Kernels via Scale Derivatives

Mira Bozzini, Milvia Rossini,
Robert Schaback, and Elena Volontè

Draft, June 17, 2013

Abstract

We generate various new kernels by taking derivatives of known kernels with respect to scale. This is different from the well-known scale mixtures used before. The resulting kernels are analyzed theoretically to some extent, and a few illustrations are provided. On the side, we provide a simple recipe that explicitly constructs new kernels from the negative Laplacian of known kernels.

1 Introduction

We consider radial *kernels* or *radial basis functions*

$$K(x,y) = \Phi(x-y) = \phi(\|x-y\|_2) \text{ for all } x,y \in \mathbf{R}^d$$

on $\mathbf{R}^d$ with a scalar function $\phi : [0,\infty) \to \mathbf{R}$. If d–variate Fourier transformability is assumed, the d–variate Fourier transform $\hat{\Phi}$ is radial again, and coincides with the Hankel transform

$$\hat{\Phi}(\omega) = \|\omega\|_2^{-(d-2)/2} \int_0^\infty \phi(r) r^{d/2} J_{(d-2)/2}(r \cdot \|\omega\|_2) dr$$

involving the Bessel function J_ν.

It is convenient [10] to rewrite everything in the new variable $s = r^2/2$, starting from rewriting the kernel as

$$\Phi(x-y) = \phi(\|x-y\|_2) =: f(\|x-y\|_2^2/2)$$

and introducing the function H_ν for $\nu := (d-2)/2$ via

$$\left(\frac{z}{2}\right)^{-\nu} J_\nu(z) =: H_\nu(z^2/4) = \sum_{k=0}^\infty \frac{(-z^2/4)^k}{k!\Gamma(k+\nu+1)}$$

to arrive at $\hat{\Phi}(\omega) = \hat{f}(\|\omega\|_2^2/2)$ with

$$\hat{f}(t) := \int_0^\infty f(s)s^\nu H_\nu(st)ds \tag{1}$$

and the inverse transform

$$f(s) = \int_0^\infty \hat{f}(t)t^\nu H_\nu(ts)dt \tag{2}$$

if the standard d–variate Fourier transform works both ways. In what follows, we shall understand Fourier transforms $f \leftrightarrow \hat{f}$ this way, observing the hidden dependence on the dimension via $\nu = (d-2)/2$.

Introducing a scaling with a positive real number z yields

$$\widehat{f(\cdot z)}(u) = z^{-\nu-1}\widehat{f(\cdot)}(u/z) \tag{3}$$

by elementary calculations, and similarly

$$\widehat{f(\cdot)}(s) = z^{\nu+1}\widehat{f(\cdot z)}(sz). \tag{4}$$

To get new interesting kernels, we consider functionals $\lambda = \lambda^u$ that act linearly on functions with respect to a variable u. If we can commute the action of the functional with the integral, we get the identity

$$(\lambda^z f(\cdot z))^\wedge(u) = \lambda^z(z^{-\nu-1}\hat{f}(u/z)) \tag{5}$$

that we shall use throughout the paper to get new radial kernels. In case of a functional acting like

$$\lambda^z f(z) = \int_I w(z)f(z)dt$$

with a positive weight function w, this is called a *scale mixture* in papers oriented on stochastic processes and geostatistics, e.g. [1, 6, 11] and plenty of others.

In (5), we can vary λ and f, and we shall do this in the following sections. In particular, we shall let f vary through several standard classes of radial kernels and let λ be a differentiation. In many cases, we get new and interesting kernels. If their Fourier transforms are nonnegative, they will be positive semidefinite. If the Fourier transforms additionally are positive on a set of positive Lebesgue measure, the kernels will be (strictly) positive definite. We shall give a few illustrations at certain places. A special case using a divided difference was treated in [3], but the general derivative case is considered here.

After an extension of (3) to generalized functions, section 3 specializes to taking λ as a derivative, and then section 4 contains examples for the Gaussian, multiquadrics, and Whittle–Matérn kernels. A final section shows that the transition $\Phi \mapsto -\Delta\Phi$ can be efficiently implemented, since it just generates linear combinations of kernels of the same family.

2 Generalized Fourier Transforms

For later use with certain conditionally positive definite radial kernels, we need the notion of *generalized* radial Fourier transforms connecting $f \leftrightarrow \hat{f}$ not by the above relations (1) and (2), but rather by

$$\int_0^\infty f(s)s^\nu v(s)ds = \int_0^\infty \hat{f}(t)t^\nu \hat{v}(t)dt \tag{6}$$

for all arbitrarily smooth test functions v that have arbitrarily fast decay at zero and infinity. Then we can introduce a scaling to get

$$\begin{aligned}
\int_0^\infty f(sz)s^\nu v(s)ds &= \int_0^\infty f(t)t^\nu z^{-\nu-1} v(t/z)dt \\
&= \int_0^\infty f(t)t^\nu z^{-\nu-1} (v(\cdot/z))(t)dt \\
&= \int_0^\infty \hat{f}(t)t^\nu z^{-\nu-1} \widehat{v(\cdot/z)}(t)dt \\
&= \int_0^\infty \hat{f}(t)t^\nu z^{-\nu-1} z^{\nu+1} \hat{v}(tz)dt \\
&= \int_0^\infty \hat{f}(t)t^\nu \hat{v}(tz)dt \\
&= \int_0^\infty \hat{f}(s/z)s^\nu z^{-\nu-1} \hat{v}(s)dt
\end{aligned}$$

where we used the standard scaling relations only on the test functions. Similarly, if we can pull the functional into the first integral,

$$\begin{aligned}
\int_0^\infty \lambda^z f(sz)s^\nu \hat{v}(s)ds &= \int_0^\infty f(t)t^\nu \lambda^z z^{-\nu-1} (\hat{v}(\cdot/z))(t)dt \\
&= \int_0^\infty f(t)t^\nu (\lambda^z v(\cdot z))^\wedge(t)dt \\
&= \int_0^\infty f(t)t^\nu \lambda^z (v(\cdot z))^\wedge(t)dt \\
&= \int_0^\infty \hat{f}(t)t^\nu \lambda^z (v(\cdot z))(t)dt \\
&= \int_0^\infty \lambda^z \hat{f}(t)t^\nu v(zt)dt \\
&= \int_0^\infty \lambda^z (z^{-\nu-1}\hat{f}(s/z))s^\nu v(s)dt,
\end{aligned}$$

and again we used the scaling relations only on the test function. This implies

Theorem 1. *The relations (3), (4), and (5) also hold for generalized Fourier transforms whenever the defining integrals (6) exist and if the functional λ commutes with the integrals.* □

3 Derivatives

As a linear functional λ that hopefully commutes with the integration, we take the k–th derivative

$$\lambda^z f(z) = \frac{d^k}{dz^k} f(z).$$

So, combining the relations (3) and (5) yields

$$\left(\frac{d^k}{dz^k} f(\cdot z)\right)^\wedge (u) = \left(f^{(k)}(\cdot z)(\cdot)^k\right)^\wedge (u) = \frac{d^k}{dz^k}\left(z^{-\nu-1}\widehat{f(\cdot)}(u/z)\right) \quad (7)$$

and for $k = 1$ we have the simple relation

$$\left(tf'(tz)\right)^{\wedge,t}(u) = \frac{d}{dz}\left(z^{-\nu-1}\widehat{f(\cdot)}(u/z)\right) \quad (8)$$

which specializes for $z = 1$ to

$$\left(tf'(t)\right)^\wedge (u) = \frac{d}{dz}_{|z=1}\left(z^{-\nu-1}\widehat{f(\cdot)}(u/z)\right) \quad (9)$$

where we used the notation $^{\wedge,t}$ to indicate that the transform acts with respect to t, if another variable is present. By Theorem 1, these relations hold for standard and generalized Fourier transforms, as long as integrals exist and the functional commutes with the integral.

In the examples of the next section, we take the first derivative with respect to scaling for a certain family of kernels. The result usually will be another positive kernel, but the interesting *new* kernel will be its Fourier transform.

4 Examples

4.1 Gaussian

For the *Gaussian* $K(x - y) = \exp(-\|x - y\|_2^2/2)$ we have

$$f(t) = \exp(-t) \quad (10)$$

and know that Fourier transforms of any dimension let the Gaussian invariant. All formulas hold without problems due to the exponential decay of f at infinity. Then (8) yields

$$\begin{array}{rcl} (tf'(tz))^{\wedge,t}(u) & = & (-t\exp(-tz))^{\wedge,t}(u) \\ & = & \frac{d}{dz}(z^{-\nu-1}\widehat{f(\cdot)}(u/z)) \\ & = & \frac{d}{dz}\left(z^{-\nu-1}\exp\left(-\frac{u}{z}\right)\right) \\ & = & (-\nu-1)z^{-\nu-2}\exp\left(-\frac{u}{z}\right) + z^{-\nu-1}\left(\frac{u}{z^2}\right)\exp\left(-\frac{u}{z}\right) \\ & = & z^{-\nu-3}\exp\left(-\frac{u}{z}\right)((-\nu-1)z+u). \end{array}$$

In the special case $z = 1$ we get

$$(-t\exp(-t))^{\wedge}(u) = \exp(-u)(u - \nu - 1). \tag{11}$$

Theorem 2. *The radial kernel*

$$\widehat{\phi}(\omega) = \exp\left(-\frac{\|\omega\|^2}{2}\right)\left(\frac{d}{2} - \frac{\|\omega\|^2}{2}\right) \tag{12}$$

of Figure 1 is strictly positive definite in $\mathbf{R}^d$.

Proof: By (11) for $\nu = (d-2)/2$, the strictly positive and bounded radial kernel

$$\phi(x) = -\left(-\frac{\|x\|^2}{2}\right)\exp\left(-\frac{\|x\|^2}{2}\right)$$

is the d-variate Fourier transform of the continuous and absolutely integrable function (12). □

This argument used the well–known

Theorem 3. *[12, Theorem 6.11, p. 74] Let* Φ *be a continuous function in* $L^1(\mathbf{R}^d)$. Φ *is strictly positive definite if and only if* Φ *is bounded and its Fourier transform is non-negative and not identically equal to zero.*

The analog for generalized Fourier transforms is

Theorem 4. *[12, Theorem 8.12, p. 105] Let* Φ *be a continuous and slowly increasing function in* $\mathbf{R}^d$ *with a generalized Fourier transform* $\hat{\Phi}$ *of order* m *which is continuous in* $\mathbf{R}^d \setminus \{0\}$. *Then* Φ *is strictly conditionally positive definite of order* m *if and only if* $\hat{\Phi}$ *is nonnegative and nonvanishing.*

4.2 Multiquadrics

Another strictly positive definite radial kernel is the *inverse multiquadric*

$$\begin{aligned} K(x) &= (1+\|x\|^2)^{-\beta} = \left(1 + 2\tfrac{\|x\|^2}{2}\right)^{-\beta} \\ f(t) &= (1+2t)^{-\beta} \end{aligned}$$

with $\beta > d/2$. It has the d-variate Fourier transform

$$\begin{aligned} \widehat{K}(\omega) &= \frac{2^{1-\beta}}{\Gamma(\beta)}\|\omega\|^{\beta-\frac{d}{2}}K_{-\beta+\frac{d}{2}}(\|\omega\|) = \\ &= \frac{2^{1-\frac{\beta}{2}-\frac{d}{4}}}{\Gamma(\beta)}\left(\frac{\|\omega\|^2}{2}\right)^{\frac{\beta}{2}-\frac{d}{4}} K_{\beta-\frac{d}{2}}\left(2^{\frac{1}{2}}\left(\frac{\|\omega\|^2}{2}\right)^{\frac{1}{2}}\right), \end{aligned}$$

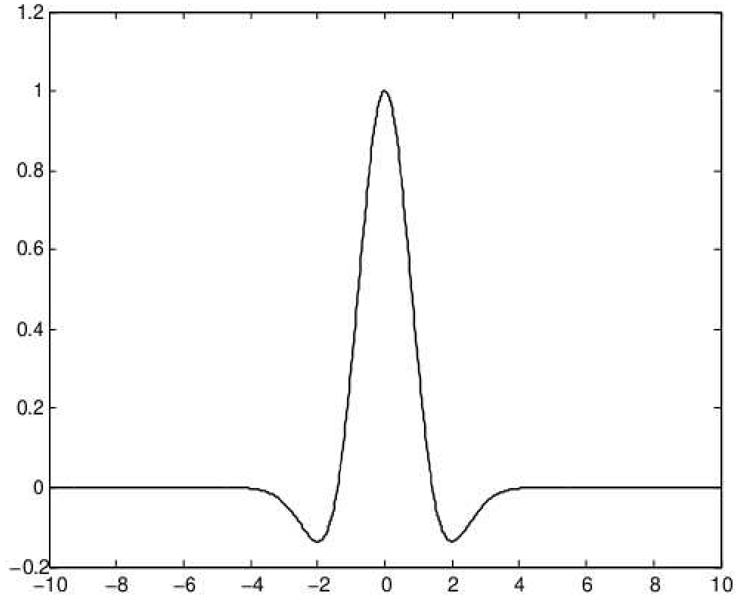

Figure 1: Radial graph of (12) for $d = 2$.

where K_ν is the modified Bessel function of second kind [12, Theorem 6.13, p. 76]. For negative β with $-\beta \notin \mathbf{N}$ we get conditionally positive definite *multiquadrics* with the same generalized Fourier transform [12, Theorem 8.15, p.109]. Due to Theorem 1, we can treat both cases simultanously here, and we check later for which cases the assumptions of the theorem are true.

In what follows, we work with

$$f(t) = (1+2t)^{-\beta}, \tag{13}$$

$$\hat{f}(u) = \frac{2^{1-\beta}}{\Gamma(\beta)}(2u)^{\frac{\beta}{2}-\frac{d}{4}} K_{\beta-\frac{d}{2}}\left((2u)^{\frac{1}{2}}\right) \tag{14}$$

and apply the relation (8) to get

$$\begin{aligned}
(tf'(tz))^{\wedge,t}(u) &= \left[-\beta\, 2t\,(1+2tz)^{-\beta-1}\right]^{\wedge,t}(u) \\
&= \tfrac{d}{dz}\left[z^{-\nu-1}\widehat{f(\cdot)}\left(\tfrac{u}{z}\right)\right] \\
&= \tfrac{d}{dz}\left[z^{-\nu-1}\tfrac{2^{1-\beta}}{\Gamma(\beta)}\left(2\tfrac{u}{z}\right)^{\frac{\beta}{2}-\frac{d}{4}} K_{\beta-\frac{d}{2}}\left(\left(2\tfrac{u}{z}\right)^{\frac{1}{2}}\right)\right] \\
&= \tfrac{2^{1-\beta}}{\Gamma(\beta)}\tfrac{d}{dz}\left[z^{-\nu-1}\left(2\tfrac{u}{z}\right)^{\frac{\beta}{2}-\frac{d}{4}} K_{\beta-\frac{d}{2}}\left(\left(2\tfrac{u}{z}\right)^{\frac{1}{2}}\right)\right].
\end{aligned}$$

Then we use the standard relation

$$\frac{d}{ds}\left(s^\nu K_\nu(s)\right) = -s^\nu K_{\nu-1}(s) \tag{15}$$

for derivatives of Bessel functions and all $\nu \geq 0$, taking $K_\nu = K_{-\nu}$ into account. We set

$$s := \left(2\frac{u}{z}\right)^{\frac{1}{2}},\ \frac{ds}{dz} = -\frac{1}{2}\frac{s}{z}$$

to continue with

$$\begin{aligned}
& \tfrac{2^{1-\beta}}{\Gamma(\beta)}\tfrac{d}{dz}\left[z^{-\nu-1}\left(\left(2\tfrac{u}{z}\right)^{\frac{1}{2}}\right)^{\beta-\frac{d}{2}} K_{\beta-\frac{d}{2}}\left(\left(2\tfrac{u}{z}\right)^{\frac{1}{2}}\right)\right] \\
= & \tfrac{2^{1-\beta}}{\Gamma(\beta)}\tfrac{d}{dz}\left[z^{-\nu-1}s^\nu K_\nu(s)\right] \\
= & \tfrac{2^{1-\beta}}{\Gamma(\beta)}\left[(-\nu-1)z^{-\nu-2}s^\nu K_\nu(s) + z^{-\nu-1}\tfrac{ds}{dz}\tfrac{d}{ds}s^\nu K_\nu(s)\right] \\
= & \tfrac{2^{1-\beta}}{\Gamma(\beta)}\left[(-\nu-1)z^{-\nu-2}s^\nu K_\nu(s) + z^{-\nu-1}\left(-\tfrac{1}{2}\tfrac{s}{z}\right)\left(-s^\nu K_{\nu-1}(s)\right)\right] \\
= & \tfrac{2^{1-\beta}}{\Gamma(\beta)}\Bigg[(-\nu-1)z^{-\nu-2}\left[\left(2\tfrac{u}{z}\right)^{\frac{1}{2}}\right]^{\beta-\frac{d}{2}} K_{\beta-\frac{d}{2}}\left(\left(2\tfrac{u}{z}\right)^{\frac{1}{2}}\right) \\
& +z^{-\nu-2}\frac{1}{2}\left(\left(2\tfrac{u}{z}\right)^{\frac{1}{2}}\right)^{\beta-\frac{d}{2}+1} K_{\beta-\frac{d}{2}-1}\left(\left(2\tfrac{u}{z}\right)^{\frac{1}{2}}\right)\Bigg].
\end{aligned}$$

If we put $z = 1$ in the relation above, we have

$$\begin{aligned}
& \left[-\beta\, 2t(1+2t)^{-\beta-1}\right]^\wedge(u) \\
= & \tfrac{2^{1-\beta}}{\Gamma(\beta)}\Bigg[(-\nu-1)\left((2u)^{\frac{1}{2}}\right)^{\beta-\frac{d}{2}} K_{\beta-\frac{d}{2}}\left((2u)^{\frac{1}{2}}\right) \\
& +\tfrac{1}{2}\left((2u)^{\frac{1}{2}}\right)^{\beta-\frac{d}{2}+1} K_{\beta-\frac{d}{2}-1}\left((2u)^{\frac{1}{2}}\right)\Bigg].
\end{aligned}$$

This proves the first statement of

Theorem 5. *The function*

$$\hat{\phi}(\omega) = \beta\|\omega\|^2(1+\|\omega\|^2)^{-\beta-1}$$

has the generalized inverse Fourier transform

$$\phi(x) = \frac{2^{1-\beta}}{\Gamma(\beta)}\left[\frac{d}{2}\|x\|^{\beta-\frac{d}{2}}K_{\beta-\frac{d}{2}}(\|x\|) - \frac{1}{2}\|x\|^2\|x\|^{\beta-\frac{d}{2}-1}K_{\beta-\frac{d}{2}-1}(\|x\|)\right]. \tag{16}$$

For inverse multiquadrics with $\beta > d/2+1$, the function ϕ is a positive definite radial kernel on $\mathbf{R}^d$ depicted in Figure 2 for $d = 2$ and $\beta = 2.5$.

Proof: In the inverse multiquadric case, the assumption $\beta > d/2+1$ is sufficient to let our derivation be valid. The function $\hat{\phi}(\omega)$ is nonnegative and not always equal to zero, and ϕ is absolutely integrable because it has exponential decay at infinity. Then we can invoke Theorem 3 to get the assertion. □

In particular we notice that ϕ is a linear combination between the Fourier transformation of inverse multiquadrics and another Fourier transformation of inverse multiquadrics multiplied by $\|\omega\|^2$.

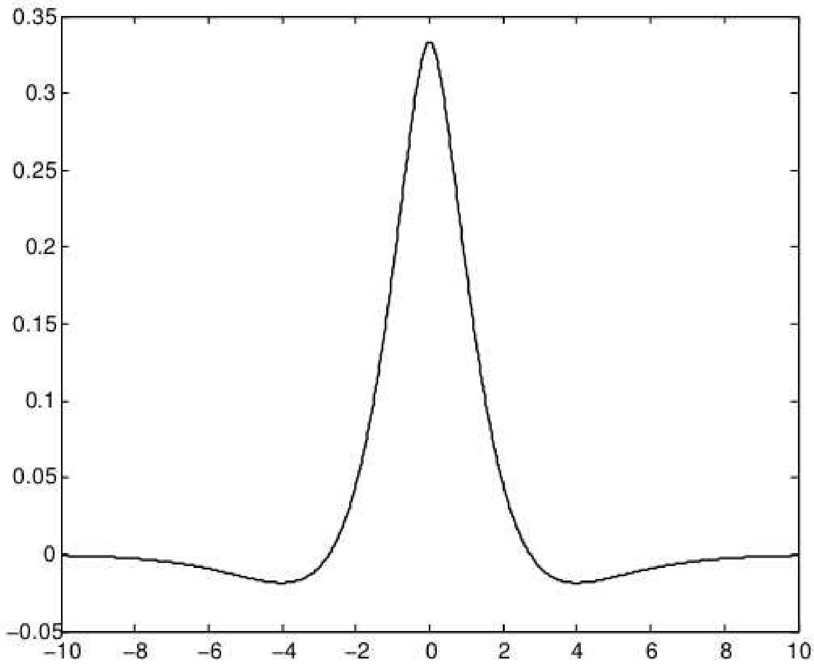

Figure 2: Radial graph of (16) for $\beta = 2.5$ and $d = 2$.

In the multiquadric case, the exponential decay of the Bessel functions let the integrals and the commutation be correct at infinity, and at the origin we need test functions with zeros of sufficiently high order. Then the generalized Fourier transform relation is valid, and we can try to invoke Theorem 4. The slowly increasing function is $\hat{\phi}$ now, but ϕ is not nonnegative, leading to no useful result.

4.3 Matérn kernel

The strictly positive definite kernels on $\mathbf{R}^d$ with positive d-variate Fourier transform $(1 + \|\omega\|^2)^{-\beta}$ for $\beta > d/2$ are the *Whittle–Matérn* or *Sobolev* kernels

$$\frac{2^{1-\beta}}{\Gamma(\beta)}\|x\|^{\beta-\frac{d}{2}}K_{\frac{d}{2}-\beta}(\|x\|) \quad = \quad \frac{2^{1-\beta}}{\Gamma(\beta)}\left(2\frac{\|x\|^2}{2}\right)^{\frac{\beta}{2}-\frac{d}{4}}K_{\beta-\frac{d}{2}}\left(\left(2\frac{\|x\|^2}{2}\right)^{\frac{1}{2}}\right)$$

with

$$f(t) = \frac{2^{1-\beta}}{\Gamma(\beta)} (2t)^{\frac{\beta}{2}-\frac{d}{4}} K_{\beta-\frac{d}{2}}\left((2t)^{\frac{1}{2}}\right), \tag{17}$$

$$\hat{f}(u) = (1+2u)^{-\beta}. \tag{18}$$

We proceed like in the multiquadric case from

$$\begin{array}{rl} & (tf'(tz))^{\wedge,t}(u) \\ = & \left[\frac{d}{dz}\left(\frac{2^{1-\beta}}{\Gamma(\beta)}(2tz)^{\frac{\beta}{2}-\frac{d}{4}} K_{\beta-\frac{d}{2}}\left((2tz)^{\frac{1}{2}}\right)\right)\right]^{\wedge,t}(u) \\ = & \left[\frac{2^{1-\beta}}{\Gamma(\beta)}\frac{d}{dz}\left(s^{\beta-\frac{d}{2}} K_{\beta-\frac{d}{2}}(s)\right)\right]^{\wedge,t}(u) \end{array}$$

with $s = (2tz)^{\frac{1}{2}}$ and $\frac{ds}{dz} = \frac{t}{s}$. Then the above formula continues with

$$\begin{array}{rl} & \frac{2^{1-\beta}}{\Gamma(\beta)}\left[\frac{ds}{dz}\frac{d}{ds}\left(s^{\beta-\frac{d}{2}} K_{\beta-\frac{d}{2}}(s)\right)\right]^{\wedge,t}(u) \\ = & -\frac{2^{1-\beta}}{\Gamma(\beta)}\left[ts^{\beta-\frac{d}{2}-1} K_{\beta-\frac{d}{2}-1}(s)\right]^{\wedge,t}(u) \\ = & -\frac{2^{1-\beta}}{\Gamma(\beta)}\left[t\left((2tz)^{\frac{1}{2}}\right)^{\beta-\frac{d}{2}-1} K_{\beta-\frac{d}{2}-1}\left((2tz)^{\frac{1}{2}}\right)\right]^{\wedge,t}(u) \\ = & -\frac{2^{1-\beta}}{\Gamma(\beta)}\left[\left(\frac{t}{2z}\right)^{\frac{1}{2}}\left((2tz)^{\frac{1}{2}}\right)^{\beta-\frac{d}{2}} K_{\beta-\frac{d}{2}-1}\left((2tz)^{\frac{1}{2}}\right)\right]^{\wedge,t}(u) \\ = & \frac{d}{dz}\left(z^{-\nu-1}\widehat{f(\cdot)}\left(\frac{u}{z}\right)\right) = \frac{d}{dz}\left(z^{-\nu-1}\left(1+2\frac{u}{z}\right)^{-\beta}\right) \\ = & (-\nu-1)z^{-\nu-2}\left(1+2\frac{u}{z}\right)^{-\beta} + z^{-\nu-1}(-\beta)\left(1+2\frac{u}{z}\right)^{-\beta-1}\left(-2\frac{u}{z^2}\right) \\ = & z^{-\nu-3}\left(1+2\frac{u}{z}\right)^{-\beta-1}\left[(-\nu-1)z\left(1+2\frac{u}{z}\right)+2\beta u\right]. \end{array}$$

If we put $z = 1$, the above computation becomes

$$\begin{array}{rl} & -\frac{2^{1-\beta}}{\Gamma(\beta)}\left[\left(\frac{t}{2}\right)^{\frac{1}{2}}\left((2t)^{\frac{1}{2}}\right)^{\beta-\frac{d}{2}} K_{\beta-\frac{d}{2}-1}\left((2t)^{\frac{1}{2}}\right)\right]^{\wedge}(u) \\ = & (1+2u)^{-\beta-1}\left[-\frac{d}{2}(1+2u)+2\beta u\right] \end{array}$$

and we get

Theorem 6. *For $\beta > 1 + d/2$, the radial kernel*

$$\phi(x) = \frac{d}{2}(1+\|x\|^2)^{-\beta} - \beta\|x\|^2(1+\|x\|^2)^{-\beta-1} \tag{19}$$

is positive definite.

Proof: Its d–variate Fourier transform is the nonnegative function

$$\begin{aligned} \hat{\phi}(\omega) &= \frac{2^{1-\beta}}{\Gamma(\beta)}\frac{\|\omega\|}{2}\|\omega\|^{\beta-\frac{d}{2}} K_{\beta-\frac{d}{2}-1}(\|\omega\|) = \\ &= \frac{2^{-\beta}}{\Gamma(\beta)}\|\omega\|^{\beta-\frac{d}{2}+1} K_{\beta-\frac{d}{2}-1}(\|\omega\|) = \\ &= \frac{2^{-\beta}}{\Gamma(\beta)}\|\omega\|^2\|\omega\|^{\beta-\frac{d}{2}-1} K_{\beta-\frac{d}{2}-1}(\|\omega\|) \end{aligned}$$

and we can apply Theorem 3. □

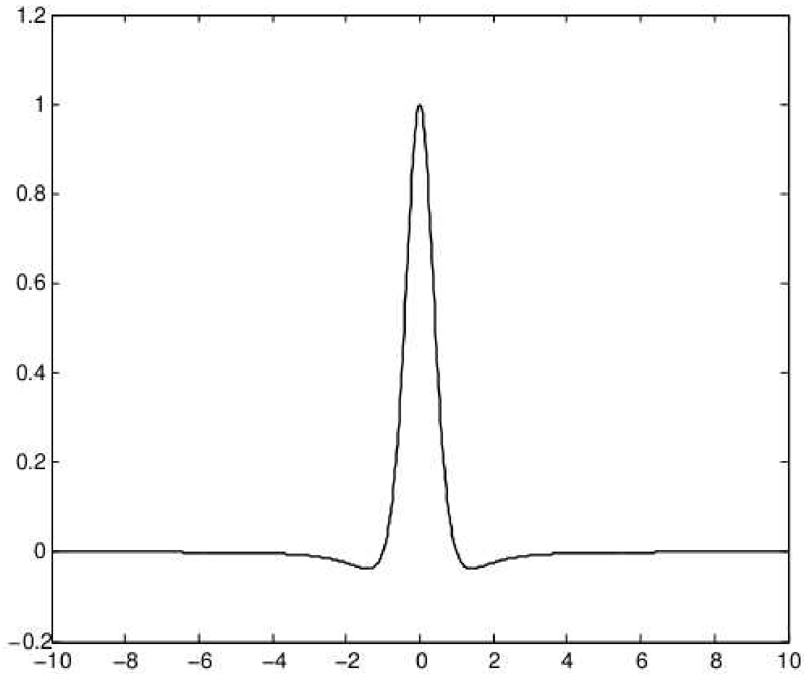

Figure 3: Radial graph of $\phi(x)$ for $\beta = 2$ and $d = 2$.

Powers and thin–plate splines are polyharmonics, and their generalized Fourier transform is a negative power, thus essentially scale–invariant. This means that there will be no new kernels when acting on a scaling of these kernels.

5 Laplacians

We now exploit the well–known fact [12, Lemma 9.15, p. 130] that the transition $\Phi \to -\Delta\Phi$ allows to generate new kernels under certain circumstances. Here, we make this construction explicit and transparent for the standard families of radial kernels. One could take the radial form of the Laplacian and apply it to any kernel, but then it is not clear a priori to which class of kernels the result belongs. We do this here in the f-form and see that the Laplacian of a kernel leads to a linear conbination of two kernels of the same family.

Kernels of the form $-\Delta K$ have applications in the Method of Particular Solutions, since one can find a special solution of $-\Delta u = f$ by interpolating f by translates of $-\Delta K$ and using the obtained coefficients to get u in terms of translates of K. See [5, 7, 2, 8, 4] for applications.

If we write a radial kernel K in f–form with $s = r^2/2 = \|x-y\|_2^2/2$, its d–variate Laplacian follows via

$$\begin{array}{rcl}\dfrac{\partial}{\partial x_j}K(x-y) & = & \dfrac{\partial s}{\partial x_j}\dfrac{d}{ds}f(s) = f'(s)(x_j-y_j)\\ \dfrac{\partial^2}{\partial x_j^2}K(x-y) & = & f''(s)(x_j-y_j)^2 + f'(s)\\ -\Delta^x K(x-y) & = & -2sf''(s) - df'(s).\end{array} \tag{20}$$

The generalized Fourier transform of $-\Delta^x K(x-y)$ is $\hat{\Phi}(\omega)\|\omega\|_2^2$ if $\Phi(x) = K(x)$. This shows that whenever $-\Delta^x K(x-y)$ exists, is continuous and Fourier transformable, the radial kernel defined by $-\Delta\Phi$ is positive definite, if Φ is positive definite. For conditionally positive definite Φ of order m, the same argument applies, but the order of conditional positive definiteness of $-\Delta\Phi$ then is $m-1$ because of the new factor in the Fourier transform.

Theorem 7. *The transition $\Phi \to -\Delta\Phi$ on radial kernels generates a radial kernel consisting of a weighted sum (20) of two radial kernels, if f is the f–form of Φ, and if the action of $-\Delta$ is valid on the kernel. If, furthermore, the class of kernels is invariant under taking pairs of forward and backward Fourier transforms in arbitrary dimensions. the resulting kernel is a weighted linear combination of two radial kernels of the* same *family.*

Proof: By the connection between Fourier transforms and derivatives in f–form [10], the hypothesis on the family of kernels implies that it is invariant under derivatives taken in f–form as far as the derivatives and Fourier transforms are valid. But then the assertion follows from (20). $\square$

The transition $\Phi \to -\Delta\Phi$ now allows to generate new kernels via the f–form, since all popular classes of radial kernels satisfy the above theorem.

For the Gaussian, $f(s) = \exp(-s)$, and the negative Laplacian is the kernel

$$(d - \|x-y\|_2^2)\exp(-\|x-y\|_2^2/2)$$

due to
$$-2sf''(s) - df'(s) \quad = \quad (d-2s)\exp(-s).$$
This coincides up to a positive factor with (12).

Inverse multiquadrics have
$$\begin{array}{rcl} f_\beta(s) & = & (1+2s)^{-\beta}, \\ f'_\beta(s) & = & -2\beta(1+2s)^{-\beta-1}, \\ f''_\beta(s) & = & -4\beta(-\beta-1)(1+2s)^{-\beta-2} \end{array}$$
for $\beta > 0$ and thus
$$-2sf''_\beta(s) - df'_\beta(s) = -8s\beta(\beta+1)(1+2s)^{-\beta-2} + 2d\beta(1+2s)^{-\beta-1}$$
leading to the new positive definite radial kernel
$$-4\beta(\beta+1)\|x-y\|_2^2(1+\|x-y\|_2^2)^{-\beta-2} + 2d\beta(1+\|x-y\|_2^2)^{-\beta-1},$$
but this is not (16). Note that no restrictions on β except $\beta > 0$ are needed. The inequality $\beta > d/2$ is required in some texts, but it is not necessary, if generalized Fourier transforms are used [12, Th. 8.15, p. 109, Th. 6.13, p. 76].

The same argument works for standard multiquadrics with $\beta < 0$ and $-\beta \notin \mathbf{N}$, which are conditionally positive definite of order $m = \max(0, \lceil -\beta \rceil)$, but the resulting kernel is conditionally positive definite of order $\max(0, m-1)$. An additional restriction on β is not necessary for standard multiquadrics [12, Th. 8.15, p. 109], but note that for small $|\beta|$ the resulting kernels may partially be inverse multiquadrics.

The Sobolev–Whittle–Matérn kernels
$$\|x-y\|_2^\nu K_\nu(\|x-y\|_2)$$
have
$$\begin{array}{rcl} f_\nu(s) & = & (2s)^{\nu/2}K_\nu(\sqrt{2s}) \\ f'_\nu(s) & = & -f_{\nu-1}(s) \\ f''_\nu(s) & = & f_{\nu-2}(s) \end{array}$$
for $\nu > 2$ following a simple calculation in [9] using the derivative rule for K_ν. Then
$$-2sf''_\nu(s) - df'_\nu(s) = -\|x-y\|_2^2 f_{\nu-2}(s) + df_{\nu-1}(s)$$
leads to the new positive definite kernel
$$-\|x-y\|_2^\nu K_{\nu-2}(\|x-y\|_2) + d\|x-y\|_2^{\nu-1}K_{\nu-1}(\|x-y\|_2)$$

for $\nu > 2$.

Wendland kernels in the notation of [9] have f–forms $\phi_{d,k}(s)$ for C^{2k} smoothness and positive definiteness in R^d with d maximal. Then, following [9],

$$\phi'_{d,k} = -\phi_{d+2,k-1}$$

and

$$-2s\phi''_{d,k}(s) - d\phi'_{d,k}(s) = -2s\phi_{d+4,k-2}(s) + d\phi_{d+2,k-1}(s)$$

leads to a new positive definite kernel for $k \geq 2$. The special case $d = k = 3$ starts from

$$\begin{array}{rcl} \phi_{3,3}(r) & = & (1-r)_+^8(32r^3 + 25r^2 + 8r + 1) \\ \phi_{5,2}(r) & = & 22(1-r)_+^7(16r^2 + 7r + 1) \\ \phi_{7,1}(r) & = & 528(1-r)_+^6(6r + 1) \end{array}$$

and leads to the new kernel

$$-528(1-r)_+^6 r^2(6r+1) + 66(1-r)_+^7(16r^2 + 7r + 1).$$

This construction is trivial for polyharmonic kernels, because the action of $-\Delta$ takes polyharmonic kernels into other polyharmonic kernels. Polyharmonic kernels can be completely characterized by the fact that their d–variate generalized Fourier transforms are of the form $\|\omega\|^{-d-\beta}$, and then the action of $-\Delta$ yields a polyharmonic kernel with d–variate generalized Fourier transform $\|\omega\|^{-d-\beta+2}$. For $\beta \notin 2\mathbf{Z}$ and $\beta > 2$ this is the transition from power kernels r^β to $r^{\beta-2}$, while for $\beta = 2k \in 2\mathbf{Z}$, $k > 1$ this goes from thin–plate splines $r^{2k}\log r$ to $r^{2k-2}\log r$.

6 Open Problems

Future work should take higher–order derivatives with respect to scale and derive recursive formulas for kernels generated this way. Furthermore, the paper [3] successfully used divided differences instead of derivatives in a special case, and it should be worth while to check if other functionals generate new kernels.

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] B. Bacchelli and M. Bozzini. Particular solution of Poisson problems using cardinal Lagrangian polyharmonic splines. In A.J.M. Ferreira, E.J. Kansa, G.E. Fasshauer, and V.M.A. Leitão, editors, *Progress on meshless methods*, pages 1–16. Springer, New York, 2009.

[3] M. Bozzini, M. Rossini, and R. Schaback. Generalized Whittle–Matérn and polyharmonic kernels. Advances in Computational Mathematics, 1-13, article in press DOI 10.1007/s10444-012-9277-9.

[4] C. S. Chen, C. M. Fan, and P. H. Wen. The method of approximate particular solutions for solving certain partial differential equations. *Numer. Methods Partial Differential Equations*, 28(2):506–522, 2012.

[5] C.S. Chen, A.S. Muleshkov, and M.A. Golberg. The numerical evaluation of particular solution for Poisson's equation - a revisit. In C.A. Brebbia and H. Power, editors, *Boundary Elements XXI*, pages 313–322. WIT Press, 1999.

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

[7] M.A. Golberg, C.S. Chen, and M. Ganesh. Particular solutions of 3D Helmholtz-type equations using compactly supported radial basis functions. *Eng. Anal. with Boundary Elements*, 24:539–547, 2000.

[8] Zi-Cai Li, Lih-Jier Young, Hung-Tsai Huang, Ya-Ping Liu, and Alexander H.-D. Cheng. Comparisons of fundamental solutions and particular solutions for Trefftz methods. *Eng. Anal. Bound. Elem.*, 34(3):248–258, 2010.

[9] R. Schaback. Matlab programming for kernel-based methods. Technical report, Institut für Numerische und Angewandte Mathematik Göttingen, 2009–. Preprint, available via http://num.math.uni-goettingen.de/schaback/research/papers/MPfKBM.pdf.

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

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

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