University of Isfahan

Meshless Approximants

Davoud Mirzaei

Polynomial interpolation in $\mathbb{R}$

For given $X = \{x_1, x_2, \ldots, x_N\} \subset \mathbb{R}$ and values $f_1, \ldots, f_N \in \mathbb{R}$, find $p \in \mathbb{P}_{N-1}$ such that

$$p(x_k) = f_k, \quad k = 1, \ldots, N$$

This problem admits a unique solution because any univariate polynomial of degree $N - 1$ has at most $N - 1$ zeros.

In a more general setting the polynomial interpolation problem has a unique solution because $\mathbb{P}_{N-1}$ is a Haar space of dimension $N$. 
Polynomial interpolation in $\mathbb{R}$

For given $X = \{x_1, x_2, \ldots, x_N\} \subset \mathbb{R}$ and values $f_1, \ldots, f_N \in \mathbb{R}$, find $p \in \mathbb{P}_{N-1}$ such that

$$p(x_k) = f_k, \quad k = 1, \ldots, N$$

This problem admits a unique solution because any univariate polynomial of degree $N - 1$ has at most $N - 1$ zeros.

In a more general setting the polynomial interpolation problem has a unique solution because $\mathbb{P}_{N-1}$ is a Haar space of dimension $N$. 
Polynomial interpolation in $\mathbb{R}$

**Definition:** Let $V \subset C(\Omega)$ be an $N$-dimensional linear space on $\Omega \subset \mathbb{R}^d$. Then $V$ is called a Haar or Chebysheff space of dimension $N$ on $\Omega$ if for arbitrary distinct points $x_1, \ldots, x_N \in \Omega$ the only function $s \in V$ which satisfies $s(x_k) = 0$ for $k = 1, \ldots, N$ be $s(x) \equiv 0$.

Examples of univariate Haar spaces:

- $\mathbb{P}_{N-1} = \text{span}\{1, x, \ldots, x^{N-1}\}$ is Haar space of dimension $N$,
- $\mathbb{E}_{N-1} = \text{span}\{1, e^x, \ldots, e^{(N-1)x}\}$ is Haar space of dimension $N$,
- $\mathbb{S}_{2n} = \text{span}\{1, \cos x, \sin x, \ldots, \cos nx, \sin nx\}$ is Haar space of dimension $N = 2n + 1$,
- $\ldots$

A negative result in $\mathbb{R}^d$, $d \geq 2$:

**Theorem** (Mairhuber–Curtis 1956, 1959):

Assume that $\Omega \subset \mathbb{R}^d$ contains an interior point and $d \geq 2$. Then there exists no Haar space on $\Omega$ of dimension $N \geq 2$. 
Polynomial interpolation in $\mathbb{R}$

**Definition:** Let $V \subset C(\Omega)$ be an $N$-dimensional linear space on $\Omega \subset \mathbb{R}^d$. Then $V$ is called a Haar or Chebyshev space of dimension $N$ on $\Omega$ if for arbitrary distinct points $x_1, \ldots, x_N \in \Omega$ the only function $s \in V$ which satisfies $s(x_k) = 0$ for $k = 1, \ldots, N$ be $s(x) \equiv 0$.

Examples of univariate Haar spaces:

- $\mathbb{P}_{N-1} = \text{span}\{1, x, \ldots, x^{N-1}\}$ is Haar space of dimension $N$,
- $\mathbb{E}_{N-1} = \text{span}\{1, e^x, \ldots, e^{(N-1)x}\}$ is Haar space of dimension $N$,
- $\mathbb{S}_{2n} = \text{span}\{1, \cos x, \sin x, \ldots, \cos nx, \sin nx\}$ is Haar space of dimension $N = 2n + 1$,
- $\ldots$

A negative result in $\mathbb{R}^d, d \geq 2$:

**Theorem** (Mairhuber–Curtis 1956, 1959):
Assume that $\Omega \subset \mathbb{R}^d$ contains an interior point and $d \geq 2$. Then there exists no Haar space on $\Omega$ of dimension $N \geq 2$. 
Polynomial interpolation in $\mathbb{R}$

**Definition:** Let $V \subset C(\Omega)$ be an $N$-dimensional linear space on $\Omega \subset \mathbb{R}^d$. Then $V$ is called a Haar or Chebysheff space of dimension $N$ on $\Omega$ if for arbitrary distinct points $x_1, \ldots, x_N \in \Omega$ the only function $s \in V$ which satisfies $s(x_k) = 0$ for $k = 1, \ldots, N$ be $s(x) \equiv 0$.

Examples of univariate Haar spaces:

- $P_{N-1} = \text{span}\{1, x, \ldots, x^{N-1}\}$ is Haar space of dimension $N$,
- $E_{N-1} = \text{span}\{1, e^x, \ldots, e^{(N-1)x}\}$ is Haar space of dimension $N$,
- $S_{2n} = \text{span}\{1, \cos x, \sin x, \ldots, \cos nx, \sin nx\}$ is Haar space of dimension $N = 2n + 1$,
- $\ldots$

A negative result in $\mathbb{R}^d$, $d \geq 2$:

**Theorem** (Mairhuber–Curtis 1956, 1959):

Assume that $\Omega \subset \mathbb{R}^d$ contains an interior point and $d \geq 2$. Then there exists no Haar space on $\Omega$ of dimension $N \geq 2$. 
Polynomial interpolation in $\mathbb{R}^d$

**A simple conclusion:** Multi-dimensional polynomial interpolation is not well-defined in general; One should care about the location of points.

**A simple example:** $\mathbb{P}_1^2 = \text{span}\{1, x, y\}$ with dimension 3. Any bivariate linear polynomial describes a plane in three dimensional space. This plane is uniquely determined by 3 points iff points are not collinear.

**Definition:** A set $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^d$ is called $V$-unisolvent if admits a unique interpolant. (The only function from $V$ that vanishes on $X$ be the zero function).
Polynomial interpolation in $\mathbb{R}$

**A simple conclusion:** Multi-dimensional polynomial interpolation is not well-defined in general; **One should care about the location of points.**

![Diagram of polynomial interpolation](image)

**A simple example:** $\mathbb{P}^2_1 = \text{span}\{1, x, y\}$ with dimension 3. Any bivariate linear polynomial describes a plane in three dimensional space. This plane is uniquely determined by 3 points iff points are not collinear.

**Definition:** A set $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^d$ is called $V$–unisolvent if admits a unique interpolant. (The only function from $V$ that vanishes on $X$ be the zero function).
Polynomial interpolation in $\mathbb{R}$

A simple conclusion: Multi-dimensional polynomial interpolation is not well-defined in general; One should care about the location of points.

A simple example: $\mathbb{P}_2 = \text{span}\{1, x, y\}$ with dimension 3. Any bivariate linear polynomial describes a plane in three dimensional space. This plane is uniquely determined by 3 points iff points are not collinear.

Definition: A set $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^d$ is called $V$–unisolvent if admits a unique interpolant. (The only function from $V$ that vanishes on $X$ be the zero function).
A conclusion from Mairhuber–Curtis Theorem

In multivariate interpolation to get unisolvency for arbitrary data we can not fix the basis functions in advance. The basis should be data–dependent.

If \( X = \{x_1, \ldots, x_N\} \subseteq \Omega \subseteq \mathbb{R}^d \):

\[
V_N = \text{span}\{\Phi(\cdot, x_1), \Phi(\cdot, x_2), \ldots, \Phi(\cdot, x_N)\}
\]

where \( \Phi \) is a kernel.

This is also known from spline theory where the space of spline functions of order \( \ell \) depends on data locations:

\[
S_X^\ell = \text{span}\left\{1, x, \ldots, x^{\ell}, (x - x_2)^\ell_+, \ldots, (x - x_{N-1})^\ell_+\right\}
\]
A conclusion from Mairhuber–Curtis Theorem

In multivariate interpolation to get unisolvency for arbitrary data we can not fix the basis functions in advance. The basis should be data–dependent.

If \( X = \{x_1, \ldots, x_N\} \subseteq \Omega \subseteq \mathbb{R}^d \):

\[
V_N = \text{span}\{\Phi(\cdot, x_1), \Phi(\cdot, x_2), \ldots, \Phi(\cdot, x_N)\}
\]

where \( \Phi \) is a kernel.

This is also known from spline theory where the space of spline functions of order \( \ell \) depends on data locations:

\[
S_X^\ell = \text{span}\left\{1, x, \ldots, x^\ell, (x - x_2)^\ell_+, \ldots, (x - x_{N-1})^\ell_+\right\}
\]
Interpolation Problem

Given data \((x_k, f_k), k = 1, 2, \ldots, N\), for \(x_k \in \mathbb{R}^d\) and \(f_k \in \mathbb{R}\), find a (continuous) function \(s = s_{f,x}\) such that

\[
s(x_k) = f_k, \quad \text{for} \; k = 1, 2, \ldots, N.
\]

Interpolation by kernels

\[
s(x) = \sum_{k=1}^{N} c_k \Phi(x, x_k),
\]

Imposing the interpolation conditions:

\[
Ac = f, \quad a_{kj} = \Phi(x_j, x_k),
\]
**Solvability:** If $\Phi$ is **positive definite** then the interpolation problem is well-defined.

A kernel $\Phi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ is semi-positive definite (SPD) on $\mathbb{R}^d$ if it is symmetric and for all $N$ and all $x_1, \ldots, x_N \in \mathbb{R}^d$ and all $\alpha \in \mathbb{C}^N$

$$\alpha^T A \overline{\alpha} = \sum_{j,k=1}^{N} \alpha_j \overline{\alpha_k} \Phi(x_j, x_k) \geq 0$$

$\Phi$ is positive definite (PD) if equality holds only for $\alpha = 0$.

**Standard kernels**

- Shift invariant basis function: $\Phi(x - y)$ for $\Phi : \mathbb{R}^d \to \mathbb{R}$
- Shift and rotation invariant (radial) basis functions: $\Phi(x, y) = \phi(\|x - y\|_2)$ for $\phi : [0, \infty) \to \mathbb{R}$
- Zonal functions: $\Phi(x, y) = \psi(x^T y)$ for $\psi : [-1, 1] \to \mathbb{R}$, (used on $S^d$)
- ...

---

**Main References**

- Approximation by Kernels
- Solving PDEs by Kernels
- Approximation by MLS
- GMLS Approximation and DMLPG Methods
**Solvability:** If $\Phi$ is **positive definite** then the interpolation problem is well-defined. A kernel $\Phi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ is semi-positive definite (SPD) on $\mathbb{R}^d$ if it is symmetric and for all $N$ and all $x_1, \ldots, x_N \in \mathbb{R}^d$ and all $\alpha \in \mathbb{C}^N$

$$\alpha^T A \overline{\alpha} = \sum_{j,k=1}^{N} \alpha_j \overline{\alpha}_k \Phi(x_j, x_k) \geq 0$$

$\Phi$ is positive definite (PD) if equality holds only for $\alpha = 0$.

**Standard kernels**

- Shift invariant basis function: $\Phi(x - y)$ for $\Phi : \mathbb{R}^d \to \mathbb{R}$
- Shift and rotation invariant (radial) basis functions: $\Phi(x, y) = \phi(||x - y||_2)$ for $\phi : [0, \infty) \to \mathbb{R}$
- Zonal functions: $\Phi(x, y) = \psi(x^T y)$ for $\psi : [-1, 1] \to \mathbb{R}$, (used on $\mathbb{S}^d$)
- ...
Example: Gaussian radial function

We will see as soon that $\Phi(x) = \exp(-\varepsilon^2 \|x\|^2)$ for a fixed $\varepsilon$ is a basic choice.

$$\text{span}\{\Phi(\cdot - x_1), \ldots, \Phi(\cdot - x_N)\}$$

Figure: 13 scattered points and the shifted Gaussians at these points

$$s_{f,x}(x) = \sum_{j=1}^{N} c_j \Phi(x - x_j), \quad x \in \Omega$$
An example in $\mathbb{R}^2$

Consider the interpolation of Franke’s function on $[0, 1]^2$ on 200 Halton points by Gaussian function $\Phi(x, y) = \exp(-\varepsilon^2 \|x - y\|^2_2)$ for $\varepsilon = 5$.

Figure: Interpolation points and the interpolant by Gaussian function
An example in $S^2$

Consider the interpolation of Franke’s function on $S^2$ on 500 scattered points by locally supported zonal function $\Phi(x, y) = (1 - \sqrt{2 - 2x^Ty})^4 (4\sqrt{2 - 2x^Ty} + 1)$.

**Figure:** Interpolation points and the interpolant on 2-sphere
Matlab code for \( \mathbb{R}^2 \) case

```matlab
1 rbf = @(r,e)exp(-e^2*r.^2); e = 5; % kernel
2 [Xe,xmm,ymm] = points(0,1,0,1,1/22); % evaluation points
3 Uex = franke(Xe(:,1),Xe(:,2)); % exact solution
4 N = 200; % number of Halton points
5 p = haltonset(2,'Skip',1e3,'Leap',1e2); % set of N Halton points
6 X = net(p,N); % distance matrix
7 r = dismat(X,X); % distance matrix
8 A = rbf(r,e); % interpolation matrix
9 b = franke(X(:,1),X(:,2)); % known right-hand side
10 c = A\b; % coefficient vector
11 re = dismat(Xe,X); % evaluation distance matrix
12 Uapp = rbf(r,e)*c; % approximate solution
```
Matlab code for $S^2$ case

1. \texttt{sbf}=@(t,e)\texttt{max}(1-e*t,0).^4.*(4*e*t+1); \% kernel
2. \texttt{[xe, ye, ze] = sphere(255);}
3. \texttt{Xe=[xe(:) ye(:) ze(:)];} \% evaluation points
4. \texttt{Uex = knownfunc(Xe);} \% exact solution
5. \texttt{N = 500;} \% number of points
6. \texttt{X = Spoints(N);} \% set of points
7. \texttt{t = real(sqrt(2-2*Y*X'));} \% distance matrix
8. \texttt{A = sbf(t,1);} \% interpolation matrix
9. \texttt{b = knownfunc(X);} \% right-hand side
10. \texttt{c = A\b;} \% coefficient vector
11. \texttt{te = real(sqrt(2-2*Xe*X'));} \% evaluation distance matrix
12. \texttt{Uapp = sbf(te,1)*c;} \% approximate solution
Positive Definite Functions

Consider this function: $\Phi(x) = e^{ix^T \omega}$, $x \in \mathbb{R}^d$, for a fixed $\omega \in \mathbb{R}^d$.

$$\alpha^T A \alpha = \sum_{j=1}^{N} \sum_{k=1}^{N} \alpha_j \alpha_k \Phi(x_j - x_k) = \sum_{j=1}^{N} \sum_{k=1}^{N} \alpha_j \alpha_k e^{i(x_j - x_k)^T \omega}$$

$$= \sum_{j=1}^{N} \alpha_j e^{ix_j^T \omega} \sum_{k=1}^{N} \alpha_k e^{ix_k^T \omega}$$

$$= \left\| \sum_{k=1}^{N} \alpha_k e^{ix_k^T \omega} \right\|_2^2 \geq 0.$$

Thus $\Phi$ is positive semi-definite.

Bochner shows that other positive definite functions are obtained as (infinite) linear combinations of this function.
Positive Definite Functions
Consider this function: $\Phi(x) = e^{ix^T \omega}$, $x \in \mathbb{R}^d$, for a fixed $\omega \in \mathbb{R}^d$.

\[
\alpha^T A \alpha = \sum_{j=1}^{N} \sum_{k=1}^{N} \alpha_j \overline{\alpha}_k \Phi(x_j - x_k) = \sum_{j=1}^{N} \sum_{k=1}^{N} \alpha_j \overline{\alpha}_k e^{i(x_j - x_k)^T \omega} = \sum_{j=1}^{N} \alpha_j e^{ix_j^T \omega} \sum_{k=1}^{N} \alpha_k e^{ix_k^T \omega} = \left\| \sum_{k=1}^{N} \alpha_k e^{ix_k^T \omega} \right\|_2^2 \geq 0.
\]

Thus $\Phi$ is positive semi-definite.
Bochner shows that other positive definite functions are obtained as (infinite) linear combinations of this function.
**Shift-invariant basis functions**

Fourier transform: If \( \Phi : \mathbb{R}^d \rightarrow \mathbb{R} \) belongs to \( L_1(\mathbb{R}^d) \) then

\[
\hat{\Phi}(\omega) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \Phi(x) e^{-ix^T\omega} dx.
\]

**Theorem:** (Bochner 1933), Integrable functions on \( \mathbb{R}^d \)

Suppose that \( \Phi \in L_1(\mathbb{R}^d) \cap C(\mathbb{R}^d) \). Then \( \Phi \) is positive definite on every \( \mathbb{R}^d \) if and only if \( \Phi \) is bounded and its Fourier transform is nonnegative and nonvanishing.

\[
\Phi(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \hat{\Phi}(\omega) e^{i\omega^T x} d\omega
\]

\[
\alpha^T A \alpha = \sum_{j,k=1}^{N} \alpha_j \alpha_k \Phi(x_j - x_k) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \left| \sum_{j=1}^{N} \alpha_j e^{i\omega^T x_j} \right|^2 \hat{\Phi}(\omega) d\omega
\]
Shift-invariant basis functions

Fourier transform: If $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ belongs to $L_1(\mathbb{R}^d)$ then

$$\hat{\Phi}(\omega) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \Phi(x) e^{-ix^T\omega} dx.$$ 

**Theorem:** (Bochner 1933), Integrable functions on $\mathbb{R}^d$

Suppose that $\Phi \in L_1(\mathbb{R}^d) \cap C(\mathbb{R}^d)$. Then $\Phi$ is positive definite on every $\mathbb{R}^d$ if and only if $\Phi$ is bounded and its Fourier transform is nonnegative and nonvanishing.

$$\Phi(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \hat{\Phi}(\omega) e^{i\omega^T x} d\omega$$

$$\alpha^T A \alpha = \sum_{j,k=1}^{N} \alpha_j \alpha_k \Phi(x_j - x_k) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \left| \sum_{j=1}^{N} \alpha_j e^{i\omega^T x_j} \right|^2 \hat{\Phi}(\omega) d\omega$$
Shift-invariant basis functions

Fourier transform: If $\Phi : \mathbb{R}^d \to \mathbb{R}$ belongs to $L_1(\mathbb{R}^d)$ then

$$\hat{\Phi}(\omega) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \Phi(x)e^{-ix^T\omega}dx.$$ 

**Theorem:** (Bochner 1933), Integrable functions on $\mathbb{R}^d$

Suppose that $\Phi \in L_1(\mathbb{R}^d) \cap C(\mathbb{R}^d)$. Then $\Phi$ is positive definite on every $\mathbb{R}^d$ if and only if $\Phi$ is bounded and its Fourier transform is nonnegative and nonvanishing.

$$\Phi(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \hat{\Phi}(\omega)e^{i\omega^T x}d\omega$$

$$\alpha^T A \alpha = \sum_{j,k=1}^{N} \alpha_j \alpha_k \Phi(x_j - x_k) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \left| \sum_{j=1}^{N} \alpha_j e^{i\omega^T x_j} \right|^2 \hat{\Phi}(\omega)d\omega$$
Examples

- **Gaussian**: $\Phi(x) = e^{-\varepsilon^2 \|x\|_2^2}$, $\alpha > 0$, $x \in \mathbb{R}^d$,

  $$\hat{\Phi}(\omega) = (2\alpha)^{-d/2} e^{-\|\omega\|_2^2/(4\varepsilon^2)}$$

- **Inverse Multiquadrics (IMQ)**: $\Phi(x) = (1 + \varepsilon^2 \|x\|_2^2)^{-\beta}$, $x \in \mathbb{R}^d$, $\varepsilon > 0$, $\beta > d/2$,

  $$\hat{\Phi}(\omega) = \frac{2^{1-\beta} \varepsilon^{2\beta}}{\Gamma(\beta)} (\varepsilon \|\omega\|_2)^{\beta-d/2} K_{d/2-\beta}(\|\omega\|_2/\varepsilon)$$

- **Matérn (Sobolev) kernel**: $\Phi(x) = \frac{2^{1-s}}{\Gamma(s)} \|x\|_2^{s-d/2} K_{d/2-s}(\|\omega\|_2)$, $x \in \mathbb{R}^d$ and $s > d/2$

  $$\hat{\Phi}(\omega) = (1 + \|\omega\|_2^2)^{-s}$$
Radial functions

Definition: A function $f$ is called completely monotone on $(0, \infty)$ if it satisfies $f \in C^\infty(0, \infty)$ and

$$(-1)^\ell f^{(\ell)}(r) \geq 0, \quad \text{for all } \ell \in \mathbb{N}_0 \text{ and all } r > 0.$$ 

If in addition $f \in C[0, \infty)$ then $f$ is called completely monotone on $[0, \infty]$.

Simple examples and motivation

- $f(r) = \text{const.},$ $\Rightarrow (-1)^\ell f^{(\ell)}(r) = 0 \geq 0,$
- $f(r) = e^{-\alpha r}, \alpha > 0,$ $\Rightarrow (-1)^\ell f^{(\ell)}(r) = \alpha^\ell e^{-\alpha r} \geq 0,$
- $f(r) = (1 + r)^{-\beta}, \beta > 0 \Rightarrow (-1)^\ell f^{(\ell)}(r) = \beta(\beta + 1) \cdots (\beta + \ell - 1)(1 + r)^{-\beta-\ell} \geq 0$

Now, if we define $\phi(r) := f(r^2)$ then $\phi$ is positive definite!
Radial functions

**Definition:** A function $f$ is called **completely monotone** on $(0, \infty)$ if it satisfies $f \in C^\infty(0, \infty)$ and

$$(−1)^\ell f^{(\ell)}(r) \geq 0, \quad \text{for all } \ell \in \mathbb{N}_0 \text{ and all } r > 0.$$ 

If in addition $f \in C[0, \infty)$ then $f$ is called completely monotone on $[0, \infty]$.

**Simple examples and motivation**

- $f(r) = \text{const.}, \quad \Rightarrow \quad (−1)^\ell f^{(\ell)}(r) = 0 \geq 0,$
- $f(r) = e^{−\alpha r}, \alpha > 0, \quad \Rightarrow \quad (−1)^\ell f^{(\ell)}(r) = \alpha^\ell e^{-\alpha r} \geq 0,$
- $f(r) = (1 + r)^{-\beta}, \beta > 0 \quad \Rightarrow \quad (−1)^\ell f^{(\ell)}(r) = \beta(\beta + 1) \cdots (\beta + \ell - 1)(1 + r)^{-\beta-\ell} \geq 0$

Now, if we define $\phi(r) := f(r^2)$ then $\phi$ is positive definite!
Radial functions

**Definition:** A function $f$ is called **completely monotone** on $(0, \infty)$ if it satisfies $f \in C^\infty(0, \infty)$ and

$$(−1)^{ℓ}f^{(ℓ)}(r) \geq 0, \quad \text{for all } ℓ \in \mathbb{N}_0 \text{ and all } r > 0.$$ 

If in addition $f \in C[0, \infty)$ then $f$ is called completely monotone on $[0, \infty]$.

**Simple examples and motivation**

- $f(r) = \text{const.}, \quad \Rightarrow \quad (−1)^{ℓ}f^{(ℓ)}(r) = 0 \geq 0,$
- $f(r) = e^{-αr}, \alpha > 0, \quad \Rightarrow \quad (−1)^{ℓ}f^{(ℓ)}(r) = α^{ℓ}e^{-αr} \geq 0,$
- $f(r) = (1 + r)^{-β}, \beta > 0 \quad \Rightarrow \quad (−1)^{ℓ}f^{(ℓ)}(r) = β(β + 1) \cdots (β + ℓ − 1)(1 + r)^{−β−ℓ} \geq 0$

Now, if we define $φ(r) := f(r^2)$ then $φ$ is **positive definite**!
Let see again IMQ:

\[ f(r) = \frac{1}{1 + r} = \int_0^\infty e^{-sr} e^{-s} ds = \mathcal{L}\nu(r) \Rightarrow \phi(r) = \frac{1}{1 + r^2} = \int_0^\infty e^{-sr^2} e^{-s} ds \]

\[ \alpha^T A \alpha = \sum_{j,k=1}^N \alpha_j \alpha_k \frac{1}{1 + \|x_j - x_k\|^2} = \int_0^\infty \left( \sum_{j,k=1}^N \alpha_j \alpha_k e^{-s\|x_j - x_k\|^2} \right) e^{-s} ds \geq 0 \]

The same is true for \( f(r) = \frac{1}{\sqrt{1 + r}} \) with different \( \nu \),

\[ \frac{1}{\sqrt{1 + r}} = \int_0^\infty e^{-sr} \frac{e^{-s}}{\sqrt{\pi s}} ds = \mathcal{L}\nu(r) \Rightarrow \frac{1}{\sqrt{1 + r^2}} = \int_0^\infty e^{-sr^2} \frac{e^{-s}}{\sqrt{\pi s}} ds \]
Let see again IMQ:

\[ f(r) = \frac{1}{1 + r} = \int_0^\infty e^{-sr} \frac{e^{-s}}{d\nu(s)} ds = \mathcal{L}\nu(r) \Rightarrow \phi(r) = \frac{1}{1 + r^2} = \int_0^\infty e^{-sr^2} e^{-s} ds \]

\[ \alpha^T A \alpha = \sum_{j,k=1}^{N} \alpha_j \alpha_k \frac{1}{1 + \|x_j - x_k\|^2_2} = \int_0^\infty \left( \sum_{j,k=1}^{N} \alpha_j \alpha_k e^{-s\|x_j - x_k\|^2_2} \right) e^{-s} ds \geq 0 \]

The same is true for \( f(r) = \frac{1}{\sqrt{1+r}} \) with different \( \nu \),

\[ \frac{1}{\sqrt{1+r}} = \int_0^\infty e^{-sr} \frac{e^{-s}}{\sqrt{\pi s}} ds = \mathcal{L}\nu(r) \Rightarrow \frac{1}{\sqrt{1 + r^2}} = \int_0^\infty e^{-sr^2} \frac{e^{-s}}{\sqrt{\pi s}} ds \]
Theorem Bernstein (1914), Hausdorff (1921), Widder (1931)
A function \( f : [0, \infty) \to \mathbb{R} \) is completely monotone on \([0, \infty)\) if and only if it is the Laplace transform of a nonnegative finite Borel measure \( \nu \), i.e. it is of the form

\[
    f(r) = \mathcal{L}\nu(r) = \int_0^\infty e^{-sr} d\nu(s)
\]

The proof of \( \Leftarrow \) is simple:

\[
    (-1)^\ell f^{(\ell)}(r) = (-1)^\ell \frac{d^\ell}{dr^\ell} \int_0^\infty e^{-sr} d\nu(s) = \int_0^\infty s^\ell e^{-sr} d\nu(s) \geq 0.
\]

The other side \( \Rightarrow \) is technical.

Theorem Schoenberg (1938)
A function \( f : [0, \infty) \to \mathbb{R} \) is completely monotone on \([0, \infty)\) if and only if \( \Phi = f(\| \cdot \|_2^2) \) is positive semi-definite on every \( \mathbb{R}^d \). Moreover \( \Phi \) is positive definite if it is not a constant.
Theorem Bernstein (1914), Hausdorff (1921), Widder (1931)

A function \( f: [0, \infty) \rightarrow \mathbb{R} \) is completely monotone on \([0, \infty)\) if and only if it is the Laplace transform of a nonnegative finite Borel measure \( \nu \), i.e. it is of the form

\[
f(r) = \mathcal{L}\nu(r) = \int_0^\infty e^{-sr} d\nu(s)
\]

The proof of \( \Leftarrow \) is simple:

\[
(-1)^\ell f^{(\ell)}(r) = (-1)^\ell \frac{d^\ell}{dr^\ell} \int_0^\infty e^{-sr} d\nu(s) = \int_0^\infty s^\ell e^{-sr} d\nu(s) \geq 0.
\]

The other side \( \Rightarrow \) is technical.

Theorem Schoenberg (1938)

A function \( f: [0, \infty) \rightarrow \mathbb{R} \) is completely monotone on \([0, \infty)\) if and only if \( \Phi = f(|| \cdot ||_2^2) \) is positive semi-definite on every \( \mathbb{R}^d \). Moreover \( \Phi \) is positive definite if it is not a constant.
Theorem Bernstein (1914), Hausdorff (1921), Widder (1931)

A function \( f : [0, \infty) \to \mathbb{R} \) is completely monotone on \([0, \infty)\) if and only if it is the Laplace transform of a nonnegative finite Borel measure \( \nu \), i.e. it is of the form

\[
f(r) = \mathcal{L}\nu(r) = \int_{0}^{\infty} e^{-sr} d\nu(s)
\]

The proof of \( \iff \) is simple:

\[
(-1)^{\ell} f^{(\ell)}(r) = (-1)^{\ell} \frac{d^{\ell}}{dr^{\ell}} \int_{0}^{\infty} e^{-sr} d\nu(s) = \int_{0}^{\infty} s^{\ell} e^{-sr} d\nu(s) \geq 0.
\]

The other side \( \Rightarrow \) is technical.

Theorem Schoenberg (1938)

A function \( f : [0, \infty) \to \mathbb{R} \) is completely monotone on \([0, \infty)\) if and only if \( \Phi = f(|| \cdot ||^{2}_{2}) \) is positive semi-definite on every \( \mathbb{R}^{d} \). Moreover \( \Phi \) is positive definite if it is not a constant.
Theorem: Bernstein (1914), Hausdorff (1921), Widder (1931), Schoenberg (1938)

For a function \( \phi : [0, \infty) \to \mathbb{R} \) the following properties are equivalent:

1. \( \phi \) is positive definite on every \( \mathbb{R}^d \);
2. \( \phi(\sqrt{\cdot}) \) is completely monotone on \([0, \infty)\) and not constant;
3. there exist a finite nonnegative Borel measure \( \nu \) on \([0, \infty)\) that is not concentrated at zero, such that

\[
\phi(r) = \int_0^\infty e^{-r^2 t} d\nu(t).
\]
Examples

- **Gaussian:** $\phi(r) = e^{-\alpha r^2}$:

  \[ \alpha = 0.5 \]
  \[ \alpha = 1.0 \]
  \[ \alpha = 2.0 \]

- **Inverse Multiquadrics:** $\phi(r) = (1 + \epsilon^2 r^2)^{-\beta}$, $\beta > 0$:

  \[ \beta = 0.5 \]
  \[ \epsilon = 2 \]
  \[ \epsilon = 5 \]
  \[ \epsilon = 8 \]
Conditionally Positive Definite (CPD) Functions

Definition
A continuous function $\Phi : \mathbb{R}^d \to \mathbb{R}$ is said to be conditionally positive semi-definite of order $m$ if, for all $N \in \mathbb{N}$, all pairwise distinct centers $x_1, \ldots, x_N \in \mathbb{R}^d$, and all $\alpha \in \mathbb{R}^N$ satisfying

$$
\sum_{j=1}^{N} \alpha_j p(x_j) = 0, \quad p \in P_{m-1}(\mathbb{R}^d),
$$

the quadratic form

$$
\sum_{k,j=1}^{N} \alpha_j \alpha_k \Phi(x_j - x_k)
$$

is nonnegative. $\Phi$ is said to be conditionally positive definite (CPD) of order $m$ if the quadratic form is positive, unless $\alpha$ is zero.
Interpolation Problem

\[ V_N = \text{span}\{ \Phi(\cdot - x_1), \Phi(\cdot - x_2), \ldots, \Phi(\cdot - x_N) \} + P_{m-1}^d \]

\[ s(x) = s_{f,x}(x) = \sum_{k=1}^{N} c_k \Phi(x - x_k) + \sum_{j=1}^{Q} b_k p_k(x), \quad Q = \text{dim}(P_{m-1}^d) \]

**Interpolation conditions:**

\[ s(x_k) = f_k, \quad k = 1, 2, \ldots, N, \quad \text{with additional conditions} \sum_{j=1}^{N} c_j p_k(x_j) = 0, \quad k = 1, \ldots, Q. \]

**The final linear system:**

\[
\begin{bmatrix}
A \\
PT \\
0
\end{bmatrix}
\begin{bmatrix}
c \\
b
\end{bmatrix} =
\begin{bmatrix}
f | X \\
0
\end{bmatrix}, \quad A = (\Phi(x_k - x_j)) \in \mathbb{R}^{N \times N}, \quad P = (p_k(x_j)) \in \mathbb{R}^{N \times Q}
\]

**Theorem:** If \( \Phi \) is CPD of order \( m \) and \( X \) is \( P_{m-1}^d \)-unisolvent the interpolation problem is well-defined.
Interpolation Problem

\[ V_N = \text{span}\{ \Phi(\cdot - x_1), \Phi(\cdot - x_2), \ldots, \Phi(\cdot - x_N) \} + \mathbb{P}_m^{d-1} \]

\[ s(x) = s_{f,X}(x) = \sum_{k=1}^{N} c_k \Phi(x - x_k) + \sum_{j=1}^{Q} b_k p_k(x), \quad Q = \text{dim}(\mathbb{P}_m^{d-1}) \]

Interpolation conditions:

\[ s(x_k) = f_k, \quad k = 1, 2, \ldots, N, \quad \text{with additional conditions} \quad \sum_{j=1}^{N} c_j p_k(x_j) = 0, \quad k = 1, \ldots, Q. \]

The final linear system:

\[
\begin{bmatrix}
A & P \\
PT & 0
\end{bmatrix}
\begin{bmatrix}
c \\
b
\end{bmatrix} =
\begin{bmatrix}
f|X \\
0
\end{bmatrix}, \quad A = (\Phi(x_k - x_j)) \in \mathbb{R}^{N \times N}, \quad P = (p_k(x_j)) \in \mathbb{R}^{N \times Q}
\]

**Theorem:** If \( \Phi \) is CPD of order \( m \) and \( X \) is \( \mathbb{P}_m^{d-1} \)-unisolvent the interpolation problem is well-defined.
Shift invariant functions

**Theorem:** Suppose $\Phi : \mathbb{R}^d \to \mathbb{R}$ is continuous, slowly increasing, and possesses a generalized Fourier transform of order $m$, which is continuous on $\mathbb{R}^d \setminus \{0\}$. Then $\Phi$ is conditionally positive definite of order $m$ if and only if $\hat{\Phi}$ is nonnegative and nonvanishing.

Radial functions

**Theorem:** Micchelli (1986), Guo, Hu, Sun (1993)

Suppose that $\phi \in C[0, \infty] \cap C^\infty(0, \infty)$ is given. Then the function $\Phi = \phi(\| \cdot \|^2_2)$ is conditionally positive definite of order $m \in \mathbb{N}_0$ on every $\mathbb{R}^d$ if and only if $(-1)^m \phi^{(m)}$ is completely monotone on $(0, \infty)$ and $\phi \notin \mathbb{P}^d_m$.

**Examples:**

- **Maltiquadrics (MQ):** $\phi(r) = (-1)^{\lceil \beta \rceil} (c + r^2)^\beta$, $c > 0$, $\beta > 0$, $\beta \notin \mathbb{N}_0$ are CPD of order $m = \lceil \beta \rceil$ on every $\mathbb{R}^d$

- **Radial powers:** $\phi(r) = (-1)^{\lceil \beta/2 \rceil} r^\beta$, $\beta > 0$, $\beta \notin 2\mathbb{N}$ are CPD of order $m = \lceil \beta/2 \rceil$ on every $\mathbb{R}^d$

- **Thin plate splines (TPS):** $\phi(r) = (-1)^{k+1} r^{2k} \log(r)$ are CPD of order $m = k + 1$ on every $\mathbb{R}^d$
Approximation by Kernels
Solving PDEs by Kernels
Approximation by MLS
GMLS Approximation and DMLPG Methods
Main References

Approximation by Kernels

Interpolation Problem
Positive Definite Kernels
Conditionally Positive Definite Functions
Compactly Supported Basis Functions
Native Spaces
Error Analysis and Stability

Shift invariant functions

**Theorem:** Suppose $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ is continuous, slowly increasing, and possesses a generalized Fourier transform of order $m$, which is continuous on $\mathbb{R}^d \setminus \{0\}$. Then $\Phi$ is conditionally positive definite of order $m$ if and only if $\hat{\Phi}$ is nonnegative and nonvanishing.

Radial functions

**Theorem:** Micchelli (1986), Guo, Hu, Sun (1993)

Suppose that $\phi \in C[0, \infty] \cap C^\infty(0, \infty)$ is given. Then the function $\Phi = \phi(\| \cdot \|_2^2)$ is conditionally positive definite of order $m \in \mathbb{N}_0$ on every $\mathbb{R}^d$ if and only if $(-1)^m \phi^{(m)}$ is completely monotone on $(0, \infty)$ and $\phi \notin \mathbb{P}_m$.

Examples:

- **Maltiquadrics (MQ):** $\phi(r) = (-1)^{\lceil \beta \rceil} (c + r^2)^\beta$, $c > 0$, $\beta > 0$, $\beta \notin \mathbb{N}_0$ are CPD of order $m = \lceil \beta \rceil$ on every $\mathbb{R}^d$
- **Radial powers:** $\phi(r) = (-1)^{\lceil \beta/2 \rceil} r^\beta$, $\beta > 0$, $\beta \notin 2\mathbb{N}$ are CPD of order $m = \lceil \beta/2 \rceil$ on every $\mathbb{R}^d$
- **Thin plate splines (TPS):** $\phi(r) = (-1)^{k+1} r^{2k} \log(r)$ are CPD of order $m = k + 1$ on every $\mathbb{R}^d$
Shift invariant functions

**Theorem:** Suppose $\Phi : \mathbb{R}^d \to \mathbb{R}$ is continuous, slowly increasing, and possesses a generalized Fourier transform of order $m$, which is continuous on $\mathbb{R}^d \setminus \{0\}$. Then $\Phi$ is conditionally positive definite of order $m$ if and only if $\hat{\Phi}$ is nonnegative and nonvanishing.

Radial functions

**Theorem:** Micchelli (1986), Guo, Hu, Sun (1993)

Suppose that $\phi \in C[0, \infty] \cap C^\infty (0, \infty)$ is given. Then the function $\Phi = \phi(\| \cdot \|_2^2)$ is conditionally positive definite of order $m \in \mathbb{N}_0$ on every $\mathbb{R}^d$ if and only if $(-1)^m \phi^{(m)}$ is completely monotone on $(0, \infty)$ and $\phi \notin \mathbb{P}_m^d$.

**Examples:**

- **Maltiquadrics (MQ):** $\phi(r) = (-1)^\lceil \beta \rceil (c + r^2)^\beta$, $c > 0$, $\beta > 0$, $\beta \notin \mathbb{N}_0$ are CPD of order $m = \lceil \beta \rceil$ on every $\mathbb{R}^d$

- **Radial powers:** $\phi(r) = (-1)^\lceil \beta/2 \rceil r^\beta$, $\beta > 0$, $\beta \notin 2\mathbb{N}$ are CPD of order $m = \lceil \beta/2 \rceil$ on every $\mathbb{R}^d$

- **Thin plate splines (TPS):** $\phi(r) = (-1)^{k+1} r^{2k} \log(r)$ are CPD of order $m = k + 1$ on every $\mathbb{R}^d$
Plots of some (I)MQ

\[ \phi(r) = (1 + \varepsilon^2 r^2)^\beta \]

\[ (-1)^{\lceil \beta \rceil} \phi(r) = (1 + \varepsilon^2 r^2)^\beta \]

- \( \beta = -0.5 \)
  - \( \varepsilon = 2 \)
  - \( \varepsilon = 5 \)
- \( \beta = 0.5 \)
  - \( \varepsilon = 2 \)
  - \( \varepsilon = 1 \)

\[ m = \lceil \beta \rceil = 0 \]
\[ m = \lceil \beta \rceil = 1 \]
Plots of polyharmonics

\[ \phi(r) = (-1)^{\left\lceil \beta/2 \right\rceil} r^\beta \]

- \( \beta = 3.0 \)
  - \( m = \left\lceil \beta/2 \right\rceil = 2 \)

- \( \beta = 0.5 \)
  - \( m = \left\lceil \beta/2 \right\rceil = 1 \)

\[ \phi(r) = (-1)^{k+1} r^{2k} \log r \]

- \( k = 1 \)
  - \( m = k + 1 = 2 \)

- \( k = 2 \)
  - \( m = 3 \)
Compactly Supported Basis Functions

Motivation

- Motivated by B-splines from univariate approximation theory
- Sparse interpolation matrix
- Fast evaluation of interpolant

History

- Schaback and Wendland 1994 (Euclid’s hat, convolution of characteristic function with itself on unit ball)
- Wu 1995 (the first general construction: Wu’s functions)
- Wendland 1995 (minimal degree functions: Wendland’s functions)
- Buhmann 1998, 2001 (a new class)
- Schaback 2010 (missing Wendland’s functions)
Compactly Supported Basis Functions

Motivation

- Motivated by B-splines from univariate approximation theory
- Sparse interpolation matrix
- Fast evaluation of interpolant

History

- Schaback and Wendland 1994 (Euclid’s hat, convolution of characteristic function with itself on unit ball)
- Wu 1995 (the first general construction: Wu’s functions)
- Wendland 1995 (minimal degree functions: Wendland’s functions)
- Buhmann 1998, 2001 (a new class)
- Schaback 2010 (missing Wendland’s functions)
Compactly Supported Basis Functions

Some Theorems

**Theorem:** A continuous and conditionally positive definite and compactly supported RBF is necessarily of order $m = 0$.

**Theorem:** A continuous, univariate, and compactly supported function $\phi$ can not be positive definite on every $\mathbb{R}^d$.

**Conclusion:** Among being compactly supported and dimension one of them should be scarified!

Piecewise polynomial functions with local support:

$$\phi(r) = \begin{cases} p(r), & 0 \leq r \leq 1, \\ 0, & r > 1. \end{cases}$$

degree of $\phi := \text{degree of } p$
Compactly Supported Basis Functions

Some Theorems

**Theorem:** A continuous and conditionally positive definite and compactly supported RBF is necessarily of order $m = 0$.

**Theorem:** A continuous, univariate, and compactly supported function $\phi$ can not be positive definite on every $\mathbb{R}^d$.

**Conclusion:** Among being compactly supported and dimension one of them should be scarified!

Piecewise polynomial functions with local support:

$$
\phi(r) = \begin{cases} 
p(r), & 0 \leq r \leq 1, \\
0, & r > 1.
\end{cases}
$$

degree of $\phi :=$ degree of $p$
Compactly Supported Basis Functions

Some Theorems

**Theorem:** A continuous and conditionally positive definite and compactly supported RBF is necessarily of order $m = 0$.

**Theorem:** A continuous, univariate, and compactly supported function $\phi$ can not be positive definite on every $\mathbb{R}^d$.

**Conclusion:** Among being compactly supported and dimension one of them should be scarified!

**Piecewise polynomial functions with local support:**

$$
\phi(r) = \begin{cases} 
p(r), & 0 \leq r \leq 1, \\
0, & r > 1.
\end{cases}
$$

degree of $\phi := \text{degree of } p$
Beautiful Wu’s and Wendland’s constructions

**Dimension Walks:**
If $\Phi = \phi(\|\cdot\|_2)$ then

$$
\hat{\Phi}(\omega) = \mathcal{F}_d \phi(r) = r^{-(d-2)/2} \int_0^\infty \phi(t) t^{d/2} J_{(d-2)/2}(rt) dt,
\quad r = \|\omega\|_2.
$$

If $\Phi$ is radial then $\hat{\Phi}$ is so.

**Definition:** *montée and descentée* operators, Matheron 1965, Wu 1995

$$
(I \phi)(r) := \int_r^\infty t \phi(t) dt, \quad t \mapsto t \phi(t) \in L_1[0, \infty), \quad r \geqslant 0,
$$

$$
(D \phi)(r) := -\frac{1}{r} \phi'(r), \quad \phi \in C^2(\mathbb{R}), \quad r \geqslant 0.
$$
Beautiful Wu’s and Wendland’s constructions

**Dimension Walks:**
If $\Phi = \phi(\| \cdot \|_2)$ then

$$\hat{\Phi}(\omega) = \mathcal{F}_d \phi(r) = r^{-(d-2)/2} \int_0^\infty \phi(t) t^{d/2} J_{(d-2)/2}(rt) dt, \quad r = \| \omega \|_2.$$ 

If $\Phi$ is radial then $\hat{\Phi}$ is so.

**Definition:** *montée* and *descentée* operators, Matheron 1965, Wu 1995

$$(\mathcal{I} \phi)(r) := \int_r^\infty t \phi(t) dt, \quad t \mapsto t \phi(t) \in L_1[0, \infty), \quad r \geq 0,$$

$$(\mathcal{D} \phi)(r) := -\frac{1}{r} \phi'(r), \quad \phi \in C^2(\mathbb{R}), \quad r \geq 0.$$
**Theorem:** If $\phi$ is continuous then

- If $t \mapsto t^{d-1} \phi(t) \in L_1[0, \infty)$ and $d \geq 3$ then $F_d(\phi) = F_{d-2}(I \phi)$.
- If $\phi \in C^2(\mathbb{R})$ is even $t \mapsto t^d \phi'(t) \in L_1[0, \infty)$ then $F_d(\phi) = F_{d+2}(D \phi)$.

**Conclusion:**

\[ \phi \text{ is positive definite in } \mathbb{R}^d \iff I \phi \text{ is positive definite in } \mathbb{R}^{d-2} \]
\[ \phi \text{ is positive definite in } \mathbb{R}^d \iff D \phi \text{ is positive definite in } \mathbb{R}^{d+2} \]

Another **Theorem:**

\[ \phi \in C^{2k}(0) \cap C^\ell(1) \Rightarrow \begin{cases} 
I \phi \in C^{2k+2}(0) \cap C^{\ell+1}(1) \\
D \phi \in C^{2k-2}(0) \cap C^{\ell-1}(1) 
\end{cases} \]
Wendland’s functions 1995

$\phi_\ell(r) := (1 - r)_{+}^\ell$, is PD on $\mathbb{R}^d$ provided that $\ell \geq \lfloor d/2 \rfloor + 1$

Thus $\phi_{\lfloor d/2 \rfloor + k + 1}$ is PD on $\mathbb{R}^{d+2k}$.

Definition:

$\phi_{d,k} := \mathcal{I}^k \phi_{\lfloor d/2 \rfloor + k + 1}$

- degree of $\phi_{d,k} = \lfloor d/2 \rfloor + 3k + 1$
- they are positive definite on $\mathbb{R}^d$ (not on every $\mathbb{R}^d$)
- $\phi_{d,k} \in C^{2k}(\mathbb{R})$
- for given dimension $d$ and smoothness $2k$ they are of minimal degree
- their Fourier transform decay polynomially!
Wendland’s functions 1995

\[ \phi_\ell(r) := (1 - r)^\ell_+ \text{, is PD on } \mathbb{R}^d \text{ provided that } \ell \geq \lfloor d/2 \rfloor + 1 \]

Thus \( \phi_{\lfloor d/2 \rfloor + k + 1} \) is PD on \( \mathbb{R}^{d+2k} \).

\[ \phi(r) = (1 - r)^2_+ \]

Definition:

\[ \phi_{d,k} := I^k \phi_{\lfloor d/2 \rfloor + k + 1} \]

- degree of \( \phi_{d,k} = \lfloor d/2 \rfloor + 3k + 1 \)
- they are positive definite on \( \mathbb{R}^d \) (not on every \( \mathbb{R}^d \))
- \( \phi_{d,k} \in C^{2k}(\mathbb{R}) \)
- for given dimension \( d \) and smoothness \( 2k \) they are of minimal degree
- their Fourier transform decay polynomially!
Wendland’s functions 1995
\[ \phi_{\ell}(r) := (1 - r)^\ell_+ \] is PD on \( \mathbb{R}^d \) provided that \( \ell \geq \lfloor d/2 \rfloor + 1 \)
Thus \( \phi_{\lfloor d/2 \rfloor + k + 1} \) is PD on \( \mathbb{R}^{d+2k} \).

\[ \phi(r) = (1 - r)^2_+ \]

**Definition:**
\[ \phi_{d,k} := \mathcal{I}^k \phi_{\lfloor d/2 \rfloor + k + 1} \]

- degree of \( \phi_{d,k} = \lfloor d/2 \rfloor + 3k + 1 \)
- they are positive definite on \( \mathbb{R}^d \) (not on every \( \mathbb{R}^d \))
- \( \phi_{d,k} \in C^{2k}(\mathbb{R}) \)
- for given dimension \( d \) and smoothness \( 2k \) they are of minimal degree
- their Fourier transform decay polynomially!
Wu’s functions 1995

Wu started with \( \psi(r) = (1 - r^2)^\ell \) which is not PD itself. He defined

\[
\psi_\ell(r) := (\psi \ast \psi)(2r)
\]

\( \psi_\ell \) belongs to \( C^{2\ell}(\mathbb{R}) \) and its degree is \( 4\ell + 1 \) and it is PD on \( \mathbb{R}^1 \).

\[
\psi_{k,\ell} := D^k \psi_\ell
\]
Wu’s functions 1995
Wu started with $\psi(r) = (1 - r^2)^{\ell}_+$ which is not PD itself. He defined

$$\psi_\ell(r) := (\psi \ast \psi)(2r)$$

$\psi_\ell$ belongs to $C^{2\ell}(\mathbb{R})$ and its degree is $4\ell + 1$ and it is PD on $\mathbb{R}^1$.

$$\psi_{k,\ell} := D^k \psi_\ell$$

- degree of $\psi_{k,\ell} = 4\ell - 2k + 1$
- they are positive definite on $\mathbb{R}^{2k+1}$ (not on every $\mathbb{R}^d$)
- $\psi_{k,\ell} \in C^{2\ell-2k}(\mathbb{R})$
- they are not of minimal degree
Wu’s functions 1995
Wu started with \( \psi(r) = (1 - r^2)_+^\ell \) which is not PD itself. He defined

\[
\psi_\ell(r) := (\psi \ast \psi)(2r)
\]

\( \psi_\ell \) belongs to \( C^{2\ell}(\mathbb{R}) \) and its degree is \( 4\ell + 1 \) and it is PD on \( \mathbb{R}^1 \).

\[
\psi_{k,\ell} := \mathcal{D}^k \psi_\ell
\]

- degree of \( \psi_{k,\ell} = 4\ell - 2k + 1 \)
- they are positive definite on \( \mathbb{R}^{2k+1} \) (not on every \( \mathbb{R}^d \))
- \( \psi_{k,\ell} \in C^{2\ell-2k}(\mathbb{R}) \)
- they are not of minimal degree
**Table:** Wendland’s functions for $d = 3$

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\phi_{3,k}$</th>
<th>smoothness</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$(1 - r)^2_+$</td>
<td>$C^0$</td>
</tr>
<tr>
<td>1</td>
<td>$(1 - r)^4_+ (4r + 1)$</td>
<td>$C^2$</td>
</tr>
<tr>
<td>2</td>
<td>$(1 - r)^6_+ (35r^2 + 18r + 3)$</td>
<td>$C^4$</td>
</tr>
<tr>
<td>3</td>
<td>$(1 - r)^8_+ (32r^3 + 25r^2 + 8r + 1)$</td>
<td>$C^6$</td>
</tr>
</tbody>
</table>
Approximation by Kernels
Solving PDEs by Kernels
Approximation by MLS
GMLS Approximation and DMLPG Methods
Main References
Interpolation Problem
Positive Definite Kernels
Conditionally Positive Definite Functions
Compactly Supported Basis Functions
Native Spaces
Error Analysis and Stability

Table: Wu’s functions for $\ell = 3$

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\psi_{k,3}$</th>
<th>smoothness</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$(1 - r)^3(5 + 35r + 101r^2 + 147r^3 + 101r^4 + 35r^5 + 5r^6)$</td>
<td>$C^6$</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>$(1 - r)^6(6 + 36r + 82r^2 + 72r^3 + 30r^4 + 5r^5)$</td>
<td>$C^4$</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>$(1 - r)^5(8 + 40r + 48r^2 + 25r^3 + 5r^4)$</td>
<td>$C^2$</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>$(1 - r)^4(16 + 29r + 20r^2 + 5r^3)$</td>
<td>$C^0$</td>
<td>7</td>
</tr>
</tbody>
</table>

Davoud Mirzaei
Meshless Approximants
Other constructions

**An idea:** If a compactly supported PD function $\Phi$ and an operator $T$ which is nonnegative in Fourier domain are at hand then

$$T\Phi$$

is PD. If $T$ respects the compact support then $T\Phi$ is compact support and PD.

**An example:** $T = \Delta$

Let $\Phi \in C^2(\mathbb{R}^d) \cap L_1(\mathbb{R}^d)$ is PD. If $\Phi$ is radial so is $\Delta \Phi$. If $\Phi$ is CS, so is $\Delta \Phi$, and finally

$$\hat{\Delta \Phi} = -\| \cdot \|_2 \hat{\Phi},$$

Thus $-\Delta \Phi$ is CS and PD.
Relation between Sobolev spaces and kernels

Let $\Omega$ be an open set in $\mathbb{R}^d$ and $k \in \mathbb{N}_0$,

$$H^k(\Omega) = \{ f \in L_2(\Omega) : D^\alpha f \in L_2(\Omega), \ |\alpha| = k \}$$

$$(f, g)_{H^k(\Omega)} := \sum_{|\alpha| \leq k} \|D^\alpha f\|_{L^2(\Omega)}^2.$$

Definition via Fourier transform on whole $\mathbb{R}^d$ for $s \geq 0$:

$$H^s(\mathbb{R}^d) = \{ f \in L_2(\mathbb{R}^d) : \hat{f}(\cdot)(1 + \|\cdot\|_2^2)^{s/2} \in L_2(\mathbb{R}^d) \}$$

$$(f, g)_{H^s(\mathbb{R}^d)} := (2\pi)^{-d/2} \int_{\mathbb{R}^d} \hat{f}(\omega)\overline{\hat{g}(\omega)}(1 + \|\omega\|_2^2)^s d\omega.$$

Sobolev embedding: If $s > d/2$ then $H^s(\mathbb{R}^d) \subseteq C(\mathbb{R}^d)$. 
Relation between Sobolev spaces and kernels

Let $\Omega$ be an open set in $\mathbb{R}^d$ and $k \in \mathbb{N}_0$,

$$H^k(\Omega) = \{ f \in L_2(\Omega) : D^\alpha f \in L_2(\Omega), \ |\alpha| = k \}$$

$$(f, g)_{H^k(\Omega)} := \sum_{|\alpha| \leq k} \|D^\alpha f\|_{L_2(\Omega)}^2.$$

Definition via Fourier transform on whole $\mathbb{R}^d$ for $s \geq 0$:

$$H^s(\mathbb{R}^d) = \{ f \in L_2(\mathbb{R}^d) : \hat{f}(\cdot)(1 + \| \cdot \|_2^2)^{s/2} \in L_2(\mathbb{R}^d) \}$$

$$(f, g)_{H^s(\mathbb{R}^d)} := (2\pi)^{-d/2} \int_{\mathbb{R}^d} \hat{f}(\omega)\overline{\hat{g}(\omega)}(1 + \|\omega\|_2^2)^s\,d\omega.$$

Sobolev embedding: If $s > d/2$ then $H^s(\mathbb{R}^d) \subseteq C(\mathbb{R}^d)$. 

Davoud Mirzaei
Meshless Approximants
Relation between Sobolev spaces and kernels

Let $\Omega$ be an open set in $\mathbb{R}^d$ and $k \in \mathbb{N}_0$,

$$H^k(\Omega) = \{ f \in L_2(\Omega) : D^\alpha f \in L_2(\Omega), \ |\alpha| = k \}$$

$$(f, g)_{H^k(\Omega)} := \sum_{|\alpha| \leq k} \|D^\alpha f\|^2_{L_2(\Omega)}.$$

Definition via Fourier transform on whole $\mathbb{R}^d$ for $s \geq 0$:

$$H^s(\mathbb{R}^d) = \{ f \in L_2(\mathbb{R}^d) : \hat{f}(\cdot)(1 + \| \cdot \|_2^2)^{s/2} \in L_2(\mathbb{R}^d) \}$$

$$(f, g)_{H^s(\mathbb{R}^d)} := (2\pi)^{-d/2} \int_{\mathbb{R}^d} \hat{f}(\omega) \overline{\hat{g}(\omega)}(1 + \|\omega\|_2^2)^s d\omega.$$

Sobolev embedding: If $s > d/2$ then $H^s(\mathbb{R}^d) \subseteq C(\mathbb{R}^d)$. 
Definition via **Fourier transform** on whole $\mathbb{R}^d$ for $s \geq 0$:

$$H^s(\mathbb{R}^d) = \{f \in L_2(\mathbb{R}^d) : \hat{f}(\cdot)(1 + \| \cdot \|_2^2)^{s/2} \in L_2(\mathbb{R}^d)\}$$

$$(f, g)_{H^s(\mathbb{R}^d)} := (2\pi)^{-d/2} \int_{\mathbb{R}^d} \hat{f}(\omega)\overline{\hat{g}(\omega)}(1 + \|\omega\|_2^2)^s d\omega.$$ 

**Sobolev embedding**: If $s > d/2$ then $H^s(\mathbb{R}^d) \subseteq C(\mathbb{R}^d)$.

For $s > d/2$ define $\Phi$ such that $\hat{\Phi}(\omega) = (1 + \|\omega\|_2^2)^{-s}$. We saw that

$$\Phi(x) = \frac{2^{1-s}}{\Gamma(s)} \|x\|_2^{s-d/2} K_{d/2-s}(\|\omega\|_2).$$

$$(f, \Phi(\cdot - x))_{H^s(\mathbb{R}^d)} = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \hat{f}(\omega) \overline{\hat{\Phi}(\omega)} e^{-ix^T\omega} \hat{\Phi}(\omega)^{-1} d\omega = f(x).$$
Definition via **Fourier transform** on whole $\mathbb{R}^d$ for $s \geq 0$:

$$H^s(\mathbb{R}^d) = \{ f \in L_2(\mathbb{R}^d) : \hat{f}(\cdot)(1 + \| \cdot \|_2^2)^{s/2} \in L_2(\mathbb{R}^d) \}$$

$$(f, g)_{H^s(\mathbb{R}^d)} := (2\pi)^{-d/2} \int_{\mathbb{R}^d} \hat{f}(\omega) \overline{\hat{g}(\omega)} (1 + \|\omega\|^2_2)^s d\omega.$$

**Sobolev embedding:** If $s > d/2$ then $H^s(\mathbb{R}^d) \subseteq C(\mathbb{R}^d)$.

For $s > d/2$ define $\Phi$ such that $\hat{\Phi}(\omega) = (1 + \|\omega\|^2_2)^{-s}$. We saw that

$$\Phi(x) = \frac{2^{1-s}}{\Gamma(s)} \|x\|_2^{s-d/2} K_{d/2-s}(\|\omega\|_2).$$

$$(f, \Phi(\cdot - x))_{H^s(\mathbb{R}^d)} = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \hat{f}(\omega) \overline{\hat{\Phi}(\omega)} e^{-ix^T\omega} \hat{\Phi}(\omega) d\omega = f(x).$$
Reproducing kernel Hilbert space (RKHS)

**Definition:**
Let \( \mathcal{H} \) be a real Hilbert space of functions \( f : \Omega \to \mathbb{R} \). A function \( \Phi : \Omega \times \Omega \to \mathbb{R} \) is called a reproducing kernel for \( \mathcal{H} \) if

- \( \Phi(\cdot, x) \in \mathcal{H} \) for all \( x \in \Omega \),
- \( f(x) = (f, \Phi(\cdot, x))_{\mathcal{H}} \) for all \( f \in \mathcal{H} \) and \( x \in \Omega \)

The reproducing kernel of a Hilbert space is unique and

\[
\delta_x \in \mathcal{H}^* \quad (|\delta_x(f)| = |f(x)| \leq M\|f\|_{\mathcal{H}}) \iff \mathcal{H} \text{ has a reproducing kernel}
\]

**Theorem:**
If \( \mathcal{H} \) is a Hilbert space with reproducing kernel \( \Phi \) then

- \( \Phi(x, y) = (\Phi(\cdot, x), \Phi(\cdot, y))_{\mathcal{H}} \),
- \( \Phi(x, y) = \Phi(y, x) \),
- convergence in \( \mathcal{H} \)-norm implies pointwise convergence,
- \( \Phi \) is semi-positive definite,
- if \( \delta_x \) are linearly independent in \( \mathcal{H}^* \) then \( \Phi \) is positive definite.
Reproducing kernel Hilbert space (RKHS)

**Definition:**
Let $\mathcal{H}$ be a real Hilbert space of functions $f : \Omega \to \mathbb{R}$. A function $\Phi : \Omega \times \Omega \to \mathbb{R}$ is called a reproducing kernel for $\mathcal{H}$ if

- $\Phi(\cdot, x) \in \mathcal{H}$ for all $x \in \Omega$,
- $f(x) = (f, \Phi(\cdot, x))_\mathcal{H}$ for all $f \in \mathcal{H}$ and $x \in \Omega$

The reproducing kernel of a Hilbert space is unique and

$$\delta_x \in \mathcal{H}^* \quad (|\delta_x(f)| = |f(x)| \leq M\|f\|_\mathcal{H}) \iff \mathcal{H} \text{ has a reproducing kernel}$$

**Theorem:**
If $\mathcal{H}$ is a Hilbert space with reproducing kernel $\Phi$ then

- $\Phi(x, y) = (\Phi(\cdot, x), \Phi(\cdot, y))_\mathcal{H}$,
- $\Phi(x, y) = \Phi(y, x)$,
- convergence in $\mathcal{H}$-norm implies pointwise convergence,
- $\Phi$ is semi-positive definite,
- if $\delta_x$ are linearly independent in $\mathcal{H}^*$ then $\Phi$ is positive definite.
Reproducing kernel Hilbert space (RKHS)

**Definition:**
Let $\mathcal{H}$ be a real Hilbert space of functions $f : \Omega \to \mathbb{R}$. A function $\Phi : \Omega \times \Omega \to \mathbb{R}$ is called a reproducing kernel for $\mathcal{H}$ if

- $\Phi(\cdot, x) \in \mathcal{H}$ for all $x \in \Omega$,
- $f(x) = (f, \Phi(\cdot, x))_{\mathcal{H}}$ for all $f \in \mathcal{H}$ and $x \in \Omega$

The reproducing kernel of a Hilbert space is unique and

$$\delta_x \in \mathcal{H}^* \ (|\delta_x(f)| = |f(x)| \leq M\|f\|_{\mathcal{H}}) \iff \mathcal{H} \text{ has a reproducing kernel}$$

**Theorem:**
If $\mathcal{H}$ is a Hilbert space with reproducing kernel $\Phi$ then

- $\Phi(x, y) = (\Phi(\cdot, x), \Phi(\cdot, y))_{\mathcal{H}}$,
- $\Phi(x, y) = \Phi(y, x)$,
- convergence in $\mathcal{H}$-norm implies pointwise convergence,
- $\Phi$ is semi-positive definite,
- if $\delta_x$ are linearly independent in $\mathcal{H}^*$ then $\Phi$ is positive definite.
Native space for PD kernels

If \( f = \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j) \) then

\[
(f, \Phi(\cdot, x))_\mathcal{H} = \left( \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j), \Phi(\cdot, x) \right)_\mathcal{H} = \sum_{j=1}^{N} \alpha_j (\Phi(\cdot, x_j), \Phi(\cdot, x))_\mathcal{H}
\]

\[
= \sum_{j=1}^{N} \alpha_j \Phi(x, x_j) = f(x)
\]

Thus the span of all \( \Phi(\cdot, x_j) \) reproduces by \( \Phi \). This motivates to define:

\[
H_\Phi(\Omega) := \text{span}\{ \Phi(\cdot, x) : x \in \Omega \}
\]

\[
(f, g)_\Phi := \sum_{j=1}^{N} \sum_{k=1}^{M} \alpha_j \beta_k \Phi(x_j, y_k), \quad f = \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j), \quad g = \sum_{k=1}^{M} \beta_k \Phi(\cdot, y_k).
\]

\[
\mathcal{N}_\Phi(\Omega) := \overline{H_\Phi(\Omega)}
\]
Native space for PD kernels

If \( f = \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j) \) then

\[
(f, \Phi(\cdot, x))_{\mathcal{H}} = \left( \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j), \Phi(\cdot, x) \right)_{\mathcal{H}} = \sum_{j=1}^{N} \alpha_j \left( \Phi(\cdot, x_j), \Phi(\cdot, x) \right)_{\mathcal{H}}
\]

\[
= \sum_{j=1}^{N} \alpha_j \Phi(x, x_j) = f(x)
\]

Thus the span of all \( \Phi(\cdot, x_j) \) reproduces by \( \Phi \). This motivates to define:

\[
H_{\Phi}(\Omega) := \text{span}\{ \Phi(\cdot, x) : x \in \Omega \}
\]

\[
(f, g)_{\Phi} := \sum_{j=1}^{N} \sum_{k=1}^{M} \alpha_j \beta_k \Phi(x_j, y_k), \quad f = \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j), \quad g = \sum_{k=1}^{M} \beta_k \Phi(\cdot, y_k).
\]

\[
\mathcal{N}_{\Phi}(\Omega) := \overline{H_{\Phi}(\Omega)}
\]
Special case $\Phi(x - y)$

Let $\Phi \in C(\mathbb{R}^d) \cap L_1(\mathbb{R}^d)$ is real-valued and PD. We have

$$\mathcal{N}_\Phi(\mathbb{R}^d) = \left\{ f \in L_2(\mathbb{R}^d) \cap C(\mathbb{R}^d) : \frac{\hat{f}}{\sqrt{\Phi}} \in L_2(\mathbb{R}^d) \right\}$$

$$(f, g)_{\mathcal{N}_\Phi(\mathbb{R}^d)} := (2\pi)^{-d/2} \left( \frac{\hat{f}}{\sqrt{\Phi}}, \frac{\hat{g}}{\sqrt{\Phi}} \right) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \frac{\hat{f}(\omega)\hat{g}(\omega)}{\hat{\Phi}} d\omega.$$

Remembering $H^s(\mathbb{R}^d) = \{ f \in L_2(\mathbb{R}^d) : \hat{f}(1 + \| \cdot \|_2^2)^{s/2} \in L_2(\mathbb{R}^d) \}$ for $s > d/2$ we have

**Conclusion:** If $\Phi \in L_1(\mathbb{R}^d) \cap C(\mathbb{R}^d)$ satisfies

$$c_1(1 + \| \omega \|_2^2)^{-s} \leq \hat{\Phi}(\omega) \leq c_2(1 + \| \omega \|_2^2)^{-s}$$

for $s > d/2$ then

$$\mathcal{N}_\Phi(\mathbb{R}^d) = H^s(\mathbb{R}^d), \quad \text{and the norms are equivalent}$$

Reproducing kernel Hilbert spaces generalize the Sobolev spaces.
Special case $\Phi(x - y)$

Let $\Phi \in C(\mathbb{R}^d) \cap L_1(\mathbb{R}^d)$ is real-valued and PD. We have

$$\mathcal{N}_\Phi(\mathbb{R}^d) = \left\{ f \in L_2(\mathbb{R}^d) \cap C(\mathbb{R}^d) : \frac{\hat{f}}{\sqrt{\hat{\Phi}}} \in L_2(\mathbb{R}^d) \right\}$$

$$(f, g)_{\mathcal{N}_\Phi(\mathbb{R}^d)} := (2\pi)^{-d/2} \left( \frac{\hat{f}}{\sqrt{\hat{\Phi}}} , \frac{\hat{g}}{\sqrt{\hat{\Phi}}} \right) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \frac{\hat{f}(\omega)\overline{\hat{g}(\omega)}}{\hat{\Phi}} d\omega.$$ 

Remembering $H^s(\mathbb{R}^d) = \left\{ f \in L_2(\mathbb{R}^d) : \hat{f}(1 + \| \cdot \|^2)^{s/2} \in L_2(\mathbb{R}^d) \right\}$ for $s > d/2$ we have

**Conclusion:** If $\Phi \in L_1(\mathbb{R}^d) \cap C(\mathbb{R}^d)$ satisfies

$$c_1 (1 + \| \omega \|^2)_s^{-s} \leq \hat{\Phi}(\omega) \leq c_2 (1 + \| \omega \|^2)_s^{-s}$$

for $s > d/2$ then

$$\mathcal{N}_\Phi(\mathbb{R}^d) = H^s(\mathbb{R}^d), \quad \text{and the norms are equivalent}$$

Reproducing kernel Hilbert spaces generalize the Sobolev spaces.
Special case $\Phi(x - y)$

Let $\Phi \in C(\mathbb{R}^d) \cap L_1(\mathbb{R}^d)$ is real-valued and PD. We have

$$\mathcal{N}_\Phi(\mathbb{R}^d) = \left\{ f \in L_2(\mathbb{R}^d) \cap C(\mathbb{R}^d) : \frac{\hat{f}}{\sqrt{\hat{\Phi}}} \in L_2(\mathbb{R}^d) \right\}$$

$$(f, g)_{\mathcal{N}_\Phi(\mathbb{R}^d)} := (2\pi)^{-d/2} \left( \frac{\hat{f}}{\sqrt{\hat{\Phi}}} , \frac{\hat{g}}{\sqrt{\hat{\Phi}}} \right) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \frac{\hat{f}(\omega)\hat{g}(\omega)}{\hat{\Phi}} d\omega.$$  

Remembering $H^s(\mathbb{R}^d) = \left\{ f \in L_2(\mathbb{R}^d) : \hat{f}(1 + \| \cdot \|^2)^{s/2} \in L_2(\mathbb{R}^d) \right\}$ for $s > d/2$ we have

**Conclusion:** If $\Phi \in L_1(\mathbb{R}^d) \cap C(\mathbb{R}^d)$ satisfies

$$c_1(1 + \|\omega\|^2_2)^{-s} \leq \hat{\Phi}(\omega) \leq c_2(1 + \|\omega\|^2_2)^{-s}$$

for $s > d/2$ then

$$\mathcal{N}_\Phi(\mathbb{R}^d) = H^s(\mathbb{R}^d), \quad \text{and the norms are equivalent}$$

Reproducing kernel Hilbert spaces generalize the Sobolev spaces.
Native space for CPD kernels

\[ H_\Phi(\Omega) := \text{span} \left\{ \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j) : N \in \mathbb{N}_0, \alpha \in \mathbb{R}^N, x_j \in \Omega \text{ with } \sum_{j=1}^{N} \alpha_j p(x_j) = 0 \text{ for all } p \in \mathbb{P}^{d}_{m-1} \right\} \]

\[ (f, g)_\Phi := \sum_{j=1}^{N} \sum_{k=1}^{M} \alpha_j \beta_k \Phi(x_j, y_k), \quad f = \sum \alpha_j \Phi(\cdot, x_j), \quad g = \sum \beta_k \Phi(\cdot, y_k). \]

The inner product is well-defined due to the side condition imposed on \( \alpha \) and \( x_j \) which implies that

\[ (f, f)_\Phi = 0 \Rightarrow \alpha = 0 \Rightarrow f = 0 \]

The native space is then technically defined as:

\[ \mathcal{N}_\Phi(\Omega) := \overline{H_\Phi(\Omega)} + \mathbb{P}^{d}_{m-1}. \]
Special cases: *native space = classical space*

- Matérn (Sobolev) kernel:

\[
\Phi_s(x) = \frac{2^{1-s}}{\Gamma(s)} \|x\|_2^{s-d/2} K_{d/2-s}(\|\omega\|_2), \quad s > d/2 \quad \Rightarrow \quad \hat{\Phi}_s(\omega) = (1 + \|\omega\|_2^2)^{-s}
\]

\[
\mathcal{N}_{\Phi_s}(\mathbb{R}^d) = H^s(\mathbb{R}^d)
\]

- Wendland’s functions:

\[
\Phi_{d,k} = \phi_{d,k}(\| \cdot \|_2) \quad \Rightarrow \quad \hat{\Phi}_{d,k}(\omega) \sim (1 + \|\omega\|_2^2)^{-(d/2+k+1/2)}
\]

\[
\mathcal{N}_{\Phi_{d,k}}(\mathbb{R}^d) = H^{d/2+k+1/2}(\mathbb{R}^d)
\]

\[
\phi_{3,1} = (1 - r)^4 + (4r + 1)
\]

Behaviour of \(F\phi_{3,1}\)
Special cases: \textit{native space = classical space}

- Matérn (Sobolev) kernel:

\[
\Phi_s(x) = \frac{2^{1-s}}{\Gamma(s)} \|x\|_2^{s-d/2} K_{d/2-s}(\|\omega\|_2), \quad s > d/2 \quad \Rightarrow \quad \hat{\Phi}_s(\omega) = (1 + \|\omega\|_2^2)^{-s}
\]

\[\mathcal{N}_{\Phi_s}(\mathbb{R}^d) = H^s(\mathbb{R}^d)\]

- Wendland’s functions:

\[
\Phi_{d,k} = \phi_{d,k}(\| \cdot \|_2) \quad \Rightarrow \quad \hat{\Phi}_{d,k}(\omega) \sim (1 + \|\omega\|_2^2)^{-(d/2+k+1/2)}
\]

\[\mathcal{N}_{\Phi_{d,k}}(\mathbb{R}^d) = H^{d/2+k+1/2}(\mathbb{R}^d)\]

\[
\phi_{3,1} = (1 - r)^4 (4r + 1)
\]

\[\text{Behaviour of } \mathcal{F}\phi_{3,1}\]

\[
d = 1 \\
d = 2 \\
d = 3
\]
Approximation by Kernels
Solving PDEs by Kernels
Approximation by MLS
GMLS Approximation and DMLPG Methods
Main References

Interpolation Problem
Positive Definite Kernels
Conditionally Positive Definite Functions
Compactly Supported Basis Functions
Native Spaces
Error Analysis and Stability

Special cases: *native space = classical space*

- Polyharmonic splines (TPS and powers): \( \Phi_{d,\ell} = \varphi_{d,\ell}(\|\cdot\|_2) \) where for \( \ell > d/2 \)

\[
\varphi_{d,\ell}(r) := \begin{cases} 
C_{d,\ell} r^{2\ell-d}, & \text{for } d \text{ odd,} \\
C'_{d,\ell} r^{2\ell-d} \log r, & \text{for } d \text{ even,}
\end{cases}
\]

is CPD of order \( m = \ell - \lfloor d/2 \rfloor + 1 \). The native space is the **Beppo-Levi space** \( \text{BL}_\ell(\mathbb{R}^d) \)
when considered as a CPD function of order \( \ell \)

\[
\mathcal{N}_{\Phi_{d,\ell}}(\mathbb{R}^d) = \text{BL}_\ell(\mathbb{R}^d) := \left\{ f \in C(\mathbb{R}^d) : D^\alpha f \in L_2(\mathbb{R}^d) \text{ for all } |\alpha| = \ell \right\}
\]

\[
= \left\{ f \in C(\mathbb{R}^d) : \hat{f}(\cdot) \| \cdot \|_2^\ell \in L_2(\mathbb{R}^d) \right\}
\]

- The null space of \( \text{BL}_\ell(\mathbb{R}^d) \) is \( \mathbb{P}_{\ell-1}^d \)
- The intersection of all \( \text{BL}_\ell(\mathbb{R}^d) \) for \( \ell \leq k \) is \( H^k(\mathbb{R}^d) \)
Special cases: *native space = classical space*

- Polyharmonic splines (TPS and powers): $\Phi_{d,\ell} = \varphi_{d,\ell}(\| \cdot \|_2)$ where for $\ell > d/2$

$$
\varphi_{d,\ell}(r) := \begin{cases} 
C_{d,\ell} r^{2\ell-d}, & \text{for } d \text{ odd,} \\
C'_{d,\ell} r^{2\ell-d} \log r, & \text{for } d \text{ even,}
\end{cases}
$$

is CPD of order $m = \ell - \lceil d/2 \rceil + 1$. The native space is the *Beppo-Levi space* $\text{BL}_\ell(\mathbb{R}^d)$ (when considered as a CPD function of order $\ell$)

$$
\mathcal{N}_{\Phi_{d,\ell}}(\mathbb{R}^d) = \text{BL}_\ell(\mathbb{R}^d) := \left\{ f \in C(\mathbb{R}^d) : D^\alpha f \in L_2(\mathbb{R}^d) \text{ for all } |\alpha| = \ell \right\}
$$

$$
= \left\{ f \in C(\mathbb{R}^d) : \hat{f}(\cdot) \| \cdot \|_2 \in L_2(\mathbb{R}^d) \right\}
$$

- The null space of $\text{BL}_\ell(\mathbb{R}^d)$ is $\mathbb{P}_{\ell-1}^d$
- The intersection of all $\text{BL}_\ell(\mathbb{R}^d)$ for $\ell \leq k$ is $H^k(\mathbb{R}^d)$
An Embedding Theorem

How about Gaussian and Multiquadrics??

- If $\Phi \subset C^{2k}(\Omega \times \Omega)$ is CPD then $\mathcal{N}_\Phi(\Omega) \subseteq C^k(\Omega)$.
- In a special case we can prove that if $s > k + d/2$ then $H^s(\mathbb{R}^d) \subset C^k(\mathbb{R}^d)$, the well-known Sobolev’s Embedding Theorem.

Extension and Restriction Theorems

- Let $\Omega_1 \subseteq \Omega_2 \subseteq \mathbb{R}^d$. Each function $f \in \mathcal{N}_\Phi(\Omega_1)$ has a natural extension to a function $Ef \in \mathcal{N}_\Phi(\Omega_1)$ and $|Ef|_{\mathcal{N}_\Phi(\Omega_2)} = |f|_{\mathcal{N}_\Phi(\Omega_1)}$.
- If $f \in \mathcal{N}_\Phi(\Omega_2)$ then $f|_{\Omega_1} \in \mathcal{N}_\Phi(\Omega_1)$.

**Conclusion:** If $\Phi \in L_1(\mathbb{R}^d)$ and $\hat{\Phi} \sim (1 + \|x\|^2)^{-k}$ for $k \in \mathbb{N}$ and $k > d/2$ and $\Omega \subset \mathbb{R}^d$ has a Lipschitz boundary then

$$\mathcal{N}_\Phi(\Omega) = H^k(\Omega)$$
Error Analysis

Some approaches

- **Using sampling inequalities**: Narcowich & Ward & Wendland (2005) and others
- **Superconvergence**: Schaback (1999, 2016)

Power function

Write the interpolant in the Lagrange form:

\[ s_{X,f}(x) = \sum_{k=1}^{N} c_k \Phi(x,x_k) = \sum_{k=1}^{N} u_k^*(x)f(x_k), \quad \text{where} \quad u_k^*(x_j) = \delta_{kj} \]

Simple calculations show

\[ Au^*(x) = b(x) \quad \text{where} \quad b(x) = [\Phi(x,x_1), \ldots, \Phi(x,x_N)]^T \]
Error Analysis

Some approaches


- **Using sampling inequalities**: Narcowich & Ward & Wendland (2005) and others


- **Superconvergence**: Schaback (1999, 2016)

Power function

Write the interpolant in the Lagrange form:

\[ s_{X,f}(x) = \sum_{k=1}^{N} c_k \Phi(x, x_k) = \sum_{k=1}^{N} u_k^*(x)f(x_k), \text{ where } u_k^*(x_j) = \delta_{kj} \]

Simple calculations show

\[ Au^*(x) = b(x) \text{ where } b(x) = [\Phi(x, x_1), \ldots, \Phi(x, x_N)]^T \]
Power function
We are looking for a bound for
\[ |f(x) - s_{X,f}(x)| \leq ?? \]

Let \( \Phi \) be the reproducing kernel for \( \mathcal{N} := \mathcal{N}_\Phi(\Omega) \). We have
\[
f(x) = \langle f, \Phi(\cdot, x) \rangle_{\mathcal{N}}
\]
\[
s_{f,X}(x) = \sum_{k=1}^{N} u_k^*(x) f(x_k) = \sum_{k=1}^{N} u_k^*(x) \langle f, \Phi(\cdot, x_k) \rangle_{\mathcal{N}} = \langle f, \sum_{k=1}^{N} u_k^*(x) \Phi(\cdot, x_k) \rangle_{\mathcal{N}}
\]

Thus we have
\[
|f(x) - s_{X,f}(x)| = \left| \langle f, \Phi(\cdot, x) - \sum_{k=1}^{N} u_k^*(x) \Phi(\cdot, x_k) \rangle_{\mathcal{N}} \right|
\]
\[
\leq \|f\|_{\mathcal{N}} \left\| \Phi(\cdot, x) - \sum_{k=1}^{N} u_k^*(x) \Phi(\cdot, x_k) \right\|_{\mathcal{N}}
\]
\[
:= P_{\Phi,X}(x)
\]
\[
= \underbrace{P_{\Phi,X}(x)}_{\text{distribution of data}} \times \underbrace{\|f\|_{\mathcal{N}}}_{\text{smoothness of data}}
\]
Power function
We are looking for a bound for

\[ |f(x) - s_{x,f}(x)| \leq ?? \]

Let \( \Phi \) be the reproducing kernel for \( \mathcal{N} := \mathcal{N}_\Phi(\Omega) \). We have

\[ f(x) = \langle f, \Phi(\cdot, x) \rangle_{\mathcal{N}} \]

\[ s_{f,X}(x) = \sum_{k=1}^{N} u_k^*(x)f(x_k) = \sum_{k=1}^{N} u_k^*(x)\langle f, \Phi(\cdot, x_k) \rangle_{\mathcal{N}} = \langle f, \sum_{k=1}^{N} u_k^*(x)\Phi(\cdot, x_k) \rangle_{\mathcal{N}} \]

Thus we have

\[ |f(x) - s_{x,f}(x)| = \left| \langle f, \Phi(\cdot, x) - \sum_{k=1}^{N} u_k^*(x)\Phi(\cdot, x_k) \rangle_{\mathcal{N}} \right| \]

\[ \leq \|f\|_{\mathcal{N}} \left\| \Phi(\cdot, x) - \sum_{k=1}^{N} u_k^*(x)\Phi(\cdot, x_k) \right\|_{\mathcal{N}} \]

\[ := P_{\Phi,X}(x) \]

\[ = P_{\Phi,X}(x) \times \|f\|_{\mathcal{N}} \]

distribution of data \hspace{1cm} smoothness of data
**Power function**

To know more about the power function define the following quadratic form:

\[ Q(u) := \Phi(x, x) - 2 \sum_{k=1}^{N} u_k \Phi(x, x_k) + \sum_{j=1}^{N} \sum_{k=1}^{N} u_j u_k \Phi(x_j, x_k) \]

\[ = \Phi(x, x) - 2u^T b(x) + u^T Au \]

Since \( A \) is positive definite, vector \( u^* \) from \( Au^* = b \) minimize \( Q(u) \).

\[ P_{\Phi, x}(x) = \left\| \Phi(\cdot, x) - \sum_{k=1}^{N} u_k^* (x) \Phi(\cdot, x_k) \right\|_N \]

\[ = (\cdots, \cdots) \]

\[ = \Phi(x, x) - 2u^* T b(x) + u^* T Au^* \]

\[ = Q(u^* (x)) \]
Power function

To know more about the power function define the following quadratic form:

\[ Q(u) := \Phi(x, x) - 2 \sum_{k=1}^{N} u_k \Phi(x, x_k) + \sum_{j=1}^{N} \sum_{k=1}^{N} u_j u_k \Phi(x_j, x_k) \]

\[ = \Phi(x, x) - 2u^T b(x) + u^T A u \]

Since \( A \) is positive definite, vector \( u^* \) from \( Au^* = b \) minimize \( Q(u) \).

\[ P_{\Phi, x}(x) = \left\| \Phi(\cdot, x) - \sum_{k=1}^{N} u_k^* (x) \Phi(\cdot, x_k) \right\|_N \]

\[ = \langle \cdots, \cdots \rangle \]

\[ = \Phi(x, x) - 2u^* T b(x) + u^* T A u^* \]

\[ = Q(u^*(x)) \]
Power function in terms of $h$
Assume $\Phi$ is shift invariant. Define the fill distance

$$h = h_{X,\Omega} := \sup_{x \in \Omega} \min_{x_j \in X} \|x - x_j\|_2$$

The power function can be bounded as

$$P_{\Phi,X}(x) \leq F(h), \quad x \in \Omega$$

We leave the technical details here and give the final results for radial functions:

<table>
<thead>
<tr>
<th>RBF</th>
<th>$\phi(r)$</th>
<th>$\sqrt{F(h)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>$\exp\left(- (\varepsilon r)^2 \right)$</td>
<td>$\exp(-c \log h</td>
</tr>
<tr>
<td>MQ</td>
<td>$(-1)^{\lceil \beta/2 \rceil} (1 + (\varepsilon r)^2)^{\beta}$, $\beta &gt; 0$, $\beta \notin \mathbb{N}$</td>
<td>$\exp(-c/h)$</td>
</tr>
<tr>
<td>IMQ</td>
<td>$(-1)^{\lceil \beta/2 \rceil} (1 + (\varepsilon r)^2)^{\beta}$, $\beta &lt; 0$</td>
<td>$\exp(-c/h)$</td>
</tr>
<tr>
<td>Powers</td>
<td>$(-1)^{\lceil \beta/2 \rceil} r^{\beta}$, $\beta &gt; 0$, $\beta \notin 2\mathbb{N}$</td>
<td>$h^{\beta/2}$</td>
</tr>
<tr>
<td>TPS</td>
<td>$(-1)^{k+1} r^{2k} \log r$, $k \in \mathbb{N}$</td>
<td>$h^k$</td>
</tr>
<tr>
<td>Wendland</td>
<td>$\phi_{d,k}$</td>
<td>$h^{k+1/2}$</td>
</tr>
</tbody>
</table>
Power function in terms of $h$
Assume $\Phi$ is shift invariant. Define the fill distance

$$h = h_{X,\Omega} := \sup_{x \in \Omega} \min_{x_j \in X} \|x - x_j\|_2$$

The power function can be bounded as

$$P_{\Phi,X}^2(x) \leq F(h), \quad x \in \Omega$$

We leave the technical details here and give the final results for radial functions:

<table>
<thead>
<tr>
<th>RBF</th>
<th>$\phi(r)$</th>
<th>$\sqrt{F(h)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>$\exp\left(-\varepsilon r^2\right)$</td>
<td>$\exp(-c</td>
</tr>
<tr>
<td>MQ</td>
<td>$(-1)\left[\beta\right]\left(1 + (\varepsilon r)^2\right)^\beta$, $\beta &gt; 0$, $\beta \notin \mathbb{N}$</td>
<td>$\exp(-c/h)$</td>
</tr>
<tr>
<td>IMQ</td>
<td>$\left(1 + (\varepsilon r)^2\right)^\beta$, $\beta &lt; 0$</td>
<td>$\exp(-c/h)$</td>
</tr>
<tr>
<td>Powers</td>
<td>$(-1)\left[\beta/2\right]r^\beta$, $\beta &gt; 0$, $\beta \notin 2\mathbb{N}$</td>
<td>$h^{\beta/2}$</td>
</tr>
<tr>
<td>TPS</td>
<td>$(-1)^{k+1}r^{2k}\log r$, $k \in \mathbb{N}$</td>
<td>$h^k$</td>
</tr>
<tr>
<td>Wendland</td>
<td>$\phi_{d,k}$</td>
<td>$h^{k+1/2}$</td>
</tr>
</tbody>
</table>
Sampling inequalities

**Theorem:** Narcowich & Ward & Wendland (2005), Wendland & Rieger (2005)

Let \( \sigma > 0 \) and \( \sigma = k + s \) where \( 0 \leq s < 1 \) and \( k \in \mathbb{N} \). Let \( m \in \mathbb{N}_0 \) satisfying \( k > m + d/2 \). If \( X \subset \Omega \) has a sufficiently small \( h = h_{X,\Omega} \) where \( \Omega \) is compact set with Lipschitz boundary then for \( u \in H^\sigma(\Omega) \) we have for \( q \in [1, \infty] \)

\[
|u|_{W^m_q(\Omega)} \leq C \left( h^{\sigma-m-d(1/2-1/q)} + |u|_{H^\sigma(\Omega)} + h^{-m} \|u|_X\|_\infty \right)
\]

If the interpolant satisfies additional stability bound \( |s_{f,X}|_{H^\sigma(\Omega)} \leq |f|_{H^\sigma(\Omega)} \) then inserting \( u = f - s_{f,X} \) we have

\[
|f - s_{f,X}|_{W^m_q(\Omega)} \leq Ch^{\sigma-m}|f|_{H^\sigma(\Omega)}
\]

This can be easily applied for RBF \( \Phi \) if \( H^\sigma \) be the native space of \( \Phi \), i.e.

\[
\hat{\Phi} \sim (1 + \| \cdot \|_2^2)^{-\sigma}
\]
Sampling inequalities

**Theorem:** Narcowich & Ward & Wendland (2005), Wendland & Rieger (2005)

Let $\sigma > 0$ and $\sigma = k + s$ where $0 \leq s < 1$ and $k \in \mathbb{N}$. Let $m \in \mathbb{N}_0$ satisfying $k > m + d/2$. If $X \subset \Omega$ has a sufficiently small $h = h_{X,\Omega}$ where $\Omega$ is compact set with Lipschitz boundary then for $u \in H^\sigma(\Omega)$ we have for $q \in [1, \infty]$

$$|u|_{W^m_q(\Omega)} \leq C \left( h^{\tau - m - d(1/2 - 1/q)} + |u|_{H^\sigma(\Omega)} + h^{-m} \|u|_X\|_\infty \right)$$

If the interpolant satisfies additional stability bound $|s_f,X|_{H^\sigma(\Omega)} \leq |f|_{H^\sigma(\Omega)}$ then inserting $u = f - s_f,X$ we have

$$|f - s_f,X|_{W^m_q(\Omega)} \leq C h^{\sigma - m} |f|_{H^\sigma(\Omega)}$$

This can be easily applied for RBF $\Phi$ if $H^\sigma$ be the native space of $\Phi$, i.e.

$$\hat{\Phi} \sim (1 + \| \cdot \|^2_2)^{-\sigma}$$
Sampling inequalities

**Theorem:** Narcowich & Ward & Wendland (2005), Wendland & Rieger (2005)

Let $\sigma > 0$ and $\sigma = k + s$ where $0 \leq s < 1$ and $k \in \mathbb{N}$. Let $m \in \mathbb{N}_0$ satisfying $k > m + d/2$. If $X \subset \Omega$ has a sufficiently small $h = h_{X,\Omega}$ where $\Omega$ is compact set with Lipschitz boundary then for $u \in H^\sigma(\Omega)$ we have for $q \in [1, \infty]$

$$|u|_{W^m_q(\Omega)} \leq C(h^{\tau - m - d(1/2 - 1/q)} + |u|_{H^\sigma(\Omega)} + h^{-m}||u|_X\infty)$$

If the interpolant satisfies additional stability bound $|s_{f,X}|_{H^\sigma(\Omega)} \leq |f|_{H^\sigma(\Omega)}$ then inserting $u = f - s_{f,X}$ we have

$$|f - s_{f,X}|_{W^m_q(\Omega)} \leq Ch^{\sigma - m}|f|_{H^\sigma(\Omega)}$$

This can be easily applied for RBF $\Phi$ if $H^\sigma$ be the native space of $\Phi$, i.e.

$$\hat{\Phi} \sim (1 + || \cdot ||_2^2)^{-\sigma}$$
Optimal recovery

Let $X$ be $\mathbb{P}_m^d$–unisolvent. Then for $f \in \mathcal{N}_\Phi(\Omega)$ we have the orthogonality

$$\langle f - s_f, x \rangle_{\mathcal{N}_\Phi} = 0, \quad \forall s \in \text{span}\{\phi(\cdot, x_j) : x_j \in X\} \cap H_\Phi(\Omega) + \mathbb{P}_m^d.$$

Conclusion:

$$|f - s_f, x|_{\mathcal{N}_\Phi}^2 + |s_f, x|_{\mathcal{N}_\Phi}^2 = |f|_{\mathcal{N}_\Phi}^2$$

$$|s_f, x|_{\mathcal{N}_\Phi} \leq |f|_{\mathcal{N}_\Phi}$$

$$|f - s_f, x|_{\mathcal{N}_\Phi} \leq |f|_{\mathcal{N}_\Phi}$$

Kernels give optimal approximation via interpolation

This can be generalized for functional approximation via kernels:

$$\lambda(f) \approx \lambda(s_f, x) = \sum_{j=1}^{N} a_j f(x_j)$$
Optimal recovery

Let $X$ be $\mathbb{P}_m^d$–unisolvent. Then for $f \in \mathcal{N}_\Phi(\Omega)$ we have the orthogonality

$$\langle f - s_f, x, s \rangle_{\mathcal{N}_\Phi} = 0, \quad \forall s \in \text{span}\{\phi(\cdot, x_j) : x_j \in X\} \cap H_\Phi(\Omega) + \mathbb{P}_m^d.$$ 

**Conclusion:**

$$|f - s_f, x|_{\mathcal{N}_\Phi}^2 + |s_f, x|_{\mathcal{N}_\Phi}^2 = |f|_{\mathcal{N}_\Phi}^2$$

$$|s_f, x|_{\mathcal{N}_\Phi} \leq |f|_{\mathcal{N}_\Phi}$$

$$|f - s_f, x|_{\mathcal{N}_\Phi} \leq |f|_{\mathcal{N}_\Phi}$$

Kernels give optimal approximation via interpolation

This can be generalized for functional approximation via kernels:

$$\lambda(f) \approx \lambda(s_f, x) = \sum_{j=1}^{N} a_j f(x_j)$$
Optimal recovery

Let $X$ be $\mathbb{P}^d_m$–unisolvent. Then for $f \in \mathcal{N}_\Phi(\Omega)$ we have the orthogonality

$$\langle f - s_f, x, s \rangle_{\mathcal{N}_\Phi} = 0, \quad \forall s \in \text{span}\{\phi(\cdot, x_j) : x_j \in X\} \cap H_\Phi(\Omega) + \mathbb{P}^d_m.$$ 

**Conclusion:**

$$|f - s_f, x|^2_{\mathcal{N}_\Phi} + |s_f, x|^2_{\mathcal{N}_\Phi} = |f|^2_{\mathcal{N}_\Phi}$$

$$|s_f, x|_{\mathcal{N}_\Phi} \leq |f|_{\mathcal{N}_\Phi}$$

$$|f - s_f, x|_{\mathcal{N}_\Phi} \leq |f|_{\mathcal{N}_\Phi}$$

Kernels give optimal approximation via interpolation

This can be generalized for functional approximation via kernels:

$$\lambda(f) \approx \lambda(s_f, x) = \sum_{j=1}^N a_j f(x_j)$$
Stationary and Nonstationary Interpolation

We are interested in **sparse interpolation matrices** when the density of data increases. To achieve this we should **scale** the RBF proportional to data density.

\[ r \rightarrow \varepsilon r, \quad \varepsilon > 0 \]

\[ (1 + (\varepsilon r)^2)^{-0.5} \quad (1 - \varepsilon r)^4 + (4\varepsilon r + 1) \]

- **Stationary:** \( \varepsilon = \varepsilon(h) = \frac{\varepsilon_0}{h} \) where \( \varepsilon_0 \) is a constant.
- **Nonstationary:** \( \varepsilon = \varepsilon_0 \) independent of \( h \).

All already established error bounds were obtained for **nonstationary** case.
Stationary and Nonstationary Interpolation

We are interested in **sparse interpolation matrices** when the density of data increases. To achieve this we should **scale** the RBF proportional to data density.

\[ r \rightarrow \varepsilon r, \quad \varepsilon > 0 \]

\[ (1 + (\varepsilon r)^2)^{-0.5} \quad \text{for Stationary case} \]

\[ (1 - \varepsilon r)^4 (4\varepsilon r + 1) \quad \text{for Nonstationary case} \]

▸ Stationary: \( \varepsilon = \varepsilon(h) = \frac{\varepsilon_0}{h} \) where \( \varepsilon_0 \) is a constant.

▸ Nonstationary: \( \varepsilon = \varepsilon_0 \) independent of \( h \).

All already established error bounds were obtained for **nonstationary** case.
In stationary case:

- Gaussian does not provide positive approximation order; the error is saturated; Buhmann (1989)
  Idea of approximate approximation by Maz’ya & Schmidt (1996) says that the level of saturation can be controlled by $\varepsilon_0$.

- Wendland’s functions do not provide positive approximation order; the error is saturated; Schaback (1997).
  Idea of multiscale interpolation by Schaback solves the problem.

- Stationary interpolation with MQ, IMQ, TPS and radial powers presents no difficulties; Schaback (1995) proved the corresponding native spaces are invariant under scaling.
In stationary case:

- Gaussian does not provide positive approximation order; the error is saturated; Buhmann (1989)
  Idea of approximate approximation by Maz’ya & Schmidt (1996) says that the level of saturation can be controlled by $\varepsilon_0$.

- Wendland’s functions do not provide positive approximation order; the error is saturated; Schaback (1997).
  Idea of multiscale interpolation by Schaback solves the problem.

- Stationary interpolation with MQ, IMQ, TPS and radial powers presents no difficulties; Schaback (1995) proved the corresponding native spaces are invariant under scaling.
In stationary case:

- Gaussian does not provide positive approximation order; the error is saturated; Buhmann (1989)
  Idea of approximate approximation by Maz’ya & Schmidt (1996) says that the level of saturation can be controlled by $\varepsilon_0$.

- Wendland’s functions do not provide positive approximation order; the error is saturated; Schaback (1997).
  Idea of multiscale interpolation by Schaback solves the problem.

- Stationary interpolation with MQ, IMQ, TPS and radial powers presents no difficulties; Schaback (1995) proved the corresponding native spaces are invariant under scaling.
Interpolation by CSRBFS

Let $X = \{x_1, \ldots, x_N\} \subset \Omega \subset \mathbb{R}^d$ and $f_1, \ldots, f_N \in \mathbb{R}$ be given. Define the scaled CSRBFS as

$$
\phi_\delta(r) := \phi \left( \frac{r}{\delta} \right), \quad \delta > 0, \quad \delta = \frac{1}{\epsilon}
$$

where $\phi$ is supported in $[0, 1]$. Then write

$$
s_{X,f}(x) = s(x) = \sum_{k=1}^{N} c_k \phi_\delta(\|x - x_k\|_2) = \sum_{k=1}^{N} c_k \phi \left( \frac{\|x - x_k\|_2}{\delta} \right)
$$

and apply the interpolation conditions $s(x_k) = f_k$.

- **Stationary Interpolation:** $\delta = C \epsilon$
- **Non-stationary interpolation:** $\delta = C$
Test example

Interpolating the Franke’s function on $[-1, 1]^2$ by $\phi_{3,1}(r) = (1 - r)^4(4r + 1)$.

Figure: Errors and condition numbers: stationary and nonstationary algorithms
Sparsity
Interpolating the Franke’s function on $[-1, 1]^2$ by $\phi_{3,1}(r) = (1 - r)^4(4r + 1)$.

Figure: Sparsity patterns: nonstationary (left), and stationary (right) algorithms
Multiscale algorithm for interpolation

Suppose \( X_1, X_2, \ldots, X_n \) is a sequence of point sets on \( \Omega \) with mesh norms \( h_1 > h_2 > \ldots > h_n \).

Let \( \delta_k = Ch_k \) and \( I_k = I_{\delta_k, x_k} \) be the interpolant operator with \( \phi_{\delta_k} \) on \( X_k \).

\[
W_k := \text{span}\{\phi_{\delta_k}(\|x - x_j\|) : x_j \in X_k\}, \quad V_n = W_1 + W_2 + \cdots + W_n
\]

1. input: function \( f \), number of levels \( n \)
2. set \( f_0 = 0 \) and \( e_0 = f \)
3. for \( k = 1, 2, \ldots, n \) do
4. \( s_k = I_k e_{k-1} \)
5. \( f_k = f_{k-1} + s_k \)
6. \( e_k = e_{k-1} - s_k \)
7. end
8. output: \( f_n \in V_n \)

\[
f_k + e_k = f_{k-1} + e_{k-1} = \cdots = f_0 + e_0 = f \Rightarrow e_k = f - f_k
\]
\[
e_k|_{x_k} = e_{k-1}|_{x_k} - s_k|_{x_k} = 0 \Rightarrow f_k|_{x_k} = f|_{x_k}
\]
\[
f_n = f_{n-1} + s_n = f_{n-2} + s_{n-1} + s_n = \cdots = s_0 + s_1 + \cdots + s_n
\]
Multiscale algorithm for interpolation

Suppose $X_1, X_2, \ldots, X_n$ is a sequence of point sets on $\Omega$ with mesh norms $h_1 > h_2 > \ldots > h_n$. Let $\delta_k = Ch_k$ and $I_k = I_{\delta_k, x_k}$ be the interpolant operator with $\phi_{\delta_k}$ on $X_k$.

$W_k := \text{span}\left\{ \phi_{\delta_k}(\|x - x_j\|) : x_j \in X_k \right\}$, \hspace{1cm} $V_n = W_1 + W_2 + \cdots + W_n$

1. input: function $f$, number of levels $n$
2. set $f_0 = 0$ and $e_0 = f$
3. for $k = 1, 2, \ldots, n$ do
   4. \hspace{1cm} $s_k = I_k e_{k-1}$
   5. \hspace{1cm} $f_k = f_{k-1} + s_k$
   6. \hspace{1cm} $e_k = e_{k-1} - s_k$
4. end
8. output: $f_n \in V_n$

\[
\begin{align*}
   f_k + e_k &= f_{k-1} + e_{k-1} = \cdots = f_0 + e_0 = f \Rightarrow e_k = f - f_k \\
   e_k|_{x_k} &= e_{k-1}|_{x_k} - s_k|_{x_k} = 0 \Rightarrow f_k|_{x_k} = f|_{x_k} \\
   f_n &= f_{n-1} + s_n = f_{n-2} + s_{n-1} + s_n = \cdots = s_0 + s_1 + \cdots + s_n
\end{align*}
\]
Multiscale algorithm for interpolation

Suppose $X_1, X_2, \ldots, X_n$ is a sequence of point sets on $\Omega$ with mesh norms $h_1 > h_2 > \ldots > h_n$.
Let $\delta_k = C h_k$ and $I_k = I_{\delta_k} x_k$ be the interpolant operator with $\phi_{\delta_k}$ on $X_k$.
$W_k := \text{span}\{ \phi_{\delta_k}(\|x - x_j\|): x_j \in X_k \}$, \hspace{1em} $V_n = W_1 + W_2 + \cdots + W_n$

1. input: function $f$, number of levels $n$
2. set $f_0 = 0$ and $e_0 = f$
3. for $k = 1, 2, \ldots, n$ do
4. \hspace{1em} $s_k = I_k e_{k-1}$
5. \hspace{1em} $f_k = f_{k-1} + s_k$
6. \hspace{1em} $e_k = e_{k-1} - s_k$
7. end
8. output: $f_n \in V_n$

\[f_k + e_k = f_{k-1} + e_{k-1} = \cdots = f_0 + e_0 = f \Rightarrow e_k = f - f_k\]
\[e_k|_{x_k} = e_{k-1}|_{x_k} - s_k|_{x_k} = 0 \Rightarrow f_k|_{x_k} = f|_{x_k}\]
\[f_n = f_{n-1} + s_n = f_{n-2} + s_{n-1} + s_n = \cdots = s_0 + s_1 + \cdots + s_n\]
Multiscale algorithm for interpolation

Suppose $X_1, X_2, \ldots, X_n$ is a sequence of point sets on $\Omega$ with mesh norms $h_1 > h_2 > \ldots > h_n$. Let $\delta_k = Ch_k$ and $I_k = I_{\delta_k, x_k}$ be the interpolant operator with $\phi_{\delta_k}$ on $X_k$.

$W_k := \text{span}\{\phi_{\delta_k}(\|x - x_j\|) : x_j \in X_k\}$, $V_n = W_1 + W_2 + \cdots + W_n$

1. input: function $f$, number of levels $n$
2. set $f_0 = 0$ and $e_0 = f$
3. for $k = 1, 2, \ldots, n$ do
4. $s_k = I_k e_{k-1}$
5. $f_k = f_{k-1} + s_k$
6. $e_k = e_{k-1} - s_k$
7. end
8. output: $f_n \in V_n$

\[ f_k + e_k = f_{k-1} + e_{k-1} = \cdots = f_0 + e_0 = f \Rightarrow e_k = f - f_k \]
\[ e_k|_{X_k} = e_{k-1}|_{X_k} - s_k|_{X_k} = 0 \Rightarrow f_k|_{X_k} = f|_{X_k} \]
\[ f_n = f_{n-1} + s_n = f_{n-2} + s_{n-1} + s_n = \cdots = s_0 + s_1 + \cdots + s_n \]
History of multiscale method

- Schaback 1995, (a side discussion)
- Floater and Iske 1996 (the first separated paper in this area)
- Narcowich, Schaback, Ward 1999 (A new setting)
- Fasshauer 1999, (more consideration with application to PDEs)
- Wendland 2010, (the first error analysis on $\mathbb{R}^d$)
- 2010-present: Interpolation and PDE solution on $\mathbb{R}^d$ and other manifolds such as $\mathbb{S}^d$.

Error analysis for pure function approximation

**Theorem:** (Wendland 2010)

Assumptions:

- $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz boundary,
- $c\mu h_k < h_{k+1} < \mu h_k$ for $\mu \in (0, 1)$ and $c \in (0, 1]$,
- $\delta_k = \nu h_k$,
- $f \in H^\tau(\Omega)$, $\tau > d/2$.

Then

$$
\|f - f_n\|_{L_2(\Omega)} \leq C h_n^{\tau - \sigma} \|f\|_{H^\tau(\Omega)}, \quad \sigma = -\log(C_1)/\log(\mu) > 0.
$$
History of multiscale method

- Schaback 1995, (a side discussion)
- Floater and Iske 1996 (the first separated paper in this area)
- Narcowich, Schaback, Ward 1999 (A new setting)
- Fasshauer 1999, (more consideration with application to PDEs)
- Wendland 2010, (the first error analysis on $\mathbb{R}^d$)
- 2010-present: Interpolation and PDE solution on $\mathbb{R}^d$ and other manifolds such as $\mathbb{S}^d$.

Error analysis for pure function approximation

**Theorem:** (Wendland 2010)

**Assumptions:**

- $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz boundary,
- $c\mu h_k < h_{k+1} < \mu h_k$ for $\mu \in (0, 1)$ and $c \in (0, 1]$,
- $\delta_k = \nu h_k$,
- $f \in H^{\tau}(\Omega)$, $\tau > d/2$.

Then

$$\|f - f_n\|_{L^2(\Omega)} \leq C h_n^{\tau - \sigma} \|f\|_{H^{\tau}(\Omega)}, \quad \sigma = -\log(C_1)/\log(\mu) > 0.$$
Coming back to previous test example

Interpolating the Franke’s function on $[-1, 1]^2$ by $\phi_{3,1}(r) = (1 - r)^4(4r + 1)$:

Figure: Errors and condition numbers: standard stationary, nonstationary and multiscale algorithms
Sparsity

Interpolating the Franke’s function on $[-1, 1]^2$ by $\phi_{3,1}(r) = (1 - r)^4 (4r + 1)$:

Figure: Sparsity patterns: nonstationary (left), and multiscale (right) algorithms
Stability

Interpolation matrix

\[
\begin{bmatrix}
A & P \\
P^T & 0
\end{bmatrix}
\begin{bmatrix}
c \\
b
\end{bmatrix} =
\begin{bmatrix}
f|_X \\
0
\end{bmatrix},
\quad A = (\Phi(x_k - x_j)) \in \mathbb{R}^{N \times N},
\quad P = (p_k(x_j)) \in \mathbb{R}^{N \times Q}
\]

\[
\lambda_{\text{min}}(A) = \inf_{\alpha \in \mathbb{R}^N \setminus \{0\}, P^T \alpha = 0} \frac{\alpha^T A \alpha}{\alpha^T \alpha}.
\]

\(\lambda_{\text{min}}(A) > 0\). If \(\Phi\) is PD then \(\lambda_{\text{min}}(A) = \text{smallest eigenvalue of } A\).

However in the case of a CPD functions this number is crucial because from \(\alpha^T A \alpha = \alpha^T f|_X\) we have

\[
\|\alpha\|_2 \leq \frac{1}{\lambda_{\text{min}}} \|f|_X\|_2.
\]

If \(A\) is PD then

\[
\text{cond}(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\lambda_{\text{max}}(A)}{\lambda_{\text{min}}(A)}
\]
Stability

Interpolation matrix

\[
\begin{bmatrix}
A & P \\
P^T & 0
\end{bmatrix}
\begin{bmatrix}
c \\
b
\end{bmatrix} =
\begin{bmatrix}
f^T |X \\
0
\end{bmatrix},
A = (\Phi(x_k - x_j)) \in \mathbb{R}^{N \times N},
P = (p_k(x_j)) \in \mathbb{R}^{N \times Q}
\]

\[
\lambda_{\text{min}}(A) = \inf_{\alpha \in \mathbb{R}^N \setminus \{0\}} \frac{\alpha^T A \alpha}{\alpha^T \alpha}.
\]

\(\lambda_{\text{min}}(A) > 0\). If \(\Phi\) is PD then \(\lambda_{\text{min}}(A) = \text{smallest eigenvalue of } A\).

However in the case of a CPD functions this number is crucial because from \(\alpha^T A \alpha = \alpha^T f|_X\) we have

\[
\|\alpha\|_2 \leq \frac{1}{\lambda_{\text{min}}} \|f|_X\|_2.
\]

If \(A\) is PD then

\[
\text{cond}(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\lambda_{\text{max}}(A)}{\lambda_{\text{min}}(A)}
\]
Stability

Interpolation matrix

\[
\begin{bmatrix}
A & P \\
PT & 0
\end{bmatrix}
\begin{bmatrix}
c \\
b
\end{bmatrix} =
\begin{bmatrix}
f|_X \\
0
\end{bmatrix}, \quad A = (\Phi(x_k - x_j)) \in \mathbb{R}^{N \times N}, \quad P = (p_k(x_j)) \in \mathbb{R}^{N \times Q}
\]

\[
\lambda_{\min}(A) = \inf_{\alpha \in \mathbb{R}^N \setminus \{0\}} \frac{\alpha^T A \alpha}{\alpha^T \alpha}.
\]

\[
\lambda_{\min}(A) > 0. \text{ If } \Phi \text{ is PD then } \lambda_{\min}(A) = \text{smallest eigenvalue of } A.
\]

However in the case of a CPD functions this number is crucial because from \(\alpha^T A \alpha = \alpha^T f|_X\) we have

\[
\|\alpha\|_2 \leq \frac{1}{\lambda_{\min}} \|f|_X\|_2.
\]

If \(A\) is PD then

\[
\text{cond}(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}
\]
Stability

Interpolation matrix

\[
\begin{bmatrix}
A & P \\
P^T & 0
\end{bmatrix}
\begin{bmatrix}
c \\
b
\end{bmatrix} =
\begin{bmatrix}
f|_X \\
0
\end{bmatrix}, \quad A = (\Phi(x_k - x_j)) \in \mathbb{R}^{N \times N}, \quad P = (p_k(x_j)) \in \mathbb{R}^{N \times Q}
\]

\[
\lambda_{\text{min}}(A) = \inf_{\alpha \in \mathbb{R}^N \setminus \{0\}} \frac{\alpha^T A \alpha}{\alpha^T \alpha}.
\]

\(\lambda_{\text{min}}(A) > 0\). If \(\Phi\) is PD then \(\lambda_{\text{min}}(A) = \text{smallest eigenvalue of } A\).

However in the case of a CPD functions this number is crucial because from \(\alpha^T A \alpha = \alpha^T f|_X\) we have

\[
\|\alpha\|_2 \leq \frac{1}{\lambda_{\text{min}}} \|f|_X\|_2.
\]

If \(A\) is PD then

\[
\text{cond}(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\lambda_{\text{max}}(A)}{\lambda_{\text{min}}(A)}
\]
Stability

By Gershgorin’s Theorem exists an index $k$ such that

$$|\lambda_{\text{max}}(A) - \Phi(x_k - x_k)| \leq \sum_{j=1}^{N} |\Phi(x_j - x_k)|$$

$$\Rightarrow \lambda_{\text{max}}(A) \leq N\|\Phi(\cdot - \cdot)\|_{\infty} \leq N\Phi(0)$$

This leads to

$$\lambda_{\text{max}}(A) \approx h_{X,\Omega}^{-d}$$

Thus $\lambda_{\text{max}}(A)$ causes no serious problem! Let see the behaviour of $\lambda_{\text{min}}(A)$. 
By Gershgorin’s Theorem exists an index $k$ such that

$$|\lambda_{\text{max}}(A) - \Phi(x_k - x_k)| \leq \sum_{\substack{j=1 \atop j \neq k}} |\Phi(x_j - x_k)|$$

$$\Rightarrow \lambda_{\text{max}}(A) \leq N\|\Phi(\cdot - \cdot)\|_{\infty} \leq N\Phi(0)$$

This leads to

$$\lambda_{\text{max}}(A) \approx h_{X,\Omega}^{-d}$$

Thus $\lambda_{\text{max}}(A)$ causes no serious problem! Let see the behaviour of $\lambda_{\text{min}}(A)$. 

Davoud Mirzaei

Meshless Approximants
Lower bound for $\lambda_{\text{min}}$

- Ball & Sivakumar & Ward (1992),

The stability bounds should be in terms of separation distance $q_X$ and not fill distance $h_X, \Omega$.

$$q_X := \frac{1}{2} \min_{j \neq k} \|x_j - x_k\|_2$$

We look for a function $G$ such that

$$\lambda_{\text{min}}(A) \geq G(q_X)$$
Lower bound for $\lambda_{\text{min}}$

- Ball & Sivakumar & Ward (1992),

The stability bounds should be in terms of separation distance $q_X$ and not fill distance $h_{X,\Omega}$.

$$q_X := \frac{1}{2} \min_{j \neq k} \| x_j - x_k \|_2$$

We look for a function $G$ such that

$$\lambda_{\text{min}}(A) \geq G(q_X)$$
Lower bound for $\lambda_{\text{min}}$

$$\lambda_{\text{min}}(A) \geq G(q_x), \quad P_{\Phi,X}^2(x) \leq F(h_X,\Omega)$$

<table>
<thead>
<tr>
<th>RBF</th>
<th>$G(q)$</th>
<th>$F(h)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>$q^{-d} \exp\left(-C/(q\varepsilon)^2\right)$</td>
<td>$\exp(-c</td>
</tr>
<tr>
<td>MQ</td>
<td>$q^{\beta-(d-1)/2} \exp\left(-C/(q\varepsilon)\right)$</td>
<td>$\exp(-c/h)$</td>
</tr>
<tr>
<td>IMQ</td>
<td>$q^{\beta-(d-1)/2} \exp\left(-C/(q\varepsilon)\right)$</td>
<td>$\exp(-c/h)$</td>
</tr>
<tr>
<td>Powers</td>
<td>$q^\beta$</td>
<td>$h^\beta$</td>
</tr>
<tr>
<td>TPS</td>
<td>$q^{2k}$</td>
<td>$h^{2k}$</td>
</tr>
<tr>
<td>Wendland</td>
<td>$q^{2k+1}$</td>
<td>$h^{2k+1}$</td>
</tr>
</tbody>
</table>

Trade-off (uncertainty) principle for standard RBF interpolation; conflict between accuracy and numerical stability; Schaback (1995)

Accuracy and smoothing vs. Stability and smoothing
Lower bound for $\lambda_{\text{min}}$

$$\lambda_{\text{min}}(A) \geq G(qx), \quad P_{\Phi, X}^2(x) \leq F(hx, \Omega)$$

<table>
<thead>
<tr>
<th>RBF</th>
<th>$G(q)$</th>
<th>$F(h)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>$q^{-d} \exp(-C/(q\varepsilon)^2)$</td>
<td>$\exp(-c</td>
</tr>
<tr>
<td>MQ</td>
<td>$q^{\beta -(d-1)/2} \exp(-C/(q\varepsilon))$</td>
<td>$\exp(-c/h)$</td>
</tr>
<tr>
<td>IMQ</td>
<td>$q^{\beta -(d-1)/2} \exp(-C/(q\varepsilon))$</td>
<td>$\exp(-c/h)$</td>
</tr>
<tr>
<td>Powers</td>
<td>$q^{\beta}$</td>
<td>$h^{\beta}$</td>
</tr>
<tr>
<td>TPS</td>
<td>$q^{2k}$</td>
<td>$h^{2k}$</td>
</tr>
<tr>
<td>Wendland</td>
<td>$q^{2k+1}$</td>
<td>$h^{2k+1}$</td>
</tr>
</tbody>
</table>

Trade-off (uncertainty) principle for standard RBF interpolation; conflict between accuracy and numerical stability; Schaback (1995)

Accuracy and smoothing vs. Stability and smoothing
Shape parameter $\varepsilon$

Usually for Gaussian and (I)MQ; infinitely smooth functions

**Theorem:** Convergence with respect to $\varepsilon$; Madych (1991):

Interpolation by Gaussian and (I)MQ for a fixed data set $X$ satisfies

$$|s_{f,X}(x) - f(x)| \leq C\gamma^{1/(\varepsilon h_{X,\Omega})}, \quad 0 < \gamma < 1$$

Trade-off principle with respect to the $\varepsilon$


- Trade-off principle for CSRBF; Schaback (1997), as we already saw in our later example!
Shape parameter $\varepsilon$

Usually for Gaussian and (I)MQ; infinitely smooth functions

**Theorem:** Convergence with respect to $\varepsilon$; Madych (1991):
Interpolation by Gaussian and (I)MQ for a fixed data set $X$ satisfies

$$|s_{f,X}(x) - f(x)| \leq C \gamma^{1/(\varepsilon h_X, \Omega)}, \quad 0 < \gamma < 1$$

Trade-off principle with respect to the $\varepsilon$


- Trade-off principle for CSRBF; Schaback (1997), as we already saw in our later example!
Shape parameter $\varepsilon$

Usually for Gaussian and (I)MQ; infinitely smooth functions

**Theorem**: Convergence with respect to $\varepsilon$; Madych (1991):
Interpolation by Gaussian and (I)MQ for a fixed data set $X$ satisfies

$$|s_{f,X}(x) - f(x)| \leq C \gamma^{1/(h_X, \Omega)}, \quad 0 < \gamma < 1$$

Trade-off principle with respect to the $\varepsilon$


- Trade-off principle for CSRBF; Schaback (1997), as we already saw in our later example!
How to avoid instability?

Standard interpolation works with space $H_\Phi$ and basis

$$\{ \Phi(\cdot - x_1), \ldots, \Phi(\cdot - x_N) \}$$

Observations:

- The error of interpolant gets smaller if $\Phi$ gets smoother,
- The stability gets worse if $\Phi$ gets smoother.

**Problem:** construct kernels with small error and small condition number!

**Result:** This is impossible; trade-off principle Schaback (1995)

**But:** The interpolation operator is uniformly bounded by the data; de Marchi & Schaback (2010)

**Consequence:** Stable basis exist, do not use $\Phi(\cdot, x_j)$

How to avoid instability?

Standard interpolation works with space $H_\Phi$ and basis

\[ \{ \Phi(\cdot - x_1), \ldots, \Phi(\cdot - x_N) \} \]

Observations:

- The error of interpolant gets smaller if $\Phi$ gets smoother,
- The stability gets worse if $\Phi$ gets smoother.

**Problem:** construct kernels with small error and small condition number!

**Result:** This is impossible; trade-off principle Schaback (1995)

**But:** The interpolation operator is uniformly bounded by the data; de Marchi & Schaback (2010)

**Consequence:** Stable basis exist, do not use $\Phi(\cdot, x_j)$

How to avoid instability?

Standard interpolation works with space $H_\Phi$ and basis

$$\{ \Phi(\cdot - x_1), \ldots, \Phi(\cdot - x_N) \}$$

Observations:

- The error of interpolant gets smaller if $\Phi$ gets smoother,
- The stability gets worse if $\Phi$ gets smoother.

Problem: construct kernels with small error and small condition number!

Result: This is impossible; trade-off principle Schaback (1995)

But: The interpolation operator is uniformly bounded by the data; de Marchi & Schaback (2010)

Consequence: Stable basis exist, do not use $\Phi(\cdot, x_j)$

How to avoid instability?

Standard interpolation works with space $H_\Phi$ and basis

$$\{ \Phi(\cdot - x_1), \ldots, \Phi(\cdot - x_N) \}$$

Observations:

- The error of interpolant gets smaller if $\Phi$ gets smoother,
- The stability gets worse if $\Phi$ gets smoother.

**Problem:** construct kernels with small error and small condition number!

**Result:** This is impossible; trade-off principle Schaback (1995)

**But:** The interpolation operator is uniformly bounded by the data; de Marchi & Schaback (2010)

**Consequence:** Stable basis exist, do not use $\Phi(\cdot, x_j)$

How to avoid instability?

Standard interpolation works with space $H_\Phi$ and basis

$$\{ \Phi(\cdot - x_1), \ldots, \Phi(\cdot - x_N) \}$$

Observations:

- The error of interpolant gets smaller if $\Phi$ gets smoother,
- The stability gets worse if $\Phi$ gets smoother.

**Problem**: construct kernels with small error and small condition number!

**Result**: This is impossible; trade-off principle Schaback (1995)

**But**: The interpolation operator is uniformly bounded by the data; de Marchi & Schaback (2010)

**Consequence**: Stable basis exist, do not use $\Phi(\cdot, x_j)$

How to avoid instability?

Standard interpolation works with space \( H_{\Phi} \) and basis

\[
\{ \Phi(\cdot - x_1), \ldots, \Phi(\cdot - x_N) \}
\]

Observations:

- The error of interpolant gets smaller if \( \Phi \) gets smoother,
- The stability gets worse if \( \Phi \) gets smoother.

**Problem:** construct kernels with small error and small condition number!

**Result:** This is impossible; trade-off principle Schaback (1995)

**But:** The interpolation operator is uniformly bounded by the data; de Marchi & Schaback (2010)

**Consequence:** Stable basis exist, do not use \( \Phi(\cdot, x_j) \)

How to avoid instability?

- If $\Phi$ is PD then $A$ is PD. Use **preconditioning** methods from linear algebra. Examples: preconditioned CG, incomplete Cholesky factorization and etc.
- **Greedy algorithms:** avoid using more points from the data site than absolutely necessary by greedily selection; Schaback & Wendland (2002), S. de Marchi et. al. (2005), Schaback (2014), D.M. (2015), etc
- **Contour-Padé method:** Fornberg & Wright (2004), limitation: small data site.
- **RBF-QR method:** Fornberg & Piret & Larsson & Flyer (2007, 2011), Application for computing differentiation matrices, stencil weights in Larsson et. al. (2013)
- ...
Summary, so far

- Kernel approximations are easily applicable for scattered data on non-trivial domains in higher dimensions,

- Numerical analysis needs functions spaces with continuous point evaluations, for example Hilbert spaces, e.g. Sobolev spaces; these have reproducing kernels,

- Kernels corresponds to Hilbert space of functions,

- The optimal approximation can be calculated via interpolation: kernels yield optimal methods for recovery of functions from data,

- Stationary interpolation does not provide positive order of approximation in some cases, but can be resolved by some tricks,

- The error of approximation gets smaller if the kernel gets smoother; even spectral accuracy is achievable,

- In parallel, the stability gets worse (even exponentially) if the kernel gets smoother,

- Instability can be overcome by certain tricks.
Unsymmetric collocation
Consider these operator equations

\[ Lu = f, \text{ in } \Omega \]
\[ Bu = g, \text{ on } \partial \Omega =: \Gamma \]

Let \( X = \{x_1, \ldots, x_N\} \subset \Omega \) be a trial set point. Approximate \( u \) by

\[ s = \sum_{j=1}^{N} c_j \Phi(\cdot, x_j) \approx u \]

Let \( Y = \{y_1, y_2, \ldots, y_M\} \) be a test set point.
Unsymmetric collocation

Consider these operator equations

\[ Lu = f, \quad \text{in } \Omega \]
\[ Bu = g, \quad \text{on } \partial \Omega =: \Gamma \]

Let \( X = \{x_1, \ldots, x_N\} \subset \Omega \) be a trial set point. Approximate \( u \) by

\[ s = \sum_{j=1}^{N} c_j \Phi(\cdot, x_j) \approx u \]

Let \( Y = \{y_1, y_2, \ldots, y_M\} \) be a test set point.
Unsymmetric collocation

Insert $s$ instead of $u$ in equations and test (collocate) them at $Y$ where $Y = Y_I \cup Y_B$ of internal and boundary points.

\[
\sum_{j=1}^{N} c_j L^y \Phi(y_k, x_j) = f(x_k), \quad y_k \in Y_I
\]

\[
\sum_{j=1}^{N} c_j B^y \Phi(y_\ell, x_j) = g(y_\ell), \quad y_\ell \in Y_B
\]

\[
\begin{bmatrix}
A_L \\
A_B
\end{bmatrix} c =
\begin{bmatrix}
f|_{Y_I} \\
g|_{Y_B}
\end{bmatrix}
\]

Some properties:

- Much simple to implement,
- Final system is of size $M \times N$ and it is unsymmetric,
- It is possibly unsolvable in some rare situations; Hon & Schaback (1998).
Unsymmetric collocation

Insert $s$ instead of $u$ in equations and test (collocate) them at $Y$ where $Y = Y_I \cup Y_B$ of internal and boundary points.

\[
\sum_{j=1}^{N} c_j L^y \Phi(y_k, x_j) = f(x_k), \quad y_k \in Y_I
\]

\[
\sum_{j=1}^{N} c_j B^y \Phi(y_\ell, x_j) = g(y_\ell), \quad y_\ell \in Y_B
\]

\[
\begin{bmatrix}
A_L \\
A_B
\end{bmatrix}
\begin{bmatrix}
c
\end{bmatrix}
= 
\begin{bmatrix}
f \mid Y_I \\
g \mid Y_B
\end{bmatrix}
\]

Some properties:

- Much simple to implement,
- Final system is of size $M \times N$ and it is unsymmetric,
- It is possibly unsolvable in some rare situations; Hon & Schaback (1998).
History

- Kansa (1990), application of MQ to PDEs. Sometimes refereed as Kansa method.
- Huge number of application to different PDE problems, 1990-present,

Some points from Schaback’s analysis:

- To overcome the insolvability, modify the setting,
- To get the error bound, there should at least be a unique solution to the modified discretization that converges to the true solution when the discretization is refined,
- Find least square solution instead,
- If $M \gg N$ (test space sufficiently larger than trial space; overtesting) then the system has full rank $N$ and stability condition is satisfied,
- Optimal convergence rate is obtained,
- The theory holds for a general cases including collocation, weak and local weak testing,
- Open problem: Connection of $M$ to $N$, How much testing is sufficient for uniform stability?
History

- Kansa (1990), application of MQ to PDEs. Sometimes refereed as Kansa method.
- Huge number of application to different PDE problems, 1990-present,

Some points from Schaback’s analysis:

- To overcome the insolvability, modify the setting,
- To get the error bound, there should at least be a unique solution to the modified discretization that converges to the true solution when the discretization is refined,
- Find least square solution instead,
- If $M \gg N$ (test space sufficiently larger than trial space; overtesting) then the system has full rank $N$ and stability condition is satisfied,
- Optimal convergence rate is obtained,
- The theory holds for a general cases including collocation, weak and local weak testing,
- **Open problem:** Connection of $M$ to $N$, How much testing is sufficient for uniform stability?
Symmetric collocation

Consider these operator equations

\[ Lu = f, \text{ in } \Omega \]
\[ Bu = g, \text{ on } \partial \Omega =: \Gamma \]

Let \( X = \{x_1, \ldots, x_N\} \subset \Omega \) be a trial and simultaneously a test set point with \( N = N_I + N_B \) points in domain and on the boundary, respectively (\( X = X_I \cup X_B \)). Approximate \( u \) by

\[ s = \sum_{j=1}^{N_I} c_j L^y \Phi(\cdot, x_j) + \sum_{\ell=1}^{N_B} d_\ell B^y \Phi(\cdot, x_\ell) \approx u \]

Approximation space \( = V_N := \text{span} \{L^y \Phi(\cdot, x_j) : x_j \in X_I\} + \text{span} \{B^y \Phi(\cdot, x_\ell) : x_\ell \in X_B\} \)

Apply \( L \) on both sides of PDE and \( B \) on both sides of boundary conditions and finally collocate at all points to get:
Symmetric collocation

Consider these operator equations

\[ Lu = f, \quad \text{in } \Omega \]
\[ Bu = g, \quad \text{on } \partial \Omega =: \Gamma \]

Let \( X = \{x_1, \ldots, x_N\} \subset \Omega \) be a trial and simultaneously a test set point with \( N = N_I + N_B \) points in domain and on the boundary, respectively \( (X = X_I \cup X_B) \). Approximate \( u \) by

\[
s = \sum_{j=1}^{N_I} c_j L^y \Phi(\cdot, x_j) + \sum_{\ell=1}^{N_B} d_\ell B^y \Phi(\cdot, x_\ell) \approx u
\]

Approximation space \( = V_N := \text{span}\{L^y \Phi(\cdot, x_j) : x_j \in X_I\} + \text{span}\{B^y \Phi(\cdot, x_\ell) : x_\ell \in X_B\} \)

Apply \( L \) on both sides of PDE and \( B \) on both sides of boundary conditions and finally collocate at all points to get:
Symmetric collocation

\[
\sum_{j=1}^{N_I} c_j L^x L^y \Phi(x_k, x_j) + \sum_{\ell=1}^{N_B} d_\ell L^x B^y \Phi(x_k, x_\ell) = f(x_k), \quad x_k \in X_I
\]

\[
\sum_{j=1}^{N_I} c_j B^x L^y \Phi(x_i, x_j) + \sum_{\ell=1}^{N_B} d_\ell B^x B^y \Phi(x_i, x_\ell) = g(x_i), \quad x_i \in X_B
\]

\[
\begin{bmatrix}
A_{LL} & A_{LB} \\
A_{BL} & A_{BB}
\end{bmatrix}
\begin{bmatrix}
c \\
d
\end{bmatrix}
= 
\begin{bmatrix}
f|_{X_I} \\
g|_{X_B}
\end{bmatrix}
\]

Some properties:

- Implementation is more involved than unsymmetric method,
- For special \( L \) and \( B \) the final system is symmetric; example \( L = \Delta, B = I_d \),
- If \( \Phi \) is positive definite, then so is the final system; discretization is solvable.
Some properties:

- Implementation is more involved than unsymmetric method,
- For special $L$ and $B$ the final system is symmetric; example $L = \Delta$, $B = Id$,
- If $\Phi$ is positive definite, then so is the final system; discretization is solvable.
Symmetric collocation

**History:**

- Introduced by Wu (1992), Narcowich & Ward (1994) referred as generalize Hermite interpolation,

- The first error analysis by Franke & Schaback (1998), introducing the generalized power function,

- Comparison with unsymmetric method and application to some PDEs by Fasshauer (1997, 1999),

- Error analysis for elliptic PDEs using sampling inequalities by Giesl & Wendland (2007),

Symmetric collocation

**Error analysis** Geisl & Wendland (2007)

Let $\Omega \subset \mathbb{R}^d$ has sufficiently smooth boundary and

$$Lu := \sum_{j=1}^{d} \kappa_{ij}(x) \partial_{i,j}u(x) + \sum_{i=1}^{d} \beta_i(x) \partial_i u(x) + \gamma(x)u(x)$$

where $\kappa(x)$, $\beta(x)$, $\gamma(x)$ are bounded, $\kappa(x)$ is symmetric and $\gamma(x) < 0$ for all $x \in \Omega$. Moreover, let $L$ be strictly elliptic, i.e. $\exists \lambda > 0$ such that for all $x \in \Omega$ and $\xi \in \mathbb{R}^d$, $\xi^T K(x) \xi \geq \lambda \|\xi\|_2^2$. Finally let the boundary conditions be Dirichlet i.e. $B = Id$. Then:

**Theorem:** Assume that $\Phi$ is a reproducing kernel for $H^\tau(\mathbb{R}^d)$ (i.e. $\widehat{\Phi} \sim (1 + \| \cdot \|_2^2)^{-\tau}$) and let $u \in H^\tau(\Omega)$ where $\tau > 2 + d/2$. Then

$$\|s - u\|_{L_\infty(\Omega)} \leq C \chi_{\mathcal{X}}^{\tau - 2 - d/2} \|u\|_{H^\tau(\Omega)}.$$
Symmetric collocation

**Wendland’s stability analysis** for special case $L = \Delta$ and $B = Id$:
Under the assumptions of the previous slide we have

$$\lambda_{\text{min}}(A) \geq Cq_X^{2\tau-d}$$

- Trade-off principle (even worse): $h_X^{\tau-2-d/2}$ for error and $q_X^{2\tau-d}$ for $\lambda_{\text{min}}$.
- The situation can be improved by an smoothing argument to get $\lambda_{\text{min}}(A) \geq Ch_X^{2\tau-d-4}$.

**More researches on symmetric collocation:**
- Symmetric collocation on spheres: Morton & Neamtu (2002),
- Multiscale symmetric collocation on $S^d$: Gia & Sloan & Wendland (2012),
- Multiscale symmetric collocation on $\mathbb{R}^d$: Farrel & Wendland (2013),
- ...
Symmetric collocation

**Wendland’s stability analysis** for special case \( L = \Delta \) and \( B = Id \):

Under the assumptions of the previous slide we have

\[
\lambda_{\text{min}}(A) \geq Cq^{2\tau - d}_X
\]

- **Trade-off principle** (even worse): \( h^{\tau - 2 - d/2}_X \) for error and \( q^{2\tau - d}_X \) for \( \lambda_{\text{min}} \).
- The situation can be improved by an smoothing argument to get

\[
\lambda_{\text{min}}(A) \geq Ch^{2\tau - d - 4}_X.
\]

**More researches on symmetric collocation:**

- Symmetric collocation on spheres: Morton & Neamtu (2002),
- Multiscale symmetric collocation on \( S^d \): Gia & Sloan & Wendland (2012),
- Multiscale symmetric collocation on \( \mathbb{R}^d \): Farrel & Wendland (2013),
- . . .
Symmetric collocation

**Wendland’s stability analysis** for special case $L = \Delta$ and $B = Id$:
Under the assumptions of the previous slide we have

$$\lambda_{\min}(A) \geq Cq_X^{2\tau-d}$$

- **Trade-off principle** (even worse): $h_X^{\tau-2-d/2}$ for error and $q_X^{2\tau-d}$ for $\lambda_{\min}$.
- The situation can be improved by an smoothing argument to get $\lambda_{\min}(A) \geq Ch_X^{2\tau-d-4}$.

**More researches on symmetric collocation:**
- Symmetric collocation on spheres: Morton & Neamtu (2002),
- Multiscale symmetric collocation on $S^d$: Gia & Sloan & Wendland (2012),
- Multiscale symmetric collocation on $\mathbb{R}^d$: Farrel & Wendland (2013),
- ...
Galerkin methods

find $u \in H$ such that $a(u, v) = \ell(v)$, for all $v \in H$

Let $s_{u,X} \approx u$ become from a kernel approximation space such as

$$V_N := \text{span}\{\Phi(\cdot, x_j) : x_j \in X\} + \mathbb{P}_m^d$$

- Few researches address this subject,
- The first paper by Wendland (1999) on Galerkin RBF for elliptic PDEs,
- Triangulation requires in test space (numerical integration), contrast with being truly meshless,
- Galerkin RBF on $S^d$ by Gia (2004), Narcowich & Rowe & Ward (2016),
Galerkin methods

\[ \text{find } u \in H \text{ such that } a(u, v) = \ell(v), \text{ for all } v \in H \]

Let \( s_{u,X} \approx u \) become from a kernel approximation space such as

\[ V_N := \text{span}\{\Phi(\cdot, x_j) : x_j \in X\} + \mathbb{P}_m^d \]

- Few researches address this subject,
- The first paper by Wendland (1999) on Galerkin RBF for elliptic PDEs,
- Triangulation requires in test space (numerical integration), contrast with being truly meshless,
- Galerkin RBF on \( S^d \) by Gia (2004), Narcowich & Rowe & Ward (2016),
- Petrov–Galerkin on \( S^d \) by D.M. (2016).
Moving Least Squares (MLS) approximation

- Shepard 1968: as an undergraduate student in Harvard suggests Shepard method,
- McLain 1974 and 1976: using the same method for surface generation,
- Lancaster & Salkauskas 1981: generalizing Shepard method and using title MLS,
- Blytschko, et. al. (1996): giving an error analysis (without some mathematical details) in multidimensional case,
- Levin (1998): connection to Backus-Gilbert optimality,
- Wendland (2001): giving much interesting error analysis,
- Other analyses: Armentano (2000), Zuppa (2004), D.M. (2015), ...

Application to PDEs:
- Nayroles et al 1992: Diffuse Element Method
- Belytschko et al 1994 Element Free Galerkin Method
- Atluri et al 1998 Meshless Local Petrov-Galerkin Method
- D.M. & Schaback 2013 Direct Meshless Local Petrov-Galerkin Method
- and many other applications ...
MLS

- Let \( \{p_1, p_2, \ldots, p_Q\} \) be the basis for \( \mathbb{P}_m^d \) where \( Q = \binom{m+d}{d} \).
- Given centers \( X = \{x_1, \ldots, x_N\} \) in \( \Omega \) and data values \( \{u(x_1), \ldots, u(x_N)\} \).
- \( \hat{u} \) is the MLS approximation of \( u \): \( \hat{u}(x) = p^T(x)b(x) = \sum_{k=1}^Q b_k(x)p_k(x) \), \( \forall x \in \Omega \).
- For \( x \in \Omega \) MLS approximates \( u(x) \) by \( \hat{u}(x) \) in a weighted square sense:

\[
\min_{b \in \mathbb{R}^Q} \sum_{j \in J(x)} \Phi_\delta(x - x_j)(p^T(x_j)b(x) - u(x_j))^2,
\]

where \( \Phi_\delta(x - y) = \phi(\|x - y\|_2 / \delta) \) is weight function and \( J(x) = \{j : \|x - x_j\|_2 \leq \delta\} \).
- The solution can also be written as

\[
\hat{u}(x) = u^T a(x) = \sum_{j=1}^N a_j(x)u(x_j), \quad a(x) = WP(P^TWP)^{-1}p(x), \quad x \in \Omega
\]

- Matrix \( A = P^T WP \) is \( Q \) by \( Q \) and positive definite.
Figure: MLS approximation: scattered points and local subdomains

Suppose that $\Omega \subset \mathbb{R}^d$ is a bounded set with a Lipschitz boundary. Let $m$ be a positive integer, $p \in [1, \infty]$, $q \in [1, \infty]$ and let $\alpha$ be a multi-index satisfying $m + 1 > |\alpha| + d/p$ for $p > 1$ and $m + 1 \geq |\alpha| + d$ for $p = 1$. If $u \in W_p^{m+1}(\Omega)$, there exist constants $C > 0$ and $h_0 > 0$ such that for all $X = \{x_1, \ldots, x_N\} \subset \Omega$ with $h_{X,\Omega} \leq \min\{h_0, 1\}$ which are quasi-uniform the estimate

$$
\|u - \hat{u}\|_{W_q^{\vert\alpha\vert}(\Omega)} \leq C h_{X,\Omega}^{m+1-\vert\alpha\vert-d(1/p-1/q)} + \|u\|_{W_p^{m+1}(\Omega)},
$$

holds. Here $(x)_+ = \max\{x, 0\}$ and $\hat{u}$ is the MLS approximation of $u$ in which the corresponding weight function $\phi : [0, \infty) \to \mathbb{R}$, is supported in $[0, 1]$ and its even extension belongs to $C^m(\mathbb{R})$. Besides the **shifted and scaled** basis polynomials

$$
\left\{ \frac{(x - z)^\alpha}{h_{X,\Omega}^{\vert\alpha\vert}} \right\}_{0 \leq \vert\alpha\vert \leq m},
$$

are employed for $\mathbb{P}_m^d$ at evaluation point $z \in \Omega$. 

---

Davoud Mirzaei  
Meshless Approximants
Note
The role of **shifted and scaled** basis functions are crucial to bound

$$\lambda_{\min}(A(x))$$

away from zero and independent of $h_{X,\Omega}$. Otherwise, experiments show that $\lambda_{\min}$ tends to zero when $h_{X,\Omega} \to 0$, thus $A(x)$ becomes ill-conditioned as the fill-distance refines. Here $A(x) = P^T W(x) P$ is the local MLS matrix. It is obviously a positive definite matrix.
**Functional approximation via MLS**

Motivated by numerical solution of PDEs, suppose that $\lambda$ is a functional in $U^*$ the dual space of function space $U$ which admits the MLS approximation. The value $\lambda(u)$ is usually approximated by $\lambda(\hat{u})$

$$\lambda(u) \approx \lambda(\hat{u}) = \sum_{j=1}^{N} \lambda(a_j)u(x_j).$$

Since shape functions $a_j$ are complicated and have no close forms, this leads to slow schemes specially when $\lambda$ is complicated. For example in weak-based meshless methods where numerical integration is required.

This is done in many meshless methods such as EFG (Blytschko, et. al.) and MLPG (Atluri, et. al.) and others.

But here we follow an alternative approach via a generalized MLS (GMLS).

Another approach to approximate \( \lambda(u) \) is a direct approximation from the data \( u(x_j) \) without detour via \( \lambda(\hat{u}) \). In this case we have

\[
\lambda(u) \approx \lambda(\hat{u}) = \sum_{j=1}^{N} a_j(\lambda)u(x_j),
\]

It should be noted that \( \lambda(a_j) \neq a_j(\lambda) \) in general.

If \( \lambda \) is finally evaluated at point \( x \in \Omega \), the same weight functions as in the classical MLS can be chosen independent of \( \lambda \). Using this assumption we can prove that

\[
a(\lambda) = WP^T (PWP^T)^{-1} \lambda(p).
\]

**Conclusion:** \( \lambda \) acts only on polynomials!

Another approach to approximate $\lambda(u)$ is a **direct approximation** from the data $u(x_j)$ without detour via $\lambda(\hat{u})$. In this case we have

$$\lambda(u) \approx \lambda(\hat{u}) = \sum_{j=1}^{N} a_j(\lambda)u(x_j),$$

It should be noted that $\lambda(a_j) \neq a_j(\lambda)$ in general.

If $\lambda$ is finally evaluated at point $x \in \Omega$, the same weight functions as in the classical MLS can be chosen independent of $\lambda$. Using this assumption we can prove that

$$a(\lambda) = WP^T (PWP^T)^{-1} \lambda(p).$$

**Conclusion:** $\lambda$ acts only on polynomials!

Another approach to approximate $\lambda(u)$ is a direct approximation from the data $u(x_j)$ without detour via $\lambda(\hat{u})$. In this case we have

$$\lambda(u) \approx \hat{\lambda}(u) = \sum_{j=1}^{N} a_j(\lambda)u(x_j),$$

It should be noted that $\lambda(a_j) \neq a_j(\lambda)$ in general.

If $\lambda$ is finally evaluated at point $x \in \Omega$, the same weight functions as in the classical MLS can be chosen independent of $\lambda$. Using this assumption we can prove that

$$a(\lambda) = WP^T (PWP^T)^{-1} \lambda(p).$$

**Conclusion:** $\lambda$ acts only on polynomials!
Theorem (special case $\lambda = D^\alpha$): D.M. (2016)

Suppose that $\Omega \subset \mathbb{R}^d$ is bounded set with a Lipschitz boundary. Let $m$ be a positive integer, $p \in [1, \infty], q \in [1, \infty]$ and let $\alpha$ be a multi-index satisfying $m + 1 > |\alpha| + d/p$ for $p > 1$ and $m + 1 \geq |\alpha| + d$ for $p = 1$. If $u \in W_p^{m+1}(\Omega)$, there exists constants $C > 0$ and $h_0 > 0$ such that for all $X = \{x_1, \ldots, x_N\} \subset \Omega$ with $h_{X,\Omega} \leq \min\{h_0, 1\}$ which are quasi-uniform the estimate

$$\left\| D^\alpha u - \hat{D}^\alpha u \right\|_{L^q(\Omega)} \leq C h_{X,\Omega}^{m+1 - |\alpha|-d(1/p-1/q)} + \|u\|_{W_q^{m+1}(\Omega)}$$

where $\hat{D}^\alpha u$ is the GMLS derivative approximation of order $\alpha$ to $u$. A $C^0$ weight function is enough for all $\alpha$ and a shifted and scaled basis functions should be applied.

Note:

$$D^\alpha (\hat{u}) \neq \hat{D}^\alpha u, \quad \text{if } u \notin \mathbb{P}_m^d$$

Standard derivatives $\neq$ GMLS derivatives
Theorem (special case $\lambda = D^\alpha$): D.M. (2016)

Suppose that $\Omega \subset \mathbb{R}^d$ is bounded set with a Lipschitz boundary. Let $m$ be a positive integer, $p \in [1, \infty], q \in [1, \infty]$ and let $\alpha$ be a multi-index satisfying $m + 1 > |\alpha| + d/p$ for $p > 1$ and $m + 1 \geq |\alpha| + d$ for $p = 1$. If $u \in W^{m+1}_p(\Omega)$, there exists constants $C > 0$ and $h_0 > 0$ such that for all $X = \{x_1, \ldots, x_N\} \subset \Omega$ with $h_{X,\Omega} \leq \min\{h_0, 1\}$ which are quasi-uniform the estimate

$$
\| D^\alpha u - \hat{D}^\alpha u \|_{L^q(\Omega)} \leq C h_{X,\Omega}^{m+1-|\alpha|-d(1/p-1/q)} + \| u \|_{W^{m+1}_q(\Omega)}
$$

where $\hat{D}^\alpha u$ is the GMLS derivative approximation of order $\alpha$ to $u$. A $C^0$ weight function is enough for all $\alpha$ and a shifted and scaled basis functions should be applied.

Note:

$$
D^\alpha (\hat{u}) \neq \hat{D}^\alpha u, \quad \text{if } u \notin \mathbb{P}^d_m
$$

Standard derivatives $\neq$ GMLS derivatives

- Assume that the discretized problem is set up with a vector
  \[ \mathbf{u} = (u(x_1), \ldots, u(x_N))^T \]
  of unknowns in “meshless” style, and all data have to be expressed in terms of this vector.

- Assume the discretized problem to consist of equations
  \[ \lambda_k(u) = \beta_k, \quad 1 \leq k \leq M, \quad M > N. \]

- Provide good estimates \( \hat{\lambda}_k \) of all \( \lambda_k \) using only values at nodes. Find real numbers \( a_j(\lambda_k) \) with
  \[ \hat{\lambda_k}(\mathbf{u}) = \sum_{j=1}^{M} a_j(\lambda_k) u(x_j) \approx \lambda_k(u) \quad \forall k, \quad 1 \leq k \leq M. \]

- Putting the \( a_j(\lambda_k) \) into an \( M \times N \) matrix \( \mathbf{A} \), one has to solve the possibly overdetermined linear system
  \[ \mathbf{A} \mathbf{u} = \mathbf{b} \]
  with \( \mathbf{b} = (\beta_1, \ldots, \beta_M)^T. \)
GMLS approximation

Shape function in GMLS:

\[ a(\lambda) = WP^T (P W P^T)^{-1} \lambda(p) \]

where \( \lambda(p) = (\lambda(p_1), \ldots, \lambda(p_Q))^T \in \mathbb{R}^Q. \)

Thus it suffices to evaluate \( \lambda \) on the space \( \mathbb{P}^d_m \), not on a certain trial space spanned by certain shape functions.

Standard examples are functionals of the form

\[ (\lambda u)(x) := \int_\Omega v_\sigma(x - y) L u(y) dy, \quad x \in \Omega \]

where \( v_\sigma \) is compactly supported function and \( L \) is a linear differential operator preserving polynomials or just the identity.
GMLS approximation

Shape function in GMLS:
\[ a(\lambda) = WP^T (P W P^T)^{-1} \lambda(p) \]

where \( \lambda(p) = (\lambda(p_1), \ldots, \lambda(p_Q))^T \in \mathbb{R}^Q \).

Thus it suffices to evaluate \( \lambda \) on the space \( \mathbb{P}_m^d \), not on a certain trial space spanned by certain shape functions.

Standard examples are functionals of the form
\[ (\lambda u)(x) := \int_{\Omega} v_\sigma(x-y) Lu(y)dy, \quad x \in \Omega \]

where \( v_\sigma \) is compactly supported function and \( L \) is a linear differential operator preserving polynomials or just the identity.
GMLS approximation

Shape function in GMLS:

\[ a(\lambda) = WP^T (PW^T)^{-1} \lambda(p) \]

where \( \lambda(p) = (\lambda(p_1), \ldots, \lambda(p_Q))^T \in \mathbb{R}^Q \).

Thus it suffices to evaluate \( \lambda \) on the space \( \mathbb{P}^d_m \), not on a certain trial space spanned by certain shape functions.

Standard examples are functionals of the form

\[ (\lambda u)(x) := \int_\Omega \nu_\sigma(x - y) Lu(y) dy, \quad x \in \Omega \]

where \( \nu_\sigma \) is compactly supported function and \( L \) is a linear differential operator preserving polynomials or just the identity.
Local weak forms

An elliptic problem

\[
\Delta u = f, \quad \text{in } \Omega, \\
u = u_D, \quad \text{on } \Gamma_D, \\
\frac{\partial u}{\partial n} = u_N, \quad \text{on } \Gamma_N.
\]

Symmetric local weak form

For example, consider the first symmetric local weak form:

\[
\lambda_{x,\sigma,v}(u) := \int_{\Gamma^x_\sigma \setminus \Gamma_N} (\nabla u \cdot n)v \, d\Gamma - \int_{\Omega^x_\sigma} \nabla u \cdot \nabla v \, d\Omega = \int_{\Omega^x_\sigma} fv \, d\Omega - \int_{\Gamma^x_\sigma \cap \Gamma_N} u_N v \, d\Gamma
\]

MLPG1/5/6 are based on this local weak form.
Local weak forms

An elliptic problem

\[ \Delta u = f, \quad \text{in } \Omega, \]

\[ u = u_D, \quad \text{on } \Gamma_D, \]

\[ \frac{\partial u}{\partial n} = u_N, \quad \text{on } \Gamma_N. \]

Symmetric local weak form

For example, consider the first symmetric local weak form:

\[ \lambda_{x,\sigma,v}(u) := \int_{\Gamma_{x,\sigma} \setminus \Gamma_N} (\nabla u \cdot n) v \, d\Gamma - \int_{\Omega_{x,\sigma}} \nabla u \cdot \nabla v \, d\Omega = \int_{\Omega_{x,\sigma}} f v \, d\Omega - \int_{\Gamma_{x,\sigma} \cap \Gamma_N} u_N v \, d\Gamma \]

MLPG1/5/6 are based on this local weak form.
Figure: Meshless points and local subdomains
Corollary

In situation of previous Theorem, suppose that

\[(\lambda u)(x) := \int_{\Omega} v_\sigma(x - y) Lu(y)dy, \quad x \in \Omega,\]

and

\[\hat{\lambda}u := \sum_{j=1}^{N} a_j(\lambda)u(x_j).\]

There exists positive constants \(C = C(\sigma)\) and \(h_0\) such that for \(u \in W_{p+1}^m(\Omega)\) and for all quasi-uniform set \(X \subset \Omega\) with \(h_{X,\Omega} \leq h_0\) we have

\[\|\lambda u - \hat{\lambda}u\|_{L^q(\Omega)} \leq Ch_{X,\Omega}^{n+1-\ell-d(1/p-1/q)} + \|u\|_{W_{p+1}^m(\Omega)}\]

Here \(\ell\) is the maximal order of derivatives involved in linear operator \(L\) and \(n, n \leq m\), is the largest integer for which \(\lambda(\mathbb{P}_n^d) \neq 0\) but \(\lambda(\mathbb{P}_k) = 0\) for \(n < k \leq m\).
Implementation

- To impose the Dirichlet BC, set $\lambda_k = \delta_{x_k}$ for $x_k \in Y_D$.
- For interior test points and points located on Neumann BC, set $\lambda_k = \lambda_{x_k,\sigma_k,v_k}$.
- Set up $\mathcal{A}$ by elements $a_j(\lambda_k)$ and solve $\mathcal{A}u = b$ for vector values $u$.

Title

Since we have direct approximations for boundary conditions and local weak forms, this method is called Direct Meshless Local Petrov-Galerkin (DMLPG) method.
Implementation

- To impose the Dirichlet BC, set $\lambda_k = \delta x_k$ for $x_k \in Y_D$.
- For interior test points and points located on Neumann BC, set $\lambda_k = \lambda x_k, \sigma_k, v_k$.
- Set up $A$ by elements $a_j(\lambda_k)$ and solve $Au = b$ for vector values $u$.

Title

Since we have direct approximations for boundary conditions and local weak forms, this method is called Direct Meshless Local Petrov-Galerkin (DMLPG) method.
DMLPG v.s. MLPG

1. In DMLPG integrations are done over low-degree and close form polynomials, while in MLPG these are done over complicated MLS shape functions (with no close forms).

2. In DMLPG for every functional $\lambda_k, 1 \leq k \leq M \geq N$, the GMLS routine is called only once, while in MLPG it is called approximatively $MK$ times where $K$ is the average number of integration points.

3. In DMLPG numerical integrations can sometimes be performed exactly.

4. DMLPG is absolutely faster than MLPG.

5. The order of convergence in both methods are the same, but numerically DMLPG turn often out to be more accurate than MLPG because of avoiding many calculations and thus roundoff errors.

6. DMLPG does not work for $m = 1$, while MLPG does. ($\lambda_k(\pi_1^d) = 0$).

7. In both methods the degree $m = 2k + 1$ cannot improve the behavior for $m = 2k$. ($n = 2k$)
Poisson equation

- We consider the Poisson equation with Dirichlet BC on $[0, 1]^2$.
- We let Franke’s function be the exact solution.
- We provide regular mesh distributions with mesh-size $h$, though the methods would work with scattered data.
- We examine with $m = 2, 3, 4$ in $P_m^2$. 
Maximum errors, ratios and CPU times used for MLPG5 and DMLPG5 with $m = 2$

<table>
<thead>
<tr>
<th>$h$</th>
<th>MLPG $|e|_{\infty}$</th>
<th>order</th>
<th>DMLPG $|e|_{\infty}$</th>
<th>order</th>
<th>Times MLPG</th>
<th>DMLPG</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>$0.44 \times 10^{-1}$</td>
<td>—</td>
<td>$0.23 \times 10^{-1}$</td>
<td>—</td>
<td>0.9 sec.</td>
<td>0.0 sec.</td>
</tr>
<tr>
<td>0.1</td>
<td>$0.15 \times 10^{-1}$</td>
<td>1.59</td>
<td>$0.72 \times 10^{-2}$</td>
<td>1.68</td>
<td>5.8</td>
<td>0.0</td>
</tr>
<tr>
<td>0.05</td>
<td>$0.73 \times 10^{-2}$</td>
<td>0.99</td>
<td>$0.20 \times 10^{-2}$</td>
<td>1.84</td>
<td>24.1</td>
<td>0.1</td>
</tr>
<tr>
<td>0.025</td>
<td>$0.24 \times 10^{-2}$</td>
<td>1.61</td>
<td>$0.58 \times 10^{-3}$</td>
<td>1.80</td>
<td>117.5</td>
<td>0.8</td>
</tr>
<tr>
<td>0.0125</td>
<td>$0.66 \times 10^{-3}$</td>
<td>1.85</td>
<td>$0.14 \times 10^{-3}$</td>
<td>1.98</td>
<td>1065.0</td>
<td>5.0</td>
</tr>
</tbody>
</table>

![Graph showing comparison between MLPG and DMLPG for different $h$ values.](image)
Maximum errors, ratios and CPU times used for MLPG5 and DMLPG5 with $m = 3$

<table>
<thead>
<tr>
<th>$h$</th>
<th>MLPG $|e|_\infty$</th>
<th>order</th>
<th>DMLPG $|e|_\infty$</th>
<th>order</th>
<th>Times</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>$0.28 \times 10^{-1}$</td>
<td></td>
<td>$0.23 \times 10^{-1}$</td>
<td></td>
<td>1.0 sec.</td>
</tr>
<tr>
<td>0.1</td>
<td>$0.13 \times 10^{-1}$</td>
<td>1.08</td>
<td>$0.74 \times 10^{-2}$</td>
<td>1.62</td>
<td>7.7</td>
</tr>
<tr>
<td>0.05</td>
<td>$0.33 \times 10^{-2}$</td>
<td>1.98</td>
<td>$0.20 \times 10^{-2}$</td>
<td>1.89</td>
<td>30.3</td>
</tr>
<tr>
<td>0.025</td>
<td>$0.78 \times 10^{-3}$</td>
<td>2.09</td>
<td>$0.58 \times 10^{-3}$</td>
<td>1.80</td>
<td>143.6</td>
</tr>
<tr>
<td>0.0125</td>
<td>$0.19 \times 10^{-3}$</td>
<td>2.06</td>
<td>$0.15 \times 10^{-3}$</td>
<td>1.98</td>
<td>1334.0</td>
</tr>
</tbody>
</table>

- $DMLPG5$  
- $MLPG5$
Maximum errors, ratios and CPU times used for MLPG5 and DMLPG5 with \( m = 4 \)

<table>
<thead>
<tr>
<th>( h )</th>
<th>( | e |_\infty )</th>
<th>( | e |_\infty )</th>
<th>Times</th>
<th>MLPG</th>
<th>DMLPG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>order</td>
<td>order</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>( 0.10 \times 10^0 )</td>
<td>( 0.12 \times 10^0 )</td>
<td>1.5 sec.</td>
<td>1.5 sec.</td>
<td>0.0 sec.</td>
</tr>
<tr>
<td>0.1</td>
<td>( 0.25 \times 10^{-1} )</td>
<td>( 0.17 \times 10^{-1} )</td>
<td>9.9</td>
<td>9.9</td>
<td>0.0</td>
</tr>
<tr>
<td>0.05</td>
<td>( 0.78 \times 10^{-2} )</td>
<td>( 0.12 \times 10^{-2} )</td>
<td>48.6</td>
<td>48.6</td>
<td>0.0</td>
</tr>
<tr>
<td>0.025</td>
<td>( 0.79 \times 10^{-3} )</td>
<td>( 0.75 \times 10^{-4} )</td>
<td>179.4</td>
<td>179.4</td>
<td>0.6</td>
</tr>
<tr>
<td>0.0125</td>
<td>( 0.55 \times 10^{-4} )</td>
<td>( 0.43 \times 10^{-5} )</td>
<td>1803.0</td>
<td>1803.0</td>
<td>4.2</td>
</tr>
</tbody>
</table>

\[ h \]

\[ \| e \|_\infty \]

\[ \| e \|_\infty \]

\[ \text{Times} \]

\[ \text{MLPG} \]

\[ \text{DMLPG} \]
An elasto-estatic problem

\[ \nabla \cdot \mathbf{\sigma} + \mathbf{b} = 0, \quad \text{in } \Omega, \]
\[ \mathbf{u} = \mathbf{\bar{u}}, \quad \text{on } \Gamma_u, \]
\[ t = \mathbf{\sigma} \cdot \mathbf{n} = \mathbf{\bar{t}}, \quad \text{on } \Gamma_t, \]

where \( \mathbf{\sigma} \) is the stress tensor, which corresponds to the displacement field \( \mathbf{u} \), and \( \mathbf{b} \) is the body force.

Cantilever beam

\[ \frac{P}{g^3}, \quad \frac{L}{g^3}, \quad \frac{D}{g^3} \]

Davoud Mirzaei
Meshless Approximants
Figure: Relative displacement and strain errors and CPU times used
**Boussinesq problem**

The Boussinesq problem can be described as a concentrated load acting on a semi-infinite elastic medium with no body force.

In numerical simulation, a finite sphere with large radius $b = 10$ is used. Due to the symmetry, a first one-eighth of the sphere is considered and symmetry boundary conditions are applied on planes $xz$ and $yz$.

In order to avoid direct encounter with the singular loading point, the theoretical displacement is applied on a small spherical surface with radius $b/40 = 0.25$. 

![Diagram of Boussinesq problem](image)
A 2 × 2 × 2 = 8—point Gaussian quadrature gives exact integrations in DMLPG, while MLPG uses 1000 = 10 × 10 × 10 integration points in this example. Thus MLS subroutine should be called $M$ times in DMLPG and 1000$M$ times in MLPG.

**CPU times used for this example:**

$MLPG : 7400 \text{ sec.}$  $DMLPG : 3.2 \text{ sec.}$
An elasto-dynamic problem

\[ \nabla \cdot \sigma + b = c\dot{u} + \rho \ddot{u}, \quad \text{in } \Omega, \]
\[ u = \bar{u}, \quad \text{on } \Gamma_u, \]
\[ t = \sigma \cdot n = \bar{t}, \quad \text{on } \Gamma_t, \]
\[ u(x, 0) = u_0(x), \quad x \in \Omega, \]
\[ \dot{u}(x, 0) = \dot{u}_0(x), \quad x \in \Omega, \]

Semi-discrete problem via DMLPG

\[ M\ddot{U}(t) + C\dot{U}(t) + KU(t) = F(t) \]
\[ BU(t) = \bar{U}(t) \]
\[ U(0) = U_0, \quad \dot{U}(0) = \dot{U}_0. \]

A Newmark time integration scheme (an implicit and unconditionally stable scheme) is applied.
An elasto-dynamic problem

\[ \nabla \cdot \sigma + b = c\dot{u} + \rho\ddot{u}, \quad \text{in } \Omega, \]
\[ u = \bar{u}, \quad \text{on } \Gamma_u, \]
\[ t = \sigma \cdot n = \bar{t}, \quad \text{on } \Gamma_t, \]
\[ u(x, 0) = u_0(x), \quad x \in \Omega, \]
\[ \dot{u}(x, 0) = \dot{u}_0(x), \quad x \in \Omega, \]

Semi-discrete problem via DMLPG

\[ M\ddot{U}(t) + C\dot{U}(t) + KU(t) = F(t) \]
\[ BU(t) = \bar{U}(t) \]
\[ U(0) = U_0, \quad \dot{U}(0) = \dot{U}_0. \]

A Newmark time integration scheme (an implicit and unconditionally stable scheme) is applied.
Exact displacement and normal stress:

\[ u_1(x, t) = \frac{8PL}{\pi^2 E} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)^2} \sin \frac{(2n-1)\pi x}{2L} (1 - \cos \omega_n t), \]

\[ \sigma(x, t) = E \frac{\partial u_1}{\partial x} = \frac{4P}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{2n-1} \cos \frac{(2n-1)\pi x}{2L} (1 - \cos \omega_n t), \]

where \( \omega_n = \frac{(2n-1)\pi}{2L} \sqrt{\frac{E}{\rho}}. \)
Figure: Time variations of displacement $u_1$ at point A (top-left) and point B (top-right), and CPU time used (bottom)
Vibration analysis of a cantilever beam

2 cases:

- $g(t) = \sin(27t)$, Harmonic load,
- $g(t) = \begin{cases} 1, & 0 \leq t \leq 1, \\ 0, & t > 0.5. \end{cases}$, Suddenly loaded and suddenly vanished.
Figure: Time variations of $u_2$ at point $A$ for harmonic load at $c = 0$ (top), and at $c = 0.4$ (bottom)
Approximation by Kernels
Solving PDEs by Kernels
Approximation by MLS
GMLS Approximation and DMLPG Methods
Main References

A GMLS approximation
DMLPG methods
Some numerical results

Davoud Mirzaei

Meshless Approximants

Figure: Time variations of \( u_2 \) at point A for suddenly load at \( c = 0 \) (top), and at \( c = 0.4 \) (down)
Main References I


Main References II


Main References III


Main References IV


Main References V


Main References VI


Main References VII


Thank You for Your Attention