Nonstandard Kernels and Their Applications

Robert Schaback*, Stefano De Marchi†

November 11, 2009

Abstract

Besides using standard radial basis functions, there are a few good reasons to look for kernels with special properties. This survey will provide several examples, starting from an introduction into kernel construction techniques. After providing the "missing" Wendland functions, we focus on kernels based on series expansions. These have some very interesting special cases, namely polynomial and periodic kernels, and "Taylor" kernels for which the reproduction formula coincides with the Taylor formula. Finally, we review the use of kernels as particular or fundamental solutions of PDEs, look at harmonic kernels and kernels generating divergence–free vector fields. Numerical examples will be provided as we are going along.

*Akademie der Wissenschaften zu Göttingen, Institut für Numerische und Angewandte Mathematik (NAM), Georg-August-Universität Göttingen, Lotzestrasse 16-18 D-37083 Göttingen

†Department of Pure and Applied Mathematics, University of Padua. Via Trieste, 63 I-35121 - Padova, Italy.

Contents

1 Kernel Basics

We start here with some notational background. Further details should be taken from the book [28] of H. Wendland. A *kernel*

$$K \ : \ \Omega \times \Omega \to \mathbb{R} \tag{1}$$

on a set $\Omega \subset \mathbb{R}^d$ is called *positive (semi-) definite* if for all finite point sets $X = \{x_1, \ldots, x_n\} \subset \Omega$ the associated *kernel matrix* $(K(x_j, x_k))_{1 \le j,k \le n}$ is *positive (semi-) definite*.

Reproducing Kernels in Hilbert Spaces $\mathcal{F}$ of functions on Ω are kernels K for which the *reproduction property*

$$(f, K(x, \cdot))_{\mathcal{F}} = f(x) \text{ for all } x \in \Omega, \ f \in \mathcal{F}. \tag{2}$$

holds. Each positive semidefinite kernel is reproducing in a unique "native" reproducing kernel Hilbert space (RKHS) associated to it, and each Hilbert space of functions has a reproducing kernel if the point evaluation functionals are continuous.

This survey focuses on some new nonstandard kernels that should get more attention, and thus we have to omit the classical kernel constructions summarized in [28] or in a somewhat more compact form in [24]. They comprise Whittle–Matérn-Sobolev kernels, polyharmonic functions, thin–plate splines, multiquadrics, Gaussians, and compactly supported kernels.

Unfortunately, space limitations force us to be very brief with certain recent interesting nonstandard constructions. We shall mention these only briefly and provide more room for the special ones we want to focus on.

For numerical analysis, the most important use of kernels is that they yield spaces of trial functions. Indeed, for each discrete set of points

$$X_n := \{x_1, \ldots, x_n\} \subset \Omega$$

the space

$$U_n := \operatorname{span} \{K(\cdot, x_j) \ : \ x_j \in X_n\} \tag{3}$$

spanned by translates of the kernel can serve for many purposes. Questions related to these spaces concern their approximation properties. In particular, one can interpolate data $f(x_k)$ of a function $f \in \mathcal{F}$ sampled at $x_k \in X_n$ by a trial function

$$s_{f,X_n} := \sum_{j=1}^{n} \alpha_j K(\cdot, x_j) \in U_n \tag{4}$$

solving the linear system

$$s_{f,X_n}(x_k) = \sum_{j=1}^{n} \alpha_j \underbrace{K(x_k, x_j)}_{\text{Kernel matrix}} = f(x_k) \text{ for all } x_k \in X_n \,. \tag{5}$$

2 Kernel–Based Error Bounds

Before we look at nonstandard kernels, we shall provide a nonstandard application of kernels. This concerns the use of kernels for obtaining error bounds for fairly general (in particular kernel–independent) interpolants. Consider a *quasi–interpolant* of the form

$$Q(f) := \sum_{j=1}^{n} f(x_j) u_j \tag{6}$$

to f on X_n using functions $u_1, \ldots, u_n$ on Ω. Note that interpolants take this form when the basis is rewritten in Lagrange form, satisfying $u_j(x_k) = \delta_{jk}$. Now consider the pointwise error functional

$$\epsilon_{Q,x}(f) := f(x) - Q(f)(x) = \left(\delta_x - \sum_{j=1}^{n} u_j(x) \delta_{x_j} \right) (f)$$

in a space $\mathcal{F}$ of functions on Ω with continuous point evaluation. Then

$$|f(x) - Q(f)(x)| = |\epsilon_{Q,x}(f)| \leq \|\epsilon_{Q,x}\|_{\mathcal{F}^*} \|f\|_{\mathcal{F}}$$

yields a bound that separates the influence of f from the influence of the quasi–interpolant. In a RKHS $\mathcal{F}$ with reproducing kernel K one has

$$(\delta_x, \delta_y)_{\mathcal{F}^*} = K(x, y) \text{ for all } x, y \in \Omega$$

and thus the norm of the error functional can be explicitly evaluated as

$$\begin{aligned} P^2_{Q,X_n,\mathcal{F}}(x) := \|\epsilon_{Q,x}\|^2_{\mathcal{F}} &= K(x,x) - 2 \sum_{j=1}^{n} u_j(x) K(x, x_j) \\ &\quad + \sum_{j=1}^{n} \sum_{k=1}^{n} u_j(x) u_k(x) K(x_k, x_j). \end{aligned} \tag{7}$$

This provides the error bound

$$|f(x) - Q(f)(x)| \leq P_{Q,X_n,\mathcal{F}}(x) \|f\|_{\mathcal{F}} \tag{8}$$

where the Q–dependent part is fully known. The function $P_{Q,X_n,\mathcal{F}}$ can be called the *Generalized Power Function*.

If one minimizes $P_{Q,X_n\mathcal{F}}(x)$ for fixed x over all quasi–interpolants Q of the form (6), it turns out that the minimum is attained when the u_j are the Lagrange basis of the kernel–based interpolant (4). This yields the standard *power function* [28] of kernel–based interpolation. We denote it by $P_{X_n,\mathcal{F}}$ and use it later.

In order to show how the above technique for error bounds works in practice, we add two examples.

EXAMPLE 1. The top left plot of Fig. 1 shows the generalized power function for linear interpolation on the triangle spanned by the three points $(0,0),(1,0),(0,1)$ in $\mathbb{R}^2$. The evaluation is done in the space $W_2^2(\mathbb{R}^2)$, using the kernel $K_1(r)r$ within the generalized power function, and where K_ν is the modified Bessel function of order ν. Note that the kernel–based interpolant using $K_1(r)r$ for interpolation must perform better, but is only slightly superior. The optimal power function is the top right plot, while the difference of the power functions is the lower plot.

EXAMPLE 2. A harder case is given in Fig. 2. In this case, we started from a grid, with spacing $h = 0.02$, in $[-1,1]^2$ and we considered only the points falling into the domain of the lower left plot, i.e. 5943 gridded points. Then, the Vandermonde matrix for interpolation by polynomials of degree 6 was formed, and a LU decomposition was calculated which selected 28 points by pivoting [21]. This results in a degree 6 polynomial interpolation method on the circled points of the lower left plot. This polynomial interpolant has an error bound of the above form with the generalized power function for $W_2^4(\mathbb{R})$ using the kernel $K_3(r)r^3$ given in the top left plot. The optimal kernel–based interpolation process for the same points in the same function space has the top right power function, while the difference of the power functions is in the lower right plot.

3 Compactly Supported Kernels

First we observe that most of the results in this session are not particularly new. In the univariate case, symmetric even–order B–splines are compactly supported positive definite kernels, because their Fourier transforms are even powers of *sinc* functions [14]. The multivariate analogues of these would be inverse Fourier transforms of even powers of Bessel functions as arising already in [12], but there are no explicit formulas available. Since 1995,

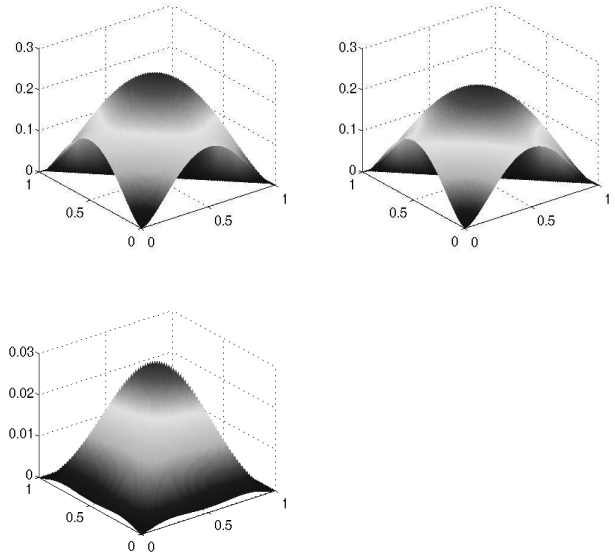

Figure 1: Error of affine-linear interpolation in W_2^2

however, there are two useful classes of compactly supported kernels due to Z.M. Wu [29] and H. Wendland [27]. The ones by Wendland, having certain advantages, we shall describe below. An addition to the zoo of compactly supported functions was given in 1998 by M.D. Buhmann [1], but we shall focus here on a very recent extension [20] to Wendland's class of functions.

We recall that Wendland's compactly supported kernels have the form

$$\Phi_{d,k}(r) = (1-r)_+^{\lfloor d/2 \rfloor + k + 1} \cdot p_{d,k}(r) \tag{9}$$

where $p_{d,k}$ is a polynomial of degree $\lfloor d/2 \rfloor + 3k + 1$ on $[0,1]$ and such that the kernel $K(x,y) = \Phi_{d,k}(\|x-y\|_2)$ is in C^{2k}, while $p_{d,k}(r)$ is of minimal degree for smoothness C^{2k}. Finally, the kernel is reproducing in a Hilbert space $\mathcal{F}$ norm–equivalent to $W_2^{d/2+k+1/2}(\mathbb{R}^d)$.

But if the dimension d is even, these functions do not generate integer–order Sobolev spaces. This calls for new compactly supported Wendland–type kernels that work for half–integer k, since it can be expected that they always generate a space equivalent to $W_2^{d/2+k+1/2}(\mathbb{R}^d)$. In general, the construction of Wendland 's functions proceeds via

$$\phi_{d,k}(r) = \psi_{\lfloor d/2 \rfloor + k + 1, k}(r) \tag{10}$$

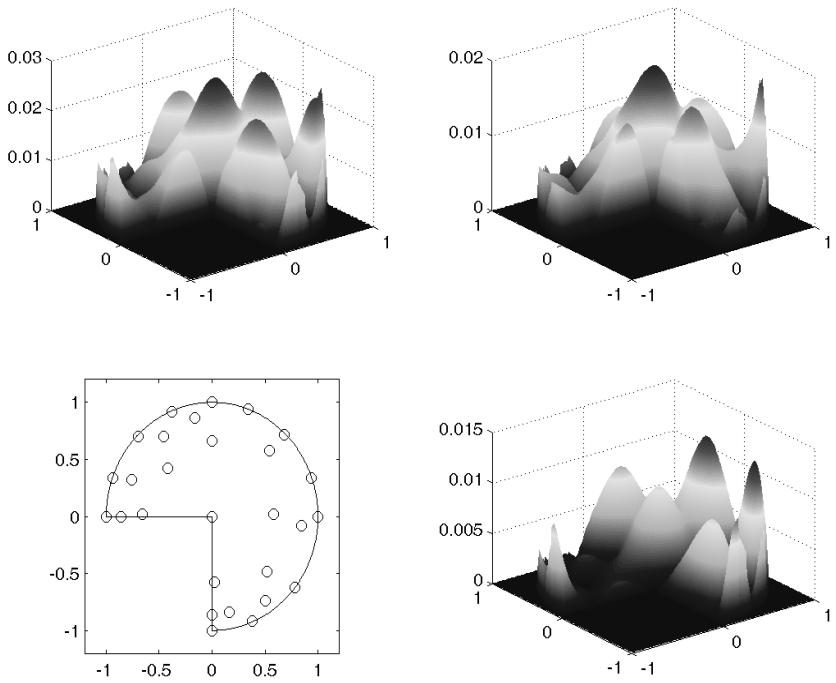

Figure 2: Polynomial interpolation, degree 6. Error in W_2^4

with

$$\psi_{\mu,k}(r) := \int_r^1 t(1-t)^\mu \frac{(t^2-r^2)^{k-1}}{\Gamma(k)2^{k-1}} dt,\ 0 \le r \le 1. \tag{11}$$

But it turns out that the above formula, used so far only for integers k in order to produce polynomials, can also be used for half–integers k. The self–explanatory MAPLE® code line

```
wend:=int(t*(1-t)^mu*(t*t-r*r)^(k-1)/(GAMMA(k)*2^(k-1)),t=r..1);
```

runs for all reasonable and fixed choices of μ and k where one half–integer is allowed, while it fails if both μ and k are genuine half–integers. A special case is

$$\phi_{2,1/2}(r) = \frac{\sqrt{2}}{3\sqrt{\pi}} \left(3r^2 \log\left(\frac{r}{1+\sqrt{1-r^2}} \right) + (2r^2+1)\sqrt{1-r^2} \right),\ 0 \le r \le 1 \tag{12}$$

plotted in Fig. 3 in the one dimensional case. It turns out that it generates a Hilbert space norm–equivalent to $W_2^2(\mathbb{R}^2)$, as expected. There are a few additional results proven in [20]:

- $\psi_{\mu,k}$ is positive definite on $\mathbb{R}^d$ for $\mu \geq \lfloor d/2 + k \rfloor + 1$;
- its d–variate Fourier transform for $\mu = \lfloor d/2 + k \rfloor + 1$ behaves like $\mathcal{O}\left(r^{-(d+2k+1)}\right)$ for $r \to \infty$;
- for $d = 2m$ and $k = n + 1/2$ the kernel $\psi_{\lfloor d/2 \rfloor + k + 1/2, k}$ generates $W_2^{m+n+1}(\mathbb{R}^{2m})$;
- the new functions have the general form

$$\psi_{2m,n-1/2}(r) = p_{m-1+n,n}(r^2) \log\left(\frac{r}{1+\sqrt{1-r^2}}\right) + q_{m-1+n,n}(r^2)\sqrt{1-r^2} \tag{13}$$

with polynomials $p_{m-1+n,n}$ and $q_{m-1+n,n}$ of degree $m - 1 + n$.

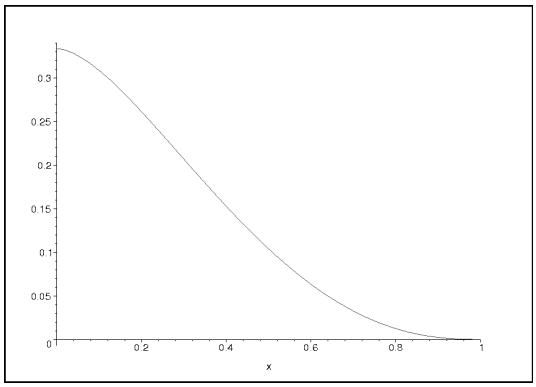

Figure 3: The compactly supported kernel $\phi_{2,1/2}(r)$

Interested readers should work out the case of integer k and half–integer μ, but these functions will not generate new classes of Sobolev spaces. To this end, we suggest to have a look at hypergeometric functions.

4 Expansion Kernels

A kernel

$$K(x,t) = \sum_{j=0}^{\infty} \lambda_j \varphi_j(x) \varphi_j(t) \tag{14}$$

based on a sequence $\{\varphi_j\}_{j\geq 0}$ of functions on Ω and a positive sequence $\{\lambda_j\}_{j\geq 0}$ of scalars can be called an *expansion kernel* if the summability

condition

$$\sum_{j=0}^{\infty} \lambda_j \varphi_j(x)^2 < \infty \text{ for all } x \in \Omega \tag{15}$$

holds. In other contexts, namely in Machine Learning, such kernels are sometimes called *Mercer* kernels due to their well–known connection to positive integral operators [18] and to the Mercer theorem. But they could also be called *Hilbert–Schmidt* kernels, because the expansion arises naturally as an eigenfunction expansion of the Hilbert–Schmidt integral operator

$$I(f)(x) := \int_{\Omega} K(x,y) f(y) dy,$$

and Mercer's theorem just asserts existence of the expansion with positive eigenvalues λ_j tending to zero for $j \to \infty$, while the eigenfunctions φ_j satisfy $I(\varphi_j) = \lambda_j \varphi_j$, are orthonormal in $L_2(\Omega)$ and orthogonal in the native Hilbert space for K. Each continuous positive definite kernel K on a bounded domain Ω has such an expansion, which, however, is hard to calculate and strongly domain–dependent.

Thus, Real Analysis allows to rewrite fairly general kernels as expansion kernels, but there also is a synthetic point of view going backwards, namely constructing a kernel from the λ_j and the φ_j under the summability condition (15).

The synthetic approach is the standard one in Machine Learning, and we shall give it a general treatment here, leaving details of kernel construction for Machine Learning to the specialized literature, in particular Part Three of the book [25] by J. Shawe–Taylor and N. Cristianini.

In Machine Learning, the domain Ω is a fairly general set of objects about which something is to be learned. The set has no structure at all, since it may consist of texts, images, or graphs, for instance. The functions φ_j associate to each object $x \in \Omega$ a certain property value, and the full set of the φ_j should map into a weighted ℓ_2 sequence space such that the summability condition is satisfied, i.e. into

$$\ell_{2,\Lambda} := \left\{ \{c_j\}_{j\geq 0} \ : \ \sum_{j=0}^{\infty} \lambda_j |c_j|^2 < \infty \right\}$$

with the inner product

$$(\{a_j\}_{j\geq 0}, \{b_j\}_{j\geq 0})_{\ell_{2,\Lambda}} := \sum_{j=0}^{\infty} \lambda_j a_j b_j.$$

The map $x \mapsto \{\varphi_j(x)\}_{j\geq 0} \in \ell_{2,\Lambda}$ is called the *feature map*. In specific applications, there will only be finitely many functions comprising the feature map, but we focus here on the infinite case. To proceed towards a Hilbert space of functions for which K is reproducing, one should look at the sequence $\Lambda\Phi(x) := \{\lambda_j\varphi_j(x)\}_{j\geq 0}$ of coefficients of the function $K(x,\cdot)$ for fixed $x \in \Omega$. This sequence lies in the space $\ell_{2,\Lambda^{-1}}$ with an inner product defined as above but using λ_j^{-1} instead of λ_j. Thus we should look at expansions into the φ_j such that the coefficient sequences lie in $\ell_{2,\Lambda^{-1}}$.

If the feature functions φ_j are linearly dependent, there are problems with non-unique coefficients for expansions into the φ_j. To handle this in general, we refer the reader to the use of frames, as done in R. Opfer's dissertation [15].

Instead, we now assume linear independence of the φ_j over Ω and define a space of functions

$$\mathcal{F} := \left\{ f(x) = \sum_{j\geq 0} c_j(f)\varphi_j(x) \ : \ \|\{c_j(f)\}_{j\geq 0}\|_{\ell_{2,\Lambda^{-1}}} := \sum_{j\geq 0} \frac{c_j^2(f)}{\lambda_j} < \infty \right\} \tag{16}$$

on the general set Ω. Clearly, all functions $K(x,\cdot)$ and all φ_j lie in this space, and due to uniqueness of the coefficients we can define the inner product

$$(f,g)_{\mathcal{F}} := \sum_{j\geq 0} \frac{c_j(f)c_j(g)}{\lambda_j} \text{ for all } f,g \in \mathcal{F}.$$

This space clearly is isometric to the Hilbert sequence space $\ell_{2,\Lambda^{-1}}$ and the kernel K is reproducing in it while the functions φ_ν are orthogonal in $\mathcal{F}$ with

$$(\varphi_j,\varphi_k)_{\mathcal{F}} = \frac{\delta_{jk}}{\lambda_k} \text{ for all } j,k \geq 0.$$

Thus, the synthesis approach recovers the orthogonality of the expansion in the native Hilbert space of the kernel, though the φ_j were chosen fairly arbitrarily.

Now it is time for some examples.

EXAMPLE 3. We start with a simple one in $\mathbb{R}$, namely

$$\exp(-(x-y)^2) = \sum_{n=0}^{\infty} \frac{2^n}{n!} \underbrace{x^n \exp(-x^2)}_{=:\varphi_n(x)} \underbrace{y^n \exp(-y^2)}_{=:\varphi_n(y)} \tag{17}$$

which easily follows by expansion of the left–hand side. Normally, the native space for this kernel would be defined via Fourier transforms, but by (16) we can write it as the Hilbert space of analytic functions with representations

$$f(x) = \exp(-x^2)\sum_{n=0}^{\infty} c_n x^n \text{ with } \sum_{n=0}^{\infty} \frac{n! c_n^2}{2^n} < \infty.$$

EXAMPLE 4. Another case is given by the well–known formula

$$\exp\left(-\frac{x^2t^2 - 2txy + y^2t^2}{2(1-t^2)}\right) = \sqrt{1-t^2}\sum_{n=0}^{\infty} H_n(x)H_n(y)\frac{t^n}{n!} \qquad (18)$$

of Mehler (cf. [26]) with $x, y \in \mathbb{R}$, the Hermite polynomials H_n and a fixed parameter $t \in (-1, 1)$. The native Hilbert space consists of functions of the form

$$f(x) = \sum_{n=0}^{\infty} c_n H_n(x) \text{ with } \sum_{n=0}^{\infty} \frac{n! c_n^2}{t^n} < \infty.$$

EXAMPLE 5. A *multiscale* expansion kernel is

$$K(x,y) = \sum_{j\geq 0} \lambda_j \underbrace{\sum_{k\in\mathbb{Z}} \varphi(2^j x - k)\varphi(2^j y - k)}_{:=\Phi_j(x,y)} \qquad (19)$$

considered by R. Opfer in his Ph.D. thesis [15, 16]. It uses a refinable function $\varphi : \mathbb{R}^d \leftrightarrow \mathbb{R}$, compactly supported, and performs a wavelet–style superposition into scale levels j and shifts k on multivariate grids. It generates certain Sobolev spaces and allows kernel–based interpolants to be rewritten in terms of wavelet bases. Thus it provides a link between grid–based wavelets and meshfree kernel–based methods. An example of such a kernel is shown in Fig. 4.

Another interesting expansion of given kernels can be obtained via the *Newton basis* (cf. [11]). The idea is as follows. Take sets $X_n := \{x_1, \ldots, x_n\} \subset \Omega \subset \mathbb{R}^d$ with fill distance

$$h_n := h(X_n, \Omega) := \sup_{y\in\Omega} \min_{1\leq j\leq n} \|y - x_j\|_2 \to 0 \text{ for } n \to \infty.$$

Then there is a unique *Newton-like* basis $v_1, v_2, \ldots, v_n, \ldots$ with

$$\begin{aligned} v_j &\in \text{Span}\,\{K(\cdot, x_k),\ 1 \leq k \leq j\} \\ v_j(x_k) &= 0,\ 1 \leq k < j \\ \|v_j\|_{\mathcal{F}} &= 1 \end{aligned}$$

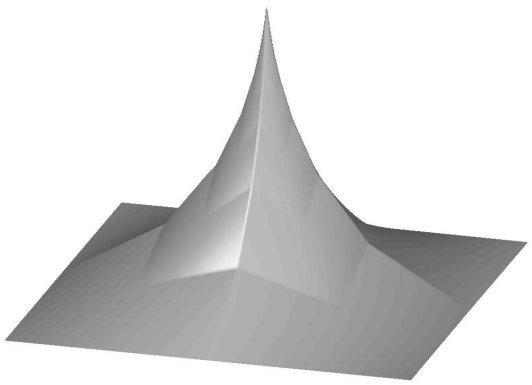

Figure 4: This multiscale kernel is taken from the title of the Ph.D. dissertation [15] by R. Opfer.

and the remarkable expansion property

$$K(x,y) = \sum_{n=1}^{\infty} v_n(x) v_n(y)\,, \tag{20}$$

showing that kernel expansions are by no means unique if they are not obtained via eigenfunction expansions. Here, we digressed from [11] by using a different but much more natural normalization.

If interpolation of a function f of the native space $\mathcal{F}$ in the points of $X_n := \{x_1, \ldots, x_n\}$ is performed by an interpolant

$$u_{f,X_n}(\cdot) = \sum_{j=1}^{n} \Delta_j(f)\; v_j(\cdot)$$

written in the new basis, the coefficients $\Delta_j(f)$ are uniquely defined linear functionals composed of the point evaluation functionals δ_{x_k} for $1 \le k \le j$, generalizing divided differences. They turn out to be orthonormal in the dual of the native space, being the Riesz representers of the v_j. Furthermore, the basis and the coefficients have the stability properties

$$\sum_{j=1}^{n} v_j^2(x) \le K(x,x),\; \sum_{j=1}^{n} \Delta_j^2(f) = \|u_{f,X_n}\|_{\mathcal{F}}^2 \le \|f\|_{\mathcal{F}}^2$$

bounded above independently of n. The power function $P_{X_n,\mathcal{F}}$ of the first section has the representation

$$P_{X_n,\mathcal{F}}^2(x) = K(x,x) - \sum_{j=1}^{n} v_j^2(x)$$

implying

$$v_n^2(x) = P^2_{X_{n-1},\mathcal{F}}(x) - P^2_{X_n,\mathcal{F}}(x) \le P^2_{X_{n-1},\mathcal{F}}(x) \to 0 \text{ for } n \to \infty.$$

Details can be found in [21].

A useful strategy for choosing interpolation points is to pick x_n as a point where $P_{X_{n-1},\mathcal{F}}$ attains its maximum. This technique goes back to [5], and by the above argument, using $P^2_{X_n,\mathcal{F}}(x_n) = 0$, it leads to

$$v_n^2(x) \le P^2_{X_{n-1},\mathcal{F}}(x) \le P^2_{X_{n-1},\mathcal{F}}(x_n) = v_n^2(x_n) \text{ for all } n \in \mathbb{N},\ x \in \Omega, \quad (21)$$

proving that the basis has no extra oscillations.

Example 6. For Fig. 5 we started with $201 \times 201 = 40401$ gridded points in $[-1,+1]^2$. From the points falling into the domain in the bottom right plot, we picked 40 points x_n adaptively by the above rule. The selected points are shown in the bottom right plot, while the other plots show the Newton basis functions v_{21}, v_{33}, and v_{37} illustrating their remarkable non–oscillation properties. In this case, the kernel was the inverse multiquadric $K(r) = (1+r^2)^{-2}$.

Example 7. Within the degree 6 polynomial interpolation of Example 2 and Fig. 2, we can form a Newton–type basis by a pivoted LU decomposition of the Vandermonde matrix. Figure 6 shows the resulting Newton basis functions v_9, v_{15}, v_{27}, and v_{28}, and now the non–oscillation is a consequence of pivoting by point permutations.

5 Polynomial Kernels

In Machine Learning, the introductory example for classification problems involves a maximal margin around a separating hyperplane. A closer look reveals that the *bilinear kernel*

$$K(x,y) := x^T y = x \cdot y = (x,y)_2 \text{ for all } x, y \in \mathbb{R}^d$$

on $\mathbb{R}^d$ is behind the scene. But since products of positive semi–definite kernels are again positive semi-definite, one can generalize to *homogeneous polynomial kernels*

$$K(x,y) := (x^T y)^n = \sum_{|\alpha|=n} \binom{n}{\alpha} x^\alpha y^\alpha \text{ for all } x, y \in \mathbb{R}^d.$$

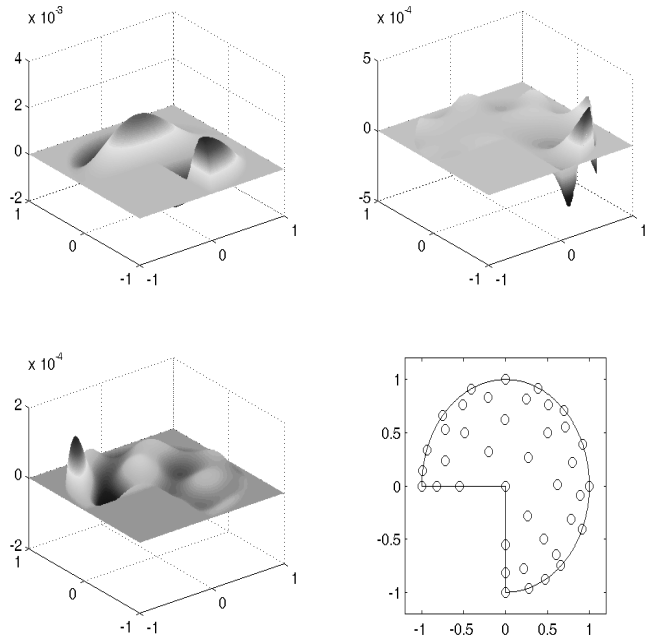

Figure 5: Selected Newton basis functions

Furthermore, positive sums of positive semidefinite kernels are positive semidefinite. This leads on to the *exponential kernel*

$$K(x,y) := \exp(x^T y) := \sum_{n=0}^{\infty} \frac{1}{n!}(x^T y)^n = \sum_{\alpha \in \mathbb{Z}^d} \frac{1}{|\alpha|!} \binom{|\alpha|}{\alpha} x^\alpha y^\alpha \text{ for all } x, y \in \mathbb{R}^d, \tag{22}$$

and, more generally, to *power series kernels*

$$K(x,y) := \sum_{\alpha \in \mathbb{Z}^d} c_\alpha x^\alpha y^\alpha \text{ for all } x, y \in \mathbb{R}^d \tag{23}$$

as recently introduced in [30] by Barbara Zwicknagl with interesting results on spectral convergence of interpolants. But we shall not deal with multivariate power series kernels in full generality here. Instead, we take a short look at the polynomial and the univariate cases.

For polynomial interpolation of scattered locations $X = \{x_1, \ldots, x_n\} \subset \mathbb{R}^d$, the result will strongly depend on the geometry of the points and the selection strategy of *admissible degree–minimal solutions*. The minimum

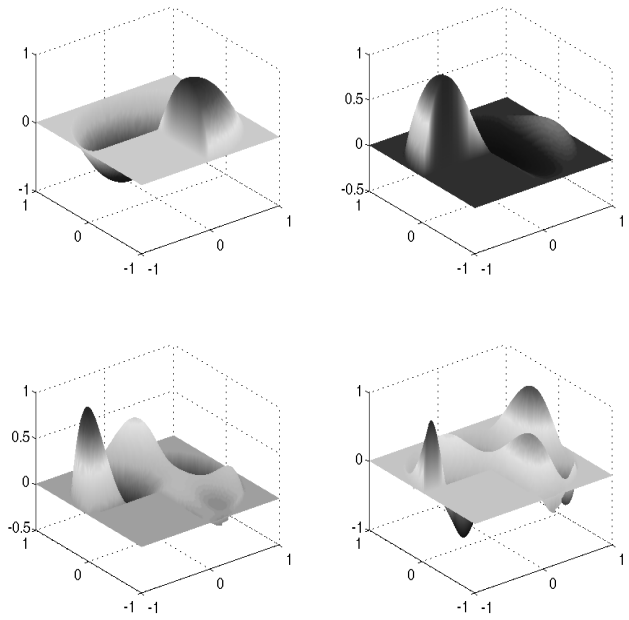

Figure 6: Selected Newton basis functions for polynomials.

necessary degree $m(X)$ is attained when the Vandermonde matrix $V_{X,k}$, with entries x_j^α, $1 \le j \le n$, $0 \le |\alpha| \le k$, has rank n for the minimal possible $k := m(X)$. Each degree–minimal polynomial interpolant to data $f(x_1), \ldots, f(x_n)$ then has coefficients c_α satisfying the linear system

$$\sum_{0 \le |\alpha| \le m(X)} c_\alpha x_j^\alpha = f(x_j),\ 1 \le j \le n$$

or in shorthand

$$V_{X,m(X)}\, c = f_X\,, \tag{24}$$

where the Vandermonde matrix will, in most cases, be non–square, making the solution non–unique. All useful solvers must have some strategy of selecting a solution. A particularly sophisticated technique is due to de Boor and Ron [3, 4], but here we want to point out how polynomial kernels come into play.

A reasonable way to regularize the (possibly) ill–conditioned system (24) is to introduce a diagonal matrix Λ with positive entries λ_α for $0 \le |\alpha| \le m(X)$ and to look for coefficient vectors of the form $c = \Lambda V_{X,m(X)}^T b$ with

$b \in \mathbb{R}^n$. This amounts to solving

$$V_{X,m(X)}\, \Lambda V_{X,m(X)}^T b = f_X \tag{25}$$

for b, and now the matrix is symmetric and positive definite. It is an easy exercise to figure out that the above technique minimizes $\|c\|_{2,\Lambda^{-1}}$ under all solutions of (24) in the terminology of the previous section, and that the matrix in (25) is the kernel matrix for the polynomial kernel

$$K_X(x,y) := \sum_{0 \le |\alpha| \le m(X)} \lambda_\alpha x^\alpha y^\alpha \text{ for all } x, y \in \mathbb{R}^d$$

which is positive semidefinite on $\mathbb{R}^d$ and positive definite on X. This approach yields a variety of minimal–degree polynomial interpolation techniques, depending on which λ_α are chosen. But it should be noted that the de Boor–Ron method is not of this simple kind, since it takes additional effort to maintain homogeneous solutions for data from homogeneous polynomials.

6 Taylor Spaces and Kernels

Now, we shall specialize power series kernels (23) to the univariate case, but we take a different starting point to arrive there.

The Taylor-MacLaurin series

$$f(x) \;=\; \sum_{n=0}^{\infty} f^{(j)}(0) \frac{x^j}{j!} \tag{26}$$

can be viewed as a reproduction formula like the ones well–known from kernel techniques. Indeed, it should take the form $f(x) = (f, K(x,\cdot))_{\mathcal{F}}$ in some suitable RKHS $\mathcal{F}$ with kernel K, but we still have to find a suitable space with a suitable kernel. Surprisingly, there are many spaces and kernels that work!

In fact, all kernels of the form

$$K(x,t) := \sum_{j \in N} \lambda_j \frac{(xt)^j}{(j!)^2},\ x,\, t \in \mathbb{R},\ \lambda_j > 0,\ N \subseteq \mathbb{N} \tag{27}$$

with the summability condition (15) taking the form

$$\sum_{j \in N} \lambda_j \frac{x^{2j}}{(j!)^2} < \infty \text{ for all } x \in \Omega$$

will be admissible here, and we allow the set $N \in \mathbb{N}_{\geq 0}$ to be infinite but otherwise arbitrary. Depending on the weights λ_j, the domain Ω can be all $\mathbb{R}$ or just an interval.

The connection to expansion kernels is via $\varphi_j(x) = x^j/j!$, and Section §4 teaches us that the native space $\mathcal{F}$ consists of all functions f of the form

$$f(x) = \sum_{j \in N} c_j \frac{x^j}{j!} \text{ with } \sum_{j \in N} \frac{c_j^2}{\lambda_j} < \infty.$$

Then $c_j = f^{(j)}(0)$ leads to the space

$$\mathcal{F} = \left\{ f \ : \ f(x) = \sum_{n \in N} f^{(j)}(0) \frac{x^j}{j!} \text{ for all } x \in \Omega, \ \sum_{n \in N} \frac{(f^{(j)}(0))^2}{\lambda_j} < \infty \right\}. \tag{28}$$

The Taylor formula as a reproduction formula now follows along the lines of Section §4 using the inner product

$$(f,g)_{\mathcal{F}} := \sum_{n \in N} \frac{f^{(j)}(0) g^{(j)}(0)}{\lambda_j} \text{ for all } f, g \in \mathcal{F}.$$

For a simultaneous discussion of all of these kernels, we go into the complex plane by introducing $z := x\overline{t}$ and look at kernels

$$\kappa(z) := \sum_{j \in N} \lambda_j \frac{z^j}{(j!)^2} \tag{29}$$

on suitable discs $D \subseteq \mathbb{C}$ around zero. Then, all power series with positive real coefficients are candidates, the coefficients determining the size of the admissible disc D, while $\Omega \subseteq \mathbb{R}$ must be in the interior of D.

In Table 1, we collect some of these *Taylor kernels* and the corresponding expansion coefficients.

Detailed results on these kernels are in [31], which is heavily relying on [30]. The common punchline of [31] is

- The native Hilbert spaces for Taylor kernels can be characterized via complex analysis.
- Interpolation of functions in native Taylor spaces by translates of Taylor kernels are exponentially convergent.

Notice that this yields a large variety of results on univariate interpolation by rational functions, exponentials, shifted logarithms, exponentials, hyperbolic cosines, etc. Here, we only illustrate a single case.

$\kappa(z) = \sum_{j\in N} \lambda_j \frac{z^j}{(j!)^2},\ z = x\bar{t}$	N	λ_j
$(1-z)^{-1}, -1 < \|z\| < 1$	$\mathbb{N}$	$(j!)^2$
$(1-z^2)^{-1}, -1 < \|z\| < 1$	$2\mathbb{N}$	$(j!)^2$
$(1-z)^{-\alpha}, \alpha \in \mathbb{N}, -1 < \|z\| < 1$	$\mathbb{N}$	$\frac{(\alpha+j-1)!j!}{(\alpha-1)!}$
$-\frac{\log(1-z)}{z}, -1 < \|z\| < 1$	$\mathbb{N}$	$\frac{(j!)^2}{j+1}$
$(1-z^2)^{-\alpha}, \alpha \in \mathbb{N}, -1 < \|z\| < 1$	$2\mathbb{N}$	$\frac{(\alpha+j-1)!j!}{(\alpha-1)!}$
$\exp(z)$	$\mathbb{N}$	$j!$
$\sinh(z)$	$2\mathbb{N}+1$	$j!$
$\sinh(z)/z$	$2\mathbb{N}$	$\frac{j!}{j+1}$
$\cosh(z)$	$2\mathbb{N}$	$j!$
$z^{-\alpha} I_\alpha(z)$	$2\mathbb{N}$	$\frac{j!}{4^j \Gamma(j+\alpha+1)}$

Table 1: Some Taylor Kernels and the corresponding expansion coefficients

Theorem 1. *(cf. [31]) The native Hilbert space $F_{\mathbb{R}}$ for the Szegö kernel $R(x,t) = (1-xt)^{-1}$ consists of real-valued functions whose complex extensions lie in the Hardy space $\mathbf{H}^2$.*

Theorem 2. *(cf. [31]) 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 $f \in F_{\mathbb{R}}$, the error between f and its rational interpolant of the form*

$$s_{f,X}(t) := \sum_{x_j \in X} \alpha_j \frac{1}{1 - t\, x_j} \tag{30}$$

on the set X, is bounded by

$$\|f - s_{f,X}\|_{L_\infty[-a,a]} \le \mathsf{e}^{-c_1/h} \|f\|_{F_R} \ . \tag{31}$$

7 Periodic Kernels

For spaces of 2π–periodic functions, there are some nice and useful kernels. We list now some interesting examples.

EXAMPLE 8. Consider

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \cos(n(x-y)) \quad = \quad \frac{1}{4}(x-y)^2 - \frac{1}{2}\pi(x-y) + \frac{1}{6}\pi^2$$

for $x - y \in [0, 2\pi]$ with periodic extension. Notice that the above series is a polynomial of degree 2 in $x - y$. An example is provided in the top left plot of Fig. 7.

EXAMPLE 9. In more generality, the functions

$$\sum_{n=1}^{\infty} \frac{1}{n^{2k}} \cos(n\ t) \tag{32}$$

represent polynomials of degree $2k$ on $[0, 2\pi]$.

To see this, consider Hurwitz-Fourier expansions

$$B_m(x) = -\frac{m!}{(2\pi i)^m} \sum_{n=-\infty,\, n\neq 0}^{+\infty} n^{-m} e^{2\pi i n x}$$

of the Bernoulli polynomials B_m of degree m on $[0, 1]$ (for details see, e.g. `http://mathworld.wolfram.com/BernoulliPolynomial.html`). If we set $t = 2\pi x$ and $m = 2k$, we get

$$\begin{aligned} B_{2k}(\tfrac{t}{2\pi}) &= (-1)^{k+1} \frac{(2k)!}{(2\pi)^{2k}} \sum_{n=-\infty,\, n\neq 0}^{+\infty} n^{-2k} (\cos(nt) + i \sin(nt)) \\ &= 2(-1)^{k+1} \frac{(2k)!}{(2\pi)^{2k}} \sum_{n=1}^{+\infty} n^{-2k} \cos(nt) \end{aligned}$$

that proves our claim.

EXAMPLE 10. But there are still some other nice cases due to A. Meyenburg [9], namely

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{1}{n!} \cos(n(x-y)) &= \cos(\sin(x-y)) \cdot \exp(\cos(x-y)) \\ \sum_{n=0}^{\infty} \frac{1}{2^n} \cos(n(x-y)) &= \frac{1 - \frac{1}{2}\cos(x-y)}{1 - \cos(x-y) + \frac{1}{4}} \end{aligned} \tag{33}$$

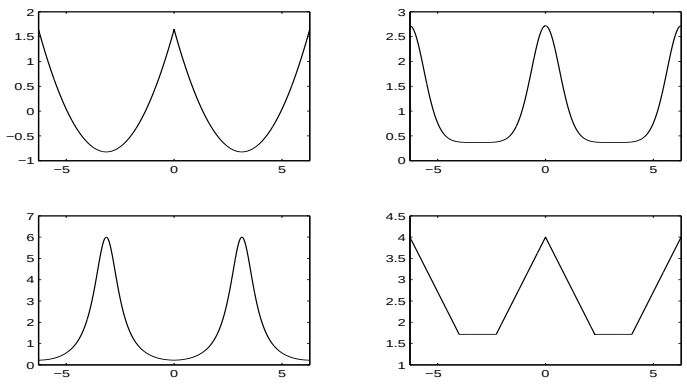

Figure 7: Periodic kernels

depicted in Fig. 7 on the top right and bottom left, and

$$\begin{aligned}
\exp(-2|x|) &= \frac{4}{\pi}\sum_{n=0}^{\infty}\frac{1-(-1)^n e^{-2\pi}}{4+n^2}\cos(nx),\ x\in[-\pi,\pi],\\
K(x) &:= \begin{cases} 4-x & 0\le x \le 2\pi-4\\ 8-2\pi & 2\pi-4< x <4\\ 4-2\pi+x & 4\le x\le 2\pi\end{cases}\\
&= \frac{16}{\pi}+\sum_{n=1}^{\infty}\frac{4\sin^2(2n)}{\pi n^2}\cos(nx).
\end{aligned}$$

The final one is plotted in the bottom right of Fig. 7.

The analysis of general kernels of the form

$$\begin{aligned}
K(x-y) &:= \frac{1}{\pi}\left(\frac{\lambda_0}{2}+\sum_{n=1}^{\infty}\lambda_n\cos(n(x-y))\right)\\
&= \frac{1}{\pi}\left(\frac{\lambda_0}{2}+\sum_{n=1}^{\infty}\lambda_n\left(\cos(nx)\cos(ny)+\sin(nx)\sin(ny)\right)\right)
\end{aligned}$$

can proceed via the Mercer theorem, since the functions

$$\frac{1}{\sqrt{2}},\ \cos(nx),\ \sin(nx),\ n\ge 1 \tag{34}$$

are orthonormal under

$$(f,g)_\pi := \frac{1}{\pi}\int_{-\pi}^{+\pi} f(t)g(t)dt.$$

Thus, for instance,

$$\begin{aligned}
& \int_{-\pi}^{+\pi} K(x-t)\cos(mt)dt \\
= & \frac{1}{\pi}\int_{-\pi}^{+\pi} \cos(mt)\left(\frac{\lambda_0}{2} + \sum_{n=0}^{\infty}\lambda_n\left(\cos(nx)\cos(nt)+\sin(nx)\sin(nt)\right)\right)dt \\
= & \lambda_m \cos(mx)
\end{aligned}$$

proves that the eigenvalues of the kernel are λ_j with the eigenfunctions being exactly the standard L_2–orthonormal basis (34). Therefore the kernel is in Mercer expansion form, and the native space $\mathcal{F}$ consists of all functions representable as Fourier series

$$f(x) = \frac{a_0}{\sqrt{2}} + \sum_{j=1}^{\infty}\left(a_j\cos(jx) + b_j\sin(jx)\right)$$

with the summability condition

$$\sum_{j=0}^{\infty}\frac{a_j^2+b_j^2}{\lambda_j} < \infty$$

and the inner product

$$(f,g)_{\mathcal{F}} := \sum_{j=0}^{\infty}\frac{a_j(f)a_j(g)+b_j(f)b_j(g)}{\lambda_j}.$$

Note that the weights $\lambda_n = n^{-2k}$ in (32) lead to Sobolev–type native spaces, while the kernels in (33) generate spaces of 2π–periodic functions with entire or meromorphic extensions into $\mathbb{C}$, respectively.

Of course, one can omit the sinus terms above, and then one gets periodic kernels that link to Chebyshev polynomials.

EXAMPLE 11. For instance, the kernel

$$K(x,y) := \sum_{n=0}^{\infty}\frac{1}{n!}T_n(x)T_n(y)$$

on $[-1,1]$ can be transformed by substitution $x=\cos(\varphi),\ y=\cos(\psi)$ into

$$\begin{aligned}
& \sum_{n=0}^{\infty}\frac{1}{n!}\cos(n\varphi)\cos(n\psi) \\
= & \frac{1}{2}\sum_{n=0}^{\infty}\frac{1}{n!}\left(\cos(n(\varphi+\psi))+\cos(n(\varphi-\psi))\right) \\
= & \frac{1}{2}\left[\cos(\sin(\varphi+\psi))\cdot\exp(\cos(\varphi+\psi))+\cos(\sin(\varphi-\psi))\cdot\exp(\cos(\varphi-\psi))\right].
\end{aligned}$$

8 Kernels for PDEs

Various well–known kernels have connections to Partial Differential Equations. To start with, we look at homogeneous differential equations $Lu = 0$ with a linear elliptic differential operator L on a bounded domain $\Omega \subset \mathbb{R}^d$. A *fundamental solution* solves the distributional differential equation $Lu = \delta_x$ for fixed x, and it thus can be written as a kernel $K(x, \cdot)$.

For example, in this sense, the *thin–plate spline* kernel

$$K(x, y) := \|x - y\|_2^2 \log(\|x - y\|_2) \text{ for all } x, y \in \mathbb{R}^d$$

is the fundamental solution of the biharmonic equation $\Delta^2 = 0$ in $\mathbb{R}^2$, and the full class of polyharmonic kernels are fundamental solutions of powers of the Laplace operator in various dimensions. The *Method of Fundamental Solutions* [2] solves homogeneous differential equations $Lu = 0$ on bounded domains by superimposing fundamental solutions $K(\cdot, x_j)$ using points x_j outside the domain in order to avoid singularities in the domain. There are various ways to select the points x_j properly (see e.g. [19]).

Unfortunately, fundamental solutions usually have a singularity, and this calls for new kernels that are singularity–free homogeneous solutions of differential operators. We now show how to do this for the Laplace operator in 2D, i.e. we construct kernels on $\mathbb{R}^2$ that are harmonic in each variable and singularity–free.

Real and imaginary parts of holomorphic complex functions are harmonic. Thus the parts $r^n \cos(n\varphi)$ and $r^n \sin(n\varphi)$ of $z^n = r^n \exp(in\varphi)$ are harmonic. We can introduce two polar coordinate variables (r, φ), (s, ψ) and write down a harmonic expansion kernel

$$\begin{aligned} & K_c((r, \varphi); (s, \psi))) \\ := & \sum_{n=0}^{\infty} \frac{1}{n!} c^{2n} r^n s^n \cos(n(\varphi - \psi)) \\ = & \exp(c^2 rs \cos(\varphi - \psi)) \cdot \cos(c^2 rs \sin(\varphi - \psi)) \end{aligned}$$

using the periodic exponential kernel from (33) again.

For error and convergence analysis and application to the solution of Poisson problems we suggest the reader to refer to the recent paper [23]. An interesting result from that paper is the following theorem.

Theorem 3. *If the boundary function $r(\varphi)$ of the domain $\Omega \subset \mathbb{R}^2$ is in C^k and if the above kernel is chosen, the native space of the kernel restricted to*

the boundary is continuously embedded in Sobolev space $W_2^k[0, 2\pi]$. Furthermore, interpolation on the boundary by harmonic kernels yields an overall error of order $h^{k-1/2}$ in the $L_\infty(\overline{\Omega})$ norm.

EXAMPLE 12. An illustration is given in Fig. 8. In that example, we look at harmonic interpolation on a few points on the boundary of a cardioid, but we do not provide fixed interpolation data. Instead, we calculate the power function which describes the pointwise norm of the error functional, as pointed out in the first section. But since harmonic functions are completely determined by their boundary values, the power function should be small in the interior of the domain even if there are no data points.

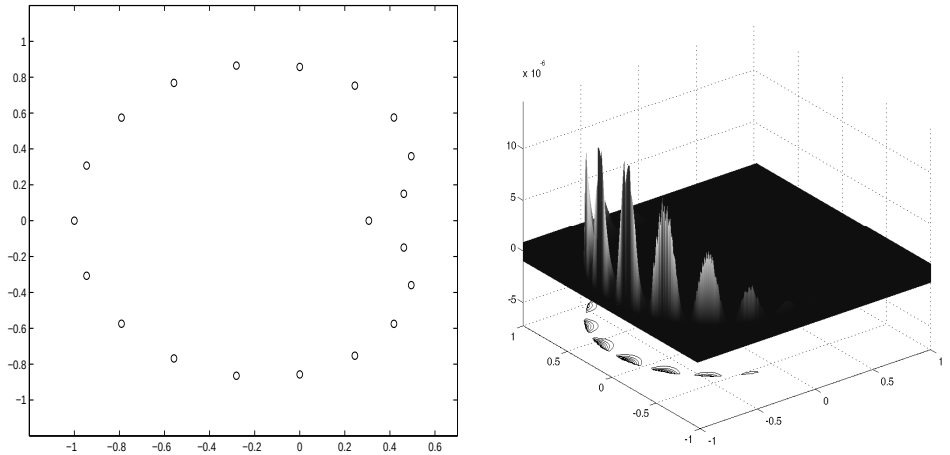

Figure 8: Harmonic interpolation.

EXAMPLE 13. One can also use the harmonic kernel to interpolate scattered data sampled from harmonic functions. An example of interpolation of $\exp(x)\cos(y)$ is displayed in Fig. 9. The interpolation points were selected adaptively from a larger scattered sample, as displayed in the bottom right figure.

Besides going for homogeneous solutions, one can also construct inhomogeneous solutions via special kernels. The idea is simple: if a PDE $Lu = f$ is to be solved, take translates $K(\cdot, x_j)$ of a smooth kernel K, apply L to these translates and approximate the function f by *particular solutions* $f_j = LK(\cdot, x_j)$. If the error in the approximation

$$f \approx \sum_{j=1}^{n} \alpha_j f_j \tag{35}$$

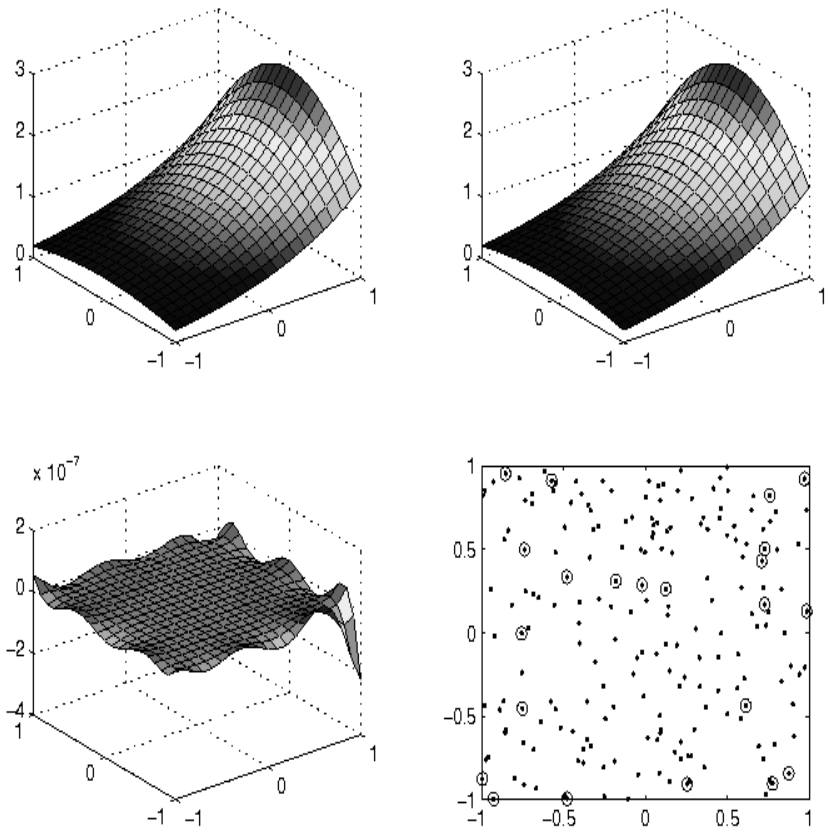

Figure 9: Interpolation with harmonic kernel

is small, the PDE $Lu = f$ is solved appproximatively by

$$u = \sum_{j=1}^{n} \alpha_j K(\cdot, x_j).$$

This does not account for any boundary conditions, but these can be attained in a second step constructing a homogeneous solution to be added to u. This two–step procedure is called the *dual reciprocity method* [17] and arose in the context of boundary element techniques.

If L is elliptic, the application of L to a positive definite translation–invariant kernel will not destroy the positive definiteness. Hence the recovery problem (35) can be done by kernel interpolation. For general operators, one can often go backwards, starting from $f_j = K(\cdot, x_j)$ and find an analytic solution u_j of $Lu_j = f_j$. But one can also use the positive definite kernel

$K_L(x,y) := L^x L^y K(x,y)$ where L is applied to both x and y, perform kernel interpolation of f with it to get coefficients solving

$$\sum_{j=1}^{n} K_L(x_j, x_k)\alpha_j = f(x_k),\ 1 \le k \le n$$

and then use

$$u := \sum_{j=1}^{n} LK(x_j, \cdot)\alpha_j$$

for an approximative solution of $Lu = f$. In the meantime, there are various papers providing kernels as particular solutions, see e.g. [10].

Note that users must be aware to handle higher–order derivatives of kernels in order to implement these procedures. We shall comment this later.

9 Special Kernels

Sometimes one runs into new kernels when looking at special solutions of PDEs. An example is the *equal width equation*

$$u_t + uu_x - \mu u_{xxt} = 0,\ x, t \in \mathbb{R}. \tag{36}$$

This nonlinear transport or wave equation has a special traveling soliton solution

$$u(x,t) = 3\,c\,\mathrm{sech}^2\left((x - x_0 - c\,t)/\sqrt{4\mu}\right). \tag{37}$$

The speed c is always one third of the amplitude, having the typical nonlinearity-based effect that higher waves travel faster. Surprisingly, the *soliton* kernel

$$K(x-y) := \mathrm{sech}^2\,(x-y) \tag{38}$$

is positive definite. This follows, because the hyperbolic secant has a positive Fourier transform which is again a hyperbolic secant, thus positive. This implies positive definiteness of the hyperbolic secant and consequently also for its square. A plot of the square of the hyperbolic secant $(1/\cosh)^2(\cdot)$ is displayed in Fig. 10.

A much more important class of PDE–dependent kernels generates *divergence–free* vector fields. Note that Computational Fluid Dynamics problems in 2D or 3D usually require the calculation of a divergence–free velocity field, for example when solving Navier–Stokes equations. But standard trial

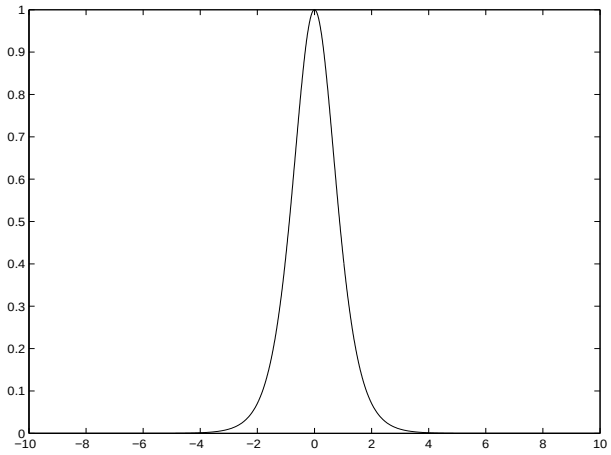

Figure 10: The function $(1/\cosh)^2$.

spaces, like finite elements, do not care for this directly. They introduce $\operatorname{div} u = 0$ as an additional condition and have to take care for Lagrange multipliers turning the PDE into a saddle–point problem.

The construction of divergence–free vector fields can be done starting from a smooth translation-invariant scalar kernel $K(x-y)$ on $\mathbb{R}^d$. Then a *matrix–valued kernel*

$$D(z) := (-\Delta \cdot I_d + \nabla\nabla^T)K(z),\ z = x - y \tag{39}$$

is formed, and each row or column is then a divergence–free vector field $\mathbb{R}^d \to \mathbb{R}^d$. The most important references on the topic are [13, 7, 8, 6].

Note that divergence–free vector fields v have the property that the integral of $v^T n$ over each closed curve (or surface in 3D) must be zero, where n is the normal. For compactly supported divergence–free 2D vector fields this means that the integral over $v^T n$ on each curve through the support must vanish. This makes it interesting to see special cases of divergence–free vector fields obtained via scalar kernels.

Example 14. Starting from the Gaussian kernel, we get Fig. 11 with two vortices, and similarly in Fig. 12 for the compactly supported Wendland kernel $K(r) = (1-r)_+^4(1+4r)$. This vector field then is compactly supported on the unit disc.

Example 15. Figure 13 shows an interpolation of a divergence–free vector field from scattered data. This is useful when recovering velocity fields from measurements or locally calculated values.

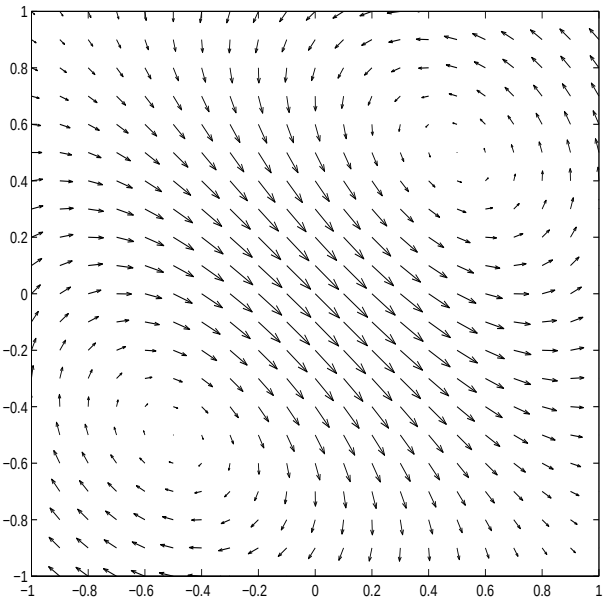

Figure 11: Divergence-free field from Gaussian kernel

At this point, readers might fear that the evaluation of higher derivatives of kernels is a cumbersome and error–prone procedure. But this is not true, due to the following technique based on rewriting radial kernels as $\phi(r) = f(r^2/2)$ in *f–form*. Then it can be observed (cf. [22]) that all standard classes of radial kernels in f–form are closed under differentiation. Typical cases are

- Sobolev/Whittle/Matérn functions $\phi_\nu(r) := K_\nu(r)r^\nu$ have $f'_\nu = -f_{\nu-1}$,
- Wendland functions $\phi_{d,k}(r)$ have $f'_{d,k} = -f_{d+2,k-1}$.

10 Conclusion

The goal of this journey through a zoo of nonstandard kernels was to encourage users to do their own *kernel engineering*. There are many ways to tailor kernels for special purposes, and there always is the technique of generalized power functions for error evaluation, no matter how strange or special the recovery process is.

Acknowledgements. This work has been done with the support of the

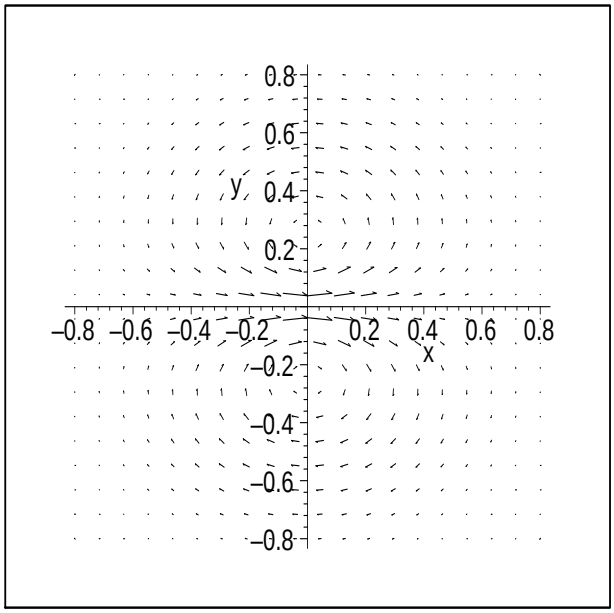

Figure 12: Divergence-free field from Wendland kernel

University of Verona, ex 60% funds, as the second author was working there. Moreover, the paper contains ideas presented at the *Second Dolomites Workshop on Constructive Approximation and Applications*, held in Alba di Canazei (Italy), 4-9 Sept. 2009. MATLAB© and MAPLE© programs are available from the research web pages of the authors.

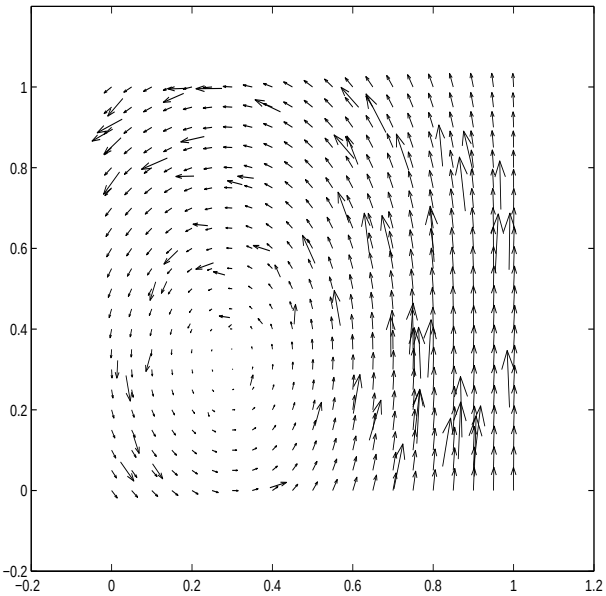

Figure 13: Divergence-free interpolation

References

[1] M.D. Buhmann. Radial functions on compact support. *Proc. Edinburgh Math. Soc.*, 41:33–46, 1998.

[2] C.S. Chen, A. Karageorghis, and Y.S. Smyrlis. *The Method of Fundamental Solutions - A Meshless Method*. Dynamic Publishers, 2008.

[3] C. de Boor and A. Ron. Computational aspects of polynomial interpolation in several variables. *Math. Comp.*, 8:705–727, 1992.

[4] C. de Boor and A. Ron. The least solution for the polynomial interpolation problem. *Math. Z.*, 210:347–378, 1992.

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

[6] E.J. Fuselier. Sobolev-type approximation rates for divergence-free and curl-free RBF interpolants. *Math. Comp.*, 77:1407–1423, 2008.

[7] S. Lowitzsch. Error estimates for matrix-valued radial basis function interpolation. *J. Approx. Theory*, 137(2):238–249, 2005.

[8] S. Lowitzsch. Matrix-valued radial basis functions: stability estimates and applications. *Adv. Comput. Math.*, 23(3):299–315, 2005.

[9] A. Meyenburg. Interpolation mit Translaten einer geraden, 2π–periodischen, stetigen Funktion. Staatsexamensarbeit, Universität Göttingen, 1996.

[10] A.S. Muleshkov, C.S. Chen, M.A. Golberg, and A.H.D. Cheng. Analytic particular solutions for inhomogeneous helmholtz-type equations. In S.N. Atluri and F.W. Brust, editors, *Adv. Comput. Eng. Sci.*, pages 27–32. Tech Science Press, 2000.

[11] St. Müller and R. Schaback. A Newton basis for kernel spaces. doi:10.1016/j.jat.2008.10.014, 2008.

[12] F.J. Narcowich and J.D. Ward. Norm of inverses and condition numbers for matrices associated with scattered data. *J. Approx. Theory*, 64:69–94, 1991.

[13] F.J. Narcowich and J.D. Ward. Generalized Hermite interpolation via matrix–valued conditionally positive definite functions. *Math. Comp.*, 63(208):661–687, 1994.

[14] E. Neuman. Moments and Fourier transforms of B-splines. *J. Comput. Appl. Math.*, 7:51–62, 1981.

[15] R. Opfer. *Multiscale Kernels*. Berichte aus der Mathematik. Shaker Verlag GmbH, Germany, 2004. Doctoral dissertation, Göttingen.

[16] R. Opfer. Multiscale kernels. *Adv. Comput. Math.*, 25(4):357–380, 2006.

[17] P.W. Partridge, C.A. Brebbia, and L.C. Wrobel. *The dual reciprocity boundary element method*. Computational Mechanics Publications, Southampton, Boston, 1992.

[18] 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.

[19] R. Schaback. An adaptive numerical solution of MFS systems. In C.S. Chen, A. Karageorghis, and Y.S. Smyrlis, editors, *The Method of Fundamental Solutions - A Meshless Method*, pages 1–27. Dynamic Publishers, 2008.

[20] R. Schaback. The missing Wendland functions. to appear in Adv. Comput. Math., 2009.

[21] R. Schaback. Newton bases. Preprint, Göttingen, 2009.

[22] R. Schaback. Programming hints for kernel-based methods. Technical report, 2009. Technical Report, Göttingen.

[23] R. Schaback. Solving the laplace equation by meshless collocation using harmonic kernels. Adv. Comp. Math. 31, 2009, 457–470.

[24] R. Schaback and H. Wendland. Characterization and construction of radial basis functions. In N. Dyn, D. Leviatan, D.and Levin, and A. Pinkus, editors, *Multivariate Approximation and Applications*, pages 1–24. Cambridge University Press, 2001.

[25] J. Shawe-Taylor and N. Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, 2004.

[26] G. N. Watson. Notes on generating functions of polynomials: (2), Hermite polynomials. *J. London Math. Soc.*, 8:194–199, 1933.

[27] H. Wendland. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Adv. Comput. Math.*, 4:389–396, 1995.

[28] H. Wendland. *Scattered Data Approximation*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, Cambridge, 2005.

[29] Z. Wu. Multivariate compactly supported positive definite radial functions. *Adv. Comput. Math.*, 4:283–292, 1995.

[30] B. Zwicknagl. Power series kernels. *Constr. Approx.*, 29(1):61–84, 2009.

[31] B. Zwicknagl and R. Schaback. Interpolation and approximation in Taylor spaces. in preparation, 2009.