DISSERTATION

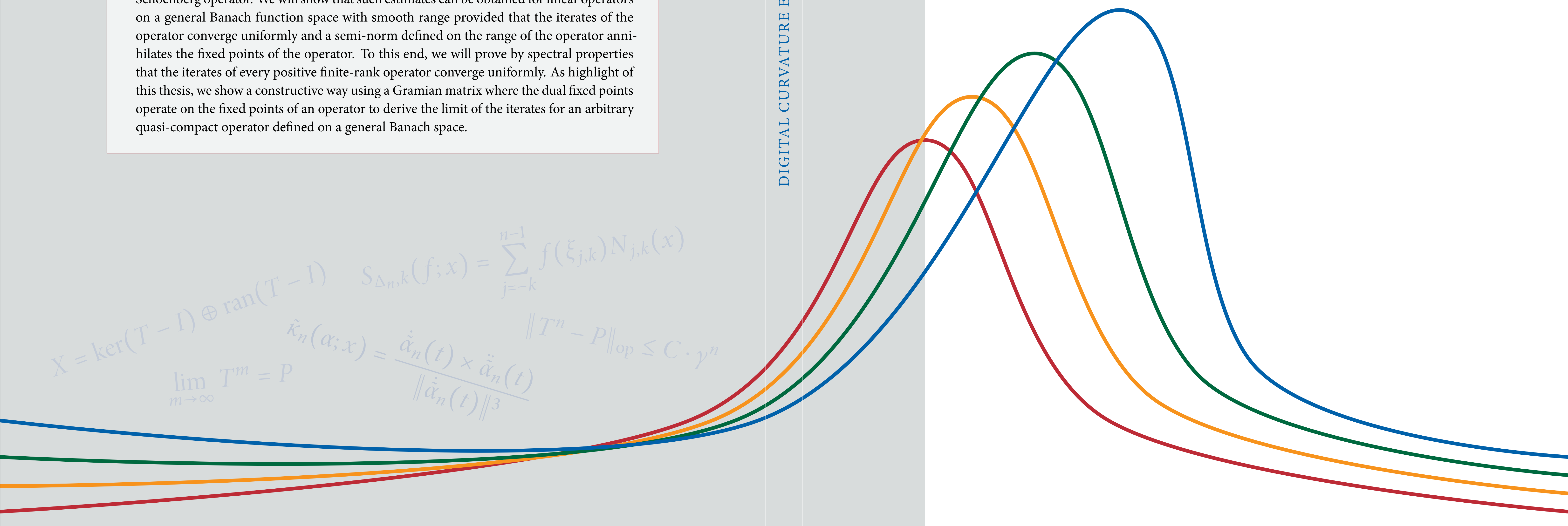
Digital Curvature Estimation: An Operator Theoretic Approach

Johannes Nagler

DIGITAL CURVATURE ESTIMATION: AN OPERATOR THEORETIC APPROACH

This thesis is divided into two parts. The first part is devoted to the curvature estimation of piecewise smooth curves using variation diminishing splines. The variation diminishing property combined with the ability to reconstruct linear functions leads to a convexity preserving approximation that is crucial if additional sign changes in the curvature estimation have to be avoided. To this end, we will first establish the foundations of variation diminishing transforms and introduce the Bernstein and the Schoenberg operator on the space of continuous functions and its generalization to the L^p-spaces. In order to be able to detect C^2-singularities in piecewise smooth curves, we establish lower estimates for the approximation error in terms of the second order modulus of smoothness for Schoenberg's variation diminishing operator. Afterwards, we consider smooth curve approximations using only finitely many samples of the curve, where the approximation, its first, and its second derivative converge uniformly to its corresponding part of the curve to be approximated. In this case, we can show that the estimated curvature converges uniformly to the real curvature if the number of samples goes to infinity. Based on the lower estimates that relates the decay rate of the approximation error with smoothness we propose a multi-scale algorithm to estimate the curvature and to detect C^2-singularities. We numerically evaluate our algorithm and compare it to others to show that our algorithm achieves competitive accuracy while our curvature estimations are significantly faster to compute.

The second part deals with generalizations of the established lower estimates for the Schoenberg operator. We will show that such estimates can be obtained for linear operators on a general Banach function space with smooth range provided that the iterates of the operator converge uniformly and a semi-norm defined on the range of the operator annihilates the fixed points of the operator. To this end, we will prove by spectral properties that the iterates of every positive finite-rank operator converge uniformly. As highlight of this thesis, we show a constructive way using a Gramian matrix where the dual fixed points operate on the fixed points of an operator to derive the limit of the iterates for an arbitrary quasi-compact operator defined on a general Banach space.



DISSERTATION

Digital Curvature Estimation: An Operator Theoretic Approach

M.Sc. Johannes Nagler

April 2015

Eingereicht an der Fakultät für Informatik und Mathematik
der Universität Passau zur Erlangung des
Grades eines Doktors der Naturwissenschaften

Betreuer/in

Prof. Dr. Brigitte Forster-Heinlein
Professur für Angewandte Mathematik
Universität Passau

Prof. Dr. Uwe Kähler
Department of Mathematics
University of Aveiro

Abstract

This thesis is divided into two parts. The first part is devoted to the curvature estimation of piecewise smooth curves using variation diminishing splines. The variation diminishing property combined with the ability to reconstruct linear functions leads to a convexity preserving approximation that is crucial if additional sign changes in the curvature estimation have to be avoided. To this end, we will first establish the foundations of variation diminishing transforms and introduce the Bernstein and the Schoenberg operator on the space of continuous functions and its generalization to the L^p-spaces. In order to be able to detect C^2-singularities in piecewise smooth curves, we establish lower estimates for the approximation error in terms of the second order modulus of smoothness for Schoenberg's variation diminishing operator. Afterwards, we consider smooth curve approximations using only finitely many samples of the curve, where the approximation, its first, and its second derivative converge uniformly to its corresponding part of the curve to be approximated. In this case, we can show that the estimated curvature converges uniformly to the real curvature if the number of samples goes to infinity. Based on the lower estimates that relates the decay rate of the approximation error with smoothness we propose a multi-scale algorithm to estimate the curvature and to detect C^2-singularities. We numerically evaluate our algorithm and compare it to others to show that our algorithm achieves competitive accuracy while our curvature estimations are significantly faster to compute.

The second part deals with generalizations of the established lower estimates for the Schoenberg operator. We will show that such estimates can be obtained for linear operators on a general Banach function space with smooth range provided that the iterates of the operator converge uniformly and a semi-norm defined on the range of the operator annihilates the fixed points of the operator. To this end, we will prove by spectral properties that the iterates of every positive finite-rank operator converge uniformly. As highlight of this thesis, we show a constructive way using a Gramian matrix where the dual fixed points operate on the fixed points of an operator to derive the limit of the iterates for an arbitrary quasi-compact operator defined on a general Banach space.

Acknowledgments

First and foremost, I want to thank my advisor Prof. Dr. Brigitte Forster-Heinlein who gave me the opportunity to start my doctoral studies under her excellent mentoring. Especially, I appreciate her patience to listen and to discuss my ideas as well as her constructive way that helped me a lot to evolve as independent researcher. Her supervision gave me the support, the confidence and the freedom to find my own way to tackle problems, to discover the topics I am interested in and to develop own perspectives. I also want to thank her for all the insights she gave me into her work and the trustful incorporation at the same eye level. Her enthusiasm for math and science as well as her open-mindedness to other people have inspired my last three years and will leave visible foot prints for my future career.

I want to express my gratitude to Paula Cerejeiras and Uwe Kähler. Especially, I wish to thank Uwe Kähler for the help to find a starting point for my research in the beginning and Paula Cerejeiras for her patience and listening skills that helped me to develop and to structure my ideas. I highly appreciate all the fruitful discussions we have had together, the joyful and relaxed atmosphere and the outstanding unconditional Portuguese hospitality both of you gave me. My thank goes also to Peter Massopust whose focused way of thinking has inspired and motivated me. I thank my colleague Thomas Fink for the careful reading of this thesis which improved the readability a lot. Besides, I would like to thank everybody else who has contributed in some sense to the development of this thesis.

Even if I am unable to express my feelings and my deep connection with words, I want to appreciate that my fiancee Veronika has crossed my path of life. She helped me to stay grounded and not to lose myself during all the intellectual work. Most notably, she guided me to be my own man and to recognize my heart over my thoughts. I am deeply grateful to meet Gerdi who helped me to face myself and to experience my limited perspectives. Finally, I am grateful to my family who gave me roots, self-reliance and everything I have needed to live my life.

Table of Contents

CHAPTER 1 Motivation

“Mathematics is not a careful march down a well-cleared highway, but a journey into a strange wilderness, where the explorer often gets lost.”

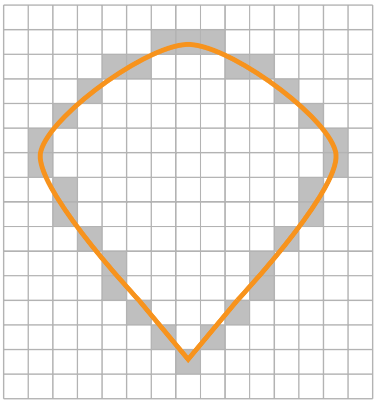
W. S. ANGLIN

THERE ARE many applications that rely on different shape properties of 2D objects in digital images. This thesis covers the problem to estimate the curvature of given 2D shapes. The curvature profile of a shape is a crucial tool in pattern recognition to match corresponding shapes, as every planar curve is uniquely determined by its curvature profile up to its orientation and translation. Accordingly, the curvature provides a natural similarity measure for 2D shapes. The theory of differential geometry provides a concrete formula of the curvature of a smooth planar curve $\alpha : I \to \mathbb{R}^2$ defined on an real interval I by

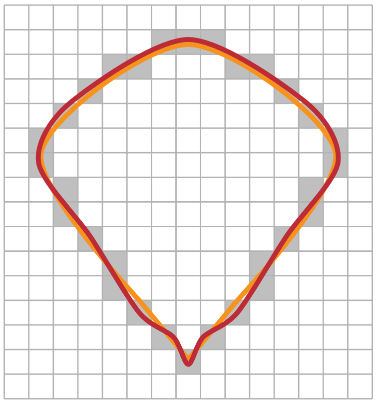
$$\kappa(t) = \frac{\det\left(\dot{\alpha}(t), \ddot{\alpha}(t)\right)}{\|\dot{\alpha}(t)\|_2^3}.$$

Despite the simple appearance of this formula, the estimation of the curvature of planar curves in digital images is a nontrivial task. Due to the digitization of real world images the continuous representation of these curves is not available, hence the exact solution is unknown. Another challenge is hidden in the denominator of this formula. The norm of the tangent vector is numerically highly unstable and the error even gets worse by the exponentiation. There is also a practical issue to handle. As curves in the real world are in general not well behaved, we should be able to detect singular points on the planar curve in order to exclude them from the curvature estimation. In the literature, the convergence of curvature estimators has already been proved recently for smooth convex shapes by several authors, see [HK07; DLF07; RL11; CLL14; LCL14].

We will show in this thesis a way to obtain an convexity preserving approximate curve representation by splines that guarantees the convergence of the approximate curvature towards the real curvature for any smooth curve, i. e., not necessarily convex. Additionally, an algorithm is presented that includes the detection of singular points on the curve by a multiscale approach. The decay rate of the approximation error



(a) exact curve



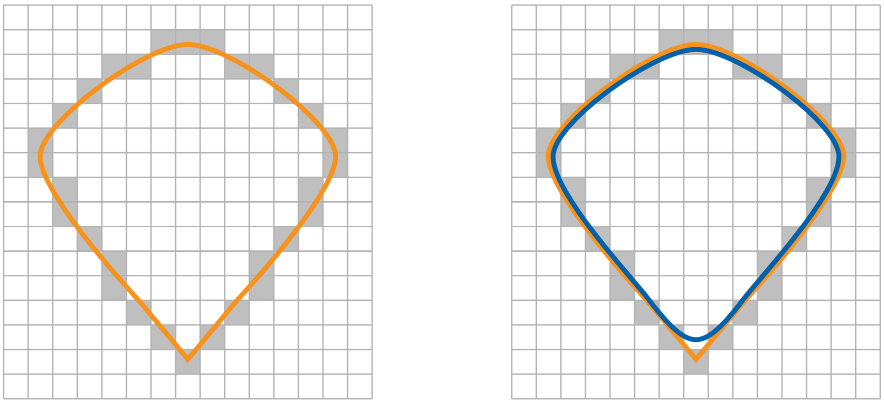
(b) smooth curve interpolation

Figure 1.1: *A planar curve in an image and its digitization. Note that the original curve is convex, whereas the interpolation introduces additional sign changes in the curvature.*

is used to weight the curvature approximation which enables a detection of singular points. Combined we provide a new algorithm that is able to estimate the curvature of the digitization of piecewise smooth curves in a stable and efficient way.

To obtain the continuous approximation of the curve there are two possible choices. Either the curve is computed by interpolation at given pixels or by an approximation scheme (quasi-interpolation). A smooth polynomial interpolation of pixel values is depicted in Figure 1.1. Clearly, it shows that this is not an appropriate choice for our case as it introduces oscillating effects around the singularity at the bottom in order to obtain a smooth interpolation. This leads to additional sign changes in the second derivative and introduces new local maxima which are often used to characterize important points on the curve. Moreover, the convexity of the original curve is not preserved. As the curvature is the deviation of the curve from a straight line, we need an approximation method which ensures that the approximated curve does not oscillate more often about any straight line as the original curve. Therefore, it is crucial to preserve the shape properties of the original curve. The shape preservation of curves can be guaranteed by only three properties, namely the preservation of the positivity, the monotonicity, and the convexity of curves. These three properties lead to the so called *shape preserving* approximation methods, where a key concept is the variation-diminishing property. An example of a shape preserving approximation method of a digitized curve is shown in Figure 1.2. Note that the shape preserving approximation method in fact preserves the convexity of the curve, but the approximation error is larger at the singularity. In fact, this observation will allow us to detect the singularity as shown later.

The estimation of a curve's curvature requires at least a C^2-smooth curve approximation that allows a stable and efficient computation of the first and the second order derivatives. The latter property is important to avoid numerical issues. A local approximation scheme would also allow for more flexibility in local changes and can avoid high computation times. If one considers all these points, there are mainly two methods that



(a) exact curve **(b)** smooth curve approximation

Figure 1.2: *A planar curve in an image and its digitization. Important to note is preservation of the shape of the approximation and the large approximation error at the singularity.*

fit these properties. An approximation by Bezier-curves based on the Bernstein polynomials [Ber12] will lead to a global approximation method that requires a uniform pixel spacing, whereas the spline approximation according to Schoenberg [Sch67] allows a nonuniform spacing and offers a local adjustment. Thus, the approximation method of our choice is the variation-diminishing spline approximation devised by Schoenberg due to its locality and flexibility with the pixel spacing. Both methods are introduced in Chapter 3.

Digitized curves are of course not naturally differentiable twice and consequently, the curvature of such curves is not defined on such points. To overcome this issue, a detection of singular point provides an elegant enhancement for piecewise smooth C^2-curves. The detection of such singular points is solved by obtaining lower estimates of the approximation error by the second order modulus of smoothness. A theory that explains the key ingredients to obtain lower estimates by moduli of smoothness or K-functionals of any order and shows how to derive these lower estimates is provided in Chapter 10. Concretely, we will show this first for Schoenberg's spline operator in Chapter 4. Accordingly, the problem of curvature estimation of discretely given points is divided into three parts: first we approximate the original continuous curve, second we compute an approximate curvature measure, and finally we detect singular points to derive a piecewise continuous approximation of the curvature. The method to approximate the curvature and to detect singular points is discussed in detail in Chapter 5. We have evaluated the curvature estimated numerically and compared the results with other recent curvature estimators.

In the second part of this thesis we will show generalizations of the lower estimates that we have shown for the Schoenberg operator. The key concept of Chapter 4 is the convergence of the iterates of the corresponding approximation operator. We will revisit the idea of the proof in Chapter 6 and outline its necessary conditions in an operator theoretic viewpoint. Accordingly, Chapter 7 shows that the iterates of positive linear operators with finite rank and a partition of unity property do always converge to a limiting operator. While the convergence is guaranteed the limiting operator has already to be known previously in this case. As highlight of this thesis, we provide in Chapter 9 a constructive method to derive the limiting operator if a basis of the fixed point space and its dual are known. Thereby, we use an approach based on operator theory to consider the iterates of quasi-compact operators on Banach spaces. This viewpoint allows the direct construction provided that the fixed point space is finite dimensional.

1.1 Outline of the thesis and highlights

We give a short overview of the organization of this thesis that provides the main results and most important references as convenience for the reader. This thesis is divided into two parts. The first part is devoted to provide a solution to the problem of curvature estimation based on variation diminishing splines. In the second part, we show generalizations based on operator theory to derive lower estimates and to construct the limiting operator of iterates.

Chapter 2 covers the variation-diminishing property introduced in 1930 by I. J. Schoenberg. A variation-diminishing transform features the property that the mapped function has no more sign changes than the function to be mapped. If additionally the positivity of functions has to be preserved, the variation-diminishing property leads to the famous and often discussed concept of total positivity, see especially the monograph of S. Karlin [Kar68]. Total positivity provides the key-role concept to construct shape preserving approximation methods.

The two operators that are important throughout the whole thesis are introduced in detail in Chapter 3: the Bernstein operator and Schoenberg's operator, while the focus lies on the spline operator devised by Schoenberg due to its local flexibility. Both operators will be introduced on the space of continuous functions and the most important properties are provided. Besides, their counterparts on the L^p-spaces are discussed. More detailed information can be found in Lorentz [Lor86] for the Bernstein operator and the Bernstein polynomials, while we refer to de Boor [dBoo01], Nürnberger [Nür89], and Schumaker [Sch07] for extensive information on splines and the Schoenberg operator. More references are given in the chapter. Both operators preserve the positivity, the monotonicity, and the convexity of functions. Derivatives of approximations by the Bernstein operator converge uniformly in all order against the real derivatives, while the derivatives of the spline approximations by Schoenberg operator converge only up to the second order derivative. For both approximation methods, the approximation error is bounded from above by the second order modulus of smoothness. While a lower estimate is known for the Bernstein operator, an analogous estimate for the Schoenberg operator is missing. We will provide such an lower estimate in Chapter 4. At the end of Chapter 3, we will also prove that the Schoenberg operator has besides 0 and 1 only distinct positive real eigenvalues.

Chapter 5 is devoted to solve the topic of this thesis. Utilizing all the results of the previous chapters, we are now able to construct approximation methods that allow an approximation of the curvature of smooth curves. We provide an estimate of the approximation error of the curvature based on the convergence rate of the first two derivatives. Therefore, for appropriate operators we are able to prove the uniform convergence of the approximated curvature towards the real curvature if the pixel spacing goes to zero. Furthermore, we present an algorithm that is able to detect singular points (corners) of piecewise smooth curves based on the lower estimate of the spline approximation error depending on the second order modulus of smoothness. In the end numerical experiments are presented that show that our algorithm outperforms the others in the sense that our spline based estimator achieves competitive accuracy but uses significantly less computation time. Besides, we demonstrate the ability to estimate the curvature of piecewise smooth curves.

Due to the importance of the proven lower estimates for the approximation with Schoenberg splines for handling piecewise smooth curves, we will start the second part of this thesis by revisiting the key concepts of the proof in Chapter 6 by an operator theoretic viewpoint. We will outline the most important conditions that are characterized by spectral properties to ensure the convergence of the iterates, the smoothness of the range and the null space of the operator.

Motivated by the spectral condition, we discuss the iterates of positive linear operators with a partition of unity property defined on a Banach function space in Chapter 7. We will show that the spectrum of such operators is not only contained in the unit circle, the only common point on the unit circle is the eigenvalue 1. By the above mentioned theorem of Katznelson and Tzafriri [KT86] we can proof the convergence of the iterates. Moreover, we show a criterion how the limiting operator can be derived. The results of this chapter have already been published in the *Journal of Mathematical Analysis and Applications* [Nag15].

The next two chapters are devoted to construct the limiting operator of the iterates of an arbitrary linear operator using techniques of Fredholm theory. To this end, we provide in Chapter 8 the necessary fundamentals of the Riesz-Schauder theory for compact operators and its generalization to Riesz operators. Finally, we introduce the class of quasi-compact operators which constitutes the necessary setting to consider the iterates of linear operators. As highlight of this thesis, we discuss in Chapter 9 the limit of the iterates of linear operators between general Banach spaces. We provide a constructive method to obtain the limiting operator. The key concept to characterize the convergence of the iterates, where the limiting operator can be provided is the concept of quasi-compactness. The main ingredience used for the results in this chapter is the work of Dunford [Dun43b] published already in 1943, the famous theorem of Katznelson and Tzafriri [KT86], and the classical results from functional analysis that can be found in Rudin [Rud91] and Heuser [Heu82].

As all the previous results have been shown on general Banach spaces, Chapter 10 provides a general framework to derive lower estimates for all operators with smooth range on general Banach functions spaces that satisfy some additional weak conditions.

We conclude this thesis in Chapter 11 with a short summary and discuss open problems.

1.2 Publications and preprints

Results of this thesis have already been partly published and submitted to different journals:

- The results of Chapter 7 have been published in the *Journal of Mathematical Analysis and Application*, see [Nag15].
- The results of Chapter 5 have been partly announced in the *Proceedings in Applied Mathematics and Mechanics*, see [Nag14].
- The results of Chapter 10 have been submitted to the *Journal of Complexity* concretely for the Schoenberg operator on $C([0,1])$ with the coauthors B. Forster and P. Cerejeiras, see [NCF14].

1.3 Notation

Let $\mathbb{R}$ be the set of real numbers and $\mathbb{C}$ the set of complex numbers. Given a positive integer n, we denote by $\mathbb{R}^n$ and $\mathbb{C}^n$ the real and complex n-dimensional vector spaces equipped with the Euclidean norm $\|\cdot\|_2$, defined for $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ (respectively $\mathbb{C}^n$), by

$$\|x\|_2 = \left(|x_1|^2 + \cdots + |x_n|^2\right)^{1/2}. \tag{1.1}$$

Throughout this thesis we will consider (if not declared otherwise) all Banach function spaces over the complex numbers $\mathbb{C}$. Some remarks on the complexification of real Banach spaces can be found in the end of this section.

IMPORTANT FUNCTION SPACES

By $C([0,1])$ we denote the Banach space of continuous functions on $[0,1]$ equipped with the norm of uniform convergence

$$\|f\|_\infty = \sup\{|f(x)| : x \in [0,1]\}.$$

The space of all polynomials of at most degree k is denoted by $\mathcal{P}_k$. For $f \in C([0,1])$, the divided difference $\Delta(x_j, \dots, x_{j+k+1})f$ is defined as the coefficient of x^k in the unique polynomial of degree k or less that interpolates the continuous function f at the points $x_j, \dots, x_{j+k+1}$. We denote the truncated power function of degree k for $x \in \mathbb{R}$ by

$$x_+^k = \begin{cases} x^k, & \text{for } x > 0, \\ 0, & \text{for } x \leq 0. \end{cases}$$

The space of measurable functions on $[0,1]$ is denoted by $L^p([0,1])$ and all functions that are k-times continuously differentiable on $[0,1]$ are denoted by $C^k([0,1])$. A detailed description of these spaces can be found in the book of Rudin [Rud91]. Note that these spaces can also be defined on a compact Hausdorff space K, while for simplicity we will act on the unit interval $[0,1]$. The space of all splines of degree k, i. e., piecewise polynomials of degree k with breakpoints at the knot sequence Δ_n, will be denoted by $\mathcal{S}(\Delta_n, k)$. A short and precise introduction to the spline space $\mathcal{S}(\Delta_n, k)$ is given in Section 3.2. A linear operator T defined on $C([0,1])$ or $L^p([0,1])$ is said to be *positive* if $Tf \geq 0$ holds whenever $f \geq 0$. Positive linear operators can also be defined on more general Banach spaces that are equipped with a partial ordering. For more details for positive linear operators defined on so called Banach lattices, we refer to the book of Schaefer [Sch74].

BANACH SPACES AND LINEAR OPERATORS

The space of bounded linear operators from X to Y is denoted by $\mathcal{L}(X,Y)$ with identity operator I and equipped with the usual operator norm $\|\cdot\|_{op}$. The *range* and the *null space* of $T \in \mathcal{L}(X,Y)$ are denoted by $\operatorname{ran}(T)$ and $\ker(T)$, respectively. We consider the following three operator topologies on $\mathcal{L}(X,Y)$. A sequence of operators $(T_n)_{n\in\mathbb{N}}$ is said to converge *uniformly* towards $T \in \mathcal{L}(X,Y)$ if it converges in the operator norm. The sequence converges *strongly* if $\|T_n x - Tx\|_Y \to 0$ converges for $n \to \infty$ and all $x \in X$. If $|\alpha^*(T_n x) - \alpha^*(Tx)| \to 0$ converges for $n \to \infty$ for all $x \in X$, $\alpha^* \in Y^*$ then the sequence $(T_n)_{n\in\mathbb{N}}$ converges in the *weak* operator topology.

SPECTRAL PROPERTIES

For $T \in \mathcal{L}(X)$, we denote by $\sigma(T)$ the spectrum of T,

$$\sigma(T) = \{\lambda \in \mathbb{C} : T - \lambda I \text{ is not invertible}\}.$$

By $\sigma_p(T)$, we denote the point spectrum of T,

$$\sigma_p(T) = \{\lambda \in \mathbb{C} : T - \lambda I \text{ is not one-to-one}\}$$

which contains all eigenvalues of T. The resolvent set $\rho(T)$ consists of all points λ in the complex plane where $T - \lambda I$ is invertible, $\rho(T) = \mathbb{C} \setminus \sigma(T)$. The open ball of radius $r > 0$ at the point $z \in \mathbb{C}$ in the complex plane will be denoted by $B(z,r) := \{\lambda \in \mathbb{C} : |\lambda - z| < r\}$ and its closure by $\overline{B(z,r)}$.

MODULI OF SMOOTHNESS AND K-FUNCTIONALS

There are different kind of functionals to measure the smoothness of a function. The K-functional introduced by J. Peetre in 1968 measures how well a function can be approximated by functions that are continuously differentiable of a certain order. The K-functional of order r is defined for $f \in C([0,1])$ and $t \geq 0$ as the value of

$$K_r(f,t) := \inf_{g \in C^r([0,1])} \left\{ \|f - g\|_\infty + t \|D^r g\|_\infty \right\},$$

where D^r denotes the differential operator on $C^r([0,1])$ of order r. One important property to note here is that for fixed $t \geq 0$ the K-functional $K_r(\cdot, t)$ is a semi-norm on $C([0,1])$ with $K_r(f,t) \leq \|f\|_\infty$.

Another way to measure the smoothness of a function is using discrete differential operators. Accordingly, the *moduli of smoothness* of order r is defined for $t \geq 0$ and $f \in C([0,1])$ by

$$\omega_r(f,t) := \sup\left\{\|\Delta_h^r f(x)\|_\infty \,:\, 0 < h < t \text{ and } x, x+rh \in [0,1]\right\}, \tag{1.2}$$

where Δ_h^r is the forward difference operator of order r,

$$\Delta_h^r f(x) := \sum_{l=0}^{r} (-1)^{r-l} \binom{r}{l} f(x+lh).$$

More properties of moduli of smoothness and K-functionals and relations between them can be found in Butzer and Berens [BB67], Johnen [Joh72], and Johnen and Scherer [JS77]. In a similar way, the modulus of smoothness can also be defined on the L^P-spaces using the p-Norm instead of the norm of uniform convergence. In this case, we denote the *integral modulus of smoothness* by $\omega_{r,p}(f,t)$. More information on moduli of smoothness and K-functionals are given in Chapter 10.

ON THE COMPLEXIFICATION OF REAL BANACH SPACES

Note that the problem to compute the curvature mentioned previously is naturally modeled on a real Banach function space X. This case will treated by its complexification $X_\mathbb{C} = X \oplus iX$ equipped with the norm

$$\|f + ig\|_\mathbb{C} = \sup_{0 \leq \varphi \leq 2\pi} \|f\cos\varphi + g\sin\varphi\|, \qquad f, g \in X.$$

Then $X_\mathbb{C}$ is a Banach space and X is continuously embedded in $X_\mathbb{C}$. The corresponding complex extension $T_\mathbb{C}$ of T is defined for all $f, g \in X$ by

$$T_\mathbb{C}(f + ig) = Tf + iTg.$$

In this way, the operator norm is consistent, i.e., $\|T\|_{op} = \|T_\mathbb{C}\|_{op}$ holds. The spectrum $\sigma(T)$ of T is defined as $\sigma(T_\mathbb{C})$. Note that the set $\sigma_p(T_\mathbb{C}) \cap \mathbb{R}$ consists of the eigenvalues of $T : X \to X$. For more details on the complexification of real Banach spaces we refer to [Rus86, pp. 7–16] and [MST99].

1.4 Related Research

There also exist many methods that interpolate discretely given data points in a shape preserving way, i. e., they preserve the positivity, monotonicity, and convexity defined by the discrete data. While most methods can construct a continuous or C^1- interpolant, only a few of them are able to interpolate with a smooth C^2-function. Note that we are interested in estimating the curvature, hence C^2-smoothness is essential here. Best known are the so called tension methods. Thereby, a smooth function is constructed as a collection of patches that depends on a set of tension parameters. The main problem here is the estimation of these parameters in order to preserve the shape of discretely given data. While local methods also require values of the derivatives of the function to interpolate, global methods rely only on function values.

Lamberti and Manni [LM01] have proposed a global tension method to obtain a C^2-shape preserving interpolant. This scheme requires the solution of a tridiagonal, diagonally dominant linear system as well as a parameter selection. The selection of the tension parameters is based on an automatic iterative algorithm depending on conditions on the monotonicity and the convexity of the curve. The main drawback of this method is that no explicit analytic expression of the final interpolant is available. Thus, computing its derivatives as we need to calculate the curvature is not possible. Recently, another approach has been developed by Goodman and Ong [GO05] using a vector subdivision scheme with cubic splines, while the shape preserving property holds only for the limiting function. An overview of various shape preserving interpolation methods has been given in the surveys of Goodman [Goo02] and Pan and Cheng [PC12].

Finally, we also want to mention other shape preserving approximation methods. Recently, Kong and Ong [KO09] have constructed a shape preserving approximation using spatial cubic splines. Shape preserving approximations based on Bernstein-type operators with fixed polynomials have been considered by Cárdenas-Morales, Garrancho, and Muñoz-Delgado [CGM06].

Before we conclude this chapter, we want to highlight the advantages of our chosen shape preserving approximation method based on Schoenberg's variation diminishing splines. As the spline operator is linear, the approximation is fast to compute and does not require to solve a linear system. Another advantage are the available explicit representations of the derivatives based on the coefficients of the operator. This allows for a fast and stable evaluation of the derivatives which is crucial to compute the curvature. Besides, we can utilize the approximation error to detect singularities of piecewise smooth curves.

PART ONE

CURVATURE ESTIMATION

CHAPTER 2 Variation Diminishing Transforms and Total Positivity

"If there is a problem you can't solve, then there is an easier problem you can solve: find it."

GEORG PÓLYA

THE FIRST ARTICLE DESCRIBING variation diminishing transforms has already been published in 1930 by I. J. Schoenberg. Given a matrix $A \in \mathbb{R}^{m \times n}$, the linear map $x \mapsto Ax$ is said to be variation diminishing if for a real vector $x \in \mathbb{R}^n$ the image $y = Ax$ has no more sign changes than the given vector x. This remarkable property is completely characterized when of all of the minors of the matrix A have the same sign. Matrices where all minors are non-negative are said to be *totally positive*. This remarkable concept has been extended on functions spaces where an approximation has less sign changes than the function to approximate. We will use this kind of transforms for the approximation of a curve to ensure that the approximation introduces no additional oscillations. For more details on total positivity we refer to the comprehensive monograph of S. Karlin [Kar68] which is influenced by I. J. Schoenberg.

Around the same time as I. J. Schoenberg (1933-1937), Gantmacher and Krein have also extensively discussed properties of totally positive matrices based on their research on vibration of mechanical systems. An extensive overview of their work can be found in the book [GK02]. In contrast to I. J. Schoenberg, they were motivated by the Perron-Frobenius Theorem for positive matrices to study the behaviour of the eigenvalues and eigenvectors of totally positive matrices. In particular, they have proved that strictly total positive matrices, i. e.matrices where all minors are positive, have distinct positive eigenvalues. Besides, they revealed an oscillatory structure of the eigenvectors. In their work they could prove these properties for a broader class of matrices, the so called *oscillatory* matrices.

We start with the precise definition in the discrete setting according to Schoenberg [Sch30] and show relations to so called sign-regular matrices. In the next section the concept of variation diminution will be specialized to positivity preserving transforms. These transformes are said to be totally positive. The work of [GK37] complements the results of I. J. Schoenberg by spectral properties. We conclude with approximation operators that preserve the shape of the function to approximate. It will be shown that variation diminishing operators which can reproduce linear functions in fact leave the positivity, monotonicity and the convexity of functions unchanged. This will serve as the framework of our choice to approximate the curvature of a curve where only discrete data points are provided.

2.1 The Variation Diminishing Property

Definition 2.1 (Variation diminishing matrix)**.** Let $v : \mathbb{R}^n \to \mathbb{R}$ count the number of sign changes of a vector. The matrix $A \in \mathbb{R}^{m\times n}$ is *variation diminishing*, if for all $x \in \mathbb{R}^n$

$$v(Ax) \leq v(x).$$

Let us consider first two examples in $\mathbb{R}^2$.

Example 2.1. The identity matrix $E_2 = \left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)$ is obviously variation diminishing. Contrarily, the matrix $A = \left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$ is not variation diminishing, as

$$\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}1\\ 1\end{pmatrix} = \begin{pmatrix}1\\ -1\end{pmatrix},$$

and $1 = v(Ax) \geq v(x) = 0$ for $x = (1,1)^T$.

For general matrices, the criterion of variation diminishing matrices used in the definition is usually hard to check, as this property has to be shown for all $x \in \mathbb{R}^n$. Another way to characterize the variation diminishing property are sign-regular matrices. We will define them next and show their relation to variation diminishing matrices.

Definition 2.2 (Sign-regular matrix)**.** A matrix $A \in \mathbb{R}^{m\times n}$ is called *sign-regular* if all minors of A have the same sign.

Example 2.2. The matrix $A = \left(\begin{smallmatrix}1 & 2\\ 0 & 1\end{smallmatrix}\right)$ is sign-regular as the minors $\det(1) = 1$, $\det(2)$, $\det(0)$, $\det(1)$ and $\det(A) = 1$ have the same sign (non-negative). In contrast, the matrix $A = \left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$ from Example 2.1 is not sign-regular.

According to Schoenberg, the following relation between sign-regular matrices and variation diminishing matrices holds:

Proposition 2.1 ([Sch30, see Satz 1 and 2])**.** *Let $A \in \mathbb{R}^{m\times n}$ with rank less or equal to n. Then the following two statements hold:*

1. *If A is sign-regular, then A is variation diminishing.*
2. *If A has rank n and A is variation diminishing, then A is sign-regular.*

The next corollary considers as special case quadratic full-rank matrices. If one combines both statements of Proposition 2.1 it immediately follows that being variation diminishing is equivalent to being sign-regular.

Corollary 2.2. *A non-singular matrix $A \in \mathbb{R}^{n\times n}$ is variation diminishing if and only if it is sign regular.*

The preceding corollary states in particular, that invertible matrices are variation diminishing if and only if all minors have the same sign. An equivalent property of general matrices has been shown by Motzkin [Mot36].

Note that it is still possible that variation diminishing matrices do not preserve the positivity of vectors.

Example 2.3. Consider the non-singular matrix $A = \left(\begin{smallmatrix} -1 & -1 \\ -2 & -1 \end{smallmatrix}\right)$. All minors are negative and hence, A is sign-regular and by Corollary 2.2 it is also variation diminishing. But the image of the positive vector $(1,2)^T$ under A yields the negative vector $(-3,-4)$. Thus, positivity is not preserved under A.

Accordingly, the concept of variation diminution will be extended to the concept of total positivity to overcome these shortcomings.

2.2 Total Positivity

In addition to the variation diminution, I. J. Schoenberg introduced the concept of total positivity in [Sch30] to describe positivity preserving transforms. Here, the non-singular real matrix is said to be *totally positive* if all of its minors, of any order, are nonnegative. If all minors are positive, then the matrix is called *strictly totally positive*. Note that in recent articles these matrices are often called totally non-negative to distinguish between strongly totally positive matrices and totally positive matrices where all minors are strictly positive. Here, we will stick to the notation which has been introduced first by I. J. Schoenberg. The reader will find more information in the gentle introduction to total positivity by T. Ando [And87], whereas even more results can be found in the recent book of A. Pinkus [Pin10]. A fundamental overview about total positivity in a variety of fields can be found in the monograph of S. Karlin [Kar68] where also applications to the theory of summability, interpolation problems and differential equations are presented.

Example 2.4. The matrix $\left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ is totally positive and also strictly totally positive, whereas the matrix $\left(\begin{smallmatrix} 2 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ is only totally positive.

An important relation between the class of strictly totally positive matrices and totally positive matrices has been shown by Whitney [Whi52]. He characterized strictly totally positive matrices as dense set in the set of totally positive matrices.

Proposition 2.3 ([Whi52, Thm. 1])**.** *Every totally positive matrix can be approximated arbitrarily closely by strictly totally positive matrices.*

During their research research on vibration of mechanical systems, Gantmacher and Krein [GK02] have also investigated properties of totally positive matrices. For the original research article, we refer to Gantmakher and Krein [GK37]. Motivated by the Perron-Frobenius Theorem they studied a generalization of strictly totally positive matrices, the so called *oscillatory* matrices. They have shown that these matrices have distinct positive eigenvalues and revealed an oscillatory structure of the eigenvectors. We will state these results in the following.

Definition 2.3 (Oscillatory matrix). A matrix $A \in \mathbb{R}^{n \times n}$ is called *oscillatory* if A is totally positive and there exists a positive integer k such that A^k is strictly totally positive.

Proposition 2.4 (Gantmacher and Krein [GK02, Thm. 6, p.87]). *Let $A \in \mathbb{R}^{n \times n}$ be an oscillatory matrix. Then the eigenvalues of A are n distinct positive real numbers, i.e., $\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0$. Besides, let v be an eigenvector for the k-th largest eigenvalue. Then v has exactly $k - 1$ sign variations.*

Gantmacher and Krein have also established a sufficient and necessary criterion for totally positive matrices to be oscillatory. This criterion characterizes oscillatory matrices as non-singular matrices, where the first diagonal below the main diagonal, and the first diagonal above the main diagonal are positive.

Proposition 2.5 ([GK02, Thm. 10, p.100]). *A totally positive matrix $A \in \mathbb{R}^{n \times n}$ is called* oscillatory *if and only if*

(i) A is non-singular and

(ii) $a_{i,i+1} > 0$ and $a_{i+1,i} > 0$ for $i \in \{1, \dots, n-1\}$.

This criterion will be of importance in the end of the next chapter to prove spectral properties of the variation-diminishing spline operator devised by Schoenberg.

Using the spectral properties of oscillatory matrices shown in Proposition 2.4 and the density of strictly totally positive matrices stated in Proposition 2.3, we can show now that totally positive matrices share the same spectral properties as symmetric, positive definite matrices.

Proposition 2.6 ([Pin10, Cor. 5.5]). *If A is a totally positive Matrix, then all the eigenvalues of A are non-negative, real numbers.*

Example 2.5 (Totally positive matrices with row sum 1). In the next chapters, we will deal with totally positive matrices that arise from functional evaluations of certain positive basis functions like polynomials or splines. In order to be able to reconstruct constants, a necessary condition on the basis functions is that they form a partition of unity. In this case, 1 is always the largest eigenvalue corresponding to the eigenvector that contains only ones. The eigenvalues are non-negative, real numbers and can be ordered as

$$1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n = 0.$$

2.3 Totally positive transforms

In the following, we will extend the concepts of total positivity to functions. It will be shown that the total positivity of a kernel guarantees the variation diminishing property for the corresponding integral transform operator. Moreover, shape preserving properties can be characterized by the kernel. For the convenience of the reader we state in this section briefly the most important results of the comprehensive book of Karlin [Kar68] that we need in our context on shape preserving approximations.

Suppose X and Y as intervals or as sets of positive integers and let $K : X \times Y \to \mathbb{R}$ be a real function. In the current section, we will discuss transforms of the form

$$T(f;x) = \int_Y K(x,y)f(y)\mathrm{d}\mu(y), \qquad x \in X, \tag{2.1}$$

where μ is a sigma-finite measure on Y. During this section, we use the Bernstein operator to illustrate the shown results. Here, $X = [0,1]$, $Y = \{1,\ldots,n\}$ and the kernel $K : X \times Y$ is given by the Bernstein polynomials of degree n:

$$K_n(x,j) = \binom{n}{j} x^j (1-x)^{n-j}, \qquad j = 0,\ldots,n.$$

In this case, (2.1) reduces to the classical Bernstein operator

$$B_n(f;x) = \sum_{j=0}^{n} f\left(\frac{j}{n}\right)\binom{n}{j} x^j (1-x)^{n-j}, \qquad x \in [0,1],$$

that approximates continuous functions by their uniform point evaluations. Originally, this kind of approximation has been used in by S. N. Bernstein [Ber12] to prove the Weierstrass approximation theorem in a short and elegant way. More details on this operator and on Schoenberg's variation diminishing spline operator are provided in the next chapter.

2.3.1 Totally positive kernels

Motivated by the integral transform (2.1) we extend the concept of total positivity to integral kernels. An example is given afterwards.

Definition 2.4. A kernel $K : X \times Y \to \mathbb{R}$ over linearly ordered sets X and Y is said to be *totally positive*, if for all integers m and ordered selections $x_1 < x_2 < \cdots < x_m$, and $y_1 < y_2 < \cdots < y_m$ in X and Y, respectively, the inequalities

$$K\begin{pmatrix} x_1 & x_2 & \cdots & x_m \\ y_1 & y_2 & \cdots & y_m \end{pmatrix} := \det \begin{pmatrix} K(x_1,y_1) & K(x_1,y_2) & \cdots & K(x_1,y_m) \\ K(x_2,y_1) & K(x_2,y_2) & \cdots & K(x_2,y_m) \\ \cdots & \cdots & & \cdots \\ K(x_m,y_1) & K(x_m,y_1) & \cdots & K(x_m,y_m) \end{pmatrix} \geq 0$$

hold true. In the case where strict inequality holds, we call K *strictly totally positive*.

In particular, $K(x,y) \geq 0$ holds for all $(x,y) \in X \times Y$. X or Y might be considered as intervals or as sets of positive integers. If both sets are finite, then $K(x,y)$ can be considered as a matrix and the notion of total positivity is understood in the sense of the preceding section.

Example 2.6 (Bernstein basis, see [Kar68, p. 287])**.** For any fixed integer n and all sequences

$$0 \le x_1 < x_2 \cdots < x_n \le 1$$

the matrix

$$\begin{pmatrix} p_{n,1}(x_1) & p_{n,1}(x_2) & \cdots & p_{n,1}(x_n) \\ p_{n,2}(x_1) & p_{n,2}(x_2) & \cdots & p_{n,2}(x_n) \\ \cdots & \cdots & & \cdots \\ p_{n,n}(x_1) & p_{n,n}(x_2) & \cdots & p_{n,n}(x_n) \end{pmatrix}$$

is totally positive, where $p_{n,j}$ are the Bernstein polynomials defined by

$$p_{n,j}(x) = \binom{n}{j} x^j (1-x)^{n-j}, \qquad j = 0, \dots, n.$$

2.3.2 Variation diminishing transforms

Let $K(x, y)$ be a totally positive kernel on $X \times Y$. We consider now the integral transform

$$T(f; x) = \int_Y K(x, y) f(y) d\mu(y), \tag{2.2}$$

where μ is a sigma-finite measure on Y. Besides, we assume that $K(x, y)$ has the necessary growth properties to ensure that this integral (2.2) exists for every bounded Borel-measurable function f defined on Y.

Clearly, T is a linear operator. Note that if Y is a set of positive integers, then T is a discrete operator. In the following, we will discuss under which conditions on the integration kernel $K(x, y)$ the operator T is variation diminishing and which assumptions T is able to preserve the monotonicity and convexity of functions.

The concept of variation diminution has been further generalized to approximation operators [Sch48; Sch50a; Sch59]. For a real-valued function f let us denote by $\nu(f)$ the number of sign changes in $[0, 1]$ of the function f,

$$\nu(f) := \sup_{\mathcal{P}_n, n \in \mathbb{N}} \nu(f(x_1), f(x_2), \dots, f(x_n)),$$

where the supremum is taken over all partitions

$$\mathcal{P}_n = \{x_1, \dots, x_n \in [0, 1] \,:\, 0 \le x_1 < x_2 < \cdots < x_n \le 1\}.$$

Then we can define an operator T to be variation diminishing in the following sense.

Definition 2.5. A linear operator T is called *variation diminishing* if

$$\nu(Tf) \leq \nu(f)$$

holds for all real-valued functions f.

Hence, if T is variation diminishing the approximation Tf has less sign changes than the original function f. Note that we use here the original notation introduced by I. Schoenberg, while other authors use the symbol S^- to measure the number of sign changes with discarded zero terms.

The following result of S. Karlin relates the concept of total positivity with the variation diminishing property.

Proposition 2.7 ([Kar68, Thm. 3.1 on p. 285]). *If $K : X \times Y \to \mathbb{R}$ is totally positive kernel, then T defined as in* (2.2) *has the variation diminishing property.*

Example 2.7 (Bernstein operator). Given a fixed integer n, the Bernstein polynomials are defined by

$$K_n(j,x) = \binom{n}{j} x^j (1-x)^{n-j}, \qquad j = 0, \ldots, n.$$

Using this kernel, the integral transform (2.1) reduces to the the Bernstein operator $B_n : C([0,1]) \to C([0,1])$ defined by

$$B_n(f;x) = \sum_{j=0}^{n} f\left(\frac{j}{n}\right) \binom{n}{j} x^j (1-x)^{n-j}, \qquad x \in [0,1].$$

By Example 2.6 B_n is variation diminishing.

2.3.3 Shape preserving transforms

If K is a totally positive kernel, i. e., T is variation diminishing, then the integral transform T has the ability to preserve the monotonicity and the convexity of functions, provided that T can reproduce constants and linear functions.

Proposition 2.8 ([Kar68, Thm. 3.4(a) on p. 285]). *If $K : X \times Y \to \mathbb{R}$ is a totally positive kernel and T is able to reproduce constants, i. e.,*

$$\int_Y K(x,y)\,d\mu(y) = 1, \quad x \in X.$$

Then the integral transform (2.1) *maps increasing functions into increasing functions.*

While the total positivity of the kernel combined with the ability to reconstruct constants leads to a monotonicity preserving transform, the additional ability to reproduce linear functions leads to a convexity preserving transform.

Proposition 2.9 ([Kar68, Thm. 3.5(a) on p. 285]). *If $K : X \times Y \to \mathbb{R}$ is a totally positive kernel and T is able to reconstruct constants and maps linear functions into linear functions, i. e.*

$$\int_Y K(x,y)d\mu(y) = 1, \qquad \int_Y K(x,y)yd\mu(y) = ax + b, \quad x \in X,$$

where a and b are real numbers and $a > 0$. Then the integral transform (2.1) *maps convex functions into convex functions.*

Example 2.8 (Bernstein operator, [Kar68, pp. 287]). Consider the Bernstein operator of Example 2.7,

$$B_n(f;x) = \sum_{j=0}^{n} f\left(\frac{j}{n}\right)\binom{n}{j}x^j(1-x)^{n-j}.$$

It has been shown Karlin [Kar68, pp. 287] that $B_n(f;x)$ is convex whenever f is convex. Moreover, $B_n(f;x) \geq f(x)$ holds for every continuous convex function f, see also Popoviciu [Pop38].

2.3.4 Shape preserving discrete transforms

To estimate the curvature in a variation diminishing way, we need an approximation method that is able to preserve the convexity as seen in the last section. Furthermore, we require an approximation method that approximates a curve given only at finitely many points on the curve. In the following, we will consider discrete approximation operators of this kind based on the integral transform (2.1). Operators of this kind are, e. g., the Bernstein operator or Schoenberg's spline operator. Both operators are discussed in the next chapter. We will show as seen in Proposition 2.9 that a general discrete operator leads to convexity preserving transform provided that the basis functions are totally positive and the operator is able to reproduce constants and linear function.

As we only consider here discrete kernels here, we will first introduce the notion of a collocation matrix in analogy to Definition 2.4 and state the most important theorems from the last sections once more in the discrete setting. Thereby, we will restrict the approximation of real-valued function onto the unit interval $[0,1]$. Finally, we will extend this discrete approximation operator to the approximation of planar curves defined on $[0,1]$. This section is along the lines of Goodman [Goo96].

Definition 2.6 (collocation matrix). Consider a sequence of real valued functions $(e_1,\dots,e_n)$. The *collocation matrix* for the points $0 \leq x_1 < x_2 < \cdots < x_n \leq 1$ is then defined as the matrix

$$M\begin{pmatrix} e_1 & e_2 & \cdots & e_n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} := \begin{pmatrix} e_1(x_1) & e_2(x_1) & \cdots & e_n(x_1) \\ e_1(x_2) & e_2(x_2) & \cdots & e_n(x_2) \\ \vdots & \vdots & & \vdots \\ e_1(x_n) & e_2(x_n) & \cdots & e_n(x_n) \end{pmatrix}.$$

The functions $e_1, \dots, e_n$ are in the following assumed to be linearly independent. Therefore, they build a basis for the n-dimensional subspace they span.

Definition 2.7 (Totally positive bases). A sequence $(e_1, \dots, e_n)$ of real valued functions over $[0,1]$ is said to be *totally positive* if for any points $0 \leq x_0 < \dots < x_m \leq 1$, the collocation matrix $M\left(\begin{smallmatrix} e_1 & e_2 & \cdots & e_n \\ x_1 & x_2 & \cdots & x_n \end{smallmatrix}\right)$ is totally positive. The sequence $(e_1, \dots, e_n)$ is *normalized totally positive* if, in addition,

$$\sum_{j=0}^{n} e_j(x) = 1, \qquad x \in [0,1] .$$

Consider, for instance, the Bernstein polynomials as shown in Example 2.6.

Based on the total positivity of the bases, we obtain the following relation to the variation diminishing property by Proposition 2.7.

Proposition 2.10 ([Goo96, Thm. 3.1]). *If $(e_1, \dots, e_n)$ are totally positive on $[0,1]$, then for any $\lambda_1, \dots, \lambda_n$ we have*

$$v(\lambda_1 e_1 + \cdots + \lambda_n e_n) \leq v(\lambda_1, \dots, \lambda_n).$$

Next, we will consider a linear operator that yields a linear combination of the basis $(e_1, \dots, e_n)$ in order to provide a variation diminishing operator. Therefore, we will assume to the rest of this section that $(e_1, \dots, e_n)$ is a normalized totally positive basis. Additionally, we assume that the basis has the property to approximate straight lines in the sense

$$x = \sum_{j=1}^{n} x_j e_j(x), \qquad x \in [0,1] .$$

We will now concretely consider the discrete approximation operator T that is defined for every real-valued function f on $[0,1]$ as

$$T(f;x) = \sum_{j=1}^{n} f(x_j) e_j(x), \qquad x \in [0,1] . \tag{2.3}$$

The next proposition characterizes the operator (2.3) as positive linear operator that is able to interpolate at the endpoints of the interval.

Proposition 2.11 ([Goo96, Thm. 3.3]). *The operator T in* (2.3) *defines a positive linear operator which reproduces linear functions. Besides, the operator T interpolates at the endpoints of* $[0,1]$*:*

$$Tf(0) = f(0) \quad \textit{and} \; Tf(1) = f(1).$$

By Proposition 2.8 and Proposition 2.9 mentioned in the last section, we obtain the following two corollaries:

Corollary 2.12. *If the function f is increasing (decreasing), then Tf is increasing (decreasing).*

Corollary 2.13. *If the function f is convex, then Tf is convex and $Tf \geq f$.*

Next, we try to generalize these properties to the approximation of curves. Recall here that convexity of planar curves can be defined as curves where any straight line has at most two intersection points. In this sense, the next proposition counts how often any straight line will cross f and Tf by the variation diminishing property:

Proposition 2.14. *For any linear function ℓ and any real-valued function f on $[0,1]$*

$$v(Tf - \ell) \leq v(f - \ell)$$

holds.

Using this property, we can consider now the approximation of curves in the following way and show a relation between the curve and the polygonal arc defined by discretely many points on the curve.

Proposition 2.15 ([Goo96, Thm. 3.2]). *Let $(e_1, \dots, e_n)$ be a normalized totally positive basis on $[0,1]$. For given points $(x_i, y_i) \in \mathbb{R}^2$ we consider the curve*

$$\alpha(t) = (x(t), y(t)) = \sum_{j=1}^{n} \begin{pmatrix} x_j \\ y_j \end{pmatrix} e_j(t), \qquad t \in [0,1]. \tag{2.4}$$

Then the number of times the curve $\alpha : I \to \mathbb{R}^2$ crosses any straight line ℓ is bounded by the number of times the polygonal arc through the points (x_j, y_j) crosses the line ℓ.

Based on this important property, we obtain the following shape preserving results:

Corollary 2.16. *If the polygonal arc through the given points (x_i, y_i) is monotonic in a given direction, then so is the curve α defined by* (2.4).

Corollary 2.17. *If the polygonal arc through the points (x_i, y_i) is convex, then so is the curve α defined by* (2.4).

To approximate the curvature of a planar curve, these two properties guarantee that the sign changes of curvature are only defined by behaviour of its polygonal arc. Furthermore, if the basis functions are locally supported, then the shape preserving properties can be considered not only globally but also locally. Thus, local monotonicity is preserved as well as local convexity in a weak sense.

2.4 Generalized discrete approximation

For the estimation of the curvature we are interested in a shape preserving approximation of a digitized curve. It was pointed out in the preceding sections, that the total positivity of a kernel or a basis is the fundamental concept for shape preserving approximations. Accordingly, we consider here a normalized totally positive basis $(e_1, \dots, e_n)$.

The integral transforms discussed in the last sections are only able to evaluate a function at certain values. We define here a generalized approximation operator T that is defined for every real-valued function f on $[0,1]$ as

$$Tf = \sum_{j=1}^{n} \alpha_j^*(f) e_j, \tag{2.5}$$

where α_k^* are positive linear functionals satisfying $\alpha_k^*(1) = 1$ and $\alpha_k^*(e_k) > 0$ for $k \in \{1, \dots, n\}$. The linear functionals are able to represent different linear digitization schemes. Note that the positivity of the functionals is crucial to assure the positivity preservation of the function f. Also note that the condition $\alpha_k^*(e_k) > 0$ can be seen as a localization property. Operators of kind (2.5) are discussed in the next chapter. We will show here only two examples of possible functionals.

Example 2.9 (Point evaluation). The Riesz representation theorem gives a characterisation of positive linear functionals on $C([0,1])$. Namely, for every positive linear functional $a^* : C([0,1]) \to \mathbb{R}$, there is a unique positive Radon measure ν such that

$$\alpha^*(f) = \int_0^1 f \mathrm{d}\nu \qquad \text{for every } f \in C([0,1]).$$

A classical example of a positive linear functional on $C([0,1])$ is the Dirac measure at a point $x \in [0,1]$ defined for $f \in C([0,1])$ by

$$\delta_x(f) = f(x).$$

Given a partition $\Delta_n = \{x_k\}_{k=1}^n$ of $[0,1]$ satisfying

$$0 = x_1 < x_2 < \dots < x_n = 1,$$

then a popular choice for the functionals are $\alpha_k^* = \delta_{x_k}$ for $k \in \{1, \dots, n\}$. In this case, the positive finite-rank operator can be written for $x \in [0,1]$ as

$$Tf = \sum_{k=1}^{n} f(x_j) e_k.$$

If the digitized values are modeled as local averages we can consider the following functionals on the space of integrable functions.

Example 2.10 (Local Averaging). Given a partition $\Delta_n = \{x_k\}_{k=1}^n$ of $[0,1]$ satisfying

$$0 = x_1 < x_2 < \dots < x_n = 1,$$

then a local averaging is provided by the functionals $\alpha_k^* = n\int_{\frac{k-1}{n}}^{\frac{k}{n}} \cdot \, \mathrm{d}t$ for $k \in \{1, \ldots, n\}$. In this case, the positive finite-rank operator can be written for $x \in [0,1]$ as

$$Tf = \sum_{k=1}^{n} \left(n \int_{\frac{k-1}{n}}^{\frac{k}{n}} f(t)\mathrm{d}t \right) e_k.$$

We have seen in the last section that this operator is shape preserving if additionally

$$Tx = \sum_{i=1}^{n} \alpha_k^*(x) e_j = x$$

holds.

CHAPTER 3 Variation Diminishing Operators

"Intuition is the undoubting conception of a pure and attentive mind,
which arises from the light of reason alone,
and is more certain than deduction."

RENÉ DESCARTES

WE INTRODUCE HERE two variation-diminishing operators on the space of continuous functions, the Bernstein operator and Schoenberg's spline operator, as well as their generalizations to the L^p-spaces. The Bernstein operator samples a continuous functions uniformly, whereas the Schoenberg operator can sample a continuous function at predefined knots. Other advantages to use spline approximations over Bernstein polynomials are the faster convergence rate and their local flexibility. While the Bernstein polynomials have their support on the whole interval $[0,1]$, the spline basis functions have support in small subintervals of $[0,1]$ whose size depends on the degree k. Thus, a local change of the function has a global influence for its approximation with Bernstein polynomials while the spline approximation only changes locally. Nevertheless, both operators can be used to estimate the curvature of discretely given functions values. The Bernstein operator preserves convexity of all orders, while the Schoenberg operator can only preserve positivity, monotonicity, and the ordinary convexity. In contrast to the point evaluations on the space of continuous functions, the generalized operators on the L^p-spaces use local integrals to evaluate measurable functions. While the operators on the continuous function space are able to reproduce constant and linear functions, the operators acting on the space of measurable functions are only able to reproduce constant functions.

The focus on this chapter lies on spline approximations based on Schoenberg's variation diminishing operator due to the previously mentioned advances. As highlight, we will prove that this operator has only non-negative distinct real eigenvalues. For this purpose, we prove that the Gramian matrix of the corresponding integral Schoenberg operator is in fact an oscillatory matrix. Finally, we show a relation between those operators to transfer these result to the Schoenberg operator on $C([0,1])$. Note that up to our knowledge a concrete formula for the eigenvalues is still missing.

3.1 Operators based on Bernstein polynomials

The Bernstein operator and the polynomials have been applied in a vast range of applications due to its simplicity and its elegant properties. Consider, e. g., the Bezier-curves as their extension to higher dimensions which are the leading concept to model curves in CAGD. To get an overview, we will state here their most important properties. The Bernstein operator as well as the other operators introduced in the following will serve as examples throughout this thesis.

For a given integer $n > 0$ the Bernstein polynomials are defined on $[0,1]$ by

$$p_{n,k}(x) := \binom{n}{k} x^k (1-x)^{n-k}, \qquad k \in \{0,\dots,n\}.$$

An illustration for $n = 8$ is shown in Figure 3.1. These polynomials have been first constructed by S. N. Bernstein [Ber12] in 1912 to prove the Weierstrass approximation theorem in a short and elegant way. Concretely, he as shown that each $f \in C([0,1])$ can be uniformly approximated in the following way:

$$\left\| f(x) - \sum_{k=0}^{n} f\left(\frac{k}{n}\right) p_{n,k}(x) \right\|_\infty \to 0, \quad \text{for } n \to \infty. \tag{3.1}$$

It is known that these functions form a partition of unity,

$$\sum_{j=0}^{n} p_{n,j}(x) = 1,$$

and they build a basis of the space of polynomials up to degree n. In the following we will discuss two operators based on these Bernstein polynomials. The classical Bernstein operator on $C([0,1])$ used in (3.1) and its extension to the $L^p([0,1])$-spaces, the Kantorovič operator. For more properties on these polynomials and corresponding operators, we refer to the comprehensive textbook of Lorentz [Lor86].

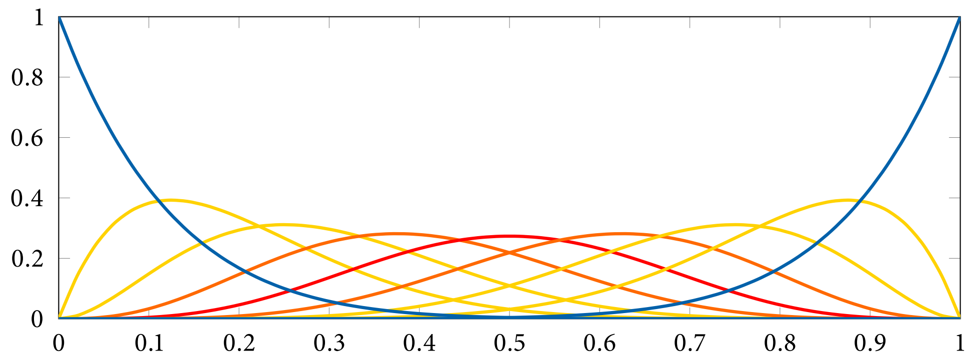


Figure 3.1: *The Bernstein polynomials for $n = 8$. Note their global support on* $[0,1]$.

3.1.1 The Bernstein operator

Given an integer $n > 0$, the *Bernstein operator* $B_n : C([0,1]) \to C([0,1])$ is defined for $f \in C([0,1])$ by

$$B_n(f;x) = \sum_{j=0}^{n} f\left(\frac{j}{n}\right)\binom{n}{j}x^j(1-x)^{n-j} = \sum_{j=0}^{n} f\left(\frac{j}{n}\right)p_{n,j}(x)$$

for all $x \in [0,1]$.

As already mentioned above, the Bernstein operator is able to reproduce constant functions due to the partition of unity. Particularly, the Bernstein operator can reproduce all linear functions as also $B_n x = x$ holds, i. e.,

$$\sum_{j=0}^{n} \frac{j}{n} p_{n,j}(x) = x, \qquad x \in [0,1],$$

and $\{1,x\}$ constitutes a basis of the linear functions. Moreover, all polynomials of degree $k \le n$ are preserved in the sense that $B_n x^k$ is always polynomial of degree k, see Lorentz [Lor86]. Due to this property [Kar68, p.286], the next proposition states the shape preserving properties of the Bernstein polynomials.

Proposition 3.1. *The Bernstein operator preserves convexity of all orders, i. e.,*

$$D^k B_n f \ge 0 \quad \text{if} \quad D^k f \ge 0, \qquad f \in C^k([0,1]).$$

Lorentz has shown that the derivatives of all orders of $B_n f$ converge uniformly to the corresponding derivative of f.

Proposition 3.2 (Lorentz [Lor37]). *If $f \in C^r([0,1])$ for some integer $r \ge 0$, then*

$$\lim_{n\to\infty} \|D^r B_n f - D^r f\|_\infty \to 0.$$

We now consider estimates for the approximation error. An upper estimate for the approximation error by the second order modulus of smoothness has first been shown by Brudnyi [Bru65] in 1962:

Proposition 3.3. *For a given integer n and $f \in C([0,1])$, there exists a constant C independent of f and n such that*

$$\|B_n f - f\|_\infty \le C \cdot \omega_2(f; \frac{1}{\sqrt{n}}).$$

Note that the second order modulus of smoothness reflects the approximation behaviour of the Bernstein operator B_n, as ω_2 annihilates linear functions as does the operator $B_n - I$. Thus, estimates using the second order modulus of smoothness is the right tool here. Inverse theorems have been considered by Berens and Lorentz [BL72], while related lower estimates have been shown only for the so called Ditzian-Totik modulus of smoothness which is a generalization of the classical modulus of smoothness. For a detailed discussion we refer to the book of Ditzian and Totik [DT87], we state here only its definition to show the lower estimates.

Definition 3.1 (Ditzian-Totik modulus of smoothness). The *Ditzian-Totik modulus of smoothness of order r* is given by

$$\omega_r^{\varphi}(f,t) := \sup\left\{ \left\| \Delta_{h\varphi(x)}^r f(x) \right\|_\infty \,:\, 0 < h < t \text{ and } x, x+rh \in [0,1] \right\},$$

where, if not stated otherwise, $\varphi(x) = \sqrt{x(1-x)}$. Here $\Delta_{h\varphi(x)}^r$ denotes the r-th forward difference operator with step $h\varphi(x)$ depending on x.

Clearly, if $\varphi(x) = 1$, then the ω_r^{φ} reduces to the usual modulus of smoothness as defined on page 8. Using this generalized modulus of smoothness, the following results has been shown.

Proposition 3.4 (Ditzian and Totik [DT87], Knoop and Zhou [KZ94]). *For any integer $n > 0$ and given $f \in C([0,1])$ there exists constants $C_1, C_2 > 0$ independent on f and n such that*

$$C_1\omega_r^{\varphi}(f, \frac{1}{\sqrt{n}}) \le \|B_n f - f\|_\infty \le C_2\omega_r^{\varphi}(f, \frac{1}{\sqrt{n}}).$$

The eigenvalues of the Bernstein operator have been revealed by the Russian author Călugăreanu. A comprehensive discussion on the corresponding eigenfunctions can be found in the work of Cooper and Waldron [CW00].

Proposition 3.5 (Călugăreanu [Căl66]). *The Bernstein operator B_n has only real eigenvalues, namely*

$$\lambda_k^{(n)} = \frac{n!}{(n-k)!}\frac{1}{n^k}, \qquad k \in \{0,1,\ldots,n\}.$$

The iterates of B_n have been also extensively discussed. The work of Kelisky and Rivlin [KR67], Nielson, Riesenfeld, and Weiss [NRW76], da Silva [dSil84], and Wenz [Wen97] have proved the limiting operator based on the theory of stochastic matrices. Karlin and Ziegler [KZ70] and Nagel [Nag80] have used theorems Korovkin-type. The authors Agratini and Rus [AR03] and Rus [Rus04] have provided an alternative way using a contraction principle. The limiting behaviour of the iterates by methods of the theory of operator semigroups has also studied by Micchelli [Mic73]. The iterates of the Bernstein operators have been recently revisited by Badea [Bad09] using methods from functional analysis and operator theory. Accordingly, the next proposition is due to these authors.

Proposition 3.6 ([KR67]). *For a given integer n and $f \in C([0,1])$ we have*

$$\lim_{m\to\infty} B_n^m(f;x) \to B_1(f;x) = f(0) + (f(1) - f(0))x$$

uniformly on $[0,1]$.

The iterates of the Bernstein operator converge uniformly against the linear interpolation operator at the endpoints of $[0,1]$. We will provide an new proof based on operator theory in Chapter 7 and Chapter 9.

3.1.2 The Kantorovič operator

As covered in the preceding sections, the Bernstein operator is only able to approximate continuous functions as it needs a point evaluation. An extension to integrable functions has been constructed by L. Kantorovič in 1930 in order to transfer the theorem of Weierstrass to the space of integrable functions (in the sense of Lebesgue).

To this end, he defined a sequence of operators $K_n : L^1([0,1]) \to C([0,1])$ as follows. Let $f \in L^1([0,1])$, then

$$K_n(f;x) := (n+1) \sum_{k=0}^{n} \int_{\frac{k}{n+1}}^{\frac{k+1}{n+1}} f(t)\mathrm{d}t \binom{n}{k} x^k (1-x)^{n-k}, \qquad x \in [0,1],$$

see [Kan30]. The variation diminishing property of the Kantorovič operator has been shown by Bardaro et al. [BBSV03, Prop. 3.3]. There is a strong relation between the Kantorovič operator and the Bernstein operator as shown in the next proposition.

Proposition 3.7 ([Lor86, p.30]). *For all $f \in C^1([0,1])$*

$$K_n(Df) = D(B_{n+1}f)$$

Cooper and Waldron [CW00] have used this property to relate eigenvalues and eigenfunction of B_n with the corresponding eigenvalues and eigenfunction of K_n.

Proposition 3.8 ([CW00, Cor. 7.1]). *Denote by $p_k^{(n)}$ the eigenfunction of B_n corresponding to the eigenvalue $\lambda_k^{(n)}$. Then $\lambda_k^{(n+1)}$ is an eigenvalue of K_n with eigenfunction $Dp_k^{(n+1)}$, i. e.,*

$$K_n\left(Dp_k^{(n+1)}\right) = D(B_{n+1}p_k^{(n+1)}) = \lambda_{k+1}^{(n+1)}\left(Dp_k^{(n+1)}\right), \qquad k \in \{0,1,\ldots,n\}.$$

We will show a similar relation for the splines in order to prove spectral properties of the Schoenberg operator that will be introduced in the next section.

Lorentz [Lor37] has proved in his dissertation that for each $f \in L^1([0,1])$ the limit

$$\lim_{n\to\infty} \int_0^1 |K_n(f;x) - f(x)|\,\mathrm{d}x \to 0$$

holds. Using a theorem of Orlicz [Orl34] for kernel based integral transforms, Lorentz could prove in 1953 that the approximations $K_n(f;x)$ converge strongly for each $f \in L^p([0,1])$ in the corresponding $\|\cdot\|_p$-norm if $n \to \infty$.

Proposition 3.9 ([Lor53, Thm. 2.1.2]). *For each $f \in L^p([0,1])$, $1 \le p < \infty$,*

$$\lim_{n\to\infty} \|K_nf - f\|_p \to 0.$$

Quantitative results based on the integral modulus of continuity have been first shown by Grundmann [Gru76] for $L^1([0,1])$. In [Mül76], Müller has considered the general case for the L^p-spaces, where $1 \leq p < \infty$.

Proposition 3.10 ([Mül76, Satz 2]). *For $f \in L^p([0,1])$, $1 \leq p < \infty$, the estimate*

$$\|K_n f - f\|_p \leq M_p \cdot \omega_{1,p}\left(f, \frac{1}{\sqrt{n}}\right)$$

holds, where M_p is a constant independent on f.

The iterates of the Kantorovič operator have been considered by Nagel using the relation of Proposition 3.7.

Proposition 3.11 ([Nag82]). *For a given integer n and $f \in L^1([0,1])$, there holds*

$$\lim_{m\to\infty} K_n^m(f;x) \to \int_0^1 f(t)\mathrm{d}t.$$

In Chapter 7 and Chapter 9, we will show a generalization for the convergence of the iterates on the L^p-spaces, $1 < p < \infty$.

3.2 Spline Operators

Even though the Bernstein operator has the beautiful properties to be variation diminishing, to preserve convexity up to order n, and the uniform convergence of the derivatives of order up to n, the approximation with Bernstein polynomials results in a slow convergence rate. In this section, we will introduce approximation operators based on splines that guarantee a faster convergence rate, while the other properties of the Bernstein operator are weakened. The Schoenberg operator discussed in Section 3.2.1 is only able to preserve the positivity, the monotonicity and the convexity and the derivatives converge only up to order 2. As main advantage, the knots can be chosen arbitrarily and for smooth functions, the approximation rate is of quadratic order. We will first introduce spline basis functions on given knots. Afterwards, we will introduce the variation diminishing operator devised by Schoenberg and present the integral Schoenberg operator analogous to the Kantarovič operator. Finally, we will prove that the Schoenberg operator has only distinct real eigenvalues (except 0 and 1).

Let $n > 0$ be an integer and $\{x_0, \ldots, x_n\}$ be a finite partition of $[0,1]$ such that

$$0 = x_0 < x_1 < \ldots < x_n = 1.$$

For $k > 0$, we consider the extended *knot sequence* $\Delta_n = \{x_j\}_{j=-k}^{n+k}$, where the first and the last knot is repeated k-times, i. e.,

$$x_{-k} := \cdots := x_{-1} := 0, \quad \text{and} \quad x_{n+1} := \cdots := x_{n+k} := 1.$$

The k-fold repetition of the first and the last knot will lead to the interpolation property at these knots, as we will see later. We will denote by $|\Delta_n|_{\min}$ and $|\Delta_n|_{\max}$ the minimal and maximal mesh gauge,

$$|\Delta_n|_{\min} := \min_{0 \le j < n} \left(x_{j+1} - x_j\right), \qquad |\Delta_n|_{\max} := \max_{0 \le j < n} \left(x_{j+1} - x_j\right).$$

If the knots are uniformly spaced then $|\Delta_n|_{\min} = |\Delta_n|_{\max}$ holds. Now we are able to define a spline function of a certain degree as a piecewise polynomial of the following form, according to Curry and Schoenberg [CS66].

Definition 3.2. A function $s(x) \in \mathcal{S}(\Delta_n, k)$ is said to be a *spline function of degree* k, provided that it satisfies the following conditions:

1. $s(x) \in C^{k-1}([0,1])$,
2. $s(x) \in \mathcal{P}_k$ in each interval $\left(x_j, x_{j+1}\right)$ for all $j \in \{0, \dots, n-1\}$.

The approximation with splines, i.e., piecewise polynomials, on knots that are uniformly distributed has first been developed by Schoenberg in 1946, see the articles [Sch46b; Sch46a]. An illustration of a basis of the spline space where the knots are uniformly distributed is shown in Figure 3.2. In the review of the very same article [Sch46b; Sch46a], Curry has already pointed out an extension to non-equidistantly spaced knots, while this generalization has not been published until 1966 by Curry and Schoenberg [CS66]. Figure 3.3 shows basis functions for non-equidistant knots.

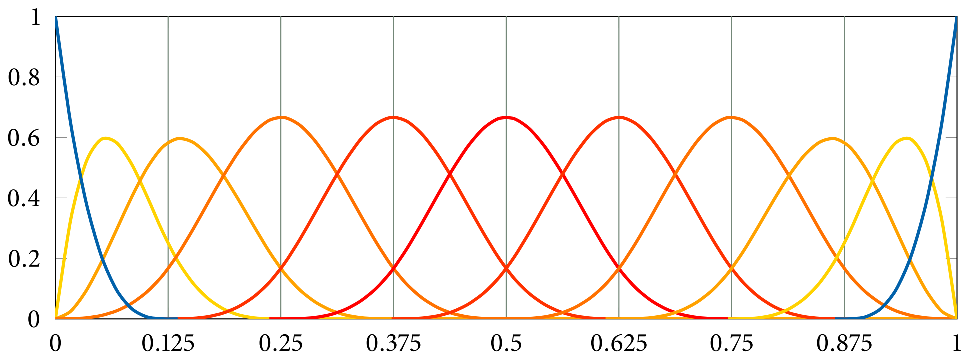


Figure 3.2: *The normalized B-splines of degree $k = 3$ and uniformly distributed knots ($n = 8$). Note that the five basis functions in the center, $N_{0,3}, \dots, N_{4,3}$, are translates of each other.*

Curry and Schoenberg have provided also a basis for the spline space, the so called B-spline basis. Let us define the *B-splines* $M_{j,k}$ for $j \in \{-k, \dots, n-1\}$ as in [CS66] by the $(k+1)$-th divided difference of the truncated power function of degree k:

$$M_{j,k}(x) := \Delta(x_j, \dots, x_{j+k+1})(k+1)(\cdot - x)_+^k. \tag{3.2}$$

These functions have finite support,

$$\operatorname{supp} M_{j,k}(x) = \left[x_j, x_{j+k+1}\right],$$

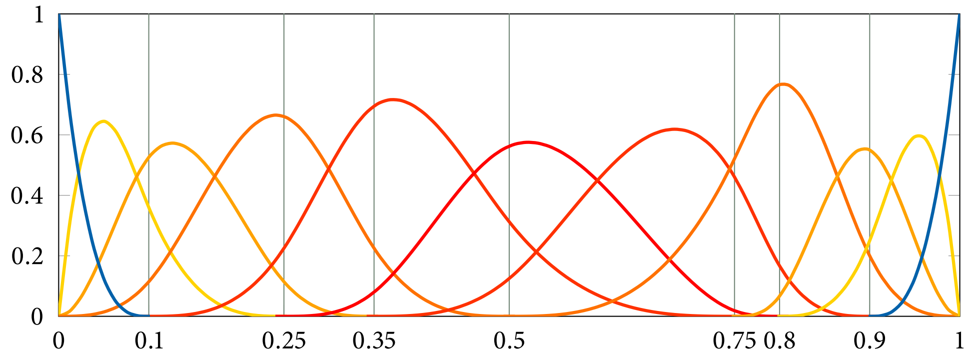


Figure 3.3: *The normalized B-splines of degree* $k = 3$ *and non-equidistant knots* ($n = 8$). *Here,* $\Delta_n = \{0, 0.1, 0.25, 0.35, 0.5, 0.75, 0.8, 0.9, 1\}$.

and are normalized to have integral one,

$$\int_0^1 M_{j,k}(x)\mathrm{d}x = 1.$$

The $n + k$ B-splines $M_{j,k}$ are linearly independent and moreover, they form a basis of the spline space $\mathcal{S}(\Delta_n, k)$. As consequence, every spline $s(x) \in \mathcal{S}(\Delta_n, k)$ can be uniquely represented in the form

$$s(x) = \sum_{j=-k}^{n-1} c_j M_{j,k}(x),$$

with coefficients $c_j \in \mathbb{R}$. For more detail, we refer to Curry and Schoenberg [CS66]. It has been shown in [Sch67, Thm. 5] that this basis is in particular variation-diminishing. Moreover, S. Karlin has shown that the B-spline basis are in fact totally positive [Kar68]. Note this results has been proved using a different technique by de Boor [dBoo76]. De Boor and DeVore have constructed a purely geometrical proof in [dBD85].

Proposition 3.12 ([Kar68, Thm. 4.1 (p.527)]). *The kernel* $M_{j,k}(x)$ *is totally positive on* $\mathbb{Z} \times \mathbb{R}$.

Recall that using Proposition 2.7, the variation diminishing property follows for every integral operator of type (2.1) as discussed in Section 2.3.2.

While the B-splines are normalized to have integral one, there exists a renormalization yielding a basis that forms a partition of unity. Accordingly, we define the *normalized B-splines* as in [Sch67, p.270 and p.274] for $x \in [0, 1]$ by

$$N_{j,k}(x) := \frac{x_{j+k+1} - x_j}{k+1} M_{j,k}(x) = (\xi_{j,k+1} - \xi_{j-1,k+1}) M_{j,k}(x). \tag{3.3}$$

We denote by $\mathcal{S}(\Delta_n, k)$ the spline space of degree k with respect to the knot sequence Δ_n,

$$\mathcal{S}(\Delta_n, k) = \left\{ \sum_{j=-k}^{n-1} c_j N_{j,k} \,:\, c_j \in \mathbb{R},\ j \in \{-k, \dots, n-1\} \right\} \subset C^{k-1}([0,1]).$$

Since $\mathcal{S}(\Delta_n, k)$ is an finite-dimensional subspace of $C([0,1])$ of dimension $n+k$, $\mathcal{S}(\Delta_n, k)$ is a Banach space with the inherited norm $\|\cdot\|_\infty$. For more information on spline spaces and splines see, e.g., the books of de Boor [dBoo01], Nürnberger [Nür89], and Schumaker [Sch07].

In the next section, we will see that these basis functions $\{N_{j,k}\}$ form a partition of unity and provide a way to preserve linear functions. These two properties are due to Schoenberg's variation-diminishing spline operator, which will be discussed in the following.

3.2.1 The Schoenberg operator

The *Schoenberg operator* $S_{\Delta_n,k} : C([0,1]) \to C([0,1])$ *of degree* k with respect to the knot sequence Δ_n is defined for the continuous function f by

$$S_{\Delta_n,k} f(x) = \sum_{j=-k}^{n-1} f(\xi_{j,k}) N_{j,k}(x), \quad x \in [0,1], \tag{3.4}$$

where $\xi_{j,k}$ are the so called *Greville nodes*, see the supplement in [Sch67], defined for all $j \in \{-k, \dots, n-1\}$ by

$$\xi_{j,k} := \frac{x_{j+1} + \dots + x_{j+k}}{k}. \tag{3.5}$$

The *normalized B-splines* $N_{j,k}$ are defined for all $j \in \{-k, \dots, n-1\}$ and $x \in [0,1]$ by

$$N_{j,k}(x) := (x_{j+k+1} - x_j)\, \Delta(x_j, \dots, x_{j+k+1})(\cdot - x)_+^k,$$

where $\Delta(()x_j, \dots, x_{j+k+1})$ denotes the divided difference operator and x_+^k denotes the truncated power function as defined in Section 1.3 on page 6. In order to guarantee the evaluation of the Schoenberg operator on the whole interval $[0,1]$, especially at the point 1, the B-splines are chosen here in such a way that the B-splines are right-continuous at the knots $x_1, \dots, x_{n-1}$, while at the point $x_n = 1$ they are chosen to be left continuous. This is due to the definition of the divided difference operator. For more details, we refer to de Boor's book [dBoo01].

The operator $S_{\Delta_n,k}$ has been first devised by I. J. Schoenberg in 1967 [Sch67]. This operator evaluates continuous functions at the so called Greville nodes. These are named after Greville who has calculated as first an exact representation of these nodes, see the supplement [Sch67, p. 286 ff.]. In this manner, Schoenberg's operator provides a variation-diminishing smooth approximation of continuous functions by a linear combination of the splines basis functions. The smoothness of the approximation is depending on the degree of the spline. The Schoenberg operator is able to reproduce constants, and hence, the normalized B-splines form a partition of unity

$$\sum_{j=-k}^{n-1} N_{j,k}(x) = 1. \tag{3.6}$$

Moreover, the Schoenberg operator can reproduce linear functions, i.e.,

$$\sum_{j=-k}^{n-1} \xi_{j,k} N_{j,k}(x) = x, \tag{3.7}$$

due to the chosen Greville nodes. For more properties of this operator, see, e.g., [Sch67; MS66; Mar70]. We note that the reference for the Greville nodes and the Schoenberg operator is dated by 1967, while the conference where the result has first been announced has been held in 1965. A comprehensive overview of direct inequalities can be found in [BGKT02].

In [Sch67, Theorem 7] it has been shown by I. J. Schoenberg that the operator $S_{\Delta_n,k}$ is in fact variation-diminishing.

Proposition 3.13 (Variation-diminishing property). *For all $f \in C([0,1])$ the approximation with the Schoenberg operator is variation-diminishing, i. e.,*

$$\nu(S_{\Delta_n,k} f) \le \nu(f).$$

Furthermore, the Schoenberg operator interpolates at the endpoints of the interval due to the k-fold end knots, i. e.,

$$S_{\Delta_n,k} f(0) = f(0), \qquad S_{\Delta_n,k} f(1) = f(1)$$

holds for all $f \in C([0,1])$ and these are the only interpolation points of $S_{\Delta_n,k}$.

The derivatives of the spline approximation derived by the Schoenberg operator can be explicitly calculated. To this end, let us define the discrete backward difference operator ∇_l working on the Greville nodes as

$$\nabla_l f(\xi_{j,k}) := \frac{f(\xi_{j,k}) - f(\xi_{j-1,k})}{\xi_{j,l} - \xi_{j-1,l}}.$$

As for all $x \in [0,1]$,

$$D N_{j,k}(x) = \frac{N_{j,k-1}(x)}{\xi_{j,k} - \xi_{j-1,k}} - \frac{N_{j+1,k-1}(x)}{\xi_{j+1,k} - \xi_{j,k}},$$

holds for all $j \in \{-k, \dots, n-1\}$, we get the following two representation according to Lemma 1 and 2 in Marsden [Mar70, pp. 32–36],

$$\begin{aligned} DS_{\Delta_n,k} f(x) &= \sum_{j=1-k}^{n-1} \frac{f(\xi_{j,k}) - f(\xi_{j-1,k})}{\xi_{j,k} - \xi_{j-1,k}} N_{j,k-1}(x) \\ &= \sum_{j=1-k}^{n-1} \nabla_k f(\xi_{j,k}) N_{j,k-1}(x), \end{aligned}$$

while the second derivative can be represented in the form

$$\begin{aligned} D^2 S_{\Delta_n,k} f(x) &= \sum_{j=2-k}^{n-1} \frac{\frac{f(\xi_{j,k})-f(\xi_{j-1,k})}{\xi_{j,k}-\xi_{j-1,k}} - \frac{f(\xi_{j-1,k})-f(\xi_{j-2,k})}{\xi_{j-1,k}-\xi_{j-2,k}}}{\xi_{j,k-1} - \xi_{j-1,k-1}} N_{j,k-2}(x) \\ &= \sum_{j=2-k}^{n-1} \nabla_{k-1} \nabla_k f(\xi_{j,k}) N_{j,k-2}(x). \end{aligned}$$

The following two important statements can be found in Marsden [Mar70]. The first lemma states that the first two derivatives of the approximation converge uniformly against the derivatives of the function, whereas for the second order derivative this only holds true on compact subsets of $(0,1)$.

Proposition 3.14 ([Mar70, Theorem 9 and 11]). *Let* $f \in C^2([0,1])$. *Then*

1. $\lim_{n\to\infty} S_{\Delta_n,k} f(x) = f(x)$ *uniformly in* $[0,1]$,
2. $\lim_{n\to\infty} DS_{\Delta_n,k} f(x) = Df(x)$ *uniformly in* $[0,1]$,
3. $\lim_{n\to\infty} D^2 S_{\Delta_n,k} f(x) = D^2 f(x)$ *uniformly on compact subsets of* $(0,1)$.

The Schoenberg operator is able to preserve positivity, monotonicity and convexity of smooth functions.

Proposition 3.15 ([Mar70, Theorem 10]). *For* $f \in C^2([0,1])$ *and* $k > 2$ *we have that*

1. *If* $f(x) \geq 0$ *on* $[0,1]$, *then* $S_{\Delta_n,k} f(x) \geq 0$ *on* $[0,1]$.
2. *If* $Df(x) \geq 0$ *on* $[0,1]$, *then* $DS_{\Delta_n,k} f(x) \geq 0$ *on* $[0,1]$.
3. *If* $D^2 f(x) \geq 0$ *on* $[0,1]$, *then* $D^2 S_{\Delta_n,k} f(x) \geq 0$ *on* $[0,1]$.

Remark. In contrast to the Bernstein operator, the above lemma does not hold for the third derivative. If $f \in C^3([0,1])$ *and* $D^3 f(x) \geq 0$ *on* $[0,1]$, *then for all* $k \geq 3$, $D^3 S_{\Delta_n,k} f(x)$ *need not to be non-negative on* $[0,1]$.

A quantitative upper estimate by the second order modulus of smoothness has been first considered by Esser [Ess76]. We show here the estimate given by Beutel et al. [BGKT02], where quantitative results on the Schoenberg operator have been extensively discussed.

Proposition 3.16 ([BGKT02, Cor. 7]). *The following uniform estimates hold for all $f \in C([0,1])$:*

$$\|S_{\Delta_n,k} f - f\|_\infty \leq \frac{5}{4} \cdot \omega_2\left(f, \frac{1}{k}\right),$$

if Δ_n is fixed and $k \to \infty$, and

$$\|S_{\Delta_n,k} f - f\|_\infty \leq \left(1 + \frac{k+1}{24}\right) \cdot \omega_2\left(f, |\Delta_n|_{\max}\right)$$

for fixed spline degree $k > 0$ and $|\Delta_n|_{\max} \to 0$.

We will prove corresponding lower estimates for the Schoenberg operator in Chapter 4. To this end, we will also show the asymptotic behaviour of the iterates of $S_{\Delta_n,k}$.

3.2.2 Integral Schoenberg operator

The integral Schoenberg operator has been introduced by M. W. Müller [Mül77] as generalization to the $L^p([0,1])$-spaces of the variation-diminishing spline approximation devised by Schoenberg [Sch67] only on the space of continuous functions. As in the last section, we assume Δ_n to be the extended knot sequence

$$x_{-k} = \dots = x_0 = 0 < x_1 < \ldots < x_n = \dots = x_{n+k} = 1.$$

Let $1 \leq p < \infty$. The generalization of the Schoenberg operator to arbitrary $L^p([0,1])$-spaces is now given for $f \in L^p([0,1])$ by

$$V_{\Delta_n,k} f(x) := D S_{\Delta_n,k+1} F(x) = \sum_{j=-k}^{n-1} \int_{\xi_{j-1,k+1}}^{\xi_{j,k+1}} f(t)\mathrm{d}t \frac{N_{j,k}(x)}{\xi_{j,k+1} - \xi_{j-1,k+1}}, \tag{3.8}$$

where $F(x) = \int_0^x f(t)\mathrm{d}t$. Note that according to (3.3) we can replace the last term

$$\left(\xi_{j,k+1} - \xi_{j-1,k+1}\right)^{-1} N_{j,k}(x)$$

by the B-splines $M_{j,k}(x)$ defined in (3.2). The following proposition gathers the convergence properties of $V_{\Delta_n,k}$.

Proposition 3.17 ([Mül77, Thm. 1]). *For $f \in L^p([0,1])$, $1 \leq p \leq \infty$, there holds for fixed spline degree k*

$$\lim_{|\Delta_n|_{\max} \to 0} \|V_{\Delta_n,k} f - f\|_p = 0.$$

If the nodes Δ_n are fixed, then similarly

$$\lim_{k \to \infty} \|V_{\Delta_n,k} f - f\|_p = 0. \tag{3.9}$$

Müller has also shown an upper estimate of the approximation error with respect to the integral modulus of continuity.

Proposition 3.18 ([Mül77, Thm. 3]). *Let $f \in L^p([0,1])$, $1 \le p \le \infty$. Then there exists $M > 0$ independent of f and p, depending only on Δ_n and k, such that*

$$\|V_{\Delta_n,k}f - f\|_p \le M \cdot \omega_{1,p}(f, |\Delta_n|_{\max}).$$

Din [Din83] has found the related upper estimate when n is fixed and k tends to infinity.

Proposition 3.19 ([Din83, Thm. 6]). *Let $f \in L^p([0,1])$, $1 \le p \le \infty$. Then there exists $M > 0$ independent of f and k, n, p such that*

$$\|V_{\Delta_n,k}f - f\|_p \le M \cdot \omega_{1,p}(f, \frac{1}{\sqrt{k+1}}).$$

We will prove a corresponding lower estimate in Chapter 10 and show the limit of the iterates of V_{Δ_n}, k in Chapter 7 and Chapter 9.

3.2.3 On the eigenvalues of the Schoenberg operator

The eigenvalues of the Bernstein operator have been revealed already in 1966 by the Russian Călugăreanu. Up to our knowledge results on the eigenvalues of the Schoenberg operator are not known explicitly. In the following, we show that 1 is a simple eigenvalue of $V_{\Delta_n,k}$ and all the other eigenvalues are distinct non-negative, real numbers. Finally, we will show that the Schoenberg operator has the same eigenvalues as $V_{\Delta_n,k}$ with the exception that 1 is not a simple eigenvalue anymore as the Schoenberg operator reproduces constants and linear functions.

Theorem 3.20. *The collocation matrix of the integral Schoenberg operator with the B-splines as defined in* (3.2)

$$\left(\int_{\xi_{i-1,k+1}}^{\xi_{i,k+1}} M_{j,k}(t)\mathrm{d}t\right)_{ij}$$

is an oscillation matrix and thus, all eigenvalues are distinct positive real numbers. In particular, 1 *is the only eigenvalue on the unit circle and* $\dim(\ker(V_{\Delta_n,k} - I)) = 1$.

Proof. Recall, that the Greville nodes $\xi_{j,k}$ are defined as the knot averages as in (3.5) by

$$\xi_{j,k} := \frac{x_{j+1} + \cdots + x_{j+k}}{k}.$$

First, note that the relations

$$x_j < \xi_{j-1,k+1} < \xi_{j,k} < \xi_{j,k+1} < x_{j+k+1}, \quad x_{j+1} < \xi_{j,k+1} < \xi_{j+1,k} < \xi_{j+1,k+1} < x_{j+k+2}$$

and

$$\operatorname{supp} M_{j,k} = \left[x_j, x_{j+k+1}\right], \qquad \operatorname{supp} M_{j+1,k} = \left[x_{j+1}, x_{j+k+2}\right]$$

hold. From the continuity of $M_{j,k}$ and $M_{j+1,k}$ and the relations

$$M_{j,k}(\xi_{j,k}) > 0, \quad M_{j+1,k}(\xi_{j,k+1}) > 0, \quad M_{j,k}(\xi_{j,k+1}) > 0,$$

we can follow that

$$\int_{\xi_{j-1,k+1}}^{\xi_{j,k+1}} M_{j,k}(t)\mathrm{d}t > 0, \int_{\xi_{j-1,k+1}}^{\xi_{j,k+1}} M_{j+1,k}(t)\mathrm{d}t > 0, \text{ and } \int_{\xi_{j,k+1}}^{\xi_{j+1,k+1}} M_{j,k}(t)\mathrm{d}t > 0$$

holds. Moreover, the matrix

$$\left(\int_{\xi_{i-1,k+1}}^{\xi_{i,k+1}} M_{j,k}(t)\mathrm{d}t\right)_{ij}$$

is non-singular as the B-splines $M_{j,k}(x)$ are linearly independent and so are the functionals $\int_{\xi_{i-1,k+1}}^{\xi_{i,k+1}} \cdot \,\mathrm{d}t$ due to their distinct support. From that we can conclude that the collocation matrix is oscillatory according to the criteria of Proposition 2.5 provided by Gantmacher and Krein [GK02]. □

Cooper and Waldron have used in [CW00] a relation between the Bernstein operator and the Kantorovič operator to deduce the eigenvalues of the Kantorovič operator. We prove here a similar relation between the Schoenberg operator and its counterpart for the L^p-spaces to characterize the eigenvalues of the Schoenberg operator.

Lemma 3.21. *For all $f \in C([0,1])$ the relation*

$$DS_{\Delta_n,k} f = V_{\Delta_n,k-1} Df$$

holds.

Proof. Follows directly by the definition of the integral Schoenberg operator, as by (3.8)

$$V_{\Delta_n,k} f(x) = DS_{\Delta_n,k+1} \int_0^x f(t)\mathrm{d}t.$$

Then a simple calculation shows

$$V_{\Delta_n,k-1} Df(x) = DS_{\Delta_n,k+1} \int_0^x Df(t)\mathrm{d}t = DS_{\Delta_n,k} \left(f(x) - f(0)\right) = DS_{\Delta_n,k} f(x).$$

In the last step, we used the linearity of $S_{\Delta_n,k}$ and that $S_{\Delta_n,k}$ can reproduce constants. □

Finally, we can use that the collocation matrix of $V_{\Delta_n,k}$ is an oscillatory matrix to prove that the eigenvalues of the Schoenberg operator are non-negative, real numbers.

The only eigenvalue with multiplicity two is 1, whereas all the others have multiplicity one.

Theorem 3.22. *The eigenvalues of the Schoenberg operator are characterized by*

$$1 = \lambda_0 = \lambda_1 > \lambda_2 > \cdots > \lambda_{n+k-1} > \lambda_{n+k} = 0.$$

Thus, besides 0 *and* 1 *the Schoenberg operator has* $n + k - 1$ *distinct positive real eigenvalues.*

Proof. We use that $V_{\Delta_n,k-1}$ has $n + k - 1$ distinct positive eigenvalues combined with the eigenvalue 0 coming from the finite-dimensional range of $V_{\Delta_n,k-1}$ and Lemma 3.21 saying that

$$DS_{\Delta_n,k} f = V_{\Delta_n,k-1} Df$$

holds for all $f \in C([0,1])$.

We show first that $0 \in \sigma_p(S_{\Delta_n,k})$. To this end, let $f \in C([0,1])$ be a function, such that

$$f(\xi_j) = 0 \qquad \text{for all } j \in \{-k, \dots, n-1\}$$

and such that there exists $x \in [0,1] \setminus \{\xi_j : j \in \{-k, \dots, n-1\}\}$ with $f(x) \neq 0$. For example, consider the polynomial $f(x) = \prod_{i=-k}^{n-1}(x - \xi_i)$. Clearly, $f \in C([0,1])$ and we obtain $S_{\Delta_n,k} f = 0 \cdot f = 0$, because for all $x \in [0,1]$

$$S_{\Delta_n,k} f(x) = \sum_{j=-k}^{n-1} \left[\prod_{i=-k}^{n-1} (\xi_j - \xi_i) \right] N_{j,k}(x) = 0.$$

We now construct the set of eigenvalues and eigenfunctions of $S_{\Delta_n,k}$ by their relation to the integral Schoenberg operator $V_{\Delta_n,k-1}$. To this end, let us consider now an eigenfunction $s \in \mathcal{S}(\Delta_n, k)$ of $S_{\Delta_n,k}$ corresponding to some eigenvalue $\lambda \in \sigma_p(S_{\Delta_n,k}) \setminus \{0,1\}$. Then we calculate

$$V_{\Delta_n,k-1} Ds = DS_{\Delta_n,k} s = \lambda Ds.$$

This states in particular that the eigenvalue $\lambda \neq 0$ of the Schoenberg operator with corresponding eigenfunction s is again an eigenvalue of $V_{\Delta_n,k-1}$ with associated eigenfunction Ds. The only exception yields the eigenfunction 1. Here, we obtain

$$V_{\Delta_n,k} D1 = D1 = 0.$$

Therefore, $0 = D1$ does not yield a new linear independent eigenfunction of $V_{\Delta_n,k-1}$. Whereas, the eigenfunction x corresponding to the eigenvalue 1 is mapped to the constant eigenfunction 1:

$$V_{\Delta_n,k} Dx = DS_{\Delta_n,k} x = Dx = 1.$$

As all the eigenfunctions $s_1, \ldots, s_{n+k-1}$ of $S_{\Delta_n,k}$ corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_{n+k-1}$ are linearly independent, so are the functions $Ds_1, \ldots, Ds_{n+k-1}$. Consequently, $\lambda_1, \ldots, \lambda_{n+k-1}$ are exactly the $n + k - 1$ distinct positive eigenvalues of $S_{\Delta_n,k}$. This concludes the proof. □

CHAPTER 4 Lower Estimates for Variation Diminishing Splines

"All truths are easy to understand once they are discovered; the point is to discover them."

GALILEO GALILEI

To be able to detect singularities of piecewise smooth functions we will establish lower estimates for Schoenberg's variation-diminishing splines which we have introduced in previous chapter. Concretely, we will show a lower bound for the approximation error by the second order modulus of smoothness or the related K-functional. Using the already known upper estimate allows us to interpret the approximation error in terms of local smoothness of the function.

In 2002, L. Beutel and her coauthors investigated in the article [BGKT02] quantitative direct approximation inequalities for the Schoenberg operator. More importantly, the authors stated an interesting conjecture regarding the equivalence of the approximation error of the Schoenberg operator on $[0,1]$ and the second order Ditzian-Totik modulus of smoothness with weight function $\varphi(x) = \sqrt{x(1-x)}$.

We show here a lower estimate for the approximation error by the classical second order modulus of smoothness. Thereby, we first characterize the asymptotic behavior of the iterates of the Schoenberg operator. Afterwards, we use this result in order to prove a lower bound of the approximation error with respect to the second order modulus of smoothness. The convergence of the iterates of the Schoenberg operator to the operator of linear interpolation at the endpoints of the interval $[0,1]$ can be also seen by the method provided by Gavrea and Ivan [GI11a] based on Korovkin-type approximation theory. However, while their methods ensure the uniform convergence of those iterates, they do not give the rate of convergence in which in fact we are interested. Therefore, our approach uses an earlier result of Badea [Bad09], where the asymptotic behavior of the iterates is characterized by their spectral properties. Moreover, these results provide a simple and elegant generalization of the results of the manuscript [NK13] to the non-uniform case by using a functional analysis based framework. Based on this framework we will show in the second part of this thesis how to transfer this method to linear operators on general Banach spaces with smooth range. We will also provide a complete characterization of the iterates of operators based on spectral properties provided that the fixed point space of the operator is finite-dimensional in Chapter 9.

Let us briefly recall the definition of the Schoenberg operator. Given integers $n > 0$, $k > 0$ and the knot sequence $\Delta_n = \{x_j\}_{j=-k}^{n+k}$ such that

$$x_{-k} = \cdots = x_0 = 0 < x_1 < \ldots < x_n = \cdots = x_{n+k} = 1$$

we consider the variation-diminishing spline operator $S_{\Delta_n,k} : C([0,1]) \to C([0,1])$

$$S_{\Delta_n,k} f(x) = \sum_{j=-k}^{n-1} f(\xi_{j,k}) N_{j,k}(x), \quad 0 \le x \le 1,$$

where $\xi_{j,k}$ are the Greville nodes and $N_{j,k}$ the normalized B-splines. More details can be found in the previous chapter in Section 3.2. For the convenience of the reader we highlight once more the partition of unity property

$$\sum_{j=-k}^{n-1} N_{j,k}(x) = 1, \tag{4.1}$$

and the ability to reproduce linear functions, i.e.,

$$\sum_{j=-k}^{n-1} \xi_{j,k} N_{j,k}(x) = x. \tag{4.2}$$

Also recall, that we denote by $|\Delta_n|_{\min}$ and $|\Delta_n|_{\max}$ the minimal and the maximal mesh gauge, respectively. Finally, we mention here that the results of this chapter can be found in the manuscript [NCF14] submitted to the *Journal of Complexity*, which is joint work with Paula Cerejeiras and Brigitte Forster.

4.1 Spectral location and the limit of the iterates

We investigate first some basic properties of the Schoenberg operator needed in order to characterize the spectrum and to show the limit of the iterates. The following fact can, e.g., be found in [dBoo73].

Theorem 4.1. *The Schoenberg operator $S_{\Delta_n,k} : C([0,1]) \to C([0,1])$ is bounded and $\|S_{\Delta_n,k}\|_{op} = 1$.*

Proof. Let $f \in C([0,1])$ with $\|f\|_\infty = 1$. Then

$$\|S_{\Delta_n,k} f\|_\infty = \left\| \sum_{j=-k}^{n-1} f(\xi_{j,k}) N_{j,k}(x) \right\|_\infty \le \|f\|_\infty \cdot \left\| \sum_{j=-k}^{n-1} N_{j,k}(x) \right\|_\infty = 1,$$

because of property (4.1). Therefore, $\|S_{\Delta_n,k}\| \leq 1$. By considering now the constant function $1 \in C([0,1])$, we get $\|S_{\Delta_n,k}1\|_\infty = 1$. Hence, the operator has norm 1, $\|S_{\Delta_n,k}\|_{op} = 1$. □

Due to the finite-dimensional image of $S_{\Delta_n,k}$, we can directly obtain the compactness of the Schoenberg operator.

Theorem 4.2. *The Schoenberg operator $S_{\Delta_n,k} : C([0,1]) \to C([0,1])$ is compact and as such* $\operatorname{ran}(S_{\Delta_n,k} - I)$ *is closed. Besides,* 1 *is not a cluster point of the spectrum $\sigma(S_{\Delta_n,k})$.*

Proof. From Theorem 4.1 it follows that the operator is bounded with $\|S_{\Delta_n,k}\|_{op} = 1$ and maps continuous functions to the spline space $\mathcal{S}(\Delta_n, k)$. Therefore, the operator has finite rank and finite rank operators are compact. For compact operators $\operatorname{ran}(T-I)$ is closed and 0 is the only possible cluster point of $\sigma(S_{\Delta_n,k})$, see [Rud91, Thm. 4.25]. □

Next, we will characterize the spectrum of the Schoenberg operator based on the results we have shown in the end of the previous chapter. We will extend Theorem 3.22 stating the point spectrum of $S_{\Delta_n,k}$ consists of only of distinct, real numbers to the spectrum of $S_{\Delta_n,k}$.

Corollary 4.3. *The spectrum of the Schoenberg operator consists only of the point spectrum and*

$$\sigma(S_{\Delta_n,k}) \subset B(0,1) \cup \{1\}.$$

Proof. Since $\|S_{\Delta_n,k}\|_{op} = 1$, for $\lambda \in \sigma(S_{\Delta_n,k})$ the inequality

$$|\lambda| \leq \|S_{\Delta_n,k}\|_{op} = 1$$

holds. Therefore, $\sigma(S_{\Delta_n,k}) \subset \overline{B(0,1)}$.

For compact operators, it is known that every $\lambda \neq 0$ in the spectrum is contained in the point spectrum of the operator. This classical result is stated, e.g., in Rudin [Rud91, Thm. 4.25]. The $n+k$ eigenvalues $\lambda_0, \ldots, \lambda_{n+k}$ of $S_{\Delta_n,k}$ can be characterized by Theorem 3.22 as

$$1 = \lambda_0 = \lambda_1 > \lambda_2 > \cdots > \lambda_{n+k-1} > \lambda_{n+k} = 0.$$

As $0 \in \sigma_p(S_{\Delta_n,k})$, it follows that

$$\sigma(S_{\Delta_n,k}) = \sigma_p(S_{\Delta_n,k}) \subset [0,1) \cup \{1\} \subset B(0,1) \cup \{1\}.$$

□

For $m \to \infty$ we investigate the asymptotic behaviour of the iterates $S^m_{\Delta_n,k}$ of the Schoenberg operator in order to prove the lower estimates, where the iterates are defined by $S^0_{\Delta_n,k} = I$ and for $m \in \mathbb{N}$ by

$$S^m_{\Delta_n,k} f(x) = S^{m-1}_{\Delta_n,k} (S_{\Delta_n,k} f)(x).$$

Concretely, we show that the iterates of the Schoenberg operator converge in the limit uniformly to the linear operator $L : C([0,1]) \to C([0,1])$, defined for $f \in C([0,1])$ by

$$(Lf)(x) = f(0) + \big(f(1) - f(0)\big)x, \qquad x \in [0,1],$$

i. e., we will show that

$$\lim_{m\to\infty} \left\| S^m_{\Delta_n,k} - L \right\|_{op} = 0.$$

In [Bad09] it has been shown that operators of a certain structure converge to this linear operator L. In fact, the Schoenberg operator $S_{\Delta_n,k} : C([0,1]) \to C([0,1])$ fulfills the following required properties:

- The operator $S_{\Delta_n,k}$ is bounded and $\operatorname{ran}(S_{\Delta_n,k} - I)$ is closed,
- $\ker(S_{\Delta_n,k} - I) = \operatorname{span}(1, x)$, i.e., the Schoenberg operator reproduces constant and linear functions,
- $S_{\Delta_n,k} f(0) = f(0)$ and $S_{\Delta_n,k} f(1) = f(1)$ for every $f \in C([0,1])$, i.e., the Schoenberg operator interpolates start and end points,
- $\sigma(S_{\Delta_n,k}) \subset B(0,1) \cup \{1\}$, and finally,
- 1 is not a cluster point of $\sigma(S_{\Delta_n,k})$, since

$$\sup\{|\lambda| : \lambda \in \sigma(S_{\Delta_n,k}) \setminus \{1\}\} < 1.$$

All these properties have been shown in the previous section. We can conclude:

Theorem 4.4. *With* $\gamma_{\Delta_n,k} := \sup\{\lambda \in \mathbb{C} : \lambda \in \sigma(S_{\Delta_n,k}) \setminus \{1\}\}$, *we obtain*

$$\left\| S^m_{\Delta_n,k} - L \right\|_{op} \le C \cdot \gamma^m_{\Delta_n,k}$$

for some suitable constant $1 \le C \le 1/(\gamma_{\Delta_n,k})$ *and therefore,*

$$\lim_{m\to\infty} \left\| S^m_{\Delta_n,k} - L \right\|_{op} = 0.$$

Proof. The result follows now immediately from [Bad09, Thm. 2.1] using the above mentioned properties of $S_{\Delta_n,k}$. □

4.2 Lower estimates for the approximation error

In this section, we show that for $r \in \mathbb{N}$, $r \geq 2$, $k > r$, there exists a uniform constant $M > 0$ such that

$$M \cdot \omega_r(f, t(\Delta_n, k)) \leq \|f - S_{\Delta_n,k} f\|_\infty$$

and $t(\Delta_n, k) \to 0$ provided that $\|f - S_{\Delta_n,k} f\|_\infty \to 0$. Recall, that the r-th modulus of smoothness $\omega_r : C([0,1]) \times \left(0, \frac{1}{r}\right] \to [0, \infty)$ is defined by

$$\omega_r(f,t) := \sup_{0<h<t} \sup \{|\Delta_h^r f(x)| : x \in [0, 1 - rh]\},$$

with the forward difference operator

$$\Delta_h^r f(x) = \sum_{l=0}^{r} (-1)^{r-l} \binom{r}{l} f(x + lh).$$

The r-th modulus of smoothness satisfies the following properties [Zyg02; Tim94]:

Lemma 4.5. *Let* $0 < t \leq \frac{1}{r}$ *be fixed.*

1. *For* $f_1, f_2 \in C([0,1])$, *the triangle inequality holds,*

$$\omega_r(f_1 + f_2, t) \leq \omega_r(f_1, t) + \omega_r(f_2, t). \tag{4.3}$$

2. *If* $f \in C([0,1])$, *then*

$$\omega_r(f,t) \leq 2^r \|f\|_\infty. \tag{4.4}$$

3. *If* $f \in C^r([0,1])$, *then*

$$\omega_r(f,t) \leq t^r \|D^r f\|_\infty. \tag{4.5}$$

Note that for $k > r$ the spline space $\mathcal{S}(\Delta_n, k) \subset C^r([0,1])$, because $S_{\Delta_n,k} f \in C^{k-1}([0,1])$. Hence, using inequalities (4.3) – (4.5), we obtain

$$\omega_r(f,t) \leq 2^r \|f - S_{\Delta_n,k} f\|_\infty + t^r \|D^r S_{\Delta_n,k} f\|_\infty. \tag{4.6}$$

This inequality shows the relation to Peetre's K-functional. Recall that $K_r(f, t^r)$, the Peetre K-functional is given by

$$K_r(f, t^r) = \inf_{g \in C^r([0,1])} \left\{ \|f - g\|_\infty + t^r \|D^r g\|_\infty \right\},$$

It has been shown by Butzer and Berens [BB67] and Johnen and Scherer [JS77] that the modulus of smoothness is equivalent to the corresponding K-functional. That is,

$$M^{-1} \cdot \omega_r(f,t) \leq K_r(f, t^r) \leq M \cdot \omega_r(f,t)$$

for a constant M independent of $0 < t \le 1$ and $f \in C([0,1])$.

In the following we will estimate the last term of (4.6) with respect to the approximation error $\|S_{\Delta_n,k} f - f\|_\infty$. To this end, we consider the minimal mesh length $|\Delta_n|_{\min}$ of the knots,

$$|\Delta_n|_{\min} := \min\left\{(x_{j+1,k} - x_{j,k}) \,:\, j \in \{0,\dots,n-1\}\right\}.$$

To establish the lower estimate, we need that the operator norm of the differential operator D^r is bounded on the range of $S_{\Delta_n,k}$ that is the spline space $\mathcal{S}(\Delta_n,k)$. We will prove first that the differential operator of order one is bounded on the spline space. Then we will extend this result to arbitrary orders.

Lemma 4.6. *The differential operator $D : \mathcal{S}(\Delta_n,k) \to \mathcal{S}(\Delta_n,k-1)$ is bounded with $\|D\|_{op} \le (2k/|\Delta_n|_{\min})d_k$, where $d_k > 0$ is a constant depending only on k.*

Proof. Let $s \in \mathcal{S}(\Delta_n,k)$, $s(x) = \sum_{j=-k}^{n-1} c_j N_{j,k}(x)$, with $\|s\|_\infty = 1$. According to [Mar70], we can calculate the derivative by

$$Ds(x) = \sum_{j=1-k}^{n-1} \frac{c_j - c_{j-1}}{\xi_{j,k} - \xi_{j-1,k}} N_{j,k-1}(x).$$

Then we obtain with the triangle inequality

$$\begin{aligned}\|Ds\|_\infty &= \left\| \sum_{j=1-k}^{n-1} \frac{c_j - c_{j-1}}{\xi_{j,k} - \xi_{j-1,k}} N_{j,k-1} \right\|_\infty \\ &= \left\| \sum_{j=1-k}^{n-1} \frac{k(c_j - c_{j-1})}{x_{j+k} - x_j} N_{j,k-1} \right\|_\infty .\end{aligned}$$

Due to the k-fold endpoints we can estimate the denominator only by

$$x_{j+k} - x_j \ge |\Delta_n|_{\min}$$

instead of $x_{j+k} - x_j \ge k\,|\Delta_n|_{\min}$ and obtain

$$\|Ds\|_\infty \le \frac{k\,(\|c\|_\infty + \|c\|_\infty)}{|\Delta_n|_{\min}} \cdot \left\| \sum_{j=1-k}^{n-1} N_{j,k-1} \right\|_\infty ,$$

where

$$\|c\|_\infty = \max\left\{|c_j| \,:\, j \in \{-k,\dots,n-1\}\right\}. \tag{4.7}$$

According to [dBoo73], there exists $d_k > 0$ depending only on k, such that

$$d_k^{-1}\,\|c\|_\infty \le \left\| \sum_{j=-k}^{n-1} c_j N_{j,k} \right\|_\infty \le \|c\|_\infty . \tag{4.8}$$

Rewriting the first inequality yields $\|c\|_\infty \le d_k$, because $\|s\|_\infty = 1$. Now we use the partition of the unity (4.1) to derive the estimate

$$\|Ds\|_\infty \le \frac{2k}{|\Delta_n|_{\min}} d_k.$$

Taking the supremum of all $s \in \mathcal{S}(\Delta_n, k)$ with $\|s\|_\infty = 1$ yields the result. □

Corollary 4.7. *For $l < k$, the differential operators $D^l : \mathcal{S}(\Delta_n, k) \to \mathcal{S}(\Delta_n, k - l)$ are bounded and*

$$\|D^l\|_{op} \le \left(\frac{2k}{|\Delta_n|_{\min}}\right)^l d_k.$$

Remark. The asymptotic behaviour of the constant d_k occuring in Lemma 4.6 is already characterized quite well in the literature. C. de Boor has conjectured that

$$d_k \sim 2^k$$

holds for all $k > 0$. In [Lyc78], T. Lyche has proved the lower bound

$$2^{-3/2}\frac{k-1}{k} \cdot 2^k \le d_k.$$

Finally, C. de Boor's conjecture was confirmed in the article [SS99] of Scherer and Shadrin up to a polynomial factor. There the authors have shown that the upper inequality

$$d_k \le k \cdot 2^k$$

holds for all $k > 0$. In our interest is the relation $d_k \to \infty$ if k tends to infinity.

Now we are able to estimate $\|D^r S_{\Delta_n,k} f\|_\infty$ in terms of the approximation error $\|f - S_{\Delta_n,k} f\|_\infty$.

Lemma 4.8. *For any $f \in C([0,1])$ there exists a constant M independent on f, such that*

$$\|D^r S_{\Delta_n,k} f\|_\infty \le M \cdot \|f - S_{\Delta_n,k} f\|_\infty .$$

Proof. We derive

$$\begin{aligned}\|D^r S_{\Delta_n,k} f\|_\infty &= \left\|D^r S_{\Delta_n,k} f - D^r S^2_{\Delta_n,k} f + D^r S^2_{\Delta_n,k} f - D^r S^3_{\Delta_n,k} f + \ldots\right\|_\infty \\ &\le \sum_{m=1}^{\infty} \left\|D^r S^m_{\Delta_n,k} (f - S_{\Delta_n,k} f)\right\|_\infty \\ &\le \|f - S_{\Delta_n,k} f\|_\infty \sum_{m=1}^{\infty} \left\|D^r S^m_{\Delta_n,k}\right\|_{op}\end{aligned}$$

$$= \|f - S_{\Delta_n,k} f\|_\infty \sum_{m=1}^{\infty} \left\|D^r(S^m_{\Delta_n,k} - L + L)\right\|_{op}$$
$$= \|f - S_{\Delta_n,k} f\|_\infty \sum_{m=1}^{\infty} \left\|D^r(S^m_{\Delta_n,k} - L)\right\|_{op},$$

as D^r annihilates linear functions and therefore, $D^r L = 0$. Then we obtain using Theorem 4.4 and Corollary 4.7

$$\begin{aligned}
\|D^r S_{\Delta_n,k} f\| &\leq \|f - S_{\Delta_n,k} f\|_\infty \|D^r\|_{op} \sum_{m=1}^{\infty} \left\|S^m_{\Delta_n,k} - L\right\|_{op} \\
&\leq \|f - S_{\Delta_n,k} f\|_\infty \|D^r\|_{op} \sum_{m=1}^{\infty} C\gamma^m_{\Delta_n,k} \\
&\leq \|f - S_{\Delta_n,k} f\|_\infty \|D^r\|_{op} \frac{C\gamma_{\Delta_n,k}}{1-\gamma_{\Delta_n,k}} \\
&\leq \frac{2^r k^r \gamma_{\Delta_n,k} d_k C}{|\Delta_n|^r_{\min}(1-\gamma_{\Delta_n,k})} \|f - S_{\Delta_n,k} f\|_\infty .
\end{aligned}$$

As $C \leq 1/\gamma_{\Delta_n,k}$, we get

$$\|D^r S_{\Delta_n,k} f\|_\infty \leq \frac{2^r k^r d_k}{|\Delta_n|^r_{\min}(1-\gamma_{\Delta_n,k})} \|f - S_{\Delta_n,k} f\|_\infty .$$

□

The following two theorems are the main result of this chapter. We will establish first the lower estimates in terms of moduli of smoothness and afterwards, we will show corresponding estimates using the K-functional. Note that both theorems are a direct consequence of the preceding lemma.

Theorem 4.9. *Let $f \in C([0,1])$ and $k > r \geq 2$. Then the spline approximation error can not be better than*

$$\frac{1}{2^{r+1}} \omega_r(f, t(\Delta_n, k)) \leq \|f - S_{\Delta_n,k} f\|_\infty ,$$

where

$$t(\Delta_n, k) = \frac{|\Delta_n|_{\min}}{k} \cdot \left(\frac{1-\gamma_{\Delta_n,k}}{d_k}\right)^{1/r}$$

given a fixed grid Δ_n and the degree k of the spline approximation. Moreover, we have that $t(\Delta_n, k) \to 0$ if the approximation error converges to zero.

Proof. The proof follows immediately by Lemma 4.8.

Applying inequality (4.6) for $0 < t \leq \frac{1}{r}$ yields

$$\omega_r(f,t) \leq 2^r \left(1 + \frac{k^r d_k}{|\Delta_n|^r_{\min}(1-\gamma_{\Delta_n,k})} t^r\right) \cdot \|f - S_{\Delta_n,k} f\|_\infty . \tag{4.9}$$

Setting

$$t(\Delta_n, k) := t = \frac{|\Delta_n|_{\min}}{k} \cdot \left(\frac{1 - \gamma_{\Delta_n,k}}{d_k}\right)^{1/r}$$

in (4.9) yields the first claim. If the spline approximation converges, then necessarily

$$\frac{|\Delta_n|_{\min}}{k} \to 0$$

holds by [Mar70] and we conclude that $t(\Delta_n, k) \to 0$ and the proof is complete. □

Using the same technique, we can state in the next theorem a uniform estimate for the K-functional in a similar way:

Theorem 4.10. *Let $f \in C([0,1])$ and $k > r \geq 2$. Then we have*

$$\frac{1}{2} \cdot K_r(f, t(\Delta_n, k)^r) \leq \|f - S_{\Delta_n,k} f\|_\infty$$

with

$$t(\Delta_n, k) := t = \frac{|\Delta_n|_{\min}}{k} \cdot \left(\frac{1 - \gamma_{\Delta_n,k}}{d_k}\right)^{1/r}.$$

Proof. Analogous to the proof of Theorem 4.9.

We obtain using the upper estimate for $\|D^r S_{\Delta_n,k} f\|_\infty$ of Lemma 4.8,

$$\begin{aligned} K_r(f, t^r) &\leq \|f - S_{\Delta_n,k}\|_\infty + t^r \|D^r S_{\Delta_n,k} f\|_\infty \\ &\leq \left(1 + \frac{2^r k^r d_k}{|\Delta_n|^r_{\min} (1 - \gamma_{\Delta_n,k})} t^r\right) \cdot \|f - S_{\Delta_n,k}\|_\infty . \end{aligned}$$

□

Finally, there is still one open question to answer. By definition of the constants, we have $d_k \to \infty$ for $k \to \infty$ and $|\Delta_n|_{\min} \to 0$ for $n \to \infty$ provided that the approximation error converges to zero. The question is whether the second largest eigenvalues of the operator can speed up the convergence in Theorem 4.9. As far as we know, the eigenvalues and eigenfunctions of the Schoenberg operator are still unknown. We conclude this chapter with the following conjecture that characterizes the behavior of the second largest eigenvalue of the Schoenberg operator.

Conjecture 4.1. Let $k > 0$ be fixed. Then

$$\gamma_{\Delta_n,k} \to 1, \quad \text{for } n \to \infty.$$

Let $n > 0$ be fixed. Then

$$\gamma_{\Delta_n,k} \to 1, \quad \text{for } k \to \infty.$$

CHAPTER 5 Curvature Estimation of Piecewise Smooth Curves

> *"In theory there is no difference between theory and practice. In practice there is."*
>
> YOGI BERRA

THE CURVATURE OF PLANAR CURVES provides a crucial tool for various applications where shape information matters. Typical problems are to match specific shapes or objects in segmented digital images or to try to understand the objects or the scene that is depicted in the image based on the shape information. The curvature profile of the contour of the shape is one of the most commonly used measures as every planar curve is uniquely determined by its curvature profile up to its orientation and translation.

This chapter is devoted to the curvature estimation of digitized planar curves as they appear in digital images in a precise mathematical way. The approximation method of our choice will be the Schoenberg's splines which we have introduced in Section 3.2. The variation-diminishing property combined with the ability to reproduce linear functions are the key to construct a convexity preserving curvature estimation that converges towards the real curvature. Even though this chapter focuses on splines due to their outstanding approximation properties, the proof for the convergence holds for general approximation operators provided that not only the curve but also their derivatives can be uniformly approximated. In this case, we can state a pointwise upper bound for the curvature approximation error. Furthermore, we are able to detect C^2-singularities using the established lower estimates of the approximation error by the second order modulus of smoothness from Chapter 4.

First, we will give an overview over some applications and related research to the estimation of the curvature. Next, we will cover the fundamentals of differential geometry to introduce the curvature of curves. We present the fundamental theorem of curves which states that each curve is uniquely defined by its curvature profile up to translation and orientation. In Section 5.3 we will deal with the problem to estimate the curvature given only discretely many points on the curve. A general proof is given for the convergence of the estimated curvature which gives necessary conditions on the used approximation operators. By the work of Goodman [Goo96] as shown in Section 2.3.4 we can state that variation-diminishing operators that are able to reproduce linear functions preserve convexity. We will show that spline approximation constructed according to Schoenberg provides a good curvature estimation that preserves the convexity of the shape. For the case of piecewise C^2-smooth curves, we present an algorithm in Section 5.4.4 that is able to detect the singularities based on lower estimates. We evaluate our algorithm numerically and compare our algorithm with the state of art estimators of recent literature at the end of this chapter. The results show that our spline based algorithm achieves competitive accuracy at significantly slower computation times.

5.1 Applications and related work

We will briefly refer to some recent applications of the curvature profile of shapes in digital images. Afterwards, we will give an overview over related research.

APPLICATIONS

Gardner et al. [GHJS05] have used the curvature for shape discrimination with a view to cancer diagnostics. More recently, Pasqualato et al. [Pas12] have proposed quantitative shape analysis based on curvature for cancer cells. Biological shape analysis by digital curvature has been discussed by Costa et al. [Cos04] and Castañón et al. [Cas07]. Another important application for shape classification is the detection of important points on the curve. In the comprehensive work of Tuytelaars and Mikolajczyk [TM07] several local invariant feature detectors have been surveyed where one of the presented features are points of high curvature. The significance of high curvature points for visual perception has been highlighted once more by Loncaric [Lon98] in their survey of shape analysis techniques. Recently, Yalim Keles and Tari [YT15] has proposed a robust method for scale independent detection of curvature-based criticalities and intersections in line drawings. Zhong and Ma have analyzed in [ZM10] the curvature scale-space and have evaluated the technique for corner detection and shape representation. A shape detection method based on curvature has been considered lately by Šukilović [Šuk15].

RELATED WORK

One way to measure the curvature in digital images is to measure the angular change of the tangent along the path of the curve. Bennett and Mac Donald [BM75] have discussed this approach with respect to the quantization noise that occurs in the discretization of the digital image. Based on the famous primal sketch model of the neuroscientist D. Marr, Asada and Brady [AB86] has introduced the curvature primal sketch that includes a set of parameterized curvature discontinuities. Medioni and Yasumoto where the first to use B-splines for corner detection based on the curvature [MY87]. Worring and Smeulders [WS93] have proposed several methods for curvature estimation in digital images. A multiscale, curvature-based shape representation for planar curves has been proposed in Mokhtarian and Mackworth [MM92], see also

Marcondes Cesar Jr. and Da Fontoura Costa [MD96] who have proposed effective shape representations based on curvature multiscale methods. The occuring numerical problems in the curvature estimation have been analyzed in Kovalevsky [Kov01]. Due to these numerical issues, Utcke [Utc03] has discussed error bounds. Recently, Hermann and Klette [HK07] evaluated in their comparative study several curvature estimators for two-dimensional curves in images. As result they recommended to use a B-spline based approximations for images with high resolution due to numerical experiments, while a quantitative result proving this result is still missing. Recently, De Vieilleville, Lachaud, and Feschet [DLF07], Roussillon and Lachaud [RL11], Coeurjolly, Lachaud, and Levallois [CLL14], and Levallois, Coeurjolly, and Lachaud [LCL14] have proved the multi-grid convergence of their estimators for convex shapes with techniques from differential geometry. We will compare our spline based curvature estimates with these estimators and confirm the observation made by Hermann and Klette in the numerical evaluations in Section 5.4.4.

5.2 Preliminaries

For the convenience of the reader, we will give here a short introduction to planar curves in order to define the curvature properly and to state the fundamental theorem for planar curves. For more details we refer to the classical book of do Carmo [dCar76].

Throughout this section, we consider $I = [a, b] \subset \mathbb{R}$ as an interval. We study planar curves $\alpha : I \to \mathbb{R}^2$ defined by two coordinate maps $x, y : I \to \mathbb{R}$ as $\alpha(t) = (x(t), y(t))$ and discuss their properties.

Definition 5.1 (Planar parametrized curves). A differentiable map $\alpha : I \to \mathbb{R}^2$ is said to be a *(parametrized) differentiable curve*. If the map α is k-times continuously differentiable, then α is called a (parametrized) planar C^k-curve.

If $\alpha : I \to \mathbb{R}^2$ is a parametrized planar curve, then the two corresponding coordinate maps $x : I \to \mathbb{R}$ and $y : I \to \mathbb{R}$ are differentiable. One important class of curves are those where the first derivative does not vanish on I.

Definition 5.2 (Regular curves). A parametrized differentiable curve $\alpha : I \to \mathbb{R}^2$ is said to be *regular* if $\dot{\alpha}(t) \neq 0$ for all $t \in I$.

For regular parametrized curves we can calculate its *arc length*. The arc length of the regular parametrized curve $\alpha : I \to \mathbb{R}^n$ is the integral value

$$L(\alpha) = \int_a^b |\dot{\alpha}(t)| \, \mathrm{d}t,$$

where

$$\|\dot{\alpha}(t)\| = \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}.$$

A special case is where $\|\dot{\alpha}(t)\| = 1$ for all $t \in I$. Then $L(\alpha) = b - a$. In that case we call the curve α *parametrized by arc length*.

Next, given a regular C^2-curve $\alpha : I \to \mathbb{R}^2$ we consider for any $t_0 \in I$ the Taylor expansion

$$\alpha(t) = \alpha(t_0) + (t - t_0)\dot{\alpha}(t_0) + \frac{(t-t_0)^2}{2}\ddot{\alpha}(t_0) + ((t-t_0)^2),$$

for $t \to t_0$. In the following we will discuss the meaning of $\dot{\alpha}(t_0)$ and $\ddot{\alpha}(t_0)$.

Definition 5.3. Let $\alpha : I \to \mathbb{R}^2$ be a parametrized curve. Then $T(t) = \dot{\alpha}(t)$ is called the *tangent* of α at the point t.

The length of the tangent vector measures the speed to travels along the curve. If the curve is parametrized by arc length, then $|\dot{\alpha}| = \|T(s)\| = 1$ holds. I.e., the curve travels along the curve with constant speed one.

The second derivative of α, i. e., the first derivative of the tangent vector, provides a perpendicular vector to the tangent vector.

Definition 5.4. Let $\alpha : I \to \mathbb{R}^2$ be a parametrized curve. Then $N(t) = \ddot{\alpha}(t)$ is called the *normal* of α at the point t. The unit normal vector will be denoted by $N_u(t) := N(t)\,\|N(t)\|$.

The length of the normal vector measures the acceleration, i. e., the rate of change of the tangent vector, to travel along the curve. Accordingly, we define the curvature as the signed length of the normal vector in the case where the curve is parametrized by arc length. The curvature measures how much the curve deviates from its tangent line.

Definition 5.5 (Curvature)**.** Let $\alpha : I \to \mathbb{R}^2$ be a curve parametrized by arc length s. Then we denote by $\kappa_\alpha(s)$ the *signed curvature* of α at the point $s \in I$, where

$$\ddot{\alpha}(s) = \kappa_\alpha(s)N_u(s).$$

Note. Straight lines have curvature 0*, while a circle of radius r has curvature* $1/r$*. The sign of the curvature provides information in which direction the curve bends. A positive sign means that the curves bends in direction of the normal vector, while a negative sign is obtained if the curve bends in the opposite direction of the normal vector.*

Proposition 5.1. *A planar curve is a line if and only if its curvature is* 0 *everywhere.*

The curvature of planar curves that are not parametrized by arc length can be calculated by a normalization:

Proposition 5.2. *Let* $\alpha : I \to \mathbb{R}^2$ *be a regular* C^2*-curve with* $\alpha(t) = (x(t), y(t))^T$ *for* $t \in I$*. Then the curvature of* α *at* $t \in I$ *is given by*

$$\kappa_\alpha(t) = \frac{\dot{\alpha}(t) \times \ddot{\alpha}(t)}{\|\dot{\alpha}(t)\|^3}.$$

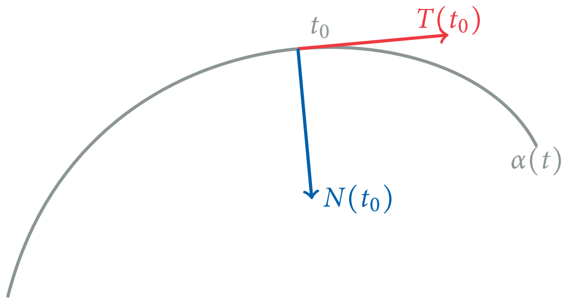


Figure 5.1: *Tangent and normal vector on a curve*

For the special case, where α is the graph of a function, the curvature term can be simplified in the following way.

Proposition 5.3. *For a function $f : I \to \mathbb{R}$, $f \in C^2(I)$, the curvature of $y = f(x)$ is given by*

$$\kappa_f(x) = \frac{D^2 f(x)}{(1 + |Df(x)|^2)^{\frac{3}{2}}},$$

where D represents the differential operator with respect to the x-coordinate.

Next, we state the fundamental theorem of curve theory in the Euclidean plane. This theorem states that a regular curve parametrized by arc length is entirely determined (up to orientation and translation) by its curvature.

Proposition 5.4 (Fundamental theorem of planar curves)**.** *Let $\tilde{\kappa} : I \to \mathbb{R}$ be a smooth function. Then there exists a planar curve α parametrized by arc length such the curvature of α equals $\tilde{\kappa}$, i. e., $\kappa_\alpha = \tilde{\kappa}$. This curve is unique up an orientation preserving isometry.*

Note that throughout the next sections we will choose w.l.o.g. the interval I as the unit interval $[0, 1]$.

5.3 Curvature estimation

Based on the fundamental theorem of planar curves, the curvature profile of a shape is an often used measure in pattern recognition to match corresponding shapes, as every planar curve is uniquely determined by its curvature profile up to its orientation and translation.

In the following, we want to approximate the curvature of a planar curve where only discrete samples are available. We want an approximation method, where the approximation does not oscillate more often about any straight line than the function to be approximated. This criterion is important, as the curvature is defined as the deviation from a straight line, namely the tangent line. This kind of approximation is called shape preserving approximations, see Section 2.3.4. There, it has been shown by Karlin [Kar68] and Goodman [Goo96] that variation diminishing operators that can reproduce constants and linear functions are able to preserve the positivity, the monotonicity, and the convexity of curves. Karlin has shown this result for smooth functions $f : [0,1] \to \mathbb{R}$, whereas this result has been extended by [Goo96] to smooth curves $\alpha : [0,1] \to \mathbb{R}^2$.

Therefore, we are considering variation diminishing operators that can reproduce linear functions. Two famous operators of this kind have been introduced and discussed in Chapter 3, the Bernstein operator and Schoenberg's spline operator. We will provide here a general framework for variation diminishing approximation operators that ensures the convergence of the curvature estimate towards the real curvature. Concretely, we will prove that the variation diminishing splines are admissible and are able to preserve the convexity of the sampled curve. We are choosing the splines here due to their local flexibility and the non-uniform spacing of the knots.

5.3.1 Uniform convergence of curvature approximations

Consider $\alpha : [0,1] \to \mathbb{R}^2$ as a regular curve where $\alpha(t) = (x(t), y(t))^T$ for $t \in [0,1]$. As shown in the last section, the curvature $\kappa(\alpha; t)$ at the point $\alpha(t)$ can be analytically computed by

$$\kappa(\alpha; t) = \frac{\dot{\alpha}(t) \times \ddot{\alpha}(t)}{\|\dot{\alpha}(t)\|^3}, \tag{5.1}$$

provided that α is twice differentiable.

Next, we consider a sequence of smooth C^2-approximation $\tilde{\alpha}_n : [0,1] \to \mathbb{R}^2$ of the given curve α where only n points of α are evaluated. Using this approximation $\tilde{\alpha}_n$, we can define an n-term curvature estimate of the curve α by

$$\tilde{\kappa}_n(\alpha; t) = \frac{\dot{\tilde{\alpha}}_n(t) \times \ddot{\tilde{\alpha}}_n(t)}{\left\|\dot{\tilde{\alpha}}_n(t)\right\|^3}, \tag{5.2}$$

We will prove in this section under additional assumptions on $\tilde{\alpha}_n$ that $\kappa_n(\alpha) \to \kappa(\alpha)$ for $n \to \infty$. First, we state the following pointwise estimate of the curvature for $n \to \infty$.

Theorem 5.5. *Let $\alpha : I \to \mathbb{R}^2$ be a regular planar C^2-curve and consider the C^2-smooth approximation $\tilde{\alpha}_n$. Suppose there exists constants $C_3, C_4 > 0$ independent on n such that*

$$\left\|\dot{\tilde{\alpha}}_n\right\|_\infty \le C_3 \|\dot{\alpha}\|_\infty, \qquad \left\|\ddot{\tilde{\alpha}}_n\right\|_\infty \le C_4 \|\ddot{\alpha}\|_\infty$$

holds and

$$\lim_{n\to\infty} \|\alpha(t) - \tilde{\alpha}_n(t)\|_\infty = 0, \qquad \lim_{n\to\infty} \left\|\dot{\alpha}(t) - \dot{\tilde{\alpha}}_n(t)\right\|_\infty = 0.$$

Then there exists constants $C_1, C_2 > 0$ independent on $\tilde{\alpha}_n$ such that

$$|\kappa(\alpha; t) - \tilde{\kappa}_n(\alpha; t)| \leq C_1 \left\|\dot{\alpha}(t) - \dot{\tilde{\alpha}}_n(t)\right\|_1 + C_2 \left\|\ddot{\alpha}(t) - \ddot{\tilde{\alpha}}_n(t)\right\|_1$$

for $n \to \infty$ and all $t \in [0,1]$.

Proof. We start adding a mixed term and use the triangle inequality:

$$\begin{aligned}
&\left|\frac{\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)}{\|\dot{\alpha}\|^3} - \frac{\dot{\tilde{x}}_n(t)\ddot{\tilde{y}}_n(t) - \ddot{\tilde{x}}_n(t)\dot{\tilde{y}}_n(t)}{\|\dot{\tilde{\alpha}}_n\|^3}\right| \\
&\quad = \left|\frac{\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)}{\|\dot{\alpha}\|^3} - \frac{\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)}{\|\dot{\tilde{\alpha}}_n\|^3}\right. \\
&\qquad \left. + \frac{\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)}{\|\dot{\tilde{\alpha}}_n\|^3} - \frac{\dot{\tilde{x}}_n(t)\ddot{\tilde{y}}_n(t) - \ddot{\tilde{x}}_n(t)\dot{\tilde{y}}_n(t)}{\|\dot{\tilde{\alpha}}_n\|^3}\right| \\
&\quad \leq \left|\frac{\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)}{\|\dot{\alpha}\|^3} - \frac{\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)}{\|\dot{\tilde{\alpha}}_n\|^3}\right| \\
&\qquad + \left|\frac{\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)}{\|\dot{\tilde{\alpha}}_n\|^3} - \frac{\dot{\tilde{x}}_n(t)\ddot{\tilde{y}}_n(t) - \ddot{\tilde{x}}_n(t)\dot{\tilde{y}}_n(t)}{\|\dot{\tilde{\alpha}}_n\|^3}\right| \\
&\quad \leq |\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)| \left|\frac{1}{\|\dot{\alpha}\|^3} - \frac{1}{\|\dot{\tilde{\alpha}}_n\|^3}\right| \\
&\qquad + \frac{1}{\|\dot{\tilde{\alpha}}_n\|^3} \left|\dot{x}(t)\ddot{y}(t) - \dot{\tilde{x}}_n(t)\ddot{\tilde{y}}_n(t) + \ddot{\tilde{x}}_n(t)\dot{\tilde{y}}_n(t) - \ddot{x}(t)\dot{y}(t)\right| \qquad (5.3)
\end{aligned}$$

In the following, we will estimate the first and the last term by the approximation error of the first and second order derivative. First, we estimate the last term by adding missing mixed terms and use the triangle inequality once more to obtain:

$$\begin{aligned}
&\left|\dot{x}(t)\ddot{y}(t) - \dot{\tilde{x}}_n(t)\ddot{\tilde{y}}_n(t) + \ddot{\tilde{x}}_n(t)\dot{\tilde{y}}_n(t) - \ddot{x}(t)\dot{y}(t)\right| \\
&\quad = \left|\dot{x}(t)\ddot{y}(t) + \dot{\tilde{x}}_n(t)\ddot{y}(t) - \dot{\tilde{x}}_n(t)\ddot{y}(t) - \dot{\tilde{x}}_n(t)\ddot{\tilde{y}}_n(t)\right. \\
&\qquad \left. + \ddot{\tilde{x}}_n(t)\dot{\tilde{y}}_n(t) - \ddot{\tilde{x}}_n(t)\dot{y}(t) + \ddot{\tilde{x}}_n(t)\dot{y}(t) - \ddot{x}(t)\dot{y}(t)\right| \\
&\quad = \left|\ddot{y}(t) \cdot [\dot{x}(t) - \dot{\tilde{x}}_n(t)] + \dot{\tilde{x}}_n(t) \cdot [\ddot{y}(t) - \ddot{\tilde{y}}_n(t)]\right. \\
&\qquad \left. + \ddot{\tilde{x}}_n(t) \cdot [\dot{\tilde{y}}_n(t) - \dot{y}(t)] + \dot{y}(t) \cdot [\ddot{\tilde{x}}_n(t) - \ddot{x}(t)]\right| \\
&\quad \leq \left|\dot{\tilde{x}}_n(t)\right| \left|\ddot{y}(t) - \ddot{\tilde{y}}_n(t)\right| + \left|\ddot{\tilde{x}}_n(t)\right| \left|\dot{y}(t) - \dot{\tilde{y}}_n(t)\right| \\
&\qquad + |\dot{y}(t)| \left|\ddot{x}(t) - \ddot{\tilde{x}}_n(t)\right| + |\ddot{y}(t)| \left|\dot{x}(t) - \dot{\tilde{x}}_n(t)\right|
\end{aligned}$$

Note that the first and the second derivative of $x(t)$ and $y(t)$ respectively are bounded on $[0,1]$. Then, with $\tilde{C}_1 := \max\{C_4 \|\ddot{x}\|_\infty, \|\ddot{y}\|_\infty\}$ and $\tilde{C}_2 := \max\{C_3 \|\dot{x}\|_\infty, \|\dot{y}\|_\infty\}$ we obtain the estimate

$$\left|\dot{x}(t)\ddot{y}(t) - \dot{\tilde{x}}_n(t)\ddot{\tilde{y}}_n(t) + \ddot{\tilde{x}}_n(t)\dot{\tilde{y}}_n(t) - \ddot{x}(t)\dot{y}(t)\right|$$

$$\leq \tilde{C}_1 \left\| \dot{\alpha}(t) - \dot{\tilde{\alpha}}_n(t) \right\|_1 + \tilde{C}_2 \left\| \ddot{\alpha}(t) - \ddot{\tilde{\alpha}}_n(t) \right\|_1 .$$

Next, we estimate the first line of (5.3) using a Taylor expansion. Concretely, we use the Taylor expansion for the function $g(x, y) = (x^2 + y^2)^{-3/2}$ evaluated at the points $\dot{\alpha}(t) = (\dot{x}(t), \dot{y}(t))$ and $\dot{\tilde{\alpha}}_n(t) = (\dot{\tilde{x}}_n(t), \dot{\tilde{y}}_n(t))$. A short calculation shows that

$$\frac{\partial}{\partial x} g(x, y) = \frac{-3x}{(x^2 + y^2)^{5/2}} \quad \text{and} \quad \frac{\partial}{\partial y} g(x, y) = \frac{-3y}{(x^2 + y^2)^{5/2}}.$$

Then we obtain for $n \to \infty$:

$$\begin{aligned} \left| \|\dot{\alpha}\|^{-3} - \left\|\dot{\tilde{\alpha}}_n\right\|^{-3} \right| &= \left| g(\dot{x}(t), \dot{y}(t)) - g(\dot{\tilde{x}}_n(t), \dot{\tilde{y}}_n(t)) \right| \\ &= \left| \frac{-3 \cdot \dot{\tilde{\alpha}}_n(t)^T}{\left(\dot{\tilde{x}}_n(t)^2 + \dot{\tilde{y}}_n(t)^2\right)^{5/2}} (\dot{\alpha}(t) - \dot{\tilde{\alpha}}_n(t)) + o\left(\left\|\dot{\alpha}(t) - \dot{\tilde{\alpha}}_n(t)\right\|_1\right) \right| \\ &\leq \tilde{C}_3 \cdot \left\|\dot{\alpha}(t) - \dot{\tilde{\alpha}}_n(t)\right\|_1 , \end{aligned}$$

where

$$\tilde{C}_3 = \frac{3 \cdot \max\left\{ \left\|\dot{\tilde{x}}_n(t)\right\|_\infty , \left\|\dot{\tilde{y}}_n(t)\right\|_\infty \right\}}{\left\| \left(\dot{\tilde{x}}_n(t)^2 + \dot{\tilde{y}}_n(t)^2\right)^{5/2} \right\|_\infty}.$$

As $|\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)|$ is bounded on $[0,1]$, we get the final result

$$\begin{aligned} &\left| \frac{\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)}{\|\dot{\alpha}\|^3} - \frac{\dot{\tilde{x}}_n(t)\ddot{\tilde{y}}_n(t) - \ddot{\tilde{x}}_n(t)\dot{\tilde{y}}_n(t)}{\left\|\dot{\tilde{\alpha}}_n\right\|^3} \right| \\ &\qquad \leq C_1 \left\|\dot{\alpha}(t) - \dot{\tilde{\alpha}}_n(t)\right\|_1 + C_2 \left\|\ddot{\alpha}(t) - \ddot{\tilde{\alpha}}_n(t)\right\|_1 . \end{aligned}$$

□

Corollary 5.6. *Let $\alpha : [0,1] \to \mathbb{R}^2$ be a regular, twice differentiable planar curve and let $(\tilde{\alpha}_n)_{n \in \mathbb{N}}$ be a sequence of C^2-smooth approximations of α such that*

- $\tilde{\alpha}_n - \alpha(t) \to 0$ *uniformly as* $n \to \infty$,
- $\dot{\tilde{\alpha}}_n - \dot{\alpha}(t) \to 0$ *uniformly as* $n \to \infty$,
- $\ddot{\tilde{\alpha}}_n - \ddot{\alpha}(t) \to 0$ *uniformly as* $n \to \infty$.

and their exists constants $C_1, C_2 > 0$ independent of n such that

- $\left\|\dot{\tilde{\alpha}}_n\right\|_\infty \leq C_1 \|\dot{\alpha}\|_\infty$,
- $\left\|\ddot{\tilde{\alpha}}_n\right\|_\infty \leq C_2 \|\ddot{\alpha}\|_\infty$.

Then for $n \to \infty$ the approximation of the curvature converges uniformly, i. e.,

$$\|\kappa(\alpha; t) - \tilde{\kappa}_n(\alpha; t)\|_\infty \to 0, \qquad \textit{for all } t \in [0,1] .$$

Note that the statements of Theorem 5.5 and Corollary 5.6 hold for all smooth approximations that satisfy the stated conditions. The rate of convergence depends on the rate of convergence of the first and the second derivatives.

5.3.2 Variation diminishing curvature estimation

In the following, we will consider variation diminishing approximation methods that guarantee preserving of convexity of curves in order to provide a meaningful curvature estimate. To this end, we consider as in Section 2.3.4 a normalized totally positive basis $\{e_1, \dots, e_n\}$ and the corresponding sequence of approximation operator $T_n : C([0,1]) \to C([0,1])$,

$$T_n(f;x) = \sum_{i=1}^{n} f(x_i) e_i(x), \qquad t \in [0,1],$$

where the evaluation points $0 \le x_1 < x_2 < \cdots < x_n \le 1$ are given. Furthermore, we assume that this operator can reproduce linear functions, i. e.,

$$\sum_{i=1}^{n} x_i e_i(x) = x$$

holds. Due to the normalization of the basis, the basis functions do also form a partition of unity. Thus, the operator T_n is able to reproduce all linear functions. Recall that by Proposition 2.11 these properties already lead to the interpolation of start and endpoints, i.e.,

$$T_n(f;0) = f(0) \quad \text{and} \quad T_n(f;1) = f(1).$$

Examples for this kind of operators are given in Chapter 3, e. g., the Schoenberg operator or the Bernstein operator. According to Proposition 2.7, the operators T_n are variation diminishing as the basis is totally positive and by Proposition 2.8 and Proposition 2.9 the monotonicity and convexity is preserved under T_n. Recall that these results are due to Karlin [Kar68]

In order to approximate the curve $\alpha : [0,1] \to \mathbb{R}^2$ we extend the operator T_n by applying it to every coordinate of $\alpha(t) = (x(t), y(t))^T$. Accordingly, for given evaluation points $0 \le t_1 < t_2 < \cdots < t_n \le 1$ we consider the operator T_n applied to the curve α as

$$T_n(\alpha;t) = \sum_{j=1}^{n} \begin{pmatrix} x(t_j) \\ y(t_j) \end{pmatrix} e_j(t), \qquad t \in [0,1]. \tag{5.4}$$

The approximation is the linear combination of the basis functions weighted with discretely many samples of the curve α. By the work of Goodman [Goo96], the shape preserving properties in the 1D-case can be transferred to the approximation of curves. Proposition 2.15 and its corollaries state that T_n preserves the monotonicity and the convexity of α.

In the following, we will show that the curvature approximation by the operator T_n converges under some additional assumptions.

Corollary 5.7. *Let $\alpha : [0,1] \to \mathbb{R}^2$ be a regular, twice differentiable planar curve and let $(T_n)_{n\in\mathbb{N}}$ be a sequence of C^2-smooth variation diminishing operators as in (5.4) that satisfy the conditions of Corollary 5.6 of the form (5.4). Let us denote by $\tilde{\alpha}_n(t) := T_n(\alpha; t)$. If*

- $\tilde{\alpha}_n(t) - \alpha(t) \to 0$ *uniformly on* $[0,1]$ *as* $n \to \infty$,
- $\dot{\tilde{\alpha}}_n(t) - \dot{\alpha}(t) \to 0$ *uniformly on* $[0,1]$ *as* $n \to \infty$,
- $\ddot{\tilde{\alpha}}_n(t) - \ddot{\alpha}(t) \to 0$ *uniformly on* $[0,1]$ *as* $n \to \infty$.

and their exists constants $C_1, C_2 > 0$ independent of n such that

- $\left\|\dot{\tilde{\alpha}}_n\right\|_\infty \le C_1 \left\|\dot{\alpha}\right\|_\infty$,
- $\left\|\ddot{\tilde{\alpha}}_n\right\|_\infty \le C_2 \left\|\ddot{\alpha}\right\|_\infty$.

Then for $n \to \infty$ the approximation of the curvature converges uniformly, i. e.,

$$\left\|\kappa(\alpha; t) - \tilde{\kappa}_n(\alpha; t)\right\|_\infty \to 0, \qquad \text{for all } t \in [0,1].$$

By the variation diminishing property and the ability to reproduce linear functions the approximations $\tilde{\alpha}_n(t)$ preserve positivity, monotonicity, and the convexity of $\alpha(t)$ for every integer n.

Also note that in this case the approximation error of the curvature is zero for constants and lines, as the operators T_n can preserve constants and linear functions.

5.4 Curvature estimation with splines

Having established the convergence results for general variation diminishing operators we will now show that the Schoenberg operator satisfies all the conditions of Corollary 5.7 and thus, provides a linear approximation scheme to estimate the curvature of finite number of samples of a smooth curve. We want to highlight the linearity here, as this will lead to a fast algorithm which is of importance if big data is considered. We will sketch the spline based multi-scale algorithm in Section 5.4.2 where the decay rate of the error is used to weight different scales. This is due to the established lower estimates of Chapter 4 which allow us to interpret the approximation error as local smoothness. Accordingly, for smooth regions coarser scales will be preferred where in more detailed regions finer scales are used.

We will evaluate our algorithm numerically in Section 5.4.3 for discrete points evaluations of a smooth curve. As the theory only guarantees the uniform convergence for point evaluations, the convergence of the curvature estimation of digitized curves is not guaranteed. We will overcome this issue in first to approximate the discrete digitized values by a smoothing spline and then we apply our algorithm to the same number of samples of the smoothing spline. We will numerically evaluate this method in Section 5.4.4 and compare our algorithm to recent curvature estimator. We will show that while we achieve comparable accuracy of the curvature estimation we outperform the other methods in the running time.

As highlight we will demonstrate in Section 5.4.4 how our multi-scale algorithm gives us the possibility to detect C^2-singularities. Thus, the algorithm will first detect the singularities for a piecewise smooth curve and then estimates the curvature on each smooth patch. As far as we know this spline based multi-scale algorithm is the first methods that is able to estimate the curvature of piecewise smooth curves. Therefore, our algorithm is not only superior to the others compared to the running time but also to the wider class of curves that can be considered.

5.4.1 Uniform convergence

Using the results of the preceding section, we will consider here the Schoenberg operator and show that all necessary condition of Corollary 5.7 are satisfied. This proves that the variation diminishing splines provide a well behaved framework to estimate the curvature of curves by only discretely given points. Recall, that the Schoenberg operator is given for any continuous function f on $[0,1]$ by

$$S_{\Delta_n,k} f(x) = \sum_{j=-k}^{n-1} f(\xi_{j,k}) N_{j,k}(x), \quad x \in [0,1],$$

where $\xi_{j,k}$ are the Greville nodes and $N_{j,k}$ are the normalized B-splines. See Section 3.2.1 on page 33 for details. In fact, it has been shown in the work of Schoenberg [Sch67] that $S_{\Delta_n,k}$ is variation diminishing and is able to preserve constants and linear functions. The uniform convergence of the first two derivatives and the operator itself has already been shown by Marsden [Mar70], see Proposition 3.14 in Chapter 3. The convexity preserving property has also been proved in [Mar70], see Proposition 3.15. Therefore, the only two missing conditions on $S_{\Delta_n,k}$ are the boundedness of the first and the second order derivative. We will show these properties in the next lemma.

Lemma 5.8. *The derivatives of the spline approximation are bounded by*

$$\|D\, S_{\Delta_n,k}(f;x)\|_\infty \le \|Df(x)\|_\infty$$

and

$$\left\|D^2\, S_{\Delta_n,k}(f;x)\right\|_\infty \le \frac{|\Delta_n|_{\max}}{|\Delta_n|_{\min}} \cdot \left\|D^2 f(x)\right\|_\infty .$$

Proof. Marsden [Mar70] has provided an explicit form for the derivatives of Schoenberg's operator applied to a continuous function. Using [Mar70, Lemma 1 on p. 32], the first derivative can be represented in the form

$$D\, S_{\Delta_n,k}(f;x) = \sum_{j=1-k}^{n-1} Df(\eta_j^{(1)}) N_{j,k-1}(x),$$

where $\eta_j^{(1)} \in \left(\xi_{j-1}, \xi_j\right)$. Similarly, we can write the second derivative by [Mar70, Lemma 2 on p. 35] as

$$D^2 S_{\Delta_n,k}(f;x) = \sum_{j=2-k}^{n-1} D^2 f(\eta_j^{(2)}) \frac{\xi_{j,k} - \xi_{j-2,k}}{2(\xi_{j,k-1} - \xi_{j-1,k-1})} N_{j,k-2}(x),$$

with $\eta_j^{(2)} \in \left(\xi_{j-1}, \xi_j\right)$. We will use both representation to directly calculate the bounds. Accordingly, the stated upper bounds for the first derivative follow by

$$\begin{aligned} \|D S_{\Delta_n,k}(f;x)\|_\infty &= \left\| \sum_{j=1-k}^{n-1} Df\left(\eta_j^{(1)}\right) N_{j,k-1}(x) \right\|_\infty \\ &\le \|Df(x)\|_\infty \cdot \left\| \sum_{j=1-k}^{n-1} N_{j,k-1}(x) \right\|_\infty = \|Df(x)\|_\infty \end{aligned}$$

and for the second derivative by

$$\begin{aligned} \left\|D^2 S_{\Delta_n,k}(f;x)\right\|_\infty &= \left\| \sum_{j=2-k}^{n-1} D^2 f\left(\eta_j^{(2)}\right) \frac{\xi_{j,k} - \xi_{j-2,k}}{2(\xi_{j,k-1} - \xi_{j-1,k-1})} N_{j,k-2}(x) \right\|_\infty \\ &\le \left\|D^2 f(x)\right\|_\infty \left\| \sum_{j=2-k}^{n-1} \frac{\xi_{j,k} - \xi_{j-2,k}}{2(\xi_{j,k-1} - \xi_{j-1,k-1})} N_{j,k-2}(x) \right\|_\infty \\ &= \left\|D^2 f(x)\right\|_\infty \left\| \sum_{j=2-k}^{n-1} \frac{(k-1)(x_{j+k} - x_j + x_{j+k-1} - x_{j-1})}{2k(x_{j+k-1} - x_j)} N_{j,k-2}(x) \right\|_\infty \\ &\le \left\|D^2 f(x)\right\|_\infty \left\| \sum_{j=2-k}^{n-1} \frac{(k-1) \cdot 2k \cdot |\Delta_n|_{\max}}{2k \cdot |\Delta_n|_{\min}} N_{j,k-2}(x) \right\|_\infty \\ &= (k-1) \cdot \frac{|\Delta_n|_{\max}}{|\Delta_n|_{\min}} \cdot \left\|D^2 f(x)\right\|_\infty . \end{aligned}$$

□

Note that we need a bound that holds independently on n. For the uniform Schoenberg operator the ratio $|\Delta_n|_{\max} / |\Delta_n|_{\min} = 1$. We assume for the nonuniform case that $|\Delta_n|_{\max} / |\Delta_n|_{\min} \le M$ for some M independent on n. In this case, we can state the following theorem.

Theorem 5.9. *Let $\alpha : [0,1] \to \mathbb{R}^2$ be a regular, twice differentiable planar curve and let us consider the approximations $\tilde{\alpha}_n(t) := S_{\Delta_n,k}(\alpha;t)$ for $k > 2$. If there exist a constant M independent on n such that $|\Delta_n|_{\max} / |\Delta_n|_{\min} \le M$ holds, then the approximation of the curvature converges uniformly for fixed k and $n \to \infty$, i. e.,*

$$\|\kappa(\alpha;t) - \tilde{\kappa}_n(\alpha;t)\|_\infty \to 0, \qquad \textit{for all } t \in [0,1].$$

Besides, the approximations $\tilde{\alpha}_n(t)$ preserve the positivity, monotonicity, and the convexity of $\alpha(t)$ for every integer n.

Proof. Follows immediately by Corollary 5.7 using Proposition 3.14 and Lemma 5.8. □

Note that the condition $|\Delta_n|_{\max} / |\Delta_n|_{\min} \leq M$ holds if we consider uniformly spaced samples of the curve α.

5.4.2 Algorithm

In the following we describe the algorithm to estimate the curvature of a piecewise smooth curve $\alpha : [0,1] \to \mathbb{R}^2$, $\alpha(t) = (x(t), y(t))^T$. Thereby, we assume that we have given a discrete set of points

$$\mathcal{P}_n = \left\{P_i \,:\, P_i = (x_i, y_i) \in \mathbb{R}^2, i = 1, \dots, n\right\}$$

that are sampled from the curve α, i. e., there exists $t_1 < \cdots < t_n \in [0,1]$ such that $\alpha(t_i) = P_i$ for all $i \in \{1, \dots, n\}$. To be able to consider different scales of the curve approximation, we assume further that we have given interleaving point sets

$$\mathcal{P}_n^{(s_1)} \subset \mathcal{P}_n^{(s_2)} \subset \cdots \subset \mathcal{P}_n^{(s_l)} \subset P_i$$

that construct the approximation $\tilde{\alpha}_n^{(s_k)}(t)$ at scale s_j or $j \in \{1, \dots, l\}$. Besides we assume that the scale numbers

$$1 \leq s_1 < \cdots < s_l \leq n$$

are equal to the number of points corresponding to their points sets. Then the algorithm works as follows.

Step 1: Compute the cubic spline approximation $\tilde{\alpha}_n, \tilde{\alpha}_n^{(s_1)}, \dots, \tilde{\alpha}_n^{(s_l)}$ by

$$\tilde{\alpha}_n(t) = S_{\Delta_n,k}\left(\alpha(\mathcal{P}_n); t\right), \qquad \tilde{\alpha}_n^{s_j}(t) = S_{\Delta_n,k}\left(\alpha(\mathcal{P}_n^{(s_j)}); t\right)$$

as shown in (5.4), where $\alpha_n(\mathcal{P}_n)$ and $\alpha_n(\mathcal{P}_n^{(s_j)})$ means that the approximation operator $S_{\Delta_n,k}$ is using only the points of the point set $\mathcal{P}_n$, respectively $\mathcal{P}_n^{(s_j)}$.

Step 2: For each scale s_j we compute the corresponding curvature estimate

$$\tilde{\kappa}_n^{(s_j)}(\alpha; t) := \kappa_n(\alpha(\mathcal{P}_n^{(s_j)}); t)$$

by (5.2).

curve	parametrization	parameters	$\kappa_{\min}$	$\kappa_{\max}$
Ellipse	$R_\phi(a\cos(t), b\sin(t))$	$a = 30$, $b = 20$, $\phi = 0.5$	0.0444	0.1500
Flower	$(\rho(t)\cdot\cos(t), \rho(t)\cdot\sin(t))$, $\rho(t) = r_1 + r_2\cdot\cos(k\cdot t)$	$r_1 = 20$, $r_2 = 5$, $k = 5$	−0.575	0.1812

Table 5.1: *The ellipse and the flower.*
Both curves are given in explicit form with their parameters and their minimal and maximal curvature. An illustration is shown in Figure 5.3 and Figure 5.4.

Step 3: For each scale s_j we measure the absolute approximation error at each point P_i corresponding to $t_i \in [0,1]$:

$$\varepsilon_j(t_i) := \left|\tilde{\alpha}_n^{s_j}(t_i) - P_i\right|,$$

where $P_i \in \mathcal{P}_n^{s_l}$. With these error numbers we define the discrete decay rates

$$d_{j-1}^{j}(t_i) := \left|\varepsilon_{s_{j-1}}(t_i) - \varepsilon_{s_j}(t_i)\right|, \qquad j \in \{2, \ldots, l\}$$

that describe the change of the error between the coarse scale s_{j-1} and the finer scale s_j.

Step 4: If the discrete decay rates

$$d_1^2(t_i), \ldots, d_{l-1}^{l}(t_i)$$

of the multi-scale errors $\varepsilon_1(t_i), \ldots, \varepsilon_l(t_i)$ are slow then the point P_i is classified as a C^2-singularity. Concretely, this is done by looking for consisting local maxima in the decay rates that are above a certain threshold.

Step 5: For the final curvature estimate at the point t_i we weight the curvature estimate at scale s_j with the corresponding approximation error and the decay rate to the next scale s_{j+1}:

$$\tilde{\kappa}_n(\alpha; t_i) := \frac{\sum_{j=1}^{n-1}\left(\varepsilon_j(t_i)\cdot d_j^{j+1}(t_i)\right)^{-1}\kappa_n^{(s_j)}(t_i)}{\sum_{j=1}^{n-1}\left(\varepsilon_j(t_i)\cdot d_j^{j+1}(t_i)\right)^{-1}}.$$

5.4.3 Numerical evaluation for samples of curves

We have implemented the algorithm in Matlab using the spline toolbox. For the evaluation of the accuracy, we have tested the algorithm for a rotated ellipse with great axis 30 and small axis 20 and a flower with k petals where an explicit parametrization as well as the explicit curvature is available. The explicit form and the used parameters for the ellipse and the flower are listed in Table 5.1. Both curves can be seen in digitized form in Figure 5.3 and Figure 5.4 together with their curvature profiles.

For both evaluations we have used the scales

$$s_1 = 8, s_2 = 4, s_3 = 2, s_4 = 1,$$

where these numbers denote the sub-sampling factors. We have evaluated both curves at $n \in [100, 20000]$ points $t_1, \ldots, t_n$ and we have computed for each n the curvature estimate $\tilde{\kappa}_n(\alpha; t_i)$. To demonstrate the convergence and to numerically estimate the convergence rate, we have measured the maximal and the average absolute error defined by

$$\varepsilon_{\text{abs}}(n) = \|\kappa(\alpha; t_i) - \kappa_n(\alpha; t_i)\|$$

and the maximal and the average relative error

$$\varepsilon_{\text{rel}}(n) = \frac{\varepsilon_{\text{abs}}(n)}{\|\kappa(\alpha; t_i)\|}.$$

The reported results for the absolute errors are plotted in Figure 5.2. We have chosen to show the absolute error over the relative error in order to better visualize the gap between the maximal and the average error. However, the measurements of both curves suggest a decay rate of the error of the curvature estimation of $\mathcal{O}(n^{-2})$ which matches the approximation rate of the spline approximation for smooth functions. Therefore, the measured results confirm the results that we have proved in theory in the preceding sections. Also note that the estimation error of the ellipse is much smaller than the curvature estimation error of the flower.

These promising results will motivate us to study the algorithm also for digitized curves, even though the uniform convergence has only been proved for point evaluations. We will curves that are digitized according to the Gauss scheme in the next section.

5.4.4 Numerical evaluation for digitized contours

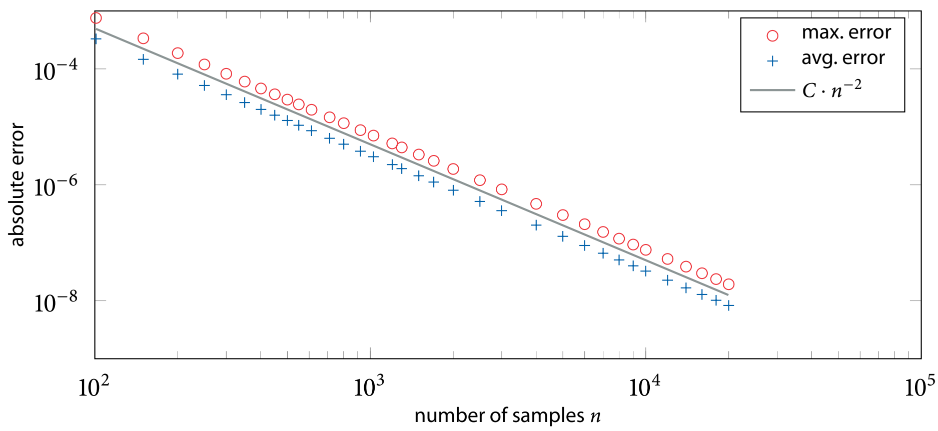
We consider now contours of shapes that have been digitized according to the Gauss scheme.

Definition 5.6 (Gauss digitization)**.** Let $X \in \mathbb{R}^d$ be a compact subset with smooth boundary δX. The *Gauss digitization* of X at grid size h is defined as

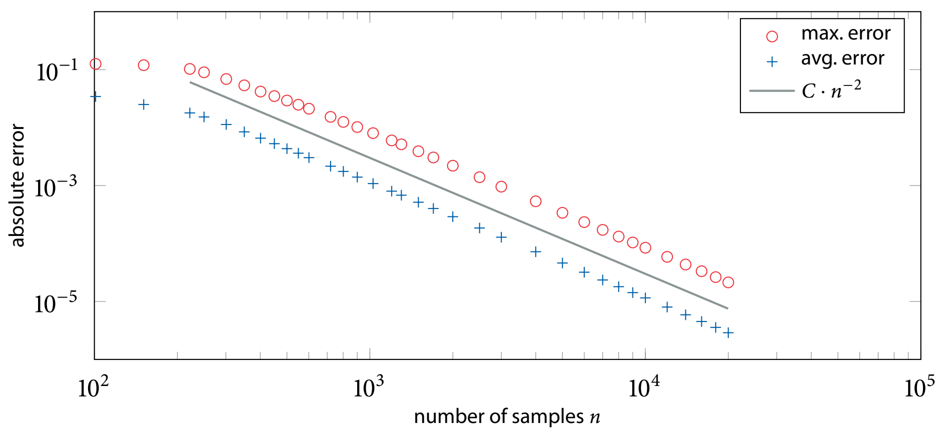
$$D_h(X) = \left(\frac{1}{h} \cdot X\right) \cap \mathbb{Z}^d,$$

where $h^{-1} \cdot X$ is the uniform scaling of X by the factor h^{-1}. We will denote by $\alpha_h(X; t)$ the corresponding digitized contour.

In the following, we will consider digitizations of the ellipse and the flower using the same parameters as in the preceding section, see also Table 5.1. Both shapes have been digitized using the Gauss scheme at three different grid steps $h \in \{1, 0.1, 0.01\}$.



(a) ellipse



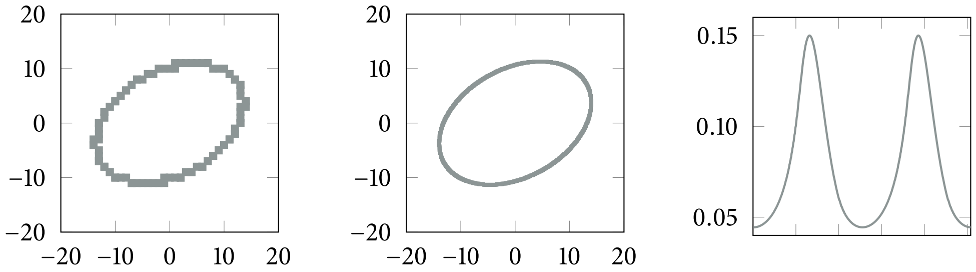
(b) flower

Figure 5.2: *The absolute maximal and average error of the curvature estimation of an ellipse and a flower at different number of point evaluations* $n \in [100, 20000]$. *The sub-sampling factors for the multi-scale approach are set to* 1, 2, 4, 8. *We have plotted the curve* n^{-2} *to illustrate the quadratic decay rate of the error. Note that the error for the ellipse is much smaller compared to the error made for the flower.*

The digitized shapes are illustrated in Figure 5.3 and Figure 5.4 for $h = 1$ and $h = 0.1$ with their curvature profile.

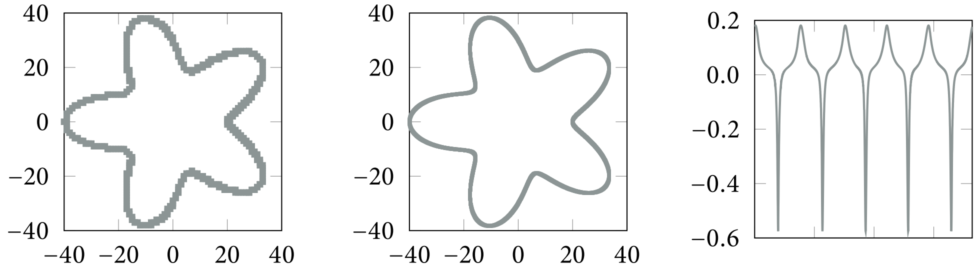
We compare our algorithm to the BC estimator [MBF08; EMC11] based on a convolution with a binomial kernel and the MDCA estimator [RL11] based on the set of maximal digital circular arcs. For the evaluation we use the available C++ implementation of the open source library DGtal [1]. Note that our algorithm has been implemented in Matlab using the spline toolbox. Thus, the run times are not directly comparable, as the computation times in Matlab are usually slower compared to a corresponding C++ implementation. Nevertheless, we show that our estimator is significantly faster to compute than the others.

[1] `www.dgtal.org`



(a) $h = 1$ (104 pixels) (b) $h = 0.1$ (1016 pixels) (c) curvature profile

Figure 5.3: *Digitized ellipse with great axis* 30 *and small axis* 20 *rotated by* 0.5rad.



(a) $h = 1$ (368 pixels) (b) $h = 0.1$ (3644 pixels) (c) curvature profile

Figure 5.4: *Digitized flower with 5 petals with great radius* 20 *and small radius* 5.

Before we will compare our algorithm with the BC- and the MDCA-estimator and discuss the results, we will briefly sketch how we have modified the original algorithm to work also on digitized contours and not only for discrete samples of continuous curves.

MODIFICATIONS OF THE ALGORITHM

In order to reduce the digitization problem to the case of point evaluations, we have first constructed a cubic smoothing spline to the digitized contour. The cubic smoothing spline $s\,[0,1] \to \mathbb{R}^2$ is the solution of the minimization problem

$$\lambda \sum_{i=1}^{n} (s(t_i) - \alpha_h(X;t_i))^2 + (1-\lambda) \int_0^1 \left(D^2 s(t)\right)^2 \mathrm{d}t.$$

With $\lambda = 1$ we obtain the cubic spline interpolant to the data $\alpha_h(X;t_i)$. If $\lambda \to 0$, the smoothing spline converges to a straight line. For the evaluation, we have used the smoothing parameter $\lambda = 0.55$ for an almost equal relation between data validity and smoothness. In order to achieve a smooth curvature estimate we also smooth the resulting curvature estimation as linear combination of the weighted scale estimates by a smoothing spline with same parameter $\lambda = 0.55$.

As second modification, we have smoothed the approximation error and the curvature estimates at the scales $s_1, \ldots, s_l$ by a convolution with a Gaussian. We have used a window size that depends on the grid size h and on the current scale. At coarse scales

	Ellipse			Flower		
grid step	$h = 1$	$h = 0.1$	$h = 0.01$	$h = 1$	$h = 0.1$	$h = 0.01$
# points	104	1016	10140	368	3644	36424
absolute max. error	0.0287	0.0081	0.0056	0.3591	0.1594	0.0580
absolute avg. error	0.0094	0.0021	0.0008	0.0305	0.0090	0.0019
relative max. error	0.1920	0.0541	0.0370	0.6246	0.2772	0.1009
relative avg. error	0.1179	0.0265	0.0103	0.3693	0.1118	0.0232

Table 5.2: *Evaluation of the spline curvature estimator.*

the window was smaller, while at fine scales more smoothing was necessary. Similarly, we have used a larger window for a small grid step h using a logarithmic scaling.

For both shapes, we have used the grid step independent scales

$$s_1 = 32, s_2 = 16, s_3 = 8, s_4 = 4, s_5 = 2,$$

where these factors are used as sub-sampling factors for the variation diminishing spline approximation. Except this modifications, we have used the algorithm as described in Section 5.4.2.

ACCURACY

We have measured the absolute average and the maximal error as well as the relative error for the digitization steps $h = 1$, $h = 0.1$, and $h = 0.01$. Moreover, we have reported the corresponding running times. The results of the spline based curvature estimator are shown in detail in Table 5.2. Table 5.3 compares the spline based estimator against the other algorithms based on the relative average error and the running time. We choose to show the relative error in order to be able to compare the error of both shapes with each other. The curvature estimations are plotted against the ground truth in Figure 5.5 for the ellipse and in Figure 5.6 for the flower.

The results show that each estimator is more accurate at the ellipse than at the flower. This phenomenon is also reflected in Figure 5.7 and Figure 5.8 where the relative average error and the relative maximal error is illustrated for both shapes. Another point to look at are the oscillations of the spline estimator in Figure 5.5. These effects are coming from the digitization of the smooth shape and its smoothing spline approximation and are better resolved in the flower shown in Figure 5.6 where the shape contains more details. However, the spline estimator achieves very good accuracy compared to the BC and the MDCA estimator as can be seen in Figure 5.7 and Figure 5.8. The BC estimator performs worse at both shapes while our spline based algorithm achieves comparable accuracy as the MDCA estimator. Also note that the accuracy of our algorithm increases more significant at finer digitizations compared to the others. Therefore, we recommend our algorithm if fine digitizations are considered. In the following, we will

	relative avg. error			run time [ms]		
	BC	MDCA	Spline	BC	MDCA	Spline
Ellipse, $h = 1$	0.0930	0.1138	0.1179	0.6	18.5	18.4
Ellipse, $h = 0.1$	0.0298	0.0311	0.0265	74.9	222.4	35.5
Ellipse, $h = 0.01$	0.0105	0.0090	0.0103	22020	2843.3	126.7
Flower, $h = 1$	0.3875	0.3354	0.3693	3.5	63.0	30.2
Flower, $h = 0.1$	0.2240	0.0907	0.1118	663.4	956.1	57.5
Flower, $h = 0.01$	0.0968	0.0266	0.0232	186821	9334.9	532.9

Table 5.3: *Comparison between curvature estimators.*
The table shows the relative average error and the corresponding run time for the curvature estimation for two shapes digitized at three different grid steps. Note that while the spline based estimator achieves competitive performance the computation times are significantly faster.

give another argument that confirms this statement as our algorithm is much faster to compute.

RUNNING TIME

All the evaluations have been run on an Intel Core i7 with 2.5Ghz and 8GB of main memory. The resulting running times are reported in Table 5.3. While at the relative average error of the curvature estimations all estimator are quite close, the spline based method is the only one that can run each evaluation significantly under a second. The longest running times have been measured for the flower shape at digitization level $h = 0.01$. The BC estimator needs more than 3 minutes for the estimation and the MDCA estimator needs 9 seconds, whereas the spline estimator only needs about 0.5 seconds. In terms of running times, the spline estimator clearly outperforms the other two which is of importance if shapes at high resolutions are considered or if the curvature of a large number of shapes have to be computed. This is due to the linear approximation scheme of the spline estimator. Also note that the BC and the MDCA estimator are implemented on C++, the spline estimator is only implemented in Matlab. Therefore, the real running times for the spline estimator can be even faster.

PIECEWISE SMOOTH CURVES

If fast computation times matter and accuracy of the curvature estimation is crucial, we have already seen that our algorithm has to be chosen over the BC and the MDCA estimator. We will give now another advantage, as the spline based algorithm is able to handle piecewise smooth curves.

The lower estimates of the spline approximation error shown in the last chapter combined with the known upper estimates allow the characterization of the smoothness of the function. A fast decay rate of the approximation error characterizes smooth regions, whereas a slow decay rate occurs around C^2-singularities. We utilize this relation for the detection of C^2-singularities. This has already described in Section 5.4.2, where we have provide a multi-scale based algorithm that decides whether a point on the curve is a singularity or not based on the decay rate. As the variation diminishing spline operator interpolates at the endpoints, we can split our curves at the singularities and reduce the problem to estimate the curvature of the piecewise smooth curve to estimate the curvature of each of the smooth patches. Then, Theorem 5.9 guarantees the convergence of the spline based curvature estimates.

We have tested the algorithm for a circle, a rectangle and for a smooth Bezier-curve that has exactly one singularity at the bottom. We have constructed the curve in such a way that a second point of the curve has about the same estimated curvature value as at the singularity. This is to show that singularities can no be detected by pure thresholding of high curvature values. The results shown in Figure 5.9, Figure 5.10, and Figure 5.11 demonstrate the exactness of the detected singularities.

5.4.5 Summary

The evaluations in the preceding section have shown that while the spline based estimator achieves competitive accuracy the running times are much faster compared to the BC and the MDCA estimator. Another highlight of the spline based solution coming from the field approximation theory is the ability to handle piecewise smooth curves due to the lower estimates by the second order modulus of smoothness if the spline degree is at least 3. Overall, our algorithm does not only outperform the others if computation times matter but is also able to handle piecewise smooth curves that occur naturally in many digital images.

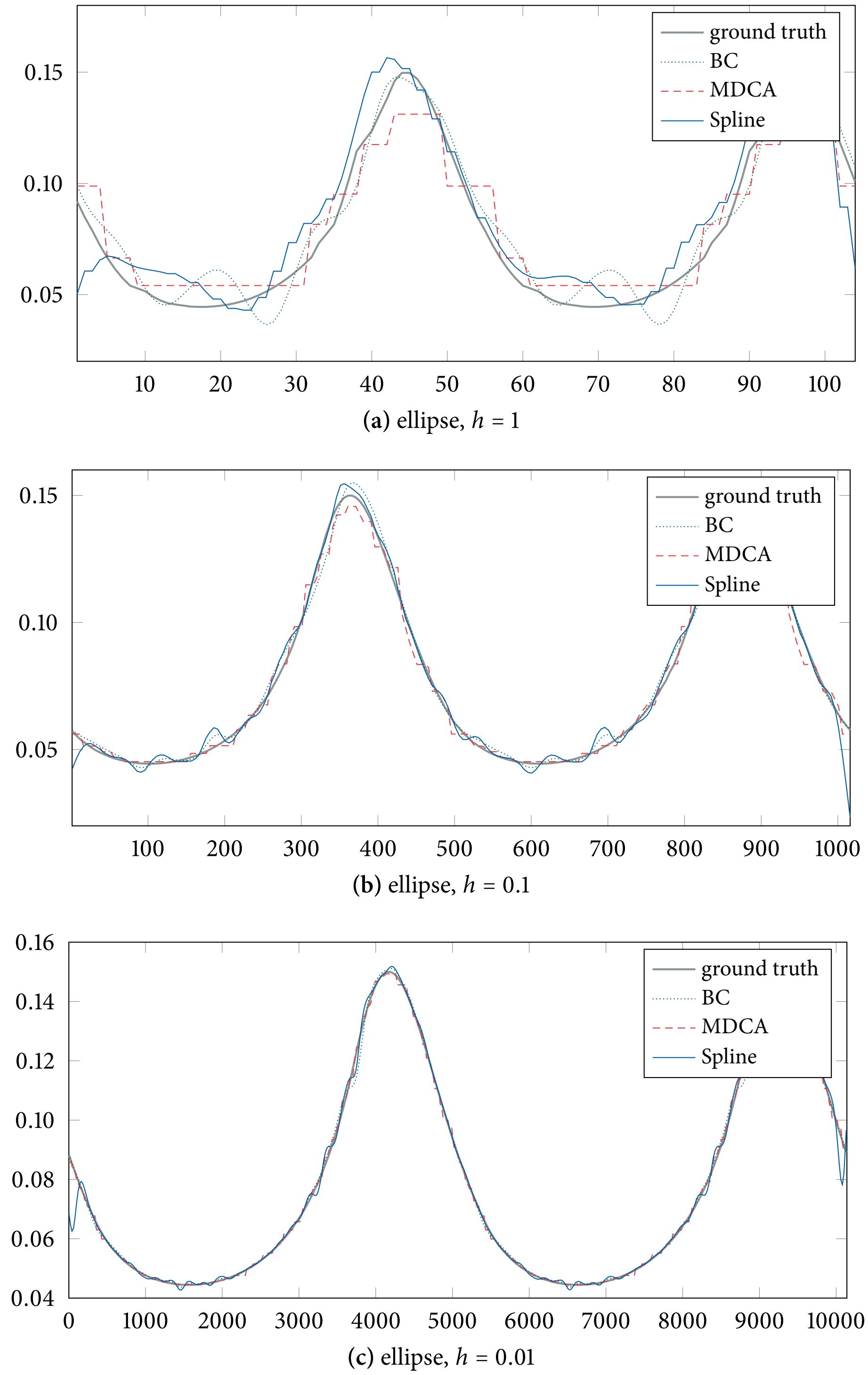
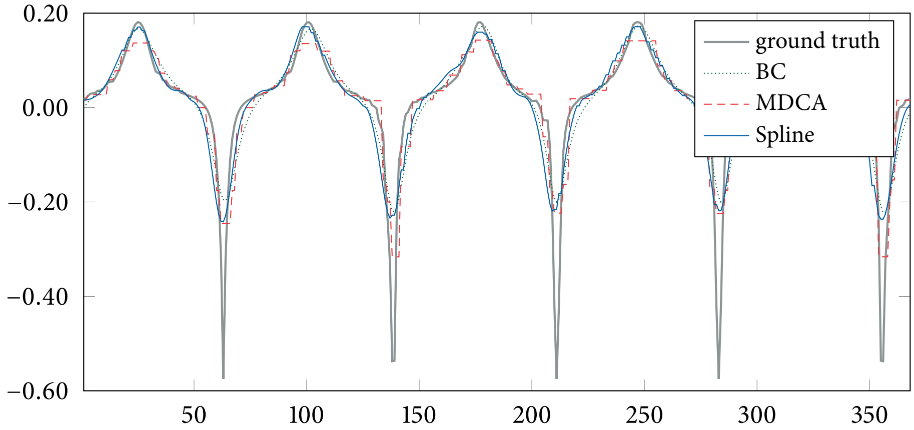
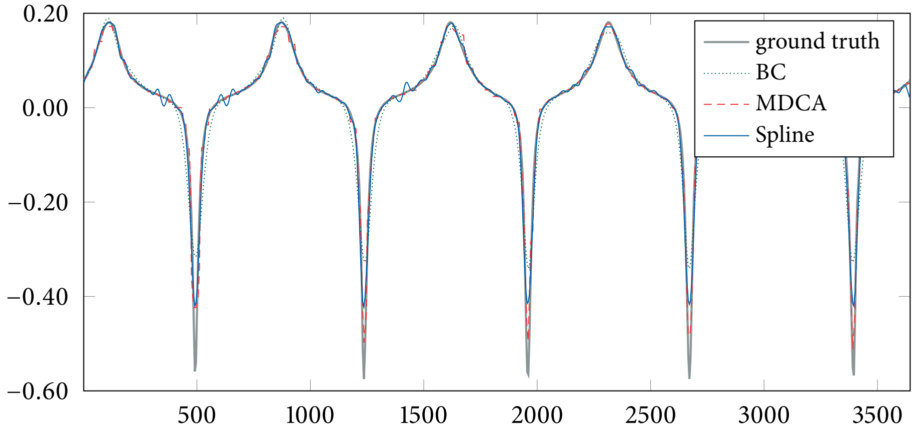


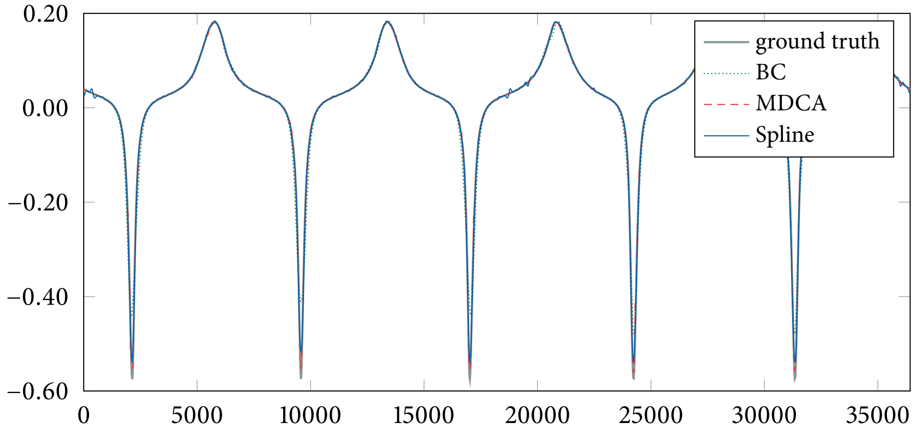
Figure 5.5: *Curvature profiles for the digitized ellipse with varying parameter* $h \in \{1, 0.1, 0.01\}$. *Note the oscillations that are made by spline estimation due to the smoothing of the digitization.*



(**a**) flower, $h = 1$

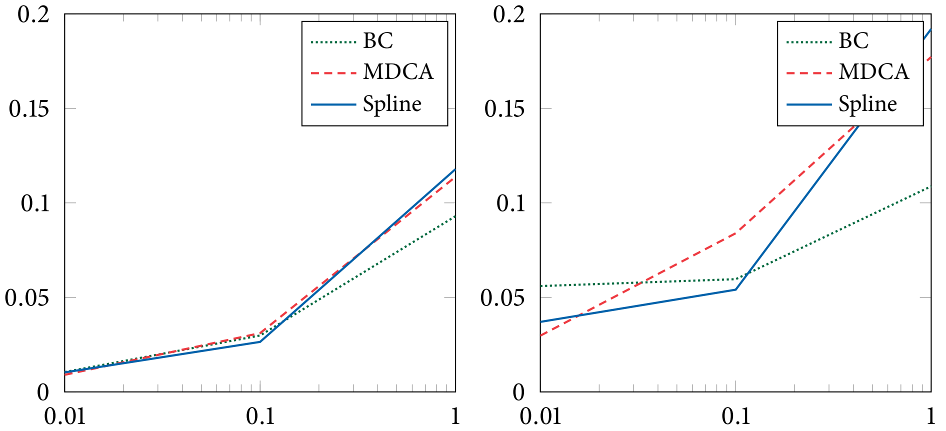


(**b**) flower, $h = 0.1$



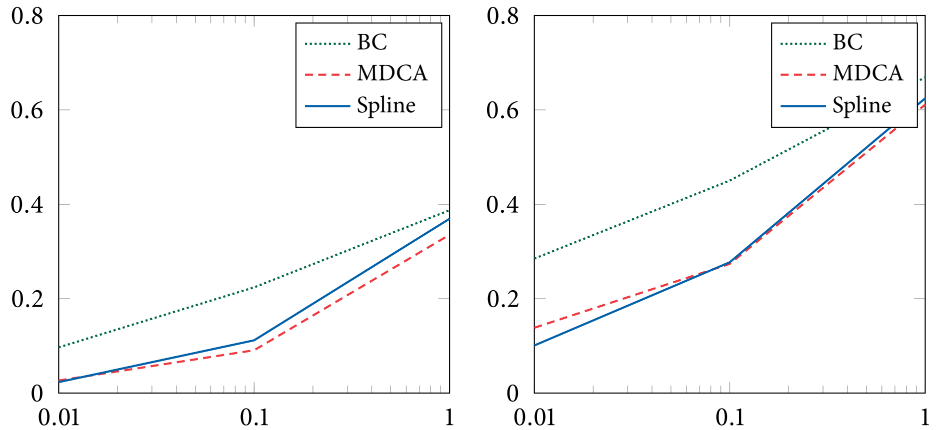
(**c**) flower, $h = 0.01$

Figure 5.6: *Curvature profiles for the digitized flower with varying parameter* $h \in \{1, 0.1, 0.01\}$.



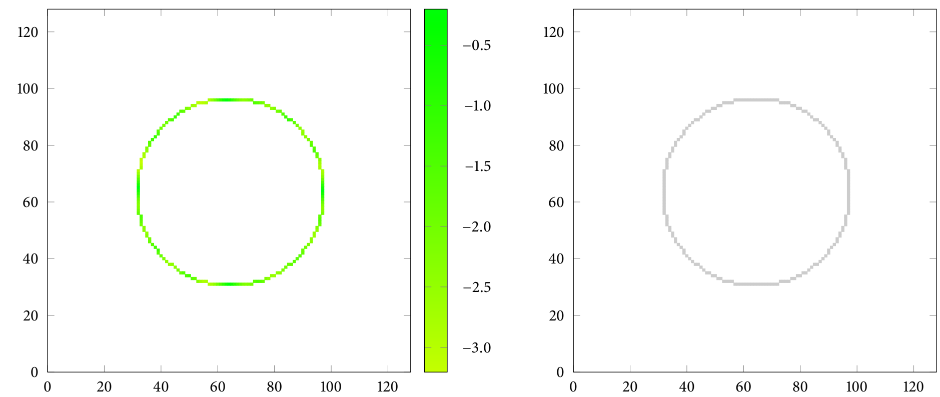
(a) Relative average error **(b)** Relative maximal error

Figure 5.7: *The relative average and maximal errors for the estimated curvature of the ellipse. While the spline estimator performs worst at the digitization step* $h = 1$*, it achieves almost the same accuracy as the MDCA estimator.*



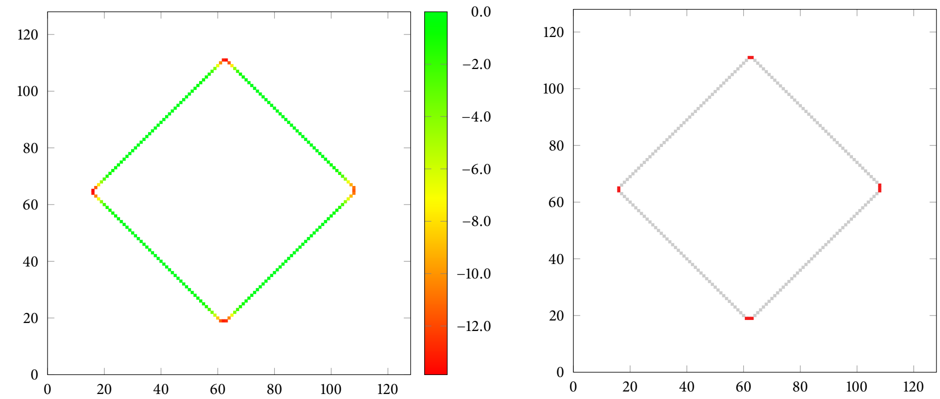
(a) Relative average error **(b)** Relative maximal error

Figure 5.8: *The relative average and maximal error for the estimated curvature of the flower. At the large grid step* $h = 1$*, the accuracy of the spline estimator is between the other estimators. The finer the grid step gets, the better the accuracy gets compared to the BC and the MDCA estimator. At* $h = 0.01$ *the spline estimator outperforms both curvature estimators.*



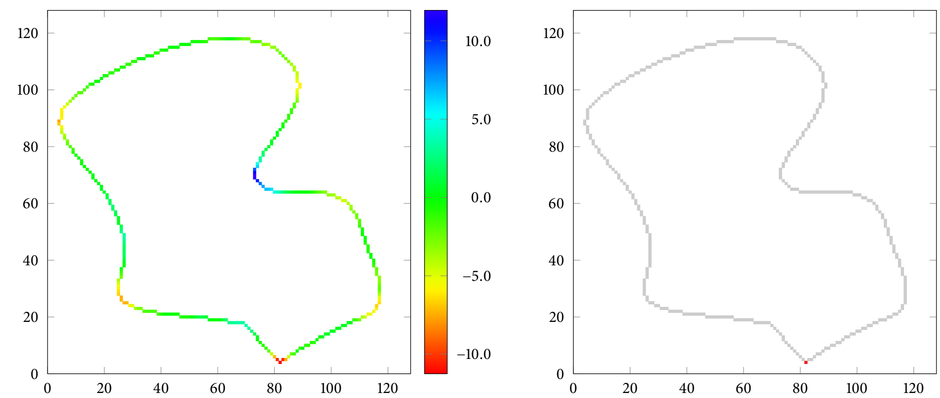
(a) Curvature (b) C^2-Singularities

Figure 5.9: *Singularity detection at a circle. The algorithm has correctly detected no singularities.*



(a) Curvature (b) C^2-Singularities

Figure 5.10: *Singularity detection at a rectangle. The detected singularities are very well localized at each corner.*



(a) Curvature (b) C^2-Singularities

Figure 5.11: *Singularity detection at smooth curve with one singularity.*
As can bee seen the singularity has been detected correctly. Note that by pure thresholding a second smooth point would have been wrongly detected.

PART TWO

GENERALIZATIONS

CHAPTER 6 Lower Estimates – An Operator Theoretic View

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

ALBERT EINSTEIN

THE ESTABLISHED LOWER ESTIMATES of the variation diminishing spline operator provide an elegant way to characterize the singularities of a continuous function depending on the chosen degree of the spline. Here, we will take an operator theoretic view on the proof of Theorem 4.9 to obtain new insights in the underlying concept for a generalization to other operators. Revisiting the proof for the lower estimate reveals mainly three important conditions. First, a differential operator of a certain order has to be bounded on the range of the operator. Second, the iterates of the operator have to converge in the uniform operator topology against a limiting operator, and finally, the differential operator has to annihilate this limiting operator. While the first condition depends only on the smoothness of the range of the operator, the other two conditions are more interesting to characterize. In the following it will be shown that the asymptotic behavior of the iterates can be established by spectral properties. Moreover, the limiting operator of the iterates is an projection operator on the fixed points of the iterated operator.

In the next section, we will revisit the proof for the lower estimate of Schoenberg's operator and highlight its key concepts in a general setting. Afterwards, a characterization of the iterates will be given in Section 6.2 by the classical work of Dunford [Dun43b] and Katznelson and Tzafriri [KT86]. Finally, we give an outlook of open questions that will be answered during the next chapters.

6.1 Lower estimates revisited

The key idea of the proof of the lower estimates for the Schoenberg operator is to find and upper bound of the norm $\|D^r S_{\Delta_n,k} f\|_\infty$ in terms of the approximation error $\|f - S_{\Delta_n,k} f\|_\infty$. This has been achieved using an extension to a telescopic series where the iterates of $S_{\Delta_n,k}$ have to be estimated. In the following we revisit the the proof of Theorem 4.9 stated on page 48 in a generalized setting.

To this end, let us consider a general bounded operator T defined on a complex Banach space $(X, \|\cdot\|_X)$ that satisfies the following conditions:

1. $\operatorname{ran}(T) \subset C^r$ *and* D^r *is bounded on the range of* T *for* $r \in \mathbb{N}$,
2. *the iterates* T^m *converge for* $m \to \infty$ *against an operator* P *with rate*
$$\|T^m - P\|_{op} \leq \gamma^{m-1},$$
where $0 < \gamma < 1$,
3. D^r *annihilates* P, *i. e.*, $D^r P = 0$.

Under these assumption on T, we can estimate the term $\|D^r Tf\|_X$ for any $f \in X$ as follows:

$$\begin{aligned}
\|D^r Tf\|_X &= \left\|D^r Tf - D^r T^2 f + D^r T^2 f - D^r T^3 f + \ldots\right\|_X \\
&\leq \sum_{m=1}^{\infty} \|D^r T^m (f - Tf)\|_X \\
&\leq \|f - Tf\|_X \sum_{m=1}^{\infty} \|D^r T^m\|_{op} \\
&= \|f - Tf\|_X \sum_{m=1}^{\infty} \|D^r (T^m - P + P)\|_{op} \\
&= \|f - Tf\|_X \sum_{m=1}^{\infty} \|D^r (T^m - P)\|_{op},
\end{aligned}$$

as D^r annihilates P. Now we can use the boundedness of D^r on the range of T to get the operator norm of D^r in front of the series:

$$\|D^r Tf\|_X \leq \|D^r\|_{op:\operatorname{ran}(T)} \|f - Tf\|_X \sum_{m=1}^{\infty} \|T^m - P\|_{op}.$$

Applying condition 2, the convergence of the iterates, yields the inequality

$$\begin{aligned}
\|D^r Tf\|_X &\leq \|D^r\|_{op:\operatorname{ran}(T)} \|f - Tf\|_X \sum_{m=1}^{\infty} \gamma^{m-1}. \\
&= \|D^r\|_{op:\operatorname{ran}(T)} \|f - Tf\|_X \sum_{m=0}^{\infty} \gamma^{m}.
\end{aligned}$$

In fact, this geometric series converges as $\gamma < 1$ and hence, we obtain final the estimate

$$\|D^r Tf\|_X \leq \frac{\|D^r\|_{op:\operatorname{ran}(T)}}{1 - \gamma} \|f - Tf\|. \tag{6.1}$$

Let us conclude with some remarks on this approach. The rate of convergence of the series depends on the real number $\gamma < 1$. We will see later that this number is equal to the second largest eigenvalue of T by its modulus. Besides, note that if $f \in \ker(T - I)$, i. e., f is a fixed point of T, then the inequality (6.1) yields $\|D^r T f\|_X \leq 0$. Therefore, $\|D^r T f\|_X = 0$ has to be true for all fixed points of T. In fact, this property is guaranteed by the condition that $D^r P = 0$ holds. It will be shown in the next section that P is necessarily the projection onto the fixed point space $\ker(T - I)$.

6.2 Iterates of linear operators

Motivated by the last section, we will consider now the asymptotic behaviour of the iterates of a linear operator. We start here with fundamental work of Dunford published already in 1943. To this end, we consider a complex Banach space X and let $T \in \mathcal{L}(X)$ a bounded linear contraction, i. e., $\|T\|_{op} \leq 1$. In a more general setting, Dunford [Dun43b] has shown the following.

Theorem 6.1 ([Dun43b, Thm. 3.16]). *Let T be an operator such that $\left\|T^{m+1} - T^m\right\|_{op} \to 0$ for $m \to \infty$. Then the following statements are equivalent.*

1. *$T^m \to P$, $P^2 = P$, $P(X) = \ker(T - I)$.*
2. *$X = \ker(T - I) \oplus \operatorname{ran}(T - I)$ and $\operatorname{ran}(T - I)$ is closed.*
3. *The point $\lambda = 1$ is either in $\rho(T)$ or else a simple pole of $R(\lambda, T)$.*
4. *$\operatorname{ran}(T - I)^2$ is closed.*

The first item guarantees the convergence of the iterates towards an projection operator P that projects onto the fixed point space of T. Furthermore, if a power of T is compact, Dunford has shown in particular that all items hold true provided that $T^{m+1} - T^m$ converges to zero in the uniform operator topology.

Proposition 6.2 ([Dun43b, Thm. 3.16]). *Let $T \in \mathcal{L}(X)$ with $\|T\|_{op} \leq 1$ be such that $\left\|T^{m+1} - T^m\right\|_{op} \to 0$ for $m \to \infty$. If T^m is compact for some integer m, then the statements of Theorem 6.1 are all true.*

As the operators of our interest are in general compact there remains only one open question. We want to know which conditions on T guarantee that $\left\|T^{m+1} - T^m\right\|_{op} \to 0$ converges if m tends to infinity. A concrete answer can be found in the beautiful work of Katznelson and Tzafriri [KT86]. They provided a sufficient and necessary criterion based on the spectral location of T.

Proposition 6.3 (Katznelson and Tzafriri [KT86, Thm. 1]). *Let T be an operator such that $\|T\|_{op} \leq 1$. Then*

$$\lim_{m\to\infty} \left\|T^{m+1} - T^m\right\|_{op} = 0$$

if and only if

$$\sigma(T) \subset B(0,1) \cup \{1\}. \tag{6.2}$$

The spectrum has to be contained in the unit ball with the only intersection at 1. Clearly, if $\|T\|_{op} < 1$, then $\|T^n\|_{op} \to 0$. Thus, we are interested in the case when $\|T\|_{op} = 1$ holds.

In the next chapter, we consider positive linear operators with finite rank and a partition of unity property as discussed in Chapter 3. In particular, we show that these operators do always fulfill the spectral location property (6.2) according to Katznelson and Tzafriri. Furthermore, we provide a sufficient criterion on the limiting projection operator and give examples based on the introduced operators of Chapter 3.

Nevertheless, there is no general method to obtain the limiting projection operator in an easy way. Using functional calculus, the limiting operator can be obtained by an integral of the form

$$P = \frac{1}{2\pi i} \int_{B(1,\varepsilon)} (T - \lambda I)^{-1} \, \mathrm{d}\lambda,$$

for some $\varepsilon > 0$ small enough such that there is no other spectral value inside the ball $B(1,\varepsilon)$, see Heuser [Heu82, pp. 204] or Section 8.5. However, this integral can not be easily solved for T in general. As highlight of this thesis, will show in Chapter 9 how to construct the limiting operator based on the fixed points of T and the fixed points of T^*. To this end, we will state the most fundamental results that are necessary to develop these results in Chapter 8.

CHAPTER 7 The Spectrum of Positive Linear Operators

"It is not enough to have a good mind. The main thing is to use it well."
RENÉ DESCARTES

WE STUDY positive linear operators defined on a general infinite-dimensional complex Banach function space X that contains the constant function 1 with norm equal to one. In addition, we assume that the associated basis functions of the positive linear operator form a partition of unity that guarantees the exact reconstruction of constant functions. Operators of this kind have already been discussed in detail in Section 2.4 and Chapter 3. Remember, e. g., the variation diminishing spline operator devised by Schoenberg or the classical Bernstein operator. The results shown here are established in a more general setting, namely on general Banach spaces. Therefore, they are applicable to all positive operators on the space of continuous functions as well as on the space of integrable functions.

We are interested in the asymptotic behavior of the iterates, as motivated in Chapter 6 by proving lower estimates. Here, we provide a functional analysis based approach using spectral properties that guarantees the existence of the limit of these iterates defined on a general Banach space. Concretely, we want to apply the famous theorem of Katznelson and Tzafriri [KT86] already stated in Proposition 6.3. This chapter is devoted to give an application of this beautiful result in the field of approximation theory for positive linear operators. Accordingly, we will show that the spectrum of a positive linear operators T with a partition of unity property is characterized by

$$\sigma(T) \subset B(0,1) \cup \{1\}.$$

This property guarantees the convergence of the iterates towards a projection operator using the work of Dunford [Dun43b]. We will finally provide a sufficient criterion to derive the limiting operator and show how this criterion can be applied based on the operators we have introduced in Chapter 3.

The asymptotic behaviour of the iterates of positive linear operators has already been extensively discussed by many authors. Kelisky and Rivlin [KR67] have been the first to consider the limit of iterates of the classical Bernstein operator on the space $C([0,1])$. This result has been extended by Karlin and Ziegler [KZ70] to a more general setting. In [Nag80; Nag82], Nagel has examined the asymptotic behaviour of the Bernstein and the Kantorovič operators. Using a contraction principle, Rus [Rus04] has shown an alternative way to prove the convergence of the iterates of the Bernstein operator. The iterates of the Bernstein operator have been also revisited by Badea [Bad09] using spectral properties. Recently, contributions have been made by Gavrea and Ivan [GI10; GI11a; GI11c; GI11b] and by Altomare [Alt13] using methods based on Korovkin-type approximation theory. However, all these results are restricted to the space of continuous functions, i.e., are not applicable for the L^p spaces, and there is no general theory that guarantees the existence of the limit of the iterates.

Note that the results shown in this chapter have already been published in the *Journal of Mathematical Analysis and Applications (JMAA)*, see [Nag15].

7.1 Setting

In the following let K be a compact Hausdorff space and let $(X, \|\cdot\|_X)$ be a complex infinite-dimensional Banach function space on K that contains the constant function 1 with $\|1\|_X = 1$. Given an integer $n > 0$ and linearly independent positive functions $e_1, \dots, e_n \in X$ that form a partition of unity, i.e.,

$$\sum_{k=1}^{n} e_k = 1. \tag{7.1}$$

we set $Y := \operatorname{span}\{e_1, \dots, e_n\}$. Clearly, Y is a finite-dimensional subspace of X with $1 \in Y$. Equipped with a norm $\|\cdot\|_Y$ that satisfies $\|1\|_Y = 1$, the space Y becomes a Banach space sharing the property of X that the constant function one is normalized. Consider, e.g, Y equipped with the inherited norm of X.

Then we define the positive finite-rank operator $T : X \to Y$ by

$$Tf = \sum_{k=1}^{n} \alpha_k^*(f) e_k, \qquad f \in X, \tag{7.2}$$

where α_k^* are positive linear functionals satisfying $\|\alpha_k^*\|_{X^*} = \alpha_k^*(1) = 1$ and $\alpha_k^*(e_k) > 0$ for $k \in \{1, \dots, n\}$.

Note that the operator, we are interested in are usually defined on real Banach spaces. However, the results of this chapter can be applied if one considers a complexificiation of the real Banach spaces $C([0,1])$, $L^p([0,1])$, as demonstrated in the first chapter of [Rus86]. In this case, there are many operators that match this definition, consider e.g., the Bernstein operator, Schoenberg's variation-diminishing spline operator that arise in many applications in approximation theory and CAGD.

7.2 Basic properties

This section discusses properties that characterize the positive finite-rank operator T. The next lemma states the positivity of T and the ability to preserve constants.

Lemma 7.1. *The linear operator T, defined by* (7.2)*, is positive and reproduces constants.*

Proof. As the α_k^* are linear positive functionals and $e_k \geq 0$, we conclude for $f \in C([0,1])$, $f \geq 0$,

$$Tf = \sum_{k=1}^{n} \alpha_k^*(f) e_k \geq 0.$$

And we obtain by applying the preconditions on T that

$$T1 = \sum_{k=1}^{n} \alpha_k^*(1) e_k = \sum_{k=1}^{n} e_k = 1.$$

□

Lemma 7.2. *The operator $T : X \to Y$ is bounded and $\|T\|_{op} = 1$.*

Proof. Let $f \in X$ such that $\|f\|_X = 1$. Then

$$\|Tf\|_Y = \left\|\sum_{k=1}^{n} \alpha_k^*(f) e_k\right\|_Y \leq \max_k |\alpha_k^*(f)| \cdot \left\|\sum_{k=1}^{n} e_k\right\|_Y \leq \|f\|_X \cdot \max_k \|\alpha_k^*\|_{X^*} = 1,$$

where we used the partition of unity (7.1) and that $\|\alpha_k^*\|_{X^*} = 1$. Using that $T1 = 1$, we conclude that $\|T\|_{op} = 1$. □

Now we will prove that the operator T is indeed a finite-rank operator and give additional basic properties.

Lemma 7.3. *The linear operator T has finite rank. Thus, the operator T is compact.*

Proof. As $\operatorname{ran}(T) = \operatorname{span}\left\{\sum_{k=1}^{n} \alpha_k^*(f) e_k : f \in X\right\}$, the range can be written as a linear combination of the n basis functions e_k and hence, $\dim(\operatorname{ran}(T)) \leq n$. Therefore, the linear operator T has finite rank and each finite-rank operator is compact. □

We can also give an explicit representation of the adjoint of T. Here, the basis of T^* are the functionals α_k^* instead of the basis functions e_k for T.

Theorem 7.4. *The adjoint $T^* : Y^* \to X^*$ of T is a finite-rank operator. It is given for $x^* \in Y^*$ as*

$$T^* x^*(f) = \sum_{k=1}^{n} x^*(e_k) \alpha_k^*(f), \qquad f \in X.$$

Proof. We calculate

$$x^*(Tf) = x^*(\sum_{k=1}^{n} \alpha_k^*(f) \cdot e_k) = \sum_{k=1}^{n} \alpha_k^*(f) x^*(e_k) = T^* x^*(f).$$

□

7.3 Spectral properties

We give here a characterization of the spectrum of the operator T. More specifically, we will show that the eigenvalue 1 is the only spectral value on the unit circle and all spectral values are in particular eigenvalues of the operator.

Theorem 7.5 (The spectrum). *The spectrum of the operator T, defined above by* (7.2), *consists only of the point spectrum and is characterized by*

$$\{0\} \subset \sigma(T) = \sigma_p(T) \subset B(0,1) \cup \{1\}.$$

Corollary 7.6. *The positive finite-rank operator T given by* (7.2) *has the following properties:*

1. $1 \in \ker(T - I)$, *i. e., 1 is an (isolated) eigenvalue of T,*
2. $\sigma(T) \subset B(0,1) \cup \{1\}$, *and finally,*
3. $\dim(\ker(T)) = \infty$.

I.e., the only eigenvalue on the peripheral spectrum is 1 *and all spectral values are eigenvalues. Moreover,* 0 *is always an eigenvalue of T corresponding to an infinite-dimensional eigenspace of T.*

Proof of Theorem 7.5. Since $\|T\|_{op} = 1$, the inequality $|\lambda| \leq \|T\| = 1$ holds for each $\lambda \in \sigma(T)$. Therefore,

$$\sigma(T) \subset \overline{B(0,1)}.$$

In the following, we show that $\sigma(T) \subset B(0,1) \cup \{1\}$, i.e., if $\lambda \in \sigma(T)$ with $|\lambda| = 1$ then $\lambda = 1$ and all the spectral values are eigenvalues of T.

Note that for compact operators it is known that every $\lambda \neq 0$ in the spectrum is contained in the point spectrum. This classical result is stated, e.g., in Rudin [Rud91, Theorem 4.25]. Therefore, if $0 \in \sigma_p(T)$, then it follows already that

$$\sigma(T) = \sigma_p(T).$$

The proof is organized as follows: in the first step, we prove that $0 \in \sigma_p(T)$. Then, we will show that $1 \in \sigma_p(T)$. Finally, we consider eigenvalues $\lambda \in \sigma_p(T) \setminus \{0,1\}$ and we show that in this case $|\lambda| < 1$ holds. Here we will use the well-known result of Gershgorin [Ger31] to describe the spectrum.

Step 1: In order to prove that $0 \in \sigma_p(T)$ we show $\ker(T) \neq \{0\}$. Using Rudin [Rud91, Theorem 4.12], we obtain that $\ker(T) = \operatorname{ran}(T^*)_\perp$. As $T(X)$ is closed in Y, so is $T^*(Y^*)$ weak*-closed in X^*. Suppose now that $\ker(T) = \{0\}$. It follows that $\operatorname{ran}(T^*)_\perp = \{0\}$ and therefore, $(\operatorname{ran}(T^*)_\perp)^\perp = X^*$. This requires that $\operatorname{ran}(T^*)$ is weak*-dense in X^*. This gives a contradiction as $\operatorname{ran}(T^*) = \operatorname{span}\{\alpha_1^*, \dots, \alpha_n^*\}$ is weak*-closed and $X^* \neq \operatorname{span}\{\alpha_1^*, \dots, \alpha_n^*\}$, because X^* is infinite-dimensional.

We conclude that $\ker(T) \neq \{0\}$ and the finite-rank operator T is not one-to-one, i. e., $0 \in \sigma_p(T)$.

Step 2: By definition of the operator T we have that $1 \in \sigma(T)$, because of the partition of unity property and the unit 1 is an eigenfunction of T corresponding to the eigenvalue 1, $T1 = 1$.

Step 3: We consider now all the other possible eigenvalues of T that are not equal to zero or one. We will prove that for all these eigenvalues $\lambda \in \sigma(T) \setminus \{0,1\}$, we have

$$|\lambda| < 1.$$

Let $\lambda \in \sigma(T) \setminus \{0,1\}$. As the operator maps continuous functions to the finite dimensional space $\operatorname{ran}(T)$, the eigenfunctions have to be in this space, too. Let $p \in \operatorname{ran}(T)$, $p = \sum_{k=1}^n c_j e_k$, be such an eigenfunction for the eigenvalue λ. Then we get the following characterization:

$$\begin{aligned}
& Tp = \lambda p \\
\iff \quad & \sum_{k=1}^n \sum_{j=1}^n c_j \alpha_k^*(e_j) e_k(x) = \lambda \sum_{k=1}^n c_j e_k(x) \\
\iff \quad & \sum_{k=1}^n \left[\sum_{j=1}^n c_j \alpha_k^*(e_j) - \lambda c_k \right] e_k(x) = 0 \\
\iff \quad & \sum_{j=1}^n c_j \alpha_k^*(e_j) = \lambda c_k, \qquad \text{for all } k \in \{0, \dots, n\}.
\end{aligned}$$

Thus, $\lambda \neq 0$ is an eigenvalue of the operator T, if and only if λ is an eigenvalue of the matrix $M \in \mathbb{R}^{n \times n}$,

$$M = \begin{pmatrix} \alpha_1^*(e_1) & \alpha_1^*(e_2) & \cdots & \alpha_1^*(e_n) \\ \alpha_2^*(e_1) & \alpha_2^*(e_2) & \cdots & \alpha_2^*(e_n) \\ \vdots & & & \\ \alpha_n^*(e_1) & \alpha_n^*(e_2) & \cdots & \alpha_n^*(e_n) \end{pmatrix}.$$

This matrix M is nonnegative as $e_k \geq 0$ and α_k^* are positive linear functionals. Moreover, every row sums up to one because of the partition of unity property.

To see this, let us calculate for some fixed row $k \in \{1,\dots,n\}$ the following sum:

$$\sum_{j=1}^{n} \alpha_k^*(e_j) = \alpha_k^*(\sum_{j=1}^{n} e_j) = \alpha_k^*(1) = 1. \tag{7.3}$$

Hence, the underlying matrix of the finite-rank operator T is a stochastic matrix. We will show next, that 1 is the only spectral value on the unit circle.

By the famous Theorem of Gershgorin [Ger31], the eigenvalues of M are contained in the union of circles,

$$\lambda \in \bigcup_{j=1}^{n} D_j,$$

where

$$D_j := \left\{ \lambda \in \mathbb{C} : \left| \lambda - \alpha_j^*(e_j) \right| \le \sum_{k=1,k\neq j}^{n} \alpha_j^*(e_k) \right\}.$$

Applying (7.3) yields

$$D_j = \left\{ \lambda \in \mathbb{C} : \left| \lambda - \alpha_j^*(e_j) \right| \le 1 - \alpha_j^*(e_j) \right\}.$$

We conclude the proof noting that $\alpha_j^*(e_j) > 0$ holds for all $j \in \{1,\dots,n\}$ and thus,

$$\bigcup_{j=1}^{n} D_j \cap \{\lambda \in \mathbb{C} : |\lambda| = 1\} = \{1\}.$$

Finally, we have proved that all spectral values of T are in particular eigenvalues and the spectrum of T is contained in $B(0,1) \cup \{1\}$.

□

7.4 The asymptotic behavior of the iterates

Using the spectral properties shown in the last section, the existence of a limit of the iterates is now guaranteed by the result of Katznelson and Tzafriri [KT86]. Furthermore, the limiting operator preserves the ability to reconstruct constants.

Theorem 7.7 (The existence of the limit of the iterates). *Let the operator T be a finite-rank operator with a partition of unity property as in* (7.2). *Then*

$$\lim_{n\to\infty} T^n = P$$

uniformly in the operator norm, where P is a compact operator on X such that $P1 = 1$ holds. Moreover, P is the unique projection operator onto $\ker(T - I)$ *that satisfies $TP = PT = P$.*

Proof. Now we will show the existence of the limit of the iterates of the finite-rank operator that has the partition of unity property.

Katznelson and Tzafriri [KT86] have shown that for every linear operator T on a Banach space X with $\|T\|_{op} \leq 1$ the limit

$$\lim_{n\to\infty} \|T^{n+1} - T^n\|_{op} = \lim_{n\to\infty} \|T^n(T - I)\|_{op} = 0$$

holds if and only if either there is no spectral value on the unit circle or the unit circle contains the single value 1. As Corollary 7.6 states, a finite-rank operator with a partition of unity property contains 1 as an eigenvalue and this is the only spectral value on the unit circle, $\sigma(T) \subset B(0,1) \cup \{1\}$.

To prove Theorem 7.7 let us consider the sequence $(T^n)_{n\in\mathbb{N}}$. By Katznelson and Tzafriri [KT86, Thm. 1] and by Dunford [Dun43b, Thm. 3.18] this sequence has a limit P in the Banach algebra $\mathcal{L}(X, Y)$. As $T^n 1 = 1$ for any positive integer n we obtain due to the uniform convergence in the operator norm the result $P1 = 1$. As a limit of the finite-rank operators T^n, the operator P is compact. Furthermore, $TP = PT = P$ holds and P is idempotent, i.e., $P = P^2$. □

The next corollary provides a sufficient (and necessary) criterion to obtain the limiting operator P. Moreover, the operator P preserves all the fixed points of T.

Corollary 7.8. *Let T be a positive linear operator as in* (7.2)*. If there exists an idempotent operator, i.e., $P^2 = P$, that commutes with T such that $TP = PT = P$ holds with range* $\operatorname{ran}(P) = \ker(T - I)$*, then*

$$\lim_{m\to\infty} \|T^m - P\|_{op} = 0.$$

Proof. We have from Dunford [Dun43b, Theorem 3.16 on page 216] that

$$X = \ker(T - I) \oplus \operatorname{ran}(T - I) \tag{7.4}$$

and $TP = PT = P$ and P is idempotent, i.e., $P = P^2$.

We now show that $\ker(P) = \operatorname{ran}(T - I)$. As $P(T - I) = PT - P = 0$, $\operatorname{ran}(T - I) \subset \ker(P)$ holds. We show that the converse holds also true. Let $0 \neq x \in \ker(P)$, then $x \notin \ker(T - I)$. By (7.4), there are $x_{\ker} \in \ker(T - I)$ and $x_{\operatorname{ran}} \in \operatorname{ran}(T - I)$ such that $x = x_{\ker} + x_{\operatorname{ran}}$. Using that $x \in \ker(P)$, we obtain

$$0 = Px = Px_{\ker} + Px_{\operatorname{ran}} = Px_{\ker},$$

as $x_{\operatorname{ran}} \in \operatorname{ran}(T-I) \subset \ker(P)$. From $x_{\ker} \in \ker(P)$ we conclude that $x_{\ker} = 0$. Therefore, $x = x_{\operatorname{ran}}$ and we get the final statement $x \in \operatorname{ran}(T - I)$. Hence, $\operatorname{ran}(P) = \ker(T - I)$

and $\ker(P) = \operatorname{ran}(T - I)$. As direct consequence we obtain Corollary 7.8 using the uniqueness of the limiting projection operator that satisfies $TP = PT = P$. □

7.5 Examples

We show concrete examples where the results of the preceding sections can be applied using the operators introduced in Chapter 3. First the space of continuous functions is considered where an operator evaluates a continuous function on finitely many given points. Here, we show the limiting operator of the Bernstein and Schoenberg operator. In the second example, we discuss the Kantorovič operator on $L^1([0,1])$ that yields an extension of the classical Bernstein operator to the space of integrable functions. Finally, we will consider the integral Schoenberg operator as L^1-extension of Schoenberg's variation diminishing spline operator.

7.5.1 Operators using point evaluations

The Riesz representation theorem gives a characterization of positive linear functionals on $C([0,1])$. Namely, for every positive linear functional $a^* : C([0,1]) \to \mathbb{R}$, there is a unique positive Radon measure ν such that

$$\alpha^*(f) = \int_0^1 f \mathrm{d}\nu \qquad \text{for every } f \in C([0,1]).$$

A classical example of a positive linear functional on $C([0,1])$ is the Dirac measure at a point $x \in [0,1]$ defined for $f \in C([0,1])$ by

$$\delta_x(f) = f(x).$$

Given a partition $\Delta_n = \{x_k\}_{k=1}^n$ of $[0,1]$ satisfying

$$0 = x_1 < x_2 < \ldots < x_n = 1,$$

then a popular choice for the functionals is $\alpha_k^* = \delta_{x_k}$ for $k \in \{1,\ldots,n\}$. In this case, the positive finite-rank operator can be written for $x \in [0,1]$ as

$$Tf(x) = \sum_{k=1}^n f(x_j)e_k(x), \qquad e_k \in C([0,1]).$$

Operators of this kind are often used to approximate continuous functions by only a finite number of samples. To apply Theorem 7.5, we also need the property that $e_k(x_k) > 0$ for all $k \in \{1,\ldots,n\}$. There are many examples where this property holds, see e.g., the Bernstein operator or the variation-diminishing Schoenberg operator discussed in Chapter 3. Let us suppose this property holds for T, then we obtain that 1 is the only spectral value on the unit circle and hence, $\lim_{m\to\infty} T^m f$ exists uniformly for all $f \in C([0,1])$ by Theorem 7.7. It has already been shown by Badea [Bad09] that the operator $Lf(x) = f(0) + (f(1) - f(0))x$ for $f \in C([0,1])$ satisfies the conditions of Corollary 7.8 for the Bernstein operator. Thus, we provide here an additional simple criterion to obtain the limit of the iterates of the Bernstein operator. Furthermore, we have shown in Chapter 4 that the conditions of Corollary 7.8 are also satisfied for the Schoenberg operator. Accordingly, the iterates of the Schoenberg operator converge uniformly towards L. The limiting operator of the Schoenberg operator has been shown the first time in Theorem 4.4 on page 44.

7.5.2 The Kantorovič operator

Let us consider the iterates of the Kantorovič operator defined on the space of integrable functions $L^1([0,1])$. Additionally, we illustrate the criterion given in Corollary 7.8 to derive the limiting operator.

Recall that the Kantorovič operator $K_n : L^1([0,1]) \to C([0,1])$ is defined as

$$K_n f(x) = (n+1) \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \int_{\frac{k}{n+1}}^{\frac{k+1}{n+1}} f(t)\mathrm{d}t, \qquad f \in L^1([0,1]), \quad x \in [0,1].$$

For more details, we refer to Section 3.1.2. First, we show that this is a finite-rank operator of the form (7.2). Then we will apply the main results of this chapter. To this end, let us denote the Bernstein polynomials by $p_{n,k}(x) := \binom{n}{k} x^k (1-x)^{n-k}$ and the functionals by

$$\alpha_{n,k}(f) := \int_{\frac{k}{n+1}}^{\frac{k+1}{n+1}} f(t)\mathrm{d}t$$

for $f \in L^1([0,1])$. Then the Kantorovič operator can be represented for $f \in L^1([0,1])$ by

$$K_n f(x) = (n+1) \sum_{k=0}^{n} \alpha_{k,n}(f) p_{k,n}(x).$$

In fact, the Bernstein polynomials form a partition of unity and $\int_{\frac{k}{n+1}}^{\frac{k+1}{n+1}} p_{k,n}(t)\mathrm{d}t > 0$ as $p_{k,n}(x) > 0$ on the open interval $\left(\frac{k}{n+1}, \frac{k+1}{n+1}\right)$ for all $k \in \{1, \dots, n\}$. Thus, each of the operators K_n is positive and has finite-rank. Using Theorem 7.5, we can characterize the spectrum of these operators by $\sigma(K_n) \subset B(0,1) \cup \{1\}$ and 1 is an isolated eigenvalue of K_n. Hence, for fixed n and as m goes to infinity the iterates K_n^m uniformly converge to an operator with the ability to reproduce constants. As this operator only preserves constant functions, we demonstrate an application of Corollary 7.8.

Let us consider the operator $L : L^1([0,1]) \to C([0,1])$,

$$L(f;x) = \int_0^1 f(t)\mathrm{d}t \cdot 1,$$

i. e., each f is mapped to the constant value of the integral over $[0,1]$. This operator L is idempotent, as

$$L^2(f;x) = \int_0^1 \left(\int_0^1 f(t)\mathrm{d}t \right) \mathrm{d}s = \int_0^1 f(t)\mathrm{d}t \cdot 1.$$

Also $K_n \circ L = L$ holds, as for all $f \in L^1([0,1])$ we obtain for $x \in [0,1]$:

$$\begin{aligned}
(K_n \circ L)(f;x) &= (n+1) \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \int_{\frac{k}{n+1}}^{\frac{k+1}{n+1}} L(f;t)\mathrm{d}t \\
&= (n+1) \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \int_{\frac{k}{n+1}}^{\frac{k+1}{n+1}} \int_0^1 f(s)\mathrm{d}s\mathrm{d}t \\
&= (n+1) \int_0^1 f(s)\mathrm{d}s \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \frac{1}{n+1} \\
&= \int_0^1 f(s)\mathrm{d}s \cdot 1.
\end{aligned}$$

In the last step we used the partition of unity property of the Kantorovič polynomials, namely that $K_n(1;x) = 1$ holds for all $x \in [0,1]$. Also $L \circ K_n = L$ holds. To prove this let $f \in C([0,1])$ and $x \in [0,1]$. Then

$$\begin{aligned}
(L \circ K_n)(f;x) &= \int_0^1 K_n(f;t)\mathrm{d}t \\
&= (n+1) \int_0^1 \sum_{k=0}^{n} \binom{n}{k} t^k (1-t)^{n-k} \int_{\frac{k}{n+1}}^{\frac{k+1}{n+1}} f(s)\mathrm{d}s\mathrm{d}t \\
&= (n+1) \sum_{k=0}^{n} \int_{\frac{k}{n+1}}^{\frac{k+1}{n+1}} f(s)\mathrm{d}s \binom{n}{k} \int_0^1 t^k (1-t)^{n-k}\mathrm{d}t \\
&= (n+1) \sum_{k=0}^{n} \int_{\frac{k}{n+1}}^{\frac{k+1}{n+1}} f(s)\mathrm{d}s \frac{1}{n+1} \\
&= \int_0^1 f(s)\mathrm{d}s = L(f;x).
\end{aligned}$$

Here, we note that $\int_0^1 t^k (1-t)^{n-k}\mathrm{d}t$ is the value of the Beta function

$$\beta(k+1, n+1-k) = \frac{k!\,(n-k)!}{(n+1)!}.$$

Therefore, the integral of the Bernstein polynomials is constant,

$$\int_0^1 \binom{n}{k} t^k (1-t)^{n-k}\mathrm{d}t = \frac{1}{n+1}.$$

This has been shown, e. g., by Kreyszig [Kre79]. The preceding results have shown that $TL = LT = L$ holds and that $L^2 = L$.

Therefore, all condition of Corollary 7.8 are satisfied and we conclude that

$$\lim_{m\to\infty} K_n^m(f;x) = L(f;x) \quad \text{uniformly on } [0,1] \text{ for all } f \in L^1([0,1]).$$

Indeed, this result has first been shown by Nagel [Nag82] but only for $f \in L^2([0,1])$. In contrast, our result not only extends the uniform convergence to the space $L^1([0,1])$ we also derive the convergence in the uniform operator norm.

7.5.3 The integral Schoenberg operator

In the same manner as for the Kantorovič operator, we will prove the limiting behaviour of the integral Schoenberg operator, introduced in Section 3.2.2. Note, that the partition of unity property $V_{\Delta_n,k}1 = 1$ holds true for this operator. Again, let us consider the operator $L : L^p([0,1]) \to C([0,1])$,

$$L(f;x) = \int_0^1 f(t)\mathrm{d}t \cdot 1.$$

We show now that $L \circ V_{\Delta_n,k} = V_{\Delta_n,k} \circ L = L$ holds. We calculate

$$\begin{aligned}
\int_0^1 V_{\Delta_n,k}(f;t)\mathrm{d}t &= \sum_{j=-k}^{n-1} \int_{\xi_{j-1,k+1}}^{\xi_{j,k+1}} f(s)\mathrm{d}s \int_0^1 \frac{N_{j,k}(t)}{\xi_{j,k+1} - \xi_{j-1,k+1}}\mathrm{d}t \\
&= \sum_{j=-k}^{n-1} \int_{\xi_{j-1,k+1}}^{\xi_{j,k+1}} f(s)\mathrm{d}s \\
&= \int_0^1 f(s)\mathrm{d}s
\end{aligned}$$

and

$$\begin{aligned}
V_{\Delta_n,k}\left(\int_0^1 f(t)\mathrm{d}t; x\right) &= \sum_{j=-k}^{n-1} \int_{\xi_{j-1,k+1}}^{\xi_{j,k+1}} \int_0^1 f(t)\mathrm{d}t\mathrm{d}s \frac{N_{j,k}(x)}{\xi_{j,k+1} - \xi_{j-1,k+1}} \\
&= \sum_{j=-k}^{n-1} \int_0^1 f(t)\mathrm{d}t \left(\xi_{j,k+1} - \xi_{j-1,k+1}\right) \frac{N_{j,k}(x)}{\xi_{j,k+1} - \xi_{j-1,k+1}} \\
&= \int_0^1 f(t)\mathrm{d}t \sum_{j=-k}^{n-1} N_{j,k}(x) \\
&= \int_0^1 f(t)\mathrm{d}t.
\end{aligned}$$

By Corollary 7.8, the iterates of the integral Schoenberg operator converge uniformly in the operator norm to the operator L, i. e.,

$$\lim_{m\to\infty} \left\| V_{\Delta_n,k}^m - L \right\|_{op} = 0.$$

7.6 Remarks

Even though the results of Theorem 7.7 and Corollary 7.8 guarantee the convergence for positive linear operators of finite-rank with a partition of unity property and also provide a criterion to derive the limiting operator, the concrete limiting operator has to be known previously. Besides, the theory of Dunford [Dun43b] and Katznelson and Tzafriri [KT86] does also work for general linear operators on complex Banach spaces. Note that complex Banach spaces are not a restriction, as real Banach spaces can be complexified as noted in Section 1.3. The complexity of the Banach space is important to consider spectral properties. Accordingly, we will generalize our setting in the next chapters to operators with finite-dimensional fixed points spaces. This viewpoint allows finally the explicit construction of the limiting operator based on an inversion of a Gramian matrix.

CHAPTER 8 Introduction to Riesz-Schauder and Fredholm Theory

"The mathematician does not study pure mathematics because it is useful; he studies it because he delights in it and he delights in it because it is beautiful."

HENRI POINCARÉ

For the convenience of the reader this chapter provides a compact overview over the most important facts on annihilators, complemented subspaces and projections. We introduce the concepts of the Riesz-Schauder theory for compact operators. Based on preserving the spectral properties of compact operators we motivate its generalization to Riesz operators by classical Fredholm theory. Finally, we consider quasi-compact operators as further generalization which constitutes the necessary setting to consider the iterates of linear operators. All results in this chapter can be found in the comprehensive books on functional analysis of Heuser [Heu82] and Rudin [Rud91].

Let us now briefly recall some notation. To this end, let X be a complex Banach space equipped with a norm $\|\cdot\|_X$. If the used norm is unambiguous we will just use the abbreviated version $\|\cdot\|$. The Banach algebra of bounded linear operators on X is denoted by $\mathcal{L}(X)$ equipped with the usual operator norm $\|\cdot\|_{op}$. The identity operator on X is $I \in \mathcal{L}(X)$. The corresponding topological dual space $\mathcal{L}(X, \mathbb{C})$ is denoted by $(X^*, \|\cdot\|_{X^*})$. The *range* and the *null space* of $T \in \mathcal{L}(X)$ is denoted by $\operatorname{ran}(T)$ and $\ker(T)$, respectively. The *closure* of $M \subset X$ is denoted by $\overline{M}$. Note that the results shown here are also applicable on real Banach spaces using a complexification as outlined in the end of Section 1.3.

8.1 Annihilators and complemented subspaces

Orthogonal complements are an important concept in Hilbert spaces. Annihilators are the corresponding generalization on Banach spaces. They are differently defined for a set on the Banach space and a set of functionals.

Definition 8.1 (Annihilator and Pre-annihilator). Let X be a Banach space and let $M \subset X$, $\Lambda \subset X^*$. The *annihilator* of M is the subspace

$$M^\perp = \{x^* \in X^* : x^*(x) = 0 \text{ for every } x \in M\} \subset X^*,$$

whereas the subspace

$$\Lambda_\perp = \{x \in X : x^*(x) = 0 \text{ for every } x^* \in \Lambda\} \subset X$$

is called the *pre-annihilator* of the set Λ.

The annihilator set $M^\perp$ contains all continuous linear functionals on X that vanish on M, while $\Lambda_\perp$ is the subset of X on which every bounded functional from Λ is zero. The next propositions state the most important properties of the annihilator and the pre-annihilator. Recall that the closure of a set $M \subset X$ is denoted by $\overline{M}$. For proofs and more properties we refer to Rudin [Rud91, 95 ff.].

Proposition 8.1. *Let X be a Banach space and let $M \subset X$ and $\Lambda \in X^*$. Then the following statements hold true:*

1. *$M^\perp$ is a closed subspace of X^*, and*

$$M^\perp = \overline{\operatorname{span}(M)}^\perp \quad \textit{and} \quad (M^\perp)_\perp = \overline{\operatorname{span}(M)}.$$

2. *$\Lambda_\perp$ is a closed subspace of X, and*

$$\Lambda_\perp = \operatorname{span}(\Lambda)_\perp \quad \textit{and} \quad \overline{\operatorname{span}(\Lambda)} \subset (\Lambda_\perp)^\perp.$$

3. *$\overline{\operatorname{span}(M)} = X$ if and only if $M^\perp = \{0\}$.*

The last item yields an elegant way of saying when a subspace $M \subset X$ is dense in X. We will use the second result later for the dual space X^* and for the pre-annihlator of a subspace $\Lambda \subset X^*$ to show that for finite many linear functionals there exists always $x \in X \setminus \{0\}$ such that $x \in \Lambda_\perp$.

Proposition 8.2. *Let X be a Banach space and let $M \subset X$ be a closed subspace. Furthermore, let $\pi : X \to X/M$ be the quotient map. Then*

$$(X/M)^* \cong M^\perp$$

via the isomorphism $\Phi : (X/M)^* \to M^\perp$, $\Phi(\alpha)(x) = \alpha(\pi(x))$.

Proposition 8.3. *Assume that X and Y are Banach spaces and let $T \in \mathcal{L}(X, Y)$. Then*

$$\ker(T^*) = T(X)^\perp, \quad \ker(T) = T^*(Y^*)_\perp.$$

Proof. See Rudin [Rud91, Thm 4.12 on p. 99]. □

Next, we will consider direct sum decompositions of a Banach space. In contrary to the finite-dimensional case, the closedness of the subspaces is necessary.

Definition 8.2 (Complemented subspace)**.** Let X be a Banach space and let $M, N \subset X$. Then X is said to be the *complemented sum* of M, denoted by $X = M \oplus N$, if for every $x \in X$ there are $m \in M$ and $n \in N$ such that $x = m + n$ has a unique representation.

We say that a closed subspace M of X is *complemented* in X, if there is a closed subspace N of X such that $X = M \oplus N$.

A natural question that arises is whether a given subspace has a complemented subspace in a Banach space. One way to answer the question is to use projections on closed subspaces. Bounded projections yield a canonical way to construct complemented subspaces by the range and the null space. To state this result, which can be found in Rudin [Rud91, 133 ff.], we will define first a projection on a Banach space X onto a given subset $M \subset X$.

Definition 8.3 (Projection)**.** Let X be a vector space. We call the linear mapping $P \in \mathcal{L}(X)$ a *projection* onto $M \subset X$, if

- $P(X) = M$ and
- $P^2 = P$ on X, i.e., P is idempotent.

The next proposition states that every bounded projection onto closed subspaces canonically yields a space decomposition, see, e. g., Rudin [Rud91, Theorem 5.16].

Proposition 8.4 (Complemented sum with projections)**.** *Let X be a Banach space and $P \in \mathcal{L}(X)$ a bounded projection onto a closed subspace of X. Then*

$$X = \operatorname{ran}(P) \oplus \ker(P).$$

This space decomposition based on a bounded projection are fundamental for the results in the next chapter. Concretely, we will construct a projection onto a finite-dimensional subspace of X. Using the projection, we have decomposed the space X into two closed subspaces.

8.2 Riesz-Schauder theory of compact operators

This section gives a brief introduction to the Riesz-Schauder theory of compact operators. For more details and proofs, we refer to Rudin [Rud91, pp. 103–111]. The Riesz-Schauder theory characterizes the class of compact operators. It will be shown that if T is a compact operator, the spectrum of T is countable with 0 as the only possible limit point, $T - \lambda I$ has closed range for all $\lambda \in \mathbb{C}$, and the dimensions of the spaces $\ker(T - \lambda I)$ and $\ker(T^* - \lambda I)$ are finite and equal. The latter property is important for the Fredholm theory where these kinds of operators get assigned the Fredholm index 0. More on Fredholm theory will be shown in the next section. Let us start with the definition of a compact operator.

Definition 8.4 (Compact operator). Let X, Y be Banach spaces and let $U \subset X$ be the open unit ball in X. $T \in \mathcal{L}(X, Y)$ is said to be *compact* if $\overline{T(U)}$ is compact in Y.

Clearly, if $\overline{\operatorname{ran}(T)}$ is compact in Y, then T is bounded, i. e., $T \in \mathcal{L}(X, Y)$. We denote the space of all compact operators from X to Y by $\mathcal{K}(X, Y)$. The next proposition gives a characterization of compact operators.

Proposition 8.5 (Rudin [Rud91, Thm. 4.18 and Thm. 4.19]). *Let X and Y be Banach spaces. Then the following statements hold true:*

1. $\mathcal{K}(X, Y)$ *is a Banach space, a closed subspace of* $\mathcal{L}(X, Y)$.
2. *If* $T \in \mathcal{L}(X, Y)$ *and* $\dim \operatorname{ran}(T) < \infty$, *then T is compact.*
3. *If* $T \in \mathcal{K}(X, Y)$ *and* $\operatorname{ran}(T)$ *is closed, then* $\dim \operatorname{ran}(T) < \infty$.
4. *If* $T \in \mathcal{K}(X)$ *and* $\lambda \neq 0$, *then* $\dim \ker(T - \lambda I) < \infty$.
5. *If* $\dim X = \infty$, $T \in \mathcal{K}(X)$, *then* $0 \in \sigma(T)$.
6. *If* $T \in \mathcal{L}(X, Y)$. *Then T is compact if and only if T^* is compact.*

The first item states in particular, that if $T_n \in \mathcal{K}(X, Y)$ converge towards T in the operator norm, then $T \in \mathcal{K}(X, Y)$. Combined with the second item, we get that a convergent series of finite-rank operators converge to a compact operators. Note that on a Hilbert space every compact operator is the limit of finite-rank operators, while this is not true on general Banach spaces which has been shown by a counterexample by Enflo [Enf73]. The property that every compact operator can be written as a limit of finite-rank operators is called the approximation property.

The following two propositions show how the operator $T - \lambda I$ behaves if t is a compact operator.

Proposition 8.6 (Rudin [Rud91, Thm. 4.23]). *If X is a Banach space, $T \in \mathcal{K}(X)$, and $\lambda \neq 0$, then $T - \lambda I$ has closed range.*

The next proposition characterizes the dimension of the eigenspaces of T and T^*. If T is compact, then $\dim \ker(T - \lambda I) = \dim \ker(T^* - \lambda I)$ holds for all $\lambda \in \mathbb{C}$. We will see later, that the difference of these dimensions defines the so called *Fredholm index*.

Proposition 8.7 (Rudin [Rud91, Thm. 4.25]). *Let X be a Banach space, $T \in \mathcal{K}(X)$. Then the space dimensions*

$$
\begin{aligned}
\alpha &= \dim \ker(T - \lambda I), \\
\beta &= \dim \left(X / \operatorname{ran}(T - \lambda I)\right), \\
\alpha^* &= \dim \ker(T^* - \lambda I), \\
\beta^* &= \dim \left(X^* / \operatorname{ran}(T^* - \lambda I)\right),
\end{aligned}
$$

are equal and finite.

Finally, compact operators posses special spectral properties. In particular, all spectral values except 0 are isolated eigenvalues in the spectrum. Moreover, the spectrum of a compact operator is always a countable set.

Proposition 8.8 (Rudin [Rud91, Thm. 4.25]). *Let $T \in \mathcal{K}(X)$, then*

1. *$\sigma(T)$ is countable,*
2. *if $\lambda \in \sigma(T) \setminus \{0\}$, then $\lambda \in \sigma_p(T)$, i. e., λ is an eigenvalue of T,*
3. *the only possible limit point of $\sigma(T)$ is 0.*

Accordingly, compact operators can be considered as an extension of matrices on finite-dimensional normed vector spaces. Analogous, we can define the geometrical and algebraic multiplicities. Note that both values are in fact finite.

Definition 8.5 (Multiplicities). The *geometrical multiplicity* of $T \in \mathcal{L}(X)$ is defined by

$$\dim \ker(T - \lambda I),$$

the dimension of the eigenspace accociated with $\lambda \in \sigma(T)$. The value of

$$\dim \bigcup_{k=1}^{\infty} \ker(T - \lambda I)^k$$

defines the *algebraic multiplicity* of $\lambda \in \sigma(T)$.

According to these results, the eigenvalues of each compact operator T can be listed by their modulus in decreasing order,

$$|\lambda_1(T)| \geq |\lambda_2(T)| \geq \cdots \geq 0,$$

where each eigenvalue is repeated with its algebraic multiplicity. In the next section, we will generalize these concepts where both multiplicities are finite.

8.3 Fredholm theory and Riesz operators

In the following, we will generalize the possibility to decompose the space X into the null space and the image of a linear operator. For these results we introduce now the most fundamental aspects of Fredholm theory of linear operators. More details can be found in the comprehensive books of Heuser [Heu82] and Ruston [Rus86].

Definition 8.6 (Ascent, Descent). Let X be a Banach space and let $T \in \mathcal{L}(X)$.

1. T is said to have *finite ascent* if there exists $k \in \mathbb{N}$ such that

   $$\ker(T^k) = \ker(T^{k+1}).$$

   Then we denote by $\operatorname{asc}(T)$ the smallest integer with this property and say $\operatorname{asc}(T)$ is the *ascent* of T.

2. T has *finite descent* if there exist $k \in \mathbb{N}$ such that

   $$\operatorname{ran}(T^k) = \operatorname{ran}(T^{k+1}).$$

Then we denote by $\mathrm{dsc}(T)$ the smallest integer with this property and call this number the *descent* of T.

The next proposition shows the relation of the ascent and the descent of a linear operator with complemented sum decompositions.

Proposition 8.9 ([Heu82, Prop. 38.1 and 38.2]). *Let $T \in \mathcal{L}(X)$ and let $m \geq 0$ be a integer. Then*

1. $\mathrm{asc}(T) \leq m < \infty$ *if and only if* $\ker(T^n) \cap \mathrm{ran}(T^m) = \{0\}$ *holds,*
2. $\mathrm{dsc}(T) \leq m < \infty$ *if and only if* $X = \ker(T^m) + \mathrm{ran}(T^n)$.

In both statements the integer $n > 0$ can be chosen arbitrarily.

If the ascent and the descent of a linear operator are finite, then they are equal, as the next proposition states. In this case, the operator T is said to have *finite chain length*.

Proposition 8.10 ([Heu82, Prop. 38.3]). *If* $\mathrm{asc}(T) < \infty$ *and* $\mathrm{dsc}(T) < \infty$, *then both values are equal, i. e.,* $\mathrm{asc}(T) = \mathrm{dsc}(T)$ *holds.*

Having the requirements of the last proposition satisfied, then such an operator yields a direct sum decomposition in the following way:

Proposition 8.11 ([Heu82, Prop. 38.4]). *Let $T \in \mathcal{L}(X)$ have finite chain length $p =$* $\mathrm{asc}(T) = \mathrm{dsc}(T) < \infty$. *Then X can be decomposed into*

$$X = \mathrm{ran}(T^p) \oplus \ker(T^p). \tag{8.1}$$

We will ask later when this space decomposition can be derived by so called spectral projections.

To better understand the properties of operators with finite chain length we now introduce the concept of Fredholm operators. First we define the nullity and the deficiency of an operator as dimension of the null space of T and T^*, respectively.

Definition 8.7 (Nullity and Deficiency). Let $T \in \mathcal{L}(X)$. We denote by

$$\alpha(T) := \dim(\ker(T))$$

the *nullity* of T and by

$$\beta(T) := \dim(\ker(T^*))$$

the *deficiency* of the operator T.

The next definition characterizes these Fredholm operators as the special class of operators, where the nullity and the deficiency are both finite.

Definition 8.8 (Fredholm Operators). Let X, Y be two Banach spaces. Then the set of Fredholm operators is defined by

$$\Phi(X,Y) = \{T \in \mathcal{L}(X,Y) \,:\, \alpha(T) < \infty \text{ and } \beta(T) < \infty\}.$$

	ran(T) closed	$\alpha(T) = \beta(T) < \infty$	asc(T) = dsc$(T) < \infty$
T Fredholm	yes	not necessarily	not necessarily
T Weyl	yes	yes	not necessarily
T Browder	yes	yes	yes

Table 8.1: *Comparison between Fredholm, Weyl and Browder operators.*

Then the *index* of $T \in \Phi(X, Y)$ is defined by $\operatorname{ind}(T) = \alpha(T) - \beta(T)$.

Now we can relate the concept of Fredholm operators, i. e., the nullity and the deficiency, with the concept of the ascent and descent of an operator $T \in \mathcal{L}(X)$.

Proposition 8.12 ([Heu82, Prop. 38.5 and 38.6]). *If T is a linear operator on a Banach space X then the following properties hold:*

1. *If* $\operatorname{asc}(T) < \infty$, *then* $\alpha(T) \le \beta(T)$.
2. *If* $\operatorname{dsc}(T) < \infty$, *then* $\alpha(T) \ge \beta(T)$.
3. *If* $\operatorname{asc}(T) = \operatorname{dsc}(T) < \infty$, *then* $\alpha(T) = \beta(T)$ *(possibly infinite).*
4. *If* $\alpha(T) = \beta(T) < \infty$, *and if either* $\operatorname{asc}(T)$ *or* $\operatorname{dsc}(T)$ *is finite, then* $\operatorname{asc}(T) = \operatorname{dsc}(T)$.

According to the first item in the last proposition, we can identify all Fredholm operators $T \in \Phi(X)$ with finite ascent, $p = \operatorname{asc}(T) < \infty$, as operators where

$$\dim(\ker(T)) \le \dim(\ker(T^*)) < \infty.$$

Thus, T has a Fredholm index less or equal than zero, $T \in \Phi_-(X)$. Operators of this kind will play an important role in the beginning of the chapter, where we consider the iterates of an operator and its corresponding fixed point space. To obtain the limiting operator, it will be finally shown that operators with index 0 are the appropriate choice.

Fredholm operators with index 0 have special spectral properties. H. Weyl [Wey09] has considered in 1909 operators where the spectrum can be partitioned into a set of isolated eigenvalues of finite multiplicity and a remaining set. Accordingly, Fredholm operators with index 0 are named after H. Weyl.

Definition 8.9 (Weyl Operator). A bounded operator $T \in \mathcal{L}(X)$ is said to be a *Weyl operator* if T is a Fredholm operator with index 0. The class of all Weyl operators on X will be denoted by $\mathcal{W}(X)$.

Definition 8.10 (Browder Operator). A bounded operator $T \in \mathcal{L}(X)$ is said to be a *Browder operator* if it is a Fredholm operator with finite chain length. We will denote the sets of all Browder operators on X by $\mathcal{W}_B(X)$.

Each Browder operator T is in fact a Weyl operator, as by definition $\operatorname{asc}(T) = \operatorname{dsc}(T) < \infty$ and $\alpha(T) < \infty, \beta(T) < \infty$ holds. By item 4 of Proposition 8.12 we conclude that

$$\alpha(T) = \beta(T) < \infty.$$

Therefore, we get the relation

$$\mathcal{W}_B(X) \subset \mathcal{W}(X).$$

A comparison between both classes is shown in Table 8.1. Consider now a Browder operator T having finite ascent p, then

1. $\dim(\ker(T)) = \dim(\ker(T^*)) < \infty$,
2. $\ker(T^p) \cap \operatorname{ran}(T^p) = \{0\}$,
3. $X = \ker(T^p) \oplus \operatorname{ran}(T^p)$.

In the next chapter, we are interested in operators $T \in \mathcal{L}(X)$ where $(T - \lambda I)$ is a Browder operator. In that case, we will prove the last item in a constructive way using a projection for the space decomposition.

Next, we introduce the concept of so called Riesz operators as generalization of compact operators. They admit similar spectral properties and share also finite multiplicities. Recall, that if $K \in \mathcal{L}(X)$ is compact, then $K - \lambda I \in \mathcal{W}(X)$, see Proposition 8.7. Furthermore, every eigenvalues except 0 is isolated. We will define a Riesz operator as an operator, where $T - \lambda I$ is a Fredholm operator for all $\lambda \in \mathbb{C} \setminus \{0\}$. We will see in the next proposition that this definition is in fact equivalent to the fact that $T - \lambda I$ is a Weyl operator.

Definition 8.11 (Riesz-Operators)**.** A bounded operator $T \in \mathcal{L}(X)$ on a Banach space X is said to be a *Riesz operator* if $T - \lambda I \in \Phi(X)$ for every $\lambda \in \mathbb{C} \setminus \{0\}$. We denote the set of all Riesz operators on X by $\mathcal{R}(X)$.

The next proposition characterizes Riesz operators. These results can be found, e. g., in Aiena [Aie04, Section 3.9].

Proposition 8.13 (Characterization of a Riesz operator)**.** *Let $T \in \mathcal{L}(X)$ where X is a Banach space. Then the following statements are equivalent:*

1. *T is a Riesz operator,*
2. *$T - \lambda I \in \mathcal{L}(X)$ for every $\lambda \in \mathbb{C} \setminus \{0\}$,*
3. *$T - \lambda I \in \mathcal{W}(X)$ for every $\lambda \in \mathbb{C} \setminus \{0\}$, and finally,*
4. *every spectral point $\lambda \neq 0$ is isolated and $\cup_{k \in \mathbb{N}} \ker(T - \lambda I)^k$ is finite-dimensional.*

The next theorem gives another characterization of Riesz operators in terms of the limit of the iterates. Namely, Ruston [Rus86] has shown that the iterates of Riesz operators can be approximated closely by compact operators.

Proposition 8.14 (Ruston Condition for Riesz operators). *Let $T \in \mathcal{L}(X)$. Then T is a Riesz operator, $T \in \mathcal{R}(X)$, if and only if*

$$\lim_{n\to\infty} \left(\inf_{K \in \mathcal{K}(X)} \|T^n - K\|_{op} \right)^{1/n} = 0.$$

This limit is called the *essential spectral radius* of T and will be discussed in more detail in the next section. Based on the essential spectral radius we will consider a further generalization, the so called quasi-compact operators. The main property is that every spectral value outside the essential spectral radius is isolated. As the essential spectral radius of every Riesz operator is zero by Proposition 8.14, every spectral value $\lambda \neq 0$ is isolated, see also the last item of Proposition 8.13. Let us consider finally two well known classes of Riesz operators.

Example 8.1 (Riesz operators).
Finite-rank operators and compact operators are special cases of Riesz-operators, and we have the inclusions $\mathcal{F}(X) \subset \mathcal{K}(X) \subset \mathcal{R}(X)$. To see this note that

1. every finite-rank operator is a compact operator, i.e., $\mathcal{F}(X) \subset \mathcal{K}(X)$, and
2. every compact operator is a Riesz operator, i.e., $\mathcal{K}(X) \subset \mathcal{R}(X)$.

The first item is already stated in Proposition 8.5. The second item follows for instance by the spectral properties of compact operators as shown in Proposition 8.8. Note that we derive both facts also by the result for compact operators of Proposition 8.7 as $T - \lambda I$ is a Weyl operator for every $\lambda \in \mathbb{C} \setminus \{0\}$.

8.4 Spectral sets and quasi-compact operators

We consider here special spectral sets that characterize isolated points in the spectrum. Of our interest will be the essential spectrum and the essential spectral radius in order to define quasi-compact operators. Quasi-compact operators have the special property that every peripheral spectral points is isolated. We start with the definition of the resolvent set and the classical spectrum of a linear operator.

For $T \in \mathcal{L}(X)$, the *resolvent set* $\rho(T)$ can be defined in the following ways:

$$\begin{aligned}
\rho(T) &:= \{\lambda \in \mathbb{C} : T - \lambda I \text{ is invertible}\} \\
&:= \{\lambda \in \mathbb{C} : \alpha(T - \lambda I) = \beta(T - \lambda I) = 0\} \\
&:= \{\lambda \in \mathbb{C} : \operatorname{asc}(T - \lambda I) = \operatorname{dsc}(T - \lambda I) = 0\}.
\end{aligned}$$

The resolvent set is an open set in $\mathbb{C}$. The *resolvent* of T corresponding to $\lambda \in \mathbb{C}$ is defined by $R(T, \lambda) := (T - \lambda I)^{-1}$. The *spectrum* of T, denoted by $\sigma(T)$, is a nonempty compact set of $\mathbb{C}$ and is defined as $\mathbb{C} \setminus \rho(T)$ and hence, contains all values $\lambda \in \mathbb{C}$ such that $T - \lambda I$ is not invertible. The value $r(T) := \sup\{|\lambda| : \lambda \in \sigma(T)\}$ is said to be the *spectral radius* of T and can be calculated using Gelfand's formula

$$r(T) = \lim_{n\to\infty} \|T^n\|_{op}^{1/n}.$$

The spectral radius $r(T)$ is bounded by $\|T\|_{op}$. The *peripheral spectrum* $\sigma_{\text{per}}(T)$ describes the boundary of the spectrum, $\sigma_{\text{per}}(T) := \sigma(T) \cap \{\lambda \in \mathbb{C} : |\lambda| = r(T)\}$. By $\sigma_p(T)$, we denote the *point spectrum* of T,

$$\sigma_p(T) := \{\lambda \in \mathbb{C} : T - \lambda I \text{ is not one-to-one}\},$$

which contains all the eigenvalues of T. Of our interest here are is also the *(Weyl) essential spectrum* of T, denoted by $\sigma_{ess}(T)$, and the *(Browder) essential spectrum*, denoted by $\sigma_b(T)$, which are defined by

$$\sigma_{\text{ess}}(T) := \{\lambda \in \mathbb{C} : T - \lambda I \notin \mathcal{W}(X)\},$$
$$\sigma_b(T) := \{\lambda \in \mathbb{C} : T - \lambda I \notin \mathcal{W}_B(X)\}.$$

Related to these essential spectra is the *essential spectral radius* $r_{\text{ess}}(T)$ which is defined as

$$r_{ess}(T) := \sup\{|\lambda| : \lambda \in \sigma_{ess}(T)\} = \lim_{n\to\infty}\left(\inf_{K\in\mathcal{K}(X)} \|T^n - K\|_{op}\right)^{1/n}.$$

In contrast to Hilbert spaces there are several ways to define the essential spectrum on Banach spaces that are not equivalent. The following two definitions are in fact equal on Hilbert spaces. If the essential spectrum is considered as the largest subset of the spectrum which remains invariant under compact perturbations, then this leads to the definition of $\sigma_{\text{ess}}(T)$, which also often said to be the *essential Weyl spectrum* according to H. Weyl [Wey09], see also M. Schechter [Sch66]. A point $\lambda \in \sigma(T)$ is in the essential spectrum if it does not have all of the following properties:

1. $\alpha(T - \lambda I) < \infty$,
2. $\beta(T - \lambda I) < \infty$,
3. $\alpha(T - \lambda I) = \beta(T - \lambda I)$, and
4. $\text{ran}(T - \lambda I)$ is closed.

However, this definition of the spectrum does not contain the limit points of the spectrum. If all these accumulation points are added to the essential spectrum, then we derive the definition of F. E. Browder [Bro61]. There, a spectral value $\lambda \in \sigma(T)$ is in the essential spectrum, if at least one of the following conditions hold:

1. $\text{ran}(T - \lambda I)$ is not closed in X,
2. λ is a limit point of the spectrum $\sigma(T)$,
3. $\bigcup_{k\in\mathbb{N}} \ker(T - \lambda I)^k$ is infinite dimensional.

These conditions are equivalent to the *essential Browder spectrum* $\sigma_b(T)$ defined above. The advantage of using σ_{ess} is the perturbation invariance, while the advantage of the Browder spectrum $\sigma_b(T)$ is that $\sigma(T) \setminus \sigma_b(T)$ is a countable set. Summing up these facts, we have the relation

$$\sigma_{ess}(T) \subseteq \sigma_b(T) = \sigma_{ess}(T) \cup \operatorname{acc} \sigma(T) \subset \sigma(T),$$

where $\operatorname{acc} \sigma(T)$ denotes all the limit points of $\sigma(T)$. Nevertheless, the essential spectral radius $r_{ess}(T)$ is the same in both definitions of the essential spectrum, i. e., all spectral limit points are on the boundary of $\sigma_{ess}(T)$ and all spectral values outside of $r_{ess}(T)$ are isolated. Besides, if $\lambda \in \sigma(T)$ with $|\lambda| > r_{ess}(T)$ then $T - \lambda I$ is a Browder operator, as $\alpha(T - \lambda I) = \beta(T - \lambda I) < infty$ and $\cup_{k\in\mathbb{N}} \ker(T - \lambda I)^k$ is finite-dimensional, thus $T - \lambda I$ has finite chain length.

We have already introduced Riesz operators in the last section as generalization of compact operators. Based on the Ruston condition, Proposition 8.14, the essential spectral radius provides an alternative way to define Riesz operators T by the relation $r_{ess}(T) = 0$. Based on the preceding notes we can derive the following spectral properties as already stated in the last item of Proposition 8.13.

Proposition 8.15. *If T is a Riesz operator, then each spectral point $\lambda \neq 0$ is isolated and the corresponding generalized eigenspace $\bigcup_{k\in\mathbb{N}} \ker(T - \lambda I)^k$ is finite-dimensional.*

In the next chapter, we discuss contraction operators T, i. e., $\|T\|_{op} \leq 1$. As shown in Section 6.2, a necessary condition for the convergence of the iterates is that $\sigma(T) \subset B(0,1) \cup \{1\}$. We are interested in operators, where 1 is an isolated point in the spectrum. Therefore, we are considering in the following operators with essential spectral radius strictly less than 1.

Definition 8.12 (Quasi-compact operator)**.** An operator $T \in \mathcal{L}(X)$ on a Banach space X is said to be *quasi-compact*, if $r_{\text{ess}}(T) < 1$.

In fact, being quasi-compact is weaker than being Riesz and thus, the following chain of implications hold:

$$\text{finite-rank} \quad \Rightarrow \quad \text{compact} \quad \Rightarrow \quad \text{Riesz} \quad \Rightarrow \quad \text{quasi-compact}$$

A comparison of these operators based on the nullity, deficient and ascent is given in Table 8.2. Note that if $\|T\|_{op} = 1$ and T is quasi-compact, then every peripheral spectral value is not contained in the essential spectrum. In the next proposition, it will be shown that such a spectral point is already an eigenvalue of T.

Proposition 8.16 (Sasser [Sas64])**.** *Let $T \in \mathcal{L}(X)$ be a quasi-compact operator. Then if $\lambda_0 \in \sigma(T)$ and $|\lambda_0| = r(T)$, then λ_0 is an isolated point in $\sigma(T)$ and is in the point spectrum.*

We will provide an explanation of this result in the next section using so called spectral projections. The main argument is that if T is quasi-compact, then $T - \lambda_0 I$ is a Browder operator if $|\lambda_0| = r(T)$. In this case, λ is an isolated eigenvalue.

	T compact	T Riesz	T quasi-compact
$r_{\mathrm{ess}}(T)$	$= 0$	$= 0$	< 1
$\alpha(T - \lambda I) = \beta(T - \lambda I) < \infty$	$\lambda \neq 0$	$\lambda \neq 0$	$\lvert\lambda\rvert > r_{\mathrm{ess}}(T)$
$\mathrm{asc}(T - \lambda I) < \infty$	$\lambda \neq 0$	$\lambda \neq 0$	$\lvert\lambda\rvert > r_{\mathrm{ess}}(T)$
$T - \lambda I \in \mathcal{W}_B(X)$	$\lambda \neq 0$	$\lambda \neq 0$	$\lvert\lambda\rvert > r_{\mathrm{ess}}(T)$

Table 8.2: *Properties of compact, Riesz and quasi-compact operators. Note that if $r(T) = \|T\|_{op} = 1$, then $T - \lambda I$ is a Browder operator for every peripheral spectral value λ, i. e., $|\lambda| = 1$, and λ is an isolated eigenvalue of T.*

8.5 Spectral projections

We conclude this chapter with the introduction of spectral projection. It has been shown in Dunford [Dun43b], that the iterates of an operator converge against a spectral projection corresponding to the spectral set 1. We will first show that the spectral projection leads to a space decomposition by its range and its null space. Then, we relate the chain length of an operator with corresponding spectral projection. Recall, that we denote by $R(T, \lambda)$ the resolvent of T, $R(T, \lambda) := (T - \lambda I)^{-1}$.

In the following X is considered as complex Banach space and $T \in \mathcal{L}(X)$. Important to define spectral projection are spectral sets. These are closed subsets of the spectrum that have always a positive distance to its complement. This is needed to create a integration path in the spectrum that contains only the spectral set.

Definition 8.13 (Spectral Set). A subset σ of $\sigma(T)$ is said to be a *spectral set* of T if σ and $\sigma(T) \setminus \sigma$ are closed.

Note that this definition is equivalent to say that there exists open sets $U_1 \supset \sigma$ and $U_2 \supset \sigma(T) \setminus \sigma$ such that $U_1 \cap U_1 = \varnothing$. Using functional calculus, see e. g., Heuser [Heu82, pp. 204], one can define spectral projections associated with a spectral set in the following way.

Definition 8.14 (Spectral projection). Let σ be a (possible empty) spectral set and let Γ_σ be a simple, closed integration path that lies in in the resolvent set $\rho(T)$. Further, we assume that Γ_σ is oriented counterclockwise and encloses σ. Then the operator

$$P_\sigma := \frac{1}{2\pi i} \int_{\Gamma_\sigma} R(T, \lambda) \mathrm{d}\lambda \tag{8.2}$$

is a bounded projection on X and is called the *spectral projection* associated with σ.

We show next that λ is a pole of the resolvent $R(T, \lambda)$ if and only if the related operator $T - \lambda I$ has finite chain length, i. e., $\mathrm{asc}(T - \lambda I) < \infty$. In this case, the null space and the range of the spectral projection associated with $\{\lambda\}$ are explicitly given. If furthermore $T - \lambda I$ is a Fredholm operator, i. e., $\mathrm{asc}(T - \lambda I) < \infty$ and $\mathrm{dsc}(T - \lambda I) < \infty$, then λ is always an isolated eigenvalue of T and the spectral projection is finite-dimensional.

Proposition 8.17 ([Heu82, Prop. 50.2]). *λ is a pole of the resolvent of T if and only if $T - \lambda I$ has positive finite chain length. The common chain-length p is the order of the pole. In this case $\lambda \in \sigma_p(T)$, i. e., λ is an eigenvalue of T. The spectral projector P corresponding to $\{\lambda\}$ satisfies*

$$\operatorname{ran}(P) = \ker(T - \lambda I)^p \quad \textit{and} \quad \ker(P) = \operatorname{ran}(T - \lambda I)^p.$$

Proposition 8.18 ([Heu82, Prop. 50.3]). *$T - \lambda I$ is a Browder operator, i. e., $T - \lambda I$ is a Fredholm operator with positive finite chain length*

$$0 < \operatorname{asc}(T - \lambda I) < \infty,$$

if and only if λ is an isolated spectral point of T and the corresponding spectral projector P is finite-dimensional. In this case λ is a pole of the resolvent $R(T, \lambda)$.

The computation of the spectral projection using the formula provided in (8.2) is in general hard to calculate. We will consider operators T where $T - \lambda I$ is a Browder operator in the next chapter and show a constructive way based on the invertibility of a Gramian matrix to derive the spectral projector P. Table 8.2 shows that quasi-compact operators with $r(T) = \|T\|_{op} = 1$ guarantee that $T - \lambda I$ is a Browder operator for every peripheral spectral value λ. By the preceding propositions, we can conclude that λ is an isolated eigenvalue and $X = \ker(T - \lambda I)^p \oplus \operatorname{ran}(T - \lambda I)^p$ where p is the ascent of $T - \lambda I$. Compare this result with Proposition 8.11 and Proposition 8.16 and note that $\ker(T - \lambda I)^p$ is finite-dimensional.

CHAPTER 9 Iterates of Quasi-Compact Operators by Fredholm Theory

"The elegance of a mathematical theorem is directly proportional to the number of independent ideas one can see in the theorem and inversely proportional to the effort it takes to see them."

GEORG PÓLYA

We discuss here the limit of the iterates of linear operators defined on a complex Banach space X and show under which conditions those iterates converge uniformly towards a limiting operator. In particular, we show as highlight of this thesis a constructive way to obtain the limit of the iterates provided that the fixed points and dual fixed points are explicitly known, which is often the case for approximation operators. We will give new necessary and sufficient criteria for the convergence of the iterates and embed the developed theory in existing results.

First, we give an overview of existing results on this field starting with fundamental work of Dunford published already in 1943. A topic closely related to the limit of the iterates uniform ergodic theory. Uniform ergodic theorems prove the desired result for the limit of the mean value of the iterates which is in fact weaker and not sufficient for our application, namely obtaining lower estimates as already outlined in Chapter 6. To show the simplicity of our approach, we will provide an example for variation-diminishing operators before we state our results. After this introduction, we discuss how finite-dimensional generalized eigenspaces of a bounded operator can be complemented in the Banach spaces by the use of projections. We will start using the classical coordinate map to show the principle of our approach. It turns out that this coordinate map extended to X is a spectral projection and naturally gives the complement of the generalized eigenspace as null space of the coordinate map. It will be shown that when using a Gramian matrix where the dual generalized eigenfunctions are operating on the generalized eigenfunctions, the range of the projection is invariant under T if and only if the Gramian matrix is invertible.

These results are applied next in order to prove the limiting behaviour of the iterates of an operator. It will be shown that if the spectrum of the operator T is contained in the unit ball where 1 is the only spectral value on the boundary and $T-I$ is a Browder operator with ascent one, then the iterates T^n converge uniformly towards the spectral projection associated with the eigenvalue 1. This result is constructive in the sense that the limit operator can be explicitly calculated by the inverse of the Gramian matrix.

9.1 Fundamentals

We will state here the fundamental results of Dunford [Dun43b; Dun43a] and show their relation to the convergence of the iterates or related ergodic theorems. We will cover both topics independently and give an overview over necessary and sufficient conditions for the convergence.

9.1.1 The work of Dunford

Let $\mathcal{H}(T)$ denote the class of all complex functions of a complex variable which are regular at every point of an open set containing the eigenvalues $\lambda_1, \dots, \lambda_k$ of T. Let us denote by D the differential operator applied to the complex variable. Recall that a sequence of operators $(T_n)_{n\in\mathbb{N}}$ on a Banach space X is said to converge *uniformly* towards $T \in \mathcal{L}(X)$ if it converges in the operator norm. A sequence of operators T_n converges strongly if $\|T_n x - Tx\|_Y \to 0$ converges for $n \to \infty$ and all $x \in X$. If $|\alpha^*(T_n x) - \alpha^*(Tx)| \to 0$ converges for $n \to \infty$ for all $x \in X$, $\alpha^* \in Y^*$ then the sequence $(T_n)_{n\in\mathbb{N}}$ converges in the weak operator topology.

Proposition 9.1 ([Dun43b, Thm. 3.13]). *Let $p(\lambda) = \prod_{i=1}^{k}(\lambda_i - \lambda)^{\nu_i}$ be a polynomial whose distinct roots are $\lambda_1, \dots, \lambda_k$. Let $f_n \in \mathcal{H}(T)$ be such that*

1. $f_n(\lambda_i) \to 1$, $D^j f_n(\lambda_i) \to 0$, $i = 1, \dots, k$, $j = 1, \dots, \nu_i - 1$.
2. $f_n(T) \to P$ *weakly,* $P^2 = P$, $\operatorname{ran}(P) = \ker(p(T))$.

Then

3. $f_n(T) \to P$ *strongly if and only if* $p(T)f_n(T) \to 0$ *strongly,*
4. $f_n(T) \to P$ *uniformly if and only if* $p(T)f_n(T) \to 0$ *uniformly and* $\operatorname{ran}(p(T))$ *is closed,*

where the limits are taken for $n \to \infty$.

Note that for the strong convergence $\operatorname{ran}(P)$ is not necessarily closed while in the case of uniform convergence $\operatorname{ran}(P)$ is closed. However,

$$X = \overline{\operatorname{ran}(P)} \oplus \ker(P)$$

holds in any case.

Of our interest is the special case $p(T) = T - I$, thus $k = 1$ and $\lambda_1 = 1$. In this case, Dunford has shown the following.

Proposition 9.2 ([Dun43b, Thm. 3.16 on p. 215]). *Let $f_n \in \mathcal{H}(T)$ satisfy $f_n(1) \to 1$ and $(T - I)f_n(T) \to 0$. Then the following statements are equivalent.*

1. $f_n(T) \to P$, $P^2 = P$, $\operatorname{ran}(P) = \ker(T - I)$.

2. *The point* $\lambda = 1$ *is either in* $\rho(T)$ *or a pole of* $R(\lambda, T)$.
3. *The point* $\lambda = 1$ *is either in* $\rho(T)$ *or a simple pole of* $R(\lambda, T)$.
4. $X = \ker(T - I) \oplus \operatorname{ran}(T - I)$ *and* $\operatorname{ran}(T - I)$ *is closed.*
5. $\operatorname{ran}(T - I)^2$ *is closed.*

Here, the convergence is understood in any of either the weak, the strong or the uniform operator topology.

See also Theorem 8 and 9 in [Dun43a, pp. 648,649]. The fourth item states particularly that $T - I$ has finite chain length one, thus both ascent and descent of $T - I$ equals one. Note that this is in fact equivalent to the third item by Proposition 8.17. If additionally P is a finite-rank projection, we can conclude that $T - I$ is a Browder operator with ascent one. We will show later how this relates to the quasi-compactness of T. Now we will discuss two popular choices of $f_n \in \mathcal{F}(T)$ for the case $p(T) = T - I$.

1. $f_n(\lambda) = n^{-1} \sum_{k=0}^{n-1} \lambda^k$, and
2. $f_n(\lambda) = \lambda^n$.

In the first option, the necessary condition $(T - I) f_n(T) \to 0$ of the preceding proposition reduces to $T^n/n \to 0$ for $n \to \infty$. To see this, we rewrite the expression to

$$(T - I)(n^{-1} \sum_{k=0}^{n-1} T^k) = n^{-1} \sum_{k=1}^{n} T^k - n^{-1} \sum_{k=0}^{n-1} T^k = n^{-1} T^n - n^{-1} I.$$

Clearly, $n^{-1}T^n - n^{-1}I$ converges to zero if and only if $n^{-1}T^n$ converges to zero. If we consider the second option, the necessary condition requires that $T^{n+1} - T^n \to 0$ uniformly for $n \to \infty$. Also note that the second option implies the first one, thus if the iterates converge uniformly then also the so called Cesáro means converge uniformly. For this reason, we consider in the following also ergodic theorems and give necessary and sufficient conditions.

9.1.2 Uniform ergodic theorems

Given an operator $T \in \mathcal{L}(X)$, we discuss the convergence of the Cesáro means

$$a_n(T) := n^{-1} \sum_{k=0}^{n-1} T^k.$$

In the literature, this kind of convergence is discussed by so called uniform ergodic theorems. Lin [Lin74] has considered the condition under which $\|T^n/n\|_{op} \to 0$ holds and could specify another criterion for the convergence of the Cesáro means $a_n(T)$. Combined with the criteria shown previously by N. Dunford [Dun43b] we get the following proposition, where the last item is due to Lin.

Proposition 9.3. *Let* T *be a bounded linear operator on a Banach space* X *satisfying* $\|n^{-1}T^n\|_{op} \to 0$ *for* $n \to \infty$. *Then the following conditions are equivalent:*

1. *There exists a bounded linear operator P such that*

$$\left\| n^{-1} \sum_{k=0}^{n-1} T^k - P \right\|_{op} \to 0,$$

   $P^2 = P$, *and* $\operatorname{ran}(P) = \ker(T - I)$ *holds.*
2. $\operatorname{ran}(T - I)$ *is closed and* $X = \ker(T - I) \oplus \operatorname{ran}(T - I)$.
3. $\operatorname{ran}(T - I)^2$ *is closed.*
4. $\operatorname{ran}(T - I)$ *is closed.*

Accordingly, if $\|n^{-1}T^n\|_{op}$ converges to 0 as $n \to \infty$ and $\operatorname{ran}(T - I)$ is closed, then T is uniformly ergodic, i. e., there exists P such that $a_n(T) \to P$ uniformly. Note that $\operatorname{ran}(T - I)$ is always closed if T is a quasi-compact operator, see for instance the definition of the essential spectrum in Section 8.4.

An interesting relation between the fixed points of T and T^* has been shown by Sine [Sin70]:

Proposition 9.4 ([Sin70]). *Let T be a contraction, i. e., $\|T\|_{op} \le 1$. Then*

$$a_n(T) := n^{-1} \sum_{k=0}^{n-1} T^k$$

converges in the strong operator topology if and only if the fixed points of T separate the fixed points of T^*.

In the proof Sine has used the fact that the Cesáro means $a_n(T)$ converge in the weak operator topology if and only if they converge in the strong operator topology provided $\|T\|_{op} \le 1$. We show in Section 9.4.4, how this point separation property relates to the invertibility of the Gramian matrix that we will consider throughout the next sections.

For positive contractions on $C(K)$, where K is a compact Hausdorff space, M. Lin has proved the following equivalent conditions in [Lin75]. Note that a bounded operator T on $C(K)$ is said to be *positive* if $Tf \ge f$ holds whenever $f \ge 0$ Besides, a bounded operator T is called a *contraction* if $\|T\|_{op} \le 1$ holds.

Proposition 9.5. *Let T be a positive contraction of $C(K)$, where K is a compact Hausdorff space. Then the following conditions are equivalent.*

1. *T is quasi-compact.*
2. *$n^{-1} \sum_{k=1}^{n} T^k$ converges for $n \to \infty$ uniformly to a finite dimensional projection.*
3. $\operatorname{ran}(T - I)$ *is closed and* $\ker(T - I)$ *is finite dimensional.*
4. $\operatorname{ran}(T^* - I)$ *is closed and the space of invariant measures* $\ker(T^* - I)$ *is finite dimensional.*
5. $\operatorname{ran}(T - I)$ *is closed and* $\ker(T^* - I)$ *is finite dimensional.*

Corollary 9.6. *Let T be a positive contraction of $C(K)$. If $n^{-1}\sum_{k=1}^{n} T^k$ converges uniformly to a finite dimensional projection, then, for every $l \geq 1$,*

$$n^{-1}\sum_{k=1}^{n} T^{kl}$$

converges uniformly.

After the preceding result stated in [Lin75], M. Lin mentioned that this corollary does not hold when the projection is infinite-dimensional.

9.1.3 Convergence of the iterates

Next, we consider the case where $f_n(T) := T^n$. We have already shown previously that Dunford's criterion reduces in this case to $T^{n+1} - T^n \to 0$ for $n \to \infty$. To fulfill the condition $T^{n+1} - T^n \to 0$ the result of Katznelson and Tzafriri [KT86] yields a sufficient and necessary criterion. Katznelson and Tzafriri state that if T is a contraction on a Banach space then $\left\|T^{n+1} - T^n\right\|_{op} \to 0$ holds if and only if the intersection of the spectrum of T and the unit circle contains at most the point 1.

Proposition 9.7 (Katznelson and Tzafriri [KT86, Thm. 1]). *Let T be an linear operator on X such that $\|T\|_{op} \leq 1$. Then*

$$\lim_{m\to\infty} \left\|T^{m+1} - T^m\right\|_{op} = 0$$

if and only if

$$\sigma(T) \subset B(0,1) \cup \{1\}. \tag{9.1}$$

Clearly, if $\|T\|_{op} < 1$, then $\|T^n\|_{op} \to 0$. Thus, of our interest will be the case where the operator norm of T is equal to one.

Using the result of Proposition 9.7, the Proposition 9.2 reduces to the following proposition.

Proposition 9.8. *Let T be an operator such that $\|T\|_{op} \leq 1$ and $\sigma_{\text{per}}(T) \subset \{1\}$. Then the following statements are equivalent.*

1. *$T^n \to P$ uniformly, $P^2 = P$, $\operatorname{ran}(P) = \ker(T - I)$.*
2. *The point $\lambda = 1$ is either in $\rho(T)$ or else a simple pole of $R(T, \lambda)$.*
3. *$X = \ker(T - I) \oplus \operatorname{ran}(T - I)$ and $\operatorname{ran}(T - I)$ is closed.*
4. *$\operatorname{ran}(T - I)^2$ is closed.*

The next sections are devoted to the construction of the limiting operator P provided that $T - I$ is a Browder operator with ascent one. In this case $\operatorname{ran}(T - I)$ is closed and we get using the spectral projection P the space decomposition

$$X = \ker(T - I) \oplus \operatorname{ran}(T - I),$$

see also Proposition 8.17. Recall, that the uniform convergence of the iterates is important to derive lower estimates as shown in Chapter 6.

9.2 An introductory example

To motivate the theory of this chapter and to demonstrate the simplicity and the elegance of our results, we will provide here an example using a variation diminishing operator. We will sketch first our results for quasi-compact operator where $T - I$ has ascent one. To this end, let $T \in \mathcal{L}(X)$ be a quasi-compact operator on a Banach space X with $\|T\|_{op} = 1$ such that $T - I$ has ascent one. Furthermore, we assume that 1 is the only spectral value on the unit circle. By Proposition 8.16 and Proposition 8.18 the fixed point space $\ker(T - I)$ is finite dimensional. The same holds true for the dual fixed point space, as $T - I$ is a Browder operator and $\alpha(T - I) = \beta(T - I)$. Accordingly, we consider the following spaces

$$M = \ker(T - I) = \{x \in X : Tx = x\},$$
$$\Lambda = \ker(T^* - I) = \{x^* \in X^* : x^*(Tx) = x^*(x) \text{ for all } x \in X\}.$$

We have that $n = \dim(M) = \dim(\Lambda)$ is a positive integer, hence these spaces have a finite basis

$$M = \operatorname{span}\{e_1, \dots, e_n\} \text{ and } \Lambda = \operatorname{span}\{e_1^*, \dots, e_n^*\},$$

and w.l.o.g. we assume that these bases are normalized.

By Katznelson and Tzafriri [KT86, Thm. 1] we get that $T^{n-1} - T^n \to 0$ in the uniform operator topology and by Dunford [Dun43b] we can conclude by Proposition 9.8 that the iterates T^n converge uniformly towards a projection P with $\operatorname{ran}(P) = \ker(T - I)$, provided that $X = \ker(T - I) \oplus \operatorname{ran}(T - I)$. In the following, we will answer the question how to calculate P if bases of the above mentioned spaces are known. Bases for these fixed-point spaces are often explicitly known for positive linear operators arising in approximation theory. Consider, for instance, the Bernstein operator or Schoenberg's spline operator. The key idea is to construct a projection P with $\operatorname{ran}(P) = \ker(T - I)$ and $\ker(P) = \operatorname{ran}(T - I)$ to achieve the above mentioned space decomposition. In this case the range and the null space of P are invariant under T.

9.2.1 The inverse of the Gram matrix

For the exact representation of the projection operator P, let us consider the Gram matrix, where the invariant functionals operate on the fixed points of T,

$$G := \begin{pmatrix} e_1^*(e_1) & \cdots & e_1^*(e_n) \\ \vdots & & \vdots \\ e_n^*(e_1) & \cdots & e_n^*(e_n) \end{pmatrix} \in \mathbb{C}^{n \times n}.$$

It will be shown in the following that if this matrix is invertible then the projection P on the space $\ker(T-I)$ has the form

$$Px = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} e_j^*(x) e_i, \qquad x \in X,$$

where $A = (a_{ij}) = G^{-1}$ and the iterates converge uniformly towards P, i. e.,

$$\lim_{m\to\infty} \|T^m - P\|_{op} = 0.$$

9.2.2 Example: variation-diminishing transforms

We now give a short example of our result on $C([0,1])$, the space of continuous functions on the interval $[0,1]$ equipped with the norm of uniform convergence. Thereby, let n be a positive integer and suppose that $\{x_j\}_{j=1}^{n}$ form a partition of $[0,1]$ such that

$$0 = x_1 < x_2 < \cdots < x_n = 1.$$

We consider the variation-diminishing operator already introduced in Section 2.3.4, see (2.3) on page 21. Let $T : C([0,1]) \to C([0,1])$ be defined for $f \in C([0,1])$ by

$$Tf(x) = \sum_{k=1}^{n} f(x_k) e_k(x), \qquad x \in [0,1],$$

where $\{e_1, \dots, e_n\}$ are positive functions of $C([0,1])$ that form a partition of unity, i.e.,

$$\sum_{k=1}^{n} e_k(x) = 1 \qquad \text{for all } x \in [0,1].$$

Besides, we assume that $Tf = f$ whenever f is a linear function. These conditions are satisfied, e. g., for the Bernstein and the Schoenberg operator, discussed in Chapter 3. It is now easy to see that in this case $T1 = 1$ and $\|T\|_{op} = r(T) = 1$, where $r(T)$ is the spectral radius of T. Using the positivity of T and the ability to reproduce constants and linear functions it follows by Goodman [Goo96, Thm. 3.5] that $e_1(0) = e_n(1) = 1$. As consequence, T interpolates at 0 and 1, as

$$Tf(0) = \sum_{k=1}^{n} f(x_k) e_k(0) = f(x_1) = f(0),$$
$$Tf(1) = \sum_{k=1}^{n} f(x_k) e_k(1) = f(x_n) = f(1).$$

Here, we used the partition of unity property, to conclude that $e_k(0) = \delta_{k,1}$ and $e_k(1) = \delta_{k,n}$ for all $k \in \{1, \dots, n\}$. The introduced operator is a finite-rank operator with $T1 = 1$ and $Tx = x$ and two linearly independent invariant functionals for T^* are given due to the interpolation at 0 and 1. If δ_0, δ_1 denote the continuous functionals that evaluate continuous functions at 0 and 1 respectively, then $\delta_0(Tf) = \delta_0(f)$ and $\delta_1(Tf) = \delta_1(f)$ for all $f \in C([0,1])$.

Now, we want to answer the question whether the limit of the iterates T^m for $m \to \infty$ exists and if so to which operator the iterates converge. In Chapter 7 ([Nag15]) it has been shown that the partition of unity property of the basis $\{e_1, \ldots, e_n\}$ guarantees the spectral location $\sigma(T) \subset B(0,1) \cup \{1\}$. To derive the limiting operator, we have to specify the eigenspaces of T and its adjoint T^* that corresponde to the eigenvalue 1. Using the partition of unity and the ability to reproduce linear functions as well as the interpolation property at the endpoints of the interval $[0,1]$, we derive the following normalized basis for the fixed-point spaces

$$\ker(T - I) = \operatorname{span}(1, x), \qquad \ker(T^* - I) = \operatorname{span}(\delta_0, \delta_1),$$

as already discussed previously. Next, we consider the Gram matrix

$$G := \begin{pmatrix} \delta_0(1) & \delta_0(x) \\ \delta_1(1) & \delta_1(x) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

where the fixed points of T^* operate on the fixed points of T. Indeed, this matrix is invertible with

$$A := G^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Using the main result of this chapter we are able to use the coefficients $a_{11} = 1$, $a_{12} = 0$, $a_{11} = -1$, $a_{12} = 1$ to conclude that

$$\lim_{m \to \infty} \|T^m - P\|_{op} = 0,$$

where the projection $P : C([0,1]) \to C([0,1])$ is defined for $f \in C([0,1])$ by

$$\begin{aligned} Pf &= (a_{11}\delta_0(f) + a_{12}\delta_1(f)) \cdot 1 + (a_{21}\delta_1(f) + a_{22}\delta_0(f)) \cdot x \\ &= \delta_0(f) \cdot 1 + \delta_1(f) - \delta_0(f) \cdot x = f(0) + (f(1) - f(0))\, x. \end{aligned}$$

The iterates converge to the linear interpolation operator that interpolates at the endpoints of $[0,1]$. Consequently, we have demonstrated our underlying framework for general variation-diminishing operators that reproduce constant and linear functions, consider e. g., the Bernstein and the Schoenberg operator. However, the convergence is guaranteed for all quasi-compact contraction operators, as we will see in the following.

9.3 Invariants of Operators

The aim of this section is to show how to construct a projection P onto a generalized eigenspace of a bounded linear operator T defined on a complex Banach space X corresponding to an eigenvalue $\lambda \in \mathbb{C}$. To this end, we consider an operator T such that $T - \lambda I$ is a Browder operator with ascent $\operatorname{asc}(T - \lambda I) = p$. In this case, the projection has the property $\ker(P) = \operatorname{ran}(T - \lambda I)^p$ which gives us generically the following space decomposition:

$$\begin{aligned} X = \ker(T - \lambda I)^p &\quad \oplus \quad \operatorname{ran}(T - \lambda I)^p \\ = \quad \operatorname{ran}(P) &\quad \oplus \quad \ker(P). \end{aligned}$$

We provide a simple criterion under which assumptions this space decomposition is possible. Before we will look at a finite-dimensional generalized eigenspace of an operator $T \in \mathcal{L}(X)$, we will construct the projection on an arbitrary finite-dimensional subspace M of a vector space X. On M we introduce the classical coordinate map defined by a basis of M and the corresponding dual basis of the dual space M^*. By the extension theorems of Hahn-Banach the coordinate map gives us a continuous projection of X onto M. In the sequel, we will discuss conditions on the functionals that can be chosen in the coordinate map to build a dual basis. Finally, we apply the results to the generalized eigenspaces of a bounded linear operator T on a Banach space X and its adjoint T^* corresponding to an eigenvalue $\lambda \in \mathbb{C}$. A necessary condition on the operator $T - \lambda I$ is being Fredholm with non-positive index. If in addition $T - \lambda I$ is a Browder operator, i. e., the index is zero and its chain length is finite, then the projection yields the previously mentioned direct sum decomposition of X.

Note that this space decomposition is already well known, see Proposition 8.17 provided $T - \lambda I$ has positive finite chain length. In contrast to existing literature we prove it using an explicitly constructed finite-rank projection P. This method uses in fact the restriction that $T - \lambda I$ has to be Weyl operator, i. e., a Fredholm operator of index zero, to guarantee that the corresponding generalized eigenspaces of T and T^* have finite dimension. This direct construction of the projection P provides an alternative way to calculate the spectral projection corresponding to the eigenvalue λ as shown in Section 8.5.

9.3.1 Dual basis and the coordinate map

Let X be a normed vector space over the complex numbers and let $M \subset X$ be a closed subspace with $0 < \dim(M) < \infty$. In the sequel, we denote its dimension by $n = \dim(M)$. Moreover, let $\{e_1, \dots, e_n\}$ be a basis for M. Then every $x \in M$ has a unique representation

$$x = \sum_{i=1}^{n} e_i^*(x) e_i, \tag{9.2}$$

where $\{e_1^*, \dots, e_n^*\}$ are appropriate continuous linear functionals on M. By definition, each e_i can also be represented by (9.2) which yields the characterization

$$e_i^*(e_k) = \delta_{ik} = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k, \end{cases} \tag{9.3}$$

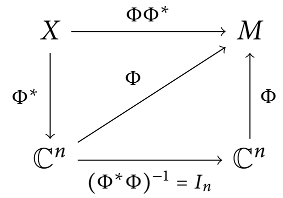


Figure 9.1: *A commutative diagram that illustrates the projection $\Phi\Phi^* : X \to M$ as composition of the synthesis operator Φ and analysis operator Φ^*. Note that the projection can also be written as $\Phi(\Phi^*\Phi)^{-1}\Phi^*$.*

for all $i, k \in \{1, \ldots, n\}$. In analogy to the construction of the frame operator on Hilbert spaces [Chr01; Chr03], we define a *synthesis operator* $\Phi : \mathbb{C}^n \to M$ by

$$\Phi(a_1, \ldots, a_n) = \sum_{i=1}^{n} a_i e_i. \tag{9.4}$$

The adjoint of this operator $\Phi^* : M \to \mathbb{C}^n$ yields the *analysis* operator

$$\Phi^*(x) = \begin{pmatrix} e_1^*(x) \\ \vdots \\ e_n^*(x) \end{pmatrix}, \qquad x \in X. \tag{9.5}$$

Combining both operators we can represent the coordinate map (9.2) by the composition $\Phi^*\Phi : M \to M$,

$$(\Phi\Phi^*)(x) = \sum_{i=1}^{n} e_i^*(x) e_i = x.$$

Note that according to (9.3) the matrix $\Phi^*\Phi \in \mathbb{C}^{n\times n}$ is the identity on $\mathbb{C}^n$:

$$\Phi^*\Phi = \begin{pmatrix} e_1^*(e_1) & \cdots & e_1^*(e_n) \\ \vdots & & \vdots \\ e_n^*(e_1) & \cdots & e_n^*(e_n) \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix} = I_n.$$

Accordingly, the basis $\{e_1^*, \ldots, e_n^*\} \subset M^*$ is said to be the *dual basis* for $\{e_1, \ldots, e_n\} \subset M$. Applying the Theorem of Hahn-Banach, the coordinate map can be extended to the whole vector space X. To simplify notation the extended functionals on X^* will be also denoted by $e_1^*, \ldots, e_n^*$ in the following.

Lemma 9.9. *The operator $\Phi\Phi^* : M \to M$ can be extended to a projection of the space X onto the closed set M and is bounded by*

$$\|(\Phi\Phi^*)(x)\| \leq \|x\| \sum_{i=1}^{n} \|e_i^*\|.$$

The matrix $(\Phi^*\Phi)_{ij} \in \mathbb{C}^{n\times n}$ *is invertible and the coordinate map* $\Phi\Phi^*\big|_M : M \to M$ *which is restricted on* M *yields an isomorphism. The space* X *can be decomposed into*

$$X = M \oplus \ker(\Phi\Phi^*).$$

Proof. The continuous functionals e_i^* can be extended by the classical Hahn-Banach Theorem to X^* with the same properties as on M. We denote the resulting extensions again as $e_i^* \in X^*$. Therefore, $\Phi\Phi^* : X \to M$ and

$$(\Phi\Phi^*)(x) = \sum_{i=1}^{n} e_i^*(x)e_i = x \qquad \text{for all } x \in M.$$

Moreover, the operator is bounded on X since for $x \in X$ we have

$$\|(\Phi\Phi^*)(x)\| = \left\|\sum_{i=1}^{n} e_i^*(x)e_i\right\| \le \sum_{i=1}^{n} \|e_i^*(x)\|\,\|e_i\| \le \|x\| \sum_{i=1}^{n} \|e_i^*\|,$$

where we used that $\|e_i\| = 1$ and the fact that $\|e_i^*(x)\| \le \|e_i^*\|\,\|x\|$. Clearly, $(\Phi^*\Phi)$ is invertible with $(\Phi^*\Phi)^{-1} = I_n$. It yields also a projection, because for every $x \in M$ we obtain $(\Phi\Phi^*)(x) = x$ and therefore, $(\Phi\Phi^*)^2 = (\Phi\Phi^*)$. As the operator $\Phi\Phi^* \in \mathcal{L}(X)$ is a bounded projection onto the closed space M, we obtain canonically the space decomposition $X = M \oplus \ker(\Phi\Phi^*)$, see Proposition 8.4. □

The key property to notice here is that $(\Phi^*\Phi)_{ij}$ is an invertible matrix and that $\Phi\Phi^*$ is a projection onto M. The commutative diagram shown in Figure 9.1 illustrates the behaviour of Φ and Φ^*.

In the following, we show which functionals $\{e_1^*, \ldots, e_n^*\}$ can be chosen instead of the dual basis such that $\Phi\Phi^*$ is still a projection where the analysis operator Φ^* now contains the new functionals. The next section shows that the matrix $\Phi^*\Phi$ must have full column rank.

9.3.2 Complemented subspaces and projections

We consider now the following problem. Given a set of linear functionals $\Lambda \subset X^*$, we ask whether it is possible to construct a projection onto the closed finite dimensional subspace $M \subset X$ with functionals chosen only from the set Λ. We give a characterization in the next theorem. As in the previous section, we consider a finite-dimensional subspace M of X. Additionally, let $\Lambda \subset X^*$ be a finite-dimensional subspace of X^*. Let us denote by $\{e_1, \ldots, e_n\}$ and $\{e_1^*, \ldots, e_m^*\}$ a basis of M and Λ, respectively. The *synthesis operator* $\Phi : \mathbb{C}^n \to X$ is constructed as in (9.4), whereas the *analysis operator* $\Phi^* : X \to \mathbb{C}^n$ is not defined as the adjoint of Φ but uses the basis functionals of Λ:

$$\Phi^*(x) = \begin{pmatrix} e_1^*(x) \\ \vdots \\ e_m^*(x) \end{pmatrix}, \qquad x \in X.$$

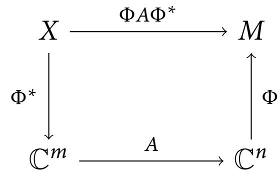


Figure 9.2: *Commutative diagram showing the projection $\Phi A \Phi^* : X \to M$. Here the matrix A is either the left inverse of the matrix $\Phi^*\Phi$ or its inverse.*

Let us assume that $\dim(\Lambda) \geq \dim(M)$ holds. Then we will show in the next theorem that again $\Phi\Phi^*$ yields a projection operator onto M provided that $\Phi^*\Phi$ has full column rank.

Theorem 9.10. *Let $\Lambda \subset X^*$ with $0 < \dim(M) \leq \dim(\Lambda) < \infty$ and let $n = \dim(M)$, $m = \dim(\Lambda)$. Then the operator $P \in \mathcal{L}(X)$ defined for $A = (a_{ij}) \in \mathbb{C}^{n\times m}$ by*

$$Px = \Phi A \Phi^*(x) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} e_j^*(x) e_i, \qquad x \in X, \tag{9.6}$$

yields a projection onto M if and only if the matrix

$$G := (\Phi^*\Phi) = \begin{pmatrix} e_1^*(e_1) & \cdots & e_1^*(e_n) \\ \vdots & & \vdots \\ e_m^*(e_1) & \cdots & e_m^*(e_n) \end{pmatrix} \in \mathbb{C}^{m\times n} \tag{9.7}$$

has full column rank n. In this case, the matrix A is determined by the left inverse of G,

$$A = (G^T G)^{-1} G^T.$$

Proof. Let us first assume that the matrix $G \in \mathbb{C}^{m\times n}$ has a left inverse

$$G_{left}^{-1} = (G^T G)^{-1} G^T \in \mathbb{C}^{n\times m},$$

and let $A = G_{left}^{-1}$. According to the definition of a left inverse matrix, the equation

$$A \cdot G = G_{left}^{-1} \cdot G = I_n \tag{9.8}$$

holds. Now, we prove that P, defined for $x \in X$ as in (9.6), is a projection onto M. To this end, we will show that $P(x) = x$ holds for all $x \in M$ by considering the basis of M. Thus, we only have to prove $P(e_k) = e_k$ for all $k \in \{1, \ldots, n\}$. The direct calculation of $P(e_k)$ yields

$$P(e_k) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} e_j^*(e_k) e_i = \sum_{i=1}^{n}\left[\sum_{j=1}^{m} a_{ij} e_j^*(e_k)\right] e_i.$$

Applying (9.8) yields $\sum_{j=1}^{m} a_{ij} e_j^*(e_k) = \delta_{ki}$. Therefore,

$$P(e_k) = \sum_{i=1}^{n} p_i \delta_{ki} = e_k$$

holds and we obtain $P(X) = M$. Furthermore, $P^2 = P$ on X as $\{e_1, \dots, e_n\}$ forms a basis for M. Finally, we show the reverse direction. To this end, let us assume that P is a projection onto M, i. e., $P(X) = M$ and $P^2 = P$ holds. Then $P(e_k) = e_k$ must hold for any $k \in \{1, \dots, n\}$, as $e_i \in M$. We calculate

$$P(e_k) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} e_j^*(e_k) e_i = \sum_{i=1}^{n} \left[\sum_{j=1} a_{ij} e_j^*(e_k) \right] e_i .$$

This yields necessary the requirement $\sum_{j=1}^{m} a_{ij} e_j^*(e_k) = \delta_{ik}$ for all $k \in \{1, \dots, n\}$. Therefore, we derive the matrix equation $A \cdot G = I_n$ with the unknown coefficient matrix $A = (a_{ij}) \in \mathbb{C}^{n \times m}$. In fact, this equation has a solution if and only if the matrix G has a left inverse G_{left}^{-1}, which concludes the proof. □

Next, we will provide an upper bound of the projection operator $P = \Phi A \Phi^*$ by the 1-norm of the matrix A.

Lemma 9.11. *Under the assumption of Theorem 9.10, the projection operator P defined by* (9.6) *has finite-rank and is bounded by*

$$\|Px\| \le \|x\| \sum_{i=1}^{n} \sum_{j=1}^{m} |a_{ij}|, \qquad x \in X.$$

Proof. Clearly, P is a finite rank operator. Let $x \in X$. For arbitrary $i \in \{1, \dots, n\}$ we obtain

$$\left\| \sum_{j=1}^{m} a_{ij} e_j^*(x) \right\| \le \|x\| \sum_{j=1}^{m} |a_{ij}|,$$

because the dual basis is normalized, i. e., $\left\| e_j^* \right\| = 1$. Using the same argument for the basis of M we get

$$\|Px\| = \|\Phi A \Phi^* x\| = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} e_j^*(x) e_i \le \|x\| \sum_{i=1}^{n} \sum_{j=1}^{m} |a_{ij}| .$$

□

9.3.3 Invariant subspaces and projections

In the following we will consider a linear operator T defined on a complex Banach space X. As in the preceding sections we are interested in the construction of a projection onto a finite-dimensional subspace of X. Here, we choose M as a generalized eigenspace of T corresponding to an eigenvalue $\lambda \in \sigma_p(T)$. It will be shown that the set of functionals is exactly given by the corresponding generalized eigenspace of the adjoint T^*.

Accordingly, given some integer $p > 0$, we consider now the following two subspaces

$$M_\lambda^p = \ker(T - \lambda I)^p = \left\{x \in X \,:\, (T - \lambda I)^p x = 0\right\} \subset X, \tag{9.9}$$

$$\Lambda_\lambda^p = \ker(T^* - \lambda I)^p = \left\{x^* \in X^* \,:\, (T^* - \lambda I)^p x^* = 0\right\} \subset X^*. \tag{9.10}$$

Note that due to the fact that $\ker(T^* - \lambda I)^p = \operatorname{ran}((T - \lambda I)^p)^\perp$ holds by Proposition 8.3 the set Λ_λ^p can also be determined as

$$\Lambda_\lambda^p = \left\{x^* \in X^* \,:\, x^*\left((T - \lambda I)^p x\right) = 0 \text{ for all } x \in X\right\}. \tag{9.11}$$

To assure that both spaces (9.9) and (9.10) are finite-dimensional and that the dimension of the functionals Λ_λ^p is greater than the dimension of M_λ^p, we assume in the following that $(T - \lambda I)^p$ is a Fredholm operator with negative index, i. e., $\operatorname{ind}(T - \lambda I)^p \le 0$. Then we have by definition

$$n = \dim(M_\lambda^p) \le \dim(\Lambda_\lambda^p) = m$$

and we can consider w.l.o.g. normalized bases of M_λ^p and Λ_λ^p:

$$M_\lambda^p = \operatorname{span}\{e_1, \dots, e_n\} \text{ and } \Lambda_\lambda^p = \operatorname{span}\{e_1^*, \dots, e_m^*\} \tag{9.12}$$

such that $\|e_i\|_X = 1$ and $\|e_i^*\|_{X^*} = 1$. If we additionally suppose we have the following finite chain of inclusions

$$\ker(T - \lambda I) \subsetneq \ker(T - \lambda I)^2 \subsetneq \cdots \subsetneq \ker(T - \lambda I)^p = \ker(T - \lambda I)^{p+1} = \cdots,$$

then the ascent of $T - \lambda I$ is specified as $p := \operatorname{asc}(T - \lambda I) < \infty$. Corresponding to the eigenvalue $\lambda \in \sigma_p(T)$, the set M_λ contains all of the generalized eigenvectors of the operator T and the set Λ_λ^p contains all the dual generalized eigenvectors. More precisely, the set Λ_λ^p contains all the generalized eigenvectors of the adjoint operator T^* to the eigenvalue λ.

Remark. Note that the assumption on T are not very restrictive. As shown in the end of the last chapter, every compact operator satisfies all of the conditions. Moreover, quasi-compact operators satisfy these condition in the case where $\lambda = 1$ is chosen. Especially, every operator where $T - \lambda I$ is a Browder operator fulfills these conditions, see Definition 8.10 and the comments below this definition on page 99. Consider also Proposition 8.7 and Proposition 8.12.

We will show next how to construct a projection P onto $\ker(T - \lambda I)^p$ to obtain the space decomposition

$$\ker(T - \lambda I)^p \oplus \ker(P)$$

such that $\ker(P) = \operatorname{ran}(T-\lambda I)^p$ holds. Note that in this case $\operatorname{ran}(T-\lambda I)^p$ is closed as it is the null space of the projection P.

First, we provide an equivalent characterization of the restrictions on T to have finite chain length of the generalized eigenspaces of T provided that $T-\lambda I$ is a Fredholm operator with $\operatorname{ind}(T-\lambda I) \leq 0$ to assure that the generalized eigenspaces of T and T^* are finite-dimensional. The next lemma shows that the ascent can be characterized by the column rank of the Gramian matrix constructed using the matrix (9.7). In the following, we will denote by $\Phi_-(X)$ all Fredholm operators defined on the Banach space X that have an index less or equal to zero.

Lemma 9.12. *Let $T \in \mathcal{L}(X)$ and $\lambda \in \sigma_p(T)$ such that $T-\lambda I \in \Phi_-(X)$. Then $T-\lambda I$ has finite ascent p, i. e., $p = \operatorname{asc}(T-\lambda I) = p < \infty$, if and only if the Gramian matrix*

$$G := (\Phi^*\Phi) = \begin{pmatrix} e_1^*(e_1) & \cdots & e_1^*(e_n) \\ \vdots & & \vdots \\ e_m^*(e_1) & \cdots & e_m^*(e_n) \end{pmatrix} \in \mathbb{C}^{m\times n}$$

has full column rank.

Proof. Suppose that $T-\lambda I$ is Fredholm operator with non-positive index. Then $(T-\lambda I)^p$ is also Fredholm with non-positive index $p\cdot\operatorname{ind}(T-\lambda I)$. This follows by the index theorem [Heu82, Thm. 23.1], as

$$\operatorname{ind}(\underbrace{(T-\lambda I)\cdots(T-\lambda I)}_{p-\text{times}}) = \sum_{i=1}^{p}\operatorname{ind}(T-\lambda I) = p\cdot\operatorname{ind}(T-\lambda I).$$

Therefore, $\operatorname{ran}(T-\lambda I)^p$ is closed [Heu82, Prop. 24.3] and

$$n = \alpha((T-\lambda I)^p) = \dim(\ker(T-\lambda I)^p) \leq \dim(\ker(T^*-\lambda I)^p) = \beta((T-\lambda I)^p) = m.$$

Note that $(\Lambda_\lambda^p)_\perp = (\operatorname{ran}((T-\lambda I)^p)^\perp)_\perp = \overline{\operatorname{ran}(T-\lambda I)^p} = \operatorname{ran}(T-\lambda I)^p$.

Let us now assume that $T-\lambda I$ has ascent p. In order to show that the columns of $G = \Phi^*\Phi$ are linearly independent, we choose $c = (c_1,\dots,c_n)^T \in \mathbb{C}^n$ such that

$$\sum_{i=1}^{n} c_i e_j^*(e_i) = 0$$

for all $j \in \{1,\dots,m\}$. Then we derive that $e_j^*(\sum_{i=0}^n c_i e_i) = 0$ for all $j \in \{1,\dots,m\}$. Therefore,

$$\sum_{i=0}^{n} c_i e_i \in \bigcap_{j=1}^{m}\ker(e_j^*) = (\Lambda_\lambda^p)_\perp = \operatorname{ran}(T-\lambda I)^p.$$

As $T-\lambda I$ has finite ascent p we can conclude with Proposition 8.9 that $\ker(T-\lambda I)^p \cap \operatorname{ran}(T-\lambda I)^p = \{0\}$ holds. As by definition also $\sum_{i=1}^n c_i e_i \in \ker(T-\lambda I)^p$ holds we derive that $\sum_{i=1}^n c_i e_i = 0$. From the linear independence of $\{e_1,\dots,e_n\}$ it follows that $c_1 = \cdots = c_n = 0$. Therefore, the matrix $\Phi^*\Phi$ has full rank, as the columns are linearly independent.

To show that the converse is also true let us suppose that the matrix G has full column rank. Hence, if $\sum_{i=1}^n c_i e_j^*(e_i) = 0$ holds it follows that every coefficient $c_i = 0$ for all $i \in \{1, \dots, n\}$. Suppose now that $x \in \ker(T - \lambda I)^p \cap \operatorname{ran}(T - \lambda I)^p$. Then x can be written as linear combination $x = \sum_{i=1}^n c_i e_i$ for some coefficients $c_i \in \mathbb{C}$. As $\operatorname{ran}(T - \lambda I)^p = (\Lambda_\lambda^p)_\perp$, we obtain for all $j \in \{1, \dots, m\}$ that

$$0 = e_j^*\left(\sum_{i=1}^n c_i e_i\right) = \sum_{i=1}^n c_i e_j^*(e_i).$$

We conclude that $c_i = 0$ for all $i \in \{1, \dots, n\}$ as the matrix G has full column rank. Finally, we have $x = 0$. Therefore, $\ker(T - \lambda I)^p \cap \operatorname{ran}(T - \lambda I)^p = \{0\}$. By the first item of Proposition 8.9 this is equivalent to the statement that the ascent of $T - \lambda I$ is p and the proof is complete. □

As the Gramian matrix has full column rank, we can construct a projection operator onto $\ker(T - \lambda I)$ according to Theorem 9.10. Consequently, we consider as in the last section the finite-rank operator $P \in \mathcal{K}(X)$ defined for $x \in X$ by

$$Px = (\Phi A \Phi^*)(x) = \sum_{i=1}^n \sum_{j=1}^m a_{ij} e_j^*(x) e_i, \tag{9.13}$$

where $e_i \in M_\lambda$, $e_j^* \in \Lambda_\lambda$ are the normalized bases and $A = (a_{ij}) \in \mathbb{C}^{n\times m}$. This time, the functionals e_j^* are explicitly chosen as basis of $\ker(T^* - \lambda I)^p$ where the coefficients a_{ij} serve as parameter. In this setting, Theorem 9.10, yields a projection operator that projects onto the generalized eigenspace M_λ^p and gives a space decomposition of X into $X = M_\lambda^p \oplus \ker(P)$.

Corollary 9.13. *Let $T \in \mathcal{L}(X)$ and $\lambda \in \sigma_p(T)$ such that $T - \lambda I \in \Phi_-(X)$ with ascent $p \in \mathbb{N}$. Then the linear operator $P \in \mathcal{K}(X)$ defined for $x \in X$ as*

$$Px = \Phi A \Phi^*(x),$$

where A is the left inverse of $(\Phi^*\Phi)$, yields a continuous projection onto $M_\lambda^p \subset X$, where $\operatorname{ran}(P) = M_\lambda^p = \ker(T - \lambda I)^p$ *is a closed subspace.*

Proof. This is a direct consequence of Lemma 9.9 and Lemma 9.12. □

Note that in the current setting, we obtain a projection P where $\operatorname{ran}(P) = \ker(T - \lambda I)^p$ is a T-invariant subspace. Accordingly, we have the space decomposition

$$X = \operatorname{ran}(P) \oplus \ker(P) = \ker(T - \lambda I)^p \oplus \ker(P).$$

In the following we are interested when also $\ker(P)$ is invariant with respect to the operator T. Then we can decompose the operator T into

$$T = \begin{pmatrix} J & 0 \\ 0 & S \end{pmatrix} \in \mathcal{L}(\ker(T - \lambda I)^p \oplus \ker(P)),$$

where J is the Jordan normal form of T on the generalized eigenspace $\ker(T - \lambda I)^p$ and $S \in \mathcal{L}(\ker(P))$ is equal to the operator T restricted to $\ker(P)$.

Remark. Even though we write the operator decomposition in matrix notation, we don't assume the Banach space X to be separable. The matrix form is only used to demonstrate the space decomposition easier where $\ker(T - \lambda I)$ *is always finite-dimensional. In this case, J is given according to some basis, whereas J is not necessarily defined by a countable dense set in X.*

Furthermore, we are not only interested when $\ker(P)$ is invariant with respect to T, we also want to know under which conditions on T the relation $\ker(P) = \operatorname{ran}(T - \lambda I)^p$ holds. It turns out that this is the exactly the case when the $T - \lambda I$ is a Browder operator, i. e., the operator $T - \lambda I$ has Fredholm index 0 and finite chain length p. We will discuss this particular case in the following. First, we show in the next lemma that the Fredholm index 0 of $T - \lambda I$ leads to the invertibility of the Gramian matrix $\Phi^*\Phi$. Finally, we will prove that in this case $\ker(P) = \operatorname{ran}(T - \lambda I)^p$ holds. We will conclude this section with an overview over related results.

Lemma 9.14. *Let $T \in \mathcal{L}(X)$ and $\lambda \in \mathbb{C}$ such that $T - \lambda I \in \Phi_-(X)$. Then $T - \lambda I$ is a Browder operator if and only if the matrix $G = \Phi^*\Phi$ is invertible.*

Proof. If G is invertible, then $T - \lambda I$ has finite ascent p by Lemma 9.12 and G is necessarily a square matrix, thus $\operatorname{ind}(T - \lambda I)^p = 0$ as

$$n = \alpha((T - \lambda I)^p) = \beta((T - \lambda I)^p) = m, \tag{9.14}$$

using the definition of the nullity $\alpha((T - \lambda I)^p) = \dim \ker(T - \lambda I)^p$ and the deficiency $\beta((T - \lambda I)^p) = \dim \ker(T^* - \lambda I)^p$. By the fourth item of Proposition 8.12 on page 99 we can conclude by $\alpha(T - \lambda I) = \beta(T - \lambda I)$ and $\operatorname{asc}(T - \lambda I) = p < \infty$ that also the descent of $T - \lambda I$ is finite. Therefore, $T - \lambda I \in \mathcal{W}_B(X)$, i. e., $T - \lambda I$ is a Browder operator with ascent p.

Assume to the contrary that $T - \lambda I$ is a Browder operator. Then $T - \lambda I$ has finite ascent p and $\operatorname{ind}(T - \lambda I) = 0$ by definition. As $\operatorname{ind}(T - \lambda I) = 0$ the matrix $G = \Phi^*\Phi$ is a $n \times n$-matrix as $n = \alpha((T - \lambda I)^p) = \beta((T - \lambda I)^p)$ using the same argument as in (9.14). As we have the conditions $\operatorname{ind}(T - \lambda I) = 0$ and $\operatorname{asc}(T - \lambda I) = p$ we can apply Lemma 9.12 to conclude that the matrix G has full rank and thus, is invertible as square matrix. □

Next, we will prove that the null space of the projection P is given by $\ker(P) = \operatorname{ran}(T - \lambda I)^p$ provided that $T - \lambda I$ is a Fredholm operators with index 0 having finite chain length p, i. e., $T - \lambda I$ is a Browder operator. Note that the invertibility of $\Phi^*\Phi$ is already sufficient for this result.

Theorem 9.15 (Space decomposition)**.** *Let $T \in \mathcal{L}(X)$ and $\lambda \in \sigma_p(T)$ such that $T - \lambda I$ is a Browder operator with ascent p. Then*

$$X = \ker(T - \lambda I)^p \oplus \operatorname{ran}(T - \lambda I)^p,$$

where $\operatorname{ran}(\Phi(\Phi^*\Phi)^{-1}\Phi^*) = \ker(T - \lambda I)^p$ *and* $\ker(\Phi(\Phi^*\Phi)^{-1}\Phi^*) = \operatorname{ran}(T - \lambda I)^p$.

Proof. Let $n = \dim(\ker(T - \lambda I)^p) = \dim(\ker(T^* - \lambda I)^p) < \infty$. As $(T - \lambda I)^p$ is a Fredholm operator, $\operatorname{ran}(T - \lambda I)^p$ is closed. We already have shown that $\operatorname{ran}(P) = \ker(T - \lambda I)^p$. In order to show $\ker(P) = \operatorname{ran}(T - \lambda I)^p$ let $x \in \ker(P)$. Then we have

$$0 = Px = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} e_j^*(x) e_i. \tag{9.15}$$

As $\{e_1, \ldots, e_n\}$ form a basis for $\ker(T - \lambda I)^p$ by (9.9) and (9.12), the relation (9.15) can only hold if

$$\sum_{j=1}^{n} a_{ij} e_j^*(x) = 0$$

for every $i \in \{1, \ldots, n\}$. Using that $A = (\Phi^*\Phi)^{-1}$ is invertible by Lemma 9.12, we obtain that $e_j^*(x) = 0$ for all $j \in \{1, \ldots, m\}$. Then it is easy to see that

$$x \in (\Lambda_\lambda^p)_\perp = (\operatorname{ran}((T - \lambda I)^p)^\perp)_\perp = \operatorname{ran}(T - \lambda I)^p,$$

because $\operatorname{ran}(T - \lambda I)^p$ is closed.

Now let $y \in \operatorname{ran}(T - \lambda I)^p$. Accordingly, there is $x \in X$ with $(T - \lambda I)^p x = y$. In this case also $y \in \ker(P)$ holds, because

$$Py = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} e_j^*((T - \lambda I)^p x) e_i = 0.$$

In the last step we used that $e_j^* \in \operatorname{ran}((T - \lambda I)^p)^\perp$. Finally, we obtain the space decomposition

$$X = \operatorname{ran}(P) \oplus \ker(P) = \ker(T - \lambda I)^p \oplus \operatorname{ran}(T - \lambda I)^p,$$

where $\operatorname{ran}(P) = \ker(T - \lambda I)^p$ and $\ker(P) = \operatorname{ran}(T - \lambda I)^p$. □

We conclude this section with a theorem that gathers all the results we have shown for a bounded operator T with eigenvalue $\lambda \in \sigma_p(T)$, where $T - \lambda I$ is a Weyl operator, i. e., a Fredholm operator with zero index. Note once more that this restriction is important for our setting where the generalized eigenspaces have to be finite-dimensional.

Theorem 9.16 (Characterization of the Browder operator $T - \lambda I$).
Let $T \in \mathcal{L}(X)$ and $\lambda \in \sigma_p(T)$ such that $T - \lambda I \in \mathcal{W}(X)$. Then the following statements are equivalent:

1. *$T - \lambda I$ is a Browder operator, $T - \lambda I \in \mathcal{W}_B(X)$,*
2. *the operator $T - \lambda I$ has finite chain length, i. e., $\operatorname{asc}(T - \lambda I) = \operatorname{dsc}(T - \lambda I) < \infty$,*
3. *the space X can be decomposed into $X = \ker(T - \lambda I)^p \oplus \operatorname{ran}(T - \lambda I)^p$,*
4. *the $n \times n$ matrix $G := (\Phi^* \Phi)$,*

$$G = \begin{pmatrix} e_1^*(e_1) & \cdots & e_1^*(e_n) \\ \vdots & & \vdots \\ e_n^*(e_1) & \cdots & e_n^*(e_n) \end{pmatrix} \in \mathbb{C}^{n \times n}$$

   is invertible, where $n = \dim\ker(T - \lambda I)^p = \dim\ker(T^* - \lambda I)^p$,

5. *the operator $P : X \to \ker(T - \lambda I)^p$ defined by*

$$Px = \Phi A \Phi^*(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} e_j^*(x) e_i, \qquad x \in X,$$

   yields a projection onto $\ker(T - \lambda I)^p$, where $A = (a_{ij}) := G^{-1}$.

My results are in particular the invertibility of the Gram matrix and the construction of the projection operator in the last item. I have shown how to construct the projection operator P by the inverse of the Gramian matrix for the space decomposition in the third point. In the next section, we will use this projection operator to derive the limiting operator of iterates.

9.4 Iterates of quasi-compact contractions

Using the preceding results, we consider in the following the limit of the iterates of an operator $T \in \mathcal{L}(X)$. Recall that T is said to be a contraction if $\|T\|_{op} \le 1$. As already mentioned previously, of our interest is the case where $\|T\|_{op} = 1$. If $\|T\|_{op} < 1$, then indeed $T^n \to 0$ uniformly for $n \to \infty$ and the limiting operator is zero. Assuming that $\|T\|_{op} = 0$, we are interested in the conditions on the operator under which the iterates converge uniformly provided that the operator T has a non-trivial fixed point space. This is of importance, as the iterates converge to a projection onto the fixed point space according to Proposition 9.8. To apply the results of the last section, we also have to assume that the fixed point space is finite-dimensional.

In order to satisfy the necessary condition of Proposition 9.8, we will consider operators where the spectrum is contained in the closed unit disk $\overline{B(0,1)}$ with spectral radius 1. There are two cases to discuss separately. The first case is where $\sigma(T) \subset B(0,1) \cup \{1\}$ and 1 is an eigenvalue of T. In the other case, we discuss operators where the peripheral spectrum $\sigma_{per} := \sigma(T) \setminus B(0,1)$ contains more eigenvalues than 1. Here, we can prove a convergence result only when σ_{per} is periodic, i. e., when there exists $\mu \in \sigma_{\text{per}}(T)$

such that $\sigma_{\text{per}}(T) = \{\mu^k : k \in \mathbb{N}\}$. This, for instance, is the case when T is a positive contraction.

We will first prove some results for quasi-compact operators. It will be shown that every quasi-compact operator T with $\|T\|_{op} = r(T) = 1$ has the property that $T - \lambda I$ is a Browder operator of ascent one if λ is a peripheral eigenvalue. Besides, there exists at least one peripheral eigenvalue. If one further assumes that T is a positive quasi-compact operator that satisfies $\|T\|_{op} = r(T) = 1$ then 1 is always an eigenvalue of T and the peripheral spectrum is cyclic. Due tho these results, we will consider operators of ascent one in Section 9.4.2 and apply the results of the preceding section. It will be shown, that this is equivalent to the property that the projection operator P constructed by the inverse of the Gramian matrix commutes with T, i. e., $TP = PT = \lambda P$ holds. Finally, we consider the case where $\lambda = 1$. In this case, it will be shown that the iterates T^n converge uniformly to P for $n \to \infty$. Note that if the iterates converge to a finite-rank operator, then this operator is quasi-compact by definition. We show further results for quasi-compact operators where the peripheral spectrum is cyclic. We conclude this chapter to show a relation to the result of Sine [Sin70] as stated in Proposition 9.4. Concretely, we will show that the fixed points of T separate the fixed points of T^* if and only if the corresponding Gramian matrix is invertible. Sine has shown that this is equivalent to the convergence of the Cesáro means $n^{-1}\sum_{k=0}^{n-1} T^k$.

9.4.1 Quasi-compact operators and the peripheral spectrum

Now, suppose $T \in \mathcal{L}(X)$ is a quasi-compact operator, i. e., the essential spectral radius is less than one. Consequently, it follows that every spectral value $\lambda \in \sigma(T)$ with modulus larger than the essential spectral radius is an isolated eigenvalue and the operator $T - \lambda I$ is a Browder operator. Therefore, there always exists an eigenvalue $\lambda \in \sigma(T)$ with modulus equal to the spectral radius $r(T)$. Moreover, there are only finitely many eigenvalues on the peripheral spectrum. The next lemma gives a characterization.

Lemma 9.17. *Let $T \in \mathcal{L}(X)$ be a quasi-compact operator with $r(T) \geq 1$. Then, there is at least one eigenvalue λ with $|\lambda| = r(T)$. Besides, every spectral value $\lambda \in \sigma(T)$ with $|\lambda| > r_{ess}(T)$ is an isolated eigenvalue of T and $T - \lambda I$ is a Browder operator. There are only finitely many eigenvalues on the peripheral spectrum of T.*

Proof. By the definition of quasi-compactness, we have $r_{ess}(T) < 1$ and all of the spectral values outside with modulus larger than $r_{ess}(T)$ are isolated. As already discussed above, $T - \lambda I$ is a Browder operator. If $\lambda \notin \sigma_b(T)$, then by [Lay68, Thm. 1], λ is a pole of the resolvent of finite order. Applying Proposition 8.17, we derive that $\operatorname{asc}(T - \lambda I)$ is positive and finite and hence, λ is an isolated eigenvalue of T. As all the cluster points of the spectrum are on the boundary of $\sigma_{ess}(T)$, there is at least one eigenvalue $\lambda \in \sigma(T)$ with $|\lambda| = r(T)$. Finally, there are only finitely many on the peripheral spectrum as otherwise there would be an accumulation point outside of the essential Browder spectrum. □

We show now that for the eigenvalues λ of quasi-compact operators $T \in \mathcal{L}(X)$ that are on the peripheral spectrum, i. e., eigenvalues λ with modulus $r(T)$, the associated Browder operator $T - \lambda I$ has ascent one. A similar result for transition probabilities has already been shown by Hennion and Hervé [HH01, Proposition V.1] and is stated here in a more general setting. We consider operators that are not restricted to $r(T) = 1$ and X can be an arbitrary complex Banach space instead a Banach function space where point evaluations are continuous linear functionals.

Lemma 9.18. *Let $T \in \mathcal{L}(X)$ be a quasi-compact operator with $r(T) \geq 1$ such that*

$$\sup_{n \in \mathbb{N}} r(T)^{-n} \|T^n\|_{op} < \infty.$$

Then for every peripheral eigenvalue $\lambda \in \sigma_{\text{per}}(T)$ the associated Browder operator $T - \lambda I$ has ascent one.

Proof. Note that the existence of a peripheral eigenvalue has been shown in the lemma above. We consider now an eigenvalue $\lambda \in \sigma_p(T)$ with $|\lambda| = r(T)$. Suppose that $x \in \ker(T - \lambda I)^2$. Then we can represent $T^n x$ for all positive integers n by

$$T^n x = (\lambda I + (T - \lambda I))^n x = \sum_{k=0}^{n} \binom{n}{k} \lambda^{n-k} (T - \lambda I)^k x = \lambda^n x - n\lambda^{n-1}(T - \lambda I)x.$$

We will show now that $(T - \lambda I)x = 0$ holds. To this end, we calculate using that $|\lambda| = r(T)$ and the existence of $B > 0$ such that $r(T)^{-m} \|T^m\|_{op} < B$ for all positive integers m:

$$\begin{aligned} \left\| n\lambda^{n-1}(T - \lambda I)x \right\| &= \|\lambda^n x - T^n x\| \leq |\lambda^n| \|x\| + \|T^n x\| \\ &\leq r(T)^n \|x\| + \|T^n\|_{op} \|x\| \\ &\leq r(T)^n (1 + B) \|x\| . \end{aligned}$$

It is now easy to see that $\|(T - \lambda I)x\| \leq \frac{r(T)(1+B)}{n} \|x\|$ and we finally conclude that $x \in \ker(T - \lambda I)$ as n was arbitrary. Thus, we have shown that $\ker(T - \lambda I)^2 = \ker(T - \lambda I)$, i. e., $T - \lambda I$ has ascent one as $\ker(T - \lambda I) \neq 0$. ☐

The following corollary considers the special case when $r(T) = \|T\|_{op}$ holds. Operators of this kind are said to be *normaloid* and have been discussed in Heuser [Heu82, Chap. 54]. Note that a similar proof of the previously lemma for normaloid operators can also be found in [Heu82, Prop. 54.2 & 54.3]. If T is a quasi-compact normaloid operator, we obtain the following corollary:

Corollary 9.19. *Let $T \in \mathcal{L}(X)$ be a quasi-compact operator with $r(T) = \|T\|_{op}$. There exists at least one eigenvalue with modulus $r(T)$. Furthermore, for every peripheral eigenvalue $T - \lambda I$ is a Browder operator with ascent one.*

Proof. For all positive integers n the inequality $r(T) \le \|T^n\|_{op}^{1/n} \le \|T\|_{op}$ holds. Therefore, $r(T)^{-n}\|T^n\|_{op} \le 1$ for all n. The result follows by Lemma 9.17 and Lemma 9.18. □

If X is a Banach lattice and T is a positive linear operator even stronger results can be made. Roughly speaking, a Banach lattice is a partially ordered Banach space, where the norm is compatible with the ordering. For more details, we refer to Schaefer [Sch74]. Note that Banach lattices are important to define positivity in general spaces due to the equipped ordering. According to H. P. Lotz [Lot68], the authors Kreĭn and Rutman [KR48] have first shown in 1948 that every positive compact operator on a Banach lattice with $r(T) = \|T\|_{op}$ has a cyclic peripheral spectrum. This result has been generalized in Lotz [Lot68, Theorem 4.10], where the peripheral spectrum of a positive operator $T \in \mathcal{L}(X)$ on a Banach lattice X is cyclic if the spectral radius $r(T)$ is a pole of the resolvent. Furthermore, H. P. Lotz concluded in [Lot68] that the peripheral spectrum of every positive compact operator is cyclic.

The next corollary sums up these results for positive quasi-compact operators.

Corollary 9.20. *Let X be a Banach lattice and let $T \in \mathcal{L}(X)$ be a positive quasi-compact operator with $r(T) = \|T\|_{op} = 1$. Then $1 \in \sigma(T)$, i. e., $T - I$ is a Browder operator of ascent one. Furthermore, the peripheral spectrum is cyclic consisting only of roots of unity.*

Proof. It has been shown by H. P. Lotz [Lot68] that $r(T) \in \sigma(T)$ if T is positive. In the case where T is a quasi-compact positive operator with $r(T) = 1$, T has the real eigenvalue one and $T - I$ is a Browder operator with ascent one.

It has been shown in Lemma 9.17 that the peripheral spectrum of the quasi-compact operator T contains only finitely many eigenvalues of T and $1 \in \sigma_{\text{per}}(T)$ by the positivity of T. Besides, the peripheral spectrum is cyclic as the spectral radius $r(T)$ is a pole of the resolvent by Proposition 8.17 and the above mentioned result in [Lot68]. Therefore, we can conclude that $\sigma_{\text{per}}(T)$ can only contain roots of unity. □

Accordingly, suppose that $T \in \mathcal{L}(X)$ is a positive quasi-compact operator on a Banach lattice X with $\|T\|_{op} = r(T) = 1$. Then clearly $\sigma(T) \subset \overline{B(0,1)}$ holds. The preceding results show that 1 is an isolated eigenvalue of T and the peripheral spectrum is cyclic. Let us denote by the positive integer l the number of spectral values in the peripheral spectrum. Then we have to discuss two cases separately:

1. $l = 1$: then $\sigma_{\text{per}}(T) = \{1\}$, otherwise
2. $\sigma_{\text{per}}(T) = \left\{e^{2\pi i \frac{k}{l}} : k \in \{1, \dots, l\}\right\}$.

The first case has already been characterized by Katznelson and Tzafriri [KT86], who have been shown that for every linear operator T on a Banach space X with $\|T\|_{op} \le 1$ the limit

$$\lim_{n\to\infty} \left\|T^{n+1} - T^n\right\|_{op} = \lim_{n\to\infty} \|T^n(T - I)\|_{op} = 0$$

holds if (and only if) $\sigma_{\mathrm{per}}(T) \subset \{1\}$, see also Proposition 9.8. In the following, we will consider quasi-compact operators and peripheral eigenvalues $\lambda \in \sigma_{\mathrm{per}}(T)$ and characterize the operators $T - \lambda I$ having ascent one. We will apply the results of the last section to construct a projection to the corresponding finite-dimensional eigenspace of T. We will show that the projection operator commutes with T and give a further characterization.

9.4.2 Operators with ascent one

The last results have shown that peripheral eigenvalues λ of quasi-compact operators, the operator $T - \lambda I$ has always ascent one. In this case the spaces M_λ^1 and Λ_λ^1 as considered in Section 9.3.3 contain only eigenvectors of T and T^* respectively. To simplify the notation let us denote them in the following by

$$M_\lambda := \ker(T - \lambda I) = \{x \in X : Tx = \lambda x\}, \tag{9.16}$$

$$\Lambda_\lambda := \ker(T^* - \lambda I) = \{x^* \in X^* : x^*(Tx) = x^*(\lambda x) \text{ for all } x \in X\}, \tag{9.17}$$

and let $n := \dim(M_\lambda) = \dim(\Lambda_\lambda)$. The result of Theorem 9.16 yields a projection $P : X \to M_\lambda$ onto the eigenspace associated with λ. Recall that the Gramian matrix where the eigenvectors of T^* are acting on the eigenvectors of T,

$$G := (\Phi^*\Phi) = \begin{pmatrix} e_1^*(e_1) & \cdots & e_1^*(e_n) \\ \vdots & & \vdots \\ e_m^*(e_1) & \cdots & e_m^*(e_n) \end{pmatrix} \in \mathbb{R}^{n\times n},$$

is invertible. Setting $A = (a_{ij}) = G^{-1}$, the projection has the form

$$Px = \Phi(\Phi^*\Phi)^{-1}\Phi^* = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}e_j^*(x)e_i, \qquad x \in X.$$

The next lemma gives a characterization of this projection.

Theorem 9.21. *Let $T \in \mathcal{L}(X)$ and $\lambda \in \sigma_p(T)$ such that $T - \lambda I$ is a Browder operator. Then the following two statements are equivalent:*

1. *$T - \lambda I$ has ascent one,*
2. *there exists a projection $P \in \mathcal{K}(X)$ such that $TP = PT = \lambda P$,*

Proof. Suppose that the first statement holds. Then we obtain a projection $P \in \mathcal{K}(X)$ from Theorem 9.16. We have the property $T \circ P = \lambda P$, because for $x \in X$ we obtain

$$(T \circ P)(x) = T\Big(\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}e_j^*(x)e_i\Big) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}e_j^*(x)T(e_i) = \lambda P(x),$$

as $Te_i = \lambda e_i$ by (9.16). Similarly, we obtain $P \circ T = \lambda P$. Namely, for $x \in X$ using $e_j^*(Tx) = e_j^*(\lambda x)$ by (9.17) it holds that

$$(P \circ T)(x) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} e_j^*(Tx) e_i = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} e_j^*(\lambda x) e_i = \lambda P(x).$$

Now we show that if there exists a projection $P \in \mathcal{K}(X)$ with $TP = PT = \lambda P$, then $T - \lambda I$ has ascent one. As $\ker(T - \lambda I) \subset \ker(T - \lambda I)^2$ and $\operatorname{ran}(P) = \ker(T - \lambda I)$, it is enough to show that $\ker(T - \lambda I)^2 \subset \operatorname{ran}(P)$. Suppose $x \in \ker(T - \lambda I)^2$. Then $(T - \lambda I)^2 x = 0$ and $(T - \lambda I)x \in \ker(T - \lambda I) = \operatorname{ran}(P)$. Therefore, there is $y \in \operatorname{ran}(P)$ such that $Py = (T - \lambda I)x$. Then

$$y = Py = P^2 y = P(T - \lambda I)x = PTx - \lambda Px = \lambda Px - \lambda Px = 0.$$

Thus, $y = 0$ and we obtain using $0 = y = Py = (T - \lambda I)x$ the final result, namely $x \in \ker(T - \lambda I) = \operatorname{ran}(P)$. □

Lemma 9.22. *Let $T \in \mathcal{L}(X)$ and $\lambda \in \sigma_p(T)$. Suppose there exists a projection $P \in \mathcal{K}(X)$ such that $TP = PT = \lambda P$. Then*

$$(T - TP)^n = T^n - T^n P = T^n - \lambda^n P$$

holds for all $n \in \mathbb{N}$.

Proof. We will use the fact that if P is a projection then $I - P$ is a projection as well. Also note that $I - P$ commutes with T, as

$$(I - P)T = T - TP = T - PT = T(I - P).$$

Now we derive the result with the following steps:

$$\begin{aligned}(T - TP)^n &= (T(I - P))^n = T^n (I - P)^n \\ &= T^n (I - P) = T^n - T^n P = T^n - \lambda^n P.\end{aligned}$$

□

9.4.3 The limit of the iterates of quasi-compact operators

Now we are able to prove the convergence of iterates of quasi-compact operators and derive the limiting operator. Let us assume in the following that T is a quasi-compact operator with $\|T\|_{op} = 1$ and $r(T) = 1$. We will restrict us first to the fixed point space of a quasi-compact operator $T \in \mathcal{L}(X)$ and assume that $\sigma(T) \subset B(0,1) \cup \{1\}$, i. e., 1 is the only peripheral eigenvalue of T. In this case, if $T - I \in \mathcal{W}_B(X)$ has ascent one and the iterates will converge to the projection operator P that projects onto the fixed point space of T. Later we will consider the case where the peripheral spectrum is cyclic.

In order to prove our main result we will need the following result that states that isolated spectral values can be removed by the projection operator on the corresponding generalized eigenspace.

Lemma 9.23. *Let $T \in \mathcal{L}(X)$ and $\lambda \in \sigma_p(T)$ such that $T - \lambda I \in \mathcal{W}_B(X)$ with ascent p. Let P be denote the projection onto* $\ker(T - \lambda I)^p$, *defined by Theorem 9.16. Then λ is an isolated spectral value and $\lambda \notin \sigma(T - TP)$.*

Proof. As usual let us denote the dimension of the generalized eigenspace $\ker(T-\lambda I)^p$ by n. We use the space decomposition $X = \ker(T - \lambda I)^p \oplus \operatorname{ran}(T - \lambda I)^p$. Then the operator T can be written as

$$\begin{pmatrix} J & 0 \\ 0 & S \end{pmatrix} \in \mathcal{L}(\ker(T - \lambda I)^p \oplus \operatorname{ran}(T - \lambda I)^p),$$

where J is the Jordan normal form of T on the finite-dimensional space $\ker(T - \lambda I)^p$ and $S \in \mathcal{L}(\operatorname{ran}(T - \lambda I)^p)$. As

$$T - TP = \begin{pmatrix} 0 & 0 \\ 0 & S \end{pmatrix},$$

it is enough to show, that $\lambda \notin \sigma(S)$. To this end, let us write $J = \lambda I_n + N$, where I_n is the identity matrix on $\mathbb{C}^n$ and N is a nilpotent matrix with $N^p = 0$. Then we can decompose the operator $(T - \lambda I)^p$ on $\ker(T - \lambda I)^p \oplus \operatorname{ran}(T - \lambda I)^p$ with basic linear algebra in the following way:

$$(T - \lambda I)^p = \begin{pmatrix} (\lambda I_n + N - \lambda I_n)^p & 0 \\ 0 & (S - \lambda I)^p \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & (S - \lambda I)^p \end{pmatrix}. \tag{9.18}$$

Let us now consider the operator $S - \lambda I$. This operator is one-to-one, because

$$\ker(S - \lambda I) \subset \ker(S - \lambda I)^p \subset \ker(T - \lambda I)^p,$$

where the last step follows from the decomposition of $(T - \lambda I)^p$ shown in (9.18). This implies in fact that $\ker(S - \lambda I)^p = \{0\}$ holds, as

$$\ker(T - \lambda I)^p \cap \operatorname{ran}(T - \lambda I)^p = \{0\}$$

by Proposition 8.9 and hence, the operator $S - \lambda I$ is one-to-one.

Next, we show that $S - \lambda I$ is onto, i. e., $\operatorname{ran}(S - \lambda I) = \operatorname{ran}(T - \lambda I)^p$. To this end, let $y \in \operatorname{ran}(T - \lambda I)^p$, i. e., there exists $x \in X$ such that $(T - \lambda I)^p x = y$. Using that $T - \lambda I = S - \lambda I$ on $\operatorname{ran}(T - \lambda I)^p$, we calculate

$$(S - \lambda I)y = (S - \lambda I)(T - \lambda I)^p x = (T - \lambda I)^{p+1} x$$

and derive that $\operatorname{ran}(S - \lambda I) = \operatorname{ran}(T - \lambda)^{p+1}$ on $\operatorname{ran}(T - \lambda I)^p$. As the chain length of $T - \lambda I$ equals p and hence $\operatorname{ran}(T - \lambda I)^{p+1} = \operatorname{ran}(T - \lambda I)^p$ holds, we conclude that $\operatorname{ran}(S - \lambda I) = \operatorname{ran}(T - \lambda I)^p$ and the proof is complete. □

Now we can state our main result that characterizes quasi-compact operators by the convergence of the iterates provided that the spectrum is located according to Katznelson and Tzafriri [KT86].

Theorem 9.24. *Let $T \in \mathcal{L}(X)$ with $r(T) = \|T\|_{op} = 1$ satisfying the spectral condition $\sigma(T) \subset B(0,1) \cup \{1\}$. Then T is quasi-compact if and only if*

$$\lim_{m \to \infty} \|T^m - P\|_{op} = 0,$$

where $P \in \mathcal{K}(X)$ is a finite-rank projection with $TP = PT = P$.

Proof. Clearly, if the iterates T^m converge to a finite-rank operator, then T is a quasi-compact operator as $r_{\text{ess}}(T) = 0$ in this case.

Now let T be quasi-compact with $\sigma(T) \subset B(0,1) \cup \{1\}$, then $r(T) = 1$ and 1 is an isolated peripheral eigenvalue. Thus, $T - I$ is Browder with ascent one. We now prove the limit of the iterates. By Theorem 9.15 the space X has the decomposition

$$X = \ker(T - I) \oplus \operatorname{ran}(T - I). \tag{9.19}$$

Therefore, we can decompose the operator T into

$$T = \begin{pmatrix} I & 0 \\ 0 & S \end{pmatrix} \in \mathcal{L}(\ker(T - I) \oplus \operatorname{ran}(T - I)),$$

with $S \in \mathcal{L}(\operatorname{ran}(T - I))$. Then

$$T - P = \begin{pmatrix} 0 & 0 \\ 0 & S \end{pmatrix},$$

where P is the projection operator defined by Theorem 9.16. Using Lemma 9.23 we derive that $1 \notin \sigma(S) \subset B(0,1) \cup \{1\}$, and hence, $\sigma(S) \subset B(0,1)$. Therefore, the spectral radius of S is strictly smaller than 1 and thus, the iterates S^m converge to 0 in the operator norm as m tends to infinity. Finally, applying Lemma 9.22 we obtain the final result

$$\lim_{m \to \infty} \|T^m - P\|_{op} = \lim_{m \to \infty} \|(T - P)^m\|_{op} = \lim_{m \to \infty} \|S^m\|_{op} = 0.$$

The iterates T^m converge in the uniform operator topology to the operator P, the projection onto the fixpoint space of T. □

Corollary 9.25 (Convergence Rate). *Let $T \in \mathcal{L}(X)$ be a quasi-compact operator with $r(T) = \|T\|_{op} = 1$ satisfying the spectral condition $\sigma(T) \subset B(0,1) \cup \{1\}$. Define*

$$\gamma := \sup\{|\gamma| : \gamma \in \sigma(T) \setminus \{1\}\}.$$

Then there exists a constant $1 \le C \le \gamma^{-1}$, such that for all $m \in \mathbb{N}$

$$\|T^m - P\|_{op} \le C \cdot \gamma^m,$$

where $P \in \mathcal{K}(X)$ is the operator defined by Theorem 9.16.

Proof. According to the proof of Theorem 9.24 we decompose

$$T = \begin{pmatrix} I & 0 \\ 0 & S \end{pmatrix} \in \mathcal{L}(\ker(T-I) \oplus \operatorname{ran}(T-I)).$$

Furthermore, we have that $\sigma(S) \subset B(0,1)$ and therefore we obtain $r(S) = \gamma < 1$. As $r(S) = \lim_{m\to\infty} \|S^m\|^{1/m}$, we obtain that there exists a constant $1 \le C \le \gamma^{-1}$ such that

$$\|S^m\| \le C \cdot \gamma^m$$

for every $m \in \mathbb{N}$. □

If a sequence of operators with the spectrum contained in $B(0,1) \cup \{1\}$ shares the same fixpoints spaces, the following limit theorem hold.

Corollary 9.26. *Let $T_n \in \mathcal{L}(X)$ be a sequence of continuous linear operators with $\sigma(T_n) \subset B(0,1) \cup \{1\}$ such that $T_n - \lambda I \in \mathcal{W}_B(X)$ has ascent one for all $n \in \mathbb{N}$. Furthermore, we assume that $\ker(T_n - I) = \ker(T_{n+1} - I)$ and $\ker(T_n^* - I) = \ker(T_{n+1}^*)$ for all $n \in \mathbb{N}$. Let $(k_n)_{n\in\mathbb{N}} \subset \mathbb{N}$ be a strictly increasing sequence of positive integers and let*

$$\gamma_n := \sup\{|\gamma| : \gamma \in \sigma(T_n) \setminus \{1\}\}.$$

If $\gamma_n^{k_n} \to 0$ for $n \to \infty$, then

$$\lim_{n\to\infty} \left\|T_n^{k_n} - P\right\|_{op} = 0,$$

where $P \in \mathcal{K}(X)$ is the operator defined by Theorem 9.16.

Proof. Follows analogously to Corollary 9.25 using the decomposition

$$T_n = \begin{pmatrix} I & 0 \\ 0 & S_n \end{pmatrix} \in \mathcal{L}(\ker(T_n - I) \oplus \operatorname{ran}(T_n - I)).$$

Then, we have that $\sigma(S_n) \subset B(0,1)$ and $r(S_n) = \gamma_n < 1$ for any $n \in \mathbb{N}$. Using Gelfand's formula,

$$r(S_n) = \lim_{m\to\infty} \|S_n^m\|_{op}^{1/m},$$

there exists a positive constant C such that

$$\|S_n^m\| \le C \cdot \gamma^m$$

for every $m \in \mathbb{N}$ and $n \in \mathbb{N}$. Consequently, we derive that

$$\left\|T_n^{k_n} - P\right\| = \left\|S_n^{k_n}\right\| \le C\gamma_n^{k_n}.$$

By the assumption $\gamma_n^{k_n} \to 0$ for $n \to \infty$ we obtain that $\left\|T_n^{k_n} - P\right\| \to 0$ if n tends to infinity. □

Finally we discuss the case, when the peripheral spectrum is cyclic.

Theorem 9.27. *Let $T \in \mathcal{L}(X)$ be a quasi-compact operator with $r(T) = \|T\|_{op} = 1$ with a non-trivial fixed point space. Furthermore, we assume the peripheral spectrum to be finite and cyclic. Then there exists $l \in \mathbb{N}$ such that*

$$\lim_{m\to\infty} \left\|T^{lm} - P\right\|_{op} = 0,$$

where $P \in \mathcal{K}(X)$ is the operator defined by Theorem 9.16 for applied to the operator T^l.

Proof. As the peripheral spectrum is finite and cyclic and $1 \in \sigma_{\text{per}}(T)$, the spectrum contains only roots of unity. Let us denote by l the number of spectral values contained in the spectrum. Then

$$\sigma_{\text{per}}(T) = \left\{\rho_l^k \;:\; k \in \{1, \ldots, l\}\right\},$$

where ρ_l is the l-th root of unity. By the spectral mapping theorem for the point spectrum, see e. g., Rudin [Rud91, Theorem 10.33] we conclude that the peripheral spectrum of T^l contains only the eigenvalue 1. As T^l is also quasi-compact, we can derive the result by Theorem 9.24 applied to T^l. □

9.4.4 Relation to ergodic theorems

We conclude this chapter by showing a relation between the theory developed in the last sections and uniform ergodic theorems. As already stated in Section 9.1.2, Sine [Sin70] has shown that if T is a contraction on a Banach space X then the Cesáro means

$$a_n(T) := n^{-1} \sum_{k=0}^{n-1} T^k$$

converge strongly for $n \to \infty$ if and only if the fixed points of T separate the fixed points of T^*.

We show here that for a contraction T where $T - I$ is a Weyl operator, i. e., a Fredholm operator of index 0, the fixed point separation property of Sine is equivalent to the property that $T - I$ has ascent one. This states in particular that $T - I$ is in fact a Browder operator.

Theorem 9.28. *Let $T \in \mathcal{L}(X)$ such that $\|T\|_{op} \le 1$ and $T - I \in \mathcal{W}(X)$. Then* $\ker(T - I)$ *separates the points of* $\ker(T^* - I)$ *if and only if the matrix*

$$G = \begin{pmatrix} e_1^*(e_1) & \cdots & e_1^*(e_n) \\ \vdots & & \vdots \\ e_n^*(e_1) & \cdots & e_n^*(e_n) \end{pmatrix} \in \mathbb{C}^{n \times n}$$

is invertible, where $n = \dim(T - \lambda I) = \dim(T^* - \lambda I)$.

Proof. We show first that if the fixed points of T separate the fixed points of T^* then the matrix G is invertible. To this end, let us assume to the contrary that the matrix G is not invertible. We will show that in this case $\ker(T - I)$ does not separate $\ker(T^* - I)$. If G is not invertible, then the rows of G are not linearly independent. Hence, we can assume there are $c_1, \ldots, c_n \in \mathbb{C}$ such that

$$\sum_{j=1}^{n} c_j e_j^*(e_i) = 0, \qquad \text{for all } i \in \{1, \ldots, n\},$$

where there is at least one coefficient with $c_k \neq 0$. Then

$$e_k^*(x) = \sum_{j \neq k} -\left(\frac{c_j}{c_k}\right) e_j^*(x)$$

for all $x \in \ker(T - I)$ as $e_1, \ldots, e_n$ form a basis. We conclude that $\ker(T - I)$ does not separate $\ker(T^* - I)$.

We prove next by contradiction that if G is invertible then $\ker(T - I)$ separates $\ker(T^* - I)$. To this end, assume that the fixed points of T do not separate the fixed points of T^*. Then there are $x_1^* \neq x_2^* \in \ker(T^* - I)$ such that for all $x \in \ker(T - I)$

$$x_1^*(x) = x_2^*(x).$$

Let $x_1^* = \sum_{j=1}^{n} c_j e_j^*$ and $x_2^* = \sum_{j=1}^{n} b_j e_j^*$. Then as well

$$x_1^*(e_i) - x_2^*(e_i) = \sum_{j=1} (c_j - b_j) e_j^*(e_i) = 0$$

holds for all $i \in \{1, \ldots, n\}$. As $c_i \neq b_i$ for at least one $i \in \{1, \ldots, n\}$ the rows of G are linearly dependent and G is not invertible. □

Finally, we extend our results of Theorem 9.16 with the result of the previous theorem.

Corollary 9.29. *Let $T \in \mathcal{L}(X)$ with $\|T\|_{op} \le 1$ such that $T - I \in \mathcal{W}(X)$. Then the following statements are equivalent:*

1. *$T - I$ has chain length one, i. e.,* $\operatorname{asc}(T - I) = \operatorname{dsc}(T - I) = 1$,

2. $X = \ker(T - I) \oplus \operatorname{ran}(T - I)$,
3. $T - I \in \mathcal{W}_B(X)$,
4. *G is invertible,*
5. $P = \Phi G^{-1} \Phi^*$ *yields a projection onto* $\ker(T - I)$,
6. *The Cesáro means* $n^{-1} \sum_{k=0}^{n-1} T^k$ *converge in the strong operator topology towards* P *for* $n \to \infty$.

CHAPTER 10 Lower Estimates for Operators with Smooth Range

“The essence of mathematics is its freedom.”

GEORG CANTOR

We prove lower bounds for the approximation error of quasi-compact contractions with smooth range in terms of classical moduli of smoothness and related K-functionals. Recall that a bounded linear operator T is said to be a contraction if $\|T\|_{op} \leq 1$. The operator T is defined here on an arbitrary open set $\Omega \subset \mathbb{R}^d$ and is not restricted to functions defined on the unit interval in contrast to Chapter 6. As underlying function spaces we consider the space of continuous functions and the L^p-spaces for $1 \leq p < \infty$. Consequently, we use the space of r-times continuously differentiable functions and the classical Sobolev spaces as their corresponding smooth subspaces. We will prove these lower estimates for linear operators based on a functional analytic framework depending on the fixed points of the operator and the smoothness of the range. The key idea is to estimate the semi-norm occuring in the K-functional by the approximation error using the convergence of the iterates of the operator. In this approach, one condition is that differential operators of a certain order annihilate the fixed points of T. Besides, these differential operators have to be bounded on the range of T. More details on the underlying concept have already been discussed in Chapter 6 in the one-dimensional setting on the unit interval.

We conclude this chapter with examples, where we show lower estimates for the variation-diminishing operators that have been introduced in Chapter 3, namely the integral Schoenberg operator, the Bernstein operator, and the Kantorovič operator. Lower estimates for the Schoenberg operator have already been proved in Chapter 4. Finally, we show how to derive lower estimates for positive linear operators with finite rank as shown in Chapter 7. The convergence of the iterates of such operators is always guaranteed. It will be shown that the degree of the modulus of smoothness or the used K-functional depends only on the smoothness of the range and the fixed points of T.

10.1 Preliminaries

As we will generalize the results of Chapter 4 to general linear operators with smooth range, we will provide here the necessary fundamentals and corresponding notation. To this end, let d be a positive integer and let Ω be an open subset of $\mathbb{R}^d$. We will first introduce the space of continuously differentiable functions on Ω and the related Sobolev spaces both equipped with a semi-norm. Afterwards, the modulus of smoothness and the K-functional are defined and a relation between them is outlined.

FUNCTION SPACES

We use the multi-index notation of Schwartz [Sch50b] to introduce derivatives. Accordingly, we denote by D^α the differential operator

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}},$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a multi-index with modulus $|\alpha| = \sum_{i=1}^n \alpha_i$. For a smooth function f, we denote its mixed partial derivative by

$$D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}, \tag{10.1}$$

In the following, we will introduce function spaces that are defined on an open set $\Omega \subset \mathbb{R}^n$. The space $C^r(\Omega)$ contains all complex valued functions f that have continuous and bounded derivatives $D^\alpha f$ up to order r, i. e., $|\alpha| \le r$. The norm on $C^r(\Omega)$ is given by

$$\|f\| := \sup_{|\alpha|=r} \|D^\alpha f\|_\infty .$$

The convergence of a sequence of functions on Ω means uniform convergence of the sequence itself and convergence of the sequence of its partial derivatives up to order r.

Next, we will define the $L^p(\Omega)$ spaces. For $1 \le p < \infty$, the space $L^p(\Omega)$ contains all Lebesgue measurable functions defined on Ω whose p-th power is integrable with respect to the measure $\mathrm{d}x = \mathrm{d}x_1 \cdots \mathrm{d}x_n = \mathrm{d}\mu$, i. e.,

$$\int_\Omega |f(x)|^p \,\mathrm{d}x < \infty$$

holds. Equipped with the norm

$$\|f\|_p = \left(\int_\Omega |f(x)|^p \,\mathrm{d}x \right)^{1/p}$$

the space $L^p(\Omega)$ becomes a Banach space. The Sobolev space $W^{p,r}(\Omega)$ corresponding to $L^p(\Omega)$ consists of all functions $f \in L^p(\Omega)$ whose derivatives $D^\alpha f \in L^p(\Omega)$ for all orders $|\alpha| \le r$. Here, the derivatives are understood in the distributional sense.

To simplify notation and to combine the previously mentioned spaces, we introduce the spaces $X^{p,r}(\Omega)$ for $1 \le p \le \infty$ and $r = 0, 1, 2, \ldots$ as follows:

$$\begin{aligned} X^{p,0}(\Omega) &:= L^p(\Omega), \quad 1 \le p < \infty; \qquad X^{\infty,0}(\Omega) := C(\Omega) \\ X^{p,r}(\Omega) &:= W^{p,r}(\Omega), \quad 1 \le p < \infty; \qquad X^{\infty,r}(\Omega) := C^r(\Omega) \end{aligned}$$

Finally, we define the semi-norms

$$|f|_{r,p} := \sup_{|\alpha|=r} \|D^\alpha f\|_p \tag{10.2}$$

for all smooth functions $f \in X^{p,r}(\Omega)$.

MODULI OF SMOOTHNESS AND K-FUNCTIONALS

Now, we will introduce the modulus of smoothness and Peetre's K-functional for the previously defined spaces according to Johnen and Scherer [JS77]. To simplify notation, let us denote for $h \in \mathbb{R}^d$ by $\Omega(h)$ the set

$$\Omega(h) := \{x \in \Omega \,:\, x + th \in \Omega \quad \text{for} \quad 0 \le t \le 1\}.$$

Then we define the r-th modulus of smoothness as follows.

Definition 10.1. The *modulus of smoothness of order r*, $\omega_{r,p} : X^{p,0}(\Omega) \times (0, \infty) \to [0, \infty)$, $1 \le p \le \infty$, is defined by

$$\omega_{r,p}(f, t) := \begin{cases} \|f\|_p, & r = 0 \\ \sup_{0<|h|\le t} \left\|\chi_{\Omega(rh)} \Delta_h^r f(x)\right\|_p, & r = 1, 2, \ldots. \end{cases}$$

where Δ_h^r is the forward difference operator into direction $h \in \mathbb{R}^d$,

$$\Delta_h^r f(x) = \sum_{l=0}^{r} (-1)^{r-l} \binom{r}{l} f(x + lh).$$

According to Peetre [Pee68], Johnen and Scherer [JS77] we define the K-functional on the spaces $X^{p,r}(\Omega)$ as follows:

Definition 10.2. The *K-functional* $K_{r,p} : X^{p,0}(\Omega) \times (0, \infty) \to [0, \infty)$, $1 \le p \le \infty$ is defined by

$$K_{r,p}(f, t^r) := \inf \left\{ \|f - g\|_p + t^r |g|_{p,r} \,:\, g \in X^{p,r}(\Omega) \right\}.$$

The next proposition shows that the modulus of smoothness can be bounded from above by the related K-functional. We provide the proof due to its brevity, as it only uses basic properties of the modulus of smoothness, see for instance (4.3) – (4.5).

Proposition 10.1 (Johnen and Scherer [JS77, Lem. 1]). *Let $\Omega \in \mathbb{R}^d$ be an open set. Then for all $0 < t < \infty$ and $f \in X^{p,0}(\Omega)$, $g \in X^{p,r}(\Omega)$, $1 \le p \le \infty$, there holds*

$$\omega_{r,p}(f,t) \le 2^r \|f - g\|_p + d^{r/2} t^r |g|_{r,p} . \tag{10.3}$$

Proof. We split f up into $f = (f - g) + g$ and obtain

$$\begin{aligned} \omega_{r,p}(f,t) &\le \omega_{r,p}(f-g,t) + \omega_{r,p}(g,t) \\ &= 2^r \|f - g\|_p + d^{r/2} t^r |g|_{r,p} . \end{aligned}$$

□

Moreover, the equivalence of the modulus of smoothness to the K-functional have been shown, see Butzer and Berens [BB67] for the one-dimensional case and Johnen and Scherer [JS77] for arbitrary Lipschitz domains.

10.2 Lower estimates

In the following, let Ω be an open subset of $\mathbb{R}^d$ and $1 \le p \le \infty$. We will consider a sequence of linear operators T_n defined on $X^{p,0}(\Omega)$ with smooth range $\operatorname{ran}(T_n) \subset X^{p,r}(\Omega)$ whose fixed point space $\ker(T_n - I)$ is annihilated by every differential operator D^α of order r that is bounded on $\operatorname{ran}(T_n)$. In this general setting, we will show that for all $s \ge r$ and $n > 0$ there is $t_n > 0$ and there are constants $M_1, M_2 > 0$ independent of n and $f \in X^{p,0}(\Omega)$, such that

$$M_1 \cdot \omega_{s,p}(f,t_n) \le \|T_n f - f\|_p \quad \text{and} \quad M_2 \cdot K_{s,p}(f,t_n^s) \le \|T_n f - f\|_p .$$

Here, $t_n \to 0$ for $n \to \infty$ provided that $\|f - T_n f\|_p \to 0$.

In order to prove these estimates, we will consider the case where the smooth function g in Proposition 10.1 is replaced by a smooth approximation $T_n f$. Then, we will estimate the semi-norm $|T_n f|_{r,p} = \sup \|D^\alpha T_n f\|_p$ with respect to the approximation error $\|T_n f - f\|_p$. The key concept – as already outlined in Chapter 6 – is to use the limiting operator of the iterates T^n. Hereby, the quasi-compactness of the operators T_n will guarantee the existence of the limiting operator. With this is mind, we can state the following lemma:

Lemma 10.2. *Let $1 \le p \le \infty$ and let $T : X^{p,0}(\Omega) \to X^{p,0}(\Omega)$ be a quasi-compact contraction, i. e., $\|T\|_{op} \le 1$. Suppose*

1. *$\sigma(T) \subset B(0,1) \cup \{1\}$,*
2. *$\operatorname{ran}(T) \subset X^{p,r}(\Omega)$ for some positive integer r,*
3. *D^α is bounded on $\operatorname{ran}(T)$ for all α with $|\alpha| = r$,*
4. *D^α annihilates $\ker(T - I)$ for all α with $|\alpha| = r$.*

Then for every $f \in X^{p,0}(\Omega)$,

$$|Tf|_{r,p} \le \frac{\sup_{|\alpha|=r} \|D^\alpha\|_{op:\mathrm{ran}(T)}}{1-\gamma} \, \|Tf - f\|_p \,,$$

where $\|D^\alpha\|_{op:\mathrm{ran}(T)}$ *is the operator norm of* D^α *on* $\mathrm{ran}(T)$ *and*

$$\gamma := \sup \{|\lambda| \,:\, \lambda \in \sigma(T) \text{ with } |\lambda| < 1\} \,.$$

Proof. As T is quasi-compact, $T - I$ is a Browder operator with ascent one. Due to the spectral property $\sigma(T) \subset B(0,1) \cup \{1\}$ and Corollary 9.25, there exists a projection P with $\mathrm{ran}(P) = \ker(T - I)$ and there exists a constant $0 \le C \le \gamma^{-1}$ such that

$$\|T^m - P\|_{op} \le C\gamma^m$$

holds for all integers $m > 0$. As the range of P is exactly the fixed point space of T, we have that $D^\alpha P = 0$ whenever $|\alpha| = r$.

Using these results we obtain

$$\begin{aligned}
|Tf|_{r,p} &= \sup_{|\alpha|=r} \|D^\alpha Tf\|_p = \sup_{|\alpha|=r} \left\|D^\alpha Tf - D^\alpha T^2 f + D^\alpha T^2 f - D^\alpha T^3 f + \dots\right\|_p \\
&\le \sup_{|\alpha|=r} \sum_{m=1}^{\infty} \|D^\alpha T^m (f - Tf)\|_p \\
&\le \|Tf - f\|_p \cdot \sup_{|\alpha|=r} \sum_{m=1}^{\infty} \|D^\alpha T^m\|_{op} \\
&= \|Tf - f\|_p \cdot \sup_{|\alpha|=r} \sum_{m=1}^{\infty} \|D^\alpha (T^m - P + P)\|_{op} \\
&= \|Tf - f\|_p \cdot \sup_{|\alpha|=r} \sum_{m=1}^{\infty} \|D^\alpha (T^m - P)\|_{op} \,,
\end{aligned}$$

as D^α annihilates $\ker(T - I)$ and therefore, $D^\alpha P = 0$. By the boundedness of D^α on $\mathrm{ran}(T)$ we get

$$\begin{aligned}
|Tf|_{r,p} &\le \|Tf - f\|_p \cdot \sup_{|\alpha|=r} \|D^\alpha\|_{op:\mathrm{ran}(T)} \sum_{m=1}^{\infty} \|T^m - P\|_{op} \\
&\le \|Tf - f\|_p \cdot \sup_{|\alpha|=r} \|D^\alpha\|_{op:\mathrm{ran}(T)} \cdot \sum_{m=1}^{\infty} C\gamma^m .
\end{aligned}$$

Using that $C \le 1/\gamma$ the series reduces to a convergent geometric series and we conclude the proof with

$$|Tf|_{r,p} \le \|Tf - f\|_p \cdot \sup_{|\alpha|=r} \|D^\alpha\|_{op:\mathrm{ran}(T)} \cdot \sum_{m=0}^{\infty} \gamma^m$$

$$\leq \frac{\sup_{|\alpha|=r} \|D^\alpha\|_{op}}{1-\gamma} \; \|Tf - f\|_p .$$

□

Note that the fourth condition of Lemma 10.2 is reflected in the shown estimate as for each $f \in \ker(T-I)$ we have that $\|Tf - f\|_p = 0$ and $|Tf|_{r,p} = 0$.

Using this lemma, we can state the main theorem of this chapter combining Proposition 10.1 and Lemma 10.2.

Theorem 10.3. *Let $1 \leq p \leq \infty$ and let $T : X^{p,0}(\Omega) \to X^{p,0}(\Omega)$ be a quasi-compact contraction that satisfies the following conditions:*

1. $\sigma(T) \subset B(0,1) \cup \{1\}$,
2. $\operatorname{ran}(T) \subset X^{p,r}(\Omega)$ *for some positive integer r,*
3. D^α *is bounded on* $\operatorname{ran}(T)$ *for all α with $|\alpha| = r$,*
4. D^α *annihilates* $\ker(T-I)$ *for all α with $|\alpha| = r$.*

Then

$$\omega_{r,p}(f,t) \leq \left(2^r + d^{r/2} t^r \frac{\sup_{|\alpha|=r} \|D^\alpha\|_{op:\operatorname{ran}(T)}}{1-\gamma}\right) \cdot \|Tf - f\|_p$$

holds for all $t \in (0,\infty)$, where $\gamma := \sup\{|\lambda| : \lambda \in \sigma(T) \text{ with } \lambda \neq 1\}$.

Proof. We apply Proposition 10.1 and get directly by Lemma 10.2 the stated result. □

Analogous we can state the result for a lower estimate by a K-functional.

Theorem 10.4. *Let $1 \leq p \leq \infty$ and let $T : X^{p,0}(\Omega) \to X^{p,0}(\Omega)$ be a quasi-compact contraction that satisfies the following conditions:*

1. $\sigma(T) \subset B(0,1) \cup \{1\}$,
2. $\operatorname{ran}(T) \subset X^{p,r}(\Omega)$ *for some positive integer r,*
3. D^α *is bounded on* $\operatorname{ran}(T)$ *for all α with $|\alpha| = r$,*
4. D^α *annihilates* $\ker(T-I)$ *for all α with $|\alpha| = r$.*

Then

$$K_{r,p}(f,t^r) \leq \left(1 + t^r \frac{\sup_{|\alpha|=r} \|D^\alpha\|_{op:\operatorname{ran}(T)}}{1-\gamma}\right) \cdot \|Tf - f\|_p$$

holds for all $t \in (0,\infty)$, where $\gamma := \sup\{|\lambda| : \lambda \in \sigma(T) \text{ with } \lambda \neq 1\}$.

Proof. Follows by the same argumentation as in the proof of Theorem 10.3. □

Corollary 10.5. *Let (T_n) be a sequence of continuous linear operators on $X^{p,0}(\Omega)$ that satisfies the conditions of Theorem 10.3 and Theorem 10.4. Besides, we assume that $\|T_n f - f\|_p \to 0$ holds for all $f \in X^{p,0}(\Omega)$ if n tends to infinity.*

Then, with setting $\gamma_n := \sup\{|\lambda| : \lambda \in \sigma(T_n) \setminus \{1\}\}$ *the uniform lower estimates*

$$\omega_{r,p}(f,\delta_n) \le (2^r + d^{r/2}) \cdot \|T_n f - f\|_p \quad \text{and} \quad K_{r,p}(f,\delta_n^r) \le 2 \cdot \|T_n f - f\|_p,$$

where

$$\delta_n = \left(\frac{1-\gamma_n}{\sup_{|\alpha|=r} \|D^\alpha\|_{op:\mathrm{ran}(T_n)}} \right)^{1/r}.$$

and $\delta_n \to 0$ *if* n *tends to infinity.*

Remark. The property that $\delta_n \to 0$ *if* n *tends to infinity follows by* $\|T_n f - f\|_p$ *for* $f \in C([0,1])$. *To assure that this property holds there are the following two options. Either the second largest eigenvalue tends in the modulus to one, i. e.,*

$$\gamma_n \to 1$$

which is satisfied as T_n *converges against the identity* I *in the strong operator topology, or* $\sup_{|\alpha|=r} \|D^r\|_{op:\mathrm{ran}(T_n)} \to \infty$. *The Bernstein operator, for example, fulfills both conditions, as we will see in the next section.*

Finally, we want to outline a generalization to derive lower estimates for a sequence of linear operators $(T_n)_{n\in\mathbb{N}}$ on arbitrary Banach spaces based on the K-functional where smoothness of the range is not necessary. The conditions depend on the underlying semi-norms defined on the range of T_n. Accordingly, the semi-norms have to annihilate the fixed points of T_n and are bounded on the range of T_n.

Theorem 10.6. *Let* $(X_1, \|\cdot\|_{X_1})$ *be a Banach space and* $(X_2, |\cdot|_{X_2})$ *be a quasi Banach space with* $X_2 \subset X_1$. *Consider* $T_n : X_1 \to X_2$ *as sequence of quasi-compact contractions. Suppose that the following conditions hold:*

1. $\sigma(T_n) \subset B(0,1) \cup \{1\}$,
2. *the semi-norm* $|\cdot|_{X_2}$ *annihilates* $\ker(T_n - I)$, *and*
3. $\sup_{f\in X_1, \|f\|_{X_1}=1} |T_n f|_{X_2} < \infty$.

Then

$$\frac{1}{2} \cdot \inf_{g\in X_2} \left(\|f-g\|_p + \delta_n^r \, |g|_{X_2} \right) \le \|T_n f - f\|_p,$$

where

$$\delta_n = \left(\frac{1-\gamma_n}{\sup_{f\in X_2, \|f\|_{X_2}=1} |T_n f|_{X_2}} \right)^{1/r}.$$

Proof. Follows directly along the lines of the proof of Theorem 10.3. □

10.3 Applications to Positive Linear Operators

We conclude this chapter with concrete examples. We prove lower estimates for the well-known operators from Chapter 3. As the lower estimate of the Schoenberg operator has already been proved in Chapter 4, we will only show the estimates for the Bernstein operator, the Kantorovič operator and the integral Schoenberg operator. Finally, we will show a general estimate for positive linear operators with finite rank as considered in Chapter 7.

10.3.1 Lower estimate for the Bernstein operator

Let $B_n : C([0,1]) \to C([0,1])$ be the Bernstein operator of order $n > 0$ defined by

$$B_n f(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f(\frac{k}{n}).$$

As shown in Chapter 3, this operator can reproduce constant and linear functions and interpolates at the endpoints of the unit interval. Therefore

$$\ker(B_n - I) = \operatorname{span}(1, x),$$

and

$$\ker(B_n^* - I) = \operatorname{span}(\delta_0, \delta_1)$$

As shown in Proposition 3.5, the eigenvalues $(\lambda_{k,n})$ of B_n are explicitly known for $k \in \{0, \dots, n\}$ by

$$\lambda_{k,n} = \frac{n!}{(n-k)!} \frac{1}{n^k}.$$

Clearly, we have $\sigma(B_n) \subset B(0,1) \cup \{1\}$, as

$$1 = \lambda_{0,n} = \lambda_{1,n} > \lambda_{2,n} > \dots > \lambda_{n,n} = \frac{n!}{n^n}.$$

Note that this property follows also by Theorem 7.5. The second largest eigenvalue γ_n of B_n is $\gamma_n := \lambda_{2,n} = \frac{n-1}{n}$. Moreover, by Theorem 9.24 we obtain the classical result of Kelisky and Rivlin [KR67], namely

$$\lim_{m\to\infty} \|B_n^m - L\|_{op} \to 0,$$

where $L = \delta_0 + (\delta_1 - \delta_0) \cdot e_1$.

The range of the Bernstein operator is given by the space of all polynomials with degree at most n. Thus, for $r < n$ we obtain using the representation of $D^r B_n f$ in

Lorentz [Lor86, p.24] the following upper bound for the operator norm of D^r on $\operatorname{ran}(B_n)$:

$$\|D^r\|_{op} \leq \frac{2^r n!}{(n-r)!}.$$

Finally, we obtain with Theorem 10.3 the lower estimate

$$\omega_r(f,t) \leq \left(2^r + t^r \frac{\frac{2^r \cdot n!}{(n-r)!}}{\frac{1}{n}}\right) \cdot \|Tf - f\|_\infty \leq 2^r \left(1 + n^{r+1} t^r\right) \cdot \|Tf - f\|_\infty$$

for all $t \in (0,\infty)$. For the case $r = 2$, we derive accordingly the following uniform estimate:

Corollary 10.7. *The approximation error of the Bernstein operator B_n can be uniformly bounded for all $f \in C([0,1])$ by*

$$\frac{1}{8}\omega_2\left(f, n^{-3/2}\right) \leq \|B_n f - f\|_\infty, \qquad n \to \infty.$$

Remark. Compared to the known lower estimate using the Ditzian-Totik modulus of smoothness as shown in Proposition 3.4 one would expect a decay rate of $n^{-1/2}$. The question arises, whether sharper estimates used in the proof can lead to this decay rate or if this is already the best possible lower estimate.

10.3.2 Lower estimate for the Kantorovič operator

Let us consider the Kantorovič operator $K_n : L^1([0,1]) \to C([0,1])$,

$$K_n f(x) = (n+1)\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}\int_{\frac{k}{n+1}}^{\frac{k+1}{n+1}} f(t)\mathrm{d}t, \qquad f \in L^1([0,1]), \quad x \in [0,1].$$

We have that $\ker(K_n - I) = \operatorname{span}\{1\}$ and similarly to the iterates of the integral Schoenberg operator

$$\lim_{m\to\infty} \|K_N^m - L\|_{op} = 0$$

holds, where $Lf = \int_0^1 f(t)\mathrm{d}t$. Clearly, $D1 = 0$, hence the differential operator D annihilates $\ker(K_n - I)$. Besides, D is bounded on $\operatorname{ran}(K_n)$ in the same way as the Bernstein operator:

$$\|DK_n f(x)\|_p = \left\|D^2 B_{n+1} F(x)\right\|_p \leq \left\|D^2\right\|_{op:\operatorname{ran}(B_{n+1})} \|f\|_1,$$

where $F(x) = \int_0^x f(t)\mathrm{d}t$. Therefore,

$$\|D\|_{op} \leq \left\|D^2\right\|_{op:\operatorname{ran}(B_{n+1})} = \frac{4(n+1)!}{(n+1-2)!} = 4(n^2+n).$$

holds. Combining these results with Theorem 10.3 we can state the lower estimate

$$\omega_{1,p}(f,t) \le \left(2 + t\frac{4(n^2+n)}{\frac{1}{n}}\right) \cdot \|Tf - f\|_\infty \le \left(2 + 4(n^3+n^2)t\right) \cdot \|Tf - f\|_\infty$$

for all $t \in (0,\infty)$. Consequently, we get the following uniform estimate:

Corollary 10.8. *The approximation error of the Kantorovič operator K_n can be uniformly bounded from below by*

$$\frac{1}{6}\,\omega_{1,p}\left(f, \frac{1}{n^3+n^2}\right) \le \|K_n f - f\|_\infty, \qquad n \to \infty,$$

for all $f \in L^1([0,1])$.

10.3.3 Lower estimate for the integral Schoenberg operator

The integral Schoenberg operator is defined by

$$V_{\Delta_n,k} f(x) := DS_{\Delta_n,k+1} F(x) = \sum_{j=-k}^{n-1} \int_{\xi_{j-1,k+1}}^{\xi_{j,k+1}} f(t)\mathrm{d}t \frac{N_{j,k}(x)}{\xi_{j,k+1} - \xi_{j-1,k+1}}, \tag{10.4}$$

where $F(x) = \int_0^x f(t)\mathrm{d}t$. More details are shown in Section 3.2.2. As shown in Section 7.5, we have that

$$\lim_{m\to\infty} \left\| V_{\Delta_n,k}^m - L \right\|_{op} = 0$$

holds, where $Lf = \int_0^1 f(t)\mathrm{d}t$. Besides, $\ker(V_{\Delta_n,k}^m - I) = \operatorname{span}\{1\}$ and $D1 = 0$ holds. In Section 3.2.3, it has been shown that

$$\sigma(V_{\Delta_n,k}) \subset B(0,1) \cup \{1\},$$

holds and 1 is an eigenvalue of the integral Schoenberg operator. The operator norm of the differential operator D, can be obtained similarly to the Kantorovič operator. We use here the relation

$$DV_{\Delta_n,k} f = D^2 S_{\Delta_n,k+1} F = \left\|D^2\right\|_{op:\operatorname{ran}(S_{\Delta_n,k+1})} \|f\|_1, \tag{10.5}$$

where $F(x) = \int_0^x f(t)\mathrm{d}t$. The operator norm of D^2 on $\operatorname{ran}(S_{\Delta_n,k+1})$ has been shown in Corollary 4.7. We get on $\operatorname{ran}(V_{\Delta_n,k})$ by (10.5) the bound:

$$\|D\|_{op:\operatorname{ran}(V_{\Delta_n,k})} \le \left\|D^2\right\|_{op:\operatorname{ran}(S_{\Delta_n,k+1})} = \left(\frac{2(k+1)}{|\Delta_n|_{\min}}\right)^2 d_{k+1}.$$

As all conditions of Corollary 10.5 are satisfied, we can state the following lower estimates:

Corollary 10.9. *Lower estimates for the integral Schoenberg opeartor $V_{\Delta_n,k}$ are given by*

$$\frac{1}{6}\omega_{1,p}(f,t(\Delta_n,k)) \le \|V_{\Delta_n,k}f - f\|_p \quad \text{and} \quad \frac{1}{5}K_{1,p}(f,t(\Delta_n,k)) \le \|V_{\Delta_n,k}f - f\|_p,$$

where

$$t(\Delta_n,k) = \frac{|\Delta_n|^2_{\min}}{(k+1)^2} \cdot \left(\frac{1-\gamma_{\Delta_n,k}}{d_{k+1}}\right).$$

10.3.4 Lower estimates for positive finite-rank operators

We generalize here the examples of the preceding sections for general positive linear operators with a partition of unity property. Let Ω be an open subset of $\mathbb{R}^d$ such that $X^{p,r}(\Omega)$ contains the constant function 1 with $\|1\|_p = 1$, for instance $X = (0,1)^d$. We consider a sequence of positive finite-rank operator $T_n : X^{p,0}(\Omega) \to X^{p,0}(\Omega)$,

$$T_n f = \sum_{k=1}^{n} \alpha_k^*(f) e_k, \qquad f \in X^{p,0}(\Omega), \tag{10.6}$$

where $e_1, \dots, e_n \in X^{p,r}(\Omega)$ are linearly independent, smooth positive functions that form a partition of unity; α_k^* are positive linear functionals satisfying $\|\alpha_k^*\| = \alpha_k^*(1) = 1$ and $\alpha_k^*(e_k) > 0$ for $k \in \{1, \dots, n\}$. By Theorem 7.5, the spectrum of T_n is characterized by

$$\sigma(T_n) \subset B(0,1) \cup \{1\}$$

and 1 is an eigenvalue of T_n due to the partition of unity property. Thus, to prove lower estimates with the technique shown in this chapter, only the last two conditions have to be checked. Thus, we can modify Corollary 10.5 as follows:

Corollary 10.10. *Let (T_n) be a sequence of continuous linear operators on $X^{p,0}(\Omega)$ of the form (10.6) such that $\|T_n f - f\|_p \to 0$ holds for all $f \in X^{p,0}(\Omega)$ if n tends to infinity. Let us denote with $\gamma_n := \sup\{|\lambda| : \lambda \in \sigma(T_n) \text{ with } \lambda \ne 1\}$. If*

(i) every differential operator of order r is bounded on $\operatorname{ran}(T_n)$, and

(ii) every differential operator of order r annihilates $\ker(T_n - I)$,

then the approximation error can be bounded from below by

$$\omega_{r,p}(f,\delta_n) \le (2^r+1) \cdot \|T_n f - f\|_p \quad \text{and} \quad K_{r,p}(f,\delta_n^r) \le 2 \cdot \|T_n f - f\|_p,$$

where

$$\delta_n = \left(\frac{1-\gamma_n}{\sup_{|\alpha|=r}\|D^\alpha\|_{op:\operatorname{ran}(T_n)}}\right)^{1/r}.$$

and $\delta_n \to 0$ if n tends to infinity.

CHAPTER 11 Conclusion

"An expert is someone who knows some of the worst mistakes that can be made in his subject, and how to avoid them."

WERNER HEISENBERG

We conclude this thesis with a short summary of the shown results. As solved problems naturally lead to new questions and problems, we will finally discuss open questions for further research.

Motivated by shape preserving properties of variation diminishing transforms, we have proved in Section 5.4 the uniform convergence of our curvature approximation using Schoenberg's splines. The curvature approximation is convexity preserving due to the variation diminishing property and features the possibility to detect C^2-singularities by the established lower estimates of Chapter 4 that relate the approximation error with local smoothness. Whereas the theory shows the convergence only for point-wise evaluations of the curve, we have also considered the task to estimate the curvature of digitized curves in Section 5.4.4. We have shown with numerical evaluations that our spline based curvature estimator achieves competitive accuracy compared to state of the art curvature estimators while our algorithm is significantly faster to compute. Another advantage of our algorithm is the ability to estimate the curvature of piecewise smooth curves and to localize occurring C^2-singularities using an multi-scale approach.

To be able to consider also other approximation operators, we have shown a general technique to prove lower estimates for operators having smooth range. We have proved in Chapter 7 that the iterates of positive finite-rank operators with a partition of unity property always converge. Using Fredholm theory and the concept of quasi-compactness, we have shown in Chapter 9 that the iterates of quasi-compact operators converge towards a finite-rank projection if the spectrum is contained in the unit ball where 1 is the only possible common point at the boundary. In particular, we have shown how to construct the limiting operator using a Gramian matrix where the dual fixed points operate on the fixed points of the operator. Using the uniform convergence of the iterates, we have provided a general technique to obtain lower estimates in terms of moduli of smoothness or K-functionals in Chapter 10.

When we have considered the spline based multi-scale algorithm for the curvature estimation and its numerical evaluation there are some open questions that may be considered in further research. An interesting problem to investigate is how the curvature estimations behave for a curve that is corrupted with noise. While the research of this thesis was restricted to spline based approximations, the proof of the curvature approximation also holds in a very general setting. Thus, we also want to consider other approximation operators. Our future research will especially investigate convolution operators that canonically lead to a multi-scale approach. Consider for instance the Gaussian scale space generated as solution of the heat equation. Another open problem is to study if a similar approach also works for the curvature estimation of surfaces.

References

[AR03] O. Agratini and I. A. Rus. "Iterates of a class of discrete linear operators via contraction principle". In: *Comment. Math. Univ. Carolin.* 44.3 (2003), pp. 555–563 (see p. 28)

[Aie04] P. Aiena. *Fredholm and local spectral theory, with applications to multipliers.* Kluwer Academic Publishers, Dordrecht, 2004 (see p. 100)

[Alt13] F. Altomare. "On some convergence criteria for nets of positive operators on continuous function spaces". In: *J. Math. Anal. Appl.* 398.2 (2013), pp. 542–552 (see p. 82).
DOI: 10.1016/j.jmaa.2012.09.010

[And87] T. Ando. "Totally positive matrices". In: *Linear Algebra Appl.* 90 (1987), pp. 165–219 (see p. 15).
DOI: 10.1016/0024-3795(87)90313-2

[AB86] H. Asada and M. Brady. "Curvature Primal Sketch." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-8.1 (1986), pp. 2–14 (see p. 52)

[Bad09] C. Badea. "Bernstein polynomials and operator theory". In: *Results Math.* 53.3-4 (2009), pp. 229–236 (see pp. 28, 41, 44, 82, 88).
DOI: 10.1007/s00025-008-0333-1

[BBSV03] C. Bardaro, P. Butzer, R. Stens, and G. Vinti. "Convergence in variation and rates of approximation for Bernstein-type polynomials and singular convolution integrals." English. In: *Analysis, München* 23.4 (2003), pp. 299–340 (see p. 29).
DOI: 10.1524/anly.2003.23.4.299

[BM75] J. R. Bennett and J. S. Mac Donald. "On the measurement of curvature in a quantized environment." In: *IEEE Trans. Comput.* 24 (1975), pp. 803–820 (see p. 52).
DOI: 10.1109/T-C.1975.224312

[BL72] H. Berens and G. G. Lorentz. "Inverse theorems for Bernstein polynomials". In: *Indiana Univ. Math. J.* 21 (1971/72), pp. 693–708 (see p. 27)

[Ber12] S. N. Bernstein. "Démonstration du théorème de *Weierstraß*, fondée sur le calcul des probabilités." French. In: *Commun. Soc. Math. Kharkow (2)* 13 (1912), pp. 1–2 (see pp. 3, 17, 26)

[BGKT02] L. Beutel, H. Gonska, D. Kacsó, and G. Tachev. "On variation-diminishing Schoenberg operators: new quantitative statements". In: *Multivariate approximation and interpolation with applications (Almuñécar, 2001).* Monogr. Real Acad. Ci. Exact. Fí s.-Quí m. Nat. Zaragoza, 20. Acad. Cienc. Exact. Fí s. Quí m. Nat. Zaragoza, Zaragoza, 2002, pp. 9–58 (see pp. 34–36, 41)

[Bro61] F. E. Browder. "On the spectral theory of elliptic differential operators. I". In: *Math. Ann.* 142 (1960/1961), pp. 22–130 (see p. 102)

[Bru65] Y. A. Brudnyi. "On a certain method for approximation of bounded functions given on a segement". Russian. In: *Studies of Contemporaray problems of the Constructive Theory of Functions (Proc. Second All—Union Conf., Baku, 1962)*. Ed. by I. Ibragimov. Izdat. Akad. Nauk Azerbeidzan. SSR. Baku, 1965, pp. 40–45 (see p. 27)

[BB67] P. L. Butzer and H. Berens. *Semi-groups of operators and approximation*. Die Grundlehren der mathematischen Wissenschaften, Band 145. Springer-Verlag New York Inc., New York, 1967 (see pp. 8, 45, 140)

[Căl66] G. Călugăreanu. "Sur les polynomes de S.N. Bernstein. Le spectre de l'operateur." Romanian. In: *Gaz. Mat., Bucur., Ser. A* 71 (1966), pp. 448–451 (see pp. 28, 37)

[CGM06] D. Cárdenas-Morales, P. Garrancho, and F. Muñoz-Delgado. "Shape preserving approximation by Bernstein-type operators which fix polynomials". In: *Applied Mathematics and Computation* 182.2 (2006), pp. 1615–1622 (see p. 9)

[Cas07] C. Castañón, J. Fraga, S. Fernandez, A. Gruber, and L. da F. Costa. "Biological shape characterization for automatic image recognition and diagnosis of protozoan parasites of the genus Eimeria". In: *Pattern Recognition* 40.7 (2007), pp. 1899–1910 (see p. 52)

[Chr01] O. Christensen. "Frames, Riesz bases, and discrete Gabor/wavelet expansions". In: *Bull. Amer. Math. Soc. (N.S.)* 38.3 (2001), 273–291 (electronic) (see p. 116). DOI: 10.1090/S0273-0979-01-00903-X

[Chr03] O. Christensen. *An introduction to frames and Riesz bases*. Applied and Numerical Harmonic Analysis. Birkhäuser Boston, Inc., Boston, MA, 2003 (see p. 116). DOI: 10.1007/978-0-8176-8224-8

[CLL14] D. Coeurjolly, J.-O. Lachaud, and J. Levallois. "Multigrid convergent principal curvature estimators in digital geometry". In: *Computer Vision and Image Understanding* 129 (2014), pp. 27–41 (see pp. 1, 53). DOI: http://dx.doi.org/10.1016/j.cviu.2014.04.013

[CW00] S. Cooper and S. Waldron. "The eigenstructure of the Bernstein operator". In: *J. Approx. Theory* 105.1 (2000), pp. 133–165 (see pp. 28, 29, 38). DOI: 10.1006/jath.2000.3464

[Cos04] L. Costa, S. dos Reis, R. Arantes, A. Alves, and G. Mutinari. "Biological shape analysis by digital curvature". In: *Pattern Recognition* 37.3 (2004), pp. 515–524 (see p. 52)

[CS66] H. B. Curry and I. J. Schoenberg. "On Pólya frequency functions. IV. The fundamental spline functions and their limits". In: *J. Analyse Math.* 17 (1966), pp. 71–107 (see pp. 31, 32)

[dSil84] M. R. da Silva. "Nonnegative order iterates of Bernstein polynomials and their limiting semigroup". In: *Portugal. Math.* 42.3 (1983/84), 225–248 (1985) (see p. 28)

[DLF07] F. De Vieilleville, J.-O. Lachaud, and F. Feschet. "Convex Digital Polygons, Maximal Digital Straight Segments and Convergence of Discrete Geometric Estimators". In: *Journal of Mathematical Imaging and Vision* 27.2 (2007), pp. 139–156 (see pp. 1, 53). DOI: 10.1007/s10851-007-0779-x

[dBD85] C. de Boor and R. DeVore. "A geometric proof of total positivity for spline interpolation". In: *Math. Comp.* 45.172 (1985), pp. 497–504 (see p. 32).
DOI: 10.2307/2008139

[dBoo73] C. de Boor. "The quasi-interpolant as a tool in elementary polynomial spline theory". In: *Approximation theory (Proc. Internat. Sympos., Univ. Texas, Austin, Tex., 1973)*. Academic Press, New York, 1973, pp. 269–276 (see pp. 42, 46)

[dBoo76] C. de Boor. "Total positivity of the spline collocation matrix". In: *Indiana Univ. Math. J.* 25.6 (1976), pp. 541–551 (see p. 32)

[dBoo01] C. de Boor. *A practical guide to splines*. Revised. Vol. 27. Applied Mathematical Sciences. Springer-Verlag, New York, 2001 (see pp. 4, 33)

[Din83] T. T. Din. "Operators of Kantorovich type and integral Schoenberg splines". In: *Acta Math. Hungar.* 42.3-4 (1983), pp. 189–203 (see p. 37).
DOI: 10.1007/BF01956766

[DT87] Z. Ditzian and V. Totik. *Moduli of smoothness*. Vol. 9. Springer Series in Computational Mathematics. Springer-Verlag, New York, 1987 (see p. 28).
DOI: 10.1007/978-1-4612-4778-4

[dCar76] M. P. do Carmo. *Differential geometry of curves and surfaces*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1976 (see p. 53)

[Dun43a] N. Dunford. "Spectral theory". In: *Bull. Amer. Math. Soc.* 49 (1943), pp. 637–651 (see pp. 108, 109)

[Dun43b] N. Dunford. "Spectral theory. I. Convergence to projections". In: *Trans. Amer. Math. Soc.* 54 (1943), pp. 185–217 (see pp. 5, 77, 79, 81, 87, 92, 104, 107–109, 112)

[Enf73] P. Enflo. "A counterexample to the approximation problem in Banach spaces." English. In: *Acta Math.* 130 (1973), pp. 309–317 (see p. 96).
DOI: 10.1007/BF02392270

[EMC11] H.-A. Esbelin, R. Malgouyres, and C. Cartade. "Convergence of binomial-based derivative estimation for 2 noisy discretized curves". In: *Theoretical Computer Science* 412.36 (2011), pp. 4805–4813 (see p. 66).
DOI: http://dx.doi.org/10.1016/j.tcs.2010.12.035

[Ess76] H. Esser. "On pointwise convergence estimates for positive linear operators on $C[a,b]$". In: *Nederl. Akad. Wetensch. Proc. Ser. A* ***79****=Indag. Math.* 38.3 (1976), pp. 189–194 (see p. 35)

[GK02] F. P. Gantmacher and M. G. Krein. *Oscillation matrices and kernels and small vibrations of mechanical systems*. Revised. AMS Chelsea Publishing, Providence, RI, 2002 (see pp. 13, 15, 16, 38)

[GK37] F. Gantmakher and M. Krein. "Sur les matrices completement non négatives et oscillatoires." French. In: *Compos. Math.* 4 (1937), pp. 445–476 (see pp. 13, 15)

[GHJS05] R. J. Gardner, A. Hobolth, E. B. V. Jensen, and F. B. Sørensen. "Shape discrimination by total curvature, with a view to cancer diagnostics". In: *Journal of Microscopy* 217.1 (2005), pp. 49–59 (see p. 52).
DOI: 10.1111/j.0022-2720.2005.01435.x

[GI10] I. Gavrea and M. Ivan. "On the iterates of positive linear operators preserving the affine functions". In: *J. Math. Anal. Appl.* 372.2 (2010), pp. 366–368 (see p. 82).
DOI: 10.1016/j.jmaa.2010.07.026

[GI11a] I. Gavrea and M. Ivan. "Asymptotic behaviour of the iterates of positive linear operators". In: *Abstr. Appl. Anal.* (2011), Art. ID 670509, 11 (see pp. 41, 82). DOI: 10.1155/2011/670509

[GI11b] I. Gavrea and M. Ivan. "On the iterates of positive linear operators". In: *J. Approx. Theory* 163.9 (2011), pp. 1076–1079 (see p. 82). DOI: 10.1016/j.jat.2011.02.012

[GI11c] I. Gavrea and M. Ivan. "The iterates of positive linear operators preserving constants". In: *Appl. Math. Lett.* 24.12 (2011), pp. 2068–2071 (see p. 82). DOI: 10.1016/j.aml.2011.05.044

[Ger31] S. Gershgorin. "Über die Abgrenzung der Eigenwerte einer Matrix." German. In: *Bull. Acad. Sci. URSS* 1931.6 (1931), pp. 749–754 (see pp. 84, 86)

[Goo96] T. N. T. Goodman. "Total positivity and the shape of curves". In: *Total positivity and its applications (Jaca, 1994).* Vol. 359. Math. Appl. Kluwer Acad. Publ., Dordrecht, 1996, pp. 157–186 (see pp. 20–22, 52, 56, 59, 113). DOI: 10.1007/978-94-015-8674-0_9

[Goo02] T. Goodman. "Shape preserving interpolation by curves". In: *Algorithms for Approximation IV. (Proc. of the 2001 International Symposion).* Ed. by J. Levesley, I. Anderson, and J. Mason. University of Huddersfield, 2002, pp. 24–35 (see p. 9)

[GO05] T. Goodman and B. Ong. "Shape-preserving interpolation by splines using vector subdivision." In: *Adv. Comput. Math.* 22.1 (2005), pp. 49–77 (see p. 9). DOI: 10.1007/s10444-005-1174-z

[Gru76] A. Grundmann. "Güteabschätzungen für den Kantorovic-Operator in der L_1-Norm." German. In: *Math. Z.* 150 (1976), pp. 45–47 (see p. 30). DOI: 10.1007/BF01213884

[HH01] H. Hennion and L. Hervé. *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness.* Vol. 1766. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2001 (see p. 127). DOI: 10.1007/b87874

[HK07] S. Hermann and R. Klette. "A Comparative Study on 2D Curvature Estimators". In: *ICCTA*. IEEE Computer Society, 2007, pp. 584–589 (see pp. 1, 53). DOI: http://doi.ieeecomputersociety.org/10.1109/ICCTA.2007.2

[Heu82] H. G. Heuser. *Functional analysis.* John Wiley & Sons, Ltd., Chichester, 1982 (see pp. 5, 80, 93, 97–99, 104, 105, 121, 127)

[Joh72] H. Johnen. "Inequalities connected with the moduli of smoothness". In: *Mat. Vesnik* 9(24) (1972), pp. 289–303 (see p. 8)

[JS77] H. Johnen and K. Scherer. "On the equivalence of the K-functional and moduli of continuity and some applications". In: *Constructive theory of functions of several variables (Proc. Conf., Math. Res. Inst., Oberwolfach, 1976).* Springer, Berlin, 1977, 119–140. Lecture Notes in Math., Vol. 571 (see pp. 8, 45, 139, 140)

[Kan30] L. Kantorovič. "Sur certains développements suivant les polynomes de la forme de S. Bernstein. I, II." Russian. In: *C. R. (Dokl.) Acad. Sci. URSS A* 1930 (1930), pp. 563–568, 595–600 (see p. 29)

[KZ70] S. Karlin and Z. Ziegler. "Iteration of positive approximation operators". In: *J. Approximation Theory* 3 (1970), pp. 310–339 (see pp. 28, 82)

[Kar68] S. Karlin. *Total positivity. Vol. I.* Stanford University Press, Stanford, Calif, 1968 (see pp. 4, 13, 15, 17–20, 27, 32, 56, 59)

[KT86] Y. Katznelson and L. Tzafriri. "On power bounded operators". In: *J. Funct. Anal.* 68.3 (1986), pp. 313–328 (see pp. 5, 77, 79–81, 86, 87, 92, 111, 112, 128, 132).
DOI: 10.1016/0022-1236(86)90101-1

[KR67] R. P. Kelisky and T. J. Rivlin. "Iterates of Bernstein polynomials". In: *Pacific J. Math.* 21 (1967), pp. 511–520 (see pp. 28, 82, 144)

[KZ94] H.-B. Knoop and X. L. Zhou. "The lower estimate for linear positive operators. II". In: *Results Math.* 25.3-4 (1994), pp. 315–330 (see p. 28).
DOI: 10.1007/BF03323413

[KO09] V. Kong and B. Ong. "Shape preserving approximation by spatial cubic splines". In: *Computer Aided Geometric Design* 26.8 (2009), pp. 888–903 (see p. 9)

[Kov01] V. Kovalevsky. "Curvature in digital 2D images". In: *International Journal of Pattern Recognition and Artificial Intelligence* 15.7 (2001), pp. 1183–1200 (see p. 53)

[KR48] M. G. Kreĭn and M. A. Rutman. "Linear operators leaving invariant a cone in a Banach space". Russian. In: *Uspehi Matem. Nauk (N. S.)* 3.1(23) (1948), pp. 3–95 (see p. 128)

[Kre79] E. Kreyszig. "Bernstein polynomials and numerical integration." In: *Int. J. Numer. Methods Eng.* 14 (1979), pp. 292–295 (see p. 90).
DOI: 10.1002/nme.1620140213

[LM01] P. Lamberti and C. Manni. "Shape-Preserving C^2 Functional Interpolation via Parametric Cubics". In: *Numerical Algorithms* 28.1-4 (2001), pp. 229–254 (see p. 9).
DOI: 10.1023/A:1014011303076

[Lay68] D. C. Lay. "Characterizations of the essential spectrum of F. E. Browder". In: *Bull. Amer. Math. Soc.* 74 (1968), pp. 246–248 (see p. 126)

[LCL14] J. Levallois, D. Coeurjolly, and J.-O. Lachaud. "Parameter-Free and Multigrid Convergent Digital Curvature Estimators". In: *Discrete Geometry for Computer Imagery*. Ed. by E. Barcucci, A. Frosini, and S. Rinaldi. Vol. 8668. Lecture Notes in Computer Science. Springer International Publishing, 2014, pp. 162–175 (see pp. 1, 53).
DOI: 10.1007/978-3-319-09955-2_14

[Lin74] M. Lin. "On the uniform ergodic theorem". In: *Proc. Amer. Math. Soc.* 43 (1974), pp. 337–340 (see p. 109)

[Lin75] M. Lin. "Quasi-compactness and uniform ergodicity of Markov operators". In: *Ann. Inst. H. Poincaré Sect. B (N.S.)* 11.4 (1975), 345–354 (1976) (see pp. 110, 111)

[Lon98] S. Loncaric. "A survey of shape analysis techniques". In: *Pattern Recognition* 31.8 (1998), pp. 983–1001 (see p. 52)

[Lor37] G. Lorentz. "Zur Theorie der Polynome von S. Bernstein." German. In: *Rec. Math. Moscou, n. Ser.* 2 (1937), pp. 543–556 (see pp. 27, 29)

[Lor53] G. G. Lorentz. *Bernstein polynomials*. Mathematical Expositions, no. 8. University of Toronto Press, Toronto, 1953 (see p. 29)

[Lor86] G. G. Lorentz. *Bernstein polynomials*. 2^{nd} ed. Chelsea Publishing Co., New York, 1986 (see pp. 4, 26, 27, 29, 145)

[Lot68] H. P. Lotz. "Über das Spektrum positiver Operatoren". In: *Math. Z.* 108 (1968), pp. 15–32 (see p. 128)

[Lyc78] T. Lyche. "A note on the condition numbers of the B-spline bases". In: *J. Approximation Theory* 22.3 (1978), pp. 202–205 (see p. 47)

[MBF08] R. Malgouyres, F. Brunet, and S. Fourey. "Binomial Convolutions and Derivatives Estimation from Noisy Discretizations". English. In: *Discrete Geometry for Computer Imagery*. Ed. by D. Coeurjolly, I. Sivignon, L. Tougne, and F. Dupont. Vol. 4992. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2008, pp. 370–379 (see p. 66). DOI: 10.1007/978-3-540-79126-3_33

[MD96] R. Marcondes Cesar Jr. and L. Da Fontoura Costa. "Towards effective planar shape representation with multiscale digital curvature analysis based on signal processing techniques". In: *Pattern Recognition* 29.9 (1996), pp. 1559–1569 (see p. 53)

[Mar70] M. J. Marsden. "An identity for spline functions with applications to variation-diminishing spline approximation". In: *J. Approximation Theory* 3 (1970), pp. 7–49 (see pp. 34, 35, 46, 49, 61, 62)

[MS66] M. Marsden and I. J. Schoenberg. "On variation diminishing spline approximation methods". In: *Mathematica (Cluj)* 8 (31) (1966), pp. 61–82 (see p. 34)

[MY87] G. Medioni and Y. Yasumoto. "Corner Detection and Curve Representation using Cubic B-Splines." In: *Computer vision, graphics, and image processing* 39.3 (1987), pp. 267–278 (see p. 52)

[Mic73] C. A. Micchelli. "The saturation class and iterates of the Bernstein polynomials". In: *J. Approximation Theory* 8 (1973), pp. 1–18 (see p. 28)

[MM92] F. Mokhtarian and A. K. Mackworth. "A theory of multiscale, curvature-based shape representation for planar curves". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14.8 (1992), pp. 789–805 (see p. 52)

[Mot36] T. Motzkin. *Beiträge zur Theorie der linearen Ungleichungen*. German. Basel: Diss. 73 S. (1936). 1936 (see p. 15)

[Mül76] M. W. Müller. "Die Gute der L_p-Approximation durch Kantorovic-Polynome." German. In: *Math. Z.* 151 (1976), pp. 243–247 (see p. 30). DOI: 10.1007/BF01214936

[Mül77] M. W. Müller. "Degree of L_p-approximation by integral Schoenberg splines". In: *J. Approximation Theory* 21.4 (1977), pp. 385–393 (see pp. 36, 37)

[MST99] G. A. Muñoz, Y. Sarantopoulos, and A. Tonge. "Complexifications of real Banach spaces, polynomials and multilinear maps". In: *Studia Math.* 134.1 (1999), pp. 1–33 (see p. 8)

[Nag80] J. Nagel. "Asymptotic properties of powers of Bernstein operators". In: *J. Approx. Theory* 29.4 (1980), pp. 323–335 (see pp. 28, 82). DOI: 10.1016/0021-9045(80)90120-3

[Nag82] J. Nagel. "Asymptotic properties of powers of Kantorovič operators". In: *J. Approx. Theory* 36.3 (1982), pp. 268–275 (see pp. 30, 82, 91).
DOI: 10.1016/0021-9045(82)90046-6

[NCF14] J. Nagler, P. Cerejeiras, and B. Forster. "Lower bounds for the approximation with variation-diminishing splines". In: *ArXiv e-prints* (Feb. 2014) (see pp. 6, 42).
arXiv: 1402.2403

[NK13] J. Nagler and U. Kähler. "A lower bound for the uniform Schoenberg operator". In: *ArXiv e-prints* (Dec. 2013) (see p. 41).
arXiv: 1312.5546

[Nag14] J. Nagler. "Short Communication: Curvature Approximation in Images". In: *PAMM* 14.1 (2014), pp. 949–950 (see p. 6).
DOI: 10.1002/pamm.201410455

[Nag15] J. Nagler. "On the spectrum of positive linear operators with a partition of unity property". In: *Journal of Mathematical Analysis and Applications* 425.1 (2015), pp. 249–258 (see pp. 5, 6, 82, 114).
DOI: http://dx.doi.org/10.1016/j.jmaa.2014.12.025

[NRW76] G. M. Nielson, R. F. Riesenfeld, and N. A. Weiss. "Iterates of Markov operators". In: *J. Approximation Theory* 17.4 (1976), pp. 321–331 (see p. 28)

[Nür89] G. Nürnberger. *Approximation by spline functions*. Springer-Verlag, Berlin, 1989 (see pp. 4, 33).
DOI: 10.1007/978-3-642-61342-5

[Orl34] W. Orlicz. "Ein Satz über die Erweiterung von linearen Operationen." German. In: *Stud. Math.* 5 (1934), pp. 127–140 (see p. 29)

[PC12] Y. Pan and M. Cheng. "A survey on shape preserving interpolation". In: *2012 International Conference on Systems and Informatics, ICSAI 2012*. 2012, pp. 2521–2523 (see p. 9)

[Pas12] A. Pasqualato, A. Palombo, A. Cucina, M. Mariggiò, L. Galli, D. Passaro, S. Dinicola, S. Proietti, F. D'Anselmi, P. Coluccia, and M. Bizzarri. "Quantitative shape analysis of chemoresistant colon cancer cells: Correlation between morphotype and phenotype". In: *Experimental Cell Research* 318 (7 2012) (see p. 52).
DOI: 10.1016/j.yexcr.2012.01.022

[Pee68] J. Peetre. *A theory of interpolation of normed spaces*. Notas de Matemática, No. 39. Instituto de Matemática Pura e Aplicada, Conselho Nacional de Pesquisas, Rio de Janeiro, 1968 (see pp. 7, 139)

[Pin10] A. Pinkus. *Totally positive matrices*. Vol. 181. Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 2010 (see pp. 15, 16)

[Pop38] T. Popoviciu. "Sur l'approximation des fonctions convexes d'ordre supérieur". In: *Bull. Sect. Sci. Acad. Roum.* 20 (1938), pp. 192–195 (see p. 20)

[RL11] T. Roussillon and J.-O. Lachaud. "Accurate Curvature Estimation along Digital Contours with Maximal Digital Circular Arcs". In: *Combinatorial Image Analysis*. Ed. by J. Aggarwal, R. Barneva, V. Brimkov, K. Koroutchev, and E. Korutcheva. Vol. 6636. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2011, pp. 43–55 (see pp. 1, 53, 66).
DOI: 10.1007/978-3-642-21073-0_7

[Rud91] W. Rudin. *Functional analysis*. 2nd ed. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, 1991 (see pp. 5, 6, 43, 84, 85, 93–97, 134)

[Rus04] I. A. Rus. "Iterates of Bernstein operators, via contraction principle". In: *J. Math. Anal. Appl.* 292.1 (2004), pp. 259–261 (see pp. 28, 82).
DOI: 10.1016/j.jmaa.2003.11.056

[Rus86] A. F. Ruston. *Fredholm theory in Banach spaces*. Vol. 86. Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1986 (see pp. 8, 82, 97, 100).
DOI: 10.1017/CBO9780511569180

[Sas64] D. W. Sasser. "Quasi-positive operators". In: *Pacific J. Math.* 14 (1964), pp. 1029–1037 (see p. 103)

[Sch74] H. H. Schaefer. *Banach lattices and positive operators*. Springer-Verlag, New York-Heidelberg, 1974 (see pp. 7, 128)

[Sch66] M. Schechter. "On the essential spectrum of an arbitrary operator. I". In: *J. Math. Anal. Appl.* 13 (1966), pp. 205–215 (see p. 102)

[SS99] K. Scherer and A. Y. Shadrin. "New upper bound for the B-spline basis condition number. II. A proof of de Boor's 2^k-conjecture". In: *J. Approx. Theory* 99.2 (1999), pp. 217–229 (see p. 47).
DOI: 10.1006/jath.1998.3310

[Sch30] I. J. Schoenberg. "Über variationsvermindernde lineare Transformationen". In: *Math. Z.* 32.1 (1930), pp. 321–328 (see pp. 4, 13–15).
DOI: 10.1007/BF01194637

[Sch46a] I. J. Schoenberg. "Contributions to the problem of approximation of equidistant data by analytic functions. Part A. On the problem of smoothing or graduation. A first class of analytic approximation formulae". In: *Quart. Appl. Math.* 4 (1946), pp. 45–99 (see p. 31)

[Sch46b] I. J. Schoenberg. "Contributions to the problem of approximation of equidistant data by analytic functions. Part B. On the problem of osculatory interpolation. A second class of analytic approximation formulae". In: *Quart. Appl. Math.* 4 (1946), pp. 112–141 (see p. 31)

[Sch50a] I. J. Schoenberg. "On Pólya frequency functions. II. Variation-diminishing integral operators of the convolution type". In: *Acta Sci. Math. Szeged* 12.Leopoldo Fejer et Frederico Riesz LXX annos natis dedicatus, Pars B (1950), pp. 97–106 (see p. 18)

[Sch59] I. J. Schoenberg. "On variation diminishing approximation methods". In: *On numerical approximation. Proceedings of a Symposium, Madison, April 21-23, 1958*. Edited by R. E. Langer. Publication no. 1 of the Mathematics Research Center, U.S. Army, the University of Wisconsin. The University of Wisconsin Press, Madison, 1959, pp. 249–274 (see p. 18)

[Sch67] I. J. Schoenberg. "On spline functions". In: *Inequalities (Proc. Sympos. Wright-Patterson Air Force Base, Ohio, 1965)*. Academic Press, New York, 1967, pp. 255–291 (see pp. 3, 32–34, 36, 61)

[Sch48] I. Schoenberg. "On variation-diminishing integral operators of the convolution type." In: *Proc. Natl. Acad. Sci. USA* 34 (1948), pp. 164–169 (see pp. 18, 19).
DOI: 10.1073/pnas.34.4.164

[Sch07] L. L. Schumaker. *Spline functions: basic theory*. 3rd ed. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2007 (see pp. 4, 33). DOI: 10.1017/CBO9780511618994

[Sch50b] L. Schwartz. *Théorie des distributions. Tome I*. French. Paris: Hermann & Cie. 148 p. (1950). 1950 (see p. 138)

[Sin70] R. Sine. "A mean ergodic theorem". In: *Proc. Amer. Math. Soc.* 24 (1970), pp. 438–439 (see pp. 110, 126, 134)

[Šuk15] T. Šukilović. "Curvature based shape detection." In: *Comput. Geom.* 48.3 (2015), pp. 180–188 (see p. 52). DOI: 10.1016/j.comgeo.2014.09.005

[Tim94] A. F. Timan. *Theory of approximation of functions of a real variable*. Dover Publications, Inc., New York, 1994 (see p. 45)

[TM07] T. Tuytelaars and K. Mikolajczyk. "Local invariant feature detectors: A survey". In: *Foundations and Trends in Computer Graphics and Vision* 3.3 (2007), pp. 177–280 (see p. 52)

[Utc03] S. Utcke. "Error-bounds on curvature estimation". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 2695 (2003), pp. 657–666 (see p. 53)

[Wen97] H.-J. Wenz. "On the limits of (linear combinations of) iterates of linear operators". In: *J. Approx. Theory* 89.2 (1997), pp. 219–237 (see p. 28). DOI: 10.1006/jath.1996.3039

[Wey09] H. Weyl. "Über beschränkte quadratische Formen, deren Differenz vollstetig ist." German. In: *Rend. Circ. Mat. Palermo* 27 (1909), pp. 373–392, 402 (see pp. 99, 102). DOI: 10.1007/BF03019655

[Whi52] A. M. Whitney. "A reduction theorem for totally positive matrices". In: *J. Analyse Math.* 2 (1952), pp. 88–92 (see p. 15)

[WS93] M. Worring and A. Smeulders. "Digital Curvature Estimation". In: *Computer Vision and Image Understanding* 58.3 (1993), pp. 366–382 (see p. 52)

[YT15] H. Yalim Keles and S. Tari. "A robust method for scale independent detection of curvature-based criticalities and intersections in line drawings". In: *Pattern Recognition* 48.1 (2015), pp. 140–155 (see p. 52)

[ZM10] B. Zhong and K.-K. Ma. "On the convergence of planar curves under smoothing". In: *IEEE Transactions on Image Processing* 19.8 (2010), pp. 2171–2189 (see p. 52)

[Zyg02] A. Zygmund. *Trigonometric series. Vols. I, II*. 3rd ed. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2002 (see p. 45)

List of Symbols

NUMBERS AND FIELDS

$\mathbb{N}$	The set of natural numbers: $0, 1, 2, 3, \dots$.
$\mathbb{R}$	The field of real numbers.
$\mathbb{C}$	The field of complex numbers.

FUNCTION SPACES

D^r	Differential operator of order $r \in \mathbb{N}$.
D^α	Mixed differential operator where α is a multi-index.
$\|\alpha\|$	The modulus of the multi-index α, $\|\alpha\| = \sum_{i=1}^n \alpha_i$.
$C(\Omega)$	The space of continuous functions $f : \Omega \to \mathbb{C}$.
$C^r(\Omega)$	The space of r-times continuously differentiable functions $f : \Omega \to \mathbb{C}$.
$L^p(\Omega)$	The space of measurable functions $f : \Omega \to \mathbb{C}$ where $\int_\Omega \|f(x)\|^p \, dx < \infty$.
$W^{p,r}(\Omega)$	The Sobolev space defined as all $f \in L^p(\Omega)$ where $D^\alpha f \in L^p(\Omega)$, $\|\alpha\| \leq r$.

BANACH SPACES AND LINEAR OPERATORS

X, Y	Banach spaces over $\mathbb{C}$.
X^*, Y^*	Corresponding dual spaces of X and Y.
$\overline{M}$	The closure of a set M.
$M^\perp$	The annihilator of a set $M \subset X$.
$\Lambda_\perp$	The pre-annihilator of a set $\Lambda \subset X^*$.
$\mathcal{L}(X, Y)$	Continuous linear operators mapping X to Y.
$\mathcal{L}(X)$	Continuous linear operators on X.
I	Identity on $\mathcal{L}(X)$.
T, S	Linear operators in $\mathcal{L}(X)$.
T^*	The adjoint of $T \in \mathcal{L}(X)$.
$\ker T$	The null space of $T \in \mathcal{L}(X)$.
$\operatorname{ran} T$	The range of $T \in \mathcal{L}(X)$.
$\\|\cdot\\|_X$	The norm on X.
$\\|\cdot\\|_{X^*}$	The norm on X^*.
$\\|\cdot\\|_{op}$	The operator norm on $\mathcal{L}(X)$.

SPECIAL CLASSES OF LINEAR OPERATORS

$\mathcal{F}(X)$	Finite-rank operators defined on X.
$\mathcal{K}(X)$	Compact operators defined on X.
$\operatorname{asc}(T)$	Ascent of T defined as smallest $k \in \mathbb{N}$ such that $\ker(T^k) = \ker(T^{k+1})$.
$\operatorname{dsc}(T)$	Descent of T defined as smallest $k \in \mathbb{N}$ such that $\operatorname{ran}(T^k) = \operatorname{ran}(T^{k+1})$.
$\alpha(T)$	Nullity of T defined as $\dim\ker(T)$.
$\beta(T)$	Deficiency of T defined as $\dim\ker(T^*)$.
$\operatorname{ind}(T)$	Index of T defined as value of $\alpha(T) - \beta(T)$.
$\Phi(X)$	Fredholm operators defined as $T \in \mathcal{L}(X)$ where $\alpha(T) < \infty$ and $\beta(T) < \infty$.
$\Phi_-(X)$	All Fredholm operators where $\operatorname{ind}(T) \leq 0$.
$\mathcal{R}(X)$	Riesz operators as operators $T \in \mathcal{L}(X)$ where $T - \lambda I \in \Phi(X)$ for all $\lambda \in \mathbb{C}$.
$\mathcal{W}(X)$	Weyl operators defined as Fredholm operators where $\operatorname{ind}(T) = 0$.
$\mathcal{W}_B(X)$	Browder operators defined as Weyl operators where $\operatorname{asc}(T) < \infty$.

SPECTRA OF LINEAR OPERATORS

$\rho(T)$	Resolvent set of T, $\rho(T) := \left\{\lambda \in \mathbb{C} : (T - \lambda I)^{-1} \text{ is invertible}\right\}$.
$R(T, \lambda)$	Resolvent of T corresponding to $\lambda \in \mathbb{C}$ defined as $(T - \lambda I)^{-1}$.
$\sigma(T)$	Spectrum of T, $\sigma(T) := \mathbb{C} \setminus \rho(T)$.
$\sigma_p(T)$	Point-spectrum of T, $\sigma_p(T) := \{\lambda \in \mathbb{C} : T - \lambda I \text{ is not one-to-one}\}$.
$\sigma_{\mathrm{per}}(T)$	Peripheral spectrum of T, $\sigma_{\mathrm{per}}(T) := \sigma(T) \cap \{\lambda \in \mathbb{C} : \lvert\lambda\rvert = r(T)\}$.
$\sigma_{\mathrm{ess}}(T)$	(Weyl) Essential spectrum of T, $\sigma_{\mathrm{ess}}(T) := \{\lambda \in \mathbb{C} : T - \lambda I \notin \mathcal{W}(X)\}$.
$\sigma_b(T)$	(Browder) Essential spectrum of T, $\sigma_b(T) := \{\lambda \in \mathbb{C} : T - \lambda I \notin \mathcal{W}_B(X)\}$.
$r(T)$	Spectral radius of T, $r(T) := \sup\{\lvert\lambda\rvert : \lambda \in \sigma(T)\}$.
$r_{\mathrm{ess}}(T)$	Essential spectral radius of T, $r_{\mathrm{ess}}(T) := \sup\{\lvert\lambda\rvert : \lambda \in \sigma_{\mathrm{ess}}(T)\}$.

Index