Generalized Whittle–Matérn and Polyharmonic Kernels

Mira Bozzini[a], Milvia Rossini[a], Robert Schaback[b,1]

[a] *Dipartimento di Matematica e Applicazioni*
Facolta' di Scienze Matematiche Fisiche e Naturali
Settore MAT/08 - Analisi Numerica
Universitá degli Studi degli Milano–Biccocca
Edificio U5
Via Roberto Cozzi 53
20125 Milano
Italy
[b] *Institut für Numerische und Angewandte Mathematik*
Fakultät für Mathematik und Informatik
Universität Göttingen
Lotzestraße 16-18
D–37073 Göttingen
Germany

Abstract

This paper simultaneously generalizes two standard classes of radial kernels, the polyharmonic kernels related to the differential operator $(-\Delta)^m$ and the Whittle–Matérn kernels related to the differential operator $(-\Delta+I)^m$. This is done by allowing general differential operators of the form $\prod_{j=1}^m(-\Delta+\kappa_j^2 I)$ with nonzero κ_j and calculating their associated kernels. It turns out that they can be explicity given by starting from scaled Whittle–Matérn kernels and taking divided differences with respect to their scale. They are positive definite radial kernels which are reproducing kernels in Hilbert spaces norm–equivalent to $W_2^m(\mathbb{R}^d)$. On the side, we prove that generalized inverse multiquadric kernels of the form $\prod_{j=1}^m(r^2+\kappa_j^2)^{-1}$ are positive definite, and we provide their Fourier transforms. Some numerical examples are added for illustration.

Keywords: radial basis functions, scattered data, kernels, matching pursuit, adaptivity, stability, duality

AMS classification: 41A05,41063, 41065, 65D05, 65D15

1. Introduction

Connections between either splines and Green's functions or radial basis functions and Green's functions have repeatedly been used and made over the

Email addresses: mira.bozzini@unimib.it (Mira Bozzini), milvia.rossini@unimib.it (Milvia Rossini), schaback@math.uni-goettingen.de (Robert Schaback)

[1]Sponsored by an invitation the University of Milano Biccocca in September 2011

past decades (see e.g. [10], [6], [11] and [4] for other references). A very special case are *Polyharmonic splines* introduced in [3] and studied by several authors because of their good properties (see e.g. [5], [7]).

The well–known *polyharmonic kernels* are fundamental solutions of elliptic equations of the form $(-\Delta)^m u = 0$ on $\mathbb{R}^d$. They come as *radial kernels* or *radial basis functions* $\phi_{2m-d}(\|x-y\|_2)$ on $\mathbb{R}^d$ in dimension–dependent form

$$\phi_{2m-d}(r) = \left\{ \begin{array}{ll} (-1)^{\lceil m-d/2 \rceil} r^{2m-d} & 2m-d \notin 2\mathbb{Z} \\ (-1)^{1+m-d/2} r^{2m-d} \log r & 2m-d \in 2\mathbb{Z} \end{array} \right\}$$

as *powers* or *thin–plate splines*, if

$$2m - d > 0 \tag{1}$$

holds. They have generalized Fourier transforms $\|\omega\|_2^{-2m}$ on $\mathbb{R}^d$ up to positive multiplicative constants, and they are conditionally positive definite of orders $\lceil m - d/2 \rceil$ and $1 + m - d/2$, respectively. See the monographs [12] and [1] for details concerning these notions.

Another prominent case under the condition (1) are the positive definite radial *Whittle–Matérn–Sobolev* kernels

$$\psi_{2m-d}(r) = r^{m-d/2} K_{m-d/2}(r)$$

involving the Bessel function K_ν of the third kind. They are reproducing kernels of Sobolev spaces $W_2^m(\mathbb{R}^d)$ and have Fourier transforms $(\|\omega\|_2^2 + 1)^{-m}$ up to positive factors, and thus are related to elliptic differential operators $(-\Delta + I)^m$ instead of the operators $(-\Delta)^m$ related to polyharmonic kernels.

We shall generalize both classes of kernels simultanously by considering fundamental solutions of more general elliptic equations of the form

$$L\,u := \prod_{j=1}^{m} (-\Delta + \kappa_j^2 I) u = 0 \tag{2}$$

with positive real numbers κ_j^2, $1 \leq j \leq m$ and assuming (1). These differential operators have positive radial Fourier transforms

$$\hat{L}(\omega) = \prod_{j=1}^{m} (\|\omega\|_2^2 + \kappa_j^2)$$

within bounds of the form

$$0 < \prod_{j=1}^{m} \kappa_j^2 \leq \hat{L}(\omega) \leq C \|\omega\|_2^{2m} \text{ for all } \omega \in \mathbb{R}^d.$$

Thus their fundamental solutions are inverse Fourier transforms of

$$\prod_{j=1}^{m} (\|\omega\|_2^2 + \kappa_j^2)^{-1}, \tag{3}$$

and they exist classically as continuous radial kernels. Our main task will be to present explicit formulas for them.

If all κ_j are equal and positive, we have to find the inverse Fourier transform of $(\|\omega\|_2^2+\kappa^2)^{-m}$, and by standard rules of Fourier transforms, the result is the *scaled* Whittle–Matérn–Sobolev kernel

$$S_{m,d,\kappa}(x,y) = \frac{2^{1-m}}{(m-1)!}\left(\frac{\|x-y\|_2}{\kappa}\right)^{m-d/2} K_{m-d/2}(\kappa\|x-y\|_2) \qquad (4)$$

for $2m > d$, $\kappa > 0$ and all $x, y \in \mathbb{R}^d$. It is a standard technique to prove conditional positive definiteness of polyharmonic kernels by starting from (4) and considering the limit $\kappa \to 0$ with appropriate multiplicative normalization.

In this paper, we shall explicitly calculate the inverse Fourier transform of (3) for general positive values of the κ_j. Since (3) is a product, this is equivalent to calculating a convolution of polyharmonic kernels S_{1,d,κ_j} with Fourier transforms $(\|\omega\|^2+\kappa_j^2)^{-1}$, which seems to be a highly nontrivial task. But we shall prove that the result can be written explicitly as a divided difference with respect to the scale parameter κ of $S_{1,d,\kappa}$. This suggests that we obtain a large new class of kernels, but it will turn out that the final result can be written as a standard Whittle–Matérn–Sobolev kernel (4) of a certain scale κ dependent on the κ_j.

We close the paper by a few numerical examples.

2. Basic Results

We shall assume (1) throughout, and treat the product in (3) by the following identity.

Lemma 2.1. *Let the m–th divided difference of a multivariate function u with respect to the variable z be written as $[t_1,\ldots,t_m]_z u(z,\ldots)$. Then the formula*

$$(-1)^{m-1}\prod_{j=1}^{m}(s+t_j)^{-1} = [t_1,\ldots,t_m]_z(s+z)^{-1}$$

holds for all $s \geq 0$ and all positive $t_1,\ldots,t_m$.

Proof: For distinct $t_1,\ldots,t_m$ the result follows by induction on m, and for coalescing t_j it follows by taking appropriate derivatives. By standard arguments for divided differences with partially coalescing arguments, it follows in general. □

Our central result is

Theorem 2.1.

1. *The inverse d–variate Fourier transform of (3) for arbitrary positive numbers $\kappa_1,\ldots,\kappa_m$ is the positive definite radial kernel*

$$\phi(r) = 2^{-m+1}(-1)^{m-1}[\kappa_1^2/2,\ldots,\kappa_m^2/2]_z\left(\frac{r}{\sqrt{2z}}\right)^{1-d/2} K_{1-d/2}(r\sqrt{2z}). \qquad (5)$$

2. *For a special value κ between $\kappa_1, \dots, \kappa_m$ it takes the form (4), i.e. it is a specially scaled Whittle–Matérn–Sobolev kernel.*
3. *The fundamental solution of a differential operator of the form (2) coincides with the fundamental solution of a differential operator $(-\Delta + \kappa^2 I)^m$ for a special value κ between $\kappa_1, \dots, \kappa_m$, i.e. it takes the form (4).*

Before we prove this, some remarks should be made.

- It is well–known that linear combinations of positive definite kernels with positive coefficients yield positive definite kernels, but divided differences have alternating factors. In spite of that, (5) is a positive definite kernel.

- A direct way to calculate the kernel would be to use convolution implied by the factorization of the Fourier transform (3). But our approach gives an explicit formula for the result of the convolution.

- For $d \geq 2$, the kernels involved in the divided differences will have singularities at zero. But the divided differences cancel these, and the resulting kernel is well–defined at zero, as indicated by the third assertion.

- All scaled Whittle–Matérn–Sobolev kernels (4) are reproducing kernels in Hilbert spaces that are norm–equivalent to Sobolev space $W_2^m(\mathbb{R}^d)$, the norm equivalence constants being dependent on m, d and the scale κ. Our approach generates new variations of "multiply scaled" Sobolev–type spaces, but the third assertion proves that we do not leave the set of norm–equivalent spaces to $W_2^m(\mathbb{R}^d)$.

- The above technique can be put upside–down, proving that generalized inverse multiquadrics of the form (3) are positive definite, their Fourier transforms being of the form (5).

- After this, one could possibly use the standard technique to go to non–inverse multiquadrics by analytic continuation and allowing generalized Fourier transforms, but we leave this to future research.

- It would be interesting to see what happens if part of the κ_j tend to zero. This would possibly yield new conditionally positive definite kernels that are products of polyharmonic and Whittle–Matérn–Sobolev kernels, but we again leave this to future research.

3. Proofs

The connection between items 1 and 2 of Theorem 2.1 suggest that we need to take derivatives of kernels $S_{m,d,\kappa}$ of (4) with respect to the scale parameter κ. This can be done by a useful technique for handling derivatives of radial kernels, as summarized in [2] and given as detailed MATLAB programming instructions in [8]. It could be bypassed for proving the first part of the theorem via Lemma

2.1 and standard Fourier transforms, but for simplicity of presentation, we shall use it throughout.

In short, a radial kernel

$$K(x,y) = \Phi(x-y) = \phi(\|x-y\|_2)$$

on $\mathbb{R}^d$ can be rewritten in the form

$$\phi(r) = f(r^2/2) \text{ for all } r \geq 0,$$

and then its d–variate Fourier transform $\hat{\Phi}$ can be recovered analogously from

$$\hat{\Phi}(\omega) = g(\|\omega\|_2^2/2) \text{ for all } \omega \in \mathbb{R}^d$$

with the function

$$g(s) = \int_0^\infty f(t)t^\nu H_\nu(ts)dt$$

and $\nu = (d-2)/2 > -1$. This is a reformulation of the standard Hankel transform of radial functions (see [9]), based on

$$\left(\frac{z}{2}\right)^{-\nu} J_\nu(z) = H_\nu(z^2/4)$$

with

$$H_\nu(t) := \sum_{n=1}^{\infty} \frac{(-t)^n}{n!\Gamma(n+\nu+1)}.$$

This reformulation of the Fourier transform has a lot of advantages, see [8] and [2]. In particular, the inverse Fourier transform is exactly the same, and there are handy rules for derivatives of kernels in this form. Furthermore, it allows Fourier transforms in spaces of fractal dimension, and it allows to take fractional derivatives [9].

In particular, the scaled Whittle–Matérn–Sobolev kernels of (4) take the form

$$f_{m,d,\kappa}(t) = \frac{2^{1-m}}{(m-1)!}\kappa^{d/2-m}(2t)^{m/2-d/4}K_{m-d/2}(\kappa\sqrt{2t})$$

after transformation $t = r^2/2 = \|x-y\|_2^2/2$. Their d–variate Fourier transforms are $\left(\kappa^2 + \|\omega\|_2^2\right)^{-m}$, and we transform them by $s = \|\omega\|_2^2/2$ to get

$$\left(\kappa^2 + \|\omega\|_2^2\right)^{-m} \quad = \quad 2^{-m}\left(\frac{\kappa^2}{2} + s\right)^{-m} =: g_{m,d,\kappa}(s) \tag{6}$$

such that the Fourier transform relations

$$g_{m,d,\kappa}(s) = \int_0^\infty f_{m,d,\kappa}(t)t^\nu H_\nu(ts)dt, \ f_{m,d,\kappa}(t) = \int_0^\infty g_{m,d,\kappa}(s)s^\nu H_\nu(ts)ds$$

hold for $\nu = (d-2)/2$.

We exploit the standard transformation to rewrite (3) in the form

$$\prod_{j=1}^m (r^2+\kappa_j^2)^{-1} \quad = \quad 2^{-m}\prod_{j=1}^m (r^2/2+\kappa_j^2/2)^{-1}$$

and consider

$$g(s) := 2^{-m}\prod_{j=1}^m (s+\kappa_j^2/2)^{-1}.$$

We want to find an explicit formula for

$$f(t) = \int_0^\infty g(s)s^\nu H_\nu(ts)ds = 2^{-m}\int_0^\infty s^\nu H_\nu(ts)\prod_{j=1}^m (s+\kappa_j^2/2)^{-1}ds \tag{7}$$

for $\nu=(d-2)/2>-1$, and we shall use Lemma 2.1. Then

$$\begin{aligned} f(t) &= 2^{-m}(-1)^{m-1}\int_0^\infty s^\nu H_\nu(ts)[\kappa_1^2/2,\ldots,\kappa_m^2/2]_z(s+z)^{-1}ds \\ &= 2^{-m}(-1)^{m-1}[\kappa_1^2/2,\ldots,\kappa_m^2/2]_z \underbrace{\int_0^\infty (s+z)^{-1}s^\nu H_\nu(ts)ds}_{=:g_{\nu,z}(t)} \end{aligned}$$

and we use that in the Sobolev case we have

$$g_{1,d,\sqrt{2z}}(s) = \frac{1}{2}(z+s)^{-1} = g_{1,d,\sqrt{2s}}(z).$$

Then

$$\begin{aligned} g_{\nu,z}(t) &= \int_0^\infty (s+z)^{-1}s^\nu H_\nu(ts)ds \\ &= 2\int_0^\infty g_{1,d,\sqrt{2z}}(s)s^\nu H_\nu(ts)ds \\ &= 2f_{1,d,\sqrt{2z}}(t) \\ &= 2\left(\frac{\sqrt{2z}}{\sqrt{2t}}\right)^\nu K_\nu(\sqrt{2z}\sqrt{2t}) \end{aligned}$$

holds, but the integral in the first line needs

$$-1<\nu<1/2 \text{ or } 0<d<4$$

in order to exist classically. However, we shall later use the more general final line which extends the first line by analytic continuation from the above values for ν. Then we arrive at

$$\begin{aligned} f(t) &= 2^{-m}(-1)^{m-1}[\kappa_1^2/2,\ldots,\kappa_m^2/2]_z g_{\nu,z}(t) \\ &= 2^{-m+1}(-1)^{m-1}[\kappa_1^2/2,\ldots,\kappa_m^2/2]_z f_{1,d,\sqrt{2z}}(t). \end{aligned}$$

We now invert the transformation $t=r^2/2$ and get

$$\begin{aligned} & \phi(r) \\ =\ & f(r^2/2) \\ =\ & 2^{-m+1}(-1)^{m-1}[\kappa_1^2/2,\ldots,\kappa_m^2/2]_z f_{1,d,\sqrt{2z}}(r^2/2) \\ =\ & 2^{-m+1}(-1)^{m-1}[\kappa_1^2/2,\ldots,\kappa_m^2/2]_z \left(\frac{r}{\sqrt{2z}}\right)^{1-d/2} K_{1-d/2}(r\sqrt{2z}) \end{aligned}$$

which is the first assertion of Theorem 2.1.

The left–hand side, rewritten in the form (7), exists for all m and $d \geq 1$ with $m > d/2$, as well as the right–hand side. Since the other parameters are fixed and positive, both sides are analytic functions of the formal parameter d. Since both sides allow all complex values of d in an open subset of $\mathbb{C}$ containing the real interval $(1/2, 2m)$, we get the above identity for $m > d/2$ by analytic continuation with respect to d.

We now want to exploit the fact that an m–th divided difference coincides with an m–th derivative at some point and up to the factor $m!$. This requires that we take the derivatives with respect to z of

$$u_{\nu,t}(z) := \left(\frac{\sqrt{2z}}{\sqrt{2t}}\right)^{\nu} K_\nu\left(\sqrt{2z}\sqrt{2t}\right)$$

for $\nu = (d-2)/2$, using $K_\nu = K_{-\nu}$ and $t = r^2/2$. We set $s := 2zt$, i.e. $\sqrt{2s} = \sqrt{2z}\sqrt{2t}$ and rewrite the expression as

$$u_{\nu,t}(z) := \left(\frac{\sqrt{2s}}{2t}\right)^{\nu} K_\nu\left(\sqrt{2s}\right) = (2t)^{-\nu} \underbrace{\left(\sqrt{2s}\right)^{\nu} K_\nu\left(\sqrt{2s}\right)}_{=:y_\nu(s)}.$$

From [8] we get the derivative relation

$$y_\nu(s)' = -y_{\nu-1}(s),$$

and then

$$\begin{aligned}
\frac{d^k}{dz^k} u_{\nu,t}(z) &= (2t)^{-\nu}(2t)^k \frac{d^k}{ds^k} y_\nu(s) \\
&= (2t)^{-\nu}(2t)^k(-1)^k y_{\nu-k}(s) \\
&= (2t)^{k-\nu}(-1)^k \left(\sqrt{2s}\right)^{\nu-k} K_{\nu-k}\left(\sqrt{2s}\right) \\
&= (2t)^{k-\nu}(-1)^k \left(\sqrt{2z}\sqrt{2t}\right)^{\nu-k} K_{\nu-k}\left(\sqrt{2z}\sqrt{2t}\right) \\
&= (-1)^k \left(\frac{\sqrt{2t}}{\sqrt{2z}}\right)^{k-\nu} K_{k-\nu}\left(\sqrt{2z}\sqrt{2t}\right)
\end{aligned}$$

leading because of $k - \nu = m - d/2$ to

$$\frac{d^k}{dz^k} u_{\nu,t}(z) = \frac{2^{-m+1}}{(m-1)!}\left(\frac{r}{\sqrt{2z}}\right)^{m-d/2} K_{m-d/2}\left(r\sqrt{2z}\right). \tag{8}$$

Taking the divided difference in (5) at values $z_j = \kappa_j^2/2$, we get the above function at some place $z = \kappa^2/2$ for some κ between the values $\kappa_1, \ldots, \kappa_m$. This means that we get a suitably scaled version of the standard kernel generating Sobolev space $W_2^m(\mathbb{R}^d)$, proving the second assertion of Theorem 2.1. The third assertion is a reformulation of the second, in terms of fundamental solutions of differential operators. □

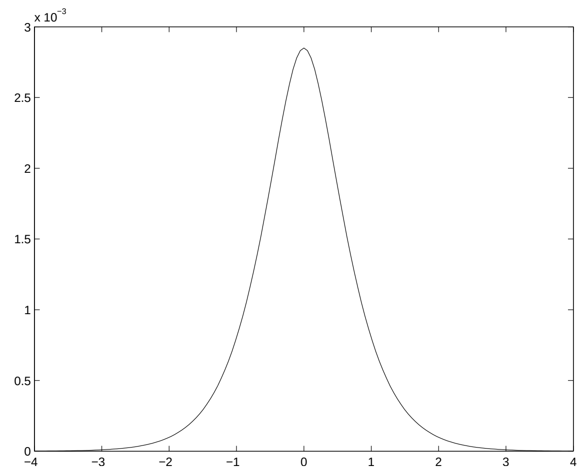

Figure 1: $m = 3$, $\kappa_1 = 2.3$, $\kappa_2 = 3$, $\kappa_3 = 4$.

4. Numerical Examples

In this section we show the 1D radial plots of (5) for $m = 3$ and $d = 2$. In Fig. 1 we have set $\kappa_1 = 2.3$, $\kappa_2 = 3$, $\kappa_3 = 4$, while in Fig. 2 we have $\kappa_1 = 9$, $\kappa_2 = 10$, and $\kappa_3 = 15$. We can observe the tension effect as the values of κ_i grow.

We have also computed experimentally the scales κ for which the considered functions agree with the form (4). In the first case $\kappa \approx 3.0516$ and with this value the maximum absolute error (computed on 151 equispaced points) between (5) and (8) is $3.3e - 5$. In the second case $\kappa = 11.1524$ provides a maximum absolute error equal to $1.8e - 7$.

[1] M. D. Buhmann. *Radial Basis Functions.* Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2004.

[2] St. De Marchi and R. Schaback. Nonstandard kernels and their applications. Preprint Göttingen, to appear, 2010.

[3] J. Duchon. Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces. *Rev. Française Automat. Informat. Rech. Opér. Anal. Numer.*, 10:5–12, 1976.

[4] G. F. Fasshauer. Green's functions: Taking another look at kernel approximation, radial basis functions and splines. Approximation Theory XIII: San Antonio 2010, to appear.

[5] A. Iske. *Multiresolution methods in scattered data modelling*, volume 37 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, Berlin, 2004.

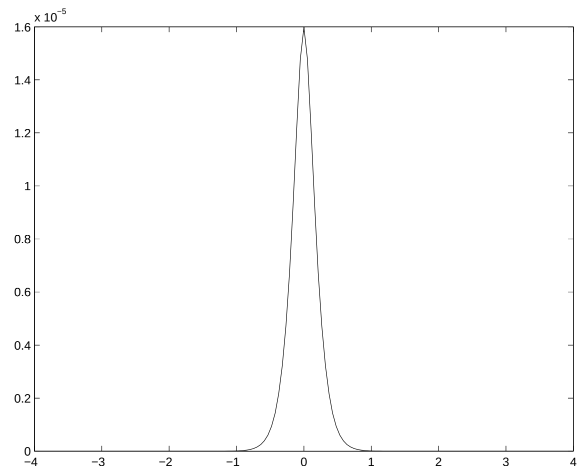

Figure 2: $m = 3$, $\kappa_1 = 9$, $\kappa_2 = 10$, $\kappa_3 = 15$.

[6] J. Kybic, T. Blu, and M. Unser. Generalized sampling: A variational approach– Part I: Theory. *IEEE Trans. Signal Proc. Networks*, 50:1965–1976, 2002.

[7] C. Rabut and M. Rossini. Polyharmonic multiresolution analysis: an overview and some new results. *Numer. Algorithms*, 48:135–160, 2008.

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

[9] R. Schaback and Z. Wu. Operators on radial basis functions. *J. Comp. Appl. Math.*, 73:257–270, 1996.

[10] L.L. Schumaker. *Spline Functions: Basic Theory.* Wiley–Interscience, 1981.

[11] M. Unser and T. Blu. Cardinal exponential splines: Part I – Theory and filtering algorithms. *IEEE Trans. Signal Proc. Networks*, 53:1425–1438, 2005.

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