Kernel Construction Techniques

Robert Schaback

Göttingen Academy of Sciences and Humanities

Institut für Numerische und Angewandte Mathematik,
Georg-August-Universität Göttingen

schaback@math.uni-goettingen.de
https://num.math.uni-goettingen.de/schaback/

Version of February 12, 2026

Abstract: This paper extends a talk at the Conference on Modern Kernel Methods and Applications in Hong Kong Baptist University in January 2026. It first reviews standard kernel constructions, but then goes over to recent extensions and presents many starting points to construct future kernels. The goal of the paper is to encourage readers to go into *kernel engineering*, in particular towards constructing C^∞ kernels with compact support. The techniques presented here use series expansions, convolution, subdivision, componentwise products and radialization. Most cases are illustrated by examples, but a deeper analysis is left for further research, together with applications.

1 Introduction

(`SecIntro`) The paper starts with the basics of kernels in Section 2(`SecBas`) , including the close connection to random fields. Mercer- and Karhunen-Loève expansions are the background for Section 3(`SecExp`) , but in the rest of the paper the expansions do not follow from the kernels. Instead, kernels are constructed from expansions into infinite series of functions. This will later cover expansions into translates of mollifiers and fundamental solutions of PDEs. Expansions into orthonormal systems in L_2 work as well, though the associated kernels have no well-defined pointwise values. Applying smoothing maps then allows to generate expansion kernels for a full scale of Sobolev spaces.

In Section 3.1(`SecExpRanFun`) , expansions facilitate the approach to paths of random fields, and then Section 3.2(`SecExpEqui`) explains the equivalence of the deterministic and the nondeterministic approaches to kernels.

The autocorrelation of general Fourier-transformable functions will have a nonnegative Fourier transform and then be positive semidefinite. This is the subject of Section 4(`SecConv`) starting from B-splines and going to C^∞ compactly supported functions generated either by autocorrelation of standard mollifiers or by infinite convolution of piecewise constants.

Another widely ignored kernel construction is by forming products of univariate kernels in Section 5(SecProd) . Except for the Gaussian, these cannot be radial, but they may come surprisingly close to radiality, as examples show.

All sections on kernel construction contain numerical examples and point out various directions for further research. A final section summarizes them for the convenience of readers.

2 Basics

(SecBas) We focus on *kernels*

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

on domains $\Omega \subseteq \mathbb{R}^d$, $d \geq 1$, ignoring complex-valued kernels and kernels on abstract sets like those arising in *kernel-based learning* [39, 40]. In most cases on $\Omega = \mathbb{R}^d$ they will be *translation-invariant* of *stationary* in the sense

$$K(x,y) = \Phi(x-y) \text{ for all } x,y \in \mathbb{R}^d \text{ with some } \Phi : \mathbb{R}^d \to \mathbb{R},$$

and in many cases they are additionally *radial* or *isotropic*, i.e.

$$K(x,y) = \Phi(x-y) = \phi(\|x-y\|_2) \text{ for all } x,y \in \mathbb{R}^d \text{ with some } \phi : \mathbb{R} \to \mathbb{R}.$$

Historically, ϕ is then called a *radial basis function* (RBF). Basic references are the books [6, 41, 18], buhmann:2003-1, wendland:2005-1, fasshauer-mccourt:2015-1 and the survey [36], schaback-wendland:2006-1.

2.1 Kernel Properties

(SecBasKerPro) Kernels will always assumed to be *symmetric* and *positive semidefinite*. Then for all finite point sets $X_n = \{x_1, \dots, x_n\}$ the associated *kernel matrices* K_{X_n,X_n} with entries $K(x_j, x_k)$, $1 \leq j \leq n$ have to be positive semidefinite. If all such matrices are nonsingular, we use the terms *positive definiteness* or *strict positive definiteness*. Readers should be aware that old papers call a kernel positive definite when all kernel matrices are positive semidefinite, a very debatable notion.

In Numerical Analysis, kernels provide *trial functions* via *kernel translates* $K(x,\cdot) : \Omega \to \mathbb{R}$, and then positive definiteness of the kernel implies uniqueness of interpolation by kernel translates. On kernel translates, one can define an inner product by the kernel itself via (eqKKK)

$$K(x,y) = (K(x,\cdot), K(y,\cdot))_{\mathscr{H}(K)} \text{ for all } x,\, y \in \Omega. \quad (1)$$

By completion, this leads to a Hilbert space $\mathscr{H}(K)$ called the *native space* of the kernel since [34]. Conversely, each Hilbert space $\mathscr{H}$ of functions on Ω with continuous point evaluations generates a positive semidefinite kernel $K(\mathscr{H})$ on $\Omega\times\Omega$ with (1, eqKKK) and the *reproduction property*

$$f(x) = (f, K(x,\cdot))_{\mathscr{H}(K)} \text{ for all } x\in\Omega,\, f\in\mathscr{H}.$$

Section 3.5(SecExpL2) will show how this extends to Hilbert spaces like L_2 where point evaluations are not continuous.

In Spatial Statistics, kernels arise as *covariance functions* of *random fields*. This connection will be in the focus of Section 3.2(SecExpEqui) , where both viewpoints are shown to be equivalent in the sense that one follows from the other. In Machine Learning, there is a background set $\mathscr{S}$ of objects that have features coded into a *feature map* $\Phi\,:\,\mathscr{S}\to V$ with values $\Phi(s)$, called *feature vectors*, in an inner-product space V. Then the kernel is (eqKPhiKPhi)

$$K(s,t) := (\Phi(s),\Phi(t))_V \text{ for all } s,t\in\mathscr{S}. \tag{2}$$

This looks similar to (1, eqKKK), and therefore kernel translates $K(x,\cdot)$ are sometimes called *feature vectors*. The similar notion of *Fourier features* will be touched in Section 6(SecFouFea) .

2.2 Kernels via Fourier Transforms

(SecClaKer)(SecKerFouTra) The main construction tool for standard translation-invariant kernels on $\mathbb{R}^d$ goes back to Bochner [3] and uses nonnegativity of the Fourier transform, provided it exists. It can be applied to a very large number of cases, surveyed extensively in [16] using hypergeometric functions. Here, we only list the most popular radial kernels, without scaling by constants in the range and the domain:

$$\begin{array}{rcl} \text{Gaussian} & : & \exp(-r^2/2)\\ \text{Matérn} & : & r^\nu K_\nu(r),\ \nu>0\\ \text{Inverse multiquadrics} & : & (1+r^2)^{-\nu},\ \nu>0\\ \text{Wendland } C^2 & : & (1-r)_+^4(1+4r) \end{array}$$

Here, K_ν is the modified Bessel function of second kind. The last case is compactly supported on the unit ball, but only positive definite in dimensions $d\le 3$. Details and plenty of other kernels can be found in [16]. A list of nonstandard kernels was given in [14], but this paper goes far beyond.

2.3 Scale Mixtures

(SecBasScaMix) Many kernels have additional parameters like a *scaling* in the sense

$$K_c(x,y) = \Phi((x-y)/c) = \phi(\|x-y\|_2/c) = \exp(-\|x-y\|_2^2/c^2) \text{ for all } x,y\in\mathbb{R}^d,\ c>0$$

for the Gaussian case. Then one can take linear combinations with positive coefficients of kernels with different scales. Or perform an integration over c with a positive weight function. This technique is called *scale mixture* [1, 20, 37], andrews-mallows:1974-1, gneiting:1997-1,schlather:2010-1. It uses the fact that linear combinations of positive definite kernels stay positive definite if the coefficients are positive. But one can also take derivatives [5], bozzini-et-al:2015-1 or finite differences [4], bozzini-et-al:2013 with respect to scales. This is not covered by the above argument. However, scale mixtures are not in the focus here.

2.4 Expansions

(SecBasExp) On bounded domains, continuous kernels have a *Mercer* expansion [25] based on the eigenvalue problem for integral operator defined by the kernel. By suitable normalization, it is orthonormal in $\mathscr{H}(K)$ and orthogonal in $L_2(\Omega)$. In the nondeterministic literature, it leads to the *Karhunen-Loève expansion* of random fields that have K as their covariance function.

Computationally much cheaper are orthonormal expansions into the *Newton basis* [26]. These are based on a suitably selected infinite point set $X_\infty \subset \Omega$ and have the property

$$N_j(x_k) = 0,\ 1 \leq k < j < \infty.$$

In both cases, there is a kernel expansion (eqKNN)

$$K(x,y) = \sum_{j=1}^{\infty} N_j(x)N_j(y) \text{ for all } x,y \in \Omega. \tag{3}$$

A useful consequence is the series expansion (eqKxx)

$$K(x,x) = \sum_{j=1}^{\infty} N_j(x)^2 \text{ for all } x \in \Omega \tag{4}$$

of the kernel on the diagonal. For the Newton basis, it allows a simple representation (eqPXnx)

$$P_{X_n}(x)^2 = \sum_{j=1}^{n+1} N_j(x)^2 \text{ for all } x \in \Omega \tag{5}$$

of the square of the *Power Function* on a set $X_n = \{x_1, \ldots, x_n\}$. By [33], the latter is defined as the norm of the error of interpolation of functions of native Hilbert space on points of X_n. Since deterministic interpolation coincides with *Kriging* on the nondeterministic side, the square of the Power Function is the variance of the error of the Kriging operator, the best linear unbiased predictor (BLUP).

The partial sum kernel

$$K_m(x,y) = \sum_{j=1}^{m} N_j(x)N_j(y) \text{ for all } x,y \in \Omega$$

approximates the full kernel at the rate of convergence of (4, eqKxx). If points are chosen by the P-greedy method [15], DeMarchi-et-al:2005-1, the rate is asymptotically best possible under all m-term linear approximations [31], santin-haasdonk:2017-1.

As a side effect, using partial sum kernels of order m, kernel matrices based on very large point sets $X_n = \{x_1, \dots, x_n\}$ can be approximated by rank m matrices, and if the Newton basis is used, the full kernel matrices factor into two $m \times n$ matrices via

$$K_m(x_i, x_k) = \sum_{j=1}^{m} N_j(x_i) N_j(x_k) \approx K(x_i, x_k).$$

Ths explains the phenomenon of rank loss [27] of large and badly conditioned kernel matrices.

All of this can be turned upside down. If one starts from an arbitrary set $\{N_j\}_{j\in\mathbb{N}}$ of functions with (eqNjL2)

$$\sum_{j=1}^{\infty} N_j(x)^2 < \infty \text{ for all } x \in \Omega, \tag{6}$$

one gets a positive semidefinite symmetric kernel by (3, eqKNN). This works also on all of $\mathbb{R}^d$, in contrast to Mercer expansions, and one gets (4, eqKxx) and (5, eqPXnx) for free. We shall use this kernel construction technique in Section 3(SecExp) .

3 Kernel Construction by Expansions

(SecExp) Here we construct expansion kernels (3, eqKNN) under the assumption (6, eqNjL2). There are a few first examples in [14], and from there we only mention the kernel

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

based on Chebyshev polynomials in $\Omega = [-1, +1]$. It illustrates how expansion kernels can be synthesized from well-known functions. By $x = \cos(\varphi)$, $y = \cos(\psi)$ it is equal to the 2π-periodic kernel

$$\frac{1}{2} \left(\cos(\sin(\varphi + \psi)) \cdot \exp(\cos(\varphi + \psi)) + \cos(\sin(\varphi - \psi)) \cdot \exp(\cos(\varphi - \psi)) \right)$$

and therefore infinitely differentiable.

3.1 Random Functions

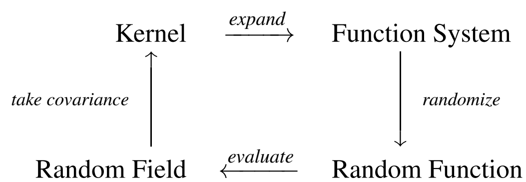
(SecExpRanFun) An expansion kernel (3, eqKNN) with (6, eqNjL2) allows an easy calculation of random functions

$$f_{V,\mathscr{D}}(x) := \sum_{j=1}^{\infty} N_j(x) v_j, \; x \in \Omega$$

by randomizing the coefficients, taking $V = \{v_j\}_{j\in\mathbb{N}}$ as a sequence of realizations v_j from standardized distributions $\mathscr{D}_j$ over $\mathbb{R}$. At each $x \in \Omega$, this yields a random variable $R_{V,\mathscr{D}}(x)$ with realizations $f_{V,\mathscr{D}}(x)$, and the covariance of this random field is the kernel K, while the paths are the above random functions.

3.2 Equivalence

(`SecExpEqui`) The paper [28], `porcu-schaback:2026-1` shows that this approach covers the paths of practically all second-order random fields having the covariance kernel K. Therefore there is a full equivalence of deterministic kernel theory and the theory of random fields, their paths and covariances.



In the above diagram, one can start anywhere and go around to come back to where started. But there are ambiguities that should be explained. Each fixed kernel K can have different expansions (3, `eqKNN`) into function systems $\{N_j\}_{j\in\mathbb{N}}$, and the randomizations can again be different, choosing different probability distributions $\{\mathscr{D}_j\}_{j\in\mathbb{N}}$. Then pointwise evaluation yields different random fields, but all of these have the original kernel as a covariance. Starting from a fixed random field R on Ω, one gets a unique covariance kernel, and the latter can have various expansions into function systems. But [28], `porcu-schaback:2026-1` proves that there are distributions $\{\mathscr{D}_j\}_{j\in\mathbb{N}}$ depending on the expansion chosen such that the randomization produces paths of R whose pointwise evaluations are the random variables $R(x)$ again.

The consequence for further research is that the nondeterministic literature on random fields, their paths and covariances will always have implications for the deterministic kernel theory, and vice versa. Readers from the deterministic community are strongly advised to be familiar with nondeterministic literature on kernels. In particular, this applies to kernel construction methods applied there.

3.3 Expansions of Mollifiers

(`SecExpMol`) Note that the functions N_j in expansion kernels need not be positive definite. Any set of functions with (6, `eqNjL2`) will work, and we are now interested in compactly supported C^∞ functions.

A natural starting point are *mollifiers* used in Real Analysis since [19], `friedrichs:1948-1`. The standard Friedrichs mollifier is the radial C^∞ function (`eqmoll`)

$$\mu(r) := \left\{ \begin{array}{cc} \exp(-1/(1-r^2)) & 0 < r < 1 \\ 0 & 1 \leq r \end{array} \right\}. \tag{7}$$

compactly supported on the unit ball for $r = \|x\|_2$. Translates of such functions can be superimposed with various weights to form a kernel by (3, `eqKNN`). The result will not be translation-invariant, but e.g.

$$K_{c,T}(x,y) := \sum_{n=1}^{\infty} \frac{1}{n^2} \mu(\|x-t_n\|_2/c)\mu(\|y-t_n\|_2/c)$$

is a positive semidefinite C^∞ kernel that vanishes for $\|x-y\|_2 > 2c$, for any choice of point sets $T_\infty = \{t_j\}_{j\in\mathbb{N}}$ and all weights that guarantee (6, `eqNjL2`). This is a first constructive step towards C^∞ compactly supported kernels, see Figure 1(`Figtestexpansionker`), but it deserves to be extended by future research.

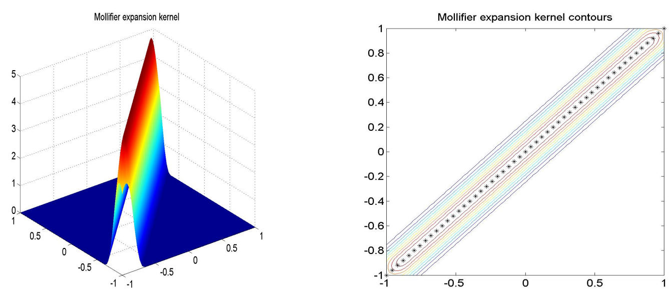


Figure 1: Expansion kernel for 51 regular points t_n in $[-1,+1]$ superimposing translates $\mu_n(x) = \exp(-1/(1-(x-t_n)^2/c^2)_+)$, $c = 0.2$ of standard mollifiers (`Figtestexpansionker`)

3.4 Expansion Kernels of Fundamental Solutions

(`SecExpFun`) In the simplest version of the Method of Fundamental Solutions, see the surveys [7, 8, 9, 23], the trial functions are known solutions u of a homogeneous PDE $Lu = 0$ on a domain $\Omega \subset \mathbb{R}^d$. In the standard case, each function $u_{(x)}$ satisfies $Lu_{(x)} = \delta_x$ on a point x outside the domain. This the standard notion of "*Fundamentallösung*" since [11]. Picking points forming a set $X_\infty = \{x_j\}_{j\in\mathbb{N}}$ on a "fictitious boundary" outside the domain leads to solutions $u_{(x_j)}$ that may need a normalization like $N_j = \nu_j u_{(x_j)},\ j \in \mathbb{N}$ to satisfy (6, `eqNjL2`) and to define a kernel that solves the PDE in both variables. Such kernels define a Hilbert space in which the N_j are orthonormal. Analyzing such spaces seems to be a widely open problem, together with choosing the normalization and proving error bounds on the approximation of boundary values.

As an example, we solve the Potential Equation $\Delta u = 0$ on the unit disc in $\mathbb{R}^2$. The fundamental solutions are of the form $\log(\|x-y\|_2)$ where y should be outside the domain and x inside. The top left plot of Figure 2(`Figfundsoltest`) shows 50 blue points on the "true" boundary and 50 red points on the "fictitious" boundary at the radius 1.3. The red star at $y = (-1.5, -1-5)$ marks the point where the true solution $\log(\|x-y\|_2)$ has a singularity. Fitting its values on the true boundary provides an approximation to the true solution in the top right plot. It does not differ from the true solution up to plot precision. In fact, the error in the lower left plot is smaller than $5\cdot 10^{-7}$, but still showing the typical peaks induced by the close-by "fictitious' points. So far, this is quite standard in the Method of Fundamental Solutions, but the lower right plot shows the associated expansion kernel. Two of the four arguments were fixed at $(0,1)$ to make the kernel plottable, and weights were taken to be one. The figure thus shows

$$K(x,(0,1)) = \sum_{j=1}^{50} \log(\|x-y_j\|_2)\log(\|(0,1)-y_j\|_2).$$

The above approach will run into problems if the extension of the true solution outside

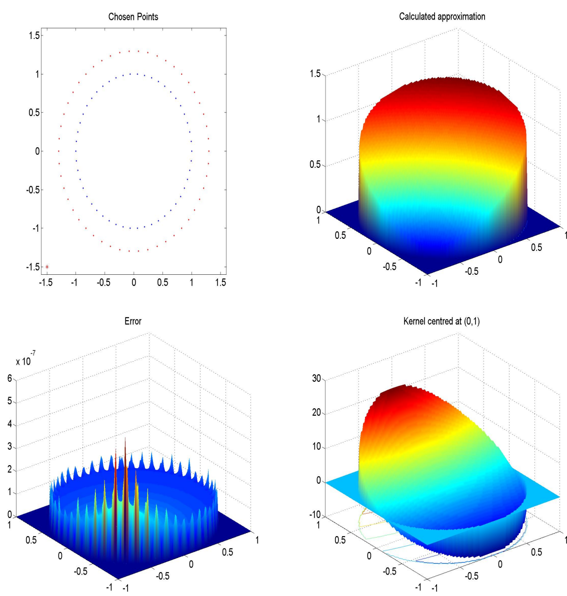


Figure 2: (`Figfundsoltest`)

the domain has singularities between the true and the fictitious boundary. This is even more serious if users take homogeneous solutions without singularities, e.g. harmonic polynomials for solving the homogeneous Dirichlet problem for the Laplacian. But there are other kernels that provide harmonic functions, see [35] in 2D and [22] in 3D, partially using expansion kernels.

3.5 Mollifying Expansion Kernels in L_2

(`SecExpL2`) In spaces $L_2(\mathbb{R}^d)$ or $L_2(\Omega)$ for $\Omega \subset \mathbb{R}^d$, point evaluations of functions are discontinuous, and therefore there are no classical kernels, though the spaces are Hilbert spaces. Anyway, [28] presents a constructive way to work with orthonormal systems in L_2, e.g. Haar wavelets. In addition, there is a theory of weighted Sobolev spaces based on orthonormal L_2 expansions that works quite like a Mercer/Karhunen/Loève expansion of Matérn kernels.

On $\mathbb{R}^d$, there are plenty of "mollifying" operators, e.g. convolutions by smooth functions decaying rapidly at infinity, or convolution with the Friedrichs mollifier of (7, `eqmoll`). Our model case here is the transition from $L_2(\mathbb{R}^d)$ to $W_2^m(\mathbb{R}^d)$ for $m > d$. The functions f in the range have Fourier transforms with behaviour like $(1+\|\omega\|_2^2)^{-m/2}$ by definition of the norm in the range, and they can be obtained from elements $g \in L_2(\mathbb{R}^d)$ by multiplying $\hat{g}(\omega)$ with $(1+\|\omega\|_2^2)^{-m/2}$, which means convolution by a Matérn kernel of order $\nu = (m-d)/2$. This "mollification" of L_2 functions is invertible because the Fourier transform of Matérn kernels does not vanish.

Let $M : L_2(\mathbb{R}^d) \to \mathscr{H}$ be an invertible linear Hilbert space valued operator generated by convolution with a function γ in $L_2(\mathbb{R}^d)$ with nonzero Fourier transform. Then an orthonormal system $U := \{u_j\}_{j\in\mathbb{N}}$ in $L_2(\mathbb{R}^d)$ goes over to a system $\{Mu_j\}_{j\in\mathbb{N}}$ in $\mathscr{H}$, and the range can be equipped by an inner product that makes the system $\{Mu_j\}_{j\in\mathbb{N}}$ orthonormal and generates the expansion kernel

$$K_{U,M}(x,y) = \sum_{j=1}^{\infty} (Mu_j)(x)(Mu_j)(y) \text{ for all } x,y \in \mathbb{R}^d.$$

It satisfies (6, `eqNjL2`) because of $\gamma \in L_2(\mathbb{R}^d)$. Positive semidefiniteness is not required for the u_j of the Mu_j. The expansion form makes the kernel positive semidefinite. This construction principle leads to a plethora of possible kernels, including compactly supported and C^∞ cases.

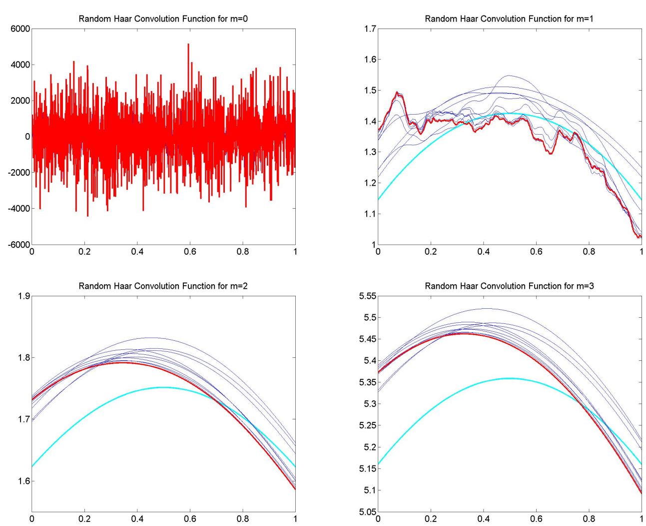


Figure 3: Random Functions based on Haar wavelets on $[0,1]$, smoothed for Sobolev order $m = 0,1,2,3$.

4 Kernels by Convolution

(`SecConv`) If a Fourier-transformable function on $\mathbb{R}^d$ can be convolved with itself (*autocorrelation*), the result is a positive semidefinite kernel, because the Fourier transform is nonnegative. This is another construction technique for kernels.

4.1 Even-Order B-Splines

(`SecConvBSpl`) Starting from (`eqbeta0`)

$$\beta^{(0)}(x) := \mathbf{1}_{[-1/2,+1/2]} = \left\{ \begin{array}{ll} 1 & -1/2 < x < 1/2 \\ 0 & 1/2 \leq |x| \end{array} \right\}, \tag{8}$$

the *symmetric regular B-spline* of order n is the n-fold convolution

$$\beta^{(n)} = \beta^{(n-1)} * \beta^{(0)}, n \geq 1$$

of $\beta^{(0)}$. It is locally a polynomial of order (= degree+1) n, has support in $[-n/2, +n/2]$ and is differentiable of order $n-2$ at its half-integer knots. The Fourier transform is

$$\widehat{\beta^{(n)}}(\omega) = \left(\frac{\sin(\omega/2)}{\omega/2} \right)^{2n},$$

and nonnegative for even n. On $\mathbb{R}^d$, one can take componentwise products like in Section 5(`SecProd`) , the result being a *tensor product B-spline* with quite some background literature [12, 13]. Taking spherical means (this is called the method of *turning bands* in the nondeterministic literature) turns these kernels into compactly supported radial basis functions. The simplest case $n = 2,\ d = 2$ was called *Euclid's hat* [32, 21] but the general case needs further research. Taking infinitely many convolutions of suitably scaled characteristic functions leads to the C^∞ compactly supported *up-function* in Section 4.3(`SecKerSubUp`) .

As another example of a C^∞ compactly supported kernel, we can convolve the standard mollifier of (7, `eqmoll`) with itself, see Figure 4(`FigtestprodkerConv`) . It is another case that needs further study. Applications may use pre-calculated approximations of it.

4.2 Convolutions with Haar Wavelets

(`SecConvHaar`) These can produce smooth compactly supported trial functions that can be superimposed to generate smooth expansion kernels. Here, we sketch a framework for calculating, encouraging readers to provide various examples. The Haar mother wavelet is

$$\psi(t) = \left\{ \begin{array}{rl} 1 & 0 \leq t \leq 1/2 \\ -1 & 1/2 \leq t \leq 1 \\ 0 & \text{otherwise,} \end{array} \right\}$$

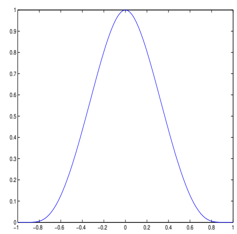


Figure 4: Normalized autocorrelation in 1D of the standard mollifier (`FigtestprodkerConv`)

and it defines the univariate L_2 orthonormal system

$$\psi_{n,k}(t) = 2^n \psi(2^n t - k)$$

where n and k vary in $\mathbb{Z}$ for $L_2(\mathbb{R})$. We consider convolutions with a function $f \in L_2(\mathbb{R})$ that either decays fast towards infinity or is compactly supported. Then the convolution

$$C_{n,k}(x) := \int_{\mathbb{R}} f(x-t)\psi_{n,k}(t)dt$$

shares similar properties and has the shift law

$$\begin{aligned} C_{n,k}(x+2^{-n}) &= \int_{\mathbb{R}} f(x+2^{-n}-t)\psi_{n,k}(t)dt \\ &= \int_{\mathbb{R}} f(x-s)\psi_{n,k-1}(s)ds \\ &= C_{n,k-1}(x) = C_{n,0}(x+k\cdot 2^{-n}),\ k \in \mathbb{Z}. \end{aligned}$$

The basic function is

$$\begin{aligned} C_{n,0}(x) &= \int_{\mathbb{R}} f(x-t)\psi_{n,0}(t)dt \\ &= 2^n \int_x^{x-2^{-n}/2} f(u)du - 2^n \int_{x-2^{-n}/2}^{x-2^{-n}} f(u)du \\ &= -2^n(F(x) - 2F(x-2^{-n}/2) + F(x-2^{-n})) \end{aligned}$$

for the integral

$$F(B) - F(A) = \int_A^B f(u)du.$$

It is smoother than f and has compact support if f has. For instance, if f is a hat function, all $C_{n.0}$ will be compactly supported C^1 piecewise quadratics.

The partial expansion kernel

$$K_n(x,y) = \sum_{k\in\mathbb{Z}} C_{n,k}(x)C_{n,k}(y) = \sum_{k\in\mathbb{Z}} C_{n,0}(x+k\cdot 2^{-n})C_{n,0}(y+k\cdot 2^{-n})$$

satisfies

$$K_n(x+k2^{-n}, y+k2^{-n}) = K_n(x,y) \text{ for all } x, y \in \mathbb{R},\ k \in \mathbb{Z}.$$

It is shift-invariant on the grid $2^{-n}\mathbb{Z}$ and all coarser dyadic grids. The final expansion kernel could be

$$K(x,y) := \sum_{k\in\mathbb{Z}} \sum_{n\in\mathbb{Z}} w_{n,k}^2 C_{n,k}(x) C_{n,k}(y)$$

for suitable weights, and then the functions $w_{n,k}C_{n,k}$ are orthonormal in the Hilbert space connected to that kernel.

Figure 5(FigHaarSplineConv) shows an example. A hat function on $[-0.1, +0.1]$ was convolved with $\psi_{2,0}$. The partial kernel K_2 would be composed of the $\mathbb{Z}/4$ shifts of this function.

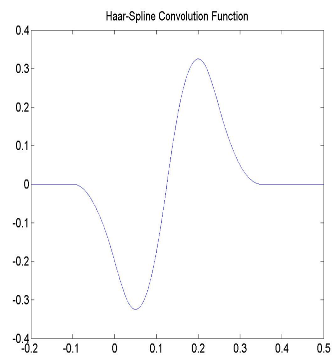


Figure 5: Piecewise quadratic function as convolution of a hat function with a Haar wavelet (FigHaarSplineConv)

4.3 The Up- and Fabius Functions

(SecKerSubUp) We can convolve scaled versions of characteristic functions like (8, eqbeta0), e.g. all functions $\mathbf{1}_{[-2^{-n},+2^{-n}]^d}, n \geq 1$. In 1D, this leads to the *up-function* of Rvachev [30], see also [2]. It can be obtained by nonstationary subdivision [10] and has support in $(0,2)$. While it is not explicitly known, it has a "refinement" equation

$$up'(x) = 2 \cdot up(2x+1) - 2 \cdot up(2x-1),\ x \in \mathbb{R}.$$

Its Fourier transform is (equpFT)

$$up^{\wedge}(\omega) = \prod_{n=1}^{\infty} \frac{\sin(\omega 2^{-n})}{\omega 2^{-n}}, \quad (9)$$

but it is an open problem to find explicit formulas for the function and the Fourier transform. It can be used for C^{∞} trial functions and superimposed by (3, eqKNN) to

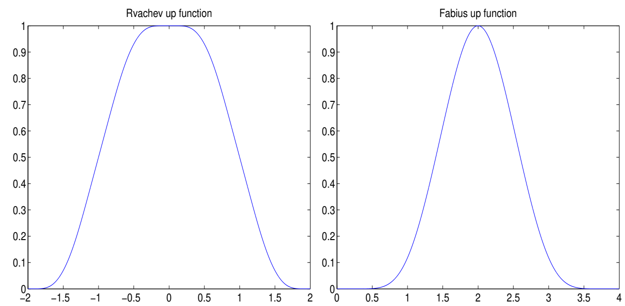


Figure 6: Rvachev and Fabius up-functions on $\mathbb{R}$ (`FigUp`).

form a C^∞ kernel that yields sparse kernel matrices. Since the Fourier transform is not nonnegative, the Ryachev up-function is not positive semidefinite. But by convolution with itself we get a positive definite function that is an infinite convolution of the hat function, the symmetric regular B-spline of order two. It is called the *Fabius up-function* [17] and yields a C^∞ kernel on $\mathbb{R}$ with support in $(0,4)$, see Figure 6(`FigUp`) . It is connected to the *Fabius function* [17], and its Fourier transform is the square of (9, `equpFT`). All of this deserves much more research.

5 Componentwise Products

(`SecProd`) It is well-known [38] that products of positive semidefinite kernels are positive semidefinite again. This allows to use products of univariate kernels to generate multivariate kernels. Unfortunately, this spoils radiality up to a single exception:

Theorem 1. *(`TheProdGau`) The Gaussian is the only nontrivial radial kernel in componentwise product form using a single scalar kernel.*

Proof: If

$$K(x_1,\ldots,x_d) = \prod_{j=1}^{d} k(x_j) = \varphi(\|x\|_2),\ x \in \mathbb{R}^d$$

holds, with $k(0) = 1$ without loss of generality, then

$$K(x_1,0,\ldots,0) = k(x_1) = \varphi(|x_1|),\ x_1 \in \mathbb{R}.$$

Then $k = \varphi$ on all positive real numbers, and we can define

$$f(t) := \varphi(\sqrt{t}) = k(\sqrt{t}),\ t \geq 0$$

to get

$$\prod_{j=1}^{d} f(x_j^2) = f(\|x\|_2^2),\ x \in \mathbb{R}^d.$$

Now $f(x_1^2)f(x_2^2) = f(x_1^2 + x_2^2)$ with $f(0) = 1$ is the functional equation solved only by exponentials $f(t) = \exp(\alpha t)$, and then $k(t) = \varphi(t) = \exp(\alpha t^2)$. $\square$

But the deviation from radiality may be small enough to be tolerated, because the stronger deviation from radiality arises only for small values, see Figure 7(`FigtestprodkerConvProd`) for the product of two univariate kernels from 4(`FigtestprodkerConv`) . The radiality of the product of Fabius up functions is similar, see Figure 8(`FigFabiusUpProd`) . In both cases, contours were plotted for levels 2^{-n}, $n = 1, \ldots, 20$, while the value ot zero was normalized to one.

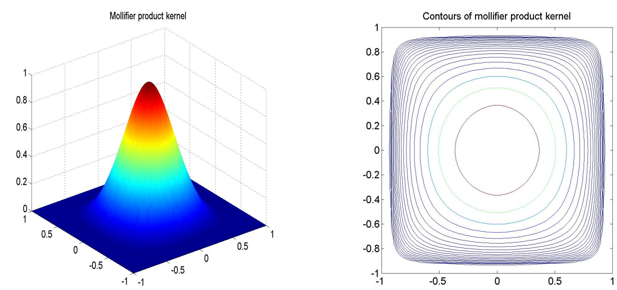


Figure 7: Product kernel of normalized autocorrelation in 1D of the standard mollifier (`FigtestprodkerConvProd`)

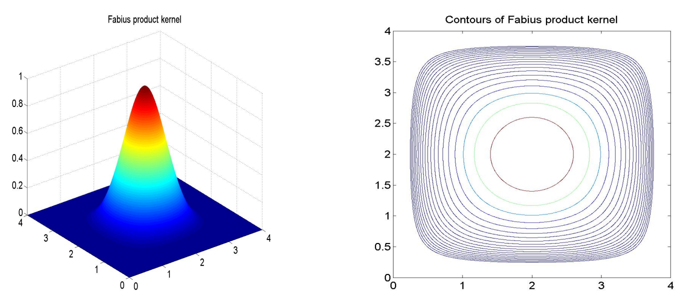


Figure 8: Fabius up-function in product form in $\mathbb{R}^2$ (`FigFabiusUpProd`).

We finally mention another tool for kernel construction. Taking *spherical means* turns non-radial kernels into radial ones. This operation applies similarly to the Fourier transform, and therefore it preserves positive definiteness in the Fourier-transformable and translation-invariant case. It is called the *turning bands* method [24], `matheron:1973-1` used for simulation in Spatial Statistics, and it was used to construct Euclid's hat in [32], `schaback:1995-2`.

6 Fourier Features Kernels

(SecFouFea) Since [29], `rahimi-recht:2007-1`, the Machine Learning community uses *random Fourier features*. A kernel-oriented approach is to take positive square-summable real numbers $c_k,\ k \in \mathbb{N}$ and vectorial frequencies $\omega_k \in \mathbb{R}^d$ to form the translation-invariant Fourier-transformable positive semidefinite kernel

$$K(x,y) = \sum_{k\in\mathbb{N}} c_k \cos(\omega_k^T (x-y)),\ x,\ y \in \mathbb{R}.$$

As an expansion kernel, one can also take

$$K(x,y) = \sum_{k\in\mathbb{N}} (a_k \cos(\omega_k^T x) + b_k \sin(\eta_k^T x))(a_k \cos(\omega_k^T y) + b_k \sin(\eta_k^T y)),\ x,\ y \in \mathbb{R}$$

for arbitrary square-summable coefficients. Such kernels may approximate other kernels well, even when only few coefficients are used, and then large kernel matrices will factorize, like in Section 2.4(`SecBasExp`) . Like others, also this kernel construction technique deserves much further study.

7 Summary and Outlook

(`SecOut`) Many construction techniques for kernels are presented here, including typical examples. But these are only starting points. The goal of the paper is to stimulate research of readers towards constructing new kernels, preferably with C^∞ smoothness and compact support.

Furthermore, even for the examples there is no theoretical analysis yet of the capabilities of the constructed kernels.

Finally, readers are encouraged to introduce new kernels in applications, wherever standard kernels are used now.

Funding

(`SecFun`) The author is very grateful for the support of Hong Kong Baptist University to attend the *Conference on Modern Kernel Methods and Applications* in January 2026.

Acknowledgement

(`SecAck`) There, discussions with various participants helped substantially to make this work possible.

References

[1] D. F. Andrews and C. L. Mallows. Scale mixtures of normal distributions. *J. Roy. Statist. Soc. Ser. B*, 36:99–102, 1974.

[2] J. Arias de Reyna. An infinitely differentiable function with compact support: Definition and properties, 2017. arXiv 1702.05442.

[3] S. Bochner. Monotone Funktionen, Stieltjes Integrale und harmonische Analyse. *Math. Ann.*, 108:378–410, 1933.

[4] M. Bozzini, M. Rossini, and R. Schaback. Generalized Whittle–Matérn and polyharmonic kernels. *Advances in Computational Mathematics*, 39:129–141, 2013. DOI 10.1007/s10444-012-9277-9.

[5] M. Bozzini, M. Rossini, and R. Schaback. Radial kernels via scale derivatives. *Adv. Comput. Math.*, 41:277–291, 2015.

[6] M.D. Buhmann. *Radial Basis Functions, Theory and Implementations*. Cambridge University Press, Cambridge,UK, 2003.

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

[8] C.S. Chen, M.D. Marcozzi, and S. Choi. The method of fundamental solutions and compactly supported radial basis functions: a meshless approach to 3D problems. In *Boundary elements, XXI (Oxford, 1999)*, volume 6 of *Int. Ser. Adv. Bound. Elem.*, pages 561–570. WIT Press, Southampton, 1999.

[9] Alexander H.D. Cheng and Yongxing Hong. An overview of the method of fundamental solutions – solvability, uniqueness, convergence, and stability. *Engineering Analysis with Boundary Elements*, 120:118–152, 2020.

[10] C. Conti, L. Gori, and F. Pitolli. Totally positive functions through nonstationary subdivision schemes. *Journal of Computational and Applied Mathematics*, 200(1):255–265, 2007.

[11] Richard Courant and David Hilbert. *Methoden der Mathematischen Physik*. Springer, Berlin, 1 edition, 1937. Reprinted 1977; English translation available as *Methods of Mathematical Physics*.

[12] C. de Boor. Multivariate piecewise polynomials. *Acta Numerica*, 2:65–109, 1993.

[13] C. de Boor. What is the inverse of a basis? *BIT*, 41:880–890, 2001.

[14] St. De Marchi and R. Schaback. Nonstandard kernels and their applications. *Dolomites Research Notes on Approximations*, 2:16–43, 2009.

[15] St. 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.

[16] X. Emery, E. Porcu, and M. Bevilacqua. Towards Unified Native Spaces in Kernel Methods. *Journal of Machine Learning Research*, 26:1–35, 2025.

[17] Jeroen Fabius. A probabilistic example of a nowhere analytic C^∞ function. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 5:173–174, 1966.

[18] G. Fasshauer and M. McCourt. *Kernel-based Approximation Methods using MATLAB*, volume 19 of *Interdisciplinary Mathematical Sciences*. World Scientific, Singapore, 2015.

[19] K.O. Friedrichs. The identity of weak and strong extensions of differential operators. *Communications on Pure and Applied Mathematics*, 1:361–372, 1948.

[20] T. Gneiting. Normal scale mixtures and dual probability densities. *Journal of Statistical Computation and Simulation*, 59:375–384, 1997.

[21] T. Gneiting. Radial positive definite functions generated by Euclid's hat. *Journal of Multivariate Analysis*, 69(1):88–119, 1999.

[22] Y.C. Hon and R. Schaback. Solving the 3D Laplace equation by meshless collocation via harmonic kernels. *Adv. in Comp. Math.*, pages 1–19, 2013.

[23] Zi-Cai Li, Hung-Tsai Huang, Yimin Wei, and Liping Zhang. *The Method of Fundamental Solutions: Theory and Applications*. Current Natural Sciences. EDP Sciences, Les Ulis, France, 2023.

[24] G. Matheron. The intrinsic random functions and their applications. *Advances in Applied Probability*, 5(3):439–468, 1973.

[25] J. Mercer. Functions of positive and negative type and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society A*, 209:441–458, 1909.

[26] St. Müller and R. Schaback. A Newton basis for kernel spaces. *Journal of Approximation Theory*, 161:645–655, 2009.

[27] A. Noorizadeghan and R. Schaback. Introducing the evaluation condition number: A novel assessment of conditioning in radial basis function methods. *Engineering Analysis with Boundary Elements*, 166, 2024.

[28] E. Porcu and R. Schaback. A computational introduction to random functions. Preprint, to appear in Num. Alg., 2026.

[29] Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. *Advances in Neural Information Processing Systems*, 20:1177–1184, 2007.

[30] V.A. Rvachev. Compactly supported solutions of functional-differential equations and their applications. *Russian Mathematical Survey*, 45:87–120, 1990.

[31] G. Santin and B. Haasdonk. Convergence rate of the data-independent *P*-greedy algorithm in kernel-based approximation. *Dolomites Res. Notes Approx.*, 10(Special Issue):68–78, 2017.

[32] R. Schaback. Creating surfaces from scattered data using radial basis functions. In T. Lyche M. Daehlen and L.L. Schumaker, editors, *Mathematical Methods for Curves and Surfaces*, pages 477–496. Vanderbilt University Press, Nashville, TN, 1995.

[33] R. Schaback. Error estimates and condition numbers for radial basis function interpolation. *Advances in Computational Mathematics*, 3:251–264, 1995.

[34] R. Schaback. Approximation by radial basis functions with finitely many centers. *Constructive Approximation*, 12:331–340, 1996.

[35] R. Schaback. Solving the Laplace equation by meshless collocation using harmonic kernels. *Adv. in Comp. Math.*, 31:457–470, 2009. DOI 10.1007/s10444-008-9078-3.

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

[37] Martin Schlather. Some covariance models based on normal scale mixtures. *Bernoulli*, 16(3):780–797, 2010.

[38] I. J. Schoenberg. Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.*, 44:522–536, 1938.

[39] B. Schölkopf and A.J. Smola. *Learning with Kernels*. MIT Press, Cambridge, 2002.

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

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