Bachelorarbeit

An application of the Krein-Milman Theorem to Markov- and Bernstein-type trinomial inequalities

ausgeführt am

Institut für
Analysis und Scientific Computing
TU Wien

unter der Anleitung von

Ao.Univ.Prof. Dr. Harald Woracek

durch

Christoph Spiess
Matrikelnummer: 11818750

Wien, Juni 2021

Contents

1 Introduction

In this thesis we investigate a Markov- and Bernstein-type inequality for trinomials of a certain form. To be more specific, we examine the space $\mathcal{P}_{m,n}(\mathbb{R})$ of all polynomials

$$p(x) = ax^m + bx^n + c \quad \text{for } (a,b,c) \in \mathbb{R}^3$$

endowed with the supremum norm $\|.\|_\infty$. For given odd numbers $m, n \in \mathbb{N}$ with $m > n$, we obtain a function $\mathcal{M}_{m,n}(x)$, which maps every $x \in [-1,1]$ to the smallest possible constant $\mathcal{M}_{m,n}(x)$, which fulfills

$$|p'(x)| \leq \mathcal{M}_{m,n}(x) \cdot \|p\|_\infty \quad \text{for all } p \in \mathcal{P}_{m,n}(\mathbb{R}).$$

After we obtained the function $\mathcal{M}_{m,n}(x)$, it will be easy to determine the smallest possible constant $M_{m,n}$, which fulfills

$$|p'(x)| \leq M_{m,n} \cdot \|p\|_\infty \quad \text{for every } x \in [-1,1] \text{ and } p \in \mathcal{P}_{m,n}(\mathbb{R}).$$

The proof mainly follows the one presented in [6] by Gustavo A. Muñoz-Fernández, Yannis Sarantopoulos and Juan B. Seoane-Sepúlveda. Also many results obtained in [7] by Gustavo A. Muñoz-Fernández and Juan B. Seoane-Sepúlveda are used. It relies on the Krein-Milman approach (Theorem 4.2), which is a direct consequence of the classical Minkowski-Carathéodory Theorem (Theorem A.17), which itself is a consequence of the Krein-Milman Theorem (Theorem 2.13). The function $\mathcal{M}_{m,n}$ can be computed for arbitrary $m, n \in \mathbb{N}$. The other cases, that m or n is even, can be settled with the same method, only require some additional computations, and we do not treat them. The reader is referred to [6]. The proofs of the Minkowski-Carathéodory Theorem and the Krein-Milman approach are provided in Appendix A. Here we follow [8] by Barry Simon. The content of the courses taught in the bachelors program of *Technische Mathematik* at the Vienna University of Technology, especially the courses *Analysis 1, Analysis 2, Analysis 3* and *Funktionalanalysis 1*, constitute the theoretical ground this work is based upon, cf. [4, 5, M. Kaltenbäck] and [9, H. Woracek, M. Kaltenbäck, and M. Blümlinger].

2 Basic Definitions and Observations

Definition 2.1. (Topological vector space)
A topological vector space is a complex vector space endowed with a topology $\mathcal{T}$ such that the operations

$$+ : \left\{ \begin{array}{ccc} X \times X & \to & X \\ (x,y) & \mapsto & x+y \end{array} \right. \qquad \text{and} \qquad \cdot : \left\{ \begin{array}{ccc} \mathbb{C} \times X & \to & X \\ (\lambda,y) & \mapsto & \lambda y \end{array} \right.$$

are continuous, where $\mathbb{C}$ is endowed with the topology $\mathcal{E}$ induced by the euclidean metric and $X \times X$ and $\mathbb{C} \times X$ are endowed with the product topologies $\mathcal{T} \times \mathcal{T}$ and $\mathcal{E} \times \mathcal{T}$ respectively. In addition, we require X to be Hausdorff, i.e. X is (T_2).

Lemma 2.2. *Let $(X,\|.\|)$ be a normed space and $\mathcal{T}$ the topology induced by $\|.\|$. Then $(X,\mathcal{T})$ is a topological vector space.*

Proof. For a proof we refer to [9, Beispiel 2.1.2, chapter 2, page 18]. □

Lemma 2.3. *Let X be a topological vector space, $\lambda \in \mathbb{C}\backslash\{0\}$ and $a \in X$. Then the operations*

$$T_a : \left\{ \begin{array}{ccc} X & \to & X \\ x & \mapsto & x+a \end{array} \right. \qquad M_\lambda : \left\{ \begin{array}{ccc} X & \to & X \\ x & \mapsto & \lambda x \end{array} \right. \qquad S_a : \left\{ \begin{array}{ccc} \mathbb{C} & \to & X \\ \lambda & \mapsto & \lambda a \end{array} \right.$$

are continuous, and T_a and M_λ are even homeomorphisms.

Proof. For a proof we refer to [9, Lemma 2.1.3, page 18]. □

Definition 2.4. (Dual space)
For a complex vector space X, we denote by X^* the set of all linear maps of X into $\mathbb{C}$ and we call X^* the *algebraic dual space* of X. The *algebraic dual space* of a real vector space is defined analogously.
For a complex topological vector space $(X,\mathcal{T})$ we denote by $(X,\mathcal{T})'$ the set of all continuous and linear maps of X into $\mathbb{C}$ and we call $(X,\mathcal{T})'$ the *topological dual space* of $(X,\mathcal{T})$. If it is clear, which topology $\mathcal{T}$ is meant, we write X and X' instead of $(X,\mathcal{T})$ and $(X,\mathcal{T})'$, respectively.

Theorem 2.5. (Hahn-Banach separation Theorem)
Let X be a topological vector space and $A, B \subseteq X$ be nonempty, disjoint and convex subsets of X. Then

(i) If A is open, then $\exists f \in X', \gamma \in \mathbb{R}$, such that

$$\mathrm{Re}(f(x)) < \gamma \leq \mathrm{Re}(f(y)) \text{ for all } x \in A, y \in B.$$

(ii) If, in addition, X is locally convex, A is compact and B is closed, then $\exists f \in X', \gamma_1, \gamma_2 \in \mathbb{R}$, such that

$$\mathrm{Re}(f(x)) \leq \gamma_1 < \gamma_2 \leq \mathrm{Re}(f(y)) \text{ for all } x \in A, y \in B.$$

Here $\mathrm{Re}(y)$ *for* $y \in \mathbb{C}$ *denotes the real part of* y*. Note that* $\mathrm{Re}(f)$ *is* $\mathbb{R}$*-linear.*

Proof. A proof can be found in [9, Satz 5.2.5, pages 77-78]. $\square$

Definition 2.6. ($\mathcal{U}(x)$)
We denote the set of all neighbourhoods of a point x in a topological space $(X, \mathcal{T})$ by $\mathcal{U}(x)$.

Definition 2.7. (A^C, A° and ∂A)
Let $(X, \mathcal{T})$ be a topological space and $A \subseteq X$.
The *complement* of a set A is denoted by $A^C := X \backslash A$.
If $x \in A$ and A is a neighbourhood of x, then x is called an *interior point* of A. The *interior* A° is defined as the set of all interior points of A:

$$A^\circ = \{x \in A : A \in \mathcal{U}(x)\}.$$

The *boundary* of a set A is denoted by ∂A and is defined as follows:

$$\partial A = \bar{A} \backslash A^\circ = \{x \in X : \forall U \in \mathcal{U}(x) : A \cap U \neq \emptyset \wedge A^C \cap U \neq \emptyset\}.$$

Definition 2.8. (Extreme subsets)
Let C be a convex subset of a vector space X. A nonempty subset $E \subseteq C$ is called extreme subset of C, if no point $x \in E$ is an interior point of any line segment whose endpoints are in C except when both endpoints are already in E, i.e.

$$x, y \in C,\ 0 < t < 1,\ tx + (1-t)y \in E \quad \Rightarrow \quad x, y \in E.$$

An extreme set that is strictly smaller than C is called a proper extreme subset of C.

Definition 2.9. (Extreme points)
Let C be a convex subset of a vector space X. A point $e \in C$ is said to be extreme in C if $\{e\}$ is an extreme subset of C, or equivalently if $x, y \in C$ and $\lambda x + (1-\lambda)y = e, \lambda \in (0,1)$, entails $x = y = e$. We denote the set of all extreme points of C by $\mathrm{Ext}(C)$.

Lemma 2.10. *Let* X *be a vector space and* C *be a convex subset of* X*. Then*

$$e \in \mathrm{Ext}(C) \Leftrightarrow C \backslash \{e\} \textit{ is convex.}$$

Proof.
"$\Rightarrow$": Let $x, y \in C \backslash \{e\}$ and $\lambda \in [0,1]$. As C is convex, it follows that $\lambda x + (1-\lambda)y = z \in C$. As e is an extreme point, we have $z \neq e$ and therefore $z \in C \backslash \{e\}$.
"$\Leftarrow$": As $C \backslash \{e\}$ is convex, for $x, y \in C \backslash \{e\}, \lambda \in (0,1)$ we have $\lambda x + (1-\lambda)y \in C \backslash \{e\}$. If w.l.o.g. $x = e$ and $\lambda x + (1-\lambda)y = e$, clearly $y = e$. $\square$

Definition 2.11. (Convex hull)
Let X be a vector space and $A \subseteq X$. The *convex hull* of A is the smallest convex set containing A and is denoted by $\mathrm{co}(A)$.
If $A = \{a_1, \dots, a_n\}$, we also write $\mathrm{co}(a_1, \dots, a_n)$ for $\mathrm{co}(A)$.

Remark 2.12. The convex hull of a set $A \subseteq X$, where X is a linear space, is the intersection of all convex supersets of A. It is also easy to see, that

$$\mathrm{co}(A) = \Big\{ \sum_{i=1}^{n} t_i x_i : n \in \mathbb{N}, x_1, \ldots, x_n \ \in A, t_1, \ldots, t_n \ \in [0,1], \sum_{i=1}^{n} t_i = 1 \Big\}$$

Theorem 2.13. (Krein-Milman Theorem)
Let X be a locally convex topological vector space and $K \subseteq X$ a nonempty, compact and convex subset of X. Then K is the closure of the convex hull of its extreme points, i.e. $K = \overline{co(\mathrm{Ext}(K))}$. In particular K possesses extreme points.

Proof. A proof can be found in [9, Satz 5.9.2, chapter 5, page 100]. □

3 An introduction to Markov- and Bernstein-type inequalities

We want to obtain explicitly a function $\mathcal{M}_{m,n}(x) : [-1,1] \to \mathbb{R}$ for odd numbers $m, n \in \mathbb{N}$ with $m > n$, which, given a trinomial $p(x) = ax^m + bx^n + c$ maps every $x \in [-1,1]$ to the smallest possible constant $\mathcal{M}_{m,n}(x)$ which fulfills

$$|p'(x)| \leq \mathcal{M}_{m,n}(x) \cdot \|p\|,$$

where $\|p\| := \max_{x\in[1,1]} |p(x)|$ is the supremum norm of p over $[-1,1]$. This and similar problems have been studied for a long time. In the following, we present some background information.
In [6], some historical facts concerning the function $\mathcal{M}_{m,n}$ are presented, whereas N.K. Govil and R.N. Mohapatra give a more detailed description of the history of Markov- and Bernstein-type inequalities in [3].

Definition 3.1. ($\|p\|_{[\alpha,\beta]}$ and $M_n(\alpha, \beta)$)
Let $n \in \mathbb{N}, \alpha < \beta \in \mathbb{R}$ and p be a polynomial function on $[\alpha, \beta]$, then

$$\|p\|_{[\alpha,\beta]} := \max_{x\in[\alpha,\beta]} |p(x)|.$$

We define $M_n(\alpha, \beta)$ to be the smallest possible constant, such that

$$|p'(x)| \leq M_n(\alpha, \beta)\|p\|_{[\alpha,\beta]} \text{ for every } x \in [\alpha, \beta], \deg(p) \leq n.$$

If $\alpha = -1$ and $\beta = 1$, then we write $\|p\|$ and M_n instead of $\|p\|_{[-1,1]}$ and $M_n(-1, 1)$, respectively.

In a more general setting the problem mentioned above has been studied since the end of the 19^{th} century. D. Mendeleev, some years after he invented the Periodic Table of the Elements, made a study of the specific gravity of a solution as a function of the percentage of the dissolved substance. This function is still of importance, as it is nowadays used for testing beer and wine for alcoholic content or testing the cooling system of an automobile for concentration of anti-freeze. This study led to mathematical problems of great interest. Mendeleev noticed that the curves, which present the function mapping the percentage of alcohol in water to the corresponding weight of the fluid could be closely approximated by a series of quadratic arcs. He wanted to find out, if his measurements were correct and if certain corners where the arcs joined were really there or just caused by errors. In particular, he was interested in the following problem:

Remark 3.2. (Question raised by D. Mendeleev)
If $p(x) = ax^2 + bx + c$ with $a, b, c \in \mathbb{R}, \alpha, \beta \in \mathbb{R}$ with $\alpha < \beta$, then what is the smallest possible constant $M_2(\alpha, \beta) > 0$ such that

$$|p'(x)| \leq M_2(\alpha, \beta) \|p\|_{[\alpha,\beta]}$$

for every $x \in [\alpha, \beta]$ and every polynomial p of the form defined above?

Remark 3.3. In the problem mentioned in Remark 3.2 we have

$$M_2(\alpha, \beta) = \frac{2}{\beta - \alpha} M_2$$

Proof. Let $p(x) = ax^2 + bx + c$ and let $x = \frac{1}{2}(\alpha + \beta - (\alpha - \beta)y)$. Then

$$q(y) := a \cdot \big(\frac{1}{2}(\alpha + \beta - (\alpha - \beta)y)\big)^2 + b \cdot \big(\frac{1}{2}(\alpha + \beta - (\alpha - \beta)y\big) + c = p(x)$$

is a polynomial on $[-1, 1]$ of the form $q(y) = \tilde{a}y^2 + \tilde{b}y + \tilde{c}$ and

$$\tilde{a} = a \cdot \frac{(\beta - \alpha)^2}{4}, \quad \tilde{b} = a \cdot \frac{\beta^2 - \alpha^2}{2} + b \cdot \frac{\beta - \alpha}{2}, \quad \tilde{c} = a \cdot \frac{(\alpha + \beta)^2}{4} + b \cdot \frac{\alpha + \beta}{2} + c$$

$$\begin{aligned} q'(y) &= a \cdot (\alpha + \beta - (\alpha - \beta)y) \cdot \frac{\beta - \alpha}{2} + b \cdot \frac{\beta - \alpha}{2} \\ &= (2ax + b) \cdot \frac{\beta - \alpha}{2} = p'(x) \cdot \frac{\beta - \alpha}{2} \end{aligned}$$

It is easy to see that the variable transformation gives us a bijective mapping of the space of quadratic polynomials on $[\alpha, \beta]$ onto the corresponding space of polynomials on $[-1, 1]$. Since $\|q\|_{[-1,1]} = \|p\|_{[\alpha,\beta]}$, we have

$$|q'(y)| \leq M_2 \|q\|_{[-1,1]} \Leftrightarrow |p'(x)| \leq \frac{2}{\beta - \alpha} M_2 \|p\|_{[\alpha,\beta]}$$

and thus we conclude $M_2(\alpha, \beta) = \frac{2}{\beta - \alpha} M_2$. □

Astonishingly, Mendeleev himself was able to solve this mathematical problem. He obtained that $M_2 = 4$. That means

$$|p'(x)| \leq 4 \cdot \|p\| \text{ for all } x \in [-1, 1],$$

whereas equality is obtained for e.g. $p(x) := 1 - 2x^2$, since $\|p\| = 1$ and $|p'(\pm 1)| = 4$. Using this result, Mendeleev was able to show that the corners in his curve were genuine and he was right, since his measurements were quite accurate and even agree with modern tables to three or more significant figures. The famous chemist then told the Russian mathematician A.A. Markov about this result, who investigated the corresponding problem in a more general setting. He was able to generalize Mendeleev's result:

Theorem 3.4. (A.A. Markov, 1889)
If p is a real polynomial with $\deg(p) \leq n \in \mathbb{N}$, then

$$|p'(x)| \leq n^2 \|p\| \text{ for } x \in [-1, 1].$$

Equality is attained at the end points of $[-1,1]$ for the n-th Chebyshev polynomial of the first kind defined by $T_n(x) = \cos(n \arccos(x))$ for $x \in [-1,1]$, in other words, $M_n = n^2$.

Proof. As Markov's original paper is written in old Russian and not readily accessible, we refer to [2, pages 169-170] for a modern proof. □

A generalization of Markov's theorem was obtained by his brother V.A. Markov. Clearly, using Theorem 3.4, we have if $|p(x)| \leq M$ on $[-1,1]$, then $|p'(x)| \leq M \cdot n^2$ on $[-1,1]$. Having found an upper bound for $|p'(x)|$, it is natural to ask for an upper bound for $|p^{(k)}(x)|$ with $k \leq n$ (if $k > n$, then $p^{(k)}$ vanishes). Iterating Markov's theorem yields that if

$$|p(x)| \leq M \text{ on } [-1,1], \text{ then } |p^{(k)}(x)| \leq n^{2k}.$$

This inequality, however, is not sharp. The results of the attempt of V.A. Markov to find exact bounds for $|p^{(k)}|$ on $[-1,1]$ included the following:

Theorem 3.5. *If p is a real polynomial with $\deg(p) \leq n$ and $\|p\| \leq 1$, then*

$$|p^{(k)}(x)| \leq \frac{n^2(n^2 - 1^2) \cdots (n^2 - (k-1)^2)}{1 \cdot 3 \cdot 5 \cdots (2k-1)} \text{ for } k \leq n \text{ and } x \in [-1,1].$$

S. Bernstein later needed the analogue of Theorem 3.4 for the unit disk in the complex plane instead of the interval $[-1,1]$. He wanted to find an upper bound for $|p'(z)|$ for a complex polynomial p (that means $p(z) = \sum_{i=0}^{n} a_i z^i$ with $a_i \in \mathbb{C}$ for all $i \in \{0, \ldots, n\}$ and $z \in \mathbb{C}$) with $|p(z)| \leq 1$ for $|z| \leq 1$. The answer is presented in the following theorem known as Bernstein's inequality.

Theorem 3.6. *If p is a complex polynomial with $\deg(p) \leq n$, then*

$$\max_{|z| \leq 1} |p'(z)| \leq n \cdot \max_{|z| \leq 1} |p(z)|.$$

The result is best possible and the equality holds for $p(z) = \lambda z^n$ for $\lambda \in \mathbb{C} \backslash \{0\}$.

However, this theorem is not of great interest to us, since we are working only with real numbers. It is important to notice that in the above definition of $M_n(\alpha, \beta)$ the equality has to hold for all $x \in [\alpha, \beta]$. Markov's result $M_n = n^2$ can be substantially improved by fixing a single point $x \in [-1, 1]$.

Definition 3.7. ($\mathcal{M}_n$)
Let $n \in \mathbb{N}$, p be polynomial function on $[-1, 1]$ and $x \in [-1, 1]$, then we define $\mathcal{M}_n(x)$ to be the smallest possible constant such that

$$|p'(x)| \leq \mathcal{M}_n(x)\|p\| \text{ for every } p \text{ with } \deg(p) \leq n.$$

S. Bernstein also managed to find an estimate on $\mathcal{M}_n(x)$ for $x \in (-1, 1)$.

Theorem 3.8. (S. Bernstein, 1912)
If p is a polynomial with $\deg(p) \leq n \in \mathbb{N}$, then $|p'(x)| \leq \frac{n}{\sqrt{1-x^2}} \cdot \|p\|$ for every $x \in (-1, 1)$. In other words, $\mathcal{M}_n(x) \leq \frac{n}{\sqrt{1-x^2}}$.

Proof. A proof can be found in [1, pages 6-11]. □

Bernstein's estimate, $\mathcal{M}_n(x) \leq \frac{n}{\sqrt{1-x^2}}$, coincides with $\mathcal{M}_n(x)$ at n points in $[-1, 1]$, but it is far from being optimal in most of the interval $[-1, 1]$. However, Bernstein's estimate gained importance as it was used in modern proofs of Markov's theorem in order to simplify the proof.

According to [6], many attempts to find $\mathcal{M}_n(x)$ for every $n \in \mathbb{N}$ have been performed. The study of E. V. Voronovskaja, who focused on polynomials p with $\|p\| = 1$ and $|p'(x)| = \mathcal{M}_n(x)$ and the properties of the function $\mathcal{M}_n(x)$, was the most successful. She even produced a method to obtain $\mathcal{M}_n(x)$ for each $n \in \mathbb{N}$. The authors of [6] state that the construction of these polynomials is not explicit in the sense that, for higher values of n, it involves solving elliptic integrals and requires numerical calculus.

4 The space $\mathcal{P}_{m,n}(\mathbb{R})$

Our aim is to give an explicit formula for the norm of point evaluation of derivatives of certain subspaces of polynomials.

Definition 4.1. ($\mathcal{P}_{m,n}(\mathbb{R})$ and $\mathcal{M}_{m,n}$)
Denote

$$\mathcal{P}_{m,n}(\mathbb{R}) := \left\{ p : \left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & ax^m + bx^n + c \end{array} \right. : a, b, c \in \mathbb{R} \right\}$$

For $x \in [-1, 1]$, we define $\mathcal{M}_{m,n}(x)$ to be the smallest possible constant, such that

$$|p'(x)| \leq \mathcal{M}_{m,n}(x)\|p\| \text{ for all } p \in \mathcal{P}_{m,n}(\mathbb{R})$$

where $\|p\| := \max_{x \in [1,1]} |p(x)|$ is the supremum norm of p over $[-1, 1]$.

We use the Krein-Milman approach to compute the functions $\mathcal{M}_{m,n}(x)$. This technique relies on the following consequence of the Minkowski-Carathéodory Theorem (Theorem A.17) which itself is a consequence of the Krein-Milman Theorem (Theorem 2.13).

Theorem 4.2. (The Krein-Milman approach)
If C is a convex body, i.e. a compact, convex, and nonempty subset, in a finite-dimensional topological vector space and $f : C \to \mathbb{R}$ is a convex function that attains its maximum on C, then there is an extreme point $e \in C$ such that $f(e) = \max\{f(p) : p \in C\}$.

Proof. Although the structure of the proof of this theorem is interesting, it is not the main aspect of this paper and requires some theory. Therefore it is deferred to the appendix, see Theorem A.19. □

Let us briefly explain that we indeed are in the situation of this theorem. The space $\mathcal{P}_{m,n}(\mathbb{R})$ endowed with $\|p\| = \max\limits_{x\in[1,1]} |p(x)|$ is a finite dimensional normed space, and its unit ball

$$B_{m,n} := \{p \in \mathcal{P}_{m,n}(\mathbb{R}) : \|p\|_{m,n} \leq 1\}$$

is a nonempty, convex, and (by finite dimensionality) compact subset. For each fixed $x_0 \in [-1, 1]$ the function $p \mapsto p'(x_0)$ is linear and hence (again by finite dimensionality) continuous. By linearity,

$$\mathcal{M}_{m,n}(x) = \max_{p\in B_{m,n}} |p'(x_0)| = \max_{p\in S_{m,n}} |p'(x_0)|,$$

where $S_{m,n}$ is the unit sphere. The function $p \mapsto |p'(x_0)|$ is convex and continuous. In particular, it attains a maximum on $B_{m,n}$.
To study $\mathcal{P}_{m,n}(\mathbb{R})$, it is useful to identify it with $\mathbb{R}^3$. From now on we use the notation $ax^m + bx^n + c$ for the function $p : \left\{\begin{array}{lcl} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & ax^m + bx^n + c \end{array}\right.$ as well as for the value $p(x)$. The meaning will be clear by context.

Remark 4.3. The mapping

$$\psi : \left\{\begin{array}{lcl} \mathcal{P}_{m,n}(\mathbb{R}) & \to & \mathbb{R}^3 \\ ax^m + bx^n + c & \mapsto & (a, b, c) \end{array}\right.$$

that assigns to each polynomial of the form $p(x) = ax^m + bx^n + c$ its coefficients (a, b, c) in the basis $\{x^m, x^n, 1\}$ of $\mathcal{P}_{m,n}(\mathbb{R})$ is a linear isomorphism.

According to Remark 4.3, we can identify $\mathcal{P}_{m,n}(\mathbb{R})$ with $\mathbb{R}^3$. We define a corresponding norm

$$\|(a, b, c)\|_{m,n} := \max_{x\in[-1,1]} |ax^m + bx^n + c|.$$

5 The extreme points of $B_{m,n}$

The crucial point is to determine $\mathrm{Ext}(B_{m,n})$.

Lemma 5.1. *Let $m, n \in \mathbb{N}$ with $m > n$.*

(i) *The equation*
$$|n + mx| = (m - n)|x|^{\frac{m}{m-n}}$$
has three solutions, one at $x = -1$, another one at some point $\lambda_0 \in (-\frac{n}{m}, 0)$ and a third at some point $\lambda_1 > 0$.

(ii) *We have*
$$|n + mx| \leq (m - n)|x|^{\frac{m}{m-n}} \tag{1}$$
if and only if $x \leq \lambda_0$ or $x \geq \lambda_1$.

Proof. Let $f(x) := |n + mx|$ and $g(x) := (m - n)|x|^{\frac{m}{m-n}}$ for $x \in \mathbb{R}$. Then the graph of f clearly consists of the two lines $L_1 := \{(x, -n - mx : x \leq -\frac{n}{m}\}$ and $L_2 := \{(x, n + mx) : x \geq -\frac{n}{m}\}$. Since $m - n > 0$ and $x \mapsto |x|^\alpha$ is a strictly convex function for $\alpha > 1$, we conclude that also g is a strictly convex function. Since every line can intersect the graph of a strictly convex function in at most two points, the graphs of f and g intersect in at most four points. Since $f(-1) = g(-1) = m - n$ and $f'(-1) = g'(-1) = -m$, the graph of g is tangent to the graph of f at the point $(-1, m - n)$. Furthermore, $f(-\frac{n}{m}) = 0 < g(-\frac{n}{m})$ and $g(0) = 0 < n = f(0)$. Since f and g are continuous, there exists a point $\lambda_0 \in (-\frac{n}{m}, 0)$ with $f(\lambda_0) = g(\lambda_0)$. Since $\lim_{x\to\infty} g(x) - f(x) = \infty$ and $f(0) > g(0)$, we conclude that there must be some $\lambda_1 > 0$ with $f(\lambda_1) = g(\lambda_1)$. Inequality (1) follows by strict convexity of g. $\square$

The results of Lemma 5.1 are visualised in the following plot.

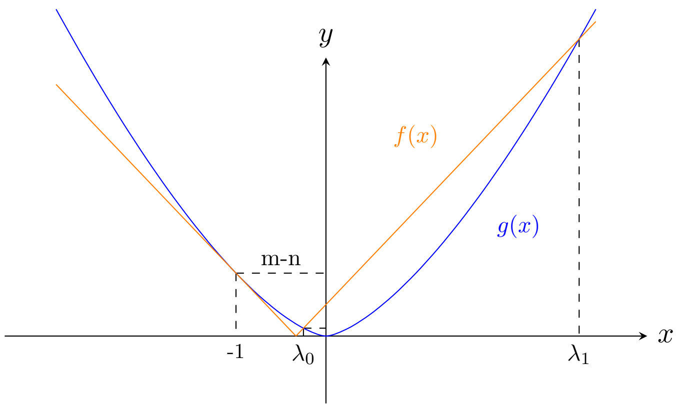

Figure 1: Solutions of $|n+mx| = (m-n)|x|^{\frac{m}{m-n}}$ with $m, n \in \mathbb{N}$ and $m > n$. Here $f(x) = |n+mx|$ and $g(x) = (m - n)|x|^{\frac{m}{m-n}}$.

What is most interesting about the results obtained above is the value of λ_0. We will use this value to further analyse the norm $\|.\|_{m,n}$.

Lemma 5.2. *Let $m, n \in \mathbb{N}$ be odd numbers with $m > n$. Then*

$$\|(a,b,c)\|_{m,n} = \|(a,b,0)\|_{m,n} + |c| \tag{2}$$

for all $(a,b,c) \in \mathbb{R}^3$. In particular, the norm $\|.\|_{m,n}$ is symmetric with respect to the ab-plane.

Proof. For $c = 0$ the statement is trivial. Let us therefore assume $c \neq 0$. Since $|ax^m + bx^n + c| \leq |ax^m + bx^n| + |c|$ for all $x \in [-1,1]$ and $(a,b,c) \in \mathbb{R}^3$, the inequality "$\leq$" is trivial. If the maximum of $|ax^m + bx^n| + |c|$ is attained at $x_0 \in [-1,1]$, then we can assume $\mathrm{sgn}(ax_0^m + bx_0^n) = \mathrm{sgn}(c)$, because $|ax_0^m + bx_0^n| = |\mathrm{sgn}(x_0)(a|x_0|^m + b|x_0|^n)| = |(a|x_0|^m + b|x_0|^n)|$, so both x_0 and $-x_0$ are maximizers and if we choose $\tilde{x_0} \in \{\pm x_0\}$ such that $\mathrm{sgn}(a\tilde{x_0}^m + b\tilde{x_0}^n) = \mathrm{sgn}(c)$, we get $|ax_0^m + bx_0^n| + |c| = |a\tilde{x_0}^m + b\tilde{x_0}^n + c|$. This shows "$\geq$". $\square$

Remark 5.3. If one of the exponents $m, n \in \mathbb{N}$ is even, equation (2) does not hold. However there is still a symmetry with respect to a coordinate plane, if m and n have different parity, namely, the ac-plane for m even and the bc-plane for m odd.

Theorem 5.4. *Let $m, n \in \mathbb{N}$ be odd numbers with $m > n$. Then*

$$\|(a,b,c)\|_{m,n} = \begin{cases} \frac{(m-n)|a|}{n} \cdot \left|\frac{nb}{ma}\right|^{\frac{m}{m-n}} + |c| & \text{if } a \neq 0 \text{ and } -1 \leq \frac{nb}{ma} \leq \lambda_0 \\ |a+b| + |c| & \text{otherwise,} \end{cases}$$

where λ_0 is the number in $(-\frac{n}{m}, 0)$ given by Lemma 5.1,(i).

Proof. Given $(a,b,c) \in \mathbb{R}^3$, define $P(x) := ax^m + bx^n$. By Lemma 5.2, it suffices to find a formula for $\|(a,b,0)\|_{m,n} = \max_{x \in [-1,1]} |P(x)|$. Since $|P|$ is symmetric with respect to the origin, it follows $\max_{x\in[-1,1]} |P(x)| = \max_{x\in[-1,0]} |P(x)|$. This maximum is attained at either the end points of the interval or the critical points of P lying in $(-1,0)$. Since $|P(x)| \geq 0 = |P(0)|$, the contribution of $|P(0)|$ to $\max_{x\in[-1,0]} |P(x)|$ is irrelevant. To find the critical points of P, we look at the equation

$$P'(x) = m \cdot ax^{m-1} + n \cdot bx^{n-1} = x^{n-1}(m \cdot ax^{m-n} + n \cdot b) = 0.$$

Since we look for critical points in $[-1,0)$, there is only one candidate $x_{m,n} := -\left|\frac{nb}{ma}\right|^{\frac{1}{m-n}}$ whenever $a \neq 0$ and $-1 \leq \frac{nb}{ma} < 0$. Thus (note that $P(1) = P(-1)$),

$$\|(a,b,0)\|_{m,n} = \begin{cases} \max\{|P(1)|, |P(x_{m,n})|\} & \text{if } a \neq 0 \text{ and } -1 \leq \frac{nb}{ma} < 0, \\ |P(1)| & \text{otherwise.} \end{cases}$$

Now using $|P(1)| = |a+b|$ and (note that $\operatorname{sgn}(b) = -\operatorname{sgn}(a)$)

$$|P(x_{m,n})| = \left|a\cdot\left|\frac{nb}{ma}\right|^{\frac{m}{m-n}} + b\cdot\left|\frac{nb}{ma}\right|^{\frac{n}{m-n}}\right| = \left|a+b\cdot\left|\frac{ma}{nb}\right|\right|\cdot\left|\frac{nb}{ma}\right|^{\frac{m}{m-n}} = \frac{(m-n)|a|}{n}\cdot\left|\frac{nb}{ma}\right|^{\frac{m}{m-n}},$$

we have

$$\|(a,b,0)\|_{m,n} = \begin{cases} \max\{|a+b|, \frac{(m-n)|a|}{n}\cdot\left|\frac{nb}{ma}\right|^{\frac{m}{m-n}}\} & \text{if } a \neq 0 \text{ and } -1 \leq \frac{nb}{ma} < 0, \\ |a+b| & \text{otherwise.} \end{cases}$$

If we multiply the inequality $|a+b| \leq \frac{(m-n)|a|}{n}\cdot\left|\frac{nb}{ma}\right|^{\frac{m}{m-n}}$ by $\frac{n}{|a|} > 0$ (for $a \neq 0$), we obtain, equivalently,

$$\left|n + \frac{nb}{a}\right| = \left|n + m\cdot\frac{nb}{ma}\right| \leq (m-n)\cdot\left|\frac{nb}{ma}\right|^{\frac{m}{m-n}}\},$$

which is the same as

$$|n + m\cdot x| \leq (m-n)|x|^{\frac{m}{m-n}},$$

where $x = \frac{nb}{ma}$. Since we look at the case $-1 \leq \frac{nb}{ma} < 0$, according to Lemma 5.1,(ii), it follows that $-1 \leq \frac{nb}{ma} \leq \lambda_0$ and hence

$$\|(a,b,0)\|_{m,n} = \begin{cases} \left|\frac{(m-n)|a|}{n}\right|\cdot\left|\frac{nb}{ma}\right|^{\frac{m}{m-n}} & \text{if } a \neq 0 \text{ and } -1 \leq \frac{nb}{ma} \leq \lambda_0, \\ |a+b| & \text{otherwise.} \end{cases}$$

□

Definition 5.5. (Projections $\pi_{ab}, \pi_{bc}, \pi_{ac}$)
We denote the projection onto the ab-plane by

$$\pi_{ab}(\tilde{a},\tilde{b},\tilde{c}) := (\tilde{a},\tilde{b}) \text{ for } (\tilde{a},\tilde{b},\tilde{c}) \in \mathbb{R}^3.$$

The projections onto the bc-plane and the ac-plane are defined correspondingly.

Our next aim is to sketch $S_{m,n}$ and to obtain the extreme points of $B_{m,n}$ for odd numbers $m,n \in \mathbb{N}$ with $m > n$. In order to do so, we have to find a parameterization of $S_{m,n}$. We will use the symmetry of $\|.\|_{m,n}$ to show that the projection of $B_{m,n}$ onto the ab-plane coincides with $\{(a,b) \in \mathbb{R}^2 : \|(a,b,0)\|_{m,n} \leq 1\}$. To make things easier and to avoid some long formulae, we need some notation. We define

$$\begin{array}{lcllcll} \Gamma(a) & := & \frac{m}{(m-n)^{\frac{m-n}{m}} n^{\frac{n}{m}}}|a|^{\frac{n}{m}} & & & & \\ V & := & \{(a,b) \in \mathbb{R}^2 : a \neq 0, -1 \leq \frac{nb}{ma} \leq \lambda_0 & \text{and} & |b| & \leq & \Gamma(a) \quad \} \\ W_1 & := & \{(a,b) \in \mathbb{R}^2 : b \geq -\frac{m}{n}\cdot a, b \geq \lambda_0\cdot\frac{ma}{n} & \text{and} & b & \leq & 1-a \quad \} \\ W_2 & := & \{(a,b) \in \mathbb{R}^2 : b \leq -\frac{m}{n}\cdot a, b \leq \lambda_0\cdot\frac{ma}{n} & \text{and} & b & \geq & -1-a \quad \} \\ W & := & W_1 \cup W_2 & & & & \end{array} \qquad (3)$$

Now let us show, that the projection $\pi_{ab}(B_{m,n})$ of $B_{m,n}$ onto the $ab-$plane is nothing but the union of V and W.

Theorem 5.6. *For odd numers $m, n \in \mathbb{N}$ with $m > n$ we have*

$$\pi_{ab}(B_{m,n}) = V \cup W$$

for V, W given by (3).

Proof. According to Lemma 5.2, we have

$$\|(a,b,c)\|_{m,n} = \|(a,b,-c)\|_{m,n} = \|(a,b,0)\|_{m,n} + |c|$$

for every $(a,b,c) \in \mathbb{R}^3$. Hence if $\|(a,b,c)\|_{m,n} \leq 1$ for any $c \neq 0$, then also $\|(a,b,0)\|_{m,n} \leq 1$. This implies that the projection $\pi_{ab}(B_{m,n})$ is the intersection of $B_{m,n}$ with the ab-plane or, in mathematical notation

$$\pi_{ab}(B_{m,n}) = \{(a,b) \in \mathbb{R}^2 : \|(a,b,0)\|_{m,n} \leq 1\}.$$

As $\pi_{ab}(B_{m,n})$ is bounded by the curve defined implicitly by $\|(a,b,0)\|_{m,n} = 1$, we take a look at the values of b at this curve in the different cases given by Lemma 5.4.
If $a \neq 0$ and $-1 \leq \frac{nb}{ma} \leq \lambda_0$, we get $\frac{(m-n)|a|}{n} \cdot \left|\frac{nb}{ma}\right|^{\frac{m}{m-n}} = 1$, and therefore

$$b = \pm\Big(\frac{n}{(m-n)|a|}\Big)^{\frac{m-n}{m}} \cdot \frac{m|a|}{n} = \pm\frac{m}{(m-n)^{\frac{m-n}{m}} n^{\frac{n}{m}}} |a|^{\frac{n}{m}} = \pm\Gamma(a).$$

However, the condition $-1 \leq \frac{nb}{ma} \leq \lambda_0$ implies that $\operatorname{sgn}(a) = -\operatorname{sgn}(b)$. Hence for $a < 0$ we have $b = \Gamma(a)$ and for $a > 0$ we have $b = -\Gamma(a)$, so in either case

$$\frac{nb}{ma} = \frac{-n \cdot \frac{m}{(m-n)^{\frac{m-n}{m}} n^{\frac{n}{m}}} |a|^{\frac{n}{m}}}{m|a|} = \frac{-n|a|^{\frac{n-m}{m}}}{(m-n)^{\frac{m-n}{m}} n^{\frac{n}{m}}} = -\Big(\frac{n}{|a|(m-n)}\Big)^{\frac{m-n}{m}}.$$

Thus,

$$-1 \leq \frac{nb}{ma} \quad \Leftrightarrow \quad 1 \geq \Big(\frac{n}{|a|(m-n)}\Big)^{\frac{m-n}{m}} \quad \Leftrightarrow \quad |a| \geq \frac{n}{m-n}$$

and

$$\begin{aligned} \lambda_0 \geq \tfrac{nb}{ma} \quad \Leftrightarrow |\lambda_0| \quad &\leq \Big(\tfrac{n}{|a|(m-n)}\Big)^{\frac{m-n}{m}} \\ \Leftrightarrow |a| \quad &\geq \tfrac{n}{(m-n)|\lambda_0|^{\frac{m}{m-n}}} \overset{5.1}{=} \tfrac{n}{|n+m\lambda_0|}. \end{aligned}$$

We conclude that

$$\begin{aligned} b &= \Gamma(a), \quad \text{if } -\tfrac{n}{n+m\lambda_0} \leq a \leq -\tfrac{n}{m-n} \\ b &= -\Gamma(a), \quad \text{if } \tfrac{n}{m-n} \leq a \leq \tfrac{n}{n+m\lambda_0}. \end{aligned} \tag{4}$$

Setting $-\frac{ma}{n} = \pm\Gamma(a)$ and $\lambda_0 \cdot \frac{ma}{n} = \pm\Gamma(a)$, we see that the bounds for a in (4) are exactly the intersecting points of the lines $b = -\frac{ma}{n}$ and $b = \lambda_0\frac{ma}{n}$ with the curves $b = \pm\Gamma(a)$.
Now if $\|(a,b,0)\|_{m,n} = 1$ and at least one of the conditions $a \neq 0$ and $-1 \leq \frac{nb}{ma} \leq \lambda_0$

does not hold, according to Lemma 5.4 we have $|a+b|=1$ and hence $b=\pm 1-a$. For $a=0$, we get $b=\pm 1$. So let us assume $a\neq 0$. Now if $b=-1-a$, it follows

$$\begin{aligned} -1 \leq \tfrac{nb}{ma} \leq \lambda_0 \quad &\Leftrightarrow \quad -\tfrac{m-n}{n} \leq -\tfrac{1}{a} \leq \tfrac{n+\lambda_0 m}{n} \\ &\Leftrightarrow \quad -\tfrac{n}{n+\lambda_0 m} \leq a \leq \tfrac{n}{m-n} \end{aligned}$$

By repeating the same steps with $b=1-a$, we get

$$\begin{aligned} b &= -1-a, \quad \text{if } -\tfrac{n}{n+m\lambda_0} \leq a \leq \tfrac{n}{m-n} \\ b &= 1-a, \quad \text{if } -\tfrac{n}{m-n} \leq a \leq \tfrac{n}{n+m\lambda_0}. \end{aligned} \tag{5}$$

Analogously to the upper case, we observe that the bounds for a in (5) are obtained by intersecting the lines $b=-\frac{ma}{n}$ and $b=\lambda_0\frac{ma}{n}$ with the lines $b=-1-a$ and $b=1-a$. Thus, we found the boundary of $\pi_{ab}(B_{m,n})$, which determines $\pi_{ab}(B_{m,n})$. $\square$

Figure 2 shows V and W. We already know, that $V\cup W = V\cup W_1 \cup W_2 = \pi_{ab}(B_{m,n})$.

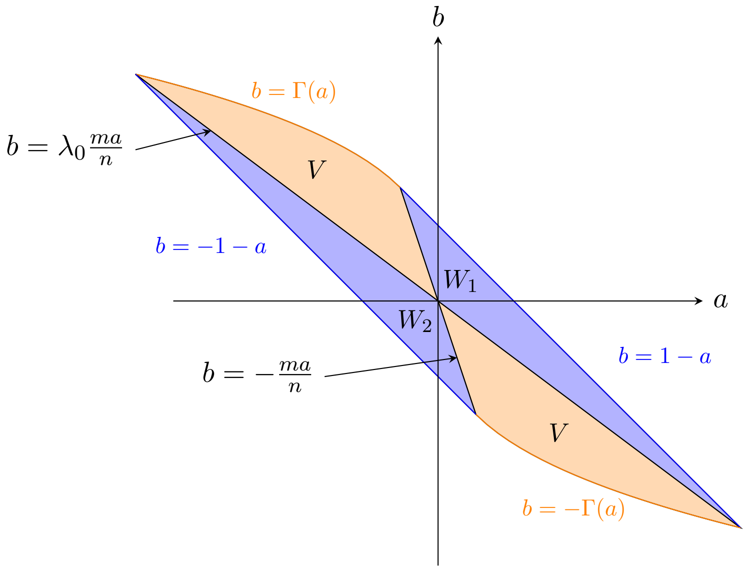

Figure 2: The projection $\pi_{ab}(B_{3,1})$ of $B_{3,1}$ onto the $ab-$plane. A similar shape of $\pi_{ab}(B_{m,n})$ for $m,n\in\mathbb{N}$ with $m>n$ is obtained in the general case. Here $\Gamma(a) := \frac{m}{(m-n)^{\frac{m-n}{m}}\, n^{\frac{n}{m}}}|a|^{\frac{n}{m}}$.

Next, we proof a result for general real normed spaces. This allows us to give a parameterization of $S_{m,n}$ and a characterization of the extreme points of $B_{m,n}$ for odd numbers $m,n\in\mathbb{N}$ with $m>n$.

Definition 5.7. (B_X and S_X) If X is a normed space, we denote the unit ball and the unit sphere of X by B_X and S_X, respectively.

$$B_X := \{x \in X : \|x\| \leq 1\}$$

$$S_X := \{x \in X : \|x\| = 1\}$$

Lemma 5.8. *Let X be a real normed vector space with norm $\|.\|_X$ and define $\tilde{X} := X \oplus \mathbb{R}$ as the space of pairs $(x, \lambda) \in X \times \mathbb{R}$ endowed with the norm given by $\|(x, \lambda)\|_{\tilde{X}} = \|x\|_X + |\lambda|$. Furthermore, define $f_+(x) := 1 - \|x\|_X$ for $x \in X$ and $f_- := -f_+$. Then*

(*i*) $S_{\tilde{X}} = \mathrm{graph}(f_+|_{B_X}) \cup \mathrm{graph}(f_-|_{B_X})$.

(*ii*) $\mathrm{Ext}(B_{\tilde{X}}) = \{(x, 0) : x \in \mathrm{Ext}(B_X)\} \cup \{\pm(0, 1)\}$.

Proof. (i): We have

$$(x, \lambda) \in S_{\tilde{X}} \Leftrightarrow \|(x, \lambda)\|_{\tilde{X}} = 1 \Leftrightarrow \|x\|_X + |\lambda| = 1 \Leftrightarrow |\lambda| = 1 - \|x\|_X$$

which means

$$(x, \lambda) \in S_{\tilde{X}} \Leftrightarrow \lambda = \pm f_+(x) \Leftrightarrow (x, \lambda) \in \mathrm{graph}(f_+|_{B_X}) \cup \mathrm{graph}(f_-|_{B_X})$$

(ii): Clearly $\mathrm{Ext}(B_{\tilde{X}}) \subseteq S_{\tilde{X}}$. Furthermore, we have $\|x\|_X, |\lambda| \leq 1$ for all $(x, \lambda) \in S_{\tilde{X}}$. Now if $\|x\| \in (0, 1)$ (note that $|\lambda| = 1 - \|x\|_X$), we have

$$(x, \lambda) = \|x\|_X \cdot (\frac{x}{\|x\|_X}, 0) + (1 - \|x\|_X) \cdot (0, \mathrm{sgn}(\lambda)).$$

Clearly $(\frac{x}{\|x\|_X}, 0), (0, \mathrm{sgn}(\lambda)) \in B_{\tilde{X}}$, so $(x, \lambda) \notin \mathrm{Ext}(B_{\tilde{X}})$, as it is a nontrivial convex combination of two elements in $B_{\tilde{X}}$. This yields

$$\mathrm{Ext}(B_{\tilde{X}}) \subseteq \{(x, 0) : x \in S_X\} \cup \{\pm(0, 1)\}.$$

If $x \in S_X, x \notin \mathrm{Ext}(B_X)$, then we find $t \in (0, 1)$ and $y, z \in B_X$ with

$$x = t \cdot y + (1 - t) \cdot z \Rightarrow (x, 0) = t \cdot (y, 0) + (1 - t) \cdot (z, 0).$$

Since $\|y\|, \|z\| \leq 1$, we have $(y, 0), (z, 0) \in B_{\tilde{X}}$ and thus $x \notin \mathrm{Ext}(B_{\tilde{X}})$.
This means

$$\mathrm{Ext}(B_{\tilde{X}}) \subseteq \{(x, 0) : x \in \mathrm{Ext}(B_X)\} \cup \{\pm(0, 1)\}.$$

To prove the other inclusion, we first take a look at the points $(0, \pm 1)$. If $(0, p_2) =: p \in \{\pm(0, 1)\}$ and $p = t \cdot (x_1, \lambda_1) + (1 - t) \cdot (x_2, \lambda_2)$ with $t \in (0, 1)$ and $(x_1, \lambda_1), (x_2, \lambda_2) \in B_X$, then clearly $\lambda_1 = \lambda_2 = p_2$, because otherwise w.l.o.g. $|\lambda_1| > 1 \Rightarrow \|(x_1, \lambda_1)\|_{\tilde{X}} > 1$. For the same reason we have $x_1 = x_2 = 0$, so $p \in \mathrm{Ext}(B_{\tilde{X}})$. Finally, if $p := (x, 0)$ with $x \in \mathrm{Ext}(B_X)$ and $p = t \cdot (x_1, \lambda_1) + (1 - t) \cdot (x_2, \lambda_2)$ with $t \in (0, 1)$ and $(x_1, \lambda_1), (x_2, \lambda_2) \in B_X$, we get $x_1 = x_2 = x$ (since $x \in \mathrm{Ext}(B_X)$) and thus $\lambda_1 = \lambda_2 = 0$ which means $p \in \mathrm{Ext}(B_X)$. $\square$

Now we are ready to characterise the extreme points of $B_{m,n}$.

Theorem 5.9. *Let $m, n \in \mathbb{N}$ be odd with $m > n$. Define $f_+(a,b) := 1 - \|(a,b,0)\|_{m,n}$ and $f_- := -f_+$ for every $(a,b) \in \mathbb{R}^2$. Then*

(i) $S_{m,n} = graph(f_+|_{V\cup W}) \cup \mathrm{graph}(f_-|_{V\cup W})$ *for V, W given by (3).*

(ii) $\mathrm{Ext}(B_{m,n}) = \{\pm(a, -\frac{m}{(m-n)^{\frac{m-n}{m}} \cdot n^{\frac{n}{m}}} a^{\frac{n}{m}}, 0) : \frac{n}{m-n} \leq a \leq \frac{n}{n+m\lambda_0}\} \cup \{\pm(0,0,1)\}$.

Proof. Consider the space $X = \mathbb{R}^2$ endowed with the norm defined by $\|(a,b)\| = \|(a,b,0)\|_{m,n}$ for every $(a,b) \in \mathbb{R}^2$. By taking a look at the calculations made in Theorem 5.6, we notice that

$$B_X = \{(a,b) \in \mathbb{R}^2 : \|(a,b,0)\|_{m,n} \leq 1\} = \pi_{ab}(B_{m,n})$$

and thus

$$Ext(B_X) = \{\pm(a, -\Gamma(a)) : \frac{n}{m-n} \leq a \leq \frac{n}{n+m\lambda_0}\},$$

where Γ is the function given by (3). This follows directly from strict concavity of Γ on $[-\frac{n}{n+m\lambda_0}, -\frac{n}{m-n}]$ and $[\frac{n}{m-n}, \frac{n}{n+m\lambda_0}]$, affinity of $b(a) = -1 - a$ and $b(a) = 1 - a$ and the fact that we only have to consider $(a,b) \in S_X$, which coincides with the boundary of $\pi_{ab}(B_{m,n})$ and is parameterized by those four functions. Taking another look at Figure 2 helps to understand that. In Lemma 5.2 we obtained that $\|(a,b,c)\|_{m,n} = \|(a,b,0)\|_{m,n} + |c|$, so by defining $\tilde{X} := X \oplus \mathbb{R}$ and endowing it with the sum norm as in Lemma 5.8, we have $\|.\|_{\tilde{X}} = \|.\|_{m,n}$. Hence we can identify $(\tilde{X}, \|.\|_{\tilde{X}})$ with $(\mathbb{R}^3, \|.\|_{m,n})$. By applying Lemma 5.8, (i) follows directly and for (ii), we get

$$\mathrm{Ext}(B_{m,n}) = \{\pm(a, -\Gamma(a)) : \frac{n}{m-n} \leq a \leq \frac{n}{n+m\lambda_0}\} \cup \{\pm(0,0,1)\}$$

which concludes the proof. $\square$

6 $\mathcal{M}_{m,n}(x)$ for odd $m, n \in \mathbb{N}$

In order to determine an explicit formula for $\mathcal{M}_{m,n}(x)$ for $x \in [-1,1]$ and $m, n \in \mathbb{N}$ odd numbers with $m > n$, we have to prove some technical lemmata.

Lemma 6.1. *Let $m, n \in \mathbb{N}$ be odd numbers with $m > n$ and let λ_0 be the real number given by Lemma 5.1. Then*

$$|\lambda_0| \cdot \frac{n}{m} \quad < \quad |\lambda_0| \cdot \frac{1 - |\lambda_0|^{\frac{n}{m-n}}}{1 - |\lambda_0|^{\frac{m}{m-n}}} \quad < \quad \frac{n}{m}$$

Proof. For the first inequality notice that according to Lemma 5.1, we have $|\lambda_0| < \frac{n}{m} < 1$ and consider the following inequality:

$$\frac{n}{m} < \frac{1 - x^n}{1 - x^m} \tag{6}$$

We will show that (6) holds for every $x \in (0, (\frac{n}{m})^{\frac{1}{m-n}})$. First, by multiplying with m and $1 - x^m$, we see that if $x \in (0,1)$, then (6) is equivalent to

$$m - n > mx^n - nx^m. \tag{7}$$

For $f(x) := mx^n - nx^m$, we have $f'(x) = nm(x^{n-1} - x^{m-1}) > 0$ for $x \in (0,1)$ and hence the right side of (7) is strictly increasing in x on $(0,1)$. Since $m - n$ is constant w.r.t. x, the graph of f meets the line $y = m - n$ at, at most, one point in $[0,1]$. Since $f(1) = m - n$ and f is strictly increasing on $(0,1)$, we obtain that f must be smaller than $m - n$ over $(0,1)$. Thus, (6) holds for every $x \in (0,1)$, in particular for every $x \in (0, (\frac{n}{m})^{\frac{1}{m-n}})$. By setting $|\lambda_0|^{\frac{1}{m-n}} = x$, the first inequality is proved.
For the second equality, by using $n + m\lambda_0 = (m-n)|\lambda_0|^{\frac{m}{m-n}}$, which follows from the definition of λ_0 in Lemma 5.1 and $n + m\lambda_0 > 0$, since $\lambda_0 \in (-\frac{n}{m}, 0)$, we get

$$\begin{aligned} |\lambda_0| \cdot \frac{1-|\lambda_0|^{\frac{n}{m-n}}}{1-|\lambda_0|^{\frac{m}{m-n}}} < \frac{n}{m} \quad &\Leftrightarrow m|\lambda_0| \cdot (1 - |\lambda_0|^{\frac{n}{m-n}}) < n \cdot (1 - |\lambda_0|^{\frac{m}{m-n}}) \\ &\Leftrightarrow m|\lambda_0| - n < (m-n)|\lambda_0|^{\frac{m}{m-n}} \\ &\Leftrightarrow m|\lambda_0| - n < n + m\lambda_0 \\ &\Leftrightarrow 2m|\lambda_0| < 2n, \end{aligned}$$

which is true, because $\lambda_0 \in (-\frac{n}{m}, 0)$. ☐

Lemma 6.2. *Let $m, n \in \mathbb{N}$ be odd numbers with $m > n$ and let λ_0 be the real number given by Lemma 5.1. Define*

$$\begin{aligned} f(x) &:= \tfrac{mn}{m-n} \cdot x^{n-1} \cdot |x^{m-n} - 1|, \qquad & x \in \mathbb{R}, \\ g(x) &:= \tfrac{mn}{n+m\lambda_0} \cdot x^{n-1} \cdot |x^{m-n} + \lambda_0|, & x \in \mathbb{R}. \end{aligned}$$

Assume

$$0 \le |x| \le \left(\frac{|\lambda_0| n}{m}\right)^{\frac{1}{m-n}} \quad \text{or} \quad \left(\frac{n}{m}\right)^{\frac{1}{m-n}} \le |x| \le 1,$$

then we have $f(x) \le g(x)$.

Proof. Since the exponents $n - 1$ and $m - n$ are even, f and g are symmetric and thus we can assume that $0 \le x \le 1$. Now for $x = 0$ the statement is trivial and for $x = 1$ we have $n + m\lambda_0 > 0$, since $\lambda_0 > -\frac{n}{m}$ according to Lemma 5.1 and hence $g(1) > 0 = f(1)$. So let us assume $0 < x < 1$. Then we have

$$\begin{aligned} f(x) = g(x) \quad \Leftrightarrow \quad x^{n-1} \cdot mn \cdot (\tfrac{|x^{m-n}-1|}{m-n} - \tfrac{|x^{m-n}+\lambda_0|}{n+m\lambda_0}) &= 0 \\ \overset{5.1}{\Leftrightarrow} \qquad |x^{m-n} - 1| \cdot |\lambda_0|^{\frac{m}{m-n}} &= |x^{m-n} + \lambda_0| \\ \Leftrightarrow \qquad |\lambda_0|^{\frac{m}{m-n}} - x^{m-n} \cdot |\lambda_0|^{\frac{m}{m-n}} &= |x^{m-n} + \lambda_0| \end{aligned}$$

and that means for $x_1^{m-n} > |\lambda_0|$, using $|\lambda_0|^{\frac{m}{m-n}} = \frac{n+m\lambda_0}{m-n}$ (Lemma 5.1) and $|\lambda_0| = -\lambda_0$, we have

$$\begin{array}{rrcl} & |\lambda_0|^{\frac{m}{m-n}} + |\lambda_0| & = & x_1^{m-n} \cdot (1 + |\lambda_0|^{\frac{m}{m-n}}) \\ \Leftrightarrow & \frac{n+m\lambda_0}{m-n} + |\lambda_0| & = & x_1^{m-n} \cdot (1 + \frac{n+m\lambda_0}{m-n}) \\ \Leftrightarrow & n \cdot (1 - |\lambda_0|) & = & x_1^{m-n} \cdot m \cdot (1 + \lambda_0) \\ \Leftrightarrow & x_1 & = & \left(\frac{n}{m}\right)^{\frac{1}{m-n}} \end{array}$$

and for $x_2^{m-n} < |\lambda_0|$ we get

$$\begin{array}{rrcl} & x_2^{m-n} \cdot (1 - |\lambda_0|^{\frac{m}{m-n}}) & = & |\lambda_0| - |\lambda_0|^{\frac{m}{m-n}} \\ \Leftrightarrow & x_2 & = & \left(|\lambda_0| \cdot \frac{1-|\lambda_0|^{\frac{n}{m-n}}}{1-|\lambda_0|^{\frac{m}{m-n}}}\right)^{\frac{1}{m-n}}. \end{array}$$

Now we can apply Lemma 6.1 to conclude that neither x_1 nor x_2 are in the intervals $\left(0, \left(\frac{|\lambda_0|n}{m}\right)^{\frac{1}{m-n}}\right)$ and $\left(\left(\frac{n}{m}\right)^{\frac{1}{m-n}}, 1\right)$. Thus, by continuity and since there is no other point $x \in (0,1)$ with $f(x) = g(x)$, we see that one of the functions must be bigger than the other one on each of the previous intervals. We already obtained $f(1) < g(1)$. On the other hand

$$\begin{array}{rrcl} & f\left(\left(\frac{|\lambda_0|n}{m}\right)^{\frac{1}{m-n}}\right) & < & g\left(\left(\frac{|\lambda_0|n}{m}\right)^{\frac{1}{m-n}}\right) \\ \Leftrightarrow & \frac{\left|\frac{|\lambda_0|n}{m} - 1\right|}{m-n} & < & \frac{\left|\frac{|\lambda_0|n}{m} + \lambda_0\right|}{n+m\lambda_0} \\ \overset{(\frac{|\lambda_0|n}{m} < |\lambda_0| < 1)}{\Leftrightarrow} & (m - |\lambda_0|n) \cdot (n + m\lambda_0) & < & (|\lambda_0|m - |\lambda_0|n) \cdot (m - n) \\ \overset{5.1}{\Leftrightarrow} & (n + \frac{1}{\lambda_0} \cdot m) \cdot |\lambda_0|^{\frac{n}{m-n}} & > & \frac{m-n}{\lambda_0} \end{array}$$

Since $\frac{1}{\lambda_0} < -1$, Lemma 5.1, (ii), delivers $n + \frac{1}{\lambda_0} \cdot m > -\frac{m-n}{|\lambda_0|^{\frac{m}{m-n}}}$, so we get

$$(n + \frac{1}{\lambda_0} \cdot m) \cdot |\lambda_0|^{\frac{n}{m-n}} > -\frac{m-n}{|\lambda_0|} = \frac{m-n}{\lambda_0}$$

which concludes the proof. $\square$

Lemma 6.3. *Let $m, n \in \mathbb{N}$ be odd numbers with $m > n$ and let λ_0 be the real number given by Lemma 5.1. Define*

$$\begin{array}{rll} f(x) & := \frac{mn}{m-n} \cdot x^{n-1} \cdot |x^{m-n} - 1|, & x \in \mathbb{R}, \\ g(x) & := \frac{mn}{n+m\lambda_0} \cdot x^{n-1} \cdot |x^{m-n} + \lambda_0|, & x \in \mathbb{R}, \\ h(x) & := n \cdot \left(\frac{n}{m}\right)^{\frac{n}{m-n}} \cdot \frac{1}{|x|}, & x \in \mathbb{R}\backslash\{0\}. \end{array}$$

Assume

$$x_1 := \left(\frac{|\lambda_0|n}{m}\right)^{\frac{1}{m-n}} \quad \leq \quad |x| \quad \leq \quad \left(\frac{n}{m}\right)^{\frac{1}{m-n}} =: x_2, \tag{8}$$

then we have $h(x) \geq \max\{f(x), g(x)\}$.

Proof. Suppose (8) holds. Since the exponents $n-1$ and $m-n$ are even and $x \mapsto |x|$ is symmetric, f, g and h are symmetric and thus we can assume that $x > 0$. Clearly, it suffices to show $h(x) \geq f(x)$ and $h(x) \geq g(x)$.
"$h(x) \geq f(x)$": Let us take a look at the function $\varphi(x) := x^n - x^m$. Since $\varphi'(x) > 0$ for all $x \in (0,1)$, we see that φ is strictly increasing over the interval $(0, x_2]$. Thus the maximum of φ over $[0, x_2]$ is attained at $x_2 = (\frac{n}{m})^{\frac{1}{m-n}}$. Its value is

$$\varphi(x_2) = x_2^n - x_2^m = \left(\frac{n}{m}\right)^{\frac{n}{m-n}} - \left(\frac{n}{m}\right)^{\frac{m}{m-n}} = \frac{(m-n) \cdot n^{\frac{n}{m-n}}}{m^{\frac{m}{m-n}}}.$$

Hence,

$$x^n - x^m \leq \frac{(m-n) \cdot n^{\frac{n}{m-n}}}{m^{\frac{m}{m-n}}} \quad \text{for} \quad x_1 \leq x \leq x_2.$$

Using this inequality and the fact that $x > 0$, we obtain

$$\begin{aligned} f(x) &= \tfrac{mn}{m-n} \cdot x^{n-1} \cdot |x^{m-n} - 1| = \tfrac{mn}{m-n} \cdot \tfrac{1}{x} \cdot (x^n - x^m) \\ &\leq \tfrac{mn}{m-n} \cdot \tfrac{1}{x} \cdot \tfrac{(m-n) \cdot n^{\frac{n}{m-n}}}{m^{\frac{m}{m-n}}} = n \cdot \left(\tfrac{n}{m}\right)^{\frac{n}{m-n}} \cdot \tfrac{1}{x} = h(x) \end{aligned}$$

for all $x \in [x_1, x_2]$.
"$h(x) \geq g(x)$": For $x \in [x_1, x_2]$, we have that the inequality

$$\frac{mn}{n + m\lambda_0} \cdot x^{n-1} \cdot |x^{m-n} + \lambda_0| \leq n \cdot \left(\frac{n}{m}\right)^{\frac{n}{m-n}} \cdot \frac{1}{|x|}$$

is equivalent to

$$\frac{m}{n + m\lambda_0} \cdot |x^m + \lambda_0 x^n| \leq \left(\frac{n}{m}\right)^{\frac{n}{m-n}}$$

For $\psi(x) := x^m + \lambda_0 x^n$, we obtain $\psi'(x) = mx^{m-1} + \lambda_0 n x^{n-1} = 0 \Leftrightarrow x = 0 \vee x = \pm\left(\frac{|\lambda_0| n}{m}\right)^{\frac{1}{m-n}} = \pm x_1$. So ψ is monotone over the interval $[x_1, x_2]$. Using $n + m\lambda_0 = (m-n)|\lambda_0|^{\frac{m}{m-n}}$ (from Lemma 5.1) and $|\lambda_0| = -\lambda_0$, we get

$$\begin{aligned} \tfrac{m}{n+m\lambda_0} \cdot |\psi(x_1)| &= \tfrac{m}{n+m\lambda_0} \cdot \left| |\lambda_0|^{\frac{m}{m-n}} \cdot \left(\left(\tfrac{n}{m}\right)^{\frac{m}{m-n}} - \left(\tfrac{n}{m}\right)^{\frac{n}{m-n}} \right) \right| \\ &= \left(\tfrac{n}{m}\right)^{\frac{n}{m-n}} \end{aligned}$$

and

$$\begin{aligned} \tfrac{m}{n+m\lambda_0} \cdot |\psi(x_2)| &= \tfrac{m}{n+m\lambda_0} \cdot \left| \left(\tfrac{n}{m}\right)^{\frac{m}{m-n}} + \lambda_0 \cdot \left(\tfrac{n}{m}\right)^{\frac{n}{m-n}} \right| \\ &= \tfrac{m}{n+m\lambda_0} \cdot \left(\tfrac{n}{m}\right)^{\frac{n}{m-n}} \cdot \left| \left(\tfrac{n}{m}\right) + \lambda_0 \right| \\ &= \left(\tfrac{n}{m}\right)^{\frac{n}{m-n}} \end{aligned}$$

Monotony of ψ over $[x_1, x_2]$ delivers

$$\psi(x) \in \Big[\min\big\{\psi(x_1), \psi(x_2)\big\}, \max\big\{\psi(x_1), \psi(x_2)\big\} \Big]$$

and thus

$$|\psi(x)| \leq \max\big\{|\psi(x_1)|, |\psi(x_2)|\big\}$$

Finally,

$$\frac{m}{n+m\lambda_0} \cdot |x^m + \lambda_0 x^n| \leq \frac{m}{n+m\lambda_0} \cdot \max\big\{|\psi(x_1)|, |\psi(x_2)|\big\} = \big(\frac{n}{m}\big)^{\frac{n}{m-n}}.$$

□

We are now ready to state and prove the main theorem of this section.

Theorem 6.4. *Let $m, n \in \mathbb{N}$ be odd numbers with $m > n$ and let λ_0 be the real number given by Lemma 5.1. Define*

$$I_{m,n} := [x_1, x_2] = \left[\Big(\frac{|\lambda_0|n}{m}\Big)^{\frac{1}{m-n}}, \Big(\frac{n}{m}\Big)^{\frac{1}{m-n}}\right],$$

then

$$\mathcal{M}_{m,n}(x) = \begin{cases} n \cdot \big(\frac{n}{m}\big)^{\frac{n}{m-n}} \cdot \frac{1}{|x|}, & \text{if} \quad |x| \in I_{m,n}, \\ \frac{mn}{n+m\lambda_0} \cdot x^{n-1} \cdot |x^{m-n} + \lambda_0|, & \text{if} \quad |x| \in [0,1] \backslash I_{m,n}, \end{cases}$$

with $\mathcal{M}_{m,n}(x)$ from Definition 4.1. In addition, $\mathcal{M}_{m,n}$ is continuous on $[-1, 1]$.

Proof. We have that

$$\mathcal{M}_{m,n}(x) = \max_{p \,\in\, \mathrm{Ext}(B_{m,n})} |p'(x)|,$$

where $B_{m,n}$ is the unit ball of the space $(\mathbb{R}^3, \|.\|_{m,n})$. So we can restrict our attention to the extreme points obtained in Theorem 5.9. Since for $p \equiv \pm 1$, it holds $|p'(x)| \equiv 0$, the contribution of $\pm(0, 0, 1)$ to $\mathcal{M}_{m,n}(x)$ is irrelevant. Thus it suffices to consider the polynomials

$$p_a(x) = \pm\Big(ax^m - \frac{m}{(m-n)^{\frac{m-n}{m}} \cdot n^{\frac{n}{m}}} a^{\frac{n}{m}} x^n\Big)$$

for $\frac{n}{m-n} \leq a \leq \frac{n}{n+m\lambda_0}$. It follows

$$\begin{aligned} \mathcal{M}_{m,n}(x) &= \sup_{\frac{n}{m-n} \leq a \leq \frac{n}{n+m\lambda_0}} \big|p_a'(x)\big| \\ &= \sup_{\frac{n}{m-n} \leq a \leq \frac{n}{n+m\lambda_0}} \Big|amx^{m-1} - \frac{m}{(m-n)^{\frac{m-n}{m}} \cdot n^{\frac{n}{m}}} a^{\frac{n}{m}} n x^{n-1}\Big| \\ &= \sup_{\frac{n}{m-n} \leq a \leq \frac{n}{n+m\lambda_0}} \Big|mx^{n-1}\Big(ax^{m-n} - \big(\tfrac{n}{m-n}\big)^{\frac{m-n}{m}} a^{\frac{n}{m}}\Big)\Big| \end{aligned}$$

We define $\varphi(a) := mx^{n-1}\big(ax^{m-n} - (\frac{n}{m-n})^{\frac{m-n}{m}} a^{\frac{n}{m}}\big)$. Clearly, the above supremum is attained at either $a_1 := \frac{n}{m-n}$, at $a_2 := \frac{n}{n+m\lambda_0}$ or at a critical point a_c of φ in the interior of the interval $\left[\frac{n}{m-n}, \frac{n}{n+m\lambda_0}\right]$. We obtain

$$\begin{array}{rcrcl}\varphi'(a_c) & = & mx^{n-1}\Big(x^{m-n} - \left(\frac{n}{m-n}\right)^{\frac{m-n}{m}} \cdot \frac{n}{m} \cdot a_c^{\frac{n-m}{m}}\Big) & = & 0 \\ \Leftrightarrow & & \frac{m}{n} \cdot \left(\frac{n}{m-n}\right)^{\frac{n-m}{m}} \cdot \frac{1}{|x|^{n-m}} & = & a_c^{\frac{n-m}{m}} \\ \Leftrightarrow & & \frac{n}{m-n} \cdot \left(\frac{n}{m}\right)^{\frac{m}{m-n}} \cdot \frac{1}{|x|^m} & = & a_c,\end{array}$$

which means we have only one critical point a_c. This leads to

$$\begin{array}{rcl}|\varphi(a_c)| & = & \frac{1}{|x|} \cdot \frac{mn}{m-n} \cdot \left|\left(\frac{n}{m}\right)^{\frac{m}{m-n}} - \left(\frac{n}{m}\right)^{\frac{n}{m-n}}\right| \\ & = & \frac{1}{|x|} \cdot \frac{mn}{m-n} \cdot \left(\frac{n}{m}\right)^{\frac{n}{m-n}}\left|\frac{n}{m} - 1\right| \\ & = & \frac{n}{|x|} \cdot \left(\frac{n}{m}\right)^{\frac{n}{m-n}}\end{array}$$

However, the condition $\frac{n}{m-n} \le a_c \le \frac{n}{n+m\lambda_0}$ and $n + m\lambda_0 = (m-n)|\lambda_0|^{\frac{m}{m-n}}$ (Lemma 5.1) imply

$$\begin{array}{rcccl} & \frac{n}{m-n} & \le \quad \frac{n}{m-n} \cdot \left(\frac{n}{m}\right)^{\frac{m}{m-n}} \cdot \frac{1}{|x|^m} & \le & \frac{n}{n+m\lambda_0} \\ \Leftrightarrow & \left(\frac{|\lambda_0|n}{m}\right)^{\frac{1}{m-n}} & \le \quad |x| & \le & \left(\frac{n}{m}\right)^{\frac{1}{m-n}}\end{array}$$

Together, we have

$$\begin{array}{rcl}\mathcal{M}_{m,n}(x) & = & \sup\limits_{\frac{n}{m-n} \le a \le \frac{n}{n+m\lambda_0}} |\varphi(a)| \\ & = & \begin{cases} \max\Big\{|\varphi(a_1)|, |\varphi(a_2)|, \frac{n}{|x|} \cdot \left(\frac{n}{m}\right)^{\frac{n}{m-n}}\Big\}, & \text{if } |x| \in I_{m,n} \\ \max\Big\{|\varphi(a_1)|, |\varphi(a_2)|\Big\}, & \text{if } |x| \in [0,1] \backslash I_{m,n}.\end{cases}\end{array}$$

We want to apply Lemma 6.2 and Lemma 6.3. Considering the functions f, g and h defined in those lemmata, we obtain

$$|\varphi(a_1)| = \frac{mn}{m-n} \cdot x^{n-1} \cdot |x^{m-n} - 1| = f(x)$$

and, using $(m-n) = (n + m\lambda_0) \cdot |\lambda_0|^{-\frac{m}{m-n}}$ (Lemma 5.1),

$$|\varphi(a_2)| = \frac{mn}{n + m\lambda_0} \cdot x^{n-1} \cdot |x^{m-n} + \lambda_0| = g(x). \tag{9}$$

Now finally

$$|\varphi(a_c)| = \frac{n}{|x|} \cdot \left(\frac{n}{m}\right)^{\frac{n}{m-n}} = h(x), \tag{10}$$

thus, the case $|x| \in [0,1]\backslash I_{m,n}$ follows from Lemma 6.2, the other case $|x| \in I_{m,n}$ from Lemma 6.3 and since we already obtained $g(x_1) = h(x_1)$ and $g(x_2) = h(x_2)$ in the proof of Lemma 6.3, $\mathcal{M}_{m,n}(x)$ is continuous. □

Definition 6.5. ($M_{m,n}$)
We define $M_{m,n}$ to be the smallest possible constant, such that

$$|p'(x)| \leq M_{m,n}\|p\| \quad \text{for all } x \in [-1,1], p \in \mathcal{P}_{m,n}(\mathbb{R}).$$

Clearly

$$M_{m,n} = \sup_{x\in[-1,1]} \mathcal{M}_{m,n}(x).$$

Proposition 6.6. *Let $m, n \in \mathbb{N}$ be odd numbers with $m > n$ and let λ_0 be the real number given by Lemma 5.1. Then*

$$M_{m,n} = \mathcal{M}_{m,n}(\pm 1) = \frac{mn(1+\lambda_0)}{n+m\lambda_0}$$

and equality is attained for the polynomials

$$p(x) = \pm\frac{nx^m + \lambda_0 m x^n}{n+m\lambda_0}.$$

Proof. According to Definition 6.5, we want to obtain $\sup_{x\in[-1,1]} \mathcal{M}_{m,n}(x)$. By Theorem 6.4, we have

$$\mathcal{M}_{m,n}(x) = \begin{cases} n \cdot \left(\frac{n}{m}\right)^{\frac{n}{m-n}} \cdot \frac{1}{|x|}, & \text{if} \quad |x| \in I_{m,n}, \\ \frac{mn}{n+m\lambda_0} \cdot x^{n-1} \cdot |x^{m-n} + \lambda_0|, & \text{if} \quad |x| \in [0,1]\setminus I_{m,n}, \end{cases}$$

where

$$I_{m,n} := [x_1, x_2] = \left[\left(\frac{|\lambda_0|n}{m}\right)^{\frac{1}{m-n}}, \left(\frac{n}{m}\right)^{\frac{1}{m-n}}\right].$$

Since $\mathcal{M}_{m,n}(x)$ is symmetric, monotonically decreasing on $I_{m,n}$, and by Theorem 6.4 continuous, we get

$$\sup_{|x|\in I_{m,n}} \mathcal{M}_{m,n}(x) = \mathcal{M}_{m,n}(x_1) = \lim_{x\nearrow x_1} \mathcal{M}_{m,n}(x)|_{[0,1]\setminus I_{m,n}}.$$

Thus, again by symmetry of $\mathcal{M}_{m,n}(x)$, we only have to find

$$\sup_{x\in[0,1]\setminus I_{m,n}} \mathcal{M}_{m,n}(x) = \sup_{x\in[0,1]\setminus I_{m,n}} \frac{mn}{n+m\lambda_0} \cdot x^{n-1} \cdot |x^{m-n} + \lambda_0|.$$

Hence, we want to find x_{max} with

$$x_{max}^{n-1} \cdot |x_{max}^{m-n} + \lambda_0| = \max_{x\in[0,x_1]\cup[x_2,1]} x^{n-1} \cdot |x^{m-n} + \lambda_0| =: \varphi(x)$$

Since $x^{m-n} < |\lambda_0|$ for $x \in [0, x_1]$ we have

$$\max_{x\in[0,x_1]} \varphi(x) = \max_{x\in[0,x_1]} |\lambda_0|x^{n-1} - x^{m-1}$$

This maximum is attained at either $x = 0$, at $x = x_1$ or at a critical point of φ in $(0, x_1)$. For $n = 1$, it is easy to see that $x_{max} = 0$, for $n > 1$ we have

$$(n-1)\cdot|\lambda_0|x^{n-2} - (m-1)x^{m-2} = 0 \quad\Leftrightarrow\quad x = 0 \quad\vee\quad x = \left(\tfrac{(n-1)|\lambda_0|}{m-1}\right)^{\frac{1}{m-n}} =: x_c$$

Since x_c is the only critical point of φ in $(0,\infty)$ and $\lim\limits_{x\to\infty} \varphi'(x) = -\infty$, it follows $\varphi'(x) < 0$ for $x > x_1$. Now since $\varphi(x_c) > \varphi(0) = 0$, it holds

$$\max_{x\in[0,x_1]} \varphi(x) = \varphi(x_c) = x_c^{n-1}(|\lambda_0| - \frac{(n-1)|\lambda_0|}{m-1}) = x_c^{n-1}\frac{(m-n)|\lambda_0|}{m-1}$$

Clearly, $\mathcal{M}_{m,n}(x)$ is monotonically increasing on $[x_2, 1]$, so we only have to compare $\varphi(1)$ with $\varphi(0)$ (if $n = 1$) and $\varphi(x_c)$ (if $n > 1$). Using Lemma 5.1, for $n = 1$ (note that $\frac{n}{m} = \frac{1}{m} < \frac{m-1}{m} = \frac{m-n}{m}$, since $m \geq 3$) we obtain

$$\varphi(0) = |\lambda_0| < \frac{n}{m} < \frac{m-n}{m} < 1 - |\lambda_0| = \varphi(1)$$

and for $n > 1$ we get

$$\varphi(x_c) = x_c^{n-1}\frac{(m-n)|\lambda_0|}{m-1} < \frac{(m-n)|\lambda_0|}{m-1} < |\lambda_0| < \varphi(1)$$

So

$$M_{m,n} = \sup_{x\in[-1,1]} \mathcal{M}_{m,n}(x) = \mathcal{M}_{m,n}(\pm 1) = \frac{mn}{n+m\lambda_0}\cdot\varphi(\pm 1) = \frac{mn(1+\lambda_0)}{n+m\lambda_0}$$

The second point follows directly by deriving p. $\square$

The results obtained in this chapter are visualised below. To notice the difference between $n = 1$ and $n > 1$, the plots of $\mathcal{M}_{m,n}(x)$ for $m = 3$ and $n = 1$ (Figure 3b) and for $m = 5$ and $n = 3$ (Figure 3a) are shown.

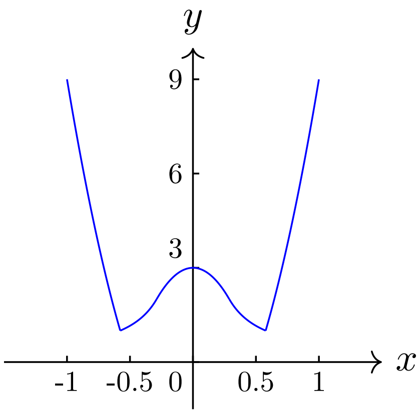

Figure 3a: $\mathcal{M}_{m,n}(x)$ for $m = 3$ and $n = 1$.
In this case $M_{3,1} = 9$.

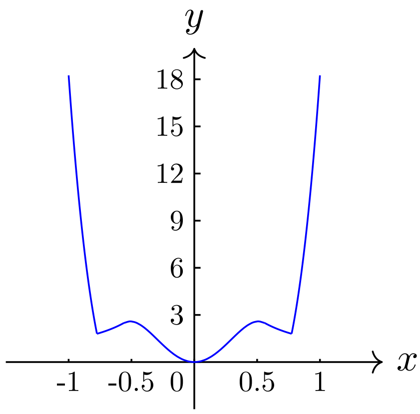

Figure 3b: $\mathcal{M}_{m,n}(x)$ for $m = 5$ and $n = 3$.
In this case $M_{5,3} = 9$.

7 Appendix

Appendix A. The Krein-Milman approach

Definition A.1. (Convex body)
A compact, convex, and nonempty set is called a convex body.

To investigate polynomial inequalities, it can be very useful to determine the set of extreme points $\operatorname{Ext}(C)$ of a convex set C. The key result connecting extreme points and inequalities is called the *Krein-Milman approach.* We are going to prove that for a convex function (like a polynomial norm) defined on a convex body C in a finite dimensional topological vector space (in our case the unit ball of a finite dimensional polynomial space) that attains its maximum on C, we can find $e \in \operatorname{Ext}(C)$ with $f(e) = \max_{x \in C} f(x)$. Therefore we can restrict our attention to the extreme points of the domain of the target function.
To prove this, we need the Minkowski-Carathéodory Theorem (Theorem A.17). The proof presented here follows [8, chapter 8, pages 122-126].

Definition A.2. (Face of C)
Let C be a convex subset of a vector space X. A (proper) convex extreme subset of C is called a (proper) face of C.

Lemma A.3. *Let A be a convex subset of a real vector space X. Let $f : A \to \mathbb{R}$ be a linear functional with*

(*i*) $\sup_{x \in A} f(x) = \gamma < \infty$

(*ii*) *The restriction $f|_A$ of f to A is not constant.*

If $F := \{x \in A : f(x) = \gamma\} \neq \emptyset$, then it is a proper face of A.

Proof. Since f is linear, F is convex. Let $x, y \in A, \lambda \in (0,1)$ and $\lambda x + (1-\lambda)y \in F$. Since f is linear, we have $\lambda f(x) + (1-\lambda)f(y) = f(\lambda x + (1-\lambda)y) = \gamma$, and $f(x), f(y) \leq \gamma$, therefore $f(x) = f(y) = \gamma$. Thus $x, y \in F$. It follows that F is a face of A. According to (ii), we can find a point $x \in A$ with $f(x) < \gamma$, which means $x \notin F$ and therefore F is a proper face of A. □

Lemma A.4. *Let X be a topological vector space and $A \subseteq X$ be convex. Then $\overline{A}$ and A° are convex.*

Proof. For a proof we refer to [9, Lemma 2.1.7, pages 19-20]. □

Lemma A.5. *Let X be a topological vector space and $A \subseteq X$ be convex. Then*

(*i*) *Every proper extreme subset of A lies in the boundary of A.*

(*ii*) *If X is locally convex and $A^\circ \neq \emptyset$, then each point $x \in A \cap \partial A$ lies in some proper face of A.*

Proof.
"(i)": Let $E \subsetneq A$ be a proper extreme subset of A, $x \in E$ and $y \in A \backslash E$. Let $z : \mathbb{R} \to X$ be the continuous function $z(\lambda) := \lambda x + (1-\lambda)y$. Since E is extremal, we have $(1, \infty) \subseteq z^{-1}(A^C)$. It follows that $x \in A \cap \overline{A^C} \subseteq \partial A$.
"(ii)": Let $x \in A \cap \partial A$ and set $B = A^\circ$. According to Lemma A.4, B is convex. The Hahn-Banach Theorem (Theorem 2.5) provides a continuous linear functional $f : X \to \mathbb{R}$ with $f(b) < \gamma \leq f(x)$ for all $b \in B$. Let $y \in A$ and choose $b \in B$. For each $\lambda \in (0,1]$ the set $\lambda B + (1-\lambda)y$ is an open subset of A, and hence contained in B. Thus $f(y) = \lim_{\lambda \to 0} f(\lambda b + (1-\lambda)y) \leq \gamma$. It follows that $f(x) = \gamma = \sup_{y \in A} f(y)$ and Lemma A.3 shows that $\{y \in A : f(y) = \gamma\}$ is a proper face of A. □

Proposition A.6. *Let X be a topological vector space, $A \subseteq X$ be a convex body and $E \subseteq A$ a face of A. Let $B \subseteq E$. Then B is a face of E if and only if it is a face of A. In particular,*

$$\mathrm{Ext}(E) = E \cap \mathrm{Ext}(A)$$

Proof. If B is a face of E and $\lambda x + (1-\lambda)y \in B$ for $x, y \in A, \lambda \in (0,1)$, then since E is a face of A, we have $x, y \in E$ and thus $x, y \in B$. It follows that B is a face of A.
If B is a face of A and $\lambda x + (1-\lambda)y \in B$ for $x, y \in E, \lambda \in (0,1)$, then $x, y \in A$ and since B is a face of A also $x, y \in B$. Thus B is a face of E. □

Definition A.7. (Affine subspac)
Let X be a vector space. An *affine subspace* is a set of the form $a + V$ with $a \in X$ and a linear subspace $V \subseteq X$. The *affine span* of a subset $A \subseteq X$ is the smallest affine subspace containing A. If $A = \{a_1, \dots, a_n\}$, then its affine span is

$$\mathrm{Aff}(a_1, \dots, a_n) = \Big\{ \sum_{i=1}^{n} \lambda_i a_i : \lambda_i \in \mathbb{R}, i \in \{1, \dots, n\}, \sum_{i=1}^{n} \lambda_i = 1 \Big\}.$$

This can be easily seen, since $\sum_{i=1}^{n} \lambda_i = 1$ implies

$$\sum_{i=1}^{n} \lambda_i a_i = a_1 + \sum_{i=2}^{n} \lambda_i (a_i - a_1),$$

which is exactly a_1 plus the vector span of $\{a_i - a_1 : i \in \{2, \dots, n\}\}$.
The set $\{a_1, \dots, a_n\}$ is called affinely independent, if and only if $\sum_{i=1}^{n} \lambda_i a_i = 0$ and $\sum_{i=1}^{n} \lambda_i = 0$ implies $\lambda_i = 0, i \in \{1, \dots, n\}$. This is the case if and only if $\{a_i - a_1 : i \in \{2, \dots, n\}\}$ is linearly independent, as $\sum_{i=1}^{n} \lambda_i = 0$ implies

$$\sum_{i=1}^{n} \lambda_i a_i = \sum_{i=2}^{n} \lambda_i a_i + (-\sum_{i=2}^{n} \lambda_i) a_1 = \sum_{i=2}^{n} \lambda_i (a_i - a_1).$$

Theorem A.8. *Let X be a real (or complex) topological vector space and Y be a linear subspace of dimension n. Then every isomorphism mapping $\mathbb{R}^n$ (or $\mathbb{C}^n$) to Y is also a homeomorphism from $(\mathbb{R}^n, \|.\|)$ (or $(\mathbb{C}^n, \|.\|)$) to $(Y, \mathcal{T})$.*

Proof. For a proof we refer to [9, Satz 2.2.1, (i), page 25]. □

Corollary A.9. (Uniqueness of finite-dimensional t.v.s.)
Let X be a finite-dimensional vector space. Then there exists a unique topology, which satisfies that $(X, \mathcal{T})$ is a topological vector space. In particular, every real or complex topological vector space of dimension n is homeomorphic to $\mathbb{R}^n$ or $\mathbb{C}^n$, respectively, and all norms on $\mathbb{R}^n$ or $\mathbb{C}^n$ are equivalent.

Proof. A proof can be found in [9, Korollar 2.2.2, pages 26-27]. □

Corollary A.10. *Every finite-dimensional real (or complex) topological vector space $(X, \mathcal{T})$ is normable.*

Proof. For $\dim(X) = n \in \mathbb{N}$, there exists an isomorphism $\psi : \mathbb{R}^n \to X$, which is also a homeomorphism (Theorem A.8). Clearly, the topology induced by $\|x\| := \|\psi^{-1}(x)\|$ equals $\mathcal{T}$. □

Proposition A.11. *Let $(X, \mathcal{T})$ be a real topological vector space and A be a finite subset of X. Then $co(A)$ has nonempty interior as a subset of $Aff(A)$, i.e. w.r.t. the relative topology $\mathcal{T}|_{Aff(A)}$.*

Proof. Write $A = \{a_1, \dots, a_n\}$. By successively eliminating dependent vectors from $B = \{a_i - a_1 : i \in \{2, \dots, n\}\}$, we can find a maximal independent subset of B. W.l.o.g. suppose it is $B' = \{a_2 - a_1, \dots, a_k - a_1\}$, so that $\{a_1, \dots, a_k\}$ is affinely independent and each $a_l - a_1$ with $l > k$ is a linear combination of vectors in B'. Then

$\operatorname{Aff}(A) = \operatorname{Aff}(a_1, \dots, a_k)$. By Theorem A.8, and since translations are homeomorphisms, the map

$$\varphi : \begin{cases} \mathbb{R}^{k-1} & \to \quad \operatorname{Aff}(A) \\ (\lambda_1, \dots, \lambda_{k-1}) & \mapsto \quad a_1 + \sum_{i=1}^{k-1} \lambda_i (a_{i+1} - a_1) \end{cases}$$

is a homeomorphism. The set

$$\big\{(\lambda_1, \dots, \lambda_{k-1}) \in \mathbb{R}^{k-1} : \lambda_i > 0, \sum_{i=1}^{k-1} < 1\big\}$$

is open, nonempty and mapped into $\operatorname{co}(A)$ by φ. □

Theorem A.12. *Let X be a finite-dimensional topological vector space and $A \subseteq X$ be convex. Then there is a unique affine subspace $W \subseteq X$ such that $A \subseteq W$, and as a subset of W, A has nonempty interior.*

Proof. Pick $a_1 \in A$ and set $B := A - a_1$. Note that $0 \in B$. Let V be the linear subspace generated by B, i.e. let $k \in \mathbb{N}$ and $\{b_1, \dots, b_{k-1}\} \subseteq B$ be a maximal linearly independent subset of B and define $V := \operatorname{span}(\{b_1, \dots, b_{k-1}\})$. Furthermore, define $a_i := b_{i-1} + a_1$ for $i \in \{2, \dots, k\}$, then $W := a_1 + V$ is the affine span of $\{a_1, \dots, a_k\}$. By construction, $B \subseteq V$ and thus $A \subseteq W = \operatorname{Aff}(a_1, \dots, a_k)$. By Proposition A.11, we know that $\operatorname{co}(a_1, \dots, a_k)$ has nonempty interior in W.
To show uniqueness note that any affine subspace Z containing A must contain $a_1, \dots, a_k$ and thus also $\operatorname{Aff}(a_1, \dots, a_k)$. If Z is an affine subspace of X containing A and its dimension is larger than $\dim(\operatorname{Aff}(a_1, \dots, a_k))$, $\operatorname{Aff}(a_1, \dots, a_k)$ has empty interior in Z and therefore also A has. □

Definition A.13. (Dimension of a convex set)
Let X be a finite-dimensional topological vector space and $A \subseteq X$ convex. The *dimension* of A is defined as the dimension of the unique affine subspace given by Theorem A.12.

Definition A.14. (A^{iint} and $\partial^i A$)
Let $(X, \mathcal{T})$ be a finite-dimensional topological vector space and $A \subseteq X$ convex.
The *intrinsic interior* of A is defined as the interior of A as a subset of the unique affine subspace W given by Theorem A.12 and is denoted by A^{iint}.
The *intrinsic boundary* $\partial^i A$ of A is defined by $\partial^i A := \overline{A}^{\mathcal{T}|_W} \setminus A^{iint}$.

Proposition A.15. *Let X be a finite-dimensional real topological vector space and $A \subseteq X$ be a convex body. Then*

(i) *The intrinsic boundary $\partial^i A$ of A is the union of all proper faces of A.*

(ii) *If $x \in \partial^i A$ and $y \in A^{iint}$, then $\{\lambda : (1-\lambda)x + \lambda y \in A\} = [0, \alpha]$ for some $\alpha > 1$.*

(iii) *If $x \in A^{iint}$ and $y \in A$, then $\{\lambda : (1-\lambda)x + \lambda y \in A\} \cap (-\infty, 0) \neq \emptyset$.*

Proof.
(i): According to Corollary A.10, X is locally convex. Consider the unique affine subspace W of X given by Theorem A.12. Write $W = x + V$ for some $x \in X$ and a linear subspace V of X. Since every subspace of a locally convex space endowed with the relative topology is locally convex, we have that V. If we view $A - x$ as a subset of V, then $A - x$ has nonempty interior and hence, Lemma A.5 shows that the union of all proper faces of $A - x$ equals $(A - x) \cap \partial^i (A - x)$. It follows that the union of all proper faces of A equals $A \cap \partial^i A$ and since A is compact and therefore closed, we get $A \cap \partial^i A = \partial^i A$.

(ii): We know that each point $x \in \partial^i A$ lies in some proper face F of A. Since $\emptyset = A^{iint} \cap \partial^i A$ and $A^{iint} \cap \partial^i A \supseteq A^{iint} \cap F$ by Lemma A.5, we know $y \notin F$. As A is convex, we know that $\Lambda := \{\lambda : (1 - \lambda)x + \lambda y \in A\}$ is an interval and $[0, 1] \subseteq \Lambda$. If $\Lambda \cap (-\infty, 0) \neq \emptyset$, then x is an interior point of a line segment in A with one end point $y \notin F$, which contradicts $x \in F$. Let $z : \mathbb{R} \to X$ be the continuous function $z(\lambda) := (1 - \lambda)x + \lambda y$. If $\Lambda \cap (1, \infty) = \emptyset$, then $y = \lim_{\lambda \searrow 1} z(\lambda)$ with $z(\lambda) \notin A$ for $\lambda > 1$, which contradicts $y \in A^{iint}$.

(iii): Similar to (ii), we obtain $\{\lambda : (1 - \lambda)x + \lambda y \in A\} \supseteq [0, 1]$. If $\{\lambda : (1 - \lambda)x + \lambda y \in A\} \cap (-\infty, 0) = \emptyset$, then $X = \lim_{\lambda \nearrow 0} z(\lambda)$, which contradicts $x \in A^{iint}$. □

Proposition A.16. *Let X be a finite-dimensional real topological vector space and $A \subseteq X$ be a convex body with dim$(A) = n$ and let F be a proper face of A. Then dim$(F) < n$.*

Proof. Let W be the unique n-dimensional affine space containing A given by Theorem A.12. If $\dim(F) = n$, then W must also be the unique $n-$dimensional affine subspace containing F, and according to Theorem A.12, F has nonempty interior as a subset of W. We are led to a contradiction since $F^{iint} \subseteq A^{iint}$, but $F \subseteq \partial^i A$ by Lemma A.5. Thus $\dim(F) < n$. □

We are now ready to prove the Minkowski-Carathéodory Theorem, which will then be used to verify the Krein-Milman approach.

Theorem A.17. (Minkowski–Carathéodory Theorem)
Let X be a finite-dimensional real topological vector space and $A \subseteq X$ be a convex body with $\dim(A) = n$. *Then every point in A is a convex combination of at most $n + 1$ extreme points. In fact, for any x, one can fix $e_0 \in \mathrm{Ext}(A)$ and find $e_1, \ldots, e_n \in \mathrm{Ext}(A)$ such that x is a convex combination of $\{e_i : 0 \leq i \leq n\}$. If $x \in A^\circ$, then $x = \sum_{i=0}^n \lambda_i e_i$ with $\lambda_0 > 0$. In particular,*

$$A = \mathrm{co}(\mathrm{Ext}(A)).$$

Proof. As A is nonempty, compact and convex and X is locally convex (Corollary A.10), the Krein-Milman Theorem (Theorem 2.13) guarantees the existence of extreme points. We use induction on the dimension of A. For $\dim(A) = 0$, which means $A = \{a\}$ for some $a \in X$, the statement is trivial.

Suppose the result holds for all sets B with $\dim(B) \leq n-1$ and let $\dim(A) = n$.
If $x \in \partial^i A$, then by means of Proposition A.15, we can find a proper face F of A with $x \in F$. Proposition A.16 implies $\dim(F) < n$, so by the induction hypothesis, $x = \sum_{i=1}^{n} \lambda_i e_i$ with $\lambda_i \geq 0, i \in \{1, \ldots, n\}$, $\sum_{i=1}^{n} \lambda_i = 1$ and $\{e_1, \ldots, e_n\} \subseteq \mathrm{Ext}(F)$. From Proposition A.6, we know $\mathrm{Ext}(F) \subseteq \mathrm{Ext}(A)$, so x is a convex combination of n extreme points of A. For an arbitrary $e_0 \in \mathrm{Ext}(A)$, if e_0 is already represented in the convex combination, we are done. Otherwise, however, we can just add e_0 to the convex combination with the factor $\lambda_0 = 0$.
If $x \in A^{iint}$ and $e_0 \in \mathrm{Ext}(A)$ (clearly $x \neq e_0$, see Proposition A.15, (i)). Proposition A.15, (ii), shows that $\{\lambda : (1-\lambda)e_0 + \lambda x \in A\} = [0, \alpha]$ for some $\alpha > 1$. Let $y = (1-\alpha)e_0 + \alpha x$. For $\lambda_0 := 1 - \alpha^{-1}$, we have $0 < \lambda_0 < 1$ and

$$y = (1-\alpha)e_0 + \alpha x \Leftrightarrow \alpha^{-1} y - \frac{1-\alpha}{\alpha} e_0 = x \Leftrightarrow x = \lambda_0 e_0 + (1-\lambda_0) y.$$

Let $z : \mathbb{R} \to X$ be the continuous function $z(\lambda) := (1-\lambda)e_0 + \lambda x$. We have $y = \lim_{\lambda \searrow \alpha} z(\lambda)$ and hence $y \in \partial^i A$ and therefore $y \in F$ for some proper face F of A. By the induction hypothesis we can write y as a convex combination $y = \sum_{i=1}^{n} \lambda_i e_i$ of points $e_1, \ldots, e_n \in \mathrm{Ext}(F)$. From that we see that x is the convex combination

$$x = \lambda_0 e_0 + \sum_{i=1}^{n} (1-\lambda_0) \lambda_i \, e_i.$$

It suffices to recall that $\mathrm{Ext}(F) \subseteq \mathrm{Ext}(A)$ by Proposition A.6. □

Now we are ready to verify the Krein-Milman approach. Let us recall the definition of a convex function.

Definition A.18. (Convex function)
Let X be a real vector space and $C \subseteq X$ be convex. A function $f : C \to \mathbb{R}$ is called convex, if

$$\forall x, y \in C, \lambda \in [0,1] : f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

Theorem A.19. (The Krein-Milman approach) *If C is a convex body in a finite-dimensional topological vector space and $f : C \to \mathbb{R}$ is a convex function that attains its maximum on C, then there is an extreme point $e \in C$ with $f(e) = \max\{f(x) : x \in \mathbb{C}\}$.*

Proof. Suppose f attains its maximum at $x \in C$. Then, according to Theorem A.17, we have $x = \sum_{i=1}^{n} \lambda_i e_i$ for some $n \in \mathbb{N}$ (w.l.o.g. we can assume $\lambda_i \neq 0$ for all $i \in \{1, \ldots, n\}$). By iterative use of the convexity of f and since $f(e_i) \leq f(x), i \in \{1 \ldots, n\}$, we get

$$f(x) = f(\sum_{i=1}^{n} \lambda_i e_i) \leq \sum_{i=1}^{n} \lambda_i f(e_i) \leq \sum_{i=1}^{n} \lambda_i f(x) = f(x)$$

and thus $f(e_i) = f(x), i \in \{1, \ldots, n\}$. $\square$

References

[1] Sergei N. Bernstein. *Sur l'ordre de la meilleure approximation des fonctions continues par des polynomes de degré donné*. Vol. 4. Académie Royale de Belgique, 1912.

[2] Ralph P. Boas. "Inequalities for the derivatives of polynomials". In: *Math. Mag.* 42 (1969), pp. 165–174.

[3] Narenda K. Govil and Rabindra N. Mohapatra. *Markov and Bernstein type Inequalities for Polynomials*. Vol. 3. 1999, pp. 349–387.

[4] Michael Kaltenbäck. *Aufbau Analysis*. 27th ed. Berliner Studienreihe zur Mathematik. Heldermann Verlag, 2021. ISBN: 978-3-88538-127-3.

[5] Michael Kaltenbäck. *Fundament Analysis*. Vol. 26. Berliner Studienreihe zur Mathematik. Heldermann Verlag, 2014. ISBN: 978-3-88538-126-6.

[6] Gustavo A. Muñoz-Fernández, Yannis Sarantopoulos, and Juan B. Seoane-Sepúlveda. "An Application of the Krein-Milman Theorem to Bernstein and Markov Inequalities". In: *Journal of Convex Analysis* 15 (2008), pp. 299–312.

[7] Gustavo A. Muñoz-Fernández and Juan B. Seoane-Sepúlveda. "Geometry of Banach spaces of trinomials". In: *Journal of Mathematical Analysis and Applications* 340 (2008), pp. 1069–1087.

[8] Barry Simon. *Convexity: An Analytic Viewpoint*. Cambridge Tracts in Mathematics. California Institute of Technology, 2011, pp. 121–126. ISBN: 978-1-10700-731-4.

[9] Harald Woracek, Michael Kaltenbäck, and Martin Blümlinger. *Funktionalanalysis*. 14th ed. Feb. 2020. URL: `https://www.asc.tuwien.ac.at/~woracek/homepage/downloads/lva/2017S_Fana1/fana.pdf`.