Interpolation and Approximation in Taylor Spaces

Barbara Zwicknagl[1] and Robert Schaback[2]

Abstract: The univariate Taylor formula without remainder allows to reproduce a function completely from certain derivative values. Thus one can look for Hilbert spaces in which the Taylor formula acts as a reproduction formula. It turns out that there are many Hilbert spaces which allow this, and they should be called **Taylor** spaces. They have certain reproducing kernels which are either polynomials or power series with nonnegative coefficients. Consequently, Taylor spaces can be spanned by translates of various classical special functions such as exponentials, rationals, hyperbolic cosines, logarithms, and Bessel functions. Since the theory of kernel–based interpolation and approximation is well–established, this leads to a variety of results. In particular, interpolation by shifted exponentials, rationals, hyperbolic cosines, logarithms, and Bessel functions provides exponentially convergent approximations to analytic functions, generalizung the classical Bernstein theorem for polynomial approximation to analytic functions.

Keywords: Univariate interpolation - nonlinear approximation - Kernels - Reproducing kernel Hilbert spaces - Error bounds - Convergence orders

Mathematics Subject Classification: 30D10 - 30E10- 30H10 - 30H20 - 41A05 - 41A10 - 41A20 - 41A25 - 41A30 - 41A58 - 68D05

1 Introduction

The theory of reproducing kernels K in Hilbert spaces H of functions on some domain Ω is well-known [4, 10], and there are plenty of applications in Numerical Analysis, Stochastics, and Machine Learning [13]. The usual reproduction formula is

$$f(x) = (f, K(x, \cdot))_H \text{ for all } x \in \Omega,\ f \in H, \tag{1}$$

but one of the most important reproduction formulas known otherwise is the *Taylor formula*

$$f(x) = \sum_{j=0}^{\infty} \frac{f^{(j)}(0)}{j!} x^j, \quad \text{for all } x \in I,\ f \in F \tag{2}$$

for functions in some space F of analytic functions in some interval I around the origin of $\mathbb{R}$.

[1]Department of Mathematical Sciences, Wean Hall 6113, Carnegie Mellon University, Pittsburgh, PA 15213, email: bzwick@andrew.cmu.edu

[2]Institut für Numerische und Angewandte Mathematik, Universität Göttingen, Lotzestr. 16-18, 37083 Göttingen, Germany, email: schaback@math.uni-goettingen.de

These formulas look different at first sight, but this paper focuses on cases where the Taylor formula (2) takes the form (1) of a reproduction formula in I. It turns out that there is a variety of interesting kernels allowing this identification, i.e. roughly all kernels $K(x,t)$ which are power series in xt with nonnegative coefficients. Each of these kernels defines a "native" Hilbert space in which it is reproducing via both (2) and (1), and allows the application of standard results concerning interpolation and approximation by spaces of the form

$$F_X := \text{Span}\{K(x,\cdot)\ :\ x \in X\}. \tag{3}$$

for fixed finite sets $X := \{x_0, \ldots, x_n\} \subset I$. These spaces can consist of exponentials, rationals, hyperbolic cosines, logarithms, and Bessel functions, for instance, depending on which kernel is chosen. Thus we can derive results on approximation and interpolation by those families from known results on kernel–based interpolation and approximation. These include optimality properties and error bounds, the latter being particularly interesting when they take the form of Bernstein–type theorems concerning spectral approximation orders for approximations of analytic functions. Such theorems come as special cases of general results of [15] on power series kernels. The paper closes with some numerical examples.

2 Taylor Spaces

We deal with Taylor spaces in the spirit of [8], and hope that there will be no confusion with the related Taylor spaces introduced by Calderon and Zygmund [6]. We fix a subset $\mathcal{N}$ of $\mathbb{N} := \{0, 1, 2, \ldots\}$ and consider functions f for which the remainder–free Taylor formula

$$f(x) = \sum_{j \in \mathcal{N}} \frac{f^{(j)}(0)}{j!} x^j, \tag{4}$$

is valid and allows reproduction of f via its derivatives at zero.

If $\mathcal{N}$ is finite and of the form $\mathcal{N} := \{0, 1, \ldots, n\}$, the admissible functions f form the space $\mathbb{P}_n$ of polynomials of degree at most n. Other finite sets $\mathcal{N}$ lead to special spaces of *lacunary* polyomials. If $\mathcal{N}$ is an infinite subset of $\mathbb{N} := \{0, 1, 2, \ldots\}$, we shall require (4) to be valid in a neighborhood of the origin with absolute convergence. This suggests to work in the complex plane $\mathbb{C}$ in all cases, allow complex-valued functions, and assume (4) to hold at least on some disc

$$D_R := \{x \in \mathbb{C}\ :\ |x| < R\}$$

of a finite radius $R > 0$ with absolute convergence. But we can also consider cases where (4) is absolutely convergent in the full complex plane, defining an entire function f there, and we shall refer to this case via $R = \infty$.

Before we take a closer look at spaces of functions satisfying (4) for infinite $\mathcal{N}$, we consider ways to turn (4) into a reproduction formula of the form (1). Here, we focus on a special technique [12] using weighted series expansions. We take positive weights λ_j for all $j \in \mathcal{N}$ satisfying

$$\begin{aligned} \sum_{j\in\mathcal{N}} \frac{R^{2j}\lambda_j}{(j!)^2} &< \infty, \qquad R < \infty \\ \sum_{j\in\mathcal{N}} \frac{\lambda_j e^j}{j!\sqrt{j}} &< \infty, \qquad R = \infty \end{aligned} \tag{5}$$

noting that Stirling's formula implies the first condition for all $R > 0$ if the second is satisfied. Then we form the weighted inner product

$$(f, g) := \sum_{j\in\mathcal{N}} \frac{f^{(j)}(0)\overline{g^{(j)}(0)}}{\lambda_j} \tag{6}$$

on the space of all functions f which have all derivatives $f^{(j)}(0)$ for $j \in \mathcal{N}$, satisfy (4) and additionally also

$$\|f\|^2 := \sum_{j\in\mathcal{N}} \frac{|f^{(j)}(0)|^2}{\lambda_j} < \infty. \tag{7}$$

We denote the completion of the space of such functions under the above inner product by F and call the resulting Hilbert space a *Taylor space*. We drop the dependence of F on the sets $\mathcal{N}$ and

$$\Lambda := \{\lambda_j \ : \ j \in \mathcal{N}\}$$

for the reader's convenience, but note that the weight set Λ will in all cases determine the inner product structure on F. Rewritten in power series form, the functions in F are

$$f(z) = \sum_{n\in\mathcal{N}} a_n z^n \text{ with } \sum_{n\in\mathcal{N}} \frac{n!^2}{\lambda_n} |a_n|^2 < \infty$$

with $a_n = f^{(n)}(0)/n!$. The inequality above is equivalent to (7), and together with (5) it implies that (4) is absolutely convergent in D_R for finite R and in $\mathbb{C}$ for $R = \infty$. In fact, in D_R we have

$$|f(z)|^2 \;\leq\; \left(\sum_{n\in\mathcal{N}} \frac{|a_n|^2 (n!)^2}{\lambda_n}\right) \cdot \left(\sum_{n\in\mathcal{N}} \frac{R^{2n}\lambda_n}{(n!)^2}\right) < \infty \tag{8}$$

by the Cauchy–Schwarz inequality.

Theorem 2.1. *If $\mathcal{N}$ is finite, the Taylor space F with respect to $\mathcal{N}$ consists of the span of the monomials x^j for $j \in \mathcal{N}$. If $\mathcal{N}$ is infinite, and if weights λ_j are chosen with* (5)*, the Taylor space F consists of real–analytic functions with power series expansions around zero which are absolutely convergent in D_R. The space will depend on Λ, consist of all functions f of the form* (4) *with* (7) *and carry the inner product* (6)*.* $\square$

We still have to turn (4) into a Hilbert space reproduction formula

$$f(x) = (f, K(x, \cdot)) \text{ for all } x \in D_R,\ f \in F \tag{9}$$

with a suitable positive (semi–) definite reproducing kernel

$$K\ :\ D_R \times D_R \to \mathbb{C}.$$

As is well–known, such a kernel must exist and is uniquely defined.

Theorem 2.2. *The kernel*

$$K(x,t) := \sum_{j\in\mathcal{N}} \lambda_j \frac{(t\overline{x})^j}{(j!)^2} =: \kappa(t\overline{x}),\ x,\, t \in I \tag{10}$$

is well–defined due to (5) *and the series is absolutely convergent for infinite $\mathcal{N}$ and for all $x, t \in D_R$. It is positive (semi–) definite and reproducing in the Taylor space, and the Taylor formula coincides with the reproduction formula* (1)*.*

Proof: The first sentence follows from arguments we have used before. Now we evaluate

$$\begin{aligned} \frac{\partial^j}{\partial t^j}_{|t=0} K(x,t) &= \lambda_j \frac{\overline{x}^j}{j!} \quad \text{and} \\ (f, K(x,\cdot)) &= \sum_{j\in\mathcal{N}} \frac{f^{(j)}(0)}{\lambda_j} \frac{\lambda_j x^j}{j!} = f(x). \quad \square \end{aligned}$$

This yields plenty of examples of kernels when starting from power series

$$\kappa(z) = \sum_{j=0}^{\infty} \lambda_j \frac{z^j}{(j!)^2}$$

which represent complex–analytic functions in a neighborhood of the origin. The connection to kernels $K(x,t)$ is via $K(x,t) = \kappa(t\overline{x})$. It allows selections of λ_j which grow like $R^j (j!)^2$ for $j \to \infty$ and arbitrary $R > 0$. Consequently, there are surprisingly many spaces which allow the Taylor formula for reproduction. Table 1 gives some examples that we partially cover later in more detail. Many of these are special instances of hypergeometric functions, i.e.

$$
\begin{aligned}
{}_2F_1(a,b;c,z) &= \sum_{n=0}^{\infty} \frac{(a)_n(b)_n}{(c)_n}\frac{z^n}{n!} \\
{}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q,z) &= \sum_{n=0}^{\infty} \frac{(a_1)_n\cdots(a_p)_n}{(b_1)_n\cdots(b_q)_n}\frac{z^n}{n!}
\end{aligned}
$$

with the Pochhammer symbol

$$(a)_n := a(a+1)\cdots(a+n-1).$$

$\kappa(z)=\sum_{j\in\mathcal{N}}\lambda_j\frac{z^j}{(j!)^2}$	$\mathcal{N}$	λ_j
$(1-z)^{-1}, 0\le\lvert z\rvert<1$	$\mathbb{N}$	$(j!)^2$
$(1-z^2)^{-1}, 0\le\lvert z\rvert<1$	$2\mathbb{N}$	$(j!)^2$
$(1-z)^{-\alpha}, \alpha\in\mathbb{N}, 0\le\lvert z\rvert<1$	$\mathbb{N}$	$\frac{(\alpha+j-1)!j!}{(\alpha-1)!}$
$(1-z^2)^{-\alpha}, \alpha\in\mathbb{N}, 0\le\lvert z\rvert<1$	$2\mathbb{N}$	$\frac{(\alpha+j-1)!j!}{(\alpha-1)!}$
$-\frac{\log(1-z)}{z}, 0\le\lvert z\rvert<1$	$\mathbb{N}$	$(-1)^{j+1}\frac{(j!)^2}{j+1}$
$\exp(z)$	$\mathbb{N}$	$j!$
$\exp(z)$	$\mathbb{N}$	$j!$
$\sinh(z)$	$2\mathbb{N}+1$	$j!$
$\sinh(z)/z$	$2\mathbb{N}$	$\frac{j!}{j+1}$
$\cosh(z)$	$2\mathbb{N}$	$j!$
$z^{-\alpha}I_\alpha(z)$	$2\mathbb{N}$	$\frac{j!}{2^{j+\alpha}\Gamma(j+\alpha+1)}$

Table 1: Some Kernels for Taylor Spaces

Other and much more general cases for **multivariate** power kernels are covered by [15]. In complex analysis, special cases of Taylor kernels and their associated native spaces are well studied, see e.g., [11, 9, 3, 2].

3 Interpolation

In what follows, we shall consider interpolation at finite sets

$$X := \{x_0,\ldots,x_n\}\subset D_R$$

of pairwise different nodes by functions from the space (3). Interpolation of *arbitrary* data on X requires nonsingularity of the *kernel matrix*

$$A_X := (K(x_j, x_k))_{0\le j,k\le n}.$$

It is hermitian and positive semidefinite because it can be written as a Gramian, but it is positive definite only if additional conditions are satisfied.

Before we focus on such conditions, we consider the case that the interpolated data on X come from a function f in the Taylor space F for K. For these special data, the system is always solvable. In fact, the Hilbert space projector Π_X from F to F_X yields a function $\Pi_X(f) = s_{f,X} \in F_X$ such that $f - \Pi_X(f)$ is orthogonal to F_X. By (1) this means that $s_{f,X}$ interpolates f in X.

For checking positive definiteness of the kernel matrix, let $a \in \mathbb{C}^{n+1}$ be a vector with

$$\begin{aligned} 0 &= a^T A_X \overline{a} \\ &= \sum_{m\in\mathcal{N}} \frac{\lambda_m}{(m!)^2} \left| \sum_{j=0}^{n} a_j \overline{x_j}^m \right|^2. \end{aligned}$$

Positive definiteness of the kernel matrix is guaranteed, if

$$0 = \sum_{j=0}^{n} a_j \overline{x_j}^m \text{ for all } m \in \mathcal{N}$$

implies $a = 0$, i.e. iff the $|\mathcal{N}| \times |X|$ Vandermonde matrix built from the $n+1 = |X|$ points in X and the $|\mathcal{N}|$ exponents in $\mathcal{N}$ has rank $n+1$. In case $\mathcal{N} = \{0, 1, \dots, N\}$ with $N \ge n$, the matrix is positive definite, but for other cases of finite $\mathcal{N}$, the above nondegeneracy condition is all we can say here. However, for $\mathcal{N} = \mathbb{N}$ we have positive definiteness for all point sets. The same holds for $\mathcal{N} = 2\mathbb{N}$ if symmetry under squaring is avoided, i.e. if all $|x_j|^2$ are different.

4 Properties of Interpolants

(secProp) The existence of interpolants being settled to some extent, we can apply known results from the theory of scattered data interpolation [14] to interpolation in Taylor spaces.

The first type of results concerns optimality. The interpolant $s_{f,X} \in F_X$ to some function $f \in F$ on a discrete set X satisfies the *norm–minimality property*

$$\|s_{f,X}\|_F \le \|g\|_F \text{ for all } g \in F \text{ with } g_{|X} = f_{|X}$$

and in particular

$$\|s_{f,X}\|_F \le \|f\|_F.$$

We now focus on error bounds and another optimality criterion. Consider all linear functionals of the form

$$\mu_{a,x} := f \mapsto f(x) - \sum_{j=0}^{n} a_j f(x_j)$$

with arbitrary vectors $a \in \mathbb{C}^{n+1}$ and points $x \in D_R$. It is well–known in the context of reproducing kernel Hilbert spaces that the solution of

$$\min_{a \in \mathbb{C}^{n+1}} \|\mu_{a,x}\|_{F^*} =: P_X(x) \tag{11}$$

exists and is attained at a vector $a = u(x) \in \mathbb{C}^{n+1}$ which solves the system

$$K(x, x_k) = \sum_{j=0}^{n} u_j(x) K(x_j, x_k),\ 0 \le k \le n$$

which is solvable because of $K(x, \cdot) \in F$. The solution is unique and satisfies Lagrange conditions $u_j(x_k) = \delta_{jk}$ if the system is nonsingular, but in general we know only that the solution may be nonunique. But we now can define the function

$$L_{f,X}(x) := \sum_{j=0}^{n} f(x_j) u_j(x) \text{ for all } x \in D_R$$

and by construction we have the standard error bound

$$|f(x) - L_{f,X}(x)| \le P_X(x) \|f\|_F \text{ for all } f \in F,\ x \in D_R \tag{12}$$

with the *power function* $P_X(x)$ defined in (11). This function can be explicitly calculated, since we have

$$\begin{aligned} P_X(x)^2 &= \|\mu_{u(x),x}\|^2_{F^*} \\ &= K(x,x) - \sum_{j=0} u_j(x) K(x, x_j) - \sum_{j=0} \overline{u_j(x)} K(x_j, x) \\ &\quad + \sum_{j,k=0}^{n} u_j(x) \overline{u_k(x)} K(x_j, x_k). \end{aligned}$$

Note that P_X is uniquely defined even if the interpolation problem is not uniquely solvable.

Since (11) implies that the optimal functional $\mu_{u(x),x}$ must be orthogonal to all functionals δ_{x_j}, we know that $L_{f,X}$ is an interpolant to f on X which will coincide with $s_{f,X}$ if the system is nonsingular. In any case, the error bound (12) is useful. For instance. one can generate new interpolation points by maximizing P_X [7], or one can assess the quality of polynomial interpolation on point sets whose cardinality is much lower than the degree. We shall provide an example in the final section.

If we are interested in asymptotic results for $n \to \infty$, we have to confine ourselves to the case of infinite $\mathcal{N}$. Furthermore, we restrict ourselves now to interpolation and evaluation on intervals $I := [-a, a]$ with $a \leq R$ and consider only functions f from F with real coefficients in their power series, and we denote the resulting space by $F_{\mathcal{R}}$. If we interpolate functions f from $F_{\mathcal{R}}$, we shall make use of the analyticity of f and apply results of [15] after some modifications. These error bounds will in all cases come out to be of spectral order in terms of the *fill distance*

$$h := \sup_{x \in I} \min_{x_j \in X} |x_j - x|$$

of X in I. Since we shall later work with univariate functions on bounded intervals only, we can assume $h \approx 1/n$ in case of n quasi–uniformly distributed data points, i.e. those with *bounded mesh ratio* in the sense used in spline theory.

5 Examples

5.1 Hardy Space

We consider the case $\kappa(z) = (1-z)^{-1}$ on the open unit disc D_1. This yields the *Szegö kernel*

$$S(\omega, z) := \frac{1}{1 - z\overline{\omega}} \text{ for all } \omega,\ z \in D_1.$$

The native Hilbert space for this kernel is the well–known *Hardy space* $\mathbf{H}^2$ which consists of those functions analytic in the unit disc whose Taylor coefficients form a square-summable series. It is known (see, e.g., [3, 2]) that the norm can also be realized as an integral

$$\|f\|^2_{\mathbf{H}^2} = \frac{1}{2\pi} \int_0^{2\pi} \left| f\left(e^{i\theta}\right) \right|^2 d\theta\ .$$

The reproducing formula then becomes the Cauchy formula. Thus the two most interesting reproduction formulas agree in this case and take the form of Hilbert space reproduction formulas.

When working in an interval $I = [-a, a]$ with $0 < a < 1$ and real-valued functions, we get

Proposition 5.1. *The native Hilbert space $F_{\mathcal{R}}$ for the rational kernel $R(x,t) = (1-xt)^{-1}$ consists of real-valued functions whose complex extensions lie in the Hardy space $\mathbf{H}^2$.*

Interpolation in point sets $X = \{x_0, \ldots, x_n\} \subset I = [-a, a]$ with $0 < a < 1$ using this kernel and the trial space (3) will be in terms of rational functions

p/q with polynomials p and q of degree up to n and $n+1$, respectively, the denominator polynomial q being fixed up to a multiplicative constant by its zeros in the points $1/x_j,\ 0 \le j \le n$. These classical rational interpolants are easy to calculate via polynomial interpolation of degree n.

For interpolation in the corresponding native Hilbert space, we obtain the following convergence results from [15]:

Theorem 5.2. *1. For each $0 < a < 1$ there are constants $c_1,\ h_0 > 0$ such that for any discrete set $X \subset I = [-a, a]$ with fill distance $h \le h_0$ and any function $f \in F_{\mathcal{R}}$, the error between f and its interpolant $s_{f,X}$ is bounded by*

$$\|f - s_{f,X}\|_{L_\infty[-a,a]} \le e^{-c_1/h} \|f\|_{\mathcal{N}_R} \ .$$

2. Suppose $n \in \mathbb{N}$ and $0 < a < 1$. Then there are constants $c_2,\ \tilde{h}_0 > 0$ such that for all discrete sets $X \subset I = [-a, a]$ with fill distance $h \le \tilde{h}_0$ and any function $f \in \mathcal{N}_R$, the error between the n-th derivative of f and its interpolant $s_{f,X}$ is bounded by

$$\left\| f^{(n)} - s_{f,X}^{(n)} \right\|_{L_\infty[-a,a]} \le e^{-c_2/\sqrt{h}} \|f\|_{\mathcal{N}_R} \ .$$

Note that for quasi–uniform data sites we get spectral or exponential convergence with respect to n.

5.2 Bergman Space

Similarly we can proceed for the *Bergman kernel*

$$B_{\mathcal{C}}(w, z) = \frac{1}{(1 - z\overline{w})^2}$$

on the unit disc. The native Hilbert space for this kernel is the Bergman space $\mathbf{B}^2$, which consists of those holomorphic functions on the unit disc that are square-summable with respect to the planar Lebesgue measure m [2]. The norm can be realized as

$$\|f\|_{\mathbf{B}^2} = \frac{1}{\pi} \int_{\mathbb{D}} |f(z)|^2 \, dm(z) \ .$$

Note that the Bergman space contains the Hardy space.

When restricting everything to the real line and real–valued functions, we get

Proposition 5.3. *The native Hilbert space $F_{\mathcal{R}}$ for the rational kernel $B(x,t) := (1 - xt)^{-2}$ on intervals $I := [-a, a]$ with $0 < a < 1$ consists of real–valued functions analytic in I whose complex extensions lie in the Bergman space $\mathbf{B}^2$.*

Now the kernel–based interpolation problem on $n+1$ points uses rational functions p/q where q has the points $1/x_j$ as double zeros and is of degree $2n+2$, while p is of degree at most $2n$. For interpolation in $F_{\mathcal{R}}$ for the Bergman kernel we obtain the following convergence results.

Theorem 5.4. *1. For all $0 < a < 1$ there are constants $c_1,\ h_0 > 0$ such that for any discrete set $X \subset [-a,a]$ with fill distance $h \le h_0$ and any function $f \in F_{\mathcal{R}}$, the error between f and its interpolant $s_{f,X}$ is bounded by*

$$\|f - s_{f,X}\|_{L_\infty[-a,a]} \le e^{-c_1/h} \|f\|_{\mathcal{N}_B} \ .$$

2. Suppose $n \in \mathbb{N}$ and $0 < a < 1$. Then there are constants $c_2,\ \tilde{h}_0 > 0$ such that for all discrete sets $X \subset [-a,a]$ with fill distance $h \le \tilde{h}_0$ and any function $f \in \mathcal{N}_R$, the error between the n-th derivative of f and its interpolant $s_{f,X}$ is bounded by

$$\left\| f^{(n)} - s_{f,X}^{(n)} \right\|_{L_\infty(-a,a)} \le e^{-c_2/\sqrt{h}} \|f\|_{\mathcal{N}_B} \ .$$

Proof:

1. By [15, Thm. 3], it suffices to check that

$$\tilde{C}^{(2k)} := \sup_{x,y\in[-a,a]} \left| D_y^{2k} \frac{1}{(1-xy)^2} \right| \le c^k k!^2$$

holds for almost all $k \in \mathbb{N}$ with some constant c independent of k, where D_y^ℓ denotes the ℓ-th derivative with respect to the variable y. Indeed, we find inductively

$$\begin{aligned} \tilde{C}^{(2k)} &:= \sup_{x,y\in[-a,a]} \left| D_y^{2k} \frac{1}{(1-xy)^2} \right| = \sup_{x,y\in[-a,a]} \left| \frac{(2k+1)!x^{2k}}{(1-xy)^{2k+2}} \right| \\ &\le \frac{(2k+1)!a^{2k}}{(1-a^2)^{2k+2}} \le \left(\frac{6a^2}{(1-a^2)^4} \right)^k k!^2 \ . \end{aligned}$$

2. By [15, Thm. 6], it suffices to check that

$$C^{(2k)} := \max_{\ell+m=2k} \sup_{x,y\in(-a,a)} \left| D_x^m D_y^\ell K(x,y) \right| \le e^{ck} k^{2k}$$

holds for some constant c independent of k. By symmetry, we may assume $m \le \ell$. Explicit calculations yield

$$\begin{aligned} C^{(2k)} &= \max_{\ell+m=2k} \sup_{x,y\in[-a,a]} \left| D_x^m D_y^\ell \sum_{n=0}^{\infty} (n+1)\, x^n y^n \right| \\ &= \max_{\ell+m=2k} \sup_{x,y\in[-a,a]} \left| \sum_{n=\ell}^{\infty} \frac{(n+1)!n!}{(n-\ell)!\,(n-m)!} x^{n-m} y^{n-\ell} \right| \\ &= a^{-2k} \max_{\ell+m=2k} \sum_{n=\ell}^{\infty} \frac{(n+1)!n!}{(n-\ell)!\,(n-m)!} a^{2n} \ . \end{aligned}$$

We claim that the maximum is attained for $\ell = m = k$. Indeed, if $\ell \ge m+2$, we find the term-wise bound

$$\begin{aligned} \sum_{n=\ell}^{\infty} \frac{(n+1)!n!}{(n-\ell)!\,(n-m)!} a^{2n} &\le \sum_{n=\ell}^{\infty} \frac{(n+1)!n!}{(n-\ell+1)!\,(n-m-1)!} a^{2n} \\ &\le \sum_{n=\ell-1}^{\infty} \frac{(n+1)!n!}{(n-(\ell-1))!\,(n-(m+1))!} a^{2n} \ , \end{aligned}$$

which implies that the maximum is attained for the symmetric situation $m = \ell = k$. Thus we have with the hypergeometric function (see [1, ch. 15])

$$\begin{aligned} C^{(2k)} &= a^{-2k} \sum_{n=k}^{\infty} \frac{(n+1)!n!}{(n-k)!^2} a^{2n} = \sum_{n=0}^{\infty} \frac{(n+k+1)!\,(n+k)!}{n!^2} a^{2n} \\ &= k!\,(k+1)! F\left(k+2, k+1; 1; a^2\right) \\ &= k!\,(k+1)! \left(1-a^2\right)^{-2-k} F\left(2+k, -k; 1; \frac{a^2}{a^2-1}\right) \quad [1, (15.3.4)] \\ &= k!\,(k+1)! \left(1-a^2\right)^{-k-1} F\left(-k-1, k+1; 1; \frac{a^2}{a^2-1}\right) \quad [1, (15.3.3)] \ . \end{aligned}$$

We apply [1, (15.4.6)] with $n = k+1$, $\alpha = 0$ and $\beta = -1$, which gives

$$C^{(2k)} = (k+1)!k! \left(1-a^2\right)^{-k-1} P_{k+1}^{(0,-1)} \left(\frac{a^2+1}{1-a^2} \right) \ .$$

By the recurrence relation [1, (22.7.16)] for the Jacobi polynomials, we have

$$(k+1)\, P_{k+1}^{(0,-1)}(x) = \left(k + \frac{1}{2}\right) (1+x)\, P_k^{(0,0)}(x) - k P_k^{(0,-1)}(x) \ .$$

Thus, we have with the Legendre polynomials $P_k := P_k^{(0,0)}$

$$\begin{aligned} \left| P_{k+1}^{(0,-1)}(x) \right| &\le |1+x|\, |P_k(x)| + \left| P_k^{(0,-1)}(x) \right| \le \dots \\ &\le |1+x| \left(|P_k(x)| + |P_{k-1}(x)| + \dots + |P_0(x)| \right) + \left| P_0^{(0,-1)}(x) \right| \ . \end{aligned}$$

In [15, Proof of Lemma 5] it is shown that

$$|P_k(x)| \le (2|x|+1)^n \ .$$

Hence,

$$\begin{aligned}
\left|P_{k+1}^{(0,-1)}(x)\right| &\le |1+x| \sum_{n=0}^{k} (2|x|+1)^n + 1 \\
&\le k|1+x|(2|x|+1)^k \le c^k \ ,
\end{aligned}$$

where the constant c depends only on $x = -\frac{a^2+1}{a^2-1}$ but not on k. Putting things together, we have

$$\begin{aligned}
C^{(2k)} &= (k+1)!k!\left(1-a^2\right)^{-k-1} P_{k+1}^{(0,-1)}\left(-\frac{a^2+1}{a^2-1}\right) \\
&\le c^k k!\,(k+1)! \le c^k k^{2k}
\end{aligned}$$

with some constant c independent of k.

□

5.3 Dirichlet Space

There are also Taylor kernels of logarithmic type. A well known one is the kernel $L(x,y) := -\frac{1}{xy}\log(1-xy)$, which can be extended via

$$L(w,z) := -\frac{1}{z\overline{w}}\log(1-z\overline{w}) \ .$$

This kernel is the reproducing kernel for the Dirichlet space $\mathcal{D}$ [2], which consists of those holomorphic functions f on the unit disc $\mathbb{D}$ whose derivative f' is in the Bergman space.

Proposition 5.5. *The native Hilbert space for the logarithmic kernel L consists of real–valued functions analytic in $[-a,a]$ whose complex extensions lie in the Dirichlet space $\mathcal{D}$.*

The interpolation on $X := \{x_0,\ldots,x_n\} \subset I = [-a,a]$ for $0 < a < 1$ is now carried out with linear combinations of functions

$$\begin{array}{ccc}
-\frac{1}{t}\log(1-tx_j) & \text{for} & x_j \ne 0, \\
1 & \text{for} & x_j = 0,
\end{array}$$

which is quite an unusual setting.

The approximation orders due to [15] are the same as for the rational case. Thus Theorem 5.2 can be reformulated exactly also for this kernel and functions from Dirichlet space.

5.4 The exponential case

In this section we consider a kernel of exponential type on $\mathbb{R}$, namely

$$E(x,t) := \exp(xt)$$

which arises when the entire function $\kappa(z) = \exp(z)$ is restricted to the real line. The native Hilbert space becomes

$$N_E := \left\{ f \;:\; \mathbb{R} \to \mathbb{R} \;\middle|\; f(z) = \sum_{n=0}^{\infty} a_n z^n \text{ with } a_n \in \mathbb{R},\ \sum_{n=0}^{\infty} n! a_n^2 < \infty \right\}$$

and consists of real–analytic functions with entire complex extensions. If an interval $I := [-a, a]$ is fixed, and if point sets $X := \{x_0, \ldots, x_n\} \subset I$ are used for interpolation, we have an interpolation with classical exponential sums, the x_j being fixed frequencies. We shall remove this coupling between frequencies and data points later.

To characterize functions from N_E more precisely, we recall some basic definitions from [5, Chapters 1 and 2].

Definition 5.6. *For an entire function f of a complex variable z, we denote by $M(r)$ the maximum modulus of $f(z)$ for $|z| = r < \infty$. We say that f is of order ρ if*

$$\limsup_{r\to\infty} \frac{\log\log M(r)}{\log r} = \rho\ .$$

By convention, a constant function has order 0.

Theorem 5.7.

1. *The native Hilbert space N_E of the exponential kernel consists of real-valued analytic functions that have entire complex extensions. In particular, it contains all polynomials.*
2. *If an entire function $f(z) = \sum_{n=0}^{\infty} a_n z^n$ with real coefficients a_n is of order $\rho < 2$, then its restriction to $\mathbb{R}$ lies in N_E.*
3. *For any function from N_E, the complex extension is of order less than or equal to* 2.

Proof:

1. Assume $f(x) = \sum_{n=0}^{\infty} a_n x^n \in N_E$. Then the natural complex extension $\tilde{f}(z) = \sum_{n=0}^{\infty} a_n z^n$ converges in the whole complex plane since

$$\overline{\lim_{n\to\infty}} \sqrt[n]{|a_n|} \le \left(\overline{\lim_{n\to\infty}} \sqrt[n]{|a_n|^2\, n!} \cdot \overline{\lim_{n\to\infty}} \sqrt[n]{\frac{1}{n!}} \right)^{1/2} = 0\ .$$

2. If an entire function $f(z) = \sum_{n=0}^{\infty} a_n z^n$ is of order $\rho = 2 - 2\epsilon$ with $1 > \epsilon > 0$, then by [5, Thm. 2.2.2],

$$\overline{\lim_{n\to\infty}} \frac{n \log n}{\log(1/|a_n|)} = 2 - 2\epsilon \ ,$$

where the expression on the right is to be taken as 0 if $a_n = 0$. Since $a_n^2 n!$ converges to 0, there is an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ we have $a_n \leq 1$. Then by definition of the lim sup, there exists an $N \in \mathbb{N}$ such that for all $n \geq N$,

$$n \log n \leq (2 - \epsilon) \log\left(\frac{1}{|a_n|}\right) \ .$$

This implies $|a_n| \leq n^{-\frac{n}{2-\epsilon}}$, and therefore

$$\sum_{n=0}^{\infty} n! a_n^2 \ \leq \ C + \sum_{n=N}^{\infty} n! n^{-\frac{2n}{2-\epsilon}} \leq C + \sum_{n=N}^{\infty} n^{-\frac{2\epsilon}{2-\epsilon}} < \infty \ .$$

If f is of order 0,

$$\overline{\lim_{n\to\infty}} \frac{n \log n}{\log(1/|a_n|)} = 0 \ ,$$

which yields that there is some $N \in \mathbb{N}$ such that for all $n \geq N$,

$$n \log n \leq \log\left(\frac{1}{|a_n|}\right) \ ,$$

and we can proceed as above, choosing $\epsilon = 1$.

3. Assume $f(x) = \sum_{n=0}^{\infty} a_n x^n \in N_E$. Then there is a constant $C > 0$, such that $a_n^2 \leq C/n!$, i.e., by Stirling's formula

$$|a_n| \leq \frac{C}{\sqrt{n!}} \leq \frac{C e^{n/2}}{n^{\frac{2n+1}{4}}} \ .$$

Thus,

$$\overline{\lim_{n\to\infty}} \frac{n \log n}{\log\left(\frac{1}{|a_n|}\right)} \leq \overline{\lim_{n\to\infty}} \frac{n \log n}{\frac{n}{2}\log n + \frac{1}{4}\log n - \log C - \frac{n}{2}} = 2 \ .$$

□

Remark 5.8. *Theorem 5.7 is sharp in the sense that there are functions of order $\rho = 2$ that lie in N_E and some that do not. For instance $f_1(x) = \sum_{n=0}^{\infty} \frac{1}{n^{n/2}} x^n$ and $f_2(x) = \sum_{n=0}^{\infty} \frac{1}{n^{1+n/2}} x^n$ both are of order 2, and $f_1 \in N_E$ while $f_2 \notin N_E$.*

Following [15, Cor. 1 and 2], we have the following approximation orders for classical interpolation in N_E.

Theorem 5.9. *1. For all $a > 0$, there are constants c_1, $h_0 > 0$ such that for any discrete set $X \subset I = [-a, a]$ with fill distance $h \le h_0$ and any function $f \in N_E$, the error between f and its interpolant $s_{f,X}$ is bounded by*

$$\|f - s_{f,X}\|_{L_\infty[-a,a]} \le e^{c_1 \log h/h} \|f\|_{N_E} .$$

2. Suppose $n \in \mathbb{N}$ andf $a > 0$. Then there are constants c_2, $\tilde{h}_0 > 0$ such that for all discrete sets $X \subset I = [-a, a]$ with fill distance $h \le \tilde{h}_0$ and any function $f \in N_E$ the error between the n-th derivative of f and its interpolant $s_{f,X}$ is bounded by

$$\left\| f^{(n)} - s_{f,X}^{(n)} \right\|_{L_\infty[-a,a]} \le e^{c_2 \log h/\sqrt{h}} \|f\|_{N_E} .$$

Due to some simple calculational tricks, similar results on spectral convergence hold when frequencies are somewhat decoupled from data points. We assume that frequencies μ_j are chosen with fill distance h in some interval $[\alpha, \beta]$, and we want to work with these frequencies, but evaluate the approximation error on $[-a, a]$. We then use the points

$$x_j := \varphi(\mu_j) := a + \frac{\mu_j - \alpha}{\beta - \alpha}(b - a) \in [-a, a]$$

for interpolation in $[-a, a]$ using scaled trial functions $\exp(ctx_j)$ as functions of $t \in [-a, a]$. Since the above error bounds will also hold for scaled kernels, we can take $c = \frac{\beta - \alpha}{b - a}$ to see that we have worked with functions

$$\begin{aligned} \exp(ctx_j) &= \exp\left(ct\left(a + \tfrac{\mu_j - \alpha}{\beta - \alpha}(b - a)\right)\right) \\ &= \exp\left(ct\left(a - \tfrac{\alpha}{\beta - \alpha}(b - a)\right)\right) \exp\left(ct\left(\tfrac{\mu_j}{\beta - \alpha}(b - a)\right)\right) \\ &= \exp\left(t(ca - \alpha)\right) \exp\left(t\mu_j\right) \end{aligned}$$

which now have the desired frequencies. The given function, however, has to be multiplied by $\exp\left(-t(ca - \alpha)\right)$ before computations start.

6 Numerical Examples

All the subsequent figures use four different interpolants:

- the solid line for the Szegö kernel,
- the dotted line for the exponential kernel,

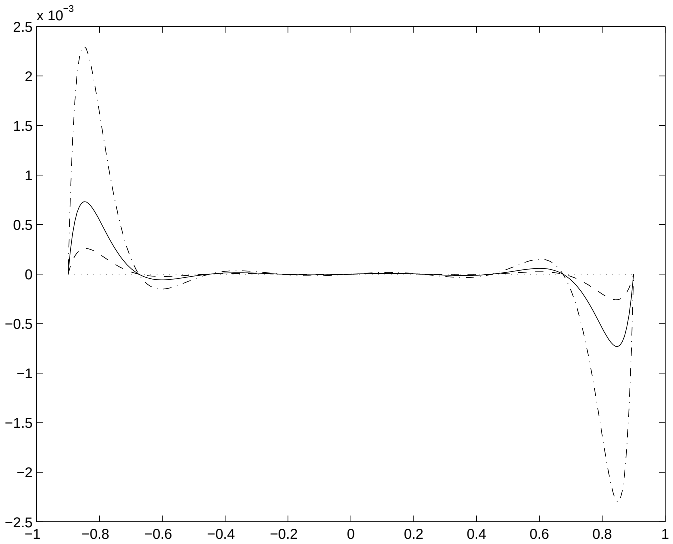

Figure 1: Errors for $f(x) = \sin(x)$

- the - - line for the log kernel, and
- the - . line for the squared Szegö kernel.

They differ in the function f providing the data. In all cases, the arising kernel matrices are severely ill–conditioned, but we did not apply any preconditioning techniques. The cases with expansions into purely even or purely odd terms are ignored and will be similar, provided that the symmetries are taken into account when setting up the problems. To allow comparisons between kernels, we fixed 9 equidistant data locations on $[-0.9, +0.9]$ in all cases. Since the scalings of the figures might be hard to read, we present the L_∞ errors in Table 2.

	Szegö	exp	log	Szegö^2
$\sin(x)$	7.31e–04	6.53e–08	2.59e–04	2.30e–03
$1/(1+x^2/25)$	8.73e–06	8.76e–10	2.97e–06	2.83e–05
$1/(1+x^2)$	1.49e–02	1.11e–03	9.67e–03	2.48e–02
$1/(1+25x^2)$	3.89e+00	9.07e–01	2.97e+00	5.38e+00
B–spline	1.23e+00	2.88e–01	9.36e–01	1.70e+00

Table 2: L_∞ errors for different examples and kernels

It is to be expected that the location of singularities of f outside the interval will be of quite some influence on the error. For the entire function $\sin(x)$, one can expect the best possible behavior, and the results are in Figure 1. We then have three examples of the Runge type, the singularities moving closer from $\pm 5i$ via $\pm i$ to $\pm 0.2i$, see Figures 2 to 4. The results get

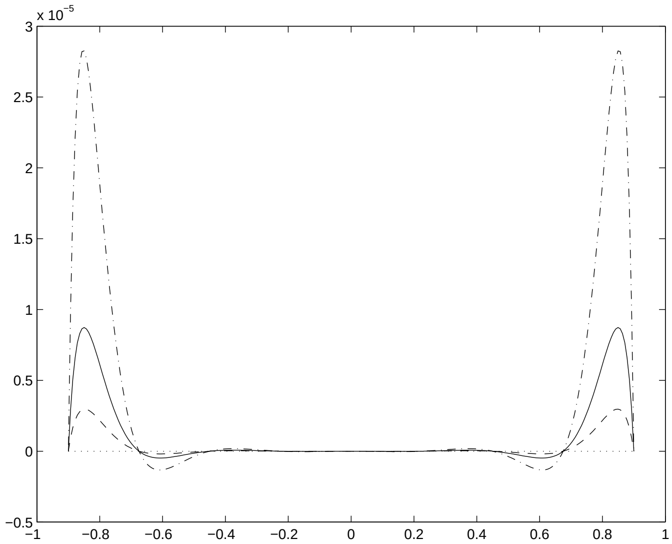

Figure 2: Errors for $f(x) = 1/(1 + x^2/25)$

dramatically worse, and for the B–spline $f(x) = 1 - |x|$ they are disastrous, but this function is too far away from any of the native Hilbert spaces of analytic functions.

In all cases, the exponential kernel performed best, followed by the log kernel, the Szegö kernel, and the squared Szegö kernel.

Acknowledgement

Special thanks go to Christian Rieger for many fruitful discussions. BZ was partly funded by a postdoctoral fellowship of the NSF under Grant No. DMS-0905778.

References

[1] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions.* Dover, New York, 1970.

[2] J. Agler and J.E. McCarthy. *Pick Interpolation and Hilbert Function Spaces.* Graduate studies in Mathematics, volume 44. American Mathematical Society, Providence, Rhode Island, 2002.

[3] D. Alpay. *The Schur algorithm, reproducing kernel spaces and system theory.* SMF/AMS Texts and Monographs, volume 5. American Mathematical Society, USA, 2001.

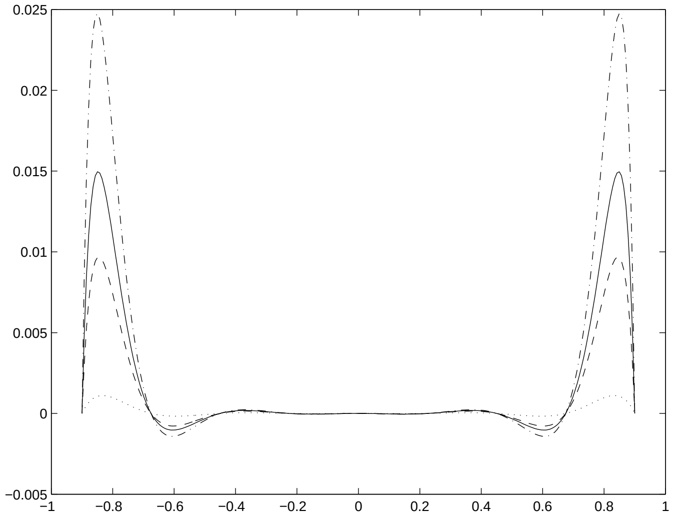

Figure 3: Errors for $f(x) = 1/(1 + x^2)$

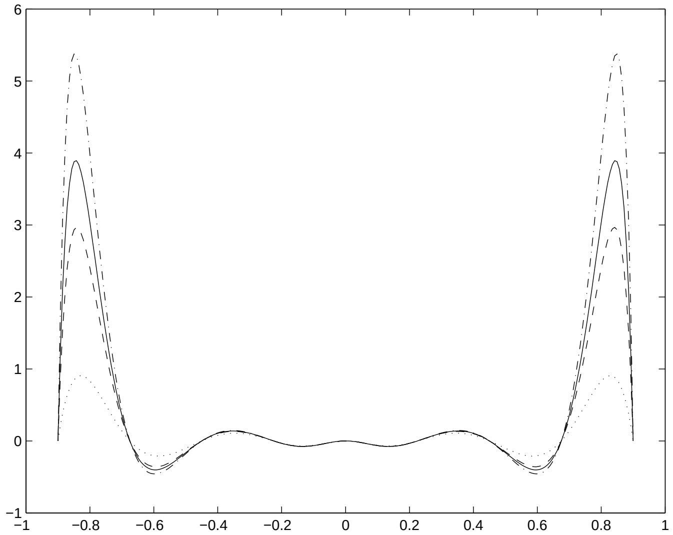

Figure 4: Errors for $f(x) = 1/(1 + 25x^2)$

[4] N. Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68:337–404, 1950.

[5] J.R.P. Boas. *Entire Functions.* A series of monographs and textbooks. Academic Press Inc., New York, N.Y., 1954.

[6] A.P. Calderón and A. Zygmund. Local properties of solutions of elliptic partial differential equations. *Studia Math.*, 20:171–225, 1961.

[7] Stefano De Marchi, R. Schaback, and H. Wendland. Near-optimal data-independent point locations for radial basis function interpolation. *Adv. Comput. Math.*, 23(3):317–330, 2005.

[8] J. Dick. A Taylor space for multivariate integration. *Monte Carlo Methods and Applications*, 12:99–112, 2006.

[9] Brian C. Hall. Holomorphic methods in analysis and mathematical physics. In *First summer school in analysis and mathematical physics*, Contemporary Mathematics, pages 1–59. American Mathematical Society, 1998.

[10] H. Meschkowski. *Hilbertsche Räume mit Kernfunktion.* Springer, Berlin, 1962.

[11] S. Saitoh. *Integral transforms, reproducing kernels and their applications.* Pitman Research Notes in Mathematics. Longman, Great Britain, 1997.

[12] R. Schaback. A unified theory of radial basis functions. Native Hilbert spaces for radial basis functions. II. *J. Comput. Appl. Math.*, 121(1-2):165–177, 2000. Numerical analysis in the 20th century, Vol. I, Approximation theory.

[13] R. Schaback and H. Wendland. Kernel techniques: from machine learning to meshless methods. *Acta Numerica*, 15:543–639, 2006.

[14] H. Wendland. *Scattered Data Approximation.* Cambridge University Press, 2005.

[15] B. Zwicknagl. Power series kernels. *Constructive Approximation*, 29(1):61–84, 2009.