dots, p + 1$ with $x_0 = a$ and $x_{p+1} = b$. Additionally, set

$$m := \frac{1}{2} \min_{i=1,\dots,p+1} M_i,$$

i.e., $m > 0$.

We construct a polynomial $g_1 \in \operatorname{lin} \mathcal{F}$ such that

(a) g_1 has the value $g(x_i) = m$ at the non-nodal zeros x_i of f with $f \geq 0$ in a neighborhood of x_i,
(b) g_1 has the values $g(x_i) = -m$ at the non-nodal zeros x_i of f with $f \leq 0$ in a neighborhood of x_i, and
(c) g_1 vanishes at all nodal zeros x_i, i.e., $g(x_i) = 0$.

After renumbering the zeros x_i we can assume $x_1, \dots, x_{k_1}$ fulfill (a), $x_{k_1+1}, \dots, x_{k_1+k_2}$ fulfill (b), and $x_{k_1+k_2+1}, \dots, x_{k_1+k_2+l}$ fulfill (c) with $k_1 + k_2 = k$. By Definition 4.2 we have $k + l \leq n$ and hence by Lemma 4.5 we have that

$$\begin{pmatrix} m \\ \vdots \\ m \\ -m \\ \vdots \\ -m \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} f_0(x_1) & \dots & f_n(x_1) \\ \vdots & & \vdots \\ f_0(x_{k_1}) & \dots & f_n(x_{k_1}) \\ f_0(x_{k_1+1}) & \dots & f_n(x_{k_1+1}) \\ \vdots & & \vdots \\ f_0(x_k) & \dots & f_n(x_k) \\ f_0(x_{k+1}) & \dots & f_n(x_{k+1}) \\ \vdots & & \vdots \\ f_0(x_{k+l}) & \dots & f_n(x_{k+l}) \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_n \end{pmatrix} \tag{4.7}$$

has at least one solution, say $\beta_0 = b_0, \dots, \beta_n = b_n$. Then $g_1 = \sum_{i=0}^n b_i \cdot f_i \in \operatorname{lin} \mathcal{F}$ fulfills (a) to (c).

Set

$$\rho := \frac{m}{2 \cdot \|g_1\|_\infty}$$

and define $g_2 := f - g_1$.

We show that to each non-nodal zero x_i of f there correspond two zeros of g_2. Let x_i be a non-nodal zero of f with $f \geq 0$ in a neighborhood of x_i. We can find a point $y_i \in (x_{i-1}, x_i)$ and a point $y_{i+1} \in (x_i, x_{i+1})$ such that

$$f(y_i) = M_i > m \qquad \text{and} \qquad f(y_{i+1}) = M_{i+1} > m.$$

Hence, $g_2(y_i) > 0$ and $g_2(y_{i+1}) > 0$. Since $g_2(x_i) = -\rho \cdot m < 0$ it follows that g_2 has a zero both in (y_i, x_i) and in (x_i, y_{i+1}).

Additionally, g_2 also vanishes at all nodal zeros of f and therefore has at least $2k + l$ distinct zeros. By Definition 4.2 we have $2k + l \leq n$. □

The previous result holds for more general sets $\mathcal{X}$.

Corollary 4.23. *Theorem 4.22 holds for sets $\mathcal{X} \subseteq \mathbb{R}$ of the form*

(i) $\mathcal{X} = (a, b)$, $[a, b)$, $(a, b]$ *with* $-\infty < a < b < \infty$,
(ii) $\mathcal{X} = (a, \infty)$, $[a, \infty)$, $(-\infty, b)$, $(-\infty, b]$ *with* $-\infty < a, b < \infty$,
(iii) $\mathcal{X} = \{x_1, \dots, x_k\} \subseteq \mathbb{R}$ *with* $k \geq n + 1$ *and* $x_1 < \dots < x_k$, *and*
(iv) countable unions of (i) to (iii).

Proof. $\mathcal{X} = [0, \infty)$: Let $0 \leq x_1 < \dots < x_k$ be the zeros of f in $[0, \infty)$. Since every T-system on $[0, \infty)$ is also a T-system on $[0, b]$ for any $b > 0$ by Corollary 4.3 the assertion follows from Theorem 4.22 with $b = x_k + 1$.

For the other assertions adapt (if necessary) the proof of Theorem 4.22. □

That non-nodal points are always inner points and have a weight of (at least) 2 in counting with multiplicities as well as that boundary points are always non-nodal and are counted (at least) once in counting the multiplicities is generalized in the following.

Definition 4.24. Let $x \in [a,b]$ with $a \leq b$. We define the *index* $\varepsilon(x)$ by

$$\varepsilon(x) := \begin{cases} 2 & \text{if } x \in (a,b), \\ 1 & \text{if } x = a \text{ or } b. \end{cases} \tag{4.8}$$

The same definition holds for sets as in Corollary 4.23.

Let $\mathcal{X} \subseteq \mathbb{R}$ be a set. We define the *index* $\varepsilon(\mathcal{X})$ *of the set* $\mathcal{X}$ by

$$\varepsilon(\mathcal{X}) := \sum_{x\in\mathcal{X}} \varepsilon(x). \tag{4.9}$$

We now want to show that for each T-system $\mathcal{F}$ not only non-negative polynomials $f \in \operatorname{lin}\mathcal{F}$ exists but we can even specify the zeros. We need the following definition.

Definition 4.25. Let $n \in \mathbb{N}_0$ and $\mathcal{F}$ be a T-system of order n on some set $\mathcal{X}$. We define

$$(\operatorname{lin}\mathcal{F})^e := \left\{ \sum_{i=0}^n a_i \cdot f_i \;\middle|\; \sum_{i=0}^n a_i^2 = 1 \right\},$$
$$(\operatorname{lin}\mathcal{F})_+ := \{ f \in \operatorname{lin}\mathcal{F} \mid f \geq 0 \text{ on } \mathcal{X} \},$$

and

$$(\operatorname{lin}\mathcal{F})_+^e := (\operatorname{lin}\mathcal{F})^e \cap (\operatorname{lin}\mathcal{F})_+.$$

With these definitions we can prove the following existence criteria for non-negative polynomials in a T-systems on $[a,b]$.

Theorem 4.26 (see [Kre51, Lem. 3.2] or e.g. [KN77, p. 35, Thm. 1.2])**.** *Let $n \in \mathbb{N}_0$, $\mathcal{F}$ be a continuous T-system on $\mathcal{X} = [a,b]$, and let $x_1,\dots,x_m \in \mathcal{X}$ be m distinct points for some $m \in \mathbb{N}$. The following are equivalent:*

(i) The points $x_1,\dots,x_m$ are zeros of a non-negative polynomial $f \in \operatorname{lin}\mathcal{F}$.

(ii) $\displaystyle\sum_{i=1}^m \varepsilon(x_i) \leq n.$

The proof is adapted from [KN77, pp. 35, Thm. 1.2].

Proof. "(i) ⇒ (ii)" is Theorem 4.22 and we therefore only have to prove "(ii) ⇒ (i)".

Case I: At first assume that $a < x_1 < \dots < x_m < b$ and $\sum_{i=0}^m \varepsilon(x_i) = 2m = n$. If $2m < n$ then add k additional points $x_{m+1},\dots,x_{m+k}$ such that $2m + 2k = n$ and $x_m < x_{m+1} < \dots < x_{m+k} < b$.

Select a sequence of points $(x_1^{(j)},\dots,x_m^{(j)}) \in \mathbb{R}^m$, $j \in \mathbb{N}$, such that

$$a < x_1 < x_1^{(j)} < \dots < x_m < x_m^{(j)} < b$$

for all $j \in \mathbb{N}$ and $\lim_{j\to\infty} x_i^{(j)} = x_i$ for all $i = 1,\dots,m$. Set

$$g_j(x) := c_j \cdot \det \begin{pmatrix} f_0 & f_1 & f_2 & \dots & f_{2m-1} & f_{2m} \\ x & x_1 & x_1^{(j)} & \dots & x_m & x_m^{(j)} \end{pmatrix} \quad \in (\operatorname{lin}\mathcal{F})^e \tag{4.10}$$

for some $c_j > 0$. Since $(\operatorname{lin}\mathcal{F})^e$ is compact we can assume that g_j converges to some $g_0 \in (\operatorname{lin}\mathcal{F})^e$. Then g_0 has $x_1, \dots, x_m$ as zeros with $\varepsilon(x_i) = 2$ and g_0 is non-negative since $g_j > 0$ on $[a, x_1)$, $(x_1^{(j)}, x_2)$, ..., $(x_{m-1}^{(j)}, x_m)$, and $(x_m^{(j)}, b]$ as well as $g_j < 0$ on $(x_1, x_1^{(j)})$, $(x_2, x_2^{(j)})$, ..., $(x_m, x_m^{(j)})$.

Case II: If $a = x_1 < x_2 < \dots < x_m < b$ with $\sum_{i=1}^m \varepsilon(x_i) = 2m - 1 = n$ the only modification required in case I is to replace (4.10) by

$$g_j(x) := -c_j \cdot \det \begin{pmatrix} f_0 & f_1 & f_2 & f_3 & \dots & f_{2m-2} & f_{2m-1} \\ x & a & x_2 & x_2^{(j)} & \dots & x_m & x_m^{(j)} \end{pmatrix} \quad \in (\operatorname{lin}\mathcal{F})^e$$

with some normalizing factor $c_j > 0$.

Case III: The procedure is similar if $x_m = b$ and $\sum_{i=1}^m \varepsilon(x_i) = n$. □

Remark 4.27. Theorem 4.26 appears in [KN77, p. 35, Thm. 1.2] in a stronger version, see also [Kre51, Lem. 3.4].

In [KN77, p. 35, Thm. 1.2] and [Kre51, Lem. 3.4] Krein claims that the $x_1, \dots, x_m$ are the *only* zeros of some non-negative $f \in \operatorname{lin}\mathcal{F}$. This holds when $n = 2m + 2p$ for some $p \geq 0$ and $x_1, \dots, x_m \in \operatorname{int}\mathcal{X}$. To see this add to $x_1, \dots, x_m$ in (4.10) points $x_{m+1}, \dots, x_{m+p} \in \operatorname{int}\mathcal{X} \setminus \{x_1, \dots, x_m\}$ and get g_0. Hence, $g_0 \geq 0$ has exactly the zeros $x_1, \dots, x_{m+p}$. Then construct in a similar way $\tilde{g}_0$ with the zeros $x_1, \dots, x_m, \tilde{x}_{m+1}, \dots, \tilde{x}_{m+p}$ with $\tilde{x}_{m+1}, \dots, \tilde{x}_{m+p} \in \operatorname{int}\mathcal{X} \setminus \{x_1, \dots, x_{m+p}\}$. Hence, $g_0 + \tilde{g}_0 \geq 0$ has only the zeros $x_1, \dots, x_m$.

A similar construction works for $n = 2m + 1$ with or without end points a or b. If $x_1, \dots, x_m$ contains no end point, i.e., all $x_i \in \operatorname{int}\mathcal{X}$, then construct a g_0 with an zero in a (and therefore $g_0(b) > 0$ since the index is odd) and a $\tilde{g}_0$ with zero in b (and therefore $\tilde{g}_0(a) > 0$). Then $g_0 + \tilde{g}_0$ has no end point as a zero.

However, Krein misses that for $n = 2m + 2p$ with $p \geq 0$ and when one end point is contained in $x_1, \dots, x_m$ then it might happen that also the other end point must appear. In [KS66, p. 28, Thm. 5.1] additional conditions are given which ensure that $x_1, \dots, x_m$ are the only zeros of some $f \geq 0$.

For example if also $\{f_i\}_{i=0}^{n-1}$ is a T-system then it can be ensured that $x_1, \dots, x_m$ are the only zeros of some non-negative polynomial $f \in \operatorname{lin}\mathcal{F}$, see [KS66, p. 28, Thm. 5.1 (b-i)], see Problem 4.7. For our main example(s), the algebraic polynomials with gaps, this holds.

The same problem appears in [KN77, p. 36, Thm. 1.3]. A weaker but correct version is given in Theorem 4.30 below.

Theorem 4.22 with the condition that $\mathcal{F}$ is an ET-system [KS66, p. 28, Thm. 5.1] is given below in Theorem 5.20. ∘

Remark 4.28. Assume that in Theorem 4.26 we have additionally that $f_0, \dots, f_n \in C^1([a, b], \mathbb{R})$. Then in (4.10) we can set $x_i^{(j)} = x_i + j^{-1}$ for all $i = 0, \dots, m$ and $j \gg 1$. For $j \to \infty$ with $c_j := j^m$ we then get

$$\begin{aligned}
g_0(x) &= \lim_{j\to\infty} j^m \cdot \det \begin{pmatrix} f_0 & f_1 & f_2 & \dots & f_{2m-1} & f_{2m} \\ x & x_1 & x_1 + j^{-1} & \dots & x_m & x_m + j^{-1} \end{pmatrix} \\
&= \lim_{j\to\infty} j^m \cdot \det \begin{pmatrix} f_0(x) & \dots & f_{2m}(x) \\ f_0(x_1) & \dots & f_{2m}(x_1) \\ f_0(x_1 + j^{-1}) & \dots & f_{2m}(x_1 + j^{-1}) \\ \vdots & & \vdots \\ f_0(x_m) & \dots & f_{2m}(x_m) \\ f_0(x_m + j^{-1}) & \dots & f_{2m}(x_m + j^{-1}) \end{pmatrix} \\
&= \lim_{j\to\infty} \det \begin{pmatrix} f_0(x) & \dots & f_{2m}(x) \\ f_0(x_1) & \dots & f_{2m}(x_1) \\ \frac{f_0(x_1+j^{-1}) - f_0(x_1)}{j^{-1}} & \dots & \frac{f_{2m}(x_1+j^{-1}) - f_{2m}(x_1)}{j^{-1}} \\ \vdots & & \vdots \\ f_0(x_m) & \dots & f_{2m}(x_m) \\ \frac{f_0(x_m+j^{-1}) - f_0(x_m)}{j^{-1}} & \dots & \frac{f_{2m}(x_m+j^{-1}) - f_{2m}(x_m)}{j^{-1}} \end{pmatrix} \qquad (4.11) \\
&= \det \begin{pmatrix} f_0(x) & \dots & f_{2m}(x) \\ f_0(x_1) & \dots & f_{2m}(x_1) \\ f_0'(x_1) & \dots & f_{2m}'(x_1) \\ \vdots & & \vdots \\ f_0(x_m) & \dots & f_{2m}(x_m) \\ f_0'(x_m) & \dots & f_{2m}'(x_m) \end{pmatrix},
\end{aligned}$$

i.e., a double zero at x_j is included by including the values $f_i'(x_j)$, $i = 0, \dots, n$. We will define that procedure and need these definitions for ET-systems in Chapter 5. ∘

Corollary 4.29. *Theorem 4.26 also holds for intervals $\mathcal{X} \subseteq \mathbb{R}$, i.e.,*

$$\mathcal{X} = (a,b),\ (a,b],\ [a,b),\ [a,b],\ (a,\infty),\ [a,\infty),\ (-\infty,b),\ (-\infty,b],\ \textit{and } \mathbb{R} \qquad (4.12)$$

with $a < b$.

Proof. We have that "(i) ⇒ (ii)" follows from Corollary 4.23. For "(ii) ⇒ (i)" we apply Theorem 4.26 on $[\min_i x_i, \max_i x_i]$. □

We will now give a sharper version of Theorem 4.22, see also Remark 4.27.

Theorem 4.30 (see e.g. [KS66, p. 30, Thm. 5.2])**.** *Let $n \in \mathbb{N}$ and $\mathcal{F}$ be a continuous T-system on $\mathcal{X} = [a,b]$. Additionally, let $x_1, \dots, x_k \in \mathcal{X}$ and $y_1, \dots, y_l \in \mathcal{X}$ be pairwise distinct points. The following are equivalent:*

(i) There exists a polynomial $f \in \operatorname{lin} \mathcal{F}$ such that

  (a) $x_1, \dots, x_k$ are the non-nodal zeros of f and
  (b) $y_1, \dots, y_l$ are the nodal zeros of f.

(ii) $2k + l \le n$.

Proof. (i) $\Rightarrow$ (ii): That is Theorem 4.22.

(ii) $\Rightarrow$ (i): Adapt the proof and especially the g_j's in (4.10) of Theorem 4.26 accordingly. Let $z_1 < \cdots < z_{k+l}$ be the x_i's and y_i's together ordered by size. Then in g_j treat every nodal z_i like the endpoint a or b, i.e., include it only once in the determinant, and insert for every non-nodal point z_i the point z_i and the sequence $z_i^{(j)} \in (z_i, z_{i+1})$ with $\lim_{j\to\infty} z_i^{(j)} = z_i$. □

Corollary 4.31. *Theorem 4.30 also holds for sets $\mathcal{X} \subseteq \mathbb{R}$ of the form*

(i) $\mathcal{X} = (a,b)$, $[a,b)$, $(a,b]$ *with* $a < b$,
(ii) $\mathcal{X} = (a,\infty)$, $[a,\infty)$, $(-\infty,b)$, $(-\infty,b]$,
(iii) $\mathcal{X} = \{x_1,\dots,x_k\} \subseteq \mathbb{R}$ *with* $k \geq n+1$ *and* $x_1 < \cdots < x_k$, *and*
(iv) finitely many unions of (i) to (iii).

Proof. In the adapted proof and the g_j's in (4.10) of Theorem 4.26 we do not need to have non-negativity, i.e., in the g_j's sign changes at the y_i's are allowed (and even required). □

Problems

4.1 Prove Corollary 4.3.

4.2 Prove Corollary 4.8.

4.3 Prove Corollary 4.9.

4.4 Prove Corollary 4.10.

4.5 Prove Example 4.18.

4.6 Why does (4.7) have at least one solution?

4.7 Assume in Theorem 4.26 we not only have that $\mathcal{F} = \{f_i\}_{i=0}^n$ is a T-system of order n, but additionally that $\{f_i\}_{i=0}^{n-1}$ is T-systems of order $n-1$. Then show that the following are equivalent:

(i) The distinct points $x_1,\dots,x_p \in [a,b]$ are the *only* zeros of some non-negative polynomial $f \in \operatorname{lin}\mathcal{F}$.
(ii) $\sum_{i=1}^p \varepsilon(x_i) \leq n$.

Chapter 5
ET- and ECT-Systems

Curiouser and curiouser!

Lewis Carroll: Alice's Adventures in Wonderland

In this chapter we introduce the concept of ET- and ECT-systems, i.e., *extended* and *extended complete* Tchebycheff systems. The sparse algebraic polynomial systems on $(0, \infty)$ are the main examples. Being an ET-system is required for certain Positiv- and Nichtnegativstellensätze in later chapters.

5.1 Definitions and Basic Properties

We remind the reader that a function $f \in C^n(\mathbb{R}, \mathbb{R})$ has a *zero* at $x_0 \in \mathbb{R}$ *of multiplicity* (at least) m if

$$f^{(k)}(x_0) = 0 \qquad \text{for all } k = 0, 1, \ldots, m-1. \tag{5.1}$$

For univariate polynomials $f \in \mathbb{R}[x]$ this translates into a factorization

$$f(x) = (x - x_0)^m \cdot g(x) \qquad \text{for some } g \in \mathbb{R}[x]. \tag{5.2}$$

While the concept of T-systems comes from the univariate polynomials, a relation like (5.2) is in general not accessible for T-systems. Hence, we rely on the more general (analytic) notion (5.1) of multiplicity but still call it *algebraic multiplicity*. At endpoints of intervals $[a, b]$ we use of course the one-sided derivatives.

Definition 5.1. Let $n \in \mathbb{N}$ and let $\mathcal{F} = \{f_i\}_{i=0}^n \subseteq C^n([a, b], \mathbb{R})$ be a T-system of order n on $[a, b]$ with $a < b$. $\mathcal{F}$ is called an *extended Tchebycheff system* (short *ET-system*) *on* $[a, b]$ if any polynomial $f \in \operatorname{lin} \mathcal{F} \setminus \{0\}$ has at most n zeros in $[a, b]$ counting algebraic multiplicities.

Remark 5.2. It is clear that every ET-system is also a T-system by only allowing multiplicity one for each zero. ∘

In Remark 4.28 eq. (4.11) we showed how double zeros can be included in the determinantal representation. Whenever we have C^1-functions in $\mathcal{F} = \{f_i\}_{i=0}^n$ and

$$x_0 < \cdots < x_i = x_{i+1} < \cdots < x_n$$

we define

$$\begin{pmatrix} f_0 & \dots & f_{i-1} & f_i & f_{i+1} & f_{i+2} & \dots & f_n \\ x_0 & \dots & x_{i-1} & (x_i & x_i) & x_{i+2} & \dots & x_n \end{pmatrix} := \begin{vmatrix} f_0(x_0) & \dots & f_n(x_0) \\ \vdots & & \vdots \\ f_0(x_{i-1}) & \dots & f_n(x_{i-1}) \\ f_0(x_i) & \dots & f_n(x_i) \\ f_0'(x_i) & \dots & f_n'(x_i) \\ f_0(x_{i+2}) & \dots & f_n(x_{i+2}) \\ \vdots & & \vdots \\ f_0(x_n) & \dots & f_n(x_n) \end{vmatrix} \tag{5.3}$$

and equivalently when $x_j = x_{j+1}, x_k = x_{k+1}, \dots$ for additional entries.

We use the additional brackets "(" and ")" to indicate that x_i is inserted in the $f_0, \dots, f_n$ and then also into $f_0', \dots, f_n'$ to distinguish (5.3) from Definition 4.4 to avoid confusion. Hence, in Definition 4.4 we have

$$\det \begin{pmatrix} f_0 & \dots & f_{i-1} & f_i & f_{i+1} & f_{i+2} & \dots & f_n \\ x_0 & \dots & x_{i-1} & x_i & x_i & x_{i+2} & \dots & x_n \end{pmatrix} = 0$$

since in two rows x_i is inserted into $f_0, \dots, f_n$, while in (5.3) we have that

$$\begin{pmatrix} f_0 & \dots & f_{i-1} & f_i & f_{i+1} & f_{i+2} & \dots & f_n \\ x_0 & \dots & x_{i-1} & (x_i & x_i) & x_{i+2} & \dots & x_n \end{pmatrix}$$

indicates that x_i is inserted in $f_0, \dots, f_n$ and then also into $f_0', \dots, f_n'$.

Extending this to zeros of multiplicity m for C^{m-1}-functions is straight forward and we leave it to the reader to write down the formulas. Similar to (5.3) we write for any $a \leq x_0 \leq x_1 \leq \cdots \leq x_n \leq b$ the matrix as

$$\begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x_0 & x_1 & \dots & x_n \end{pmatrix}^* \tag{5.4}$$

when $f_0, \dots, f_n$ are sufficiently differentiable.

We often want to express polynomials $f \in \operatorname{lin} \mathcal{F}$ as determinants (4.10) only by knowing their zeros $x_1, \dots, x_k$. If arbitrary multiplicities appear we only have $x_1 \leq x_2 \leq \cdots \leq x_n$ where we include zeros multiple times according to their algebraic multiplicities. Hence, for

$$x_0 = \cdots = x_{i_1} \ < \ x_{i_1+1} = \cdots = x_{i_2} \ < \ \dots \ < \ x_{i_k+1} = \cdots = x_n$$

we introduce a simpler notation to write down (5.3):

$$\begin{pmatrix} f_0 & f_1 & f_2 & \dots & f_n \\ x & x_1 & x_2 & \dots & x_n \end{pmatrix} := \begin{pmatrix} f_0 & f_1 & \dots & f_{i_1} & f_{i_1+1} & \dots & f_{i_2} & \dots & f_{i_k+1} & \dots & f_{i_k+1} \\ x & (x_1 & \dots & x_{i_1}) & (x_{i_1+1} & \dots & x_{i_2}) & \dots & (x_{i_k+1} & \dots & x_n) \end{pmatrix}. \tag{5.5}$$

Clearly (5.5) $\in \operatorname{lin} \mathcal{F}$. For (5.5) to be well-defined we need $\mathcal{F} \subseteq C^{m-1}$ where m is the largest multiplicity of any zero.

We see here why we require in Definition 5.1 $\mathcal{F} = \{f_i\}_{i=0}^n \subseteq C^n([a,b],\mathbb{R})$. In the case of $x_0 = x_1 = \dots = x_n$ the functions f_i need to be $C^n([a,b],\mathbb{R})$, not just $C^{n-1}([a,b],\mathbb{R})$.

Similar to Lemma 4.5 we have the following.

Theorem 5.3 ([Kre51] or e.g. [KN77, p. 37, P.1.1])**.** *Let $n \in \mathbb{N}$ and $\mathcal{F} = \{f_i\}_{i=0}^n \subseteq C^n([a,b],\mathbb{R})$ with $a < b$. Then the following are equivalent:*

(i) $\mathcal{F}$ is an ET-system.
(ii) We have

$$\det \begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x_0 & x_1 & \dots & x_n \end{pmatrix}^* \neq 0$$

for every $a \leq x_0 \leq x_1 \leq \dots \leq x_n \leq b$.

Proof. Let $x_0, \dots, x_n \in [a,b]$ with

$$a \ \leq \ x_0 = \dots = x_{i_1} \ < \ x_{i_1+1} = \dots = x_{i_2} \ < \ \dots \ < \ x_{i_k} = \dots = x_n \ \leq \ b$$

be the zeros of some $f = \sum_{i=0}^n a_n f_i \in \operatorname{lin} \mathcal{F}$. We get the coefficients $a_0, \dots, a_n$ from the system

$$0 = \begin{pmatrix} f(x_0) \\ f'(x_0) \\ \vdots \\ f^{(i_1)}(x_0) \\ f(x_{i_1+1}) \\ \vdots \\ f^{(n-i_k)}(x_{i_k}) \end{pmatrix} = \underbrace{\begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x_0 & x_1 & \dots & x_n \end{pmatrix}^*}_{=:M} \cdot \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}. \tag{5.6}$$

Hence, since $x_0, \dots, x_n$ are arbitrary we have (i) $\mathcal{F}$ is an ET-systems $\Leftrightarrow a_0 = \dots = a_n = 0 \Leftrightarrow$ (5.6) has only the trivial solution $\Leftrightarrow M$ has full rank $\Leftrightarrow$ (ii). □

Remark 5.4. Similar to Remark 4.7 for T-systems we can assume after a sign change in f_n that for every ET-system $\mathcal{F} = \{f_i\}_{i=0}^n$ on $[a,b]$ we have that

$$\det \begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x_0 & x_1 & \dots & x_n \end{pmatrix}^* > 0$$

holds for all $a \leq x_0 \leq x_1 \leq \dots \leq x_n \leq b$ since $\mathcal{F} \subseteq C^n([a,b],\mathbb{R})$. ∘

An even more special case of ET-systems and therefore T-systems are the ECT-systems which we define now.

Definition 5.5. Let $n \in \mathbb{N}_0$ and let $f_0, \dots, f_n \in C^n([a,b],\mathbb{R})$ with $a < b$. The family $\mathcal{F} = \{f_0\}_{i=0}^n$ is called an *extended complete Tchebycheff system* (short *ECT-system*) *on* $[a,b]$ if $\{f_i\}_{i=0}^k$ is an ET-system on $[a,b]$ for all $k = 0, \dots, n$.

5.2 Wronskian Determinant

To handle and work with ECT-systems it is useful to introduce the following determinant.

Definition 5.6. Let $n \in \mathbb{N}_0$ and let $f_0, \dots, f_n \in C^n([a,b],\mathbb{R})$ be with $a < b$. For each $k = 0, \dots, n$ we define the *Wronskian determinant* (short *Wronskian*) $\mathcal{W}(f_0, \dots, f_k)$ of $f_0, \dots, f_k$ to be

$$\mathcal{W}(f_0, f_1, \dots, f_k) := \det \begin{pmatrix} f_0 & f_0' & \dots & f_0^{(k)} \\ f_1 & f_1' & \dots & f_1^{(k)} \\ \vdots & \vdots & & \vdots \\ f_k & f_k' & \dots & f_k^{(k)} \end{pmatrix}. \tag{5.7}$$

The Wronskian is a common tool in the theory of ordinary differential equations. In the previous definition (5.7) we could also shortly write

$$\mathcal{W}(f_0, \dots, f_k)(x) := \det \begin{pmatrix} f_0 & f_1 & \dots & f_k \\ x & x & \dots & x \end{pmatrix}^*$$

for all $x \in [a,b]$.

Let $m_1, \dots, m_k \in \mathbb{N}$ with $m_1 + \dots + m_k = n+1$ and $x_1 < \dots < x_k$. Then the first m_j columns of $\mathcal{W}(f_0, \dots, f_n)$ are the m_j columns in

$$\begin{pmatrix} f_0 & \dots & f_{m_1-1} & f_{m_1} & \dots & f_{m_1+m_2-1} & f_{m_1+m_2} & \dots & f_n \\ x_1 & \dots & x_1 & x_2 & \dots & x_2 & x_3 & \dots & x_k \end{pmatrix}^*$$

involving x_j.

Lemma 5.7. *Let $n \in \mathbb{N}_0$, let $\mathcal{F} = \{f_i\}_{i=0}^n$ be an ET-system on $[a,b]$ with $a < b$, and let $g \in C^n([a,b],\mathbb{R})$ with $g > 0$. Then*

$$\mathcal{G} := \{g_i\}_{i=0}^n \quad \text{with} \quad g_i := g \cdot f_i$$

is an ET-system and we have

$$\mathcal{W}(g_0, \dots, g_n) = g^{n+1} \cdot \mathcal{W}(f_0, \dots, f_n).$$

Proof. See Problem 5.1. □

Lemma 5.8. *Let $n \in \mathbb{N}_0$, let $\mathcal{F} = \{f_i\}_{i=0}^n$ be an ET-system on $[c,d]$, and $g \in C^n([a,b],[c,d])$ with $g' > 0$ on $[a,b]$. Then*

$$\mathcal{G} := \{f_i \circ g\}_{i=0}^n \quad \textit{with} \quad g_i := f_i \circ g$$

is an ET-system on $[a,b]$ with

$$\mathcal{W}(g_0,\dots,g_n) = (g')^{\frac{n(n+1)}{2}} \cdot \mathcal{W}(f_0,\dots,f_n) \circ g.$$

Proof. See Problem 5.2. □

For the Wronskian the following reduction property holds.

Lemma 5.9 (see e.g. [KS66, p. 377])**.** *Let $n \in \mathbb{N}_0$ and let $f_0,\dots,f_n \in C^n([a,b],\mathbb{R})$ be with $a < b$ and $f_0 > 0$. Then for the* reduced system $g_0,\dots,g_{n-1} \in C^{n-1}([a,b],\mathbb{R})$ *defined by*

$$g_i := \left(\frac{f_{i+1}}{f_0}\right)' \tag{5.8}$$

for all $i = 0,\dots,n-1$ we have

$$\mathcal{W}(f_0,\dots,f_n) = f_0^{n+1} \cdot \mathcal{W}(g_0,\dots,g_{n-1}). \tag{5.9}$$

Proof. See Problem 5.3. □

Remark 5.10. Since $f_0,\dots,f_n \in C^n([a,b],\mathbb{R})$ we have that $\mathcal{W}(f_0,\dots,f_k)(x)$ is continuous in $x \in [a,b]$ and hence after adjusting the signs of $f_0,\dots,f_n$ we have that (5.7) being non-zero on $[a,b]$ is equivalent to $\mathcal{W}(f_0,\dots,f_k) > 0$ on $[a,b]$ for all $k = 0,\dots,n$, see also Remark 4.7 and Remark 5.4. ∘

Lemma 5.11 (see e.g. [KS66, pp. 242–245, Lem. 5.1 - 5.3])**.** *Let $n \in \mathbb{N}_0$ and let $f_1,\dots,f_n \in C^n([a,b],\mathbb{R})$ be such that*

$$\mathcal{W}(f_0) > 0, \quad \dots, \quad \mathcal{W}(f_0,\dots,f_n) > 0$$

on $[a,b]$. Define functions $g_0,\dots,g_n : [a,b] \to \mathbb{R}$ by

$$\begin{aligned} g_0 &:= f_0 \\ g_1 &:= D_0 f_1 \\ g_2 &:= D_1 D_0 f_2 \\ &\ \ \vdots \\ g_n &:= D_{n-1} \dots D_1 D_0 f_n \end{aligned}$$

with

$$D_j f := \left(\frac{f}{g_j}\right)', \quad i.e., \quad D_j = \frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{g_j}. \tag{5.10}$$

Then

(i) $g_i \in C^{n-i}([a,b],\mathbb{R})$ *are well defined with*

$$g_1 = \frac{\mathcal{W}(f_0,f_1)}{f_0^2} \qquad \text{and} \qquad g_i = \frac{\mathcal{W}(f_0,\dots,f_i)\cdot\mathcal{W}(f_0,\dots,f_{i-2})}{\mathcal{W}(f_0,\dots,f_{i-1})^2}$$

for all $i = 2,\dots,n$,

(ii) $g_i > 0$ *on* $[a,b]$ *for all* $i = 0,\dots,n$,

(iii) for any $g_{n+1} \in C([a,b],\mathbb{R})$ *with* $g_{n+1} > 0$ *on* $[a,b]$ *we define*

$$f_{n+1}(x) := g_0(x)\int_a^x g_1(y_1)\int_a^{y_1} g_2(y_2)\cdots\int_a^{y_n} g_{n+1}(y_{n+1})\,\mathrm{d}y_{n+1}\dots\mathrm{d}y_1$$

and we get

$$g_{n+1} = D_n\dots D_1D_0f_{n+1},$$

(iv) for all $k = 0,\dots,n+1$ *we have*

$$\mathcal{W}(f_0,\dots,f_k) = g_0^{k+1}g_1^k\cdots g_k$$

with g_{n+1} *and* f_{n+1} *from (iii),*

(v) there exists a $f_{n+1} \in C^{n+1}([a,b],\mathbb{R})$ *such that*

$$\mathcal{W}(f_0,\dots,f_n,f_{n+1}) > 0$$

on $[a,b]$, *and*

(vi) for all $k = 0,\dots,n+1$ *the families* $\{f_i\}_{i=0}^k$ *are T-systems on* $[a,b]$.

Proof. (i) and (ii): Since $\mathcal{W}(f_0) > 0$ we have $f_0 > 0$ and hence $g_1 = (f_1/f_0)'$ is well-defined and we have

$$\frac{\mathcal{W}(f_0,f_1)}{f_0^2} = f_0^{-2}\cdot\det\begin{pmatrix} f_0 & f_0' \\ f_1 & f_1'\end{pmatrix} = \frac{f_0f_1' - f_1f_0'}{f_0^2} = \left(\frac{f_1}{f_0}\right)' = g_1,$$

i.e., $g_1 > 0$ on $[a,b]$. The relations for g_i for all $i = 2,\dots,n$ follow by induction from Sylvester's identity [Syl51, AAM96].

(iii): From the definition of f_{n+1} we get immediately $g_{n+1} = D_n\dots D_1D_0f_{n+1}$.

(iv): Follows immediately from (i).

(v): Take the f_{n+1} from (iii).

(vi): For $k = 0$ it is clear that $\{f_i\}_{i=0}^0$ is a T-system since $f_0 > 0$ on $[a,b]$. So assume that for any $f_0,\dots,f_{n-1}$ with $\mathcal{W}(f_0,\dots,f_k) > 0$ on $[a,b]$ for all $k = 0,\dots,n-1$ we have that all $\{f_i\}_{i=0}^k$ with $k = 0,\dots,n-1$ are T-systems. We show that $\{f_i\}_{i=0}^n$ is also a T-system. So let $x_0,\dots,x_n \in [a,b]$ with $x_0 < x_1 < \dots < x_n$. We then have

$$\det\begin{pmatrix} f_0 & \dots & f_n \\ x_0 & \dots & x_n\end{pmatrix} = \det(f_i(x_j))_{i,j=0}^n$$

and factoring out $f_0(x_j) > 0$ in each column gives

$$= \prod_{j=0}^{n} f_0(x_j) \cdot \det \left(\tilde{f}_i(x_j)\right)_{i,j=0}^{n}$$

with $\tilde{f}_i := f_i/f_0$ for all $i = 0, \dots, n$ and substracting from each row its predecessor (the row above) gives

$$= \prod_{j=0}^{n} f_0(x_j) \cdot \det \left(\delta_{0,j}, \tilde{f}_1(x_j) - \tilde{f}_1(x_{j-1}), \dots, \tilde{f}_n(x_j) - \tilde{f}_n(x_{j-1})\right)_{j=0}^{n}.$$

Expanding along the first column and applying the theorem of the mean gives

$$= \prod_{j=0}^{n} f_0(x_j) \cdot \prod_{i=0}^{n-1} (x_{i+1} - x_i) \cdot \det \left(\hat{f}_i(y_j)\right)_{i,j=0}^{n-1}$$

for some $y_0, \dots, y_{n-1}$ with $x_0 < y_0 < x_1 < y_1 < \cdots < y_{n-1} < x_n$ and $\hat{f}_i := (f_{i+1}/f_0)'$ for all $i = 0, \dots, n-1$. The family $\{\hat{f}_i\}_{i=0}^{n-1}$ is the reduced system from Lemma 5.9 and hence by (5.9) we have

$$\mathcal{W}(\hat{f}_0, \dots, \hat{f}_{k-1}) = \frac{\mathcal{W}(f_0, \dots, f_k)}{f_0^{k+1}} > 0$$

on $[a, b]$ for all $k = 1, \dots, n$. By the induction hypothesis we have that $\{\hat{f}_i\}_{i=0}^{n-1}$ is a T-system, i.e.,

$$\det \left(\hat{f}_i(y_j)\right)_{i,j=0}^{n-1} \neq 0 \quad \Rightarrow \quad \det \begin{pmatrix} f_0 & \dots & f_n \\ x_0 & \dots & x_n \end{pmatrix} \neq 0$$

and $\{f_i\}_{i=0}^{n}$ is a T-system which ends the proof. $\square$

The previous lemma is used to characterize all ECT-systems.

5.3 Characterizations of ECT-Systems

We have the following characterization of ECT-systems.

Theorem 5.12 (see e.g. [KS66, p. 376, Thm. 1.1]). *Let $n \in \mathbb{N}_0$ and let $f_0, \dots, f_n \in C^n([a, b], \mathbb{R})$ be with $a < b$. The following are equivalent:*

(i) $\mathcal{F} = \{f_i\}_{i=0}^{n}$ is an ECT-system.
(ii) For all $k = 0, \dots, n$ we have that $\mathcal{W}(f_0, \dots, f_k) \neq 0$ on $[a, b]$.

After adjusting the signs of $f_0, \dots, f_n$ by Remark 5.10 we can in Theorem 5.12 (ii) also assume that $\mathcal{W}(f_0, \dots, f_k) > 0$ on $[a, b]$ for all $k = 0, \dots, n$.

The following proof is adapted from [KS66, pp. 376–379].

Proof. (i) $\Rightarrow$ (ii): Since every ECT-system is also an ET-system the statement is Theorem 5.3 (i) $\Rightarrow$ (ii) because

$$\mathcal{W}(f_0,\dots,f_k)(x) = \begin{pmatrix} f_0 & f_1 & \dots & f_k \\ x & x & \dots & x \end{pmatrix}^*$$

for all $x \in [a,b]$.

(ii) $\Rightarrow$ (i): To show that $\mathcal{F}$ is an ECT-system we have to show that $\{f_i\}_{i=0}^k$ is an ET-system for all $k = 0,\dots,n$. And to show that $\{f_i\}_{i=0}^k$ is an ET-system it is by Theorem 5.3 sufficient to show

$$\det \begin{pmatrix} f_0 & f_1 & \dots & f_k \\ x_0 & x_1 & \dots & x_k \end{pmatrix}^* \neq 0$$

for every $a \leq x_0 \leq x_1 \leq \dots \leq x_k \leq b$. We make two case distinctions:

Case I: All $x_0,\dots,x_k$ are pairwise distinct: $x_0 < x_1 < \dots < x_n$.
Case II: At least once we have $x_j = x_{j+1}$ for some $j = 0,\dots,n-1$.

After renaming $x_0,\dots,x_k$ we can assume $a \leq x_1 < x_2 < \dots < x_l \leq b$ and $m_1,\dots,m_l \in \mathbb{N}$ are the algebraic multiplicities with $m_1 + \dots + m_l = n+1$ for some $l \in \mathbb{N}_0$.

Case I: We have $m_0 = \dots = m_k = 1$ and that is Lemma 5.11 (vi).

Case II: We assume $m_j \geq 2$ for some j. We show that we can reduce the system. We show this reduction by induction over n.

Induction beginning ($n = 0$): Since $\mathcal{W}(f_0)(x) \neq 0$ it is an ET- and an ECT-system. We can assume by changing the sign of f_0 that $f_0 > 0$ on $[a,b]$.

Induction step ($n-1 \to n$): By the induction beginning ($n = 0$) we can assume $f_0 > 0$ on $[a,b]$. Then we have to show that

$$\det \begin{pmatrix} f_0 & f_1 & \dots & f_{m_1-1} & f_{m_1} & \dots & f_n \\ x_1 & x_1 & \dots & x_1 & x_2 & \dots & x_l \end{pmatrix}^* \tag{5.11}$$

is non-zero. To show this we factor $f_0(x_j) > 0$ out of the m_j rows containing x_j in (5.11) for each $j = 0,\dots,l$ to get

$$\det \begin{pmatrix} 1 & \frac{f_0'}{f_0}(x_1) & \dots & \frac{f_0^{(m_1-1)}}{f_0}(x_1) & 1 & \dots & \frac{f_0^{(m_l-1)}}{f_0}(x_l) \\ \frac{f_1}{f_0}(x_1) & \frac{f_1'}{f_0}(x_1) & \dots & \frac{f_1^{(m_1-1)}}{f_0}(x_1) & \frac{f_1}{f_0}(x_2) & \dots & \frac{f_1^{(m_l-1)}}{f_0}(x_l) \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \frac{f_n}{f_0}(x_1) & \frac{f_n'}{f_0}(x_1) & \dots & \frac{f_n^{(m_1-1)}}{f_0}(x_1) & \frac{f_n}{f_0}(x_2) & \dots & \frac{f_n^{(m_l-1)}}{f_0}(x_l) \end{pmatrix}.$$

Then subtract from each of the columns containing x_j a linear combination of its predecessors to obtain for these m_j columns the first m_j columns of $\mathcal{W}(1, f_1/f_0,\dots,f_n/f_0)$ evaluated at x_j:

$$\det\begin{pmatrix} 1 & \left(\frac{f_0}{f_0}\right)'(x_1) & \dots & \left(\frac{f_0}{f_0}\right)^{(m_1-1)}(x_1) & 1 & \dots & \left(\frac{f_0}{f_0}\right)^{(m_l-1)}(x_l) \\ \frac{f_1}{f_0}(x_1) & \left(\frac{f_1}{f_0}\right)'(x_1) & \dots & \left(\frac{f_1}{f_0}\right)^{(m_1-1)}(x_1) & \frac{f_1}{f_0}(x_2) & \dots & \left(\frac{f_1}{f_0}\right)^{(m_l-1)}(x_l) \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \frac{f_n}{f_0}(x_1) & \left(\frac{f_n}{f_0}\right)'(x_1) & \dots & \left(\frac{f_n}{f_0}\right)^{(m_1-1)}(x_1) & \frac{f_n}{f_0}(x_2) & \dots & \left(\frac{f_n}{f_0}\right)^{(m_l-1)}(x_l) \end{pmatrix}$$

$$= \det\begin{pmatrix} 1 & 0 & \dots & 0 & 1 & \dots & 0 \\ \frac{f_1}{f_0}(x_1) & \left(\frac{f_1}{f_0}\right)'(x_1) & \dots & \left(\frac{f_1}{f_0}\right)^{(m_1-1)}(x_1) & \frac{f_1}{f_0}(x_2) & \dots & \left(\frac{f_1}{f_0}\right)^{(m_l-1)}(x_l) \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \frac{f_n}{f_0}(x_1) & \left(\frac{f_n}{f_0}\right)'(x_1) & \dots & \left(\frac{f_n}{f_0}\right)^{(m_1-1)}(x_1) & \frac{f_n}{f_0}(x_2) & \dots & \left(\frac{f_n}{f_0}\right)^{(m_l-1)}(x_l) \end{pmatrix}.$$

The Leibniz rule on differentiation, here for us explicitly

$$\left(\frac{f_i}{f_0}\right)^{(k)} = \sum_{j=0}^{k} \binom{k}{j} \cdot f_i^{(k-j)} \cdot \left(\frac{1}{f_0}\right)^{(j)},$$

ensures that this is always possible.

We then subtract from each column which starts with a 1 its predecessor which also starts with a 1 and apply the mean value theorem to get apart from the positive factor $(x_{j+1} - x_j)$

$$\det\begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \frac{f_1}{f_0}(x_1) & \left(\frac{f_1}{f_0}\right)'(x_1) & \dots & \left(\frac{f_1}{f_0}\right)^{(m_1-1)}(x_1) & \left(\frac{f_1}{f_0}\right)'(y_2) & \left(\frac{f_1}{f_0}\right)'(x_2) & \dots & \left(\frac{f_1}{f_0}\right)^{(m_l-1)}(x_l) \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \\ \frac{f_n}{f_0}(x_1) & \left(\frac{f_n}{f_0}\right)'(x_1) & \dots & \left(\frac{f_n}{f_0}\right)^{(m_1-1)}(x_1) & \left(\frac{f_n}{f_0}\right)'(y_2) & \left(\frac{f_1}{f_0}\right)'(x_2) & \dots & \left(\frac{f_n}{f_0}\right)^{(m_l-1)}(x_l) \end{pmatrix}$$

with $x_1 < y_2 < x_2 < \dots < x_l$ and expanding by the first row gives

$$\det\begin{pmatrix} \left(\frac{f_1}{f_0}\right)'(x_1) & \dots & \left(\frac{f_1}{f_0}\right)^{(m_1-1)}(x_1) & \left(\frac{f_1}{f_0}\right)'(y_2) & \left(\frac{f_1}{f_0}\right)'(x_2) & \dots & \left(\frac{f_1}{f_0}\right)^{(m_l-1)}(x_l) \\ \vdots & & \vdots & \vdots & \vdots & & \\ \left(\frac{f_n}{f_0}\right)'(x_1) & \dots & \left(\frac{f_n}{f_0}\right)^{(m_1-1)}(x_1) & \left(\frac{f_n}{f_0}\right)'(y_2) & \left(\frac{f_1}{f_0}\right)'(x_2) & \dots & \left(\frac{f_n}{f_0}\right)^{(m_l-1)}(x_l) \end{pmatrix}. \tag{5.12}$$

In (5.12) we now have the reduced system $g_i := (f_{i+1}/f_0)'$ with $i = 0, \dots, n-1$ from (5.8) in Lemma 5.9. By (5.9) in Lemma 5.9 and since the reduced systems is of dimension $n-1$ where the inductions hypotheses holds we have that (5.12) is non-zero and hence also (5.11) is non-zero which we wanted to prove. □

Remark 5.13 (see e.g. [KS66, p. 379, Rem. 1.2]). We find the following complete characterization of ECT-systems which requires the additional property (5.13). Fortunately, this seemingly additional property can always be generated by a change of basis vectors, i.e., for any vector space spanned by an ECT-system a suitable basis with (5.13) can be found. ∘

Theorem 5.14 (see e.g. [KS66, p. 379, Thm. 1.2]). *Let $n \in \mathbb{N}_0$ and let $f_0, \dots, f_n \in C^n([a,b],\mathbb{R})$ be such that*

$$f_j^{(k)}(a) = 0 \tag{5.13}$$

holds for all $k = 0,\dots,j-1$ and $j = 1,\dots,n$. After suitable sign changes in $f_0,\dots,f_n$ the following are equivalent:

(i) There exist $g_0,\dots,g_n$ with $g_i \in C^{n-i}([a,b],\mathbb{R})$ and $g_i > 0$ on $[a,b]$ for all $i = 0,\dots,n$ such that

$$\begin{aligned}
f_0(x) &= g_0(x)\\
f_1(x) &= g_0(x)\cdot \int_a^x g_1(y_1)\,\mathrm{d}y_1\\
f_2(x) &= g_0(x)\cdot \int_a^x g_1(y_1)\cdot \int_a^{y_1} g_2(y_2)\,\mathrm{d}y_2\,\mathrm{d}y_1\\
&\ \ \vdots\\
f_n(x) &= g_0(x)\cdot \int_a^x g_1(y_1)\cdot \int_a^{y_1} g_2(y_2)\dots \int_a^{y_{n-1}} g_n(y_n)\,\mathrm{d}y_n\ \dots\ \mathrm{d}y_2\,\mathrm{d}y_1.
\end{aligned}$$

(ii) $\{f_i\}_{i=0}^n$ is an ECT-system on $[a,b]$.
(iii) $\mathcal{W}(f_0,\dots,f_k) > 0$ on $[a,b]$ for all $k = 0,\dots,n$.

If one and therefore all of the equivalent conditions (i) – (iii) hold then the g_i in (i) are given by

$$g_0 := f_0 \quad \text{and} \quad g_i := D_{i-1}\dots D_1 D_0 f_i \quad \text{with} \quad D_i := \frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{f_0}$$

for all $i = 1,\dots,n$ or equivalently by

$$g_0 := f_0, \quad g_1 := \frac{\mathcal{W}(f_0,f_1)}{f_0^2}, \quad \text{and} \quad g_i := \frac{\mathcal{W}(f_0,\dots,f_i)\cdot\mathcal{W}(f_0,\dots,f_{i-2})}{\mathcal{W}(f_0,\dots,f_{i-1})^2}$$

for all $i = 2,\dots,n$.

Proof. "(ii) ⇔ (iii)" is Theorem 5.12, "(iii) ⇒ (i)" is Lemma 5.11 (i) – (iii), and "(i) ⇒ (iii)" is Lemma 5.11 (iv). □

Condition (ii) in Theorem 5.14 is of course to be understood after suitable sign changes in $f_0,\dots,f_n$.

The partial statement Theorem 5.14 (i) ⇒ (ii) can be found e.g. in [KS66, p. 19, Exm. 12] and [KN77, pp. 39–40, P.2.4].

5.4 Examples of ET- and ECT-Systems

An equivalent result as Corollary 4.3 for T-systems, i.e., restricting the domain $\mathcal{X}$ of a T-system leads again to a T-system, also holds for ET- and ECT-systems. We leave that to the reader, see Problem 5.4. Hence, it is sufficient to give (examples of) ET- and ECT-systems with the largest possible domain $\mathcal{X} \subseteq \mathbb{R}$.

While the condition of being an ET-system or being even an ECT-system seems very restrictive, several examples are known.

Example 5.15. Let $n \in \mathbb{N}_0$ and $\mathcal{F} = \{x^i\}_{i=0}^n$. Then $\mathcal{F}$ on $\mathbb{R}$ is an ECT-system. ∘

Proof. Clearly, $\mathcal{F} \subset C^\infty(\mathbb{R},\mathbb{R})$ and every non-trivial $f \in \operatorname{lin}\mathcal{F} = \mathbb{R}[x]_{\leq n}$ has at most n real zeros counting multiplicities by the fundamental theorem of algebra, i.e., $\mathcal{F}$ is an ET-systems. Besides that we have that

$$\mathcal{W}(1,x,x^2,\dots,x^k)(x) = \det\begin{pmatrix} 1 & 0 & 0 & \dots & 0\\ x & 1 & 0 & \dots & 0\\ x^2 & 2x & 2 & \dots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ x^k & kx^k & k(k-1)x^{k-1} & \dots & k! \end{pmatrix} \geq 1$$

holds for all $x \in \mathbb{R}$ and $k = 0,\dots,n$ which shows that $\mathcal{F}$ is also an ECT-system. □

Example 5.16. Let $\mathcal{F} = \{1,x,x^3\}$ on $[0,b]$ with $b > 0$. Then $\mathcal{F}$ is a T-system (see Example 4.15) but not an ET-system. To see this let $x_0 = x_1 = x_2 = 0$, then

$$\begin{pmatrix} f_0 & f_1 & f_2\\ 0 & 0 & 0\end{pmatrix}^* = \begin{pmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}.$$

This shows that $\mathcal{F}$ is not an ET-system. ∘

In the previous example the position $x = 0$ prevents the T-system to be an ET-system. If $x = 0$ is removed then it is even an ECT-system.

Example 5.17. Let $\alpha_0,\dots,\alpha_n \in \mathbb{N}_0$ with $\alpha_0 < \alpha_1 < \dots < \alpha_n$. Then $\mathcal{F} = \{x^{\alpha_i}\}_{i=0}^n$ on $(0,\infty)$ is an ECT-system. For $n = 2m$ and $0 < x_1 < x_2 < \dots < x_m$ we often encounter a specific polynomial structure and hence we write it down explicitly once:

$$\begin{aligned}
&\det\begin{pmatrix} x^{\alpha_0} & x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m-1}} & x^{\alpha_{2m}}\\ x & (x_1 & x_1) & \dots & (x_m & x_m)\end{pmatrix}\\
&= \lim_{\varepsilon\to 0} \varepsilon^{-m}\cdot\det\begin{pmatrix} x^{\alpha_0} & x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m-1}} & x^{\alpha_{2m}}\\ x & x_1 & x_1+\varepsilon & \dots & x_m & x_m+\varepsilon\end{pmatrix}\\
&= \lim_{\varepsilon\to 0}\left[\prod_{i=1}^m (x_i - x)(x_i+\varepsilon - x)\right]\cdot\left[\prod_{1\leq i<j\leq m} (x_j - x_i)^2(x_j - x_i - \varepsilon)(x_j+\varepsilon - x_i)\right]
\end{aligned} \tag{5.14}$$

$$\times\, s_\alpha(x, x_1, x_1+\varepsilon, \ldots, x_m, x_m+\varepsilon)$$
$$= \prod_{i=1}^m (x_i - x)^2 \cdot \prod_{1\leq i<j\leq m} (x_j - x_i)^4 \cdot s_\alpha(x, x_1, x_1, \ldots, x_m, x_m)$$

where s_α is the Schur polynomial of $\alpha = (\alpha_0, \ldots, \alpha_n)$ [Mac95]. Hence,

$$s_\alpha(x, x_1, x_1, \ldots, x_m, x_m)$$

is not divisible by any $(x_i - x)$. ◦

Proof. Combine the induction

$$f^{(m+1)}(x) = \lim_{h\to 0} \frac{f^{(m)}(x+h) - f^{(m)}(x)}{h}$$

and

$$\det\begin{pmatrix} x^{\alpha_0} & \ldots & x^{\alpha_n} \\ x_0 & \ldots & x_n \end{pmatrix} = \prod_{0\leq i<j\leq n} (x_j - x_i) \cdot s_\alpha(x_0, \ldots, x_n)$$

where s_α is the Schur polynomial of $\alpha = (\alpha_0, \ldots, \alpha_n)$. □

With Theorem 5.14 the previous example can be generalized.

Examples 5.18 (Examples 4.16 and 4.17 continued)**.** Let $n \in \mathbb{N}_0$ and let

$$-\infty < \alpha_0 < \alpha_1 < \cdots < \alpha_n < \infty$$

be reals. Then

(a) $\mathcal{F} = \{x^{\alpha_0}, \ldots, x^{\alpha_n}\}$ on $\mathcal{X} = (0, \infty)$ (Example 4.16) and
(b) $\mathcal{G} = \{e^{\alpha_0 x}, \ldots, e^{\alpha_n x}\}$ on $\mathcal{Y} = \mathbb{R}$ (Example 4.17)

are ECT-systems. ◦

Proof. See Problem 5.6. □

In Problem 5.5 we will see that also Example 4.19 are ET- and ECT-systems.

5.5 Representation as a Determinant, Zeros, and Non-Negativity

Similar to Theorem 4.20 we have the following for ET-systems, i.e., knowing n zeros of a polynomial f counting multiplicities determines f uniquely up to a scalar.

Theorem 5.19. *Let $n \in \mathbb{N}_0$ and let $\mathcal{F} = \{f_i\}_{i=0}^n \subseteq C^n([a,b], \mathbb{R})$ be an ET-system. Let $x_1, \ldots, x_n \in [a,b]$ with*

$$x_1 = \cdots = x_{i_1} \;<\; x_{i_1+1} = \cdots = x_{i_1+i_2} \;<\; \ldots \;<\; x_{i_1+\cdots+i_{k-1}+1} = \cdots = x_{i_1+\cdots+i_k=n}$$

for some $k, i_1, \dots, i_k \in \mathbb{N}$ and let $f \in \operatorname{lin} \mathcal{F}$. The following are equivalent:

(i) $f^{(l)}(x_j) = 0$ *for all* $j = 1, \dots, k$ *and* $l = 0, \dots, i_j - 1$.
(ii) There exists a constant $c \in \mathbb{R}$ *such that*

$$f(x) = c \cdot \det \left(\begin{array}{c|cccc} f_0 & f_1 & f_2 & \dots & f_n \\ x & x_1 & x_2 & \dots & x_n \end{array} \right).$$

Proof. (ii) $\Rightarrow$ (i): Clear.

(i) $\Rightarrow$ (ii): If $f = 0$ then $c = 0$ so the assertion holds. If $f \neq 0$ then there exists a point $x_0 \in \mathcal{X} \setminus \{x_1, \dots, x_n\}$ such that $f(x_0) \neq 0$ since $\mathcal{F}$ is an ET-system. Then also the determinant in (ii) is non-zero and we can choose c such that both f and the scaled determinant coincide also in x_0. Since $\mathcal{F}$ is an ET-system we have by Theorem 5.3 that

$$\begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x_0 & x_1 & \dots & x_n \end{pmatrix}^*$$

has full rank, i.e., the coefficients of f and

$$c \cdot \det \left(\begin{array}{c|cccc} f_0 & f_1 & f_2 & \dots & f_n \\ x & x_1 & x_2 & \dots & x_n \end{array} \right)$$

coincide. □

The following result is a strengthened version of Theorem 4.26. It is a small extension of e.g. [KS66, p. 28, Thm. 5.1] with explicit multiplicities of the zeros of a non-negative polynomial.

Theorem 5.20. *Let $n \in \mathbb{N}_0$ and let $\mathcal{F} = \{f_i\}_{i=0}^n$ be an ET-system on $[a, b]$ with $a < b$. Let $x_1 < \dots < x_k$ in $[a, b]$ and let $m_1, \dots, m_k \in \mathbb{N}$ for some $k \in \mathbb{N}$. The following hold:*

(a) If $m_1 + \dots + m_k \leq n$ and $m_i \in 2\mathbb{N}$ for all $x_i \in (a, b)$ then there exists a $f \in \operatorname{lin} \mathcal{F}$ such that

(i) $f \geq 0$ on $[a, b]$,
(ii) f has precisely the zeros $x_1, \dots, x_k$,
(iii) the zeros $x_i \in (a, b)$ of f have multiplicity m_i,
(iv) if $x_1 = a$ then $x_1 = a$ has multiplicity m_1 or $m_1 + 1$, and
(v) if $x_k = b$ then $x_k = b$ has multiplicity m_k or $m_k + 1$.

(b) If $\mathcal{F}$ is an ECT-system or $m_1 + \dots + m_k = n$ then there exists a $f \in \operatorname{lin} \mathcal{F}$ such that

(i) $f \geq 0$ on $[a, b]$,
(ii) f has precisely the zeros $x_1, \dots, x_k$, and
(iii) the zeros x_i of f have multiplicity exactly m_i.

Proof. (a): Set $m := m_1 + \dots + m_k$. If all $x_1, \dots, x_k \in (a, b)$ and $n = m + p$ for some $p \in \mathbb{N}_0$ then the polynomial

$$f(x) = (-1)^p \cdot \det \left(\begin{array}{c|ccccccc} f_0 & f_1 & \dots & f_p & f_{p+1} & \dots & f_{p+m_1} & \dots & f_n \\ x & (a & \dots & a) & (x_1 & \dots & x_1) & \dots & x_k) \end{array} \right)$$
$$+ \det \left(\begin{array}{c|ccccccc} f_0 & f_1 & \dots & f_{m_1} & \dots & f_m & f_{m+1} & \dots & f_n \\ x & (x_1 & \dots & x_1) & \dots & x_k) & (b & \dots & b) \end{array} \right)$$

fulfills the requirements. If $x_1 = a$ and/or $x_k = b$ then include $x_1 = a$ with multiplicity m_1 or $m_1 + 1$ and $x_k = b$ with multiplicity m_k or $m_k + 1$. Use the choice m_1 or $m_1 + 1$ resp. m_k or $m_k + 1$ to let $p \in 2\mathbb{N}_0$ and add y and z with $x_{k-1} < y < z < x_k$. Once construct a polynomial with the zeros $x_1, \dots, x_k, y$ with the corresponding multiplicities and add another polynomial with the zeros $x_1, \dots, x_k, z$ with the corresponding multiplicities to it as above.

(b): Use $\{f_i\}_{i=0}^m$ as the ET-system in (a). □

Problems

5.1 Prove Lemma 5.7.

5.2 Prove Lemma 5.8.

5.3 Prove Lemma 5.9.

5.4 **(a)** Let $n \in \mathbb{N}_0$ and let $\mathcal{F} = \{f_i\}_{i=0}^n$ be an ET-system on $[a, b]$ for some $a < b$. Show that $\mathcal{F}$ on $[a', b']$ with $a < a' < b' < b$ is also an ET-system.

(b) Show (a) for ECT-systems.

5.5 Prove that Example 4.19 is an ECT-system.

5.6 Prove that the Examples 5.18 are ECT-systems.

5.7 Let

$$\mathcal{F} := \{1, x^2, x^3, x^5, x^8, x^{11}, x^{13}, x^{42}\}$$

on $[0, \infty)$. Give an algebraic polynomial $f \in \operatorname{lin} \mathcal{F}$ such that

(a) f is non-negative on $[0, \infty)$,
(b) f has $x_1 = 1$ as a zero with multiplicity $m_1 = 2$,
(c) f has $x_2 = 3$ as a zero with multiplicity $m_2 = 4$, and
(d) f has no zeros in $[0, \infty)$ other than x_1 and x_2.

Chapter 6
Generating ET-Systems from T-Systems by Using Kernels

Life is a short affair;
we should try to make it smooth, and free from strife.

Euripides: The Suppliant Women [Eur13, p. 175]

We have seen that ET- and especially ECT-systems have much nicer properties than T-systems. Therefore, especially for technical reasons, it is desirable to smoothen a T-system into an ET-system. Usually, a function is smoothed by convolution with e.g. the Gaussian kernel. This procedure is also used for T-systems.

The smoothing of T-systems into ET-systems is used in the proof of the main theorem, Karlin's Theorem 7.1. Therein, at first the result is proven for ET-systems and then in a second step the T-system is smoothened into an ET-systems and a limit procedure gives then the statement also for the T-system. Readers only interested in the polynomial cases can skip this chapter, go directly to Chapter 7, and use only the first part of the proof of Karlin's Theorem 7.1 since the polynomials are already ET-systems.

6.1 Kernels

Let $\mathcal{X}$ and $\mathcal{Y}$ be sets and

$$K : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$$

be a bivariate function, also called *kernel*. A family $\{f_i\}_{i=0}^n$ on $\mathcal{Y}$ can then be seen as a special case of K with $\mathcal{X} = \{0, 1, \dots, n\}$, i.e., $f_i = K(i, \cdot)$ for all $i \in \mathcal{X}$. For a kernel K we define the short hand notation

$$K\begin{pmatrix} x_0 & x_1 & \dots & x_n \\ y_0 & y_1 & \dots & y_n \end{pmatrix} := \det(K(x_i, y_j))_{i,j=0}^n. \tag{6.1}$$

Definition 6.1. Let $k \in \mathbb{N}_0$, $\mathcal{X}$ and $\mathcal{Y}$ be ordered sets, and $K : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ be a kernel. The kernel K is called *totally positive* (*of order* k), short (TP_k) property, if for all $i = 0, 1, \dots, k$ we have

$$K\begin{pmatrix} x_1 & x_2 & \dots & x_i \\ y_1 & y_2 & \dots & y_i \end{pmatrix} \geq 0$$

for all $x_1 < x_2 < \cdots < x_i$, $y_1 < y_2 < \cdots < y_i$, and $(x_l, y_m) \in \mathcal{X} \times \mathcal{Y}$ for all $l, m = 1, \ldots, i$. The kernel K is called *strictly totally positive* (*of order* k), short (STP_k), if we always have

$$K\begin{pmatrix} x_1 & x_2 & \ldots & x_i \\ y_1 & y_2 & \ldots & y_i \end{pmatrix} > 0.$$

For more on sign regular kernels see e.g. [Kar68] and [GM96].

Corollary 6.2 (see e.g. [KS66, p. 10, Exm. 3])**.** *Let $n \in \mathbb{N}_0$, let K be a STP_{n+1} kernel with $\mathcal{X} = [a, b]$, $\mathcal{Y} = [c, d]$ and $K(x, \cdot) \in C([c, d], \mathbb{R})$ for all $x \in \mathcal{X}$, and let $x_0 < x_1 < \cdots < x_n$ in $\mathcal{X}$.*

Then $\{K(x_i, \cdot)\}_{i=0}^k$ is a continuous T-system on $\mathcal{Y} = [c, d]$ for all $k = 0, \ldots, n$.

Proof. Follows immediately from Lemma 4.5. □

Definition 6.3. Let $k \in \mathbb{N}$, $\mathcal{X} = [a, b]$, $\mathcal{Y} = [c, d]$, and $K : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ be a kernel such that $K(x, \cdot) \in C^k(\mathcal{Y}, \mathbb{R})$ for all $x \in \mathcal{X}$. We define

$$K^*\begin{pmatrix} x_1 & x_2 & \ldots & x_k \\ y_1 & y_2 & \ldots & y_k \end{pmatrix} := \det \begin{pmatrix} K(x_1, \cdot) & K(x_2, \cdot) & \ldots & K(x_k, \cdot) \\ y_1 & y_2 & \ldots & y_k \end{pmatrix}^* \tag{6.2}$$

for all $x_1 < x_2 < \cdots < x_k$ in $\mathcal{X}$ and $y_1 \leq y_2 \leq \cdots \leq y_k$ in $\mathcal{Y}$.

We say K is *extended totally positive* (*of order* k), short ETP_k, if for all $i = 1, 2, \ldots, k$ we have

$$K^*\begin{pmatrix} x_1 & x_2 & \ldots & x_i \\ y_1 & y_2 & \ldots & y_i \end{pmatrix} > 0$$

for all $x_1 < x_2 < \cdots < x_i$ in $\mathcal{X}$ and $y_1 \leq y_2 \leq \cdots \leq y_i$ in $\mathcal{Y}$.

Corollary 6.4 (see e.g. [KS66, p. 10, Exm. 3])**.** *Let $n \in \mathbb{N}_0$, let K be an ETP_{n+1} kernel with $\mathcal{X} = [a, b]$, $\mathcal{Y} = [c, d]$ and $K(x, \cdot) \in C^n([c, d], \mathbb{R})$ for all $x \in \mathcal{X}$, and let $x_0 < x_1 < \cdots < x_n$ in $\mathcal{X}$.*

Then $\{K(x_i, \cdot)\}_{i=0}^n$ is an ECT-system on $\mathcal{Y} = [c, d]$.

Proof. Follows immediately from Theorem 5.3. □

Example 6.5. Let $\mathcal{X} = \mathbb{R}$, $\mathcal{Y} = [a, b] \subset (0, \infty)$, and $K(x, y) = y^x$. Then K is ETP_k for all $k \in \mathbb{N}_0$. ◦

Proof. Follows immediately from Examples 5.18. □

Example 6.6 (see e.g. [KS66, p. 11, Exm. 5])**.** For any $\sigma > 0$ the Gaussian kernel

$$K_\sigma(x, y) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{x - y}{\sigma}\right)^2\right) \qquad \text{on } \mathcal{X} \times \mathcal{Y} = \mathbb{R}^2 \tag{6.3}$$

is ETP_k for any $k \in \mathbb{N}$.

The proof is adapted from [KS66, p. 11].

Proof. It is sufficient to show that $K(x,y) = e^{-(x-y)^2}$ is ETP$_k$ for all $k \in \mathbb{N}_0$.

In Example 5.18 (b) we have seen that $\{e^{\alpha_i x}\}_{i=0}^n$ is an ECT-system on $\mathbb{R}$ for all $n \in \mathbb{N}_0$ and all $\alpha_0 < \alpha_1 < \cdots < \alpha_n$ in $\mathbb{R}$. Hence, by writing

$$f_n(x) := \sum_{i=0}^n a_n \cdot e^{-(x_i - x)^2} \qquad \text{as} \qquad f_n(x) = e^{-x^2} \cdot \sum_{i=0}^n a_i \cdot e^{-x_i^2} \cdot e^{2x_i x}$$

we see that f_n has at most n zeros (counting multiplicities) in $\mathbb{R}$ if $a_0, \dots, a_n \in \mathbb{R}$ with $a_0^2 + \cdots + a_n^2 > 0$. □

6.2 The Basic Composition Formulas

The following equations (6.4) and (6.6) are the *basic composition formulas*.

Lemma 6.7 (see e.g. [KS66, pp. 13–14, Exm. 8])**.** *Let $K : [a,b] \times [c,d] \to \mathbb{R}$ and $L : [c,d] \times [e,f] \to \mathbb{R}$ be kernels. Let μ be a σ-finite measure such that $M(x,z)$ defined by*

$$M : [a,b] \times [e,f] \to \mathbb{R}, \quad M(x,z) := \int_c^d K(x,y) \cdot L(y,z) \,\mathrm{d}\mu(y)$$

exists for all $(x,z) \in [a,b] \times [e,f]$. The following hold:

(i) M is a kernel.
(ii) For all $k \in \mathbb{N}$, $x_1 < \cdots < x_k$ in $[a,b]$, and $z_1 < \cdots < z_k$ in $[e,f]$ we have

$$M\begin{pmatrix} x_1 & \dots & x_k \\ z_1 & \dots & z_k \end{pmatrix} = \int \cdots \int_{c \le y_1 < \cdots < y_k \le d} K\begin{pmatrix} x_1 & \dots & x_k \\ y_1 & \dots & y_k \end{pmatrix} \cdot L\begin{pmatrix} y_1 & \dots & y_k \\ z_1 & \dots & z_k \end{pmatrix} \,\mathrm{d}\mu(y_1) \dots \mathrm{d}\mu(y_k). \tag{6.4}$$

(iii) If $L(y, \cdot) \in C^{k-1}([e,f], \mathbb{R})$ for some $k \in \mathbb{N}$ and

$$\partial_z^i M(x,z) := \int_c^d K(x,y) \cdot \partial_z^i L(y,z) \,\mathrm{d}\mu(y) \tag{6.5}$$

holds for all $i = 0, \dots, k-1$ then

$$M^*\begin{pmatrix} x_1 & \dots & x_k \\ z_1 & \dots & z_k \end{pmatrix} = \int \cdots \int_{c \le y_1 < \cdots < y_k \le d} K\begin{pmatrix} x_1 & \dots & x_k \\ y_1 & \dots & y_k \end{pmatrix} \cdot L^*\begin{pmatrix} y_1 & \dots & y_k \\ z_1 & \dots & z_k \end{pmatrix} \,\mathrm{d}\mu(y_1) \dots \mathrm{d}\mu(y_k) \tag{6.6}$$

for all $x_1 < \cdots < x_k$ *in* $[a, b]$*, and* $z_1 \leq \cdots \leq z_k$ *in* $[e, f]$*.*

Proof. (i) is clear, (ii) follows by straight forward calculations, see e.g. [PS70, p. 48, No. 68], and (iii) follows from (ii) with (6.5). □

6.3 Smoothing T-Systems into ET-Systems

With the Gaussian kernel from Example 6.6 we get from Lemma 6.7 the following smoothing result.

Corollary 6.8 (see e.g. [KS66, p. 15])**.** *Let* $n \in \mathbb{N}_0$ *and* $\mathcal{F} = \{f_i\}_{i=0}^n$ *be a continuous T-system on* $[a, b]$*. For any* $\sigma > 0$ *let*

$$K_\sigma(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right) \qquad \text{on } \mathcal{X} = \mathbb{R}$$

be the Gaussian kernel and define $f_{i,\sigma} := f_i * K_\sigma$ *for all* $i = 0, \ldots, n$*. Then* $\mathcal{F}_\sigma := \{f_{i,\sigma}\}_{i=0}^n$ *is an ET-system.*

Proof. See Problem 6.1. □

If $\mathcal{F}$ is a continuous T-system on $[a, b]$ then

$$\lim_{\sigma \searrow 0} f_{i,\sigma}(x) = f_i(x)$$

for all $x \in (a, b)$ and $i = 0, \ldots, n$.

Corollary 6.9. *If* $\{f_i\}_{i=0}^k$ *in Corollary 6.8 is a T-system for all* $k = 0, \ldots, n$ *then* $\mathcal{F}_\sigma$ *is an ECT-system.*

Proof. Apply Corollary 6.8 for every $k = 0, 1, \ldots, n$. □

Approximating a T-system by ET-systems with the Gaussian kernel is often used [GK02, Sch53, Kar68], see also [KS66, p. 16]. We will need it in the proof of Karlin's Theorem 7.1.

Problems

6.1 Prove Corollary 6.8 from Lemma 6.7.

Part III
Karlin's Positivstellensätze and Nichtnegativstellensätze

Chapter 7
Karlin's Positivstellensatz and Nichtnegativstellensatz on $[a, b]$

Beauty is the first test: there is no permanent place in this world for ugly mathematics.

Godfrey Harold Hardy [Har69, §10, p. 85]

We now come to the main result (Karlin's Theorem 7.1) and its variations: Karlin's Positivstellensatz 7.3 for T-systems on $[a, b]$ and Karlin's Nichtnegativstellensatz 7.6 for ET-systems on $[a, b]$. Earlier versions were already developed in [KS53]. Both results are used in the following chapters to prove Karlin's Positivstellensatz 8.1 for T-systems on $[0, \infty)$, Karlin's Nichtnegativstellensatz 8.3 for ET-systems on $[0, \infty)$, Karlin's Positivstellensatz 8.4 for T-systems on $\mathbb{R}$, and finally Karlin's Nichtnegativstellensatz 8.5 for ET-systems on $\mathbb{R}$.

The main applications and examples will be the various sparse algebraic Positivstellensätze and sparse algebraic Nichtnegativstellensätze in Part IV.

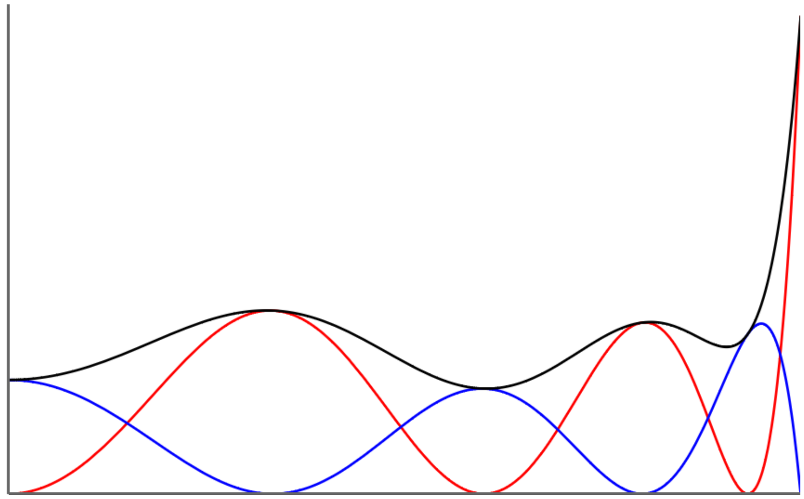
7.1 Karlin's Positivstellensatz for T-Systems on $[a, b]$

For the following main result we remind the reader what it means that a set has an index, see Definition 4.24: If $x \in (a, b)$ then its index is 2 and if $x = a$ or b then its index is 1. The following result is due to Karlin and we name it therefore after him.

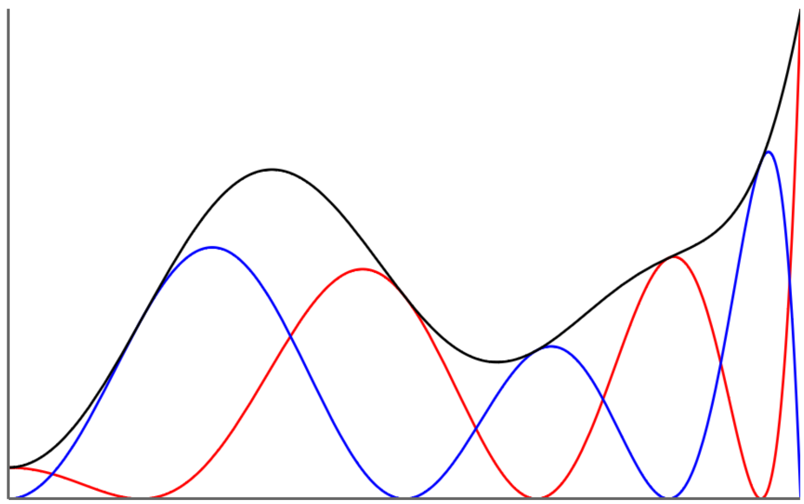
Karlin's Theorem 7.1 (for $f > 0$ on $[a, b]$; [Kar63, Thm. 1] or e.g. [KS66, p. 66, Thm. 10.1])**.** *Let $n \in \mathbb{N}_0$, $\mathcal{F} = \{f_i\}_{i=0}^n$ be a continuous T-system of order n on $[a, b]$ with $a < b$, and let $f \in C([a, b], \mathbb{R})$ with $f > 0$ on $[a, b]$ be a strictly positive continuous function. The following hold:*

(i) There exists a unique polynomial $f_* \in \operatorname{lin} \mathcal{F}$ such that

(a) $f(x) \geq f_*(x) \geq 0$ for all $x \in [a, b]$,
(b) f_* vanishes on a set with index n,
(c) the function $f - f_*$ vanishes at least once between each pair of adjacent zeros of f_*,
(d) the function $f - f_*$ vanishes at least once between the larges zero of f_* and the end point b, and
(e) $f_*(b) > 0$.



(a) $n = 5$



(b) $n = 6$

Fig. 7.1: The functions $f \in C([a, b], \mathbb{R})$ with $f > 0$ (black), $f_* \in \operatorname{lin} \mathcal{F}$ (red), and $f^* \in \operatorname{lin} \mathcal{F}$ (blue) from the Karlin's Positivstellensatz 7.3 with $n = 5$ and $n = 6$.

(ii) There exists a unique polynomial $f^* \in \operatorname{lin} \mathcal{F}$ *which satisfies the conditions (a) to (d) of (i) and*

(e') $f^*(b) = 0$.

Examples of f_* and f^* are depicted in Figure 7.1 for an odd and an even n.

The proof is taken from [KS66, pp. 68–71]. The proof constructs the polynomials f_* and f^* by using the Fixed Point Theorem of Brouwer [Bro11, Satz 4], see also e.g. [Zei86, Prop. 2.6].[1]

Proof. We distinguish three different cases.

Case 1: Let $n = 2m$ and let $\mathcal{F}$ be an ET-system. We construct f_* in (i) as follows. For each point $\xi = (\xi_0, \dots, \xi_m)$ in the m-dimensional simplex

$$\Xi^m := \left\{ (\xi_0, \dots, \xi_m) \in \mathbb{R}^{m+1} \,\middle|\, \xi_i \geq 0,\ i = 0, 1, \dots, m,\ \sum_{i=0}^{m} \xi_i = b - a \right\} \tag{7.1}$$

set

$$x_i := a + \sum_{k=0}^{i-1} \xi_k$$

for all $i = 0, \dots, m$ and define

$$f_\xi(x) := c_\xi \cdot \det \begin{pmatrix} f_0 & f_1 & f_2 & \dots & f_{n-1} & f_n \\ x & x_1 & x_1 & \dots & x_m & x_m \end{pmatrix} \tag{7.2}$$

with $c_\xi \in \mathbb{R}$ such that $f_\xi = \sum_{i=0}^{n} a_i f_i \geq 0$ on $[a, b]$ with $a_0^2 + \dots + a_n^2 = 1$. If p of the points x_i coincide, this common point is to have multiplicity $2p$.

Define

$$\delta_i(\xi) := \min\{\delta \geq 0 \,|\, \delta \cdot f \geq u_\xi \text{ on } [x_i, x_{i+1}]\} \tag{7.3}$$

for all $i = 0, \dots, m$ with $x_0 = a$ and $x_{m+1} = b$. The coefficients $a_i(\xi)$ are continuous in ξ and hence the functions $\delta_i(\xi)$ are continuous in ξ.

Next, define

$$F_i(\xi) := \delta_i(\xi) - \min_k \delta_k(\xi) \tag{7.4}$$

for all $i = 0, \dots, m$ and set $F_{m+1}(\xi) := F_0(\xi)$. If there does not exist a point ξ such that $F_i(\xi) = 0$ for all $i = 0, \dots, m$, then $\sum_{i=0}^{m} F_i(\xi) > 0$ for all $\xi \in \Xi^m$. In this event the continuous mapping

$$\cdot' : \Xi^m \to \Xi^m,\ \xi \mapsto \xi' \quad \text{with} \quad \xi_i' := \frac{F_{i+1}(\xi)}{\sum_{k=0}^{m} F_k(\xi)} \cdot (b - a)$$

for all $i = 0, \dots, m$ is well-defined. The Fixed Point Theorem of Brouwer affirms the existence of a point $\xi^* \in \Xi^m$ for which

$$\xi_i^* := \frac{F_{i+1}(\xi^*)}{\sum_{k=0}^{m} F_k(\xi^*)} \cdot (b - a) \tag{7.5}$$

[1] Note that in [Zei86] the work *Über Abbildungen von Mannigfaltigkeiten* [Bro11] is incorrectly dated in the references and Proposition 2.6 on p. 52 to the year 1912 while the paper actually appeared in 1911 in the *Mathematische Annalen*. However, we also want to point out that Zeidler gives three proofs of the Fixed Point Theorem of Brouwer, including a constructive one in [Zei86, pp. 254–255, Problem 6.7e].

for all $i = 0, \dots, m$. By (7.4) we have that for any $\xi \in \Xi^m$ we have $F_i(\xi) = 0$ for some i. Suppose $F_j(\xi^*) = 0$ for some fixed $j = 0, \dots, m$. Then (7.5) implies $\xi^*_{j-1} = 0$. By (7.3) and (7.4) imply $F_{j-1}(\xi^*) = 0$. Continuing in this way we get $F_i(\xi^*) = 0$ for all $i = 0, \dots, j$ and since $F_{m+1}(\xi) = F_0(\xi)$ we have $F_i(\xi^*) = 0$ for all $i = 0, \dots, m$. But this contradicts our assumption $\sum_{i=0}^m F_i(\xi^*) > 0$. Therefore, there exists at least one point $\xi^* \in \Xi^m$ such that $\delta_i(\xi^*) = \delta$ for all $i = 0, \dots, m$. Since $f_\xi \neq 0$ it follows that $\delta > 0$ and hence all x_i are distinct, i.e.,

$$a = x_0 < x_1 < \cdots < x_m = b.$$

Hence, $f_* := \delta^{-1} \cdot f_{\xi^*}$ by the nature of its construction fulfills the requirements (a) – (e) of (i).

For f^* we let $x_0 = a$ and $x_m = b$ and we define similar to (7.2) the polynomial

$$g_\xi(x) := d_\xi \cdot \det \left(\begin{array}{c|cccccccc} f_0 & f_1 & f_2 & f_3 & \dots & f_{n-2} & f_{n-1} & f_n \\ x & a & x_1 & x_1 & \dots & x_{m-1} & x_{m-1} & b \end{array} \right).$$

Repeating the arguments from above we get f^* which fulfills (a) – (d) and (e') in (ii).

Case 2: Let $n = 2m + 1$ and let $\mathcal{F}$ be an ET-system. Similar to case 1, we define the polynomials

$$f_\xi(x) := d_\xi \cdot \det \left(\begin{array}{c|ccccccc} f_0 & f_1 & f_2 & f_3 & \dots & f_{n-1} & f_n \\ x & a & x_1 & x_1 & \dots & x_m & x_m \end{array} \right).$$

and

$$g_\xi(x) := d_\xi \cdot \det \left(\begin{array}{c|cccccc} f_0 & f_1 & f_2 & \dots & f_{n-2} & f_{n-1} & f_n \\ x & x_1 & x_1 & \dots & x_m & x_m & b \end{array} \right).$$

Repeating the procedure of case 1 gives the statement.

Case 3: Let $n = 2m$ and $\mathcal{F}$ be a T-systems. Then we consider the functions

$$f_i(x;\sigma) := \int_a^b K_\sigma(x, y) \cdot f_i(y)\, \mathrm{d}y$$

where

$$K_\sigma(x, y) := \frac{1}{\sigma \cdot \sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{x - y}{\sigma} \right)^2 \right]$$

with $\sigma > 0$, see Chapter 6. By Corollary 6.8 we have that $\mathcal{F}_\sigma := \{f_i(\,\cdot\,;\sigma)\}_{i=0}^n$ is an ET-system on $[a, b]$ and hence also on any subinterval $[a', b']$ with $a < a' < b' < b$. The need to restrict the system $\mathcal{F}_\sigma$ to the proper interval $[a', b']$ is due to the annoyance that at the end points $x = a$ and $x = b$ we have

$$\lim_{\sigma \searrow 0} f_i(x;\sigma) = \frac{1}{2} f_i(x)$$

for all $i = 0, \dots, n$ while for $x \in (a, b)$ we have

$$\lim_{\sigma \searrow 0} f_i(x; \sigma) = f_i(x).$$

From the cases 1 and 2 we find that for any $\sigma > 0$ we have a polynomial $f_{*,\sigma}$ satisfying conditions (a) – (e) of (i) on the interval $[a', b']$. If

$$f_{*,\sigma} = \sum_{i=0}^{n} a_i(\sigma) \cdot f_i(\,\cdot\,, \sigma)$$

we can chose a sequence $\sigma_k \searrow 0$ and let $x_1^{(k)}, \dots, x_m^{(k)}$ be the zeros of f_{*,σ_k}. Additionally, let $y_1^{(k)}, \dots, y_{m+1}^{(k)}$ be the points which interlace with $\{x_i^{(k)}\}_{i=0}^m$, i.e., $a' < y_1^{(k)} < x_1^{(k)} < \dots < x_m^{(k)} < y_{m+1}^{(k)} \leq b'$ and satisfying $f(y_i^{(k)}) = f_{*,\sigma_k}(y_i^{(k)})$ for all $i = 0, \dots, m+1$.

Since $f(x) \geq f_{*,\sigma} \geq 0$ on $[a', b']$ and solving the system of equations

$$f_{*,\sigma}(x_j) = \sum_{i=0}^{n} a_i(\sigma) \cdot f_i(x_j; \sigma)$$

for $i = 0, \dots, n$ we find that these quantities are uniformly bounded. We now select a subsequence $\{\sigma_{k'}\}$ from $\{\sigma_k\}$ with the property that as $k' \to \infty$ we obtain

$$\begin{aligned}
a_i(\sigma_{k'}) &\to a_i && \text{for all } i = 0, \dots, n,\\
y_j^{(k')} &\to y_j && \text{for all } j = 1, \dots, m+1,\\
x_l^{(k')} &\to x_l && \text{for all } l = 1, \dots, m
\end{aligned}$$

and

$$a' \leq y_1 \leq x_1 \leq \dots \leq x_m \leq y_{m+1} \leq b'.$$

The function $f_{*,a',b'} := \sum_{i=0}^n a_i \cdot f_i$ vanishes at all x_l, $l = 1, \dots, m$, and equals f at all y_j, $j = 1, \dots, m+1$. Therefore, since $f_{*,a',b'}$ is continuous we see that

$$a' \leq y_1 < x_1 < \dots < x_m < y_{m+1} \leq b'.$$

Hence, $f_{*,a',b'}$ satisfies (a) – (e) of (i) on the interval $[a', b']$.

Performing a last limiting procedure letting $a' \searrow a$ and $b' \nearrow b$ we obtain a polynomial f_* satisfying (a) – (e) in (i) on the full interval $[a, b]$.

For f^* the same procedure gives the desired polynomial satisfying the conditions (a) – (d) and (e').

Uniqueness of f_* and f^*: Let $n = 2m$. Observe that if another polynomial $\tilde{f}_*$ with properties (a) – (e) exists then it must have m interior zeros $\tilde{x}_1, \dots, \tilde{x}_m$. Denote by $x_1, \dots, x_m$ the zeros of f_*. Without loss of generality we can assume that either $\tilde{x}_1 < x_1$ or $\tilde{x}_1 = x_1$ and $f_* - \tilde{f}_*$ is non-negative in some interval $(x_1 - \varepsilon, x_1)$. Otherwise we interchange the roles of f_* and $\tilde{f}_*$. We count the zeros of $g := f_* - \tilde{f}_*$. We say g has a zero in the closed interval $[c, d]$ if

- $g(t_0) = 0$ for $t_0 \in (c, d)$,

- $g(c) = 0$ and $g \geq 0$ on $(c, c + \varepsilon)$, or
- $g(d) = 0$ and $g \geq 0$ on $(d - \varepsilon, d)$.

Counting zeros in this fashion we see that g has at least two zeros in each of the intervals $[x_{i-1}, x_i]$ for $i = 1, \dots, m$ where $x_0 = a$ and at least one in the interval $[x_m, b]$. In total g vanishes at least $n + 1$ times. Notice, that certain non-nodal zeros of g have been counted twice and hence by Theorem 4.22 we have $g = 0$.

In a similar way we get uniqueness of f^* and also in the case $n = 2m + 1$. □

Note, in the previous result we do not need to have $f \in \operatorname{lin} \mathcal{F}$. The function f only needs to be continuous and strictly positive on $[a, b]$.

An earlier version of (or at least connected to) Karlin's Theorem 7.1 combined with Theorem 4.22 (which was used in the proof of Karlin's Theorem 7.1) is a lemma by Markov [Mar84], see also [ST43, p. 80].

Lemma 7.2 ([Mar84], see also [ST43, p. 80]). *Let $m \in \mathbb{N}$ and let $f \in C^{n+1}([a, b], \mathbb{R})$ be such that $f > 0$ and $f^{(k)} \geq 0$ for all $k = 1, \dots, m + 1$ in $[a, b]$. Let $p_m \in \mathbb{R}[x]_{\leq m}$ and $c \in (a, b)$. Let $m_1 \in \mathbb{N}$ be the number of zeros in (a, c) of the function $f - p_m$ and m_2 be the number of zeros of p_m in (c, b), both counted with multiplicity. Then $m_1 + m_2 \leq m + 1$.*

Karlin's Theorem 7.1 is of course much more general. As a consequence of Karlin's Theorem 7.1 we get Karlin's Positivstellensatz for T-systems on $[a, b]$.

Karlin's Positivstellensatz 7.3 (for T-Systems on $[a, b]$; see [Kar63, Cor. 1] or e.g. [KS66, p. 71, Cor. 10.1(a)]). *Let $n \in \mathbb{N}_0$, let $\mathcal{F}$ be a continuous T-system of order n on $[a, b]$ with $a < b$, and let $f \in \operatorname{lin} \mathcal{F}$ with $f > 0$ on $[a, b]$. Then there exists a unique representation*

$$f = f_* + f^*$$

with $f_*, f^* \in \operatorname{lin} \mathcal{F}$ such that

(i) $f_*, f^* \geq 0$ on $[a, b]$,
(ii) the zeros of f_* and f^* each are sets of index n,
(iii) the zeros of f_* and f^* strictly interlace,
(iv) $f_*(b) = f(b) > 0$, and
(v) $f^*(b) = 0$.

Proof. Let f_* be the unique f_* from Karlin's Theorem 7.1(i). Then $f - f_* \in \operatorname{lin} \mathcal{F}$ is a polynomial and fulfills (a) – (d), and (e') of f^* in Karlin's Theorem 7.1. But since also f^* is unique we have $f - f_* = f^*$. □

7.2 The Snake Theorem: An Interlacing Theorem

In Karlin's Theorem 7.1 a polynomial $f_* \in \operatorname{lin} \mathcal{F}$ was found with $0 \leq f_* \leq f$ for some given $f \in C([a, b], \mathbb{R})$ with $f > 0$ on $[a, b]$. This can be extended to find a function $f_* \in \operatorname{lin} \mathcal{F}$ between some $g_1, g_2 \in C([a, b], \mathbb{R})$ as the following result shows. In [KS66, p. 368, Thm. 6.1] M. G. Krein and A. A. Nudel'man called it the *Snake Theorem* which is an accurate description of its graphical representation, see Figure 7.2.

Snake Theorem 7.4 ([Kar63, Thm. 2] or e.g. [KS66, p. 72, Thm. 10.2] and [KN77, p. 368, Thm. 6.1])**.** *Let $n \in \mathbb{N}_0$, $\mathcal{F} = \{f_i\}_{i=0}^n$ be a continuous T-system of order n on $[a, b]$ with $a < b$, and let $g_1, g_2 \in C([a, b], \mathbb{R})$ be two continuous functions on $[a, b]$ such that there exists a function $g \in \operatorname{lin} \mathcal{F}$ with*

$$g_1 < g < g_2$$

on $[a, b]$. Then the following hold:

(i) There exists a unique polynomial $f_* \in \operatorname{lin} \mathcal{F}$ such that

(a) $g_1(x) \leq f_*(x) \leq g_2(x)$ for all $x \in [a, b]$, and
(b) there exist $n + 1$ points $x_1 < \cdots < x_{n+1}$ in $[a, b]$ such that

$$f_*(x_{n+1-i}) = \begin{cases} g_1(x_{n+1-i}) & \text{for } i = 1, 3, 5, \dots, \\ g_2(x_{n+1-i}) & \text{for } i = 0, 2, 4, \dots. \end{cases}$$

(ii) There exists a unique polynomial $f^* \in \operatorname{lin} \mathcal{F}$ such that

(a') $g_1(x) \leq f^*(x) \leq g_2(x)$ for all $x \in [a, b]$, and
(b') there exist $n + 1$ points $y_1 < \cdots < y_{n+1}$ in $[a, b]$ such that

$$f^*(y_{n+1-i}) = \begin{cases} g_2(y_{n+1-i}) & \text{for } i = 1, 3, 5, \dots, \\ g_1(y_{n+1-i}) & \text{for } i = 0, 2, 4, \dots. \end{cases}$$

The functions g_1, g_2, g, f_*, and f^* of the Snake Theorem 7.4 are illustrated in Figure 7.2. The following proof is taken from [KS66, p. 73].

Proof. Let $n = 2m$ and $\mathcal{F}$ be an ET-system. We proceed as in the proof of Karlin's Theorem 7.1. For each $\xi = (\xi_0, \dots, \xi_n) \in \Xi^n$ and $\sum_{i=0}^n \xi_i = b - a$ we construct the polynomial

$$f_\xi(x) = \sum_{i=0}^n a_i(\xi) \cdot f_i(x) = c_\xi \cdot \det \begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x & x_1 & \dots & x_n \end{pmatrix}$$

which vanishes at each of the points

$$x_i := a + \sum_{k=0}^{i-1} \xi_k$$

for all $i = 0, \dots, n$ and let $c_\xi \in \mathbb{R}$ be such that $a_0(\xi)^2 + \dots + a_n(\xi)^2 = 1$ and $f_\xi \geq 0$ on $[x_i, x_{i+1}]$ if i is even.

For $i = 0, 2, 4, \dots, n$ we define

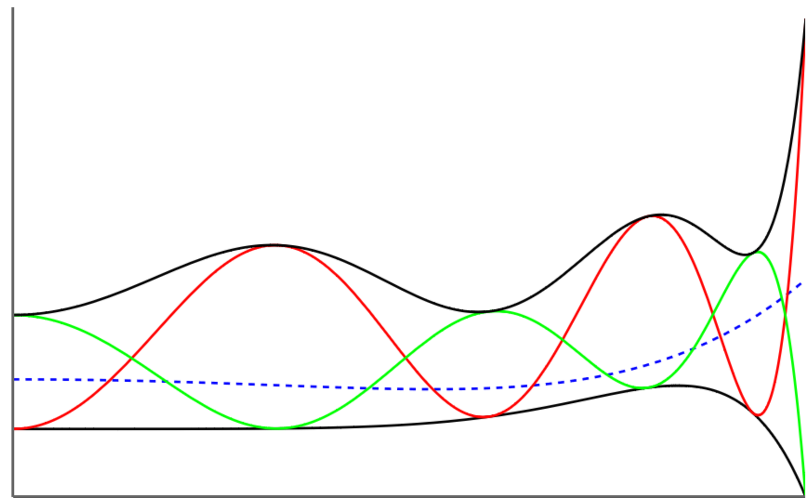


Fig. 7.2: The functions $g_1, g_2 \in C([a,b],\mathbb{R})$ (black, g_1 bottom, g_2 top), $g \in \operatorname{lin}\mathcal{F}$ (blue, dashed), $f_* \in \operatorname{lin}\mathcal{F}$ (red), and $f^* \in \operatorname{lin}\mathcal{F}$ (green) from the Snake Theorem 7.4.

$$\delta_i(\xi) := \min\left\{\delta \geq 0 \,\middle|\, \delta\cdot(g_2 - g) \geq f_\xi \text{ on } [x_i, x_{i+1}]\right\},$$

where $x_0 = a$ and $x_{n+1} = b$, while for $i = 1, 3, \ldots, n-1$ we define

$$\delta_i(\xi) := \min\left\{\delta \geq 0 \,\middle|\, f_\xi \geq \delta\cdot(g - g_1) \text{ on } [x_i, x_{i+1}]\right\}.$$

As in Karlin's Theorem 7.1 we define $F_k(\xi) := \delta_k(\xi) - \min_i \delta_i(\xi)$. And as before assuming $\sum_{k=0}^n F_k(\xi) > 0$ for all $\xi \in \Xi^n$ leads to a contradiction. Therefore, there exists a $\xi^* \in \Xi^n$ for which $\delta_i(\xi^*) = \delta$ for all $i = 0, \ldots, n$. It is clear that $\delta > 0$ and that the polynomial $f_* := \delta^{-1}\cdot f_{\xi^*} + g$ satisfies the conditions of the theorem.

The polynomial f^* is constructed employing the same line of arguments.

The extension encompassing the case where $\mathcal{F}$ is merely a T-system and the proof of the uniqueness proceed as in the proof of Karlin's Theorem 7.1. □

7.3 Karlin's Nichtnegativstellensatz for ET-Systems on $[a, b]$

While Karlin's Theorem 7.1 with $f > 0$ can be proved for T-systems, an equivalent version allowing zeros in $f \in C([a,b],\mathbb{R})$, i.e., $f \geq 0$ but not $f > 0$, needs to assume that $\mathcal{F}$ is an ET-system.

Karlin's Theorem 7.5 (for $f \geq 0$ on $[a, b]$; [Kar63, Thm. 3] or e.g. [KS66, p. 74, Thm. 10.3])**.** *Let $n \in \mathbb{N}_0$, let $\mathcal{F} = \{f_i\}_{i=0}^n$ be a continuous ET-system of order n on*

$[a,b]$ with $a < b$, and let $f \in C^n([a,b],\mathbb{R})$ be such that $f \geq 0$ on $[a,b]$ and f has $r < n$ zeros (counting multiplicities). The following hold:

(i) There exists a unique polynomial $f_* \in \operatorname{lin}\mathcal{F}$ such that

(a) $f(x) \geq f_*(x) \geq 0$ for all $x \in [a,b]$,
(b) f_* has n zeros counting multiplicities,
(c) if $x_1 < \cdots < x_{n-r}$ in (a,b) are the zeros of f_* which remain after removing the r zeros of f then $f - f_*$ vanishes at least twice more (counting multiplicities) in each open interval (x_i, x_{i+1}), $i = 1,\ldots,n-r-1$, and at least once more in each of the intervals $[a, x_1)$ and $(x_{n-r}, b]$,
(d) the zeros $x_1,\ldots,x_{n-r}$ of (c) are a set of index $n-r$, and
(e) $x_{n-r} < b$.

(ii) There exists a unique polynomial $f^* \in \operatorname{lin}\mathcal{F}$ which satisfies the conditions (a) to (d) and

(e') $x_{n-r} = b$.

The proof is taken from [KS66, pp. 74–75].

Proof. Let $z_1,\ldots,z_p$ be the distinct zeros of f with multiplicities $m_1,\ldots,m_p$ where $\sum_{i=1}^p m_i = r \leq n-1$ and set $n' := n - r$. The proof is now similar to the proof of Karlin's Theorem 7.1 where n is replaced by n'. Since the odd and the even cases are again somewhat the same and for the sake of some slight variety we treat now the odd case $n' = 2m' + 1$. The construction of f_* in part (i) proceeds as follows. For each $\xi \in \Xi^{m'}$ we construct the polynomial

$$\begin{aligned} f_\xi(x) &:= \sum_{i=0}^n a_i(\xi)\cdot f_i \\ &= c_\xi \cdot \det\begin{pmatrix} f_0 & f_1 \ldots f_{m_1} & \ldots & f_{m_1+m_{p-1}+1} & \ldots f_r & f_{r+1} & f_{r+2} & \ldots & f_{n-2} & f_{n-1} & f_n \\ x & z_1 \ldots z_1 & \ldots & z_p & \ldots z_p & x_1 & x_1 & \ldots & x_{n'} & x_{n'} & a \end{pmatrix} \end{aligned} \tag{7.6}$$

where $c_\xi \in \mathbb{R}$ is chosen such that $a_1(\xi)^2 + \cdots + a_n(\xi)^2 = 1$ and

$$x_i := a + \sum_{k=0}^i \xi_k$$

for all $i = 1,\ldots,m'$ are the zeros of multiplicity two and a is a zero of multiplicity one. Now we define

$$\delta_i(\xi) := \min\left\{\delta \geq 0 \,\middle|\, \delta \geq \frac{f_\xi}{f} \text{ on } [x_i, x_{i+1}]\right\}$$

for $i = 1,\ldots,m'+1$ with $x_{m'+2} = b$ where the ratio is evaluated by l'Hopital's rule at the zeros $z_1,\ldots,z_p$ of f.

By examining $\frac{f_\xi}{f}$ first in the neighborhood of each of the points $z_1, \ldots, z_p$ and then over the remaining part we find that if $\xi^{(k)} \to \xi$ then

$$\frac{f_{\xi^{(k)}}}{f} \to \frac{f_\xi}{f}$$

uniformly on $[a, b]$. Consequently, each of th δ_i is continuous in ξ and $\delta_i(\xi) = 0$ if and only $\xi_i = 0$.

The same arguments used in the proof of Karlin's Theorem 7.1 now show that for some $\xi^* \in \operatorname{int} \Xi^{m'}$ we have $\delta_i(\xi^*) = \delta > 0$ for all $i = 1, \ldots, m' + 1$. It is simple to see that $f_* := \delta^{-1} \cdot f_{\xi^*}$ possesses the properties (a), (b), (d), and (e) in (i). To show property (c) observe that if $x_i = z_j$ for some j then f_{ξ^*} has a zero at z_j with multiplicity exceeding that of f so that δ is strictly greater than $f_{\xi^*} \cdot f^{-1}$ in some neighborhood of z_j. This implies the equality $\delta = f_{\xi^*}(x) \cdot f(x)^{-1}$ for some x in each of the open intervals $(x_1, x_2), \ldots, (x_{m'}, x_{m'+1})$ and somewhere in $(x_{m'+1}, b]$. Thus, in each (x_i, x_{i+1}), either $f(x) - \delta^{-1} \cdot f_{\xi^*}(x)$ vanishes somewhere other than at the zeros of f or the multiplicity of one of the common zeros of f and $\delta^{-1} \cdot f_{\xi^*}$ is increased by two. In the interval $(x_{m'+1}, b]$ the function $f - \delta^{-1} \cdot f_\xi$ may vanish at b with multiplicity only one greater than the zero of f at this point. This concludes that f_* also fulfills (c) in (i).

The polynomial f^* when $n' = 2m' + 1$ is constructed in the same manner by replacing a in (7.6) by b.

Uniqueness: Assume another polynomial g satisfies the same properties as f_*. Without loss of generality we can assume that the first zero of $f - g$ other than the zeros of f is less than or equal to first zero of $f - f_*$. Define $h := \frac{f_* - g}{f}$. A zero of h occurring at one of the values x_i, $i = 2, \ldots, n' + 1$ is necessarily at least a double zero. In this case we assign one zero to each of the intervals $[x_{i-1}, x_i]$ and $[x_i, x_{i+1}]$ with $x_{m'+2} = b$. Under this counting procedure, and taking account of the oscillation properties of f_* and g, we deduce that h has at least three zeros in $[a, x_1]$, at least two zeros in each of the intervals $[x_i, x_{i+1}]$, $i = 2, \ldots, m'$, and at least one zero in $[x_{m'+1}, b]$. Clearly, all of these zeros are other than the r zeros of f, so that $f_* - g$ has at least $3 + 2(m' - 1) + 1 + r = n + 1$ zeros (counting multiplicities). Hence, $h = 0$ and $f_* = g$. □

If $f \in \operatorname{lin} \mathcal{F}$ in Karlin's Theorem 7.5 we get similar to Karlin's Positivstellensatz 7.3 the following Nichtnegativstellensatz on $[a, b]$ due to Karlin.

Karlin's Nichtnegativstellensatz 7.6 (for ET-Systems on $[a, b]$; [Kar63, p. 603, Cor. after Thm. 3] or e.g. [KS66, p. 76, Cor. 10.3])**.** *Let $n \in \mathbb{N}_0$, $\mathcal{F} = \{f_i\}_{i=0}^n$ be an ET-system of order n on $[a, b]$ with $a < b$, and let $f \in \operatorname{lin} \mathcal{F}$ be such that $f \geq 0$ on $[a, b]$ and f has $r < n$ zeros $a \leq z_1 \leq z_2 \leq \cdots \leq z_r \leq b$ (counting multiplicities). Then there exists a unique representation*

$$f = f_* + f^*$$

with $f_*, f^* \in \operatorname{lin} \mathcal{F}$ such that

(i) $f_*, f^* \geq 0$ *on* $[a, b]$,
(ii) *for* f_* *and* f^* *the sets of zeros counting algebraic multiplicities is after removing the zeros of* f *with algebraic multiplicity a set of index* $n - r$ *which strictly interlace, and*
(iii) *the set of zeros of* f^* *contains after removing the zeros of* f *with algebraic multiplicities the point* b.

Proof. Let f_* be the polynomial from Karlin's Theorem 7.5 and set $g := f - f_*$. Then g fulfills the conditions of f^* in Karlin's Theorem 7.5 and by its uniqueness we have $g = f^*$ which proves the statement. □

Remark 7.7. Since Karlin's Nichtnegativstellensatz 7.6 (ii) might be a little bit confusing we explain it more detailed.

Let $\mathcal{F}$ be an ET-system of order $n \in \mathbb{N}_0$ on $[a, b]$ with $a < b$ and let $f \in \operatorname{lin} \mathcal{F}$ be such that $f \geq 0$ on $[a, b]$ and f has the zeros $z_1, \dots, z_l$ with algebraic multiplicities $m_1, \dots, m_l$, $m_1 + \dots + m_l =: r < n$.

(i) If $n - r = 2m$ is *even* then the zeros of f_* from Karlin's Nichtnegativstellensatz 7.6 are $x_1, \dots, x_m$ all with algebraic multiplicity 2 and the zeros of f^* are $y_0, y_1, \dots, y_m$ where y_0 and y_m have algebraic multiplicity 1 and otherwise the y_i have algebraic multiplicity 2. They interlace, i.e., we have

$$a = y_0 < x_1 < y_1 < \dots < x_m < y_m = b.$$

The f_* and f^* are then given by

$$f_*(x) = c_* \cdot \det \begin{pmatrix} f_0 & f_1 & f_2 & \dots & f_{2m-1} & f_{2m} & f_{2m+1} & \dots & f_n \\ x & x_1 & x_1 & \dots & x_m & x_m & z_1 & \dots & z_l \end{pmatrix}$$

and

$$f^*(x) = c^* \cdot \det \begin{pmatrix} f_0 & f_1 & f_2 & f_3 & \dots & f_{2m-2} & f_{2m-1} & f_{2m} & f_{2m+1} & \dots & f_n \\ x & a & y_1 & y_1 & \dots & y_{m-1} & y_{m-1} & b & z_1 & \dots & z_l \end{pmatrix}$$

where $c_*, c^* \in \mathbb{R} \setminus \{0\}$ and the signs are such that $f_*, f^* \geq 0$ on $[a, b]$. The zeros $z_1, \dots, z_l$ are included with their corresponding algebraic multiplicities $m_1, \dots, m_l$, i.e., z_1 is included m_1-times, ..., z_l is included m_l-times.

(ii) If $n - r = 2m + 1$ is *odd* then the zeros of f_* from Karlin's Nichtnegativstellensatz 7.6 are $x_0, \dots, x_m$ where x_0 has algebraic multiplicity 1 and the other algebraic multiplicity 2. For f^* we have the zeros $y_0, \dots, y_m$ where $y_0, \dots, y_{m-1}$ have algebraic multiplicity 2 and y_m has algebraic multiplicity 1. They interlace, i.e., we have

$$x_0 = a < y_0 < \dots < x_m < y_m = b.$$

The f_* and f^* are then given by

$$f_*(x) = c_* \cdot \det \begin{pmatrix} f_0 & f_1 & f_2 & f_3 & \dots & f_{2m} & f_{2m+1} & f_{2m+2} & \dots & f_n \\ x & a & x_1 & x_1 & \dots & x_m & x_m & z_1 & \dots & z_l \end{pmatrix}$$

and

$$f^*(x) = c^* \cdot \det \begin{pmatrix} f_0 & f_1 & f_2 & \dots & f_{2m-1} & f_{2m} & f_{2m+1} & f_{2m+2} & \dots & f_n \\ x & y_0 & y_0 & \dots & y_{m-1} & y_{m-1} & b & z_1 & \dots & z_l \end{pmatrix}$$

where $c_*, c^* \in \mathbb{R} \setminus \{0\}$ and the signs are such that $f_*, f^* \geq 0$ on $[a, b]$. The zeros $z_1, \dots, z_l$ are included with their corresponding algebraic multiplicities $m_1, \dots, m_l$, i.e., z_1 is included m_1-times, ..., z_l is included m_l-times. ∘

With the proof of Karlin's Theorem 7.5 one can prove a similar interlacing theorem as the Snake Theorem 7.4 when $g_2 - g_1$ has a certain number of zeros, see [KS66, p. 76, Rem. 10.1].

We stated here Karlin's Positivstellensatz 7.3 and Karlin's Nichtnegativstellensatz 7.6 for functions on $[a, b]$. There are also similar statements for periodic functions, see [Kar63, Thm. 6 and 7]. The cases on $[0, \infty)$ and $\mathbb{R}$ are given in the next chapter.

Problems

7.1 Examine the proof of Karlin's Theorem 7.5 more closely. In the statement of the theorem it is required that $\mathcal{F}$ is an ET-system on $[a, b]$. But for a given $f \geq 0$ where does the family $\mathcal{F}$ actually only needs to be an ET-system?

Chapter 8
Karlin's Positivstellensätze and Nichtnegativstellensätze on $[0, \infty)$ and $\mathbb{R}$

Look at the conclusion! And try to think of a familiar theorem having the same or a similar conclusion.

George Pólya [Pól45, p. 25]

In this chapter we extend the results of the previous chapter, i.e., we extend Karlin's Positivstellensatz 7.3 on $[a, b]$ to $[0, \infty)$ in Karlin's Positivstellensatz 8.1 and to $\mathbb{R}$ in Karlin's Positivstellensatz 8.4 as well as we extend Karlin's Nichtnegativstellensatz 7.6 on $[a, b]$ to $[0, \infty)$ in Karlin's Nichtnegativstellensatz 8.3 and to $\mathbb{R}$ in Karlin's Nichtnegativstellensatz 8.5.

8.1 Karlin's Positivstellensatz for T-Systems on $[0, \infty)$

By a transformation of $[a, b]$ to $[0, \infty]$ and then restriction to $[0, \infty)$ we get from Karlin's Positivstellensatz 7.3 the following.

Karlin's Positivstellensatz 8.1 (for T-Systems on $[0, \infty)$; see [Kar63, Thm. 9] or e.g. [KS66, p. 169, Thm. 8.1]). *Let $n \in \mathbb{N}_0$ and $\mathcal{F} = \{f_i\}_{i=0}^n$ be a continuous T-system of order n on $[0, \infty)$ such that*

(a) there exists a $C > 0$ such that $f_n(x) > 0$ for all $x \geq C$,
(b) $\lim\limits_{x\to\infty} \dfrac{f_i(x)}{f_n(x)} = 0$ for all $i = 0, \dots, n-1$, and
(c) $\{f_i\}_{i=0}^{n-1}$ is a continuous T-system on $[0, \infty)$.

Then for any $f = \sum_{i=0}^n a_i f_i \in \operatorname{lin} \mathcal{F}$ with $f > 0$ on $[0, \infty)$ and $a_n > 0$ there exists a unique representation

$$f = f_* + f^*$$

with $f_*, f^* \in \operatorname{lin} \mathcal{F}$ and $f_*, f^* \geq 0$ on $[0, \infty)$ such that the following hold:

(i) If $n = 2m$ the polynomials f_* and f^* each possess m distinct zeros $\{x_i\}_{i=1}^m$ and $\{y_i\}_{i=0}^{m-1}$ satisfying

$$0 = y_0 < x_1 < y_1 < \cdots < y_{m-1} < x_m < \infty.$$

All zeros except y_0 are double zeros.

(ii) If $n = 2m+1$ the polynomials f_* and f^* each possess the zeros $\{x_i\}_{i=1}^{m+1}$ and $\{y_i\}_{i=1}^{m}$ satisfying

$$0 = x_1 < y_1 < x_2 < \dots < y_m < x_{m+1} < \infty.$$

All zeros except x_1 are double zeros.

(iii) The coefficient of f_n in f_* is equal to a_n.

The proof is adapted from [KS66, pp. 168].

Proof. By (a) there exists a function $w \in C([0,\infty),\mathbb{R})$ such that $w > 0$ on $[0,\infty)$ and $\lim_{x\to\infty} \frac{f_n(x)}{w(x)} = 1$. By (b) we define

$$v_i(x) := \begin{cases} \frac{f_i(x)}{w(x)} & \text{if } x \in [0,\infty),\\ \delta_{i,n} & \text{if } x = \infty \end{cases}$$

for all $i = 0,1,\dots,n$. Then by (c) and Corollary 4.9 we have that $\{v_i\}_{i=0}^n$ is a T-system on $[0,\infty]$. With $t(x) := \tan(\pi x/2)$ we define $g_i(x) := v_i \circ t$ for all $i = 0,1,\dots,n$. Hence, $\mathcal{G} = \{g_i\}_{i=0}^n$ is a T-system on $[0,1]$ by Corollary 4.8. We now apply Karlin's Positivstellensatz 7.3 to $\mathcal{G}$. Set $g := \left(\frac{f}{w}\right) \circ t$.

(i): Let $n = 2m$. Then by Karlin's Positivstellensatz 7.3 there exist points

$$0 = y_0 < x_1 < y_1 < \dots < x_m < y_m = 1$$

and unique functions g_* and g^* such that $g = g_* + g^*$, $g_*, g^* \geq 0$ on $[0,1]$, $x_1,\dots,x_m$ are the zeros of g_*, and $y_0,\dots,y_m$ are the zeros of g^*. Then $f_* := (g_* \circ t^{-1}) \cdot w$ and $f^* := (g^* \circ t^{-1}) \cdot w$ are the unique components in the decomposition $f = f_* + f^*$.

(ii): Similar to (i).

(iii): From (i) (and (ii) in a similar way) we have $g_i(1) = 0$ for $i = 0,\dots,n-1$ and $g_n(1) = 1$. Hence, we get with $g^*(y_m = 1) = 0$ that g_n is not contained in g^*, i.e., g_* has the only g_n contribution because $\mathcal{G}$ is linearly independent. This is inherited by f_* and f^* which proves (iii). □

The transformation $g_i = v_i \circ t$ with t the tan-function is due to Krein [Kre51].

If $\mathcal{F}$ in Karlin's Positivstellensatz 8.1 is an ET-system then the f_* and f^* can be written down explicitly. For that we only need $\mathcal{F}$ to be an ET-system on $(0,\infty)$ not on all $[0,\infty)$ since at $x = 0$ a possible zero in f_* or f^* only has multiplicity one.

Corollary 8.2. *If in Karlin's Positivstellensatz 8.1 we have additionally that $\mathcal{F}$ is an ET-system on $(0,\infty)$ then the unique f_* and f^* are given*

(i) for $n = 2m$ by

$$f_*(x) = c^* \cdot \det \begin{pmatrix} f_0 & f_1 & f_2 & \dots & f_{2m-1} & f_{2m} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) \end{pmatrix}$$

and

$$f^*(x) = -c_* \cdot \det \begin{pmatrix} f_0 & f_1 & f_2 & f_3 & \dots & f_{2m-2} & f_{2m-1} \\ x & y_0 & (y_1 & y_1) & \dots & (y_{m-1} & y_{m-1}) \end{pmatrix},$$

(ii) and for $n = 2m+1$ by

$$f_*(x) = -c_* \cdot \det \begin{pmatrix} f_0 & f_1 & f_2 & f_3 & \dots & f_{2m} & f_{2m+1} \\ x & x_1 & (x_2 & x_2) & \dots & (x_{m+1} & x_{m+1}) \end{pmatrix}$$

and

$$f^*(x) = c^* \cdot \det \begin{pmatrix} f_0 & f_1 & f_2 & \dots & f_{2m-1} & f_{2m} \\ x & (y_1 & y_1) & \dots & (y_m & y_m) \end{pmatrix}$$

for some $c_*, c^* > 0$.

Proof. Combine Karlin's Positivstellensatz 8.1 with Remark 4.28 and note that since 0 is never a multiple zero we only need $\mathcal{F}$ to be an ET-system on $(0,\infty)$. □

8.2 Karlin's Nichtnegativstellensatz for ET-Systems on $[0,\infty)$

In Karlin's Positivstellensatz 8.1 we needed to transform the domain $[a,b]$ into $[0,\infty]$ of a T-system. For Karlin's Nichtnegativstellensatz 8.3 we needed an ET-system because of the additional zeros from $f \geq 0$.

With the same technique as in the proof of Karlin's Positivstellensatz 8.1 and Lemma 5.8 we get from Karlin's Nichtnegativstellensatz 7.6 the following.

Karlin's Nichtnegativstellensatz 8.3 (for ET-Systems on $[0,\infty)$)**.** *Let $n \in \mathbb{N}_0$ and $\mathcal{F} = \{f_i\}_{i=0}^n$ be an ET-system of order n on $[0,\infty)$ such that*

(a) there exists a $C > 0$ such that $f_n(x) > 0$ for all $x \geq 0$,

(b) $\lim\limits_{x\to\infty} \dfrac{f_i(x)}{f_n(x)} = 0$ for all $i = 0,\dots,n-1$, and

(c) $\{f_i\}_{i=0}^{n-1}$ is an ET-system.

Then for any $f = \sum_{i=0}^n a_i f_i \in \operatorname{lin}\mathcal{F}$ such that $f \geq 0$ on $[0,\infty)$, $a_n > 0$, and f has $r < n$ zeros counting multiplicity there exists a unique representation

$$f = f_* + f^*$$

with $f_*, f^* \in \operatorname{lin}\mathcal{F}$ such that the following hold:

(i) $f_*,\ f^* \geq 0$ on $[0,\infty)$,
(ii) f_* has n zeros (counting multiplicities),
(iii) f^* has $n-1$ zeros (counting multiplicities),
(iv) the zeros of f_* and f^* strictly interlace if the zeros of f are removed, and
(v) the coefficient of f_n in f_* is equal to a_n.

Proof. The conditions (a) – (c) are such that $\mathcal{F}$ on $[0,\infty]$, i.e., including ∞, is an ET-system.

With the same argument as in the proof of Karlin's Positivstellensatz 8.1 we transform $\mathcal{F}$ on $[0, \infty]$ into $\mathcal{G}$ on $[0, 1]$ with the tan-function. Here Lemma 5.8 ensures that also $\mathcal{G}$ is an ET-system.

Application of Karlin's Nichtnegativstellensatz 7.6 on $[0, 1]$ gives the desired decomposition $g = g_* + g^*$ with the observation that $x = 1$ is a zero of at most multiplicity one by (a) and (b). Backwards transformation into $\mathcal{F}$ on $[0, \infty]$ resp. $[0, \infty)$ then gives the assertion. □

8.3 Karlin's Positivstellensatz for T-Systems on $\mathbb{R}$

We have seen that from Karlin's Positivstellensatz 7.3 on $[a, b]$ we get Karlin's Positivstellensatz 8.1 on $[0, \infty)$ with the transformation $t(x) = \tan(\pi x/2)$ from $[0, 1]$ to $[0, \infty]$ and only need to pay attention to the end point $x = 1$ resp. $x = \infty$. The same transformation however also applies going from $[-1, 1]$ to $[-\infty, \infty]$ now paying attention to both end points.

Karlin's Positivstellensatz 8.4 (for T-Systems on $\mathbb{R}$; see [Kar63, Thm. 10] or e.g. [KS66, p. 198, Thm. 8.1])**.** *Let $m \in \mathbb{N}_0$ and $\mathcal{F} = \{f_i\}_{i=0}^{2m}$ be a continuous T-system of order $2m$ on $\mathbb{R}$ such that*

(a) there exists a $C > 0$ such that $f_{2m}(x) > 0$ for all $x \in (-\infty, -C] \cup [C, \infty)$,
(b) $\lim\limits_{|x| \to \infty} \dfrac{f_i(x)}{f_{2m}(x)} = 0$ for all $i = 0, \dots, 2m - 1$, and
(c) $\{f_i\}_{i=0}^{2m-1}$ is a continuous T-system of order $2m - 1$ on $\mathbb{R}$.

Let $f = \sum_{i=0}^{2m} a_i f_i$ be such that $f > 0$ on $\mathbb{R}$ and $a_{2m} > 0$. Then there exists a unique representation

$$f = f_* + f^*$$

with $f_*, f^* \in \operatorname{lin} \mathcal{F}$ and $f_*, f^* \geq 0$ on $\mathbb{R}$ such that

(i) the coefficient of f_{2m} in f_* is a_{2m}, and
(ii) f_* and f^* are non-negative polynomials having zeros $\{x_i\}_{i=1}^{m}$ and $\{y_i\}_{i=1}^{m-1}$ with

$$-\infty < x_1 < y_1 < x_2 < \cdots < y_{m-1} < x_m < \infty.$$

Proof. See Problem 8.1. □

8.4 Karlin's Nichtnegativstellensatz for ET-Systems on $\mathbb{R}$

On $\mathbb{R}$ we have the following Nichtnegativstellensatz for ET-systems.

Karlin's Nichtnegativstellensatz 8.5 (for ET-Systems on $\mathbb{R}$)**.** *Let $m \in \mathbb{N}_0$ and $\mathcal{F} = \{f_i\}_{i=0}^{2m}$ be an ET-system of order $2m$ on $\mathbb{R}$ such that* *(a)* *there exists a* $C > 0$ *such that* $f_{2m} > 0$ *for all* $x \in (-\infty, -C] \cup [C, \infty)$,

(b) $\lim_{|x|\to\infty} \frac{f_i(x)}{f_{2m}(x)} = 0$ *for all* $i = 0, \ldots, 2m-1$,

(c) $\{f_i\}_{i=0}^{n-1}$ *is an ET-system of order* $n-1$ *on* $\mathbb{R}$.

Let $f = \sum_{i=0}^{2m} a_i f_i \in \operatorname{lin} \mathcal{F}$ *be such that* $f \geq 0$, $a_{2m} > 0$, *and* f *has* $r < n$ *zeros counting multiplicities. Then there exists a unique representation*

$$f = f_* + f^*$$

with $f_*, f^* \in \operatorname{lin} \mathcal{F}$ *such that the following hold:*

(i) $f_*, f^* \geq 0$ *on* $\mathbb{R}$,
(ii) f_* *has* $2m$ *zeros counting multiplicity,*
(iii) f^* *has* $2m-2$ *zeros counting multiplicity,*
(iv) *the zeros of* f_* *and* f^* *strictly interlace if the zeros of* f *are removed, and*
(v) *the coefficient of* f_n *in* f_* *is equal to* a_n.

Proof. See Problem 8.2. □

Problems

8.1 Prove Karlin's Positivstellensatz 8.4, i.e., adapt the proof of Karlin's Positivstellensatz 8.1 such that both interval ends a and b of $[a, b]$ are mapped to $-\infty$ and $+\infty$, respectively.

8.2 Prove Karlin's Nichtnegativstellensatz 8.5, i.e., adapt the proof of Karlin's Nichtnegativstellensatz 8.3 such that both interval ends a and b of $[a, b]$ are mapped to $-\infty$ and $+\infty$, respectively.

Part IV
Non-Negative Algebraic Polynomials on $[a, b]$, $[0, \infty)$, and $\mathbb{R}$

Chapter 9
Non-Negative Algebraic Polynomials on $[a, b]$

I hold that it is only when we can prove everything we assert that we understand perfectly the thing under consideration.

Gottfried Wilhelm Leibniz [Lei89]

We developed in the previous chapters the Positiv- and Nichtnegativestellensätze for general T- and ET-systems due to Karlin. We will now apply these to the algebraic polynomials, i.e., we will plug in Example 5.15 and Example 5.17.

9.1 Sparse Algebraic Positivstellensatz on $[a, b]$

At first let us have a look how all sparse strictly positive polynomials on some interval $[a, b] \subseteq (0, \infty)$ look like.

Theorem 9.1 (Sparse Algebraic Positivstellensatz on $[a, b]$ with $0 < a < b$)**.** *Let $n \in \mathbb{N}_0$, $\alpha_0, \dots, \alpha_n \in \mathbb{R}$ be real numbers with $\alpha_0 < \alpha_1 < \dots < \alpha_n$, and let $\mathcal{F} = \{x^{\alpha_i}\}_{i=0}^n$. Then for any $f = \sum_{i=0}^n a_i x^{\alpha_i} \in \operatorname{lin} \mathcal{F}$ with $f > 0$ on $[a, b]$ and $a_n > 0$ there exists a unique decomposition*

$$f = f_* + f^*$$

with $f_*, f^* \in \operatorname{lin} \mathcal{F}$ such that

(i) for $n = 2m$ there exist points $x_1, \dots, x_m, y_1, \dots, y_{m-1} \in [a, b]$ with

$$a < x_1 < y_1 < \dots < x_m < b$$

and constants $c_*, c^* > 0$ with

$$f_*(x) = c_* \cdot \det \begin{pmatrix} x^{\alpha_0} & x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m-1}} & x^{\alpha_{2m}} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) \end{pmatrix} \geq 0 \tag{9.1}$$

and

$$f^*(x) = -c^* \cdot \det \begin{pmatrix} x^{\alpha_0} & x^{\alpha_1} & x^{\alpha_2} & x^{\alpha_3} & \dots & x^{\alpha_{2m-2}} & x^{\alpha_{2m-1}} & x^{\alpha_{2m}} \\ x & a & (y_1 & y_1) & \dots & (y_{m-1} & y_{m-1}) & b \end{pmatrix} \geq 0 \tag{9.2}$$

for all $x \in [a,b]$, *or*

(ii) for $n = 2m+1$ *there exist points* $x_1, \dots, x_m, y_1, \dots, y_m \in [a,b]$ *with*

$$a < y_1 < x_1 < \dots < y_m < x_m < b$$

and $c_*, c^* > 0$ *with*

$$f_*(x) = -c_* \cdot \det \begin{pmatrix} x^{\alpha_0} & x^{\alpha_1} & x^{\alpha_2} & x^{\alpha_3} & \dots & x^{\alpha_{2m}} & x^{\alpha_{2m+1}} \\ x & a & (x_1 & x_1) & \dots & (x_m & x_m) \end{pmatrix} \geq 0 \tag{9.3}$$

and

$$f^*(x) = c^* \cdot \det \begin{pmatrix} x^{\alpha_0} & x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m-1}} & x^{\alpha_{2m}} & x^{\alpha_{2m+1}} \\ x & (y_1 & y_1) & \dots & (y_m & y_m) & b \end{pmatrix} \geq 0 \tag{9.4}$$

for all $x \in [a,b]$.

Proof. By Example 5.17 we have that $\mathcal{F}$ on $[a,b]$ is an ET-system. Hence, Karlin's Positivstellensatz 7.3 applies. We check both cases $n = 2m$ and $n = 2m+1$ separately.

$n = 2m$: By Karlin's Positivstellensatz 7.3 we have that the zero set $\mathcal{Z}(f^*)$ of f^* has index $2m$ and contains b with index 1, i.e., $a \in \mathcal{Z}(f^*)$ and all other zeros have index 2. Hence, $\mathcal{Z}(f^*) = \{a = y_0 < y_1 < \dots < y_{m-1} < y_m = b\}$. By Karlin's Positivstellensatz 7.3 we have that $\mathcal{Z}(f_*)$ also has index $2m$ and the zeros of f_* and f^* interlace. Then the determinantal representations of f_* and f^* follow from Remark 4.28.

$n = 2m+1$: By Karlin's Positivstellensatz 7.3 we have that $b \in \mathcal{Z}(f^*)$ and since the index of $\mathcal{Z}(f^*)$ is $2m+1$ we have that there are only double zeros $y_1, \dots, y_m \in (a,b)$ in $\mathcal{Z}(f^*)$. Similar we find that $a \in \mathcal{Z}(f_*)$ since its index is odd and only double zeros $x_1, \dots, x_m \in (a,b)$ in $\mathcal{Z}(f_*)$ remain. By Karlin's Positivstellensatz 7.3 (iii) the zeros x_i and y_i strictly interlace and the determinantal representation of f_* and f^* follow again from Remark 4.28. □

Note, if $\alpha_0, \dots, \alpha_n \in \mathbb{N}_0$ then by Example 5.17 equation (5.14) the algebraic polynomials f_* and f^* in (9.1) – (9.4) can be written down with Schur polynomials.

Remark 9.2. The condition $a_n > 0$ in Theorem 9.1 is no restriction. The result also holds for $a_n < 0$ as long as $f > 0$ on $[a,b]$. Since $[a,b]$ is compact the polynomials x^{α_i} are bounded. In the definition of a T-system the order of the functions f_i can be altered since only any linear combination has to have at most n zeros. Hence, in a $f > 0$ at least one coefficient a_i is larger then zero and we interchange f_i with f_n. A possible sign change in the f_* and f^* in (9.1) – (9.4) might appear. ∘

Theorem 9.1 does not hold for $a = 0$ and $\alpha_0 > 0$ or $\alpha_0, \dots, \alpha_k < 0$. In case $\alpha_0 > 0$ the determinantal representations of f^* for $n = 2m$ and f_* for $n = 2m+1$ are the zero polynomials. In fact, in this case $\mathcal{F}$ is not even a T-system since in Lemma 4.5 the determinant contains a zero column if $x_0 = 0$. We need to have $\alpha_0 = 0$ ($x^{\alpha_0} = 1$) to let $a = 0$. For $\alpha_0, \dots, \alpha_k < 0$ we have singularities at $x = 0$ and hence no T-system.

Corollary 9.3. *If* $\alpha_0 = 0$ *in Theorem 9.1 then Theorem 9.1 also holds with* $a = 0$.

Proof. The determinantal representations of f_* for $n = 2m + 1$ and f^* for $n = 2m$ in Theorem 9.1 continuously depend on a. It is sufficient to show that these representations are non-trivial (not the zero polynomial) for $a = 0$. We show this for f_* in case (ii) $n = 2m + 1$. The other cases are equivalent.

We have that $\mathcal{F}$ is a T-system on $[0, b]$ with $b > 0$. For $\varepsilon > 0$ small enough we set

$$\begin{aligned} g_\varepsilon(x) &= -\varepsilon^{-m} \cdot \det \begin{pmatrix} 1 & x^{\alpha_1} & x^{\alpha_2} & x^{\alpha_3} & \dots & x^{\alpha_{2m}} & x^{\alpha_{2m+1}} \\ x & 0 & x_1 & x_1 + \varepsilon & \dots & x_m & x_m + \varepsilon \end{pmatrix} \\ &= -\varepsilon^{-m} \cdot \det \begin{pmatrix} 1 & x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m+1}} \\ 1 & 0 & 0 & \dots & 0 \\ 1 & x_1^{\alpha_1} & x_1^{\alpha_2} & \dots & x_1^{\alpha_{2m+1}} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & (x_m+\varepsilon)^{\alpha_1} & (x_m+\varepsilon)^{\alpha_2} & \dots & (x_m+\varepsilon)^{\alpha_{2m+1}} \end{pmatrix} \end{aligned}$$

develop with respect to the second row

$$\begin{aligned} &= \varepsilon^{-m} \cdot \det \begin{pmatrix} x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m-1}} \\ x_1^{\alpha_1} & x_1^{\alpha_2} & \dots & x_1^{\alpha_{2m-1}} \\ \vdots & \vdots & & \vdots \\ (x_m+\varepsilon)^{\alpha_1} & (x_m+\varepsilon)^{\alpha_2} & \dots & (x_m+\varepsilon)^{\alpha_{2m+1}} \end{pmatrix} \\ &= \varepsilon^{-m} \cdot \det \begin{pmatrix} x^{\alpha_1} & x^{\alpha_2} & x^{\alpha_3} & \dots & x^{\alpha_{2m}} & x^{\alpha_{2m+1}} \\ x & x_1 & x_1 + \varepsilon & \dots & x_m & x_m + \varepsilon \end{pmatrix}. \end{aligned}$$

Then $x_1, x_1 + \varepsilon, \dots, x_m, x_m + \varepsilon \in (0, b]$, i.e., $\{x^{\alpha_i}\}_{i=1}^n$ is an ET-system on $[a', b]$ with $0 = a < a' < x_1$, see Example 5.17. By Remark 4.28 the limit $\varepsilon \searrow 0$ is not the zero polynomial which ends the proof. $\square$

Remark 9.4. It is clear that if $\alpha_0 > 0$ then we can just factor out x^{α_0}

$$f(x) = a_0 x^{\alpha_0} + a_1 x^{\alpha_1} + \dots + a_n x^{\alpha_n} = x^{\alpha_0} \cdot (\underbrace{a_0 + a_1 x^{\alpha_1 - \alpha_0} + \dots + a_n x^{\alpha_n - \alpha_0}}_{=:\tilde{f}(x)})$$

and apply Theorem 9.1 or Corollary 9.3 to $\tilde{f}$. $\circ$

We now prove a stronger version of (3.5). We only need the sparse algebraic Positivstellensatz on $[a, b]$ (Theorem 9.1) but not the sparse algebraic Nichtnegativestellensatz (Theorem 9.10) even for $p \geq 0$ on $[a, b]$. This result was already proved in [KS53]. Later the T-system approach was developed in [Kar63] and summarized and expanded in [KS66].

We now get the strengthened version of the Lukács–Markov Theorem 3.7. Earlier versions are due to Markov [Mar06] and Lukács [Luk18], see the Lukács–Markov Theorem 3.7 in Section 3.1 and the discussion around it.

Lukács–Markov Theorem 9.5 (see [KS53, Thm. 10.3] or [KN77, p. 373, Thm. 6.4]). *Let $p \in \mathbb{R}[x]$ with $p \geq 0$ on $[a, b]$ with $-\infty < a < b < \infty$ and let $z_1, \dots, z_r \in [a, b]$ be the zeros of p in $[a, b]$ with algebraic multiplicities $m_1, \dots, m_r \in \mathbb{N}$.*

(i) If $\deg p - m_1 - \dots - m_r = 2m$, $m \in \mathbb{N}_0$, is even then there exist points $x_1, \dots, x_m$ and $y_1, \dots, y_{m-1}$ with

$$a < x_1 < y_1 < \dots < y_{m-1} < x_m < b$$

and constants $\alpha, \beta > 0$ such that

$$p(x) = (x - z_1)^{m_1} \cdots (x - z_r)^{m_r} \cdot \left(\alpha \cdot \prod_{i=1}^{m} (x - x_i)^2 \right.$$
$$\left. + \beta \cdot (x - a) \cdot (b - x) \cdot \prod_{i=1}^{m-1} (x - y_i)^2 \right).$$

(ii) If $\deg p - m_1 - \dots - m_r = 2m + 1$, $m \in \mathbb{N}_0$, is odd then there exist points $x_1, \dots, x_m$ and $y_0, \dots, y_{m-1}$ with

$$a < y_0 < x_1 < y_1 < \dots < y_{m-1} < x_m < b$$

and constants $\alpha, \beta > 0$ such that

$$p(x) = (x - z_1)^{m_1} \cdots (x - z_r)^{m_r} \cdot \left(\alpha \cdot (x - a) \cdot \prod_{i=1}^{m} (x - x_i)^2 \right.$$
$$\left. + \beta \cdot (b - x) \cdot \prod_{i=0}^{m-1} (x - y_i)^2 \right).$$

Proof. We have $p(x) = (x - z_1)^{m_1} \cdots (x - z_r)^{m_r} \cdot \tilde{p}(x)$ with $\tilde{p} \in \mathbb{R}[x]$ and $\tilde{p} > 0$ on $[a, b]$. By a translation $p(\,\cdot\, + a)$ we can assume $a = 0$ and the assertion follows from Corollary 9.3. $\square$

Note, in Theorem 9.1 (and Theorem 9.10) we need $a \geq 0$. But in the Lukács–Markov Theorem 9.5 we can allow for arbitrary $a \in \mathbb{R}$ since by $p \in \mathbb{R}[x]_{\leq \deg p}$ the translation $p(\,\cdot\, + a)$ remains in $\mathbb{R}[x]_{\leq \deg p}$. We see here also why in Theorem 9.1 and Corollary 9.3 we have the restriction $a \geq 0$ since a translation can produce monomials which are not in the family $\{x^{\alpha_i}\}_{i=0}^n$.

Additionally, note that in Lukács–Markov Theorem 9.5 we can have $z_i = a$ or b for some i.

9.2 Sparse Hausdorff Moment Problem

Theorem 9.1 is a complete description of $\operatorname{int}(\operatorname{lin}\mathcal{F})_+$. Since $\mathcal{F}$ is continuous on the compact interval $[a,b]$ and $x^{\alpha_0} > 0$ on $[a,b]$, we have that the truncated moment cone is closed. Hence, $(\operatorname{lin}\mathcal{F})_+$ and the moment cone are dual to each other. With Theorem 9.1 we can now write down the conditions for the sparse truncated Hausdorff moment problem on $[a,b]$ with $a > 0$. A first but insufficient attempt was done in [Hau21b] since Hausdorff did not have access to the sparse Positivstellensatz by Karlin and therefore Theorem 9.1.

Theorem 9.6 (Sparse Truncated Hausdorff Moment Problem on $[a,b]$ with $a > 0$)**.** *Let $n \in \mathbb{N}_0$, $\alpha_0, \dots, \alpha_n \in [0,\infty)$ with $\alpha_0 < \dots < \alpha_n$, and a, b with $0 < a < b$. Set $\mathcal{F} = \{x^{\alpha_i}\}_{i=0}^n$. Then the following are equivalent:*

(i) $L : \operatorname{lin}\mathcal{F} \to \mathbb{R}$ *is a truncated $[a,b]$-moment functional.*
(ii) $L(p) \geq 0$ *holds for all*

$$p(x) := \begin{cases} \det\begin{pmatrix} x^{\alpha_0} & x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m-1}} & x^{\alpha_{2m}} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) \end{pmatrix} \\ -\det\begin{pmatrix} x^{\alpha_0} & x^{\alpha_1} & x^{\alpha_2} & x^{\alpha_3} & \dots & x^{\alpha_{2m-2}} & x^{\alpha_{2m-1}} & x^{\alpha_{2m}} \\ x & a & (x_1 & x_1) & \dots & (x_{m-1} & x_{m-1}) & b \end{pmatrix} \end{cases} \qquad \text{if } n = 2m$$

and

$$p(x) := \begin{cases} -\det\begin{pmatrix} x^{\alpha_0} & x^{\alpha_1} & x^{\alpha_2} & x^{\alpha_3} & \dots & x^{\alpha_{2m}} & x^{\alpha_{2m+1}} \\ x & a & (x_1 & x_1) & \dots & (x_m & x_m) \end{pmatrix} \\ \det\begin{pmatrix} x^{\alpha_0} & x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m-1}} & x^{\alpha_{2m}} & x^{\alpha_{2m+1}} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) & b \end{pmatrix} \end{cases} \qquad \text{if } n = 2m+1$$

and all $x_1, \dots, x_m$ with $a < x_1 < \dots < x_m < b$.

Proof. The implication (i) $\Rightarrow$ (ii) is clear since all given polynomials p are non-negative on $[a,b]$. It is therefore sufficient to prove (ii) $\Rightarrow$ (i).

Since $a > 0$ we have that $x^{\alpha_0} > 0$ on $[a,b]$ and since $[a,b]$ is compact we have that the moment cone $((\operatorname{lin}\mathcal{F})_+)^*$ as the dual of the cone of non-negative (sparse) polynomials $(\operatorname{lin}\mathcal{F})_+$ is a closed pointed cone.

To establish $L \in ((\operatorname{lin}\mathcal{F})_+)^*$ it is sufficient to have $L(f) \geq 0$ for all $f \in (\operatorname{lin}\mathcal{F})_+$. Let $f \in (\operatorname{lin}\mathcal{F})_+$. Then for all $\varepsilon > 0$ we have $f_\varepsilon := f + \varepsilon \cdot x^{\alpha_n} > 0$ on $[a,b]$, i.e., by Theorem 9.1 f_ε is a conic combination of the polynomials p in (ii) and hence $L(f) + \varepsilon \cdot L(x^{\alpha_n}) = L(f_\varepsilon) \geq 0$ for all $\varepsilon > 0$. Since $x^{\alpha_n} > 0$ on $[a,b]$ we also have that x^{α_n} is a conic combination of the polynomials p in (ii) and therefore $L(x^{\alpha_n}) \geq 0$. Then $L(f) \geq 0$ follows from $\varepsilon \to 0$ which proves (i). $\square$

Corollary 9.7. *If $\alpha_0 = 0$ in Theorem 9.6 then Theorem 9.6 also holds with $a = 0$, i.e., the following are equivalent:*

(i) $L : \operatorname{lin} \mathcal{F} \to \mathbb{R}$ *is a truncated* $[0, b]$*-moment functional.*
(ii) $L(p) \geq 0$ *holds for all*

$$p(x) := \begin{cases} \det \begin{pmatrix} 1 & x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m-1}} & x^{\alpha_{2m}} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) \end{pmatrix} & \text{if } n = 2m \\ \det \begin{pmatrix} x^{\alpha_1} & x^{\alpha_2} & x^{\alpha_3} & \dots & x^{\alpha_{2m-2}} & x^{\alpha_{2m-1}} & x^{\alpha_{2m}} \\ x & (x_1 & x_1) & \dots & (x_{m-1} & x_{m-1}) & b \end{pmatrix} & \end{cases}$$

and

$$p(x) := \begin{cases} \det \begin{pmatrix} x^{\alpha_1} & x^{\alpha_2} & x^{\alpha_3} & \dots & x^{\alpha_{2m}} & x^{\alpha_{2m+1}} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) \end{pmatrix} & \text{if } n = 2m+1 \\ \det \begin{pmatrix} 1 & x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m-1}} & x^{\alpha_{2m}} & x^{\alpha_{2m+1}} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) & b \end{pmatrix} & \end{cases}$$

and all $x_1, \dots, x_m$ *with* $a < x_1 < \dots < x_m < b$.

Proof. Follows immediately from Corollary 9.3. □

For the following we want to remind the reader of the *Müntz–Szász Theorem* [Mün14, Szá16]. It states that for real exponents $\alpha_0 = 0 < \alpha_1 < \alpha_2 < \dots$ the vector space $\operatorname{lin} \{x^{\alpha_i}\}_{i \in \mathbb{N}_0}$ of finite linear combinations is dense in $C([0, 1], \mathbb{R})$ with respect to the sup-norm if and only if $\sum_{i \in \mathbb{N}} \frac{1}{\alpha_i} = \infty$.

We state the following only for the classical case of the interval $[0, 1]$. Other cases $[a, b] \subseteq [0, \infty)$ are equivalent. Hausdorff required $\alpha_i \to \infty$. The Müntz–Szász Theorem does not require $\alpha_i \to \infty$. The conditions $\alpha_0 = 0$ and $\sum_{i \in \mathbb{N}} \frac{1}{\alpha_i} = \infty$ already appear in [Hau21b, eq. (17)]. We can remove here the use of the Müntz–Szász Theorem and therefore the condition $\sum_{i \in \mathbb{N}} \frac{1}{\alpha_i} = \infty$ for the existence of a representing measure. We need it only for uniqueness. Additionally, we allow negative exponents. The following is an improvement of [Hau21b] and we are not aware of a reference for this result.

Theorem 9.8 (General Sparse Hausdorff Moment Problem on $[a, b]$ with $0 \leq a < b$)**.** *Let* $I \subseteq \mathbb{N}_0$ *be an index set (finite or infinite), let* $\{\alpha_i\}_{i \in I}$ *be such that* $\alpha_i \neq \alpha_j$ *for all* $i \neq j$ *and*

(a) if $a = 0$ *then* $\{\alpha_i\}_{i \in I} \subset [0, \infty)$ *with* $\alpha_i = 0$ *for an* $i \in I$*, or*
(b) if $a > 0$ *then* $\{\alpha_i\}_{i \in I} \subset \mathbb{R}$.

Let $\mathcal{F} = \{x^{\alpha_i}\}_{i \in I}$*. Then the following are equivalent:*

(i) $L : \operatorname{lin} \mathcal{F} \to \mathbb{R}$ *is a Hausdorff moment functional.*
(ii) $L(p) \geq 0$ *holds for all* $p \in (\operatorname{lin} \mathcal{F})_+$.
(iii) $L(p) \geq 0$ *holds for all* $p \in \operatorname{lin} \mathcal{F}$ *with* $p > 0$.
(iv) $L(p) \geq 0$ *holds for all*

$$p(x) = \begin{cases} \det \begin{pmatrix} x^{\alpha_{i_0}} & x^{\alpha_{i_1}} & x^{\alpha_{i_2}} & \dots & x^{\alpha_{i_{2m-1}}} & x^{\alpha_{2m}} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) \end{pmatrix}, & \text{if } |I| = 2m \text{ or } \infty, \\ \det \begin{pmatrix} x^{\alpha_{i_1}} & x^{\alpha_{i_2}} & x^{\alpha_{i_3}} & \dots & x^{\alpha_{i_{2m-2}}} & x^{\alpha_{i_{2m-1}}} & x^{\alpha_{i_{2m}}} \\ x & (x_1 & x_1) & \dots & (x_{m-1} & x_{m-1}) & b \end{pmatrix}, & \text{if } |I| = 2m \text{ or } \infty, \\ \det \begin{pmatrix} x^{\alpha_{i_1}} & x^{\alpha_{i_2}} & x^{\alpha_{i_3}} & \dots & x^{\alpha_{i_{2m}}} & x^{\alpha_{i_{2m+1}}} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) \end{pmatrix}, & \text{if } |I| = 2m+1 \text{ or } \infty, \\ \det \begin{pmatrix} x^{\alpha_{i_0}} & x^{\alpha_{i_1}} & x^{\alpha_{i_2}} & \dots & x^{\alpha_{i_{2m-1}}} & x^{\alpha_{i_{2m}}} & x^{\alpha_{i_{2m+1}}} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) & b \end{pmatrix}, & \text{if } |I| = 2m+1 \text{ or } \infty, \end{cases}$$

for all $m \in \mathbb{N}$ if $|I| = \infty$, all $0 < x_1 < x_2 < \dots < x_m < b$, and all $\alpha_{i_0} < \alpha_{i_1} < \dots < \alpha_{i_m}$ with $\alpha_{i_0} = 0$ if $a = 0$.

If additionally $\sum_{i:\alpha_i \neq 0} \frac{1}{|\alpha_i|} = \infty$ then L is determinate.

Proof. The case $|I| < \infty$ is Theorem 9.6. We therefore prove the case $|I| = \infty$. The choice $\alpha_{i_0} < \alpha_{i_1} < \dots < \alpha_{i_m}$ with $\alpha_{i_0} = 0$ if $a = 0$ makes $\{x^{\alpha_{i_j}}\}_{j=0}^m$ a T-system. The implications "(i) $\Rightarrow$ (ii) $\Leftrightarrow$ (iii)" are clear and "(iii) $\Leftrightarrow$ (iv)" is Theorem 9.1. It is therefore sufficient to show "(ii) $\Rightarrow$ (i)". But the space $\operatorname{lin} \mathcal{F}$ is an adapted space and the assertion follows therefore from the Basic Representation Theorem 2.9.

For the determinacy of L split $\{\alpha_i\}_{i \in I}$ into positive and negative exponents. If $\sum_{i:\alpha_i \neq 0} \frac{1}{|\alpha_i|} = \infty$ then the corresponding sum over at least one group is infinite. If the sum over the positive exponents is infinite apply the Müntz–Szász Theorem. If the sum over the negative exponents is infinite apply the Müntz–Szász Theorem to $\{(x^{-1})^{-\alpha_i}\}_{i \in I:\alpha_i < 0}$ since $a > 0$. □

Note, since $[a, b]$ is compact the fact that $\{x^{\alpha_i}\}_{i \in I}$ is an adapted space is trivial.

Remark 9.9. If in Theorem 9.8 we have $a = 0$ and $\alpha_0 > 0$ then we can of course factor out x^{α_0} and instead of determining $\mathrm{d}\mu(x)$ of the linear functional L we determine $\mathrm{d}\tilde{\mu}(x) = x^{\alpha_0}\, \mathrm{d}\mu(x)$. ∘

9.3 Sparse Algebraic Nichtnegativstellensatz on $[a, b]$

The non-negative polynomials are described in the following result.

Theorem 9.10 (Sparse Algebraic Nichtnegativstellensatz on $[a, b]$ with $0 < a < b$)**.** *Let $n \in \mathbb{N}_0$, $\alpha_0, \dots, \alpha_n \in \mathbb{R}$ be real numbers with $\alpha_0 < \alpha_1 < \dots < \alpha_n$, and let $\mathcal{F} = \{x^{\alpha_i}\}_{i=0}^n$. Let $f \in \operatorname{lin} \mathcal{F}$ with $f \geq 0$ on $[a, b]$. Then there exist points $x_1, \dots, x_n, y_1, \dots, y_n \in [a, b]$ (not necessarily distinct) with $y_n = b$ which include the zeros of f with multiplicities such that*

$$f = f_* + f^*$$

with $f_*, f^* \in \operatorname{lin} \mathcal{F}$, $f_*, f^* \geq 0$ on $[a, b]$. The polynomials f_* and f^* are given by

$$f_*(x) = c_* \cdot \det \begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x & x_1 & \dots & x_n \end{pmatrix} \qquad \textit{and} \qquad f^*(x) = c_* \cdot \det \begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x & y_1 & \dots & y_n \end{pmatrix}$$

for all $x \in [a, b]$ and some constants $c_*, c^* \in \mathbb{R}$

Removing the zeros of f from $x_1, \dots, x_n, y_1, \dots, y_n$ we can assume that the remaining x_i and y_i are disjoint and when grouped by size the groups strictly interlace:

$$a \leq x_{i_1} = \dots = x_{i_k} < y_{j_1} = \dots = y_{j_l} < \dots < x_{i_p} = \dots = x_{i_q} < y_{j_r} = \dots = y_{j_s} = b.$$

Each such group in (a, b) has an even number of members.

Proof. By Example 5.17 we have that $\mathcal{F}$ on $[a, b]$ is an ET-system. We then apply Karlin's Nichtnegativstellensatz 7.6 similar to the proof of Theorem 9.1. □

Remark 9.11. The signs of c_* and c^* are determined by x_1 and y_1 and their multiplicity. If $x_1 = \dots = x_k < x_{k+1} \leq \dots \leq x_n$ then $\operatorname{sgn} c_* = (-1)^k$. The same holds for c^* from y_1. ∘

Example 9.12. Let $\alpha \in (0, \infty)$ and let $\mathcal{F} = \{1, x^\alpha\}$ on $[0, 1]$. Then we have $1 = 1_* + 1^*$ with $1_* = x^\alpha$ and $1^* = 1 - x^\alpha$. ∘

In Theorem 9.10 we can let $a = 0$ if $\alpha_0 = 0$ and $f(0) > 0$.

Theorem 9.13 (Sparse Algebraic Nichtnegativstellensatz on $[0, b]$ with $0 < b$)**.** *Let $n \in \mathbb{N}_0$, $\alpha_0, \dots, \alpha_n \in \mathbb{R}$ be real numbers with $0 = \alpha_0 < \alpha_1 < \dots < \alpha_n$, and let $\mathcal{F} = \{x^{\alpha_i}\}_{i=0}^n$ on $[0, b]$ with $b > 0$. Let $f \in \operatorname{lin} \mathcal{F}$ with $f \geq 0$ on $[0, b]$ and $f(0) > 0$. Then there exist points $x_1, \dots, x_n, y_1, \dots, y_n \in [0, b]$ (not necessarily distinct) with $y_n = b$ which include the zeros of f with multiplicities such that*

$$f = f_* + f^*$$

with $f_*, f^* \in \operatorname{lin} \mathcal{F}$, $f_*, f^* \geq 0$ on $[0, b]$ and the points $x_1, \dots, x_n$ are the zeros of f_* and $y_1, \dots, y_n$ are the zeros of f^*. Removing the zeros of f from $x_1, \dots, x_n, y_1, \dots, y_n$ we can assume that the remaining x_i and y_i are disjoint and when grouped by size the groups strictly interlace:

$$0 \leq x_{i_1} = \dots = x_{i_k} < y_{j_1} = \dots = y_{j_l} < \dots < x_{i_p} = \dots = x_{i_q} < y_{j_r} = \dots = y_{j_s} = b.$$

Each such group in (a, b) has an even number of members.

Proof. See Problem 9.1. □

Problems

9.1 Prove Theorem 9.13, i.e., show that Theorem 9.10 can be extended to the case $a = 0$, i.e., on $[0, b]$ with $b > 0$.

Chapter 10
Non-Negative Algebraic Polynomials on $[0,\infty)$ and on $\mathbb{R}$

Mathematics is the tool specially suited for dealing with abstract concepts of any kind and there is no limit to its power in this field.

Paul Adrien Maurice Dirac [Dir58, p. viii]

We went a long way to arrive here. But by using Karlin's Positivstellensatz 8.1 and Karlin's Nichtnegativstellensatz 8.3 on the interval $[0,\infty)$ we can now describe all sparse algebraic strictly positive and non-negative polynomials on $[0,\infty)$ and on $\mathbb{R}$.

10.1 Sparse Algebraic Positivstellensatz on $[0,\infty)$

For the sparse algebraic Positivstellensatz on $[a,b]$ (Theorem 9.1) we had a lot of freedom in the exponents α_i for $a > 0$. We no longer have such a large range of freedom on $[0,\infty)$. If we now plug Example 4.16 into Karlin's Positivstellensatz 8.1 we get the following.

Theorem 10.1 (Sparse Algebraic Positivstellensatz on $[0,\infty)$)**.** *Let $n \in \mathbb{N}_0$, $\alpha_0,\dots,\alpha_n \in [0,\infty)$ be real numbers with $\alpha_0 = 0 < \alpha_1 < \dots < \alpha_n$, and let $\mathcal{F} = \{x^{\alpha_i}\}_{i=0}^n$ on $[0,\infty)$. Then for any $f = \sum_{i=0}^n a_i f_i \in \operatorname{lin}\mathcal{F}$ with $f > 0$ on $[0,\infty)$ and $a_n > 0$ there exists a unique decomposition*

$$f = f_* + f^*$$

with $f_*, f^* \in \operatorname{lin}\mathcal{F}$ and $f_*, f^* \geq 0$ on $[0,\infty)$ such that the following hold:

(i) If $n = 2m$ then the polynomials f_* and f^* each possess m distinct zeros $\{x_i\}_{i=1}^m$ and $\{y_i\}_{i=0}^{m-1}$ satisfying

$$0 = y_0 < x_1 < y_1 < \dots < y_{m-1} < x_m < \infty.$$

The polynomials f_* and f^* are given by

$$f_*(x) = c_* \cdot \det\begin{pmatrix} 1 & x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m-1}} & x^{\alpha_{2m}} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) \end{pmatrix}$$

and

$$f^*(x) = c^* \cdot \det \begin{pmatrix} x^{\alpha_1} & x^{\alpha_2} & x^{\alpha_3} & \dots & x^{\alpha_{2m-2}} & x^{\alpha_{2m-1}} \\ x & (y_1 & y_1) & \dots & (y_{m-1} & y_{m-1}) \end{pmatrix}$$

for some $c_*, c^* > 0$.

(ii) *If $n = 2m+1$ then f_* and f^* have zeros $\{x_i\}_{i=1}^{m+1}$ and $\{y_i\}_{i=1}^{m}$ respectively which satisfy*

$$0 = x_1 < y_1 < x_2 < \dots < y_m < x_{m+1} < \infty.$$

The polynomials f_* and f^* are given by

$$f_*(x) = c_* \cdot \det \begin{pmatrix} x^{\alpha_1} & x^{\alpha_2} & x^{\alpha_3} & \dots & x^{\alpha_{2m}} & x^{\alpha_{2m+1}} \\ x & (x_2 & x_2) & \dots & (x_{m+1} & x_{m+1}) \end{pmatrix}$$

and

$$f^*(x) = c^* \cdot \det \begin{pmatrix} 1 & x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m-1}} & x^{\alpha_{2m}} \\ x & (y_1 & y_1) & \dots & (y_m & y_m) \end{pmatrix}$$

for some $c_*, c^* > 0$.

Proof. We have that $\mathcal{F}$ fulfills conditions (a) and (b) of Karlin's Positivstellensatz 8.1 and by Example 4.15 we known that $\mathcal{F}$ on $[0,\infty)$ is also a T-system, i.e., (c) in Karlin's Positivstellensatz 8.1 is fulfilled. We can therefore apply Karlin's Positivstellensatz 8.1.

(i) $n = 2m$: By Karlin's Positivstellensatz 8.1 (i) the unique f_* and f^* each possess m distinct zeros $\{x_i\}_{i=1}^{m}$ and $\{y_i\}_{i=0}^{m-1}$ with $0 \leq y_0 < x_1 < \dots < y_{m-1} < x_m < \infty$. Since $x_1, \dots, x_m \in (0,\infty)$ and $\mathcal{F}$ on $[x_1/2,\infty)$ is an ET-system we immediately get the determinantal representation of f_* by Corollary 8.2 (combine Karlin's Positivstellensatz 8.1 with Remark 4.28). For f^* we have $y_0 = 0$ and by Example 5.16 this is no ET-system. Hence, we prove the representation of f^* by hand, similar as in the proof of Corollary 9.3.

Let $\varepsilon > 0$ be such that $0 = y_0 < y_1 < y_1 + \varepsilon < \dots < y_{m-1} < y_{m-1} + \varepsilon$ holds. Then

$$\begin{aligned} g_\varepsilon(x) &= -\varepsilon^{-m+1} \cdot \det \begin{pmatrix} 1 & x^{\alpha_1} & x^{\alpha_2} & x^{\alpha_3} & \dots & x^{\alpha_{2m-2}} & x^{\alpha_{2m-1}} \\ x & 0 & y_1 & y_1+\varepsilon & \dots & y_{m-1} & y_{m-1}+\varepsilon \end{pmatrix} \\ &= -\varepsilon^{-m+1} \cdot \det \begin{pmatrix} 1 & x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m-1}} \\ 1 & 0 & 0 & \dots & 0 \\ 1 & y_1^{\alpha_1} & y_1^{\alpha_2} & \dots & y_1^{\alpha_{2m-1}} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & (y_{m-1}+\varepsilon)^{\alpha_1} & (y_{m-1}+\varepsilon)^{\alpha_2} & \dots & (y_{m-1}+\varepsilon)^{\alpha_{2m-1}} \end{pmatrix} \end{aligned}$$

expand by the second row

$$= \varepsilon^{-m+1}\cdot\det\begin{pmatrix} x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m-1}}\\ y_1^{\alpha_1} & y_1^{\alpha_2} & \dots & y_1^{\alpha_{2m-1}}\\ \vdots & \vdots & & \vdots\\ (y_{m-1}+\varepsilon)^{\alpha_1} & (y_{m-1}+\varepsilon)^{\alpha_2} & \dots & (y_{m-1}+\varepsilon)^{\alpha_{2m-1}}\end{pmatrix}$$

$$= \varepsilon^{-m+1}\cdot\det\begin{pmatrix} x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m-2}} & x^{\alpha_{2m-1}}\\ x & y_1 & y_1+\varepsilon & \dots & y_{m-1} & y_{m-1}+\varepsilon\end{pmatrix}$$

is non-negative on $[0, y_1]$ and every $[y_i+\varepsilon, y_{i+1}]$. Now $y_0 = 0$ is removed and all $y_i, y_i+\varepsilon > 0$. Hence, we can work on $[y_1/2,\infty)$ where $\{x^{\alpha_i}\}_{i=1}^{2m}$ is an ET-system and we can go to the limit $\varepsilon \searrow 0$ as in Remark 4.28. Then Corollary 8.2 proves the representation of f^*.

(ii) $n = 2m+1$: Similar to the case (i) with $n = 2m$. □

If all $\alpha_i \in \mathbb{N}_0$ then we can express the f_* and f^* in Theorem 10.1 also with Schur polynomials, see (5.14) in Example 5.17.

We now prove a stronger version of (3.4), i.e., $p = f^2 + x\cdot g^2$ for any $p \geq 0$ on $[0,\infty)$. It is sufficient to have only the sparse algebraic Positivstellensatz (Theorem 10.1). A previous version already appeared in [KS53].

Corollary 10.2 (see [KS66, p. 169, Cor. 8.1])**.** *Let $p \in \mathbb{R}[x]$ with $p \geq 0$ on $[0,\infty)$. Let $z_1,\dots,z_r \in [0,\infty)$ be the zeros of p in $[0,\infty)$ and let $m_1,\dots,m_r \in \mathbb{N}$ be the corresponding algebraic multiplicities.*

(i) If $\deg p - m_1 - \dots - m_r = 2m$, $m \in \mathbb{N}_0$, is even then there exist points $\{x_i\}_{i=1}^m$ and $\{y_i\}_{i=1}^{m-1} \subseteq (0,\infty)$ with

$$0 < x_1 < y_1 < \dots < y_{m-1} < x_m < \infty$$

and constants $a, b > 0$ such that

$$p(x) = \prod_{i=1}^r (x-z_i)^{m_i}\cdot\left(a\cdot\prod_{i=1}^m (x-x_i)^2 + b\cdot x\cdot\prod_{i=1}^{m-1}(x-y_i)^2\right).$$

The constant a is the leading coefficient of p.

(ii) If $\deg p - m_1 - \dots - m_r = 2m+1$, $m \in \mathbb{N}_0$, is odd then there exist points $\{x_i\}_{i=1}^m$ and $\{y_i\}_{i=1}^m \subset (0,\infty)$ with

$$0 < x_1 < y_1 < \dots < x_m < y_m < \infty$$

and constants $a, b > 0$ such that

$$p(x) = \prod_{i=1}^r (x-z_i)^{m_i}\cdot\left(a\cdot\prod_{i=1}^m (x-x_i)^2 + b\cdot x\cdot\prod_{i=1}^{m}(x-y_i)^2\right).$$

The constant b is the leading coefficient of p.

Proof. Since $z_1,\dots,z_r$ are the zeros of p in $[0,\infty)$ with multiplicities $m_1,\dots,m_r$ we have that $p(x) = (x-z_1)^{m_1}\cdots(x-z_r)^{m_r}\cdot\tilde{p}(x)$ with $\tilde{p}\in\mathbb{R}[x]$ and $\tilde{p}>0$ on $[0,\infty)$. Applying Theorem 10.1 to $\tilde{p}$ gives the assertion. □

Note, in the previous result we were able to factor out the zeros of p and were only left with $\tilde{p}>0$ on $[0,\infty)$ since we are working in $\mathbb{R}[x]_{\leq \deg p}$ where all monomials $1,x,\dots,x^{\deg p}$ are present. In sparse systems we are not able to factor out the zeros since we no longer know which monomials in $\tilde{p}$ will appear.

Remark 10.3. Working in the sparse setting, i.e., in T-systems, gives us an additional information. In (3.4) we only have $p(x) = x\cdot f^2 + g^2$. But this also includes that f and g might contain factors $((x-y_i)^2+\delta_i)$ with $\delta_i>0$, i.e., a pair of complex conjugated zeros can be present. In Corollary 10.2 we see that this is not necessary. The polynomials f and g can always be chosen such that they decompose into linear factors having only real zeros. A similar results holds on $\mathbb{R}$, see Theorem 10.7. ∘

10.2 Sparse Stieltjes Moment Problem

In Section 3.2 we have seen that Boas already investigated the sparse Stieltjes moment problem [Boa39a]. However, the description was complicated and is even incomplete since Boas did not had access to Karlin's Positivstellensatz 8.1 and therefore Theorem 10.1. We get the following complete and simple description. It fully solves [Boa39a]. We are not aware of a reference for the following result.

Theorem 10.4 (Sparse Stieltjes Moment Problem)**.** *Let $\{\alpha_i\}_{i\in\mathbb{N}_0}\subseteq[0,\infty)$ be such that $\alpha_0 = 0 < \alpha_1 < \alpha_2 < \dots$ and let $\mathcal{F} = \{x^{\alpha_i}\}_{i\in\mathbb{N}_0}$. Then the following are equivalent:*

(i) $L:\operatorname{lin}\mathcal{F}\to\mathbb{R}$ *is a* $[0,\infty)$*-moment functional.*
(ii) $L(p)\geq 0$ *for all* $p\in\operatorname{lin}\mathcal{F}$ *with* $p\geq 0$.
(iii) $L(p)\geq 0$ *for all* $p\in\operatorname{lin}\mathcal{F}$ *with* $p>0$.
(iv) $L(p)\geq 0$ *for all*

$$p(x) = \begin{cases} \det\begin{pmatrix} 1 & x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m-1}} & x^{\alpha_{2m}} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) \end{pmatrix}, \\ \det\begin{pmatrix} x^{\alpha_1} & x^{\alpha_2} & x^{\alpha_3} & \dots & x^{\alpha_{2m-2}} & x^{\alpha_{2m-1}} \\ x & (x_1 & x_1) & \dots & (x_{m-1} & x_{m-1}) \end{pmatrix}, \\ \det\begin{pmatrix} x^{\alpha_1} & x^{\alpha_2} & x^{\alpha_3} & \dots & x^{\alpha_{2m}} & x^{\alpha_{2m+1}} \\ x & (x_2 & x_2) & \dots & (x_{m+1} & x_{m+1}) \end{pmatrix}, \text{ and} \\ \det\begin{pmatrix} 1 & x^{\alpha_1} & x^{\alpha_2} & \dots & x^{\alpha_{2m-1}} & x^{\alpha_{2m}} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) \end{pmatrix} \end{cases}$$

for all $m\in\mathbb{N}_0$ *and* $0<x_1<\dots<x_m$.

Proof. The implications "(i) $\Rightarrow$ (ii) $\Leftrightarrow$ (iii)" are clear and "(iii) $\Leftrightarrow$ (iv)" is Theorem 10.1. It is therefore sufficient to prove "(ii) $\Rightarrow$ (i)".

We have $\operatorname{lin}\mathcal{F} = (\operatorname{lin}\mathcal{F})_+ - (\operatorname{lin}\mathcal{F})_+$, we have $1 = x^{\alpha_0} \in \operatorname{lin}\mathcal{F}$, and for any $g = \sum_{i=0}^m a_i \cdot x^{\alpha_i} \in (\operatorname{lin}\mathcal{F})_+$ we have $\lim_{x\to\infty} \frac{g(x)}{x^{\alpha_{m+1}}} = 0$, i.e., there exists a $f \in (\operatorname{lin}\mathcal{F})_+$ which dominates g. Hence, $\operatorname{lin}\mathcal{F}$ is an adapted space on $[0,\infty)$ and the assertion follows from the Basic Representation Theorem 2.9. □

In the previous result we did needed $0 = \alpha_0 < \alpha_1 < \alpha_2 < \dots$. We did not needed $\alpha_i \to \infty$. Hence, Theorem 10.4 also includes the case $\sup_{i\in\mathbb{N}_0} \alpha_i < \infty$.

Theorem 10.4 also holds with $\alpha_0 > 0$ since we can factor out x^{α_0} and therefore determine $x^{\alpha_0}\,\mathrm{d}\mu(x)$ instead of $\mathrm{d}\mu(x)$.

10.3 Sparse Algebraic Nichtnegativstellensatz on $[0,\infty)$

For $\{1,x,x^3\}$ we have seen in Example 5.16 that this is not an ET-systen on $[0,\infty)$, or on any other $[0,b]$. If we remove the point $x=0$ and work on $(0,\infty)$ then it is an ET-system and even an ECT-system (Examples 5.18). For a Nichtnegativstellensatz we therefore have to exclude zeros at $x=0$ in a sparse polynomial $p \geq 0$.

Theorem 10.5 (Sparse Algebraic Nichtnegativstellensatz on $[0,\infty)$)**.** *Let $n \in \mathbb{N}_0$, $\alpha_0,\dots,\alpha_n \in [0,\infty)$ be real numbers with $\alpha_0 = 0 < \alpha_1 < \dots < \alpha_n$, and let $\mathcal{F} = \{x^{\alpha_i}\}_{i=0}^n$. Let $f = \sum_{i=0}^n a_i x^{\alpha_i} \geq 0$ on $[0,\infty)$ with $a_n > 0$ and $f(0) = a_0 > 0$. Then there exist points $x_1,\dots,x_n,y_1,\dots,y_{n-1} \in [0,\infty)$ (not necessarily distinct) which include the zeros of f with multiplicities and there exist constants $c_*, c^* \in \mathbb{R}$ such that*

$$f = f_* + f^*$$

with $f_*, f^* \in \operatorname{lin}\mathcal{F}$, $f_*, f^* \geq 0$ on $[0,\infty)$, and the polynomials f_* and f^* are given by

$$f_*(x) = c_* \cdot \det\begin{pmatrix} 1 & x^{\alpha_1} & \dots & x^{\alpha_n} \\ x & x_1 & \dots & x_n \end{pmatrix} \qquad \textit{and} \qquad f^*(x) = c_* \cdot \det\begin{pmatrix} 1 & x^{\alpha_1} & \dots & x^{\alpha_{n-1}} \\ x & y_1 & \dots & y_{n-1} \end{pmatrix}$$

for all $x \in [0,\infty)$.

Proof. See Problem 10.1. □

Remark 10.6. Note, if $f(0) = a_0 = 0$ in Theorem 10.5 then

$$f(x) = a_i x^{\alpha_i} + a_{i+1}x^{\alpha_{i+1}} + \dots + a_n x^{\alpha_n} = x^{\alpha_i} \cdot (\underbrace{a_i + a_{i+1}x^{\alpha_{i+1}-\alpha_i} + \dots + a_n x^{\alpha_n - \alpha_i}}_{=:\tilde{f}(x)})$$

where a_i is the first non-zero coefficient and it fulfills $a_i > 0$ since $f \geq 0$. Then apply Theorem 10.5 to $\tilde{f}$ to get $\tilde{f} = \tilde{f}_* + \tilde{f}^*$ and hence $f = x^{\alpha_i} \cdot (\tilde{f}_* + \tilde{f}^*)$. ∘

10.4 Algebraic Positiv- and Nichtnegativstellensatz on $\mathbb{R}$

Since we treat $\mathcal{F} = \{x^i\}_{i=0}^n$ we need only Karlin's Positivstellensatz 8.4 on $\mathbb{R}$ but not Karlin's Nichtnegativstellensatz 8.5 on $\mathbb{R}$ as we will see in the next result.

Theorem 10.7 (Algebraic Positiv- and Nichtnegativstellensatz on $\mathbb{R}$, see [KS53,] or e.g. [KS66, p. 198, Cor. 8.1])**.** *Let $p \in \mathbb{R}[x]$ with $p \geq 0$ on $\mathbb{R}$ and let $z_1, \dots, z_r \in \mathbb{R}$ be the zeros of p with algebraic multiplicities $m_1, \dots, m_r \in 2\mathbb{N}$. Then there exist pairwise distinct points $\{x_i\}_{i=1}^m, \{y_i\}_{i=1}^{m-1} \subseteq \mathbb{R}$ with $2m = \deg p - m_1 - \dots - m_r$ and*

$$-\infty < x_1 < y_1 < \dots < y_{m-1} < x_m < \infty$$

as well as constants $a, b > 0$ such that

$$p(x) = \prod_{i=1}^{r} (x - z_i)^{m_i} \cdot \left(a \cdot \prod_{i=1}^{m} (x - x_i)^2 + b \cdot \prod_{i=1}^{m-1} (x - y_i)^2 \right). \tag{10.1}$$

The constant a is the leading coefficient of p.

Proof. We have $p(x) = (x - z_1)^{m_1} \cdots (x - z_r)^{m_r} \cdot \tilde{p}(x)$ for some $\tilde{p} \in \mathbb{R}[x]$ with $\tilde{p} > 0$ on $\mathbb{R}$. Applying Karlin's Positivstellensatz 8.4 to $\tilde{p}$ gives the assertion. $\square$

Like in the case on $[0, \infty)$ in Corollary 10.2 a factorization

$$p(x) = (x - z_1)^{m_1} \cdots (x - z_r)^{m_r} \cdot \tilde{p}(x)$$

is not possible in T-systems or sparse algebraic systems on $\mathbb{R}$. But since we are working in $\mathbb{R}[x]_{\leq \deg p}$ all monomials $1, x, \dots, x^{\deg p}$ are present.

Remark 10.8. Similar to Remark 10.3 we see that Theorem 10.7 gives a stronger version of (3.2), i.e., $p = f^2 + g^2$. By applying only the Fundamental Theorem of Algebra f and g might contain pairs of complex conjugated zeros, see e.g. [Mar08, Prop. 1.2.1]. But by working in the T-system framework of Karlin's Positivstellensatz 8.4 on $\mathbb{R}$ we see that f and g can be chosen to have only real zeros. $\circ$

On $[0, \infty)$ we have seen that for any $0 = \alpha_0 < \alpha_1 < \alpha_2 < \dots < \alpha_n$ we have that $\mathcal{F} = \{x^{\alpha_i}\}_{i=0}^n$ is a T-system. On $\mathbb{R}$ this is no longer the case.

Example 10.9. Let $\mathcal{F} = \{1, x^2, x^4, x^6, x^8\}$. Then $\mathcal{F}$ on $\mathbb{R}$ is no T-system. Let $p \in \operatorname{lin} \mathcal{F}$ be non-negative on $[0, \infty)$ with zeros at $x = 1$ and 2. By symmetry $p \geq 0$ on $\mathbb{R}$ with double zeros at $x = \pm 1$ and ± 2 which contradicts Theorem 4.22. $\circ$

Problems

10.1 Use Karlin's Nichtnegativstellensatz 8.3 to prove Theorem 10.5.

10.2 Show that a in (10.1) in Theorem 10.7 is the leading coefficient of p.

Part V
Applications of T-Systems

Chapter 11
Moment Problems for continuous T-Systems on $[a, b]$

Long is the way and hard, that out of Hell leads up to light.

John Milton: Paradise Lost

In this chapter we demonstrate how e.g. Karlin's Positivstellensatz 7.3 for general T-systems on $[a, b]$ can be used to prove moment problems which do not live on the algebraic polynomials $\mathbb{R}[x]$.

11.1 General Moment Problems for continuous T-Systems on $[a, b]$

For T-system $\mathcal{F}$ on $[a, b]$ Karlin's Positivstellensatz 7.3 describes all polynomials $f \in \operatorname{lin} \mathcal{F}$ with $f > 0$.

Theorem 11.1. *Let $n \in \mathbb{N}$, let $\mathcal{F} = \{f_i\}_{i=0}^n$ be a continuous T-system on $[a, b]$ with $a < b$. The following are equivalent:*

(i) $L : \operatorname{lin} \mathcal{F} \to \mathbb{R}$ *is an $[a, b]$-moment functional.*
(ii) $L(f) \geq 0$ *for all $f \in \operatorname{lin} \mathcal{F}$ such that*

  (a) $f \geq 0$ *on $[a, b]$ and*
  (b) the zero set of f has index n.

Proof. The implication (i) $\Rightarrow$ (ii) is clear since $f \geq 0$. It is therefore sufficient to prove (ii) $\Rightarrow$ (i).

Since $\mathcal{F}$ is a continuous T-system there exists a polynomial $e \in \operatorname{lin} \mathcal{F}$ with $e > 0$ on $[a, b]$. Since $[a, b]$ is compact, $\mathcal{F}$ is continuous and finite dimensional, and there exists a $e > 0$ we have that the moment cone $((\operatorname{lin} \mathcal{F})_+)^*$ is closed. Therefore, to show that L is a moment functional it is sufficient to show that $L(f) \geq 0$ for all $f \in (\operatorname{lin} \mathcal{F})_+$.

By Karlin's Positivstellensatz 7.3 there are $e_*, e^* \in \operatorname{lin} \mathcal{F}$ with $e_*, e^* \geq 0$ and the zero sets of e_* and of e^* have index n. Hence, $L(e) = L(e_*) + L(e^*) \geq 0$.

Let $f \in (\operatorname{lin} \mathcal{F})_+$ and $\varepsilon > 0$. Then $f_\varepsilon = f + \varepsilon \cdot e > 0$ on $[a, b]$, i.e., by Karlin's Positivstellensatz 7.3 there exist $(f_\varepsilon)_*, (f_\varepsilon)^* \in (\operatorname{lin} \mathcal{F})_+$ each with zero sets of index

n. Assumption (ii) then implies $L(f+\varepsilon\cdot e) = L((f_\varepsilon)_*)+L((f_\varepsilon)^*) \geq 0$ for all $\varepsilon > 0$, i.e., $L(f) \geq 0$. That proves the assertion. □

Note, that a continuous T-system on $[a, b]$ is always an adapted space. Additionally, the use of Basic Representation Theorem 2.9 is not necessary since we only need to check in this case $L \in ((\operatorname{lin}\mathcal{F})_+)^*$ since the moment cone is $((\operatorname{lin}\mathcal{F})_+)^*$ and hence it is closed.

If in the previous theorem we additionally have that $\mathcal{F}$ is an ET-system then we can write down f_* and f^* explicitly in the similar way as in Theorem 9.6.

Theorem 11.2. *Let $n \in \mathbb{N}$, let $\mathcal{F} = \{f_i\}_{i=0}^n$ be an ET-system on $[a, b]$ with $a < b$. The following are equivalent:*

(i) $L : \operatorname{lin}\mathcal{F} \to \mathbb{R}$ *is a moment functional.*
(ii) $L(f) \geq 0$ *holds for all*

$$f(x) := \begin{cases} \det\begin{pmatrix} f_0 & f_1 & f_2 & \dots & f_{2m-1} & f_{2m} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) \end{pmatrix} \\ -\det\begin{pmatrix} f_0 & f_1 & f_2 & f_3 & \dots & f_{2m-2} & f_{2m-1} & f_{2m} \\ x & a & (x_1 & x_1) & \dots & (x_{m-1} & x_{m-1}) & b \end{pmatrix} \end{cases} \qquad \text{if } n = 2m$$

and

$$f(x) := \begin{cases} -\det\begin{pmatrix} f_0 & f_1 & f_2 & f_3 & \dots & f_{2m} & f_{2m+1} \\ x & a & (x_1 & x_1) & \dots & (x_m & x_m) \end{pmatrix} \\ \det\begin{pmatrix} f_0 & f_1 & f_2 & \dots & f_{2m-1} & f_{2m} & f_{2m+1} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) & b \end{pmatrix} \end{cases} \qquad \text{if } n = 2m+1$$

and all $x_1, \dots, x_m$ with $a < x_1 < \dots < x_m < b$.

Proof. Follows from Theorem 11.1 with Theorem 5.3. □

11.2 A Non-Polynomial Example

In Example 4.18 we have seen that

$$\mathcal{F} = \left\{\frac{1}{x+\alpha_0}, \frac{1}{x+\alpha_1}, \dots, \frac{1}{x+\alpha_n}\right\}$$

with $n \in \mathbb{N}$ and $\alpha_0 < \alpha_1 < \dots < \alpha_n$ reals is a continuous T-system on any $[a, b]$ with $-\alpha_0 < a < b$, see Problem 4.5 for the proof. But in the proof of Example 4.18 we actually showed that this $\mathcal{F}$ is an ET-system since we multiplied with $(x+\alpha_0)\cdots(x+\alpha_n)$ which has no zeros on $[a, b]$ and hence the multiplicities of the zeros do not change. Multiplicity restriction from the fundamental theorem of algebra then shows that $\mathcal{F}$ is an ET-system.

Corollary 11.3. *Let $n \in \mathbb{N}$ and $\alpha_0 < \alpha_1 < \cdots < \alpha_n$ be reals. Then*

$$\mathcal{F} = \left\{ \frac{1}{x+\alpha_0}, \frac{1}{x+\alpha_1}, \dots, \frac{1}{x+\alpha_n} \right\}$$

is an ET-system on any $[a, b]$ with $-\alpha_0 < a < b$.

From Theorem 11.2 and Corollary 11.3 we therefore get the following.

Corollary 11.4. *Let $n \in \mathbb{N}$, let $\alpha_0 < \alpha_1 < \cdots < \alpha_n$ be reals, and let*

$$\mathcal{F} = \left\{ f_0(x) = \frac{1}{x+\alpha_0}, f_1(x) = \frac{1}{x+\alpha_1}, \dots, f_n(x) = \frac{1}{x+\alpha_n} \right\}$$

on $[a, b]$ with $-\alpha_0 < a < b$. Then the following are equivalent:

(i) $L : \operatorname{lin} \mathcal{F} \to \mathbb{R}$ *is a $[a, b]$-moment functional.*
(ii) $L(f) \geq 0$ *holds for all*

$$f(x) := \begin{cases} \det \begin{pmatrix} f_0 & f_1 & f_2 & \dots & f_{2m-1} & f_{2m} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) \end{pmatrix} \\ -\det \begin{pmatrix} f_0 & f_1 & f_2 & f_3 & \dots & f_{2m-2} & f_{2m-1} & f_{2m} \\ x & a & (x_1 & x_1) & \dots & (x_{m-1} & x_{m-1}) & b \end{pmatrix} \end{cases} \qquad \text{if } n = 2m$$

and

$$f(x) := \begin{cases} -\det \begin{pmatrix} f_0 & f_1 & f_2 & f_3 & \dots & f_{2m} & f_{2m+1} \\ x & a & (x_1 & x_1) & \dots & (x_m & x_m) \end{pmatrix} \\ \det \begin{pmatrix} f_0 & f_1 & f_2 & \dots & f_{2m-1} & f_{2m} & f_{2m+1} \\ x & (x_1 & x_1) & \dots & (x_m & x_m) & b \end{pmatrix} \end{cases} \qquad \text{if } n = 2m+1$$

and all $x_1, \dots, x_m$ with $a < x_1 < \cdots < x_m < b$.

In a similar way many other T-system moment problems can be proven from Theorem 11.1.

Chapter 12
Polynomials of Best Approximation and Optimization over Linear Functionals

The rest is silence.

William Shakespeare: Hamlet (Act 5, Scene 2)

This last chapter is devoted to best approximation polynomials and optimization over linear functionals.

We started in Chapter 1 with moments and moment functionals, went to the theory of T-systems in Part II, proved Karlin's Theorems in Part III, and applied them to algebraic polynomials in Part IV. Now we finish our lecture by closing the circle. We apply the previous results to best approximation in Section 12.1 and to optimization over linear (moment) functionals in Section 12.2.

12.1 Polynomials of Best Approximation

A classical question is:

> How to approximate a given function $f \in C([a,b],\mathbb{R})$ in the sup-norm by a finite linear combination $\sum_{i=0}^n a_i f_i$ of some given $f_0,\dots,f_n \in C([a,b],\mathbb{R})$?

Definition 12.1. Let $n \in \mathbb{N}_0$, let $f, f_0, \dots, f_n \in C([a,b],\mathbb{R})$, and let $\mathcal{F} := \{f_i\}_{i=0}^n$. The polynomial $\underline{f} \in \operatorname{lin}\mathcal{F}$ which solves

$$\min_{a_0,\dots,a_n} \left\| f - \sum_{i=0}^n a_i f_i \right\|_\infty \tag{12.1}$$

is called the *polynomial of best approximation*.

Approximations (12.1) with the sup-norm are called *Tchebycheff approximations*.

The connection between polynomials of best approximation and T-systems is revealed in the following result.

Theorem 12.2 (see [Haa18], [Ber26]; or e.g. [Ach56, p. 74, §48], [KS66, p. 280, Thm. 1.1])**.** *Let $n \in \mathbb{N}_0$, let $a, b \in \mathbb{R}$ with $a < b$, and let $\mathcal{F} := \{f_i\}_{i=0}^n \subseteq C([a,b],\mathbb{R})$ be a family of continuous functions. The following hold:*

(i) The following are equivalent:

(a) The polynomial minimizing

$$\min_{a_0,\ldots,a_n} \left\| f - \sum_{i=0}^{n} a_i f_i \right\|_\infty \tag{12.2}$$

is uniquely determined for every $f \in C([a,b],\mathbb{R})$.

(b) The family $\mathcal{F}$ *is a continuous T-system on* $[a,b]$.

(ii) If $\mathcal{F}$ *is a T-system then for each* $f \in C([a,b],\mathbb{R})$ *the unique polynomial*

$$\underline{f} = \sum_{i=0}^{n} \underline{a}_i f_i$$

minimizing (12.2) is characterized by the property that there exist $n+2$ *points*

$$a \le \quad x_1 < x_2 < \cdots < x_{n+2} \quad \le b$$

such that

$$(-1)^i \cdot \delta \cdot (f(x_i) - \underline{f}(x_i)) = \max_{a \le x \le b} \left| f(x) - \underline{f}(x) \right|$$

holds for all $i = 1, 2, \ldots, n+2$ *with* $\delta = +1$ *or* -1.

Statement (i) of the previous theorem is essentially due to A. Haar [Haa18]. The following proof significantly differs from Haar's proof and it is more general. It is taken from [KS66, pp. 284–286], see also [Ach56, pp. 75–76].

Proof. (a) $\Rightarrow$ (b): We prove $\neg$(b) $\Rightarrow$ $\neg$(a).

Assume $\mathcal{F}$ is not a T-system. There exist $n+1$ distinct points $a \le x_0 < x_1 < \cdots < x_n \le b$ such that

$$\det \left(f_i(x_j)\right)_{i,j=0}^{n} = 0. \tag{12.3}$$

Hence, there exist real coefficients $c_0, c_1, \ldots, c_n$ with $\sum_{i=0}^n c_i^2 > 0$ with $\sum_{i=0}^n c_i f_j(x_i) = 0$ for all $j = 0, \ldots, n$. That implies

$$\sum_{i=0}^{n} c_i p(x_i) = 0 \tag{12.4}$$

for all $p \in \operatorname{lin} \mathcal{F}$.

The relation (12.3) also implies the existence of a non-trivial polynomial $\tilde{p} = \sum_{i=0}^n b_i f_i \in \operatorname{lin} \mathcal{F}$ which vanishes at the points $x_0, x_1, \ldots, x_n$.

Let $g \in C([a,b],\mathbb{R})$ be such that $\|g\|_\infty \le 1$ and

$$g(x_i) = \frac{c_i}{|c_i|}$$

for all $i = 0, 1, \ldots, n$ with $c_i \neq 0$.

Let $\lambda > 0$ be such that $\|\lambda\tilde{p}\|_\infty < 1$ then $f := g \cdot (1 - |\lambda\tilde{p}|)$ has the same signs at the points x_i with $c_i \neq 0$ as g.

We will now construct an infinite number of polynomials of the same minimum deviation from f.

If

$$\left\| f - \sum_{i=0}^{n} a_i f_i \right\|_\infty < 1$$

for some $a_0, a_1, \ldots, a_n$ then

$$-1 < g(x_j) \cdot (1 - |\lambda\tilde{p}(x_j)|) - \sum_{i=0}^{n} a_i f_i(x_j) < 1$$

for all $j = 0, 1, \ldots, n$ which reduces to

$$-1 < g(x_j) - \sum_{i=0}^{n} a_i f_i(x_j) < 1$$

for all $j = 0, 1, \ldots, n$. Hence, if $c_j \neq 0$ the value of $\sum_{i=0}^{n} a_i f_i(x_j)$ has the sign of the c_j so that $\sum_{j=0}^{n} c_j \sum_{i=0}^{n} a_i f_i(x_j) \neq 0$ which contradicts (12.4). Therefore,

$$\left\| f - \sum_{i=0}^{n} a_i f_i \right\|_\infty \geq 1.$$

If now $|\delta| < 1$ then

$$\begin{aligned} |f(x) - \delta\lambda\tilde{p}(x)| &\leq |f(x)| + |\delta\lambda\tilde{p}(x)| \\ &\leq |g(x)| \cdot (1 - |\lambda\tilde{p}(x)) + |\delta\lambda\tilde{p}(x)| \\ &\leq 1 - (1 - |\delta|) \cdot |\lambda\tilde{p}(x)| \\ &\leq 1 \end{aligned}$$

so that $\delta\lambda\tilde{p}$ minimizes the distance to f independent of $\delta \in (-1, 1)$. Hence, we proved $\neg$(a).

We now prove (ii) which will also establish (b) $\Rightarrow$ (a). Let $\mathcal{F}$ be a T-system. At least one minimal polynomial exists since $\operatorname{lin}\mathcal{F}$ is finite dimensional. Assume $g = \sum_{i=0}^{n} b_i f_i$ fulfills

$$\|f - g\|_\infty = m = \min_{a_0,\ldots,a_n} \left\| f - \sum_{i=0}^{n} a_i f_i \right\|_\infty$$

and $f - g$ takes on the values $\pm m$ alternatively at only $k \leq n + 1$ points. We suppose for definiteness that $f - g$ assumes the values $+m$ before it takes the value $-m$. In this case there exist $k - 1$ points

$$a \leq \quad y_1 < \cdots < y_{k-1} \quad \leq b$$

such that

$$f(y_i) - g(y_i) = 0$$

for all $i = 1, 2, \dots, k-1$ and for some $d > 0$ we have

$$\begin{aligned} m \;&\geq\; f - g \;\geq\; -m + d \qquad && \text{on } [a, y_1] \cup [y_2, y_3] \cup \dots \\ m - d \;&\geq\; f - g \;\geq\; -m \qquad && \text{on } [y_1, y_2] \cup [y_3, y_4] \cup . \end{aligned}$$

By Theorem 4.30 and Remark 4.27 there exists a polynomial h whose *only* zeros on the *open* interval (a, b) are the nodal zeros $y_1, \dots, y_{k-1}$ and additionally $h \leq 0$ on $[a, y_1]$. Let $\delta > 0$ be such that $|\delta h| \leq d/2$ then

$$|f - g + \delta h| < m \tag{12.5}$$

on (a, b).

Equality in (12.5) is possible at the end point a only if $f(a) - g(a) = m$ and $h(a) = 0$ and at b only if $|f(a) - g(b)| = m$ and $h(b) = 0$. To repair the situation at the points a and b let $\tilde{h}$ be such that $\tilde{h} \cdot (f - g) > 0$ at a and b. Then for sufficient small η we have

$$|f - g + \delta h - \eta \tilde{h}| < m$$

on $[a, b]$. Hence, by continuity on the compact interval $[a, b]$ we have

$$\min_{a_0, \dots, a_n} \left\| f - \sum_{i=0}^{n} a_i f_i \right\|_\infty < m$$

contradicting the fact that m is the minimum deviation. That proves (ii) including uniqueness in (i). ◻

In the previous theorem we have seen the close connection between the best approximation polynomials from the minimum problem (12.2) and T-systems. The next result shows that the connection is even closer, i.e., the solution of (12.2) is connected to the Snake Theorem 7.4.

Theorem 12.3 (see e.g. [KS66, p. 283, Thm. 2.1]). *Let $n \in \mathbb{N}_0$ and let $f_0, \dots, f_n, f \in C([a, b], \mathbb{R})$ be such that $\{f_0, \dots, f_n\}$ and $\{f_0, \dots, f_n, f\}$ are continuous T-systems on $[a, b]$ with $a < b$. Let*

$$f^* = c \cdot f + \sum_{i=0}^{n} c_i \cdot f_i$$

be the f^* from the Snake Theorem 7.4 with $g_1 = -1$ and $g_2 = 1$, i.e., f^* is uniquely characterized by the following conditions:

(a) $-1 \leq f^* \leq 1$ on $[a, b]$, and
(b) there exist $n + 2$ points $x_1 < x_2 < \cdots < x_{n+2}$ in $[a, b]$ such that

$$f^*(x_i) = (-1)^{n+1-i}$$

for all $i = 1, \dots, n+2$.

Then $c \neq 0$ and the polynomial

$$\underline{f} := -\frac{1}{c} \cdot \sum_{i=0}^{n} c_i f_i$$

is the unique minimizer of

$$d = \min_{a_0,\dots,a_n} \left\| f - \sum_{i=0}^{n} a_i f_i \right\|_\infty$$

and the minimum deviation is $d = |c|^{-1}$.

The proof is taken from [KS66, pp. 283–284].

Proof. The coefficient c can not be zero. Otherwise the polynomial $\sum_{i=0}^{n} c_i f_i$ vanishes at $n+1$ points in the T-system $\{f_0, \dots, f_n\}$ by (b) and would therefore be equal to zero by Lemma 4.5.

From (a) we get

$$\left\| f - \left(-\frac{1}{d} \sum_{i=0}^{n} c_i f_i \right) \right\|_\infty \leq \frac{1}{|d|}.$$

Since $\underline{f}$ fulfills (b) we get from Theorem 12.2 (ii) uniqueness of $\underline{f}$ and $d = |c|^{-1}$. $\square$

Finding approximations is also done with respect to the $\mathcal{L}^p$-norms

$$\min_{a_0,\dots,a_n} \int \left| f(x) - \sum_{i=0}^{n} a_i f_i(x) \right|^p \mathrm{d}\mu(x) \tag{12.6}$$

with a fixed measure μ and $p \geq 1$. For $p = 2$ this leads to the well-studied *orthogonal polynomials*, a special branch of the theory of moments.

For $p = 1$ in (12.6) this also is connected to T-systems. D. Jackson [Jac24] showed that if $\mathcal{F} = \{f_0, \dots, f_n\}$ is a T-system then the best approximation of (12.6) is unique, see also [Ach56, p. 77].

12.2 Optimization over Linear Functionals

In optimization one often encounters the problem of having only a linear functional $L : \mathcal{V} \to \mathbb{R}$, e.g. a moment functional, and one wants to minimize $L(f)$ over $\mathcal{V}_+$. By removing the dependency on the scaling of f we get the following result.

Theorem 12.4 (see e.g. [KS66, p. 312, Thm. 9.1])**.** *Let $n \in \mathbb{N}_0$, let $\mathcal{F} = \{f_i\}_{i=0}^n$ be an ET-system on $[a,b]$ with $a < b$, and let $L, S : \operatorname{lin}\mathcal{F} \to \mathbb{R}$ be two linear functionals such that S is strictly positive on $(\operatorname{lin}\mathcal{F})_+$, i.e., $S(f) > 0$ for all $f \in \operatorname{lin}\mathcal{F} \setminus \{0\}$ with $f \geq 0$. Then*

$$\min_{f\in(\operatorname{lin}\mathcal{F})_+\setminus\{0\}} \frac{L(f)}{S(f)} \qquad \textit{and} \qquad \max_{f\in(\operatorname{lin}\mathcal{F})_+\setminus\{0\}} \frac{L(f)}{S(f)} \tag{12.7}$$

are attained at non-negative polynomials possessing n zeros counting multiplicities.

The proof is taken from [KS66, p. 312].

Proof. Since $\operatorname{lin}\mathcal{F}$ is finite dimensional the values in (12.7) are attained.

It is sufficient to prove the statement for the maximum. But maximizing $\frac{L(f)}{S(f)}$ over $(\operatorname{lin}\mathcal{F})_+ \setminus \{0\}$ is equivalent to maximize $L(f)$ over $f \in (\operatorname{lin}\mathcal{F})_+ \setminus \{0\}$ with $S(f) = 1$.

Let $f \geq 0$ be such that $S(f) = 1$ and suppose f has at most $n-1$ zeros counting multiplicities. Then by Karlin's Nichtnegativstellensatz 7.6 there is a unique decomposition $f = f_* + f^*$ where f_* and f^* differ, are non-negative, and both have n zeros counting multiplicities. Set $\alpha := S(f_*)$ and $\beta := S(f^*)$. Then $\alpha, \beta > 0$ since S is strictly positive and $\alpha + \beta = S(f_*) + S(f^*) = S(f) = 1$. Then

$$f = \alpha \cdot \frac{f_*}{\alpha} + \beta \cdot \frac{f^*}{\beta}$$

and by linearity

$$L(f) \leq \max\left(\frac{L(f_*)}{\alpha}, \frac{L(f^*)}{\beta}\right)$$

which proves the statement. □

More results on best approximation and optimization over linear functionals can already be found in [Ber26], [Ach56], and [KS66]. Let alone the enormous literature after that.

Appendices

Solutions

Problems of Chapter 1

1.1 The Stone–Weierstrass Theorem 0.3 states that for a compact set $K \subset \mathbb{R}^n$ the polynomials $\mathbb{R}[x_1, \dots, x_n]$ are dense in $C(K, \mathbb{R})$ with respect to the sup-norm. Let $A \in \mathfrak{B}(K)$ be a Borel measurable set, let $\varepsilon > 0$, and let μ_1 and μ_2 be two representing measures of L. Set $A_\delta := (A + B_\delta(0)) \cap K$ for all $\delta > 0$. Then for any $\varepsilon > 0$ there exists a $\delta = \delta(\varepsilon) > 0$ such that $\mu_1(A_\delta \setminus A), \mu_2(A_\delta \setminus A) < \varepsilon$.

By Urysohn's Lemma 0.2 there exists a $\varphi_\varepsilon \in C(K, [0,1])$ such that

$$\varphi_\varepsilon(x) = \begin{cases} 1 & \text{for } x \in A \\ 0 & \text{for } x \in K \setminus A_\varepsilon \end{cases}$$

and since $\mathbb{R}[x_1, \dots, x_n]$ is dense in $C(K, \mathbb{R})$ there exists a family of polynomials $(p_i^\varepsilon)_{i \in \mathbb{N}} \subseteq \mathbb{R}[x_1, \dots, x_n]$ such that

$$\|p_i^\varepsilon - \varphi_\varepsilon\|_\infty \xrightarrow{i \to \infty} 0 \quad \text{and hence} \quad \int_K p_i^\varepsilon(x)\, \mathrm{d}\mu_j(x) \xrightarrow{i \to \infty} \int_K \varphi_\varepsilon(x)\, \mathrm{d}\mu_j(x)$$

for $j = 1, 2$. Then we have

$$\begin{aligned} \mu_1(A) &= \lim_{\varepsilon \searrow 0} \int_K \varphi_\varepsilon(x)\, \mathrm{d}\mu_1(x) \\ &= \lim_{\varepsilon \searrow 0} \lim_{i \to \infty} \int_K p_i^\varepsilon(x)\, \mathrm{d}\mu_1(x) \\ &= \lim_{\varepsilon \searrow 0} \lim_{i \to \infty} L(p_i^\varepsilon) \\ &= \lim_{\varepsilon \searrow 0} \lim_{i \to \infty} \int_K p_i^\varepsilon(x)\, \mathrm{d}\mu_2(x) \\ &= \lim_{\varepsilon \searrow 0} \int_K \varphi_\varepsilon(x)\, \mathrm{d}\mu_2(x) \qquad = \mu_2(A). \end{aligned}$$

Since $A \in \mathfrak{B}(K)$ was arbitrary we have $\mu_1 = \mu_2$, i.e., L has a unique representing measure and is therefore determinate.

1.2 Proof of Corollary 1.3
Let $\mu_1, \mu_2 \in \mathcal{M}(L)$ and $\lambda \in [0, 1]$. Then

$$\begin{aligned}\int p(x)\ \mathrm{d}[\lambda\mu_1 + (1-\lambda)\mu_2](x) &= \lambda \int p(x)\ \mathrm{d}\mu_1(x) + (1-\lambda) \int p(x)\ \mathrm{d}\mu_2(x)\\ &= \lambda L(p) + (1-\lambda)L(p)\\ &= L(p)\end{aligned}$$

and hence $\lambda\mu_1 + (1-\lambda)\mu_2 \in \mathcal{M}(L)$ which proves convexity.

1.3 Proof of Corollary 1.10
Let $\mu_0, \mu_1 \in \mathcal{M}(L)$ with $\mu_0 \neq \mu_1$, i.e., there exists a $A \in \mathfrak{A}$ such that $\mu_0(A) \neq \mu_1(A)$ and without loss of generality we have $\mu_0(A) < \mu_1(A)$. Hence, for all $\lambda \in [0, 1]$ we set $\mu_\lambda := \lambda\mu_1 + (1-\lambda)\mu_0$ and we have

$$\mu_{\lambda_0}(A) < \mu_{\lambda_1}(A)$$

for all $0 \leq \lambda_0 < \lambda_1 \leq 1$ which proves that $\mu_{\lambda_0} \neq \mu_{\lambda_1}$ for all $\lambda_0 \neq \lambda_1$. Hence, we have at least $|[0, 1]| = |\mathbb{R}|$ many representing measures for L.

Problems of Chapter 2

2.1 Proof of Lemma 2.1
The proof is taken from [Cho69, Vol. 2, p. 268].

(i) $\Rightarrow$ (ii): If $F+C$ is a vector space then $-(F+C) = (F+C)$ and $-(F+C) = F-C$ since $-F = F$.

(ii) $\Rightarrow$ (iii): If $x \in F + C$, i.e., $x = y' + z$ for some $y' \in F$ and $z \in C$, then $x \geq y'$. Similarly, if $x = y - w$ then $y \geq x$.

(iii) $\Rightarrow$ (i): First note that $F + C$ is a convex cone. So if suffices to show that $F + C = -(F + C)$, i.e., $F + C = F - C$. But if $x \in F + C$ and $x \leq y$ then $x = y - z$ for some $z \in C$, or $x \in F - C$. Similarly, if $x \in F - C$ and $x = y' + w$ for some $w \in C$ then $x \in F + C$.

2.2 Proof of Lemma 2.6

(i) $\Rightarrow$ (ii): Set $K_\varepsilon = \operatorname{supp} h_\varepsilon$.

(ii) $\Rightarrow$ (iii): Chose by Urysohn's Lemma 0.2 a $\eta_\varepsilon \in C_c(\mathcal{X}, \mathbb{R})$ with $\eta_\varepsilon|_{K_\varepsilon} = 1$.

(iii) $\Rightarrow$ (i): Take $h_\varepsilon = \eta_\varepsilon \cdot g \in C_c(\mathcal{X}, \mathbb{R})$.

2.3 Since $\mathcal{X}$ is compact for every $f \in E$ we have $m_f := \min_{x\in\mathcal{X}} f(x) > -\infty$ and $M_f := \max_{x\in\mathcal{X}} f(x) < \infty$, especially for $f = e > 0$ we have $m_e > 0$. Then for every f there exists a $d_f > 0$ such that $f = (f + d_f e) - d_f e$ such that $f + d_f e, d_f e \in E_+$ and hence $E = E_+ - E_+$ proving (i) in Definition 2.7.

Since $e > 0$ we also have (ii) in Definition 2.7.

For (iii) in Definition 2.7 it is sufficient to note that $\mathcal{X}$ is compact, i.e., for every g there is a $c_g > 0$ such that $g \leq c_g e$.

2.4 Let $E = \mathbb{R}[x_1, \ldots, x_n]$ on $\mathcal{X}$. Then (i) $E = E_+ - E_+$ follows immediately from the fact that for every $f \in E$ there is a $g \in E_+$ such that $f = f + g - g$ with $f + g \in E_+$.

For (ii) we take $f = 1 > 0$ on $\mathcal{X}$.

For (iii) take the g from (i).

2.5 Since E is finite dimensional we can equip it with a norm, e.g. the l^2-norm in the coefficients of f. Assume $\mathcal{X}$ is not compact then there exists an unbounded sequence $(x_i)_{i\in\mathbb{N}_0}$ and a $f \in E$ with $\|f\| \leq 1$ such that $(f(x_i))_{i\in\mathbb{N}_0}$ grows faster than any other $(g(x_i))_{i\in\mathbb{N}_0}$. Hence, f can not be dominated by any g.

2.6 Proof of Lemma 2.8
Since $K = \operatorname{supp} g$ is compact and E is an adapted space, i.e., there exists a $f \in E_+$ with $f > 0$ we have that $\min_{x\in K} f(x) > 0$ and hence there exists a $c > 0$ such that $cf > g$ on K and hence on all $\mathcal{X}$.

Problems of Chapter 3

3.1 Proof of Stieltjes' Theorem 3.1
We have (iii) $\Leftrightarrow$ (iv) $\Leftrightarrow$ (v) by the definition of the Hankel matrix and also (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii). Additionally, we have (iii) $\Rightarrow$ (ii) by (3.4) since $L(p) = L(f^2) + L(xg^2) \geq 0$. At last (ii) $\Rightarrow$ (i) holds by the Basic Representation Theorem 2.9 since $\mathbb{R}[x]$ on $[0, \infty)$ is an adapted space.

3.2 Proof of Hamburger's Theorem 3.2
We have (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) and additionally (iii) $\Leftrightarrow$ (iv) $\Leftrightarrow$ (v) by the definition of the Hankel matrix. The implication (iii) $\Rightarrow$ (ii) follows from Equation (3.2) by $L(p) = L(f^2 + g^2) \geq 0$. At last (ii) $\Rightarrow$ (i) holds by the Basic Representation Theorem 2.9 since $\mathbb{R}[x]$ on $\mathbb{R}$ is an adapted space.

3.3 Proof of Hausdorff's Theorem 3.3
We have (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) and additionally (iii) $\Leftrightarrow$ (iv) $\Leftrightarrow$ (v) by the definition of the Hankel matrix. The implication (iii) $\Rightarrow$ (ii) follows from (3.9) since it is sufficient to look only at $f(x)^2 + xg(x)^2 + (1-x)h(x)^2$. At last (ii) $\Rightarrow$ (i) holds by the Basic Representation Theorem 2.9 since $\mathbb{R}[x]$ on $[0, 1]$ is an adapted space.

3.4 Proof of Haviland's Theorem 3.4
Since (i) $\Rightarrow$ (ii) is clear it is sufficient to show (ii) $\Rightarrow$ (i). But since $E = \mathbb{R}[x_1, \ldots, x_n]$ on K, is an adapted space (see Problem 2.4) and since $E_+ = \operatorname{Pos}(K)$ by definition the Basic Representation Theorem 2.9 applies and gives the assertion.

3.5 Proof of Corollary 3.6

We have that (ii) $\Rightarrow$ (i) is clear since $x^k \cdot (1-x)^l > 0$ on $(0,1)$ and at least one $c_{k',l'} > 0$. It remains to prove (i) $\Rightarrow$ (ii).

Let $f \in \mathbb{R}[x] \setminus \{0\}$ with $f > 0$ on $(0,1)$ then we can write f as

$$f(x) = x^p \cdot (1-x)^q \cdot \tilde{f}(x)$$

with $\tilde{f} \in \mathbb{R}[x]$, $\tilde{f} > 0$ on $[0,1]$, and $p, q \in \mathbb{N}_0$, i.e., by the fundamental theorem of algebra we can factor out the zeros at $x = 0$ and at $x = 1$. Applying Bernstein's Theorem 3.5 (ii) to $\tilde{f}$ then gives the assertion.

3.6 Proof of Lemma 3.9

Since the moment cone $\mathcal{S}_\mathcal{F}$ and the hyperplane H are convex we have that $\mathcal{S}_\mathcal{F} \cap H$ is a convex cone, i.e., it is a moment cone and there exists a family $\mathcal{G} \subsetneq \operatorname{lin} \mathcal{F}$ of $m < n$ elements which spans $\mathcal{S}_\mathcal{F} \cap H$. It is sufficient to show that $\mathcal{G}$ lives on $(\mathcal{Y}, \mathfrak{A}|_\mathcal{Y})$ for some $\mathcal{Y} \subseteq \mathcal{X}$.

For the hyperplane H there exists a function $h \in \operatorname{lin} \mathcal{F}$ such that $L_s(h) \geq 0$ for all $s \in \mathcal{S}_\mathcal{F}$. Note, that $\mathcal{N} = \cap_{k \in \mathbb{N}} \{x \in \mathcal{X} \mid f_1(x)^2 + \cdots + f_n(x)^2 \geq k\}$ has measure zero for any representing measure μ_s on $\mathcal{X}$ of a moment sequence $s \in \mathcal{S}_\mathcal{F}$ since the moments are finite, i.e., the f_i are μ_s-integrable. Without loss of generality we can therefore work on $\mathcal{X} \setminus \mathcal{N}$. Hence, all δ_x with $x \in \mathcal{X} \setminus \mathcal{N}$ are moment measures and $L_s(h) \geq 0$ implies $h \geq 0$ on $\mathcal{X} \setminus \mathcal{N}$.

Then $s \in \mathcal{S}_\mathcal{F} \cap H \Leftrightarrow L_s(h) = 0$ implies that all representing measures μ of all $s \in \mathcal{S}_\mathcal{F} \cap H$ have the support in $\mathcal{Y} := \{x \in \mathcal{X} \setminus \mathcal{N} \mid h(x) = 0\}$.

3.7 Let $\mathcal{F} = \{f_1, \ldots, f_n\}$ be measurable functions on $(\mathcal{X}, \mathfrak{A})$ which are not necessarily bounded. Set

$$I := \bigcap_{k \in \mathbb{N}} \{x \in \mathcal{X} \mid |f_i(x)| > k \text{ for all } i = 1, \ldots, n\}.$$

Then I is measurable. Let s be a moment sequence with representing measure μ. Since all f_i are μ-measurable we have $\mu(I) = 0$. Therefore, by working on $\mathcal{X} \setminus I$ we can assume without loss of generality that $|f_i(x)| < \infty$ for all $x \in \mathcal{X}$.

Define $\mathcal{G} = \{g_1, \ldots, g_n\}$ with $g_i := \frac{g_i}{f}$ and $f := 1 + \sum_{i=1}^n f_i^2$.

At first we note that from

$$\int_\mathcal{X} f_i(x)\, \mathrm{d}\mu(x) = \int_\mathcal{X} g_i(x) \cdot f(x)\, \mathrm{d}\mu = \int_\mathcal{X} g_i(x)\, \mathrm{d}\nu(x), \tag{S.1}$$

we have that every sequence $s = (s_1, \ldots, s_n)$ is a moment sequence with respect to $\mathcal{G}$ if and only if it is moment sequence with respect to $\mathcal{F}$.

Since all g_i are bounded we have by Rosenbloom's Theorem that there is a k-atomic representing measure $\nu = \sum_{i=1}^k c_i \cdot \delta_{x_i}$ which represents the moment sequence s. Then by (S.1) we find that $\mu = \sum_{i=1}^k c_i \cdot f(x_i)^{-1} \cdot \delta_{x_i}$ is a representing measure of s with respect to $\mathcal{F}$ which proves the statement.

Problems of Chapter 4

4.1 Proof of Corollary 4.3

Let $f \in \operatorname{lin} \mathcal{F}$. Then f has at most n zeros in $\mathcal{X}$ and hence $f|_{\mathcal{Y}}$ has at most n zeros in $\mathcal{Y} \subset \mathcal{X}$. Since for any $g \in \operatorname{lin} \mathcal{G}$ there is a $f \in \operatorname{lin} \mathcal{F}$ such that $g = f|_{\mathcal{Y}}$ we have the assertion.

4.2 Proof of Corollary 4.8

Let $w_0, \dots, w_n \in \mathcal{W}$ be pairwise distinct. Since g is injective we have that also $g(w_0), \dots, g(w_n) \in \mathcal{X}$ are pairwise distinct. Hence,

$$\det \begin{pmatrix} g_0 & g_1 & \dots & g_n \\ w_0 & w_1 & \dots & w_n \end{pmatrix} = \det \begin{pmatrix} f_0 & f_1 & \dots & f_n \\ g(w_0) & g(w_1) & \dots & g(w_n) \end{pmatrix} \neq 0$$

and the statement follows from Lemma 4.5.

4.3 Proof of Corollary 4.9

Let $x_0, \dots, x_n \in \mathcal{X}$ be pairwise distinct. Then

$$\det \begin{pmatrix} g_0 & g_1 & \dots & g_n \\ x_0 & x_1 & \dots & x_n \end{pmatrix} = \det \begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x_0 & x_1 & \dots & x_n \end{pmatrix} \cdot g(x_1) \cdots g(x_n) \neq 0$$

and the statement follows from Lemma 4.5.

4.4 Proof of Corollary 4.10

(i) Assume $f_0, \dots, f_n$ are linearly dependent, i.e., there are $a_0, \dots, a_n \in \mathbb{R}$ not all zero such that $a_0 f_0 + \dots + a_n f_n$ is the zero polynomial. Hence, f has at least $n + 1$ zeros. But since $\mathcal{F}$ is a T-system this is a contradiction.

(ii) Let $x_0, \dots, x_n \in \mathcal{X}$ be $n+1$ pairwise distinct points. Then by Definition 4.4 we have

$$\begin{pmatrix} f(x_0) \\ \vdots \\ f(x_n) \end{pmatrix} = \underbrace{\begin{pmatrix} f_0 & \dots & f_n \\ x_0 & \dots & x_n \end{pmatrix}}_{=:M} \cdot \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix}$$

and since $\mathcal{F}$ is a T-system we have that M has full rank by Lemma 4.5. Hence, the coefficients $a_0, \dots, a_n$ are unique.

4.5 Proof of Example 4.18

Set $f_i(x) := (x + \alpha_i)^{-1}$ and $g(x) = (x + \alpha_0) \cdots (x + \alpha_n)$. Then $g > 0$ on $[a, b]$ since $-\alpha_0 < a < b$. Hence, $\mathcal{F}$ is a T-system on $[a, b]$ if and only if $\mathcal{G} = \{g_i := g \cdot f_i\}_{i=0}^n$ is a T-system on $[a, b]$ by Corollary 4.9.

We have $g_i(x) = (x + \alpha_0) \cdots (x + \alpha_{i-1}) \cdot (x + \alpha_{i+1}) \cdots (x + \alpha_n)$ and $\deg g_i = n$. It is now sufficient to show that $\mathcal{G}$ is a T-system on $\mathbb{R}$ by Corollary 4.3 since then it will also be a T-system on $[a, b]$.

Since $g_i(\alpha_j) = 0$ for all $i \neq j$ we have that the g_i are linearly independent. Hence, $\operatorname{lin} \mathcal{G} = \mathbb{R}[x]_{\leq n}$. But since $\{x^i\}_{i=0}^n$ is a T-system so is $\mathcal{G}$ since every non-trivial $f \in \operatorname{lin} \mathcal{G} = \mathbb{R}[x]_{\leq n}$ has at most n zeros.

In summary, we have that $\{x^i\}_{i=0}^n$ is a T-system on $\mathbb{R} \Rightarrow \mathcal{G}$ on $\mathbb{R}$ is a T-system $\Rightarrow$ $\mathcal{G}$ on $[a,b]$ is a T-system $\Rightarrow \mathcal{F}$ on $[a,b]$ is a T-system.

4.6 To the points $x_0, \ldots, x_{k+l} \in [a,b]$ add pairwise distinct points $x_{k+l+1}, \ldots, x_n \in [a,b] \setminus \{x_0, \ldots, x_{k+l}$. Then the matrix

$$\begin{pmatrix} f_0(x_0) & \ldots & f_n(x_0) \\ \vdots & & \vdots \\ f_0(x_n) & \ldots & f_n(x_n) \end{pmatrix} \tag{S.2}$$

has full rank since $\mathcal{F}$ is a T-system, i.e., every vector, especially

$$(m, \ldots, m, -m, \ldots, -m, 0, \ldots, 0, *, \ldots, *)^T \in \mathbb{R}^{n+1}$$

is in its image. But the matrix

$$\begin{pmatrix} f_0(x_1) & \ldots & f_n(x_1) \\ \vdots & & \vdots \\ f_0(x_{k+l}) & \ldots & f_n(x_{k+l}) \end{pmatrix}$$

in (4.7) only contains the first $k + l$ rows of (S.2), i.e., (4.7) has at least one solution.

4.7 By Remark 4.27 only the case $n = 2m + 2p$ and one end point is contained. But then we can apply Theorem 4.26 to $\tilde{\mathcal{F}} = \{f_i\}_{i=0}^{n-1}$ which ensures by the same arguments in Remark 4.27 that $x_1, \ldots, x_p$ are the only zeros of some $f \geq 0$.

Problems of Chapter 5

5.1 Proof of Lemma 5.7

Set $g_i := g \cdot f_i$. Then we have to check that

$$\mathcal{W}(g_0, \ldots, g_k)(x) = \det \begin{pmatrix} g_0(x) & g_1(x) & \ldots & g_n(x) \\ g_0'(x) & g_1'(x) & \ldots & g_n'(x) \\ \vdots & \vdots & & \vdots \\ g_0^{(n)}(x) & g_1^{(n)}(x) & \ldots & g_n^{(n)}(x) \end{pmatrix} \neq 0$$

holds for all $x \in [a,b]$. Since $g_i = g \cdot f_i$ we apply the product rule and get

$$\mathcal{W}(g_0,\dots,g_k)(x) = g^2 \cdot \det\begin{pmatrix} f_0(x) & f_1(x) & \dots & f_n(x) \\ f_0'(x) & f_1'(x) & \dots & f_n'(x) \\ g_0''(x) & g_1''(x) & \dots & g_n''(x) \\ \vdots & \vdots & & \vdots \\ g_0^{(n)}(x) & g_1^{(n)}(x) & \dots & g_n^{(n)}(x) \end{pmatrix}$$

since in the first line we factored out g and then subtracted g'-times the first line from the second, and factored out g from the remaining second line. For the second derivatives in the third line we have

$$(g \cdot f_i)'' = g'' \cdot f_i + 2g' \cdot f_i' + g \cdot f_i''$$

and hence subtracting g''-times the first row, $2g'$-times the second row, and finally factoring out g from the remaining third row we get

$$\mathcal{W}(g_0,\dots,g_k)(x) = g^3 \cdot \det\begin{pmatrix} f_0(x) & f_1(x) & \dots & f_n(x) \\ f_0'(x) & f_1'(x) & \dots & f_n'(x) \\ f_0''(x) & f_1''(x) & \dots & f_n''(x) \\ g_0'''(x) & g_1'''(x) & \dots & g_n'''(x) \\ \vdots & \vdots & & \vdots \\ g_0^{(n)}(x) & g_1^{(n)}(x) & \dots & g_n^{(n)}(x) \end{pmatrix}.$$

Proceeding in this manner we arrive at

$$\mathcal{W}(g_0,\dots,g_k)(x) = g^{n+1} \cdot \mathcal{W}(f_0,\dots,f_n)(x) \neq 0$$

for all $x \in [a,b]$ which proves the statement.

5.2 Proof of Lemma 5.8

We proceed similar to Problem/Solution 5.1 but now with the rule of differentiation for $f_i \circ g$. We have

$$(f_i \circ g)' = g' \cdot (f_i' \circ g)$$

and hence

$$\mathcal{W}(g_0,\dots,g_n) = g' \cdot \det\begin{pmatrix} f_0 \circ g & \dots & f_n \circ g \\ f_0' \circ g & \dots & f_n' \circ g \\ (f_0 \circ g)'' & \dots & (f_n \circ g)'' \\ \vdots & & \vdots \\ (f_0 \cdot g)^{(n)} & \dots & (f_n \circ g)^{(n)} \end{pmatrix}$$

by factoring out g' from the second row. Then we have

$$(f_i \circ g)'' = (g' \cdot (f_i' \circ g))' = g'' \cdot (f_i' \circ g) + (g')^2 \cdot (f_i'' \circ g),$$

i.e., we subtract g''-times the second row and factor out $(g')^2$ to get

$$\mathcal{W}(g_0,\dots,g_n) = (g')^3 \cdot \det\begin{pmatrix} f_0\circ g & \dots & f_n\circ g \\ f_0'\circ g & \dots & f_n'\circ g \\ f_0''\circ g & \dots & f_n''\circ g \\ (f_0\circ g)''' & \dots & (f_n\circ g)''' \\ \vdots & & \vdots \\ (f_0\cdot g)^{(n)} & \dots & (f_n\circ g)^{(n)} \end{pmatrix}.$$

Proceeding in this manner with

$$(f_i\circ g)^{(k)} = (g')^{(k)}\cdot (f_i^{(k)}\circ g) + \dots + g^{(k)}\cdot (f_i'\circ g)$$

we get

$$\mathcal{W}(g_0,\dots,g_n) = (g')^{\frac{n(n+1)}{2}}\cdot \mathcal{W}(f_0,\dots,f_n)\circ g$$

with proves the assertion.

5.3 Proof of Lemma 5.9
Set $\mathcal{H} = \{h_i\}_{i=0}^n$ with $h_i := \frac{f_i}{f_0}$. Then by Lemma 5.7 we have

$$\mathcal{W}(f_0,\dots,f_n) = f_0^{n+1}\cdot \mathcal{W}(h_0,\dots,h_n)$$

and since $h_0 = 1$ we have $h_0' = h_0'' = \dots = 0$ and

$$= f_0^{n+1}\cdot \det\begin{pmatrix} 1 & h_1 & \dots & h_n \\ 0 & h_1' & \dots & h_n' \\ \vdots & \vdots & & \vdots \\ 0 & h_1^{(n)} & \dots & h_n^{(n)} \end{pmatrix}$$

which gives by expanding along the first column

$$\begin{aligned} &= f_0^{n+1}\cdot \det\begin{pmatrix} h_1' & \dots & h_n' \\ \vdots & & \vdots \\ h_1^{(n)} & \dots & h_n^{(n)} \end{pmatrix} \\ &= f_0^{n+1}\cdot \mathcal{W}(h_1',\dots,h_n') \end{aligned}$$

and with $g_i = h_{i+1}'$ for $i = 0,\dots,n-1$ we get

$$= f_0^{n+1}\cdot \mathcal{W}(g_0,\dots,g_{n-1})$$

which proves the statement.

5.4 (a) Since $\mathcal{F}$ is an ET-system on $[a,b]$ we have

$$\mathcal{W}(f_0,\dots,f_n)(x) \neq 0$$

for all $x \in [a,b]$, i.e., also for all $x \in [a',b'] \subseteq [a,b]$ and hence it is an ET-system on $[a',b']$.
(b) Since $\mathcal{F}$ is an ECT-system on $[a,b]$ we have

$$\mathcal{W}(f_0,\dots,f_k)(x) \neq 0$$

for all $x \in [a,b]$ and $k = 0,\dots,n$, i.e., also for all $x \in [a',b'] \subseteq [a,b]$ and $k = 0,\dots,n$ and hence it is an ECT-system on $[a',b']$.

5.5 Proof of Example 4.19
We already know that $\{1,x,x^2,\dots,x^k\}$ is an ET-system for any $k = 0,1,\dots,n$ since

$$cW(1,x,\dots,x^k)(x) = 1 \cdot 1! \cdot \dots \cdot k! > 0.$$

From the Wronskian determinant

$$\mathcal{W}(1,x,\dots,x^n,f)(x) = 1 \cdot 1! \cdot 2! \cdot \dots \cdot n! \cdot f^{(n)}(x) > 0$$

we then get that $\mathcal{F}$ is an ECT-system on $[a,b]$ by Theorem 5.12.

5.6 Proof of Examples 5.18
By Lemma 5.8 we only need to prove the statement for one case, say case (b) $\mathcal{G} = \{e^{\alpha_i x}\}_{i=0}^n$. Let $k \in \{0,1,\dots,n\}$. Then

$$\mathcal{W}(g_0,\dots,g_k) = \det\begin{pmatrix} g_0 & g_1 & \dots & g_k \\ g_0' & g_1' & \dots & g_k' \\ \vdots & \vdots & & \vdots \\ g_0^{(k)} & g_1^{(k)} & \dots & g_k^{(k)} \end{pmatrix}$$

and with $g_i^{(j)} = \alpha_i^j \cdot g_i$ we get

$$\begin{aligned} &= \deg\begin{pmatrix} g_0 & g_1 & \dots & g_k \\ \alpha_0 g_0 & \alpha_1 g_1 & \dots & \alpha_k g_k \\ \vdots & \vdots & & \vdots \\ \alpha_0^k g_0 & \alpha_1^k g_1 & \dots & \alpha_k^k g_k \end{pmatrix} = g_0 \cdot g_1 \cdots g_n \cdot \det\begin{pmatrix} 1 & 1 & \dots & 1 \\ \alpha_0 & \alpha_1 & \dots & \alpha_k \\ \vdots & \vdots & & \vdots \\ \alpha_0^k & \alpha_1^k & \dots & \alpha_k^k \end{pmatrix} \\ &= g_0 \cdot g_1 \cdots g_k \cdot \prod_{0 \le i < j \le k} (\alpha_j - \alpha_i) \neq 0 \end{aligned}$$

which proves the statement.

5.7 To construct the non-negative polynomial on $[0,\infty)$ with the double zero $x_1 = 1$ and the zero $x_2 = 2$ with algebraic multiplicity $m_2 = 4$ we need 7 monomials. We chose $f_0(x) = 1, f_1(x) = x^2, f_2(x) = x^3, f_3(x) = x^5, f_4(x) = x^8, f_5(x) = x^{11}, f_6(x) = x^{13}$ and leave out x^{42}. With (5.5) we get

$$
\begin{aligned}
f(x) &= \det\begin{pmatrix} f_0 & f_1 & f_2 & f_3 & f_4 & f_5 & f_6 \\ x & 1 & 1 & 2 & 2 & 2 & 2 \end{pmatrix}\\
&= \det\begin{pmatrix}
f_0(x) & f_1(x) & f_2(x) & f_3(x) & f_4(x) & f_5(x) & f_6(x)\\
f_0(x_1) & f_1(x_1) & f_2(x_1) & f_3(x_1) & f_4(x_1) & f_5(x_1) & f_6(x_1)\\
f_0'(x_1) & f_1'(x_1) & f_2'(x_1) & f_3'(x_1) & f_4'(x_1) & f_5'(x_1) & f_6'(x_1)\\
f_0(x_2) & f_1(x_2) & f_2(x_2) & f_3(x_2) & f_4(x_2) & f_5(x_2) & f_6(x_2)\\
f_0'(x_2) & f_1'(x_2) & f_2'(x_2) & f_3'(x_2) & f_4'(x_2) & f_5'(x_2) & f_6'(x_2)\\
f_0''(x_2) & f_1''(x_2) & f_2''(x_2) & f_3''(x_2) & f_4''(x_2) & f_5''(x_2) & f_6''(x_2)\\
f_0'''(x_2) & f_1'''(x_2) & f_2'''(x_2) & f_3'''(x_2) & f_4'''(x_2) & f_5'''(x_2) & f_6'''(x_2)\\
f_0^{(4)}(x_2) & f_1^{(4)}(x_2) & f_2^{(4)}(x_2) & f_3^{(4)}(x_2) & f_4^{(4)}(x_2) & f_5^{(4)}(x_2) & f_6^{(4)}(x_2)
\end{pmatrix}\\
&= \det\begin{pmatrix}
1 & x^2 & x^3 & x^5 & x^8 & x^{11} & x^{13}\\
1 & 1 & 1 & 1 & 1 & 1 & 1\\
0 & 2 & 3 & 5 & 8 & 11 & 13\\
1 & 4 & 8 & 32 & 256 & 2\,048 & 8\,192\\
0 & 4 & 12 & 80 & 1\,024 & 11\,264 & 53\,248\\
0 & 2 & 12 & 160 & 3\,584 & 56\,320 & 319\,488\\
0 & 0 & 6 & 240 & 10\,752 & 253\,440 & 1\,757\,184
\end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
f(x) = 48\cdot\Big(&14\,980\,788x^{13} - 184\,325\,420x^{11} + 2\,421\,354\,616x^8 - 26\,336\,028\,160x^5\\
&+112\,945\,898\,496x^3 - 112\,347\,781\,120x^2 + 23\,485\,900\,800\Big).
\end{aligned}
$$

The function f is shown in Figure S.1.

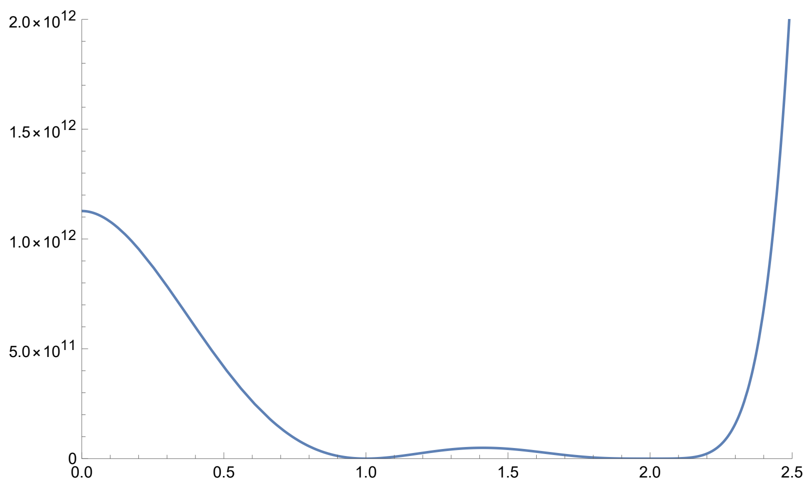


Fig. S.1: The function f from the solution of Problem 5.7.

This function f we gave here is not unique. Of course every multiple of f also fulfills the requirements but we also made the restrictions to use all monomials except

x^{42}. We get another polynomial when we e.g. leave out x^{13} (or any other monomial except 1) instead of x^{42}. Then any conic linear combination of these functions also fulfills the requirements.

We can not leave out 1 since any linear combination has the additional zero $x = 0$.

Problems of Chapter 6

6.1 Proof of Corollary 6.8

Since $\mathcal{F}$ is a continuous T-system we can assume that

$$\det\begin{pmatrix} f_0 & f_1 & \dots & f_n \\ x_0 & x_1 & \dots & x_n \end{pmatrix} > 0$$

for all $a \leq x_0 < x_1 < \dots < x_n \leq b$. Since the Gaussian kernel is ETP_k for every $k \in \mathbb{N}_0$, see Example 6.6, we have

$$K_\sigma^* \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \end{pmatrix} > 0$$

for all $x_1 < x_2 < \dots < x_n$ and $y_1 \leq y_2 \leq \dots \leq y_n$ in $\mathbb{R}$ as well as $\sigma > 0$. Hence, in $\mathcal{W}(f_{\sigma,0}, f_{\sigma,1}, \dots, f_{\sigma,n})(x) = (6.6)$ in Lemma 6.7 we are integrating over a non-negative functions with respect to the Lebesgue measure $\mu = \lambda$, i.e., $\mathcal{W}(f_{\sigma,0}, f_{\sigma,1}, \dots, f_{\sigma,n})(x) > 0$ for all $x \in [a, b]$ which proves the statement.

Problems of Chapter 7

7.1 The family $\mathcal{F}$ on $[a, b]$ needs for a fixed $f \geq 0$ only be an ET-system around the zeros of f but otherwise the proof of Karlin's Theorem 7.1 is employed, i.e., there we only need $\mathcal{F}$ to be a T-system.

Problems of Chapter 8

8.1 Proof of Karlin's Positivstellensatz 8.4 on $\mathbb{R}$

By (a) there exists a function $w \in C(\mathbb{R}, \mathbb{R})$ such that $w > 0$ on $\mathbb{R}$ and

$$\lim_{x\to\infty} \frac{f_n(x)}{w(x)} = 1.$$

By (b) we define

$$v_i(x) := \begin{cases} \frac{f_i(x)}{w(x)} & \text{if } x \in \mathbb{R}, \\ \delta_{i,n} & \text{if } x = \pm\infty \end{cases}$$

for all $i = 0, 1, \ldots, n$. Then by (c) and Corollary 4.9 we have that $\{v_i\}_{i=0}^n$ is a T-system on $[0, \infty]$. With $t(x) := \tan(\pi x/2)$ we define

$$g_i(x) := v_i \circ t$$

for all $i = 0, 1, \ldots, n$. Hence, $\mathcal{G} = \{g_i\}_{i=0}^n$ is a T-system on $[-1, 1]$ by Corollary 4.8. We now apply Karlin's Positivstellensatz 7.3 to $\mathcal{G}$. Set $g := (\frac{f}{w}) \circ t$.

(ii): By Karlin's Positivstellensatz 7.3 on $[a, b]$ there exist points

$$-1 = y_0 < x_1 < y_1 < \cdots < x_m < y_m = 1$$

and unique functions g_* and g^* such that $g = g_* + g^*$, $g_*, g^* \geq 0$ on $[-1, 1]$, $x_1, \ldots, x_m$ are the zeros of g_*, and $y_0, \ldots, y_m$ are the zeros of g^*. Then $f_* := (g_* \circ t^{-1}) \cdot w$ and $f^* := (g^* \circ t^{-1}) \cdot w$ are the unique components in the decomposition $f = f_* + f^*$.

(i): Since $g^*(y_0) = g^*(y_m) = 0$ we have that g^* contains no g_{2m} and hence the coefficient of g_{2m} in g_* is a_{2m}.

8.2 Proof of Karlin's Nichtnegativstellensatz 8.5 on $\mathbb{R}$

Similar to the proof of Karlin's Nichtnegativstellensatz 8.3 on $[0, \infty)$ and hence Problem/Solution 8.1.

The conditions (a) – (c) are such that $\mathcal{F}$ on $[-\infty, \infty]$, i.e., including $\pm\infty$, is an ET-system.

With the same argument as in the proof of Karlin's Positivstellensatz 8.1 we transform $\mathcal{F}$ on $[-\infty, \infty]$ into $\mathcal{G}$ on $[-1, 1]$ with the tan-function. Here Lemma 5.8 ensures that also $\mathcal{G}$ is an ET-system.

Application of Karlin's Nichtnegativstellensatz 7.6 on $[-1, 1]$ gives the desired decomposition $g = g_* + g^*$ with the observation that $x = \pm 1$ is a zero of at most multiplicity one by (a) and (b). Backwards transformation into $\mathcal{F}$ on $[-\infty, \infty]$ resp. $[-\infty, \infty)$ then gives the assertion.

Problems of Chapter 9

9.1 Proof of Theorem 9.13

Theorem 9.10 can in general not be extended to $[0, b]$ since $\{x^{\alpha_0}, \ldots, x^{\alpha_n}\}$ is not an ET-system. This fails at $x = 0$. But on $(0, b]$ it is an ET-system. We can therefore factor out the zeros of $f \geq 0$ at $x = 0$

$$f(x) = a_i x^{\alpha_i} + a_{i+1} x^{\alpha_{i+1}} + \cdots + a_n x^{\alpha_n} = x^{\alpha_i} \cdot (\underbrace{a_i + a_{i+1} x^{\alpha_{i+1} - \alpha_i} + \cdots + a_n x^{\alpha_n - \alpha_i}}_{=: \tilde{f}(x)})$$

to get some $\tilde{f}$ with $\tilde{f} \geq 0$ on $[0, b]$ and $\tilde{f}(0) > 0$. To $\tilde{f}$ we can then apply Theorem 9.10 with $a = 0$.

In summary, Theorem 9.10 on $[0, b]$ holds if $f(0) > 0$, see also Theorem 10.5 and Remark 10.6 for the corresponding version on $[0, \infty)$.

Problems of Chapter 10

10.1 Proof of Theorem 10.5
To prove Theorem 10.5 we have to note that $\mathcal{F} = \{x^{\alpha_i}\}_{i=0}^n$ with $\alpha_0 = 1$ is an ET-system on $(0, \infty)$. The only difficulty is $x = 0$ where $\mathcal{F}$ fails to be a ET-system.

But looking closely at the proof of Karlin's Theorem 7.5 (see Problem/Solution 7.1) the ET-system property is only required in a neighborhood of the zeros of f and otherwise it is the proof of Karlin's Theorem 7.1 for T-systems. Since $f(0) > 0$ we have no zero at $x = 0$ where $\mathcal{F}$ fails to be a T-system. In fact, we have $f(x) > 0$ for all $x \in [0, \varepsilon)$ for some $\varepsilon > 0$. Hence, we can apply Karlin's Nichtnegativstellensatz 8.3 since its proof requires for our f with $f(0) > 0$ only that $\mathcal{F}$ to be an ET-system on $(0, \infty)$ which is fulfilled.

10.2 By expanding

$$\prod_{i=1}^{r} (x - z_i)^{m_i} \cdot \left(a \cdot \prod_{i=1}^{m} (x - x_i)^2 + b \cdot \prod_{i=1}^{m-1} (x - y_i)^2 \right)$$

we see that $a \cdot x^{m_1 + \cdots + m_r + 2m}$ is the monomial with the highest degree $m_1 + \cdots + m_r + 2m = \deg p$ and the coefficient is a.

References

[AAM96] A. G. Akritas, E. K. Akritas, and G. I. Malaschonok, *Various proofs of Sylvester's (determinant) identity*, Math. Comput. Simulat. **42** (1996), 585–593.

[Ach56] N. I. Achieser, *Theory of Approximation*, Frederick Ungar, New York, 1956.

[AK62] N. I. Ahiezer and M. Kreĭn, *Some Questions in the Theory of Moments*, American Mathematical Society, Providence, Rhode Island, 1962.

[Ban29a] S. Banach, *Sur les fonctionnelles linéaires I*, Stud. Math. **1** (1929), 211–216.

[Ban29b] ———, *Sur les fonctionnelles linéaires II*, Stud. Math. **1** (1929), 223–239.

[BCJ79] C. Berg, J. P. R. Christensen, and C. U. Jensen, *A remark on the multidimensional moment problem*, Math. Ann. **243** (1979), 163–169.

[Ber12] S. N. Bernstein, *Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités*, Comm. Kharkiv Math. Soc. **13** (1912), 1–2.

[Ber15] ———, *Sur la représentation des polynômes positifs*, Comm. Kharkiv Math. Soc., 2nd series **14** (1915), 227–228.

[Ber26] ———, *Leçons sur les Propriétés extrémales et la Meilleure Approximation des Fonctions Analytique d'une Variable rèelle*, Gauthier–Villars, Paris, 1926.

[Ber37] ———, *Extremal properties of polynomials and the best approximation of continuous functions of one real variable*, vol. 1, ONTI, Leningrad-Moscow, 1937.

[Blo53] M. Bloom, *On the total variation of solutions of the bounded variation moment problem*, Proc. Amer. Math. Soc. **4** (1953), 118–126.

[Boa39a] R. P. Boas, *On a generalization of the Stieltjes moment problem*, Trans. Amer. Math. Soc. **46** (1939), 142–150.

[Boa39b] ———, *The Stieltjes moment problem for functions of bounded variation*, Bull. Amer. Math. Soc. **45** (1939), 399–404.

[Bog07] V. I. Bogachev, *Measure Theory*, Springer-Verlag, Berlin, 2007.

[Bor95] E. Borel, *Sur quelques points de la théorie des fonctions*, Ann. Ecole Norm. Sup. **12** (1895), 9–55.

[Bri59] L. Brickman, *A new generalization of a problem of F. Lukács*, Compos. Math. **14** (1959), 195–227.

[Bro11] L. E. J. Brouwer, *Über Abbildungen von Mannigfaltigkeiten*, Math. Ann. **71** (1911), 97–115.

[Car11] C. Carathéodory, *Über den Variabilitätsbereich der Fourier'schen Konstanten von positiven harmonischen Funktionen*, Rend. Circ. Mat. Palermo **32** (1911), 193–217.

[CdD22] R. Curto and P. J. di Dio, *Time-dependent moments from the heat equation and a transport equation*, Int. Math. Res. Notices (2022), https://doi.org/10.1093/imrn/rnac244.

[Cho69] G. Choquet, *Lectures on Analysis*, W. A. Benjamin, Inc., New York, Amsterdam, 1969, 3 volumes.

[Cur59] P. C. Curtis, Jr., *n-parameter families and best approximation*, Pacific J. Math. **9** (1959), 1013–1027.

[Dan18] P. J. Daniell, *A General Form of Integral*, Ann. Math. **19** (1918), 279–294.

[Dan20] ———, *Further Properties of the General Integral*, Ann. Math. **21** (1920), 203–220.

[dD23] P. J. di Dio, *The Early History of Moment Problems and Non-Negative Polynomials with Gaps: Sparse Moment Problems, Sparse Positivstellensätze, and Sparse Nichtnegativstellensätze from a T-System Point of View*, unpublished manuscript, arXiv:2309.03864.

[dDS22] P. J. di Dio and K. Schmüdgen, *The multidimensional truncated moment problem: The moment cone*, J. Math. Anal. Appl. **511** (2022), 126066, 38 pages.

[Dir58] P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th ed., Oxford University Press, Oxford, 1958, with revisions in 1967.

[Dur89] A. J. Duran, *The Stieltjes Moments Problem for Rapidly Decreasing Functions*, Proc. Amer. Math. Soc. **107** (1989), 731–741.

[Edw78] D. A. Edwards, *On the existence of probablity measures with given marginales*, Ann. Inst. Fourier, Grenoble **28** (1978), no. 4, 53–78.

[Ein35] A. Einstein, *The Late Emmy Noether. Professor Einstein Writes in Appreciation of a Fellow-Mathematician.*, The New York Times (May 4, 1935), 12.

[Eur13] Euripides, *Euripides II: Andromache, Hecuba, the Suppliant Women, Electra*, 3rd ed., The University of Chicago Press, Chicago, London, 2013.

[Fed69] H. Federer, *Geometric Measure Theory*, Springer-Verlag, Berlin, 1969.

[Fil64] V. A. Fil'štinkiĭ, *The power moment problem on the entire axis with a finite number of empty intervals in the spectrum*, Zap. Meh.-Mat. Fak. i Harkow. Mat. Obšč **40** (1964), 186–200, in Russian.

[FP01] C. F. Floudas and P. M. Pardalos (eds.), *Encyclopedia of Optimization*, vol. 1, Kluwer Academic Publishers, Dordrecht, 2001.

[Gau15] C. F. Gauß, *Methodus nova integralium valores per approximationem inveniendi*, Comm. Soc. Sci. Göttingen Math. **3** (1815), 29–76.

[GK02] V. Gantmacher and M. Krein, *Oscillation matrices and kernels and small vibrations of mechanical systems*, American Mathematical Society, Providence, RI, 2002, originally published in Russian, Moskow, 1950.

[GM96] M. Gasca and C. A. Micchelli (eds.), *Total Positivity and Its Applications*, Mathematics and Its Applications, no. 359, Springer Science+Business Media, Dordrecht, 1996.

[Gon00] V. L. Goncharov, *The Theory of Best Approximation of Functions*, J. Approx. Theo. **106** (2000), 2–57.

[Gru09] G. Grubb, *Distributions and Operators*, Spinger, New York, 2009.

[Haa18] A. Haar, *Die Minkowskische Geometrie und die Annäherung an stetige Funktionen*, Math. Ann. **78** (1918), 294–311.

[Hah27] H. Hahn, *Über lineare Gleichungssysteme in linearen Räumen*, J. reine angew. Math. **157** (1927), 214–229.

[Ham20] H. L. Hamburger, *Über eine Erweiterung des Stieltjesschen Momentenproblems*, Math. Ann. **81** (1920), 235–319.

[Har69] G. H. Hardy, *A Mathematician's Apology*, Cambridge University Press, London, 1969, reprint of the first edition from 1940, with a forword of C. P. Snow.

[Hau21a] F. Hausdorff, *Summationsmethoden und Momentenfolgen I*, Math. Z. **9** (1921), 74–109.

[Hau21b] ———, *Summationsmethoden und Momentenfolgen II*, Math. Z. **9** (1921), 280–299.

[Hau23] ———, *Momentprobleme für ein endliches Intervall*, Math. Z. **16** (1923), 220–248.

[Hav35] E. K. Haviland, *On the momentum problem for distribution functions in more than one dimension*, Amer. J. Math. **57** (1935), 562–572.

[Hav36] ———, *On the momentum problem for distribution functions in more than one dimension II*, Amer. J. Math. **58** (1936), 164–168.

[Hel12] E. Helly, *Über lineare Funktionaloperatoren*, Sitzungsber. Akad. Wiss. Wien **121** (1912), 265–297.

[Hil88] D. Hilbert, *Über die Darstellung definiter Formen als Summe von Formenquadraten*, Math. Ann. **32** (1888), 342–350.

[Hil32] T. H. Hildebrandt, *On the moment problem for a finite interval*, Bull. Amer. Math. Soc. **38** (1932), 269–270.

[Hoi92] L. Hoischen, *A Generalization of Theorems of Eidelheit and Carleman Concerning Approximation and Interpolation*, J. Approx. Theo. **71** (1992), 154–174.

[Hor77] R. A. Horn, *On the Moments of Complex Measures*, Math. Z. **156** (1977), 1–11.

[HS33] T. H. Hildebrandt and I. J. Schoenberg, *On Linear Functional Operations and the Moment Problem for a Finite Interval in One or Several Dimensions*, Ann. Math. **34** (1933), 317–328.

[Jac24] D. Jackson, *A general class of problems in approximation*, Am. J. Math. **XLVI** (1924), 215–234.

[Kak41] S. Kakutani, *Concrete representation of abstract (M)-spaces (A characterization of the space of continuous functions)*, Ann. Math. Series 2 **42** (1941), 994–1024.

[Kar63] S. Karlin, *Representation Theorems for Positive Functions*, J. Math. Mech. **12** (1963), 599–618.

[Kar68] ______, *Total Positivity*, vol. 1, Stanford University Press, Stanford, California, 1968.

[Kem68] J. H. B. Kemperman, *The General Moment Problem, a Geometric Approach*, Ann. Math. Stat. **39** (1968), 93–122.

[KN77] M. G. Kreĭn and A. A. Nudel'man, *The Markow Moment Problem and Extremal Problems*, American Mathematical Society, Providence, Rhode Island, 1977, translation of the Russian original from 1973.

[Kow84] M. A. Kowalski, *A Note on the General Multivariate Moment Problem*, Constr. Theo. Funct. (1984), 493–499.

[Kre51] M. G. Kreĭn, *The ideas of P. L. Čebyšev and A. A. Markov in the theory of limiting values of integrals and their further development*, Uspehi Mat. Nauk **6** (1951), no. 4 (44), 3–120, English transl.: Amer. Math. Soc. Transl. (2) **12** (1959), 1–121.

[KS53] S. Karlin and L. S. Shapley, *Geometry of moment spaces*, Mem. Amer. Math. Soc. **12** (1953).

[KS66] S. Karlin and W. J. Studden, *Tchebycheff Systems: With Applications in Analysis and Statistics*, John Wiley & Sons, Interscience Publishers, New York, NY, 1966.

[Lau09] M. Laurent, *Sums of squares, moment matrices and optimization over polynomials*, Emerging application of algebraic geometry, IMA Vol. Math. Appl., vol. 149, Springer, New York, 2009, pp. 157–270.

[Lei89] G. W. Leibniz, *Philosophical Essays*, Hackett Publishing Company, Inc., Indianapolis, Indiana, 1989, p. 1, letter to Foucher in 1675, in French.

[Lew43] K. Lewin, *Psychology and the Process of Group Living*, J. Soc. Psych. S.P.S.S.I. Bull. **17** (1943), 113–131.

[Lin99] T. Lindvall, *On Strassen's Theorem on stochastic Domination*, Elect. Comm. in Probab. **4** (1999), 51–59.

[Lor86] G. G. Lorentz, *Bernstein Polynomials*, AMS Chelsea Publishing, Providence, Rhode Island, 1986.

[Luk18] F. Lukács, *Verschärfung des ersten Mittelwertsatzes der Integralrechnung für rationale Polynome*, Math. Z. **2** (1918), 295–305.

[Mac95] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, 2nd ed., Oxford Science Publications, Oxford, 1995.

[Mai56] J. C. Mairhuber, *On Haar's theorem concering Chebyshev approximation problems having unique solutions*, Proc. Amer. Math. Soc. **7** (1956), 609–615.

[Mar84] A. A. Markov, *On certain applications of algebraic continued fractions*, Thesis (in Russian), St. Petersburg, 1884, 131 pp.

[Mar95] ______, *Deux démonstrations de la convergence de certaines fractions continues*, Acta Math. **19** (1895), 93–104.

[Mar06] ______, *Lectures on functions deviating least from zero*, Mineographed Notes, St. Petersburg, 1906, in Russian; reprinted in A. A. Markov, *Selected papers on continued fractions and the theory of functions deviating least from zero*, OGIZ, Moscow, 1948, p. 244–291.

[Mar38] A. Markov, *On mean values and exterior densities*, Rec. Math. Moscou. N.S. **4** (1938), 165–190.

[Mar08] M. Marshall, *Positive Polynomials and Sums of Squares*, Mathematical Surveys and Monographs, no. 146, American Mathematical Society, Rhode Island, 2008.

[Mün14] C. H. Müntz, *Mathematische Abhandlungen Hermann Amandus Schwarz zu seinem fünfzigjährigen Doktorjubiläum am 6. August 1914 gewidmet von Freunden und Schülern.*, ch. Über den Approximationssatz von Weierstrass, pp. 303–312, Springer, Berlin, 1914.

[NW72] S.-B. Ng and S. Warner, *Continuity of positive and multiplicative functionals*, Duke Math. J. **39** (1972), 281–284.

[Phe01] R. R. Phelps, *Lectures on Choquet's Theorem*, Springer-Verlag, Berlin, Heidelberg, 2001.

[Pie07] A. Pietsch, *History of Banach Spaces and Linear Operators*, Birkhäuser, Boston, 2007.

[Pól38] G. Pólya, *Sur l'indétermination d'un théoreme voisin du probléme des moments*, C. R. Acad. Sci. Paris **207** (1938), 708–711.

[Pól45] ———, *How To Solve It. A New Aspect of Mathematical Method*, Princeton University Press, Princeton, NJ, 1945, reprinted 2022 in Pinguin Books.

[PS64] G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, 3rd ed., vol. 2, Springer-Verlag, Berlin, 1964.

[PS70] ———, *Aufgaben und Lehrsätze aus der Analysis*, 4th ed., vol. 1, Springer-Verlag, Berlin, 1970.

[Rad80] R. Radau, *Étude sur les formules d'approximation qui servant à calculer la valeur numérique d'une intégrale définie*, J. Math. Pures Appl., 3^{me} série **6** (1880), 283–336.

[Ric57] H. Richter, *Parameterfreie Abschätzung und Realisierung von Erwartungswerten*, Bl. Deutsch. Ges. Versicherungsmath. **3** (1957), 147–161.

[Rie09] F. Riesz, *Sur les opérations fonctionnelles linéaires*, C. R. Acad. Sci. Paris **144** (1909), 1409–1411.

[Roc72] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1972.

[Rog58] W. W. Rogosinski, *Moments of non-negative mass*, Proc. R. Soc. Lond. A **245** (1958), 1–27.

[Ros52] P. C. Rosenbloom, *Quelques classes de problème extrémaux. II*, Bull. Soc. Math. France **80** (1952), 183–215.

[San05] G. Santayana, *Reason in Common Sense*, The Life of Reason, vol. 1, London Constable, London, 1905.

[Sch53] I. J. Schoenberg, *On smoothing operations and their generating functions*, Bull. Amer. Math. Soc. **59** (1953), 199–230.

[Sch78] K. Schmüdgen, *Uniform Topologies on Enveloping Algebras*, J. Funct. Anal. **39** (1978), 57–66.

[Sch90] ———, *Unbounded Operator Algebras and Representation Theory*, Birkhäuser, Basel, Boston, Berlin, 1990.

[Sch91] ———, *The K-moment problem for closed semi-algebraic sets*, Math. Ann. **289** (1991), 203–206.

[Sch14] R. Schneider, *Convex Bodies: The Brunn–Minkowski Theory*, Cambridge University Press, Cambridge, 2014.

[Sch17] K. Schmüdgen, *The Moment Problem*, Springer, New York, 2017.

[She64] T. Sherman, *A Moment Problem on* $\mathbb{R}^n$, Rend. Circ. Mat. Palermo **13** (1964), 273–278.

[Sie58] K. Sieklucki, *Topological properties of sets admitting the Tschebycheff systems*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astr. Phys. **6** (1958), 603–606.

[Ska93] J. Skala, *The Existence of probability Measures with given Marginals*, Ann. Prob. **21** (1993), 136–142.

[ST43] J. A. Shohat and J. D. Tamarkin, *The Problem of Moments*, Amer. Math. Soc., Providence, R.I., 1943.

[Sti94] T. J. Stieltjes, *Recherches sur les fractions continues*, Ann. Fac. Sci. Toulouse **8** (1894), no. 4, J1–J122.

[Sto37] M. H. Stone, *Applications of the theory of Boolean rings to general topology*, Trans. Amer. Math. Soc. **41** (1937), 375–481.

[Sto01] J. Stochel, *Solving the truncated moment problem solves the moment problem*, Glasgow J. Math. **43** (2001), 335–341.

[Str65] V. Strassen, *The existence of probability measures with given marginales*, Ann. Math. Stat. **36** (1965), 423–439.

[Šve39] K. I. Švenco, *On Hamburger's moment problem with supplementary requirement that masses are absent on a given interval*, Commun. Soc. Math. Kharkov **16** (1939), 121–128, in Russian.

[SW99] H. H. Schaefer and M. P. Wolff, *Topological Vector Spaces*, Graduate Texts in Mathematics, no. 3, Springer Science+Business Media, New York, 1999.

[Syl51] J. J. Sylvester, *On the relation between the minor determinants of linearly equivalent quadratic functions*, Phil. Mag. **1** (1851), 295–305.

[Szá16] O. Szász, *Über die Approximation stetiger Funktionen durch lineare Aggregate von Potenzen*, Math. Ann. **77** (1916), 482–496.

[Tch74] P. L. Tchebycheff, *Sur les values limites des intégrales*, J. Math. Pures Appl. (2) **19** (1874), 157–160.

[Tch57] M. V. Tchakaloff, *Formules de cubatures méchaniques a coefficients non négatifs*, Bull. Sci. Math. **81** (1957), 123–134.

[Toe11] O. Toeplitz, *Über allgemeine lineare Mittelbildung*, Prace Mat.-Fiz. **22** (1911), 113–119.

[Trè67] F. Trèves, *Topological Vector Spaces, Distributions and Kernels*, Academic Press, New York, 1967.

[Ury25] P. Urysohn, *Über die Mächtigkeit der zusammenhängenden Mengen*, Math. Ann. **94** (1925), 262–295.

[Wal39] A. Wald, *Limits of distribution function determined by absolute moment and inequalities satisfied by absolute moments*, Trans. Amer. Math. Soc. **46(2)** (1939), 280–306.

[Wei85] K. Weierstrass, *Über die analytische Darstellbarkeit sogenannter willkürlicher Functionen reeller Argumente*, Sitzungsber. Königl. Preuss. Akad. Wiss. (9. Juli 1885), 633–639.

[Wer07] D. Werner, *Funktionalanalysis*, Springer-Verlag, Berlin, 2007.

[Wie88] N. Wiener, *The Human Use of Human Beings: Cybernetics and Society*, Da Capo Press, Cambridge, Massachusetts, 1988.

[Xia59] D. Xia, *Positive functional on algebras*, Dokl. Akad. Nauk SSSR **124** (1959), 1223–1225, Izv. *Akad. Nauk, SSS4, Ser. Mat.* **23** (1959), 509–528; in Russian.

[Yos68] K. Yosida, *Functional Analysis*, 2nd ed., Springer-Verlag, Berlin, 1968.

[Zei86] E. Zeidler, *Nonlinear Functional Analysis and its Applications I: Fixed-Point Theorems*, Springer, New York, 1986.

List of Symbols

Matrices

Determinants

Further Mathematical Symbols

Index