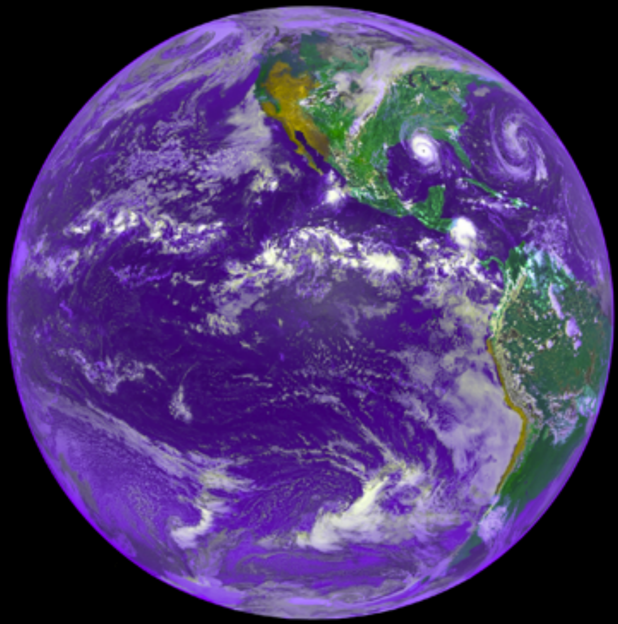
# 2014 Montestigliano Workshop

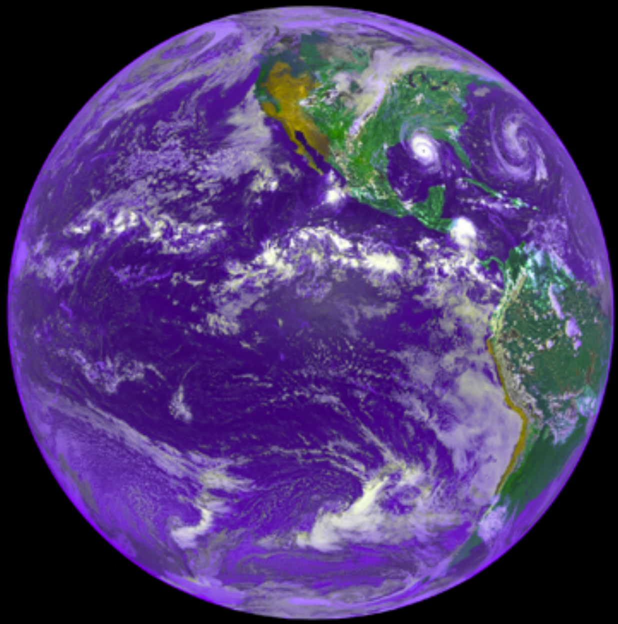


# Radial Basis Functions for Scientific Computing

Grady B. Wright
Boise State University

*This work is supported by NSF grants DMS 0934581

# 2014 Montestigliano Workshop



## Part I: Introduction

*Supplementary lecture slides*


Grady B. Wright

Boise State University



*This work is supported by NSF grants DMS 0934581

# Overview



- Scattered data interpolation in $\mathbb{R}^d$
  - Positive definite radial kernels: radial basis functions (RBF)
  - Some theory
- Scattered data interpolation on the sphere $\mathbb{S}^2$
  - Positive definite (PD) zonal kernels
  - Brief review of spherical harmonics
  - Characterization of PD zonal kernels
  - Conditionally positive definite zonal kernels
  - Examples
- Error estimates:
  - Reproducing kernel Hilbert spaces
  - Sobolev spaces
  - Native spaces
  - Geometric properties of node sets
- Optimal nodes on the sphere

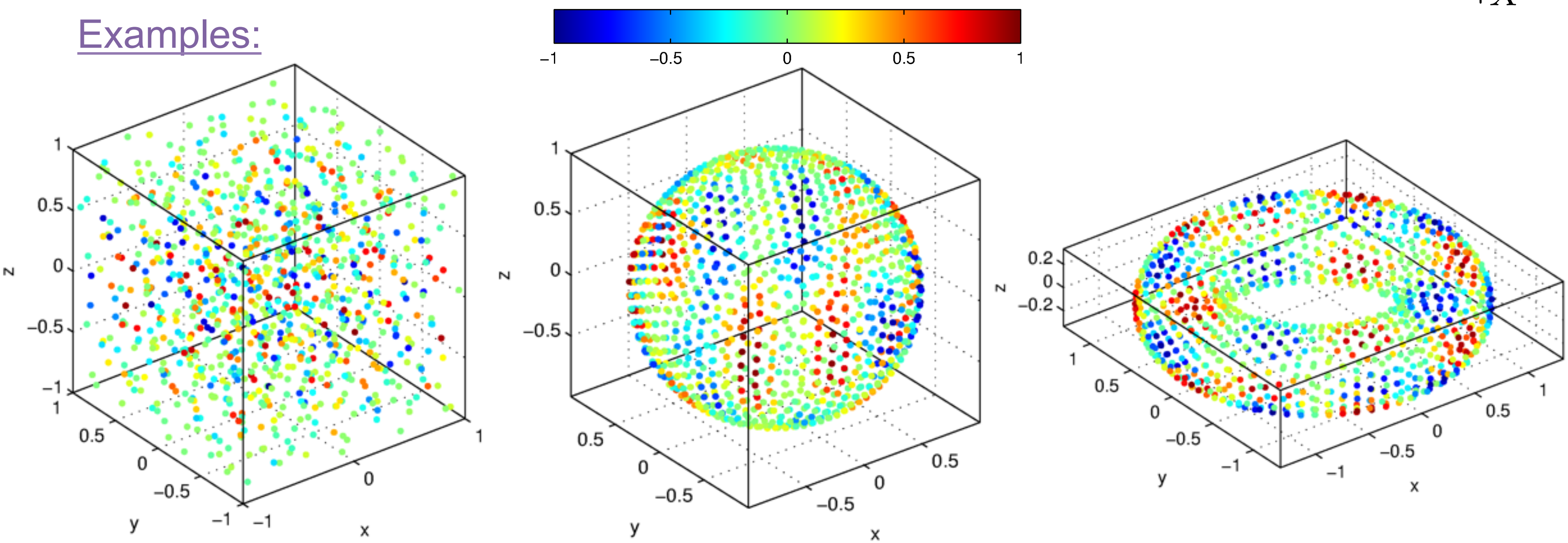
# Interpolation with kernels

*Supplementary material*

- Let $\Omega \subset \mathbb{R}^d$ and $X = \{\mathbf{x}_j\}_{j=1}^N$ a set of nodes on $\Omega$.

- Consider a continuous target function $f : \Omega \to \mathbb{R}$ sampled at $X$: $f\big|_X$.

Examples:



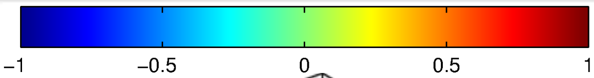

$\Omega = [-1,1]^3$ $\Omega = \mathbb{S}^2$ $\Omega = \mathbb{T}^2$

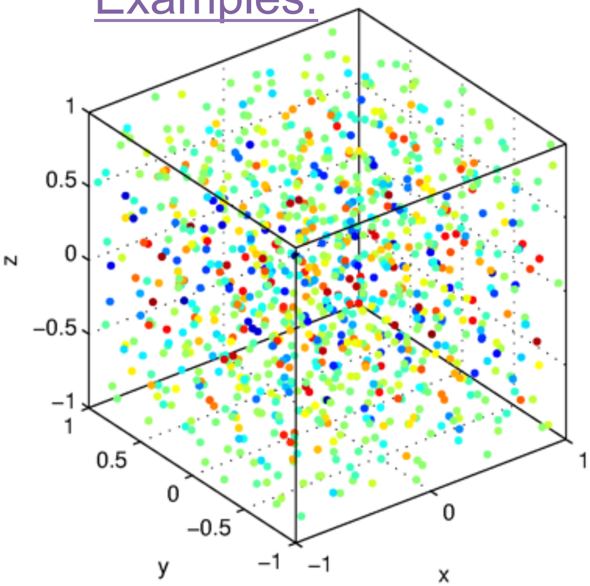
- Kernel interpolant to $f\big|_X$:
$$I_X f = \sum_{j=1}^{N} c_j \Phi(\cdot, \mathbf{x}_j)$$
where $\Phi : \Omega \times \Omega \to \mathbb{R}$ and $c_j$ come from requiring $I_X f\big|_X = f\big|_X$

# Interpolation with kernels

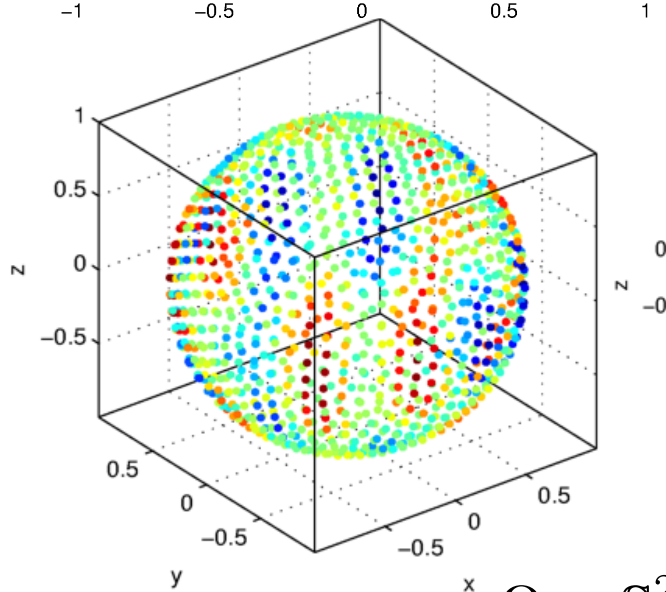


Examples:

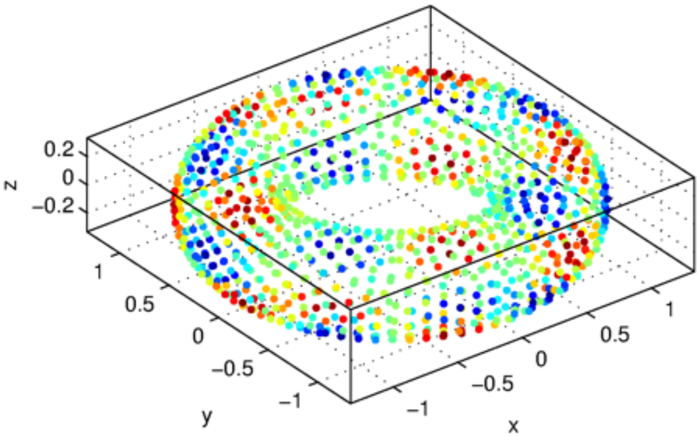



$\Omega = [-1,1]^3$




$\Omega = \mathbb{S}^2$




$\Omega = \mathbb{T}^2$

- Kernel interpolant to $f\big|_X$: $\qquad I_X f = \sum_{j=1}^{N} c_j \Phi(\cdot, \mathbf{x}_j)$

- Definition: $\Phi$ is a positive definite kernel on $\Omega$ if the matrix $A = \{\Phi(\mathbf{x}_i, \mathbf{x}_j)\}$ is positive definite for any distinct $X = \{\mathbf{x}_j\}_{j=1}^N \subset \Omega$, i.e.

$$\sum_{i=1}^{N} \sum_{j=1}^{N} b_i \Phi(\mathbf{x}_i, \mathbf{x}_j) b_j > 0, \text{ provided } \{b_i\}_{i=1}^N \not\equiv 0.$$

- In this case $c_j$ are uniquely determined by $X$ and $f\big|_X$.

# Interpolation with kernels



- Kernel interpolant to $f\big|_X$: $\qquad I_X f = \sum_j c_j \Phi(\cdot, \mathbf{x}_j)$.
- Some considerations for choosing the kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$

  1. The kernel should be easy to compute.
  2. The kernel interpolant should be uniquely determined by $X$ and $f\big|_X$.
  3. The kernel interpolant should accurately reconstruct $f$.

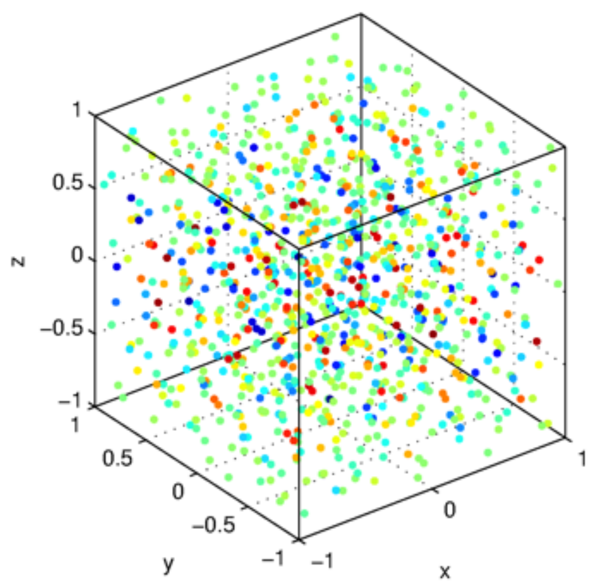
# Interpolation with kernels



- Kernel interpolant to $f\big|_X$: $\quad I_X f = \sum_j c_j \Phi(\cdot, \mathbf{x}_j)$.
- Some considerations for choosing the kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$

  1. The kernel should be easy to compute.
  2. The kernel interpolant should be uniquely determined by $X$ and $f\big|_X$.
  3. The kernel interpolant should accurately reconstruct $f$.

- For problems like




$\Omega = [-1, 1]^3$

Obvious choice: $\phi$ is a (conditionally) positive definite radial kernel

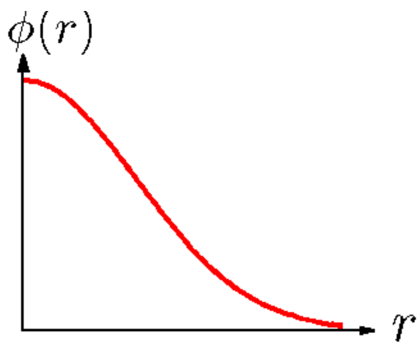
$$\Phi(\mathbf{x}, \mathbf{x}_j) = \phi(\|\mathbf{x} - \mathbf{x}_j\|_2) = \phi(r)$$

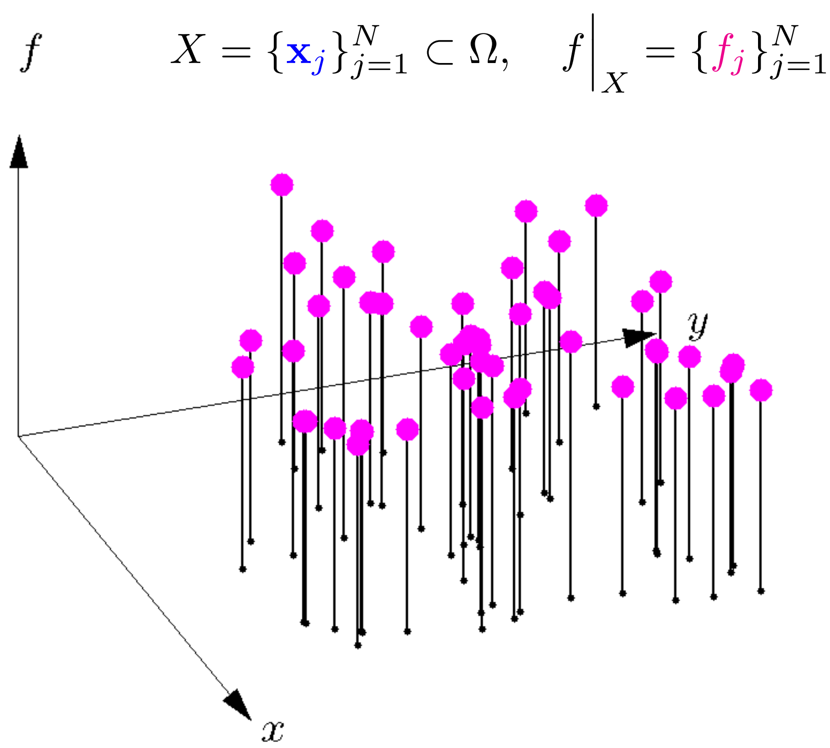
- Leads to radial basis function (RBF) interpolation.

# Radial basis function (RBF) interpolation

*Supplementary material*

Key idea: linear combination of translates and rotations of a single radial kernel:




Basic RBF Interpolant for $\Omega \subseteq \mathbb{R}^2$

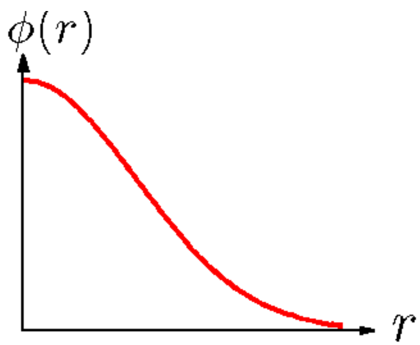
$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \phi(\|\mathbf{x} - \mathbf{x}_j\|)$$

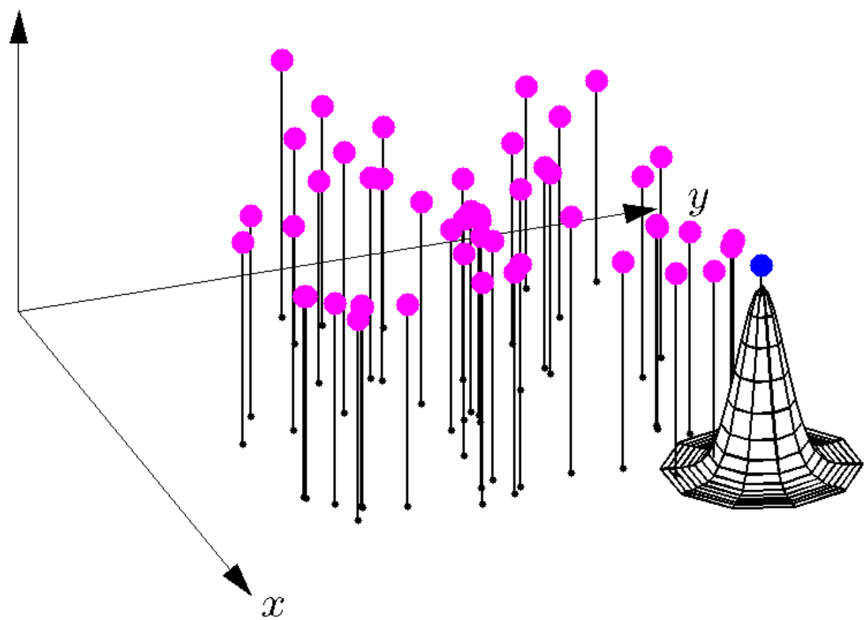
where $\|\mathbf{x} - \mathbf{x}_j\| = \sqrt{(x - x_j)^2 + (y - y_j)^2}$

# Radial basis function (RBF) interpolation

*Supplementary material*

Key idea: linear combination of translates and rotations of a single radial kernel:




Basic RBF Interpolant for $\Omega \subseteq \mathbb{R}^2$

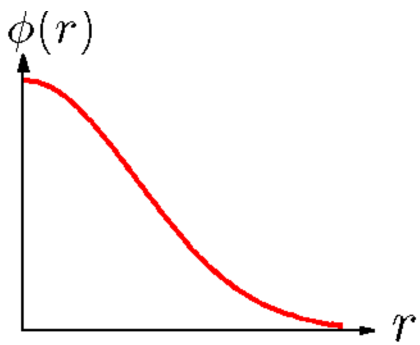
$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \phi(\|\mathbf{x} - \mathbf{x}_j\|)$$

$$\text{where } \|\mathbf{x} - \mathbf{x}_j\| = \sqrt{(x - x_j)^2 + (y - y_j)^2}$$

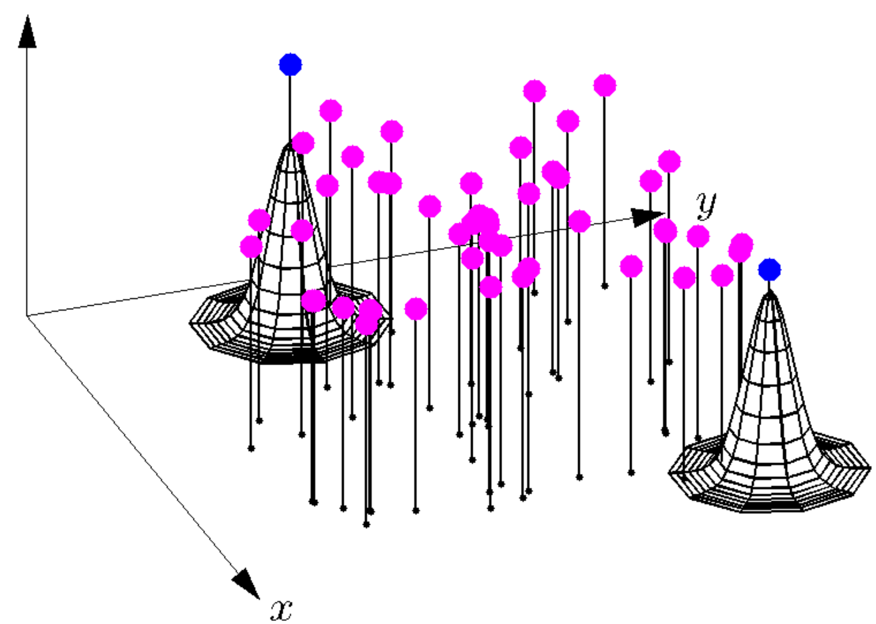
# Radial basis function (RBF) interpolation

*Supplementary material*

Key idea: linear combination of translates and rotations of a single radial kernel:




$f \qquad X = \{\mathbf{x}_j\}_{j=1}^{N} \subset \Omega, \quad f\Big|_X = \{f_j\}_{j=1}^{N}$




Basic RBF Interpolant for $\Omega \subseteq \mathbb{R}^2$

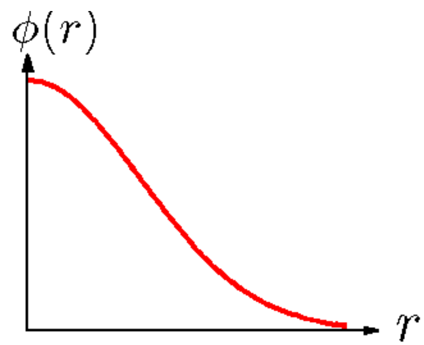
$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \phi(\|\mathbf{x} - \mathbf{x}_j\|)$$

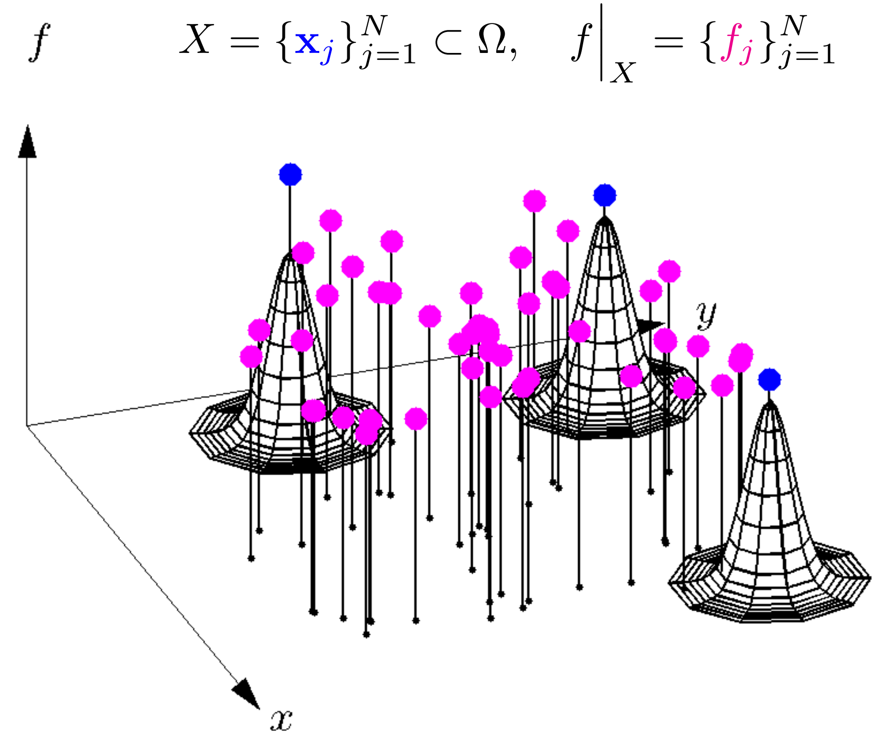
$$\text{where } \|\mathbf{x} - \mathbf{x}_j\| = \sqrt{(x - x_j)^2 + (y - y_j)^2}$$

# Radial basis function (RBF) interpolation



Key idea: linear combination of translates and rotations of a single radial kernel:




$X = \{\mathbf{x}_j\}_{j=1}^N \subset \Omega, \quad f\Big|_X = \{f_j\}_{j=1}^N$

Basic RBF Interpolant for $\Omega \subseteq \mathbb{R}^2$

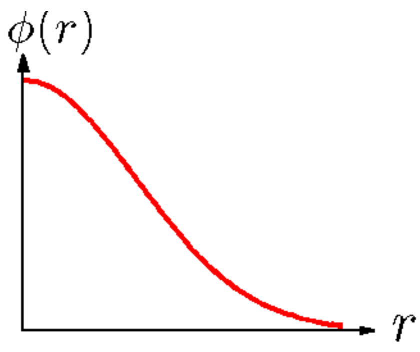
$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \phi(\|\mathbf{x} - \mathbf{x}_j\|)$$

where $\|\mathbf{x} - \mathbf{x}_j\| = \sqrt{(x - x_j)^2 + (y - y_j)^2}$

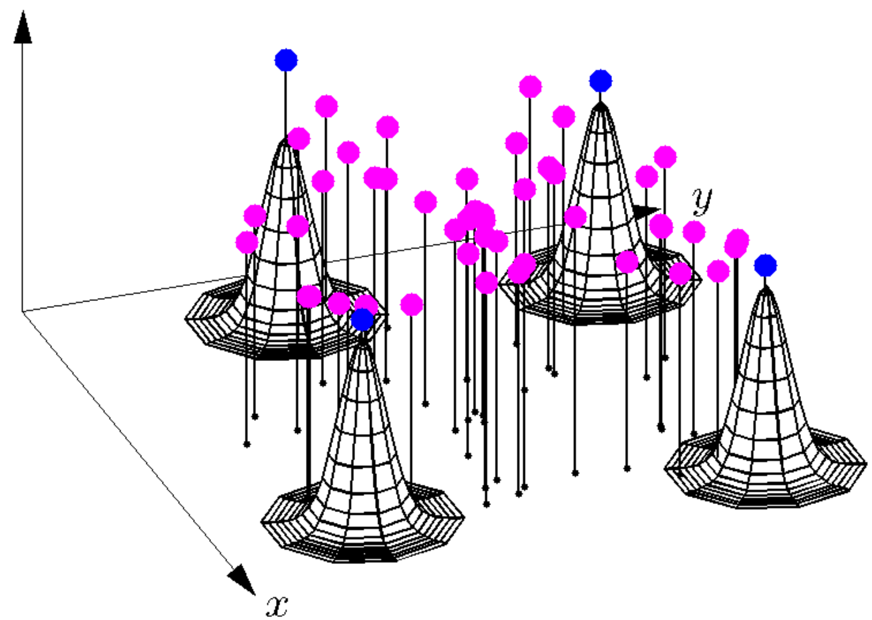
# Radial basis function (RBF) interpolation



Key idea: linear combination of translates and rotations of a single radial kernel:




$$X = \{\mathbf{x}_j\}_{j=1}^{N} \subset \Omega, \quad f\Big|_X = \{f_j\}_{j=1}^{N}$$




Basic RBF Interpolant for $\Omega \subseteq \mathbb{R}^2$

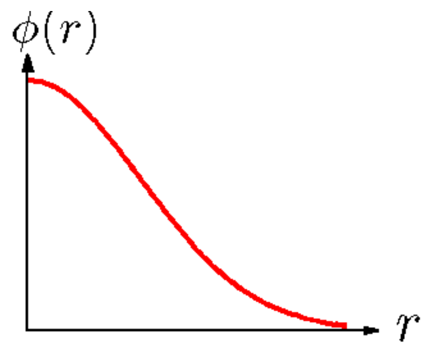
$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \phi(\|\mathbf{x} - \mathbf{x}_j\|)$$

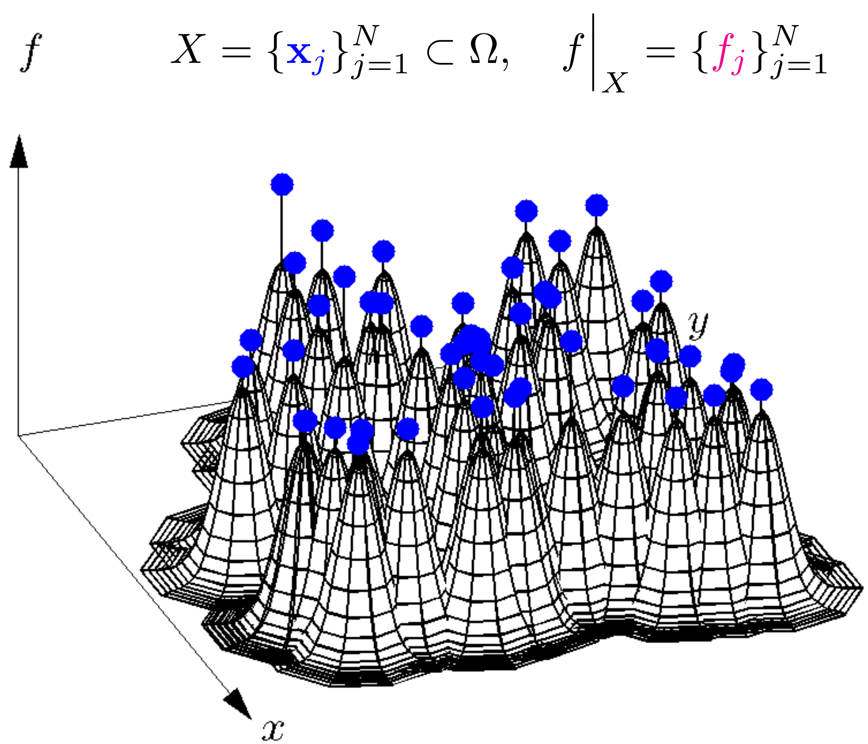
$$\text{where } \|\mathbf{x} - \mathbf{x}_j\| = \sqrt{(x - x_j)^2 + (y - y_j)^2}$$

# Radial basis function (RBF) interpolation



<u>Key idea</u>: linear combination of translates and rotations of a single radial kernel:




<u>Basic RBF Interpolant for $\Omega \subseteq \mathbb{R}^2$</u>

$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \phi(\|\mathbf{x} - \mathbf{x}_j\|)$$

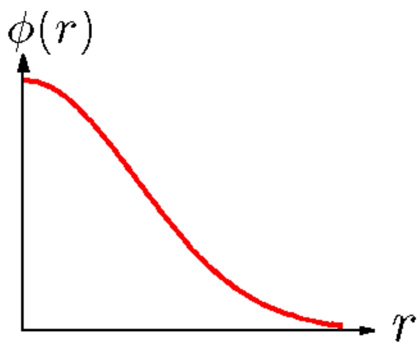
$$\text{where } \|\mathbf{x} - \mathbf{x}_j\| = \sqrt{(x - x_j)^2 + (y - y_j)^2}$$

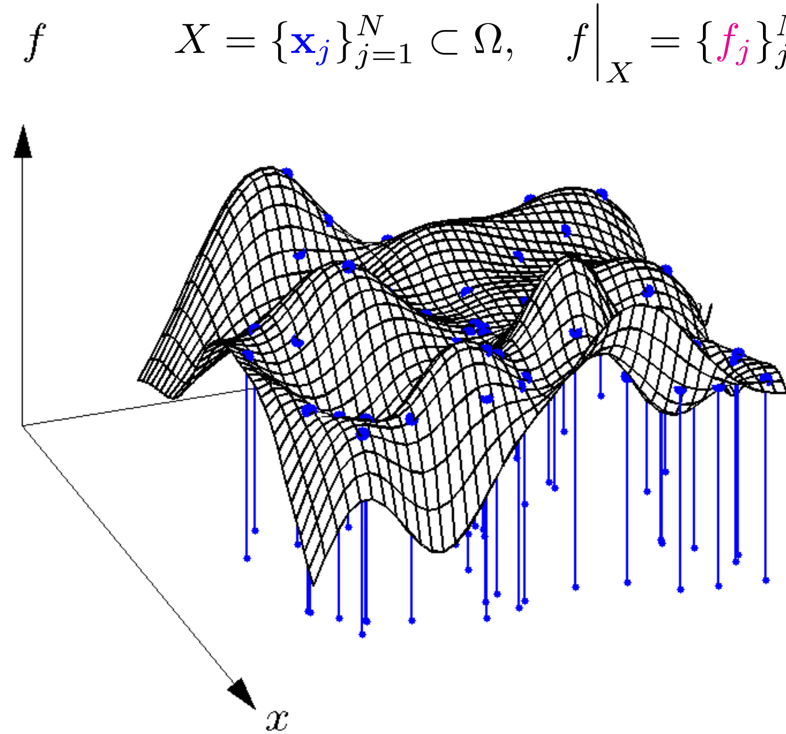
# Radial basis function (RBF) interpolation

*Supplementary material*

<u>Key idea</u>: linear combination of translates and rotations of a single radial kernel:

$$X = \{\mathbf{x}_j\}_{j=1}^N \subset \Omega, \quad f\Big|_X = \{f_j\}_{j=1}^N$$




<u>Basic RBF Interpolant for $\Omega \subseteq \mathbb{R}^2$</u>

$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \phi(\|\mathbf{x} - \mathbf{x}_j\|)$$

<u>Linear system for determining the interpolation coefficients</u>

$$\underbrace{\begin{bmatrix} \phi(\|\mathbf{x}_1 - \mathbf{x}_1\|) & \phi(\|\mathbf{x}_1 - \mathbf{x}_2\|) \cdots & \phi(\|\mathbf{x}_1 - \mathbf{x}_N\|) \\ \phi(\|\mathbf{x}_2 - \mathbf{x}_1\|) & \phi(\|\mathbf{x}_2 - \mathbf{x}_2\|) \cdots & \phi(\|\mathbf{x}_2 - \mathbf{x}_N\|) \\ \vdots & \vdots \quad \ddots & \vdots \\ \phi(\|\mathbf{x}_N - \mathbf{x}_1\|) & \phi(\|\mathbf{x}_N - \mathbf{x}_2\|) \cdots & \phi(\|\mathbf{x}_N - \mathbf{x}_N\|) \end{bmatrix}}_{A_X} \underbrace{\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix}}_{\underline{c}} = \underbrace{\begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_N \end{bmatrix}}_{\underline{f}}$$

$A_X$ is guaranteed to be positive definite if $\phi$ is positive definite.

# Positive definite radial kernels



- Some results on positive definite radial kernels.

**Theorem.** If $\phi \in C[0, \infty)$ with $\phi(0) > 0$ and $\phi(\rho) < 0$ for some $\rho > 0$, then $\phi$ cannot be positive definite in $\mathbb{R}^d$ for all $d$.

# Positive definite radial kernels

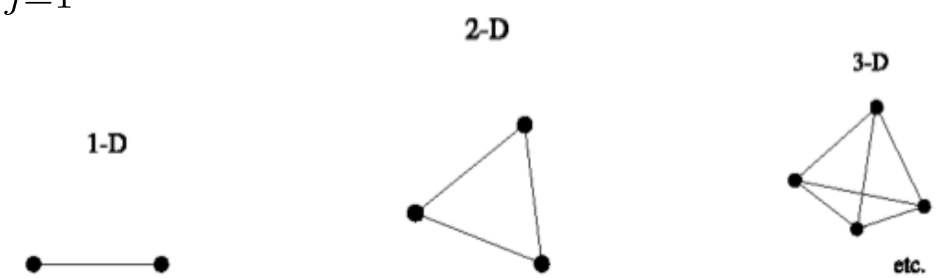


- Some results on positive definite radial kernels.

**Theorem.** If $\phi \in C[0,\infty)$ with $\phi(0) > 0$ and $\phi(\rho) < 0$ for some $\rho > 0$, then $\phi$ cannot be positive definite in $\mathbb{R}^d$ for all $d$.

Proof

Consider $X$ to be the vertices of an $m$ dimensional simplex with spacing $\rho$, i.e. $X = \{\mathbf{x}\}_{j=1}^{m+1} \subset \mathbb{R}^m$




Then

$$\sum_{i=1}^{m+1}\sum_{j=1}^{m+1} \phi(\|\mathbf{x}_i - \mathbf{x}_j\|) = \sum_{i=1}^{m+1} \phi(0) + \sum_{i=1}^{m+1} \sum_{j=1, j\neq i}^{m+1} \phi(\rho)$$
$$= (m+1)[\phi(0) + m\phi(\rho)].$$

Given $\phi(0) > 0$, we can find a $\rho$ for which $\phi(\rho) < 0$ and an $m$ to make this sum zero.

# Positive definite radial kernels



- Some results on positive definite radial kernels.

**Definition.** A function $\Phi : [0, \infty) \to \mathbb{R}$ is said to be completely monotone on $[0, \infty)$ if

$$(1)\ \Phi \in C[0,\infty), \quad (2)\ \Phi \in C^{\infty}(0,\infty), \quad (3)\ (-1)^k \Phi^{(k)}(t) \geq 0,\ t > 0,\ k = 0, 1, \ldots$$

# Positive definite radial kernels



- Some results on positive definite radial kernels.

**Definition.** A function $\Phi : [0, \infty) \to \mathbb{R}$ is said to be completely monotone on $[0, \infty)$ if

$$(1)\ \Phi \in C[0,\infty), \quad (2)\ \Phi \in C^{\infty}(0,\infty), \quad (3)\ (-1)^k \Phi^{(k)}(t) \geq 0,\ t > 0,\ k = 0, 1, \ldots$$

**Theorem** (Hausdorff-Bernstien-Widder). A function $\Phi$ is completely monotone if and only if it can be written in the form

$$\Phi(t) = \int_0^{\infty} e^{-st} d\gamma(s),$$

where $\gamma(s)$ is bounded, non-decreasing, and not concentrated at zero.

# Positive definite radial kernels



- Some results on positive definite radial kernels.

**Definition.** A function $\Phi : [0, \infty) \to \mathbb{R}$ is said to be completely monotone on $[0, \infty)$ if

$$(1)\ \Phi \in C[0,\infty), \quad (2)\ \Phi \in C^{\infty}(0,\infty), \quad (3)\ (-1)^k \Phi^{(k)}(t) \geq 0,\ t > 0,\ k = 0, 1, \ldots$$

**Theorem** (Hausdorff-Bernstien-Widder). A function $\Phi$ is completely monotone if and only if it can be written in the form

$$\Phi(t) = \int_0^{\infty} e^{-st} d\gamma(s),$$

where $\gamma(s)$ is bounded, non-decreasing, and not concentrated at zero.

**Theorem** (Schoenberg 1938). Let $\phi : [0, \infty) \to \mathbb{R}$ be a radial kernel and $\Phi(r) = \phi(\sqrt{r})$. Then $\phi$ is positive definite on $\mathbb{R}^d$, for all $d$, if and only if $\Phi$ is completely monotone on $[0, \infty)$ and not constant.

Proof: Use Bernstein-Hausdorff-Widder result and the fact the Gaussian is positive definite.

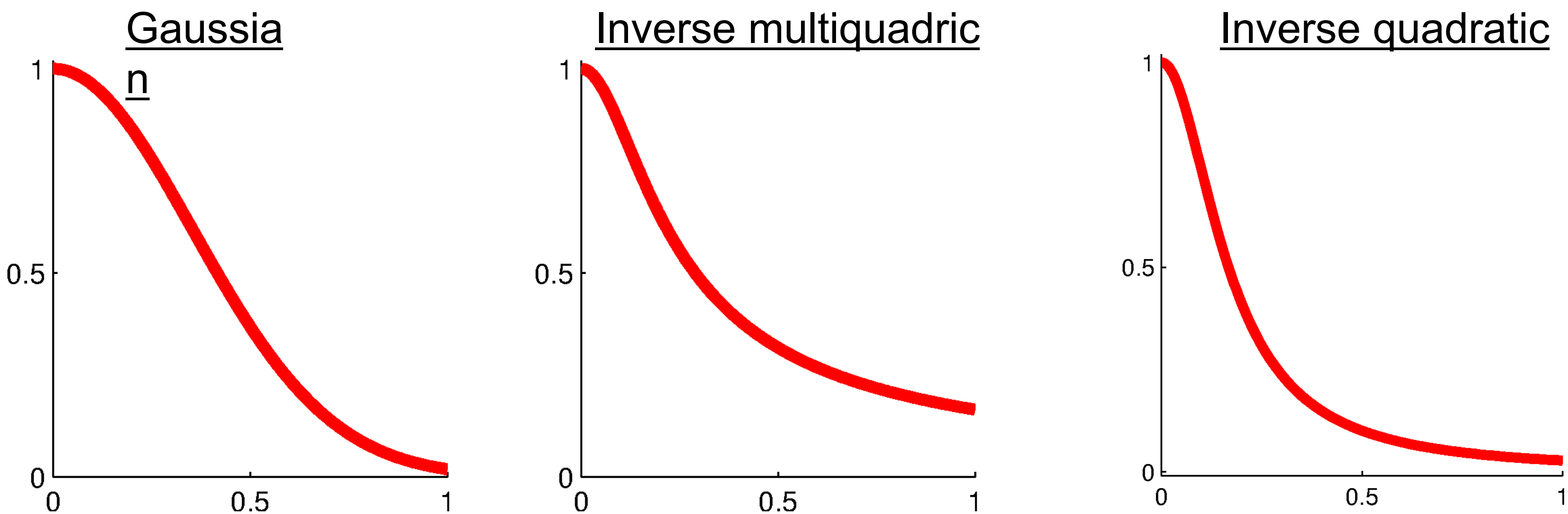
# Positive definite radial kernels



- Some results on positive definite radial kernels.

**Theorem** (Schoenberg 1938). Let $\phi : [0,\infty) \to \mathbb{R}$ be a radial kernel and $\Phi(r) = \phi(\sqrt{r})$. Then $\phi$ is positive definite on $\mathbb{R}^d$, for all $d$, if and only if $\Phi$ is completely monotone on $[0,\infty)$ and not constant.

Examples:




$$\phi(r) = \exp(-(\varepsilon r)^2) \qquad \phi(r) = \frac{1}{\sqrt{1+(\varepsilon r)^2}} \qquad \phi(r) = \frac{1}{1+(\varepsilon r)^2}$$

Here $\varepsilon$ is called the shape parameter (more on this later).

# Positive definite radial kernels



- Results on dimensions specific positive definite radial kernels:

**Theorem** (General kernel)**.** Let $\phi$ be a continuous kernel in $L_1(\mathbb{R}^d)$. Then $\phi$ is positive definite if and only if $\phi$ is bounded and its $d$-dimensional Fourier transform $\hat{\phi}(\boldsymbol{\omega})$ is non-negative and not identically equal to zero.

Remark: Related to Bochner's theorem (1933). Theorem and proof can be found in Wendland (2005).

- To make the result specific to radial kernels, we apply the $d$-dimensional Fourier transform and use radial symmetry to get (Hankel transform):

$$\hat{\phi}(\boldsymbol{\omega}) = \hat{\phi}(\|\boldsymbol{\omega}\|_2) = \frac{1}{\|\boldsymbol{\omega}\|_2^{\nu}} \int_0^{\infty} \phi(t) t^{d/2+1} J_{\nu}(\|\boldsymbol{\omega}\|_2 t) dt,$$

where $\nu = d/2 - 1$ and $J_\nu$ is the $J$-Bessel function of order $\nu$.

- Note that if $\phi$ is positive definite on $\mathbb{R}^d$ then it is positive definite on $\mathbb{R}^k$ for any $k \leq d$.

# Positive definite radial kernels



- Examples

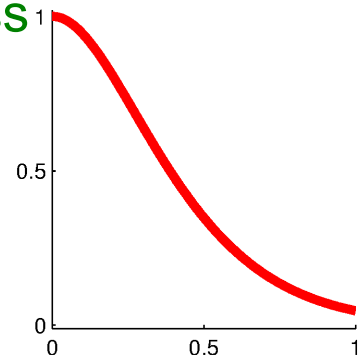
## Finite-smoothness

### Matérn

$$(\varepsilon r)^{\nu-d/2}K_{\nu-d/2}(\varepsilon r)$$

PD for $2\nu > d$

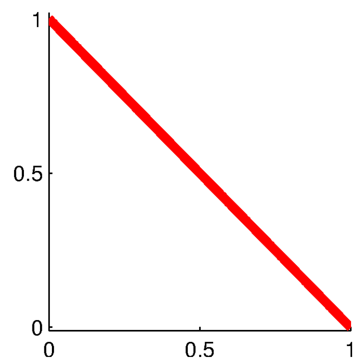
Ex: $e^{-r}(r^2+3r+3)$




### Truncated powers

$$(1-\varepsilon r)_+^{\ell}$$

PD for $\ell \geq \lfloor d/2 \rfloor + 1$

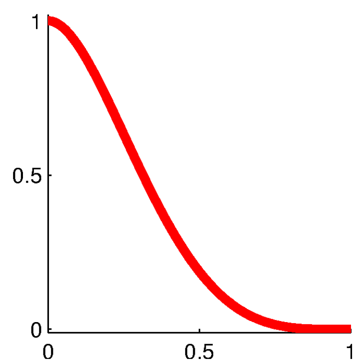



### Wendland (1995)

$$(1-\varepsilon r)_+^{k} p_{d,k}(\varepsilon r)$$

$p_{d,k}$ is a polynomial whose degree depends on $d$ and $k$.

Ex: $(1-\varepsilon r)_+^4(4\varepsilon r+1)$

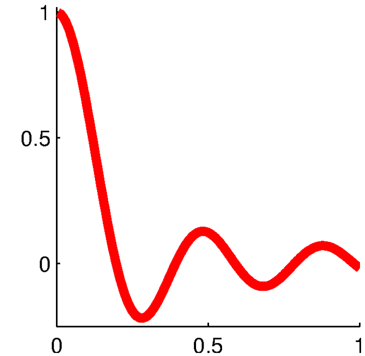



## Infinite-smoothness

### J-Bessel

$$\frac{J_{d/2-1}(\varepsilon r)}{(\varepsilon r)^{d/2}}$$

Ex $(d=3)$: $\dfrac{\sin(\varepsilon r)}{\varepsilon r}$

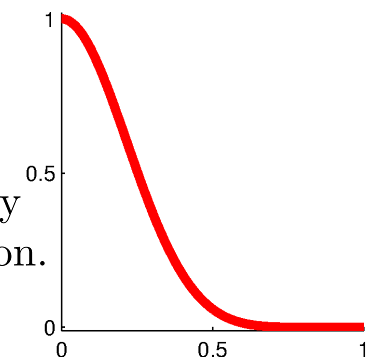



### Platte

$$(\varphi * \varphi)(r)$$

$\varphi$ is a $C^{\infty}(\mathbb{R})$ compactly supported radial function.

PD dimension depends on convolution dimension.

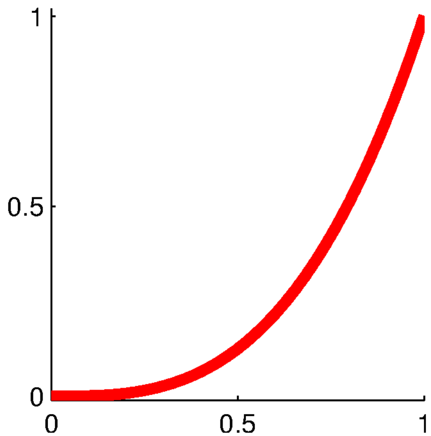


# Conditionally positive definite kernels



- Discussion thus far does not cover many important radial kernels:

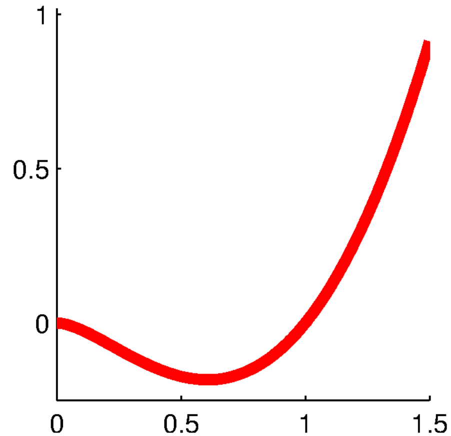
Cubic




$$\phi(r) = r^3$$

Cubic spline in 1-D

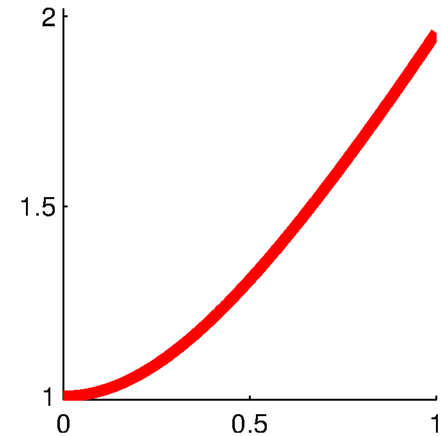
Thin plate spline




$$\phi(r) = r^2 \log r$$

Generalization of energy minimizing spline in 2D

Multiquadric




$$\phi(r) = \sqrt{1 + (\varepsilon r)^2}$$

Popular kernel and first used in any RBF application; Hardy 1971

- These can covered under the theory of conditionally positive definite kernels.

- CPD kernels can be characterized similar to PD kernels but, using generalized Fourier transforms. We will not take this approach; see Ch. 8 Wendland 2005 for details.

# Conditionally positive definite kernels

*Supplementary material*

**Definition.** A continuous kernel $\phi : [0, \infty) \to \mathbb{R}$ is said to be conditionally positive definite of order $k$ on $\mathbb{R}^d$ if, for any distinct $X = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{R}^d$, and all $\mathbf{b} \in \mathbb{R}^N \backslash \{\mathbf{0}\}$ satisfying

$$\sum_{j=1}^{N} b_j p(\mathbf{x}_j) = 0$$

for all $d$-variate polynomials of degree $< k$, the following is satisfied:

$$\sum_{i=1}^{N} \sum_{j=1}^{N} b_i \phi(\|\mathbf{x}_i - \mathbf{x}_j\|) b_j > 0.$$

# Conditionally positive definite kernels



**Definition.** A continuous kernel $\phi : [0,\infty) \to \mathbb{R}$ is said to be conditionally positive definite of order $k$ on $\mathbb{R}^d$ if, for any distinct $X = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{R}^d$, and all $\mathbf{b} \in \mathbb{R}^N \backslash \{\mathbf{0}\}$ satisfying

$$\sum_{j=1}^{N} b_j p(\mathbf{x}_j) = 0$$

for all $d$-variate polynomials of degree $< k$, the following is satisfied:

$$\sum_{i=1}^{N}\sum_{j=1}^{N} b_i \phi(\|\mathbf{x}_i - \mathbf{x}_j\|) b_j > 0.$$

- Alternatively, $\phi$ is positive definite on the subspace $V_{k-1} \subset \mathbb{R}^N$:

$$V_{k-1} = \left\{ \mathbf{b} \in \mathbb{R}^N \left| \sum_{j=1}^{N} b_j p(\mathbf{x}_j) = 0 \text{ for all } p \in \Pi_{k-1}(\mathbb{R}^d) \right. \right\},$$

  where $\Pi_m(\mathbb{R}^d)$ is the space of all $d$-variate polynomials of degree $\leq m$.

- The case $k = 0$, corresponds to standard positive definite kernels on $\mathbb{R}^d$.

# Conditionally positive definite kernels



**Definition.** A continuous kernel $\Phi : [0, \infty) \to \mathbb{R}$ is said to be completely monotone *of order k* on $(0, \infty)$ if $(-1)^k \Phi^{(k)}$ is completely monotone on $(0, \infty)$.

Examples:

*k*=1

$$\Phi(t) = \sqrt{t} \qquad \Phi(t) = \sqrt{1+t}$$

*k*=2

$$\Phi(t) = t^{3/2} \qquad \Phi(t) = \frac{1}{2} t \log t$$

# Conditionally positive definite kernels



**Definition.** A continuous kernel $\Phi : [0,\infty) \to \mathbb{R}$ is said to be completely monotone *of order* $k$ on $(0,\infty)$ if $(-1)^k \Phi^{(k)}$ is completely monotone on $(0,\infty)$.

Examples:

| $k=1$ | | $k=2$ | |
|---|---|---|---|
| $\Phi(t) = \sqrt{t}$ | $\Phi(t) = \sqrt{1+t}$ | $\Phi(t) = t^{3/2}$ | $\Phi(t) = \frac{1}{2} t \log t$ |

**Theorem** (Micchelli (1986); Guo, Hu, & Sun (1993)). The radial kernel $\phi : [0,\infty)$ is conditionally positive definite on $\mathbb{R}^d$, for all $d$, if and only if $\Phi = \phi(\sqrt{\cdot})$ is completely monotone of order $k$ on $(0,\infty)$ and $\Phi^{(k)}$ is not constant.

Remark:

- This is one of the BIG theorems that launched the RBF field.
- It says, for example, that linear, cubic, thin-plate splines, and the multiquadric are conditionally positive definite on $\mathbb{R}^d$ for any $d$.
- Next, its consequences on RBF interpolation of scattered data...

# Conditionally positive definite kernels



**Definition.** Let $\phi : [0, \infty) \to \mathbb{R}$ be continuous and $\{p_i(\mathbf{x})\}_{i=1}^{n}$ be a basis for $\Pi_{k-1}(\mathbb{R}^d)$ $(k > 1)$. The general RBF interpolant for the distinct nodes $X = \{\mathbf{x}_j\}_{j=1}^{N} \subset \mathbb{R}^d$ and some target, $f$, sampled on $X$, $\{f_j\}_{j=1}^{N}$ is

$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \phi(\|\mathbf{x} - \mathbf{x}_j\|) + \sum_{\ell=1}^{n} d_\ell p_\ell(\mathbf{x}),$$

where $I_X f(\mathbf{x}_i) = f_i$, $i = 1, \ldots, N$ and $\sum_{j=1}^{N} c_j p_\ell(\mathbf{x}_j) = 0$, $\ell = 1, \ldots, n$.

In linear system form, these constraints are

$$\begin{bmatrix} A & P \\ P^T & 0 \end{bmatrix} \begin{bmatrix} \underline{c} \\ \underline{d} \end{bmatrix} = \begin{bmatrix} \underline{f} \\ \underline{0} \end{bmatrix}, \text{ where } a_{i,j} = \phi(\|\mathbf{x}_i - \mathbf{x}_j\|),\ p_{i,\ell} = p_k(\mathbf{x}_i)$$

# Conditionally positive definite kernels



**Definition.** Let $\phi : [0,\infty) \to \mathbb{R}$ be continuous and $\{p_i(\mathbf{x})\}_{i=1}^{n}$ be a basis for $\Pi_{k-1}(\mathbb{R}^d)$ $(k > 1)$. The general RBF interpolant for the distinct nodes $X = \{\mathbf{x}_j\}_{j=1}^{N} \subset \mathbb{R}^d$ and some target, $f$, sampled on $X$, $\{f_j\}_{j=1}^{N}$ is

$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \phi(\|\mathbf{x} - \mathbf{x}_j\|) + \sum_{\ell=1}^{n} d_\ell p_\ell(\mathbf{x}),$$

where $I_X f(\mathbf{x}_i) = f_i$, $i = 1, \ldots, N$ and $\sum_{j=1}^{N} c_j p_\ell(\mathbf{x}_j) = 0$, $\ell = 1, \ldots, n$.

In linear system form, these constraints are

$$\begin{bmatrix} A & P \\ P^T & 0 \end{bmatrix} \begin{bmatrix} \underline{c} \\ \underline{d} \end{bmatrix} = \begin{bmatrix} \underline{f} \\ \underline{0} \end{bmatrix}, \text{ where } a_{i,j} = \phi(\|\mathbf{x}_i - \mathbf{x}_j\|),\ p_{i,\ell} = p_k(\mathbf{x}_i)$$

**Theorem** (Micchelli (1986)). The above linear system is invertible for any distinct $X$, provided

- $\text{rank}(P) = n$ (i.e. $X$ is unisolvent on $\Pi_{k-1}(\mathbb{R}^d)$),
- $\Phi = \phi(\sqrt{\cdot})$ is completely monotone of order $k$ on $(0,\infty)$,
- $\Phi^{(k)}$ is not constant.

# Conditionally positive definite kernels



**Definition.** Let $\phi : [0,\infty) \to \mathbb{R}$ be continuous and $\{p_i(\mathbf{x})\}_{i=1}^{n}$ be a basis for $\Pi_{k-1}(\mathbb{R}^d)$ $(k > 1)$. The general RBF interpolant for the distinct nodes $X = \{\mathbf{x}_j\}_{j=1}^{N} \subset \mathbb{R}^d$ and some target, $f$, sampled on $X$, $\{f_j\}_{j=1}^{N}$ is

$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \phi(\|\mathbf{x}-\mathbf{x}_j\|) + \sum_{\ell=1}^{n} d_\ell p_\ell(\mathbf{x}),$$

where $I_X f(\mathbf{x}_i) = f_i$, $i = 1,\ldots,N$ and $\sum_{j=1}^{N} c_j p_\ell(\mathbf{x}_j) = 0$, $\ell = 1,\ldots,n$.

In linear system form, these constraints are

$$\begin{bmatrix} A & P \\ P^T & 0 \end{bmatrix} \begin{bmatrix} \underline{c} \\ \underline{d} \end{bmatrix} = \begin{bmatrix} \underline{f} \\ \underline{0} \end{bmatrix}, \text{ where } a_{i,j} = \phi(\|\mathbf{x}_i - \mathbf{x}_j\|),\ p_{i,\ell} = p_k(\mathbf{x}_i)$$

**Example** (Thin plate spline, $\mathbb{R}^2$). Let

- $\phi(r) = r^2 \log(r)$
- $p_1(x,y) = 1$, $p_2(x,y) = x$, and $p_3(x,y) = y$.

The system has a unique solution provided the nodes are not collinear.

# Conditionally positive definite kernels



**Theorem** (Micchelli (1986)). Suppose $\Phi = \phi(\sqrt{\cdot})$ is completely monotone of order 1 on $(0,\infty)$ and $\Phi'$ is not constant. Then for any distinct set of nodes $X = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{R}^d$, and any $d$, the matrix $A$ with entries $a_{i,j} = \phi(\|\mathbf{x}_i - \mathbf{x}_j\|)$, $i,j = 1,\ldots,N$, has $N-1$ positive eigenvalues and 1 negative eigenvalue. Hence it is invertible.

Remark:

- This theorem means that for kernels like the popular multiquadric $\phi(r) = \sqrt{1+(\varepsilon r)^2}$ the basic RBF interpolant

$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \phi(\|\mathbf{x} - \mathbf{x}_j\|)$$

  has a unique solution for any distinct set of nodes $X = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{R}^d$ and sampled target function $f$ on $X$.

- Augmenting the RBF interpolant with polynomials is not necessary to guarantee uniqueness for order 1 CPD kernels.

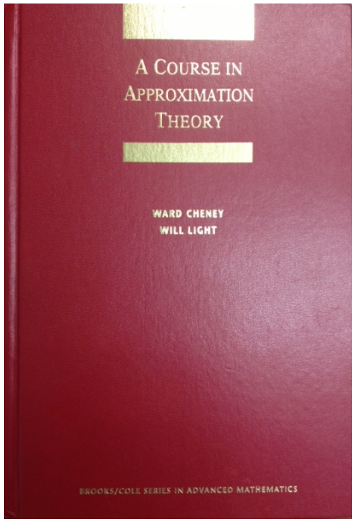
- This theorem answered a conjecture from Franke (1983) regarding the multiquadric.

# Radial basis function (RBF) interpolation

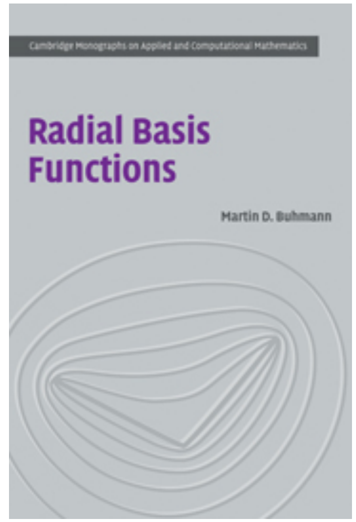


- Many good books to consult further on RBF theory and applications:

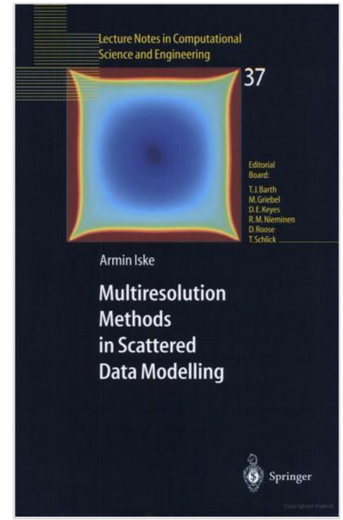
1999




2003

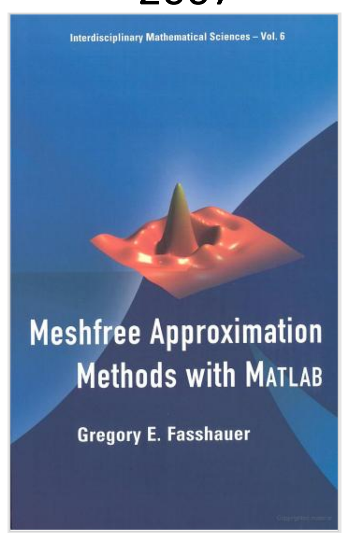



2004




2005


2007




2014: SIAM

A Primer on Radial Basis Functions with Applications to the Geosciences

Bengt Fornberg
Natasha Flyer

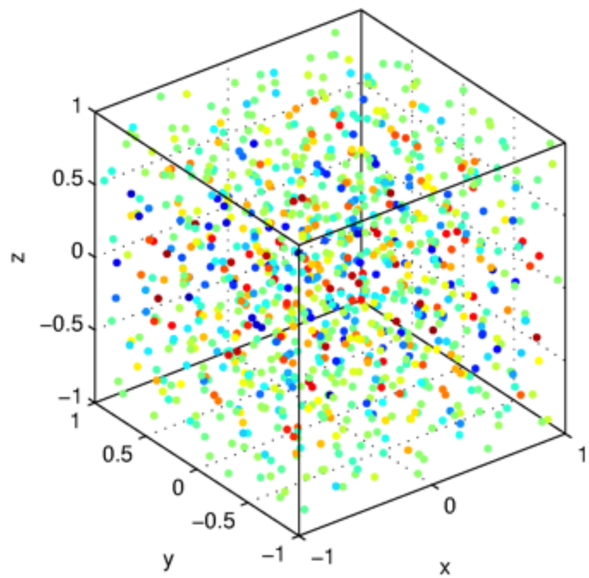
# Interpolation with kernels (revisited)

*Supplementary material*

- Kernel interpolant to $f\big|_X$: $\quad I_X f = \sum_j c_j \Phi(\cdot, \mathbf{x}_j)$.
- Some considerations for choosing the kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$

  1. The kernel should be easy to compute.
  2. The kernel interpolant should be uniquely determined by $X$ and $f\big|_X$.
  3. The kernel interpolant should accurately reconstruct $f$.

- For problems like




$\Omega = [-1, 1]^3$

Obvious choice: $\phi$ is a (conditionally) positive definite radial kernel

$$\Phi(\mathbf{x}, \mathbf{x}_j) = \phi(\|\mathbf{x} - \mathbf{x}_j\|_2) = \phi(r)$$

- Leads to radial basis function (RBF) interpolation.

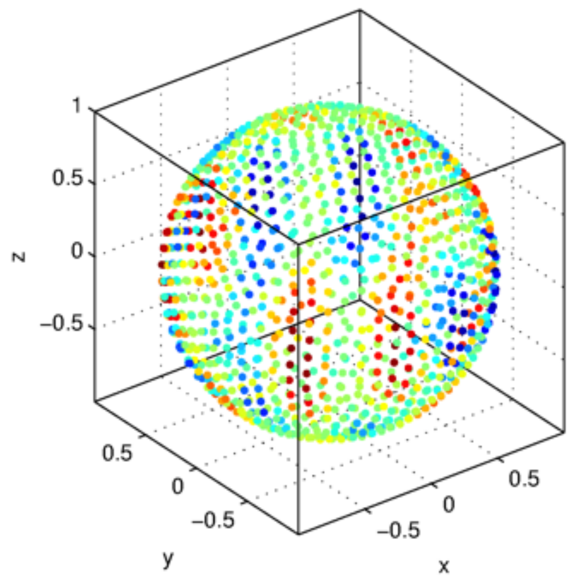
# Interpolation with kernels on the sphere



- Kernel interpolant to $f\big|_X$: $\quad I_X f = \sum_j c_j \Phi(\cdot, \mathbf{x}_j)$.
- Some considerations for choosing the kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$

  1. The kernel should be easy to compute.
  2. The kernel interpolant should be uniquely determined by $X$ and $f\big|_X$.
  3. The kernel interpolant should accurately reconstruct $f$.

- For problems like




$\Omega = \mathbb{S}^2$

Obvious(?) choice: $\Phi$ is a (conditionally) positive definite zonal kernel:

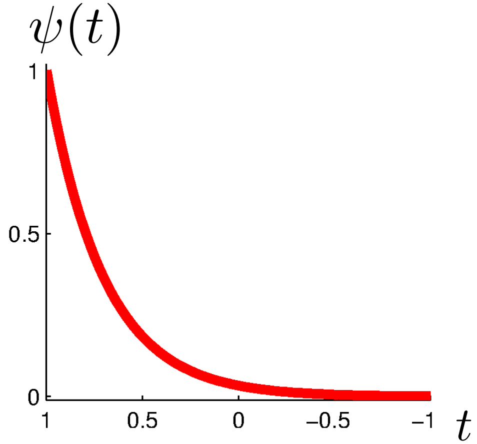
$$\Phi(\mathbf{x}, \mathbf{x}_j) = \psi(\mathbf{x}^T \mathbf{x}_j) = \psi(t), \; t \in [-1, 1]$$

- Analog of RBF interpolation for the sphere: SBF interpolation.

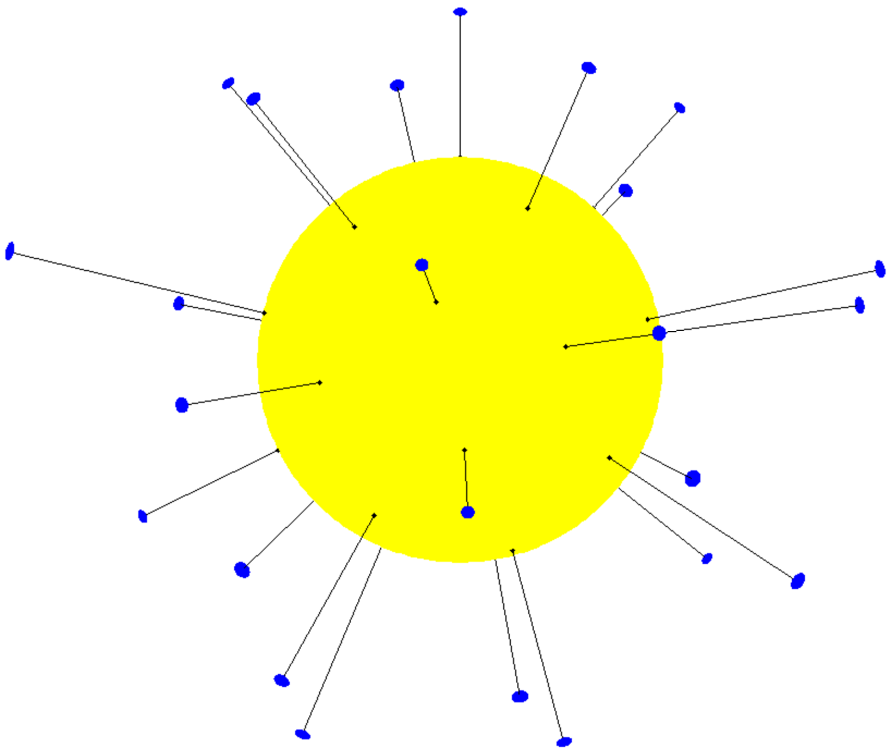
# SBF interpolation



Key idea: linear combination of translates and rotations of a single zonal kernel on $\mathbb{S}^2$




$$X = \{\mathbf{x}_j\}_{j=1}^N \subset \Omega, \quad f\Big|_X = \{f_j\}_{j=1}^N$$



Basic SBF Interpolant for $\mathbb{S}^2$

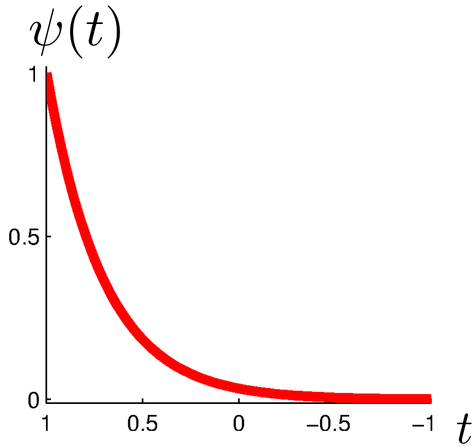
$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \psi(\mathbf{x}^T \mathbf{x}_j)$$

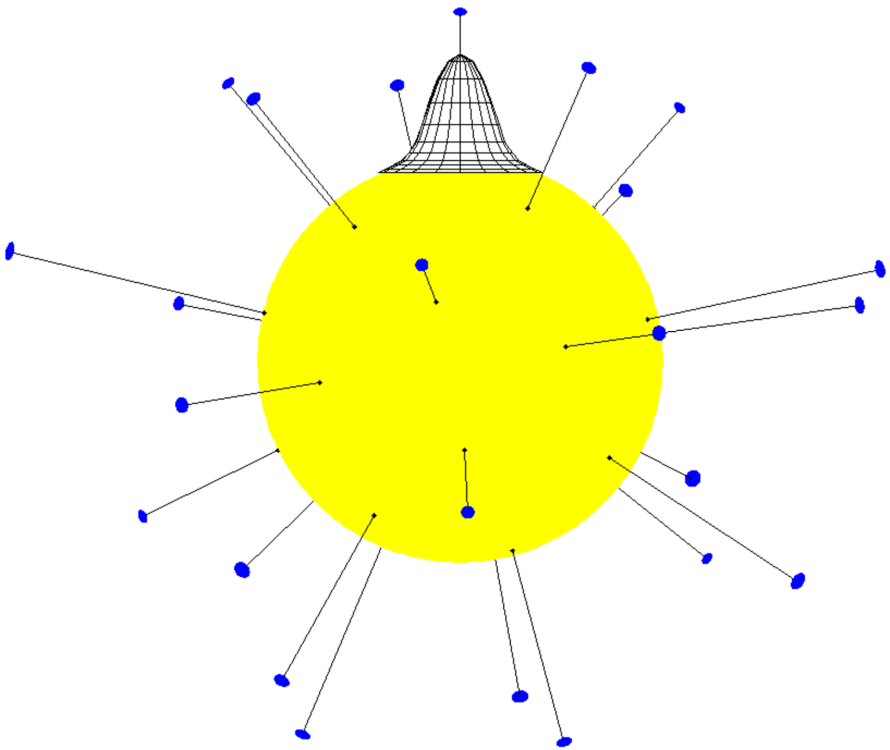
# SBF interpolation



Key idea: linear combination of translates and rotations of a single zonal kernel on $\mathbb{S}^2$

$$X = \{\mathbf{x}_j\}_{j=1}^N \subset \Omega, \quad f\Big|_X = \{f_j\}_{j=1}^N$$



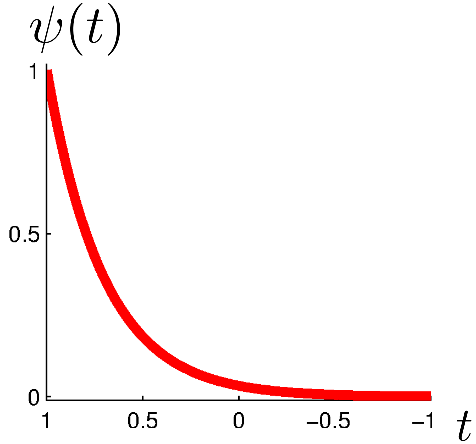
Basic SBF Interpolant for $\mathbb{S}^2$

$$I_X f(\mathbf{x}) = \sum_{j=1}^N c_j \psi(\mathbf{x}^T \mathbf{x}_j)$$

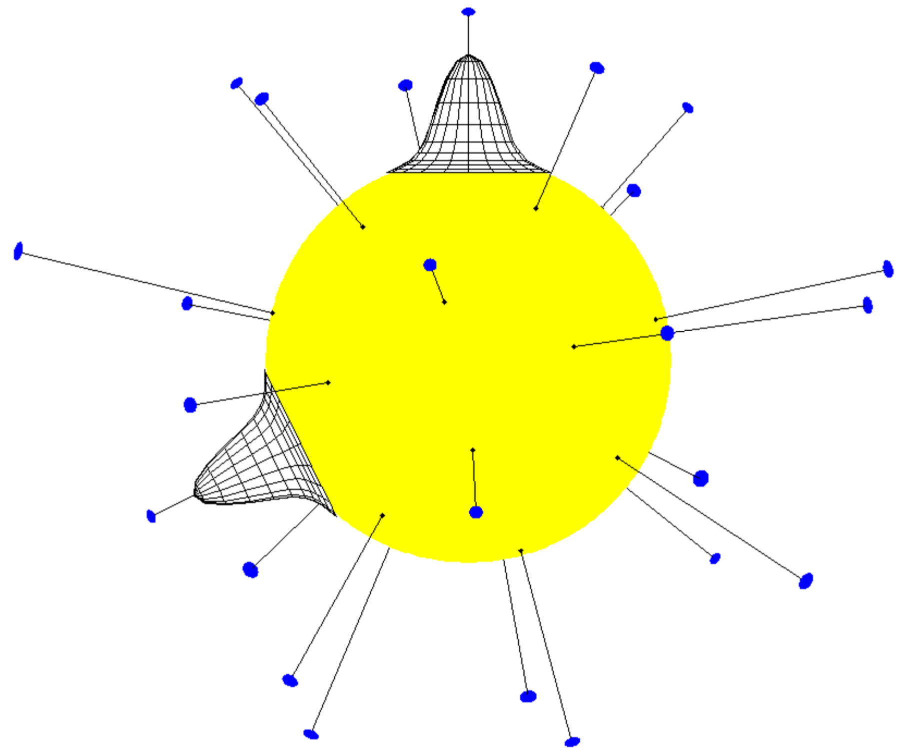
# SBF interpolation



Key idea: linear combination of translates and rotations of a single zonal kernel on $\mathbb{S}^2$




$$X = \{\mathbf{x}_j\}_{j=1}^N \subset \Omega, \quad f\Big|_X = \{f_j\}_{j=1}^N$$



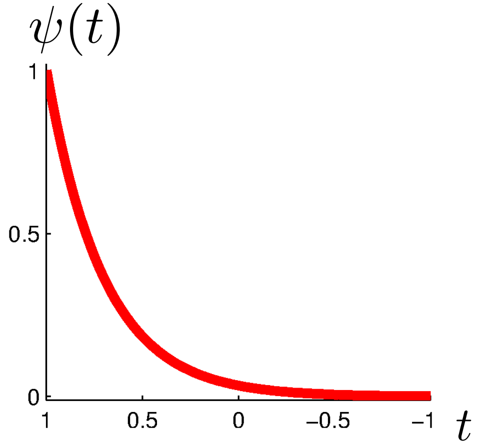
Basic SBF Interpolant for $\mathbb{S}^2$

$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \psi(\mathbf{x}^T \mathbf{x}_j)$$

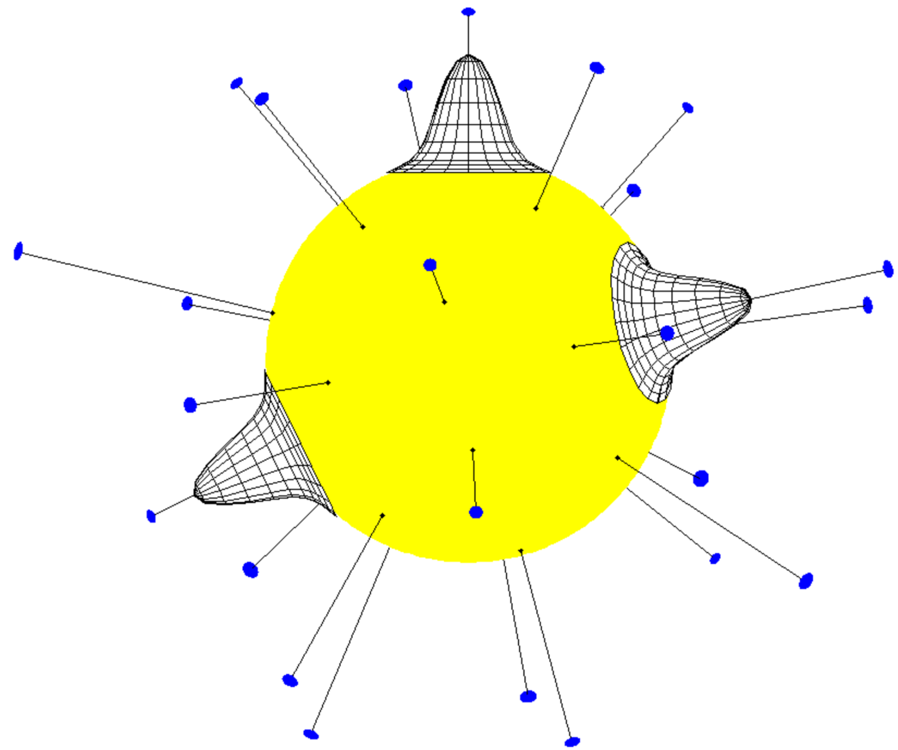
# SBF interpolation



<u>Key idea</u>: linear combination of translates and rotations of a single zonal kernel on $\mathbb{S}^2$




$$X = \{\mathbf{x}_j\}_{j=1}^N \subset \Omega, \quad f\Big|_X = \{f_j\}_{j=1}^N$$



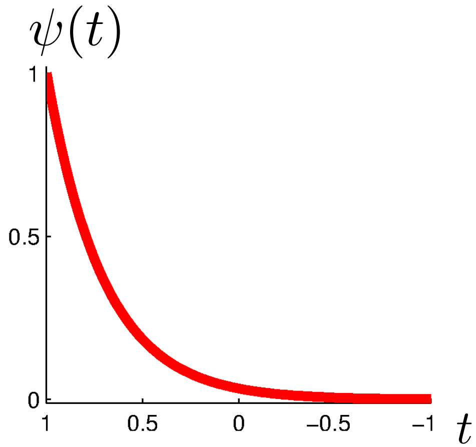
<u>Basic SBF Interpolant for $\mathbb{S}^2$</u>

$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \psi(\mathbf{x}^T \mathbf{x}_j)$$

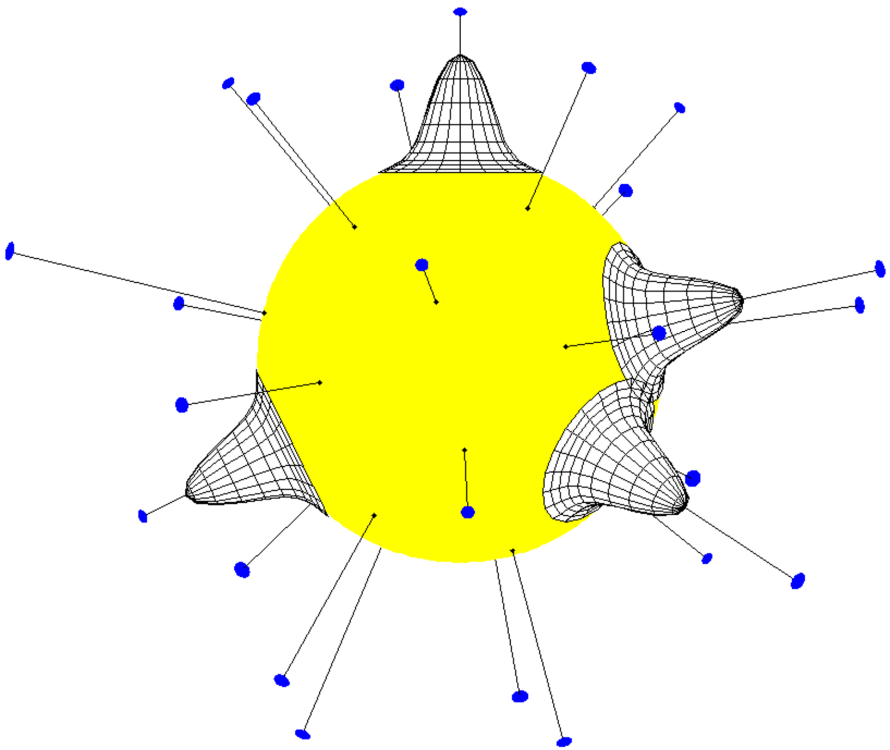
# SBF interpolation



Key idea: linear combination of translates and rotations of a single zonal kernel on $\mathbb{S}^2$




$$X = \{\mathbf{x}_j\}_{j=1}^N \subset \Omega, \quad f\Big|_X = \{f_j\}_{j=1}^N$$



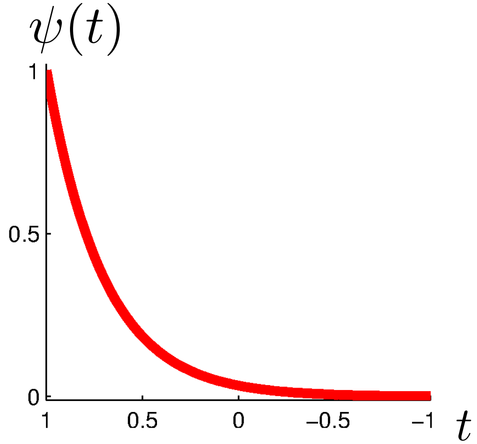
Basic SBF Interpolant for $\mathbb{S}^2$

$$I_X f(\mathbf{x}) = \sum_{j=1}^N c_j \psi(\mathbf{x}^T \mathbf{x}_j)$$

# SBF interpolation



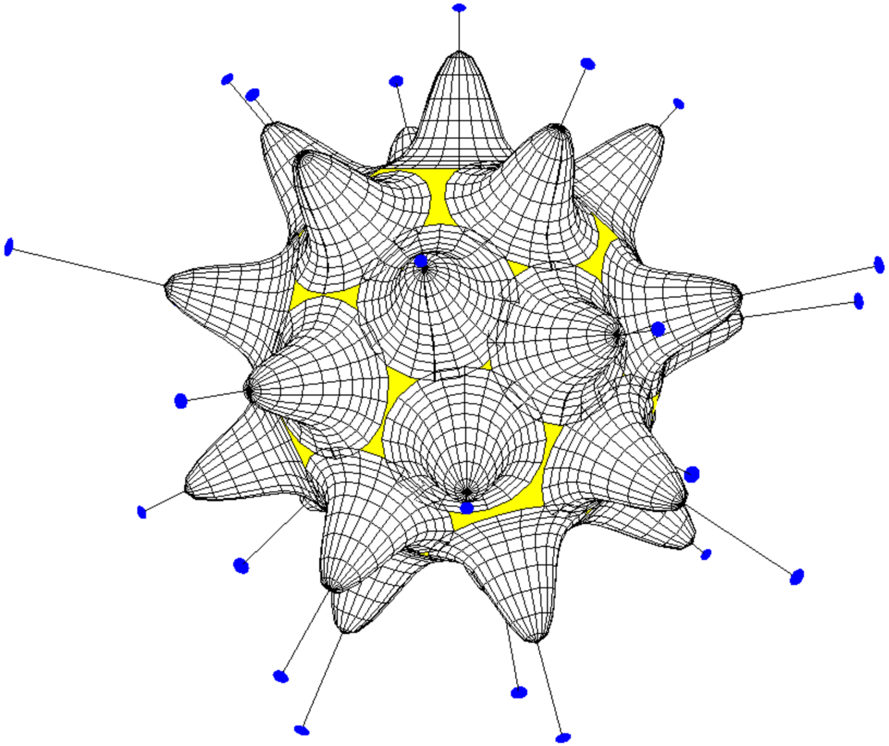
Key idea: linear combination of translates and rotations of a single zonal kernel on $\mathbb{S}^2$




Basic SBF Interpolant for $\mathbb{S}^2$

$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \psi(\mathbf{x}^T \mathbf{x}_j)$$

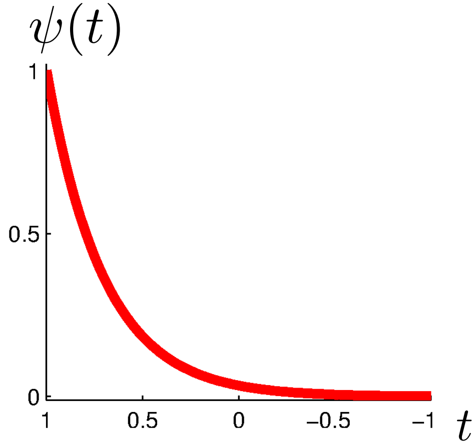
$X = \{\mathbf{x}_j\}_{j=1}^N \subset \Omega, \quad f\Big|_X = \{f_j\}_{j=1}^N$



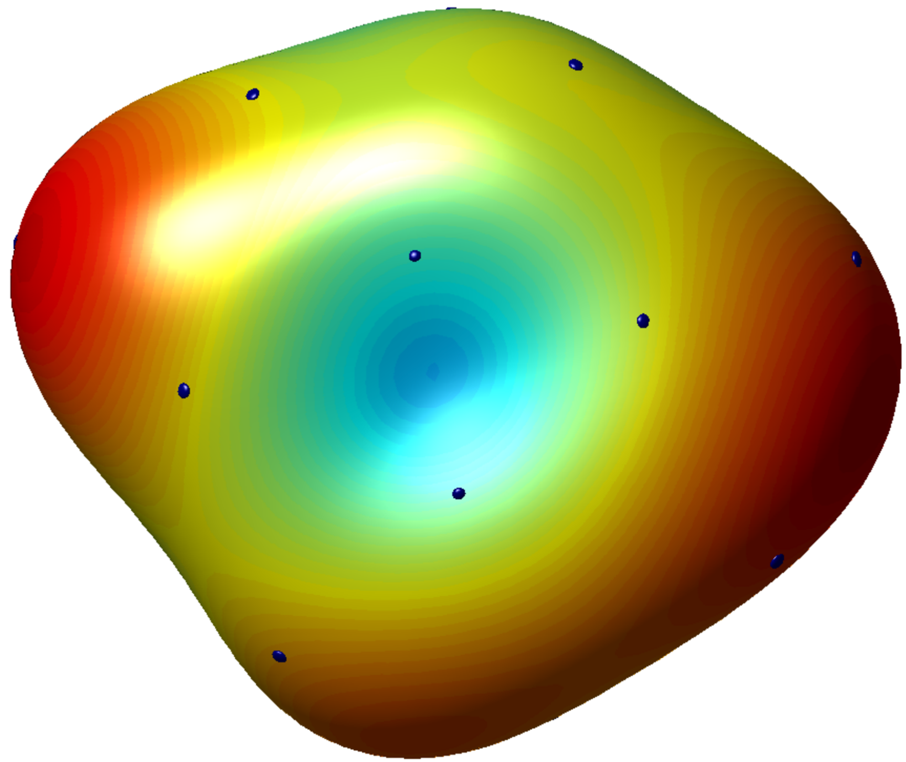
# SBF interpolation



Key idea: linear combination of translates and rotations of a single zonal kernel on $\mathbb{S}^2$




$X = \{\mathbf{x}_j\}_{j=1}^N \subset \Omega, \quad f\big|_X = \{f_j\}_{j=1}^N$



Basic SBF Interpolant for $\mathbb{S}^2$

$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \psi(\mathbf{x}^T \mathbf{x}_j)$$

Linear system for determining the interpolation coefficients

$$\underbrace{\begin{bmatrix} \psi(\mathbf{x}_1^T\mathbf{x}_1) & \psi(\mathbf{x}_1^T\mathbf{x}_2) \cdots \psi(\mathbf{x}_1^T\mathbf{x}_N) \\ \psi(\mathbf{x}_2^T\mathbf{x}_1) & \psi(\mathbf{x}_2^T\mathbf{x}_2) \cdots \psi(\mathbf{x}_2^T\mathbf{x}_N) \\ \vdots & \vdots \quad \ddots \quad \vdots \\ \psi(\mathbf{x}_N^T\mathbf{x}_1) & \psi(\mathbf{x}_N^T\mathbf{x}_2) \cdots \psi(\mathbf{x}_N^T\mathbf{x}_N) \end{bmatrix}}_{A_X} \underbrace{\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix}}_{\underline{c}} = \underbrace{\begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_N \end{bmatrix}}_{\underline{f}}$$

$A_X$ is guaranteed to be positive definite if $\psi$ is a positive definite zonal kernel

# Positive definite zonal kernels



**Definition.** A kernel $\Psi : \mathbb{S}^{d-1} \times \mathbb{S}^{d-1} \rightarrow \mathbb{R}$ is called radial or zonal on $\mathbb{S}^{d-1}$ if $\Psi(\mathbf{x}, \mathbf{y}) = \psi(\mathbf{x}^T\mathbf{y})$, where $\psi : [-1, 1] \rightarrow \mathbb{R}$. In this case, $\psi$ is simply referred to as the zonal kernel and no reference is made to $\Psi$.

# Positive definite zonal kernels



**Definition.** A kernel $\Psi : \mathbb{S}^{d-1} \times \mathbb{S}^{d-1} \to \mathbb{R}$ is called radial or zonal on $\mathbb{S}^{d-1}$ if $\Psi(\mathbf{x}, \mathbf{y}) = \psi(\mathbf{x}^T\mathbf{y})$, where $\psi : [-1, 1] \to \mathbb{R}$. In this case, $\psi$ is simply referred to as the zonal kernel and no reference is made to $\Psi$.

**Definition.** A zonal kernel $\psi : [-1, 1] \to \mathbb{R}$ is said to be a positive definite zonal kernel on $\mathbb{S}^{d-1}$ if for any distinct set of nodes $X = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{S}^{d-1}$ and $\underline{b} \in \mathbb{R}^N \backslash \{0\}$ the matrix $A = \{\psi(\mathbf{x}_i^T\mathbf{x}_j)\}$ is positive definite, i.e.

$$\sum_{i=1}^{N} \sum_{j=1}^{N} b_i \psi(\mathbf{x}_i^T \mathbf{x}_j) b_j > 0.$$

Remark: PD zonal kernels are sometimes called spherical basis functions (SBFs).

# Positive definite zonal kernels



**Definition.** A kernel $\Psi : \mathbb{S}^{d-1} \times \mathbb{S}^{d-1} \to \mathbb{R}$ is called radial or zonal on $\mathbb{S}^{d-1}$ if $\Psi(\mathbf{x}, \mathbf{y}) = \psi(\mathbf{x}^T\mathbf{y})$, where $\psi : [-1, 1] \to \mathbb{R}$. In this case, $\psi$ is simply referred to as the zonal kernel and no reference is made to $\Psi$.

**Definition.** A zonal kernel $\psi : [-1, 1] \to \mathbb{R}$ is said to be a positive definite zonal kernel on $\mathbb{S}^{d-1}$ if for any distinct set of nodes $X = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{S}^{d-1}$ and $\underline{b} \in \mathbb{R}^N \backslash \{0\}$ the matrix $A = \{\psi(\mathbf{x}_i^T\mathbf{x}_j)\}$ is positive definite, i.e.

$$\sum_{i=1}^{N}\sum_{j=1}^{N} b_i\psi(\mathbf{x}_i^T\mathbf{x}_j)b_j > 0.$$

Remark: PD zonal kernels are sometimes called spherical basis functions (SBFs).

- The study of positive definite kernels on $\mathbb{S}^{d-1}$ started with Schoenberg (1940).
- Extension of this work, including to conditionally positive definite kernels, began in the 1990s (Cheney and Xu (1992)), and continues today.
- Our interest is strictly in $\mathbb{S}^2$ and we will only present results for this case.

# Positive definite zonal kernels



- Any positive definite radial kernel $\phi$ on $\mathbb{R}^3$ is also positive definite on $\mathbb{S}^2$.
- In fact, they are positive definite zonal kernels, since for $\mathbf{x}, \mathbf{y} \in \mathbb{S}^2$

$$\phi(\|\mathbf{x}-\mathbf{y}\|) = \phi\left(\sqrt{2 - 2\mathbf{x}^T\mathbf{y}}\right) = \psi(\mathbf{x}^T\mathbf{y})$$

- So, standard RBF methods can be used for problems on the sphere $\mathbb{S}^2$.
- Cheney (1995) appears to have been the first to mathematically study the specialization of RBFs to the sphere.
- Many others have followed suit, e.g.

  Fasshauer & Schumaker (1998); Baxter & Hubbert (2001); Levesley & Hubbert (2001); Hubbert & Morton (2004); zu Castel & Filbir (2005); Narcowich, Sun, & Ward (2007); Narcowich, Sun, Ward, & Wendland (2007); Fornberg & Piret (2007); Narcowich, Ward, & W (2007); Fuselier, Narcowich, Ward, & W (2009); Fuselier & W (2009)

# Positive definite zonal kernels



- Any positive definite radial kernel $\phi$ on $\mathbb{R}^3$ is also positive definite on $\mathbb{S}^2$.
- In fact, they are positive definite zonal kernels, since for $\mathbf{x}, \mathbf{y} \in \mathbb{S}^2$

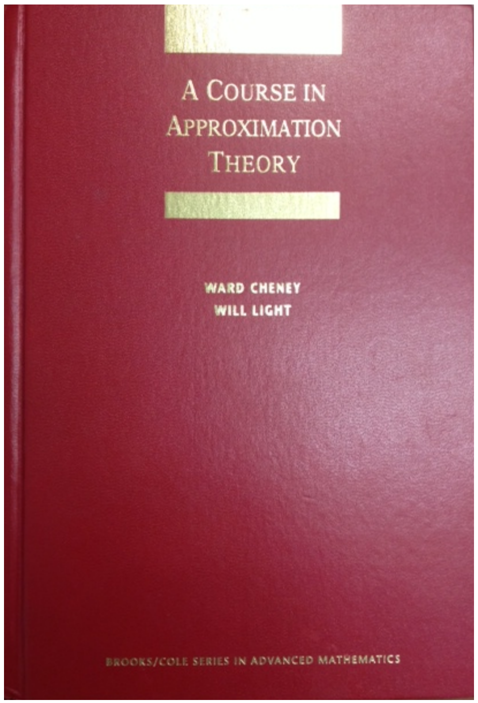
$$\phi(\|\mathbf{x}-\mathbf{y}\|) = \phi\left(\sqrt{2-2\mathbf{x}^T\mathbf{y}}\right) = \psi(\mathbf{x}^T\mathbf{y})$$

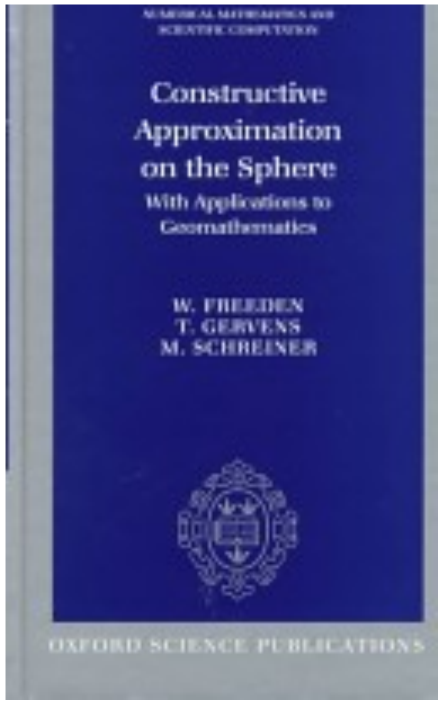
- So, standard RBF methods can be used for problems on the sphere $\mathbb{S}^2$.
- Cheney (1995) appears to have been the first to mathematically study the specialization of RBFs to the sphere.
- Many others have followed suit, e.g.
  Fasshauer & Schumaker (1998); Baxter & Hubbert (2001); Levesley & Hubbert (2001); Hubbert & Morton (2004); zu Castel & Filbir (2005); Narcowich, Sun, & Ward (2007); Narcowich, Sun, Ward, & Wendland (2007); Fornberg & Piret (2007); Narcowich, Ward, & W (2007); Fuselier, Narcowich, Ward, & W (2009); Fuselier & W (2009)
- Open question: Are there any advantages to using a purely PD zonal kernel to a restricted PD radial kernel? (Baxter & Hubbert (2001))
- Personally, I have always used restricted radial kernels.

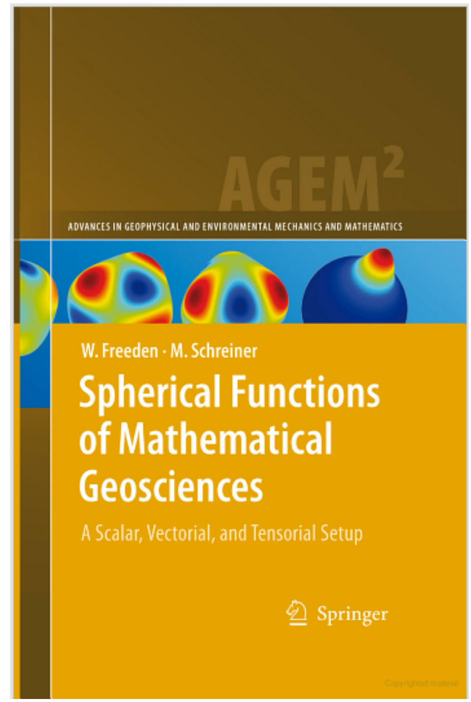
# Positive definite zonal kernels

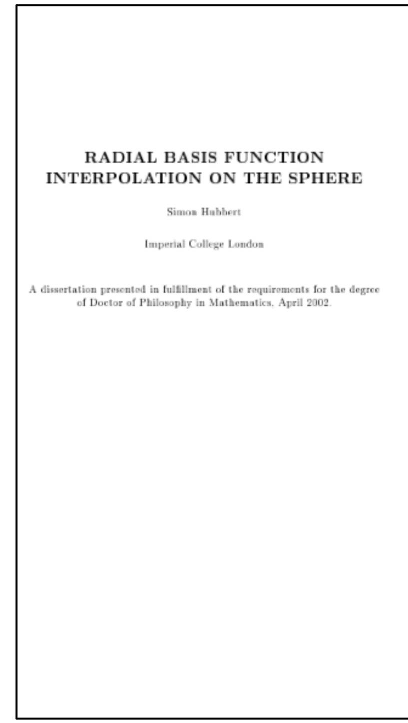
*Supplementary material*

- Some references for the material to come:





RADIAL BASIS FUNCTION INTERPOLATION ON THE SPHERE

Simon Hubbert

Imperial College London

A dissertation presented in fulfillment of the requirements for the degree of Doctor of Philosophy in Mathematics. April 2002.

# Spherical harmonics



- A good understanding of functions on the sphere requires one to be well-versed in spherical harmonics.
- Spherical harmonics are the analog of 1-D Fourier series for approximation on spheres of dimension 2 and higher.
- Several ways to introduce spherical harmonics (Freeden & Schreiner 2008)
- We will use the eigenfunction approach and restrict our attention to the 2-sphere.
- Following this we review some important results about spherical harmonics.

# Overview of spherical harmonics

*Supplementary material*

- Laplacian in spherical coordinates ($x = r\cos\theta\cos\varphi,\ y = r\cos\theta\sin\varphi,\ z = r\sin\theta$)

$$\Delta = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\underbrace{\left\{\frac{\partial^2}{\partial\theta^2} - \tan\theta\frac{\partial}{\partial\theta} + \frac{1}{\cos^2\theta}\frac{\partial^2}{\partial\varphi^2}\right\}}_{\Delta_s = \text{Laplace-Beltrami operator}}$$

# Overview of spherical harmonics



- Laplacian in spherical coordinates ($x = r\cos\theta\cos\varphi$, $y = r\cos\theta\sin\varphi$, $z = r\sin\theta$)

$$\Delta = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\underbrace{\left\{\frac{\partial^2}{\partial\theta^2} - \tan\theta\frac{\partial}{\partial\theta} + \frac{1}{\cos^2\theta}\frac{\partial^2}{\partial\varphi^2}\right\}}_{\Delta_s = \text{Laplace-Beltrami operator}}$$

- Spherical harmonics: Set of all functions bounded at $\theta = \pm\frac{\pi}{2}$ or $z = \pm 1$ such that $\Delta_s Y = \lambda Y$.

- Solve using separation of variables to arrive at:

$$Y_\ell^m(\theta,\varphi) = a_\ell^{|m|}P_\ell^{|m|}(\cos\theta)e^{im\varphi},\ \ell = 0,1,\ldots,\ \ m = -\ell,-\ell+1,\ldots,\ell-1,\ell.$$

- Here $P_\ell^k$, for $k = 0,1,\ldots$, $\ell = k,k+1,\ldots$, are the Associated Legendre functions, given by Rodrigues' formula

$$P_\ell^k(z) = (1-z^2)^{k/2}\frac{d^k}{dz^k}\left(P_\ell(z)\right),$$

where $P_\ell$ is the standard Legendre polynomial of degree $\ell$.

- The $a_\ell^k$ are normalization factors (e.g. $a_\ell^k = \sqrt{((2\ell+1)(\ell-k)!)/(4\pi(\ell+m)!})$

# Overview of spherical harmonics

*Supplementary material*

- Each spherical harmonic satisfies $\Delta_s Y_\ell^m = -\ell(\ell+1) Y_\ell^m$.
- For each $\ell = 0, 1, \ldots$, there are $2\ell+1$ harmonics with eigenvalue $-\ell(\ell+1)$.

# Overview of spherical harmonics



- Each spherical harmonic satisfies $\Delta_s Y_\ell^m = -\ell(\ell+1)Y_\ell^m$.
- For each $\ell = 0, 1, \ldots,$ there are $2\ell+1$ harmonics with eigenvalue $-\ell(\ell+1)$.
- Real-form of spherical harmonics:

$$Y_\ell^m(\theta,\varphi) = Y_\ell^m(z,\varphi) = \begin{cases} \sqrt{2} a_\ell^m P_\ell^m(z)\cos(m\varphi) & m > 0, \\ a_\ell^0 P_\ell(z) & m = 0, \\ \sqrt{2} a_\ell^{|m|} P_\ell^{|m|}(z)\sin(m\varphi) & m < 0. \end{cases}$$

# Overview of spherical harmonics



- Each spherical harmonic satisfies $\Delta_s Y_\ell^m = -\ell(\ell+1) Y_\ell^m$.
- For each $\ell = 0, 1, \ldots$, there are $2\ell+1$ harmonics with eigenvalue $-\ell(\ell+1)$.
- Real-form of spherical harmonics:

$$Y_\ell^m(\theta,\varphi) = Y_\ell^m(z,\varphi) = \begin{cases} \sqrt{2} a_\ell^m P_\ell^m(z)\cos(m\varphi) & m > 0, \\ a_\ell^0 P_\ell(z) & m = 0, \\ \sqrt{2} a_\ell^{|m|} P_\ell^{|m|}(z)\sin(m\varphi) & m < 0. \end{cases}$$

- Can also be expressed purely in Cartesian coordinates ($\mathbf{x} = (x,y,z) \in \mathbb{S}^2$):

$$Y_\ell^m(\mathbf{x}) = Y_\ell^m(x,y,z) = \begin{cases} \sqrt{2} a_\ell^m Q_\ell^m(z)\frac{1}{2}\left((x+iy)^m + (x-iy)^m\right) & m > 0, \\ a_\ell^0 P_\ell(z) & m = 0, \\ \sqrt{2} a_\ell^{|m|} Q_\ell^{|m|}(z)\frac{1}{2i}\left((x+iy)^{-m} - (x-iy)^{-m}\right) & m < 0. \end{cases}$$

where $Q_\ell^m(z) = (-1)^m \dfrac{\partial^m}{\partial z^m} P_\ell(z)$.

- We will sometimes switch notation from $Y_\ell^m(\theta,\varphi)$ to $Y_\ell^m(\mathbf{x})$.

# Overview of spherical harmonics

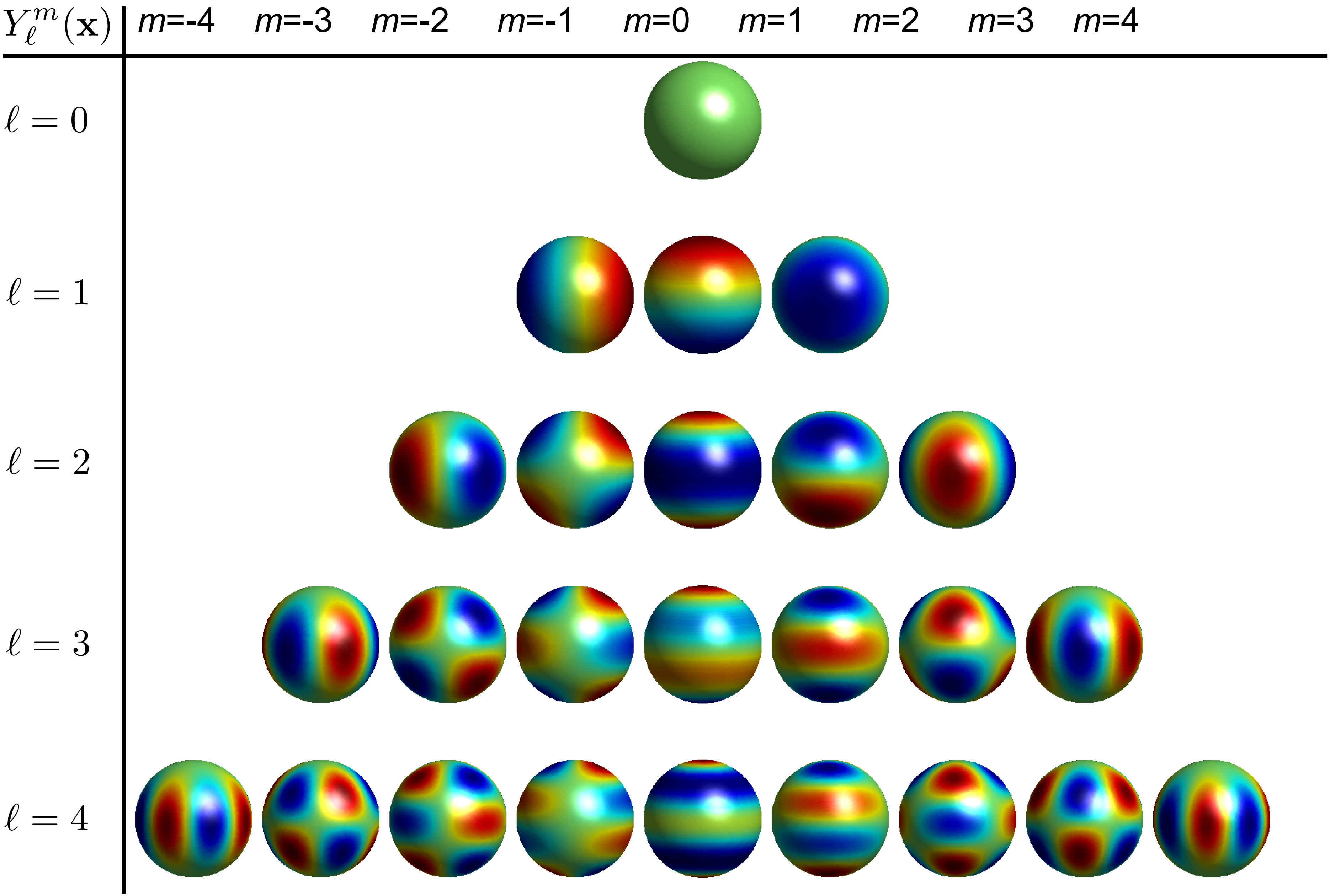
*Supplementary material*

- Spherical harmonics $Y_\ell^m(\mathbf{x})$ in Cartesian form, for $\ell = 0, 1, 2, 3$.

| *m*=-3 | *m*=-2 | *m*=-1 | *m*=0 | *m*=1 | *m*=2 | *m*=3 |
|---|---|---|---|---|---|---|
| | | | $\frac{1}{2\sqrt{\pi}}$ | | | |
| | | $-\frac{1}{2}\sqrt{\frac{3}{2\pi}}\,y$ | $\frac{1}{2}\sqrt{\frac{3}{\pi}}\,z$ | $-\frac{1}{2}\sqrt{\frac{3}{2\pi}}\,x$ | | |
| | $\frac{1}{2}\sqrt{\frac{15}{2\pi}}\,x\,y$ | $-\frac{1}{2}\sqrt{\frac{15}{2\pi}}\,y\,z$ | $\frac{1}{4}\sqrt{\frac{5}{\pi}}\,(3z^2-1)$ | $-\frac{1}{2}\sqrt{\frac{15}{2\pi}}\,x\,z$ | $\frac{1}{4}\sqrt{\frac{15}{2\pi}}\,(x-y)(x+y)$ | |
| $-\frac{1}{8}\sqrt{\frac{35}{\pi}}\,(3x^2y-y^3)$ | $\frac{1}{2}\sqrt{\frac{105}{2\pi}}\,x\,y\,z$ | $-\frac{1}{4}\sqrt{\frac{7}{3\pi}}\,y\left(\frac{15z^2}{2}-\frac{3}{2}\right)$ | $\frac{1}{2}\sqrt{\frac{7}{\pi}}\left(\frac{5z^3}{2}-\frac{3z}{2}\right)$ | $-\frac{1}{4}\sqrt{\frac{7}{3\pi}}\,x\left(\frac{15z^2}{2}-\frac{3}{2}\right)$ | $\frac{1}{4}\sqrt{\frac{105}{2\pi}}\,z(x^2-y^2)$ | $-\frac{1}{8}\sqrt{\frac{35}{\pi}}\,(x^3-3xy^2)$ |

# Overview of spherical harmonics





# Overview of spherical harmonics



- Spherical harmonics satisfy the $L_2(\mathbb{S}^2)$ orthogonality condition:

$$\int_{\mathbb{S}^2} Y_\ell^m(\mathbf{x}) Y_k^n(\mathbf{x}) d\mu(\mathbf{x}) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{-\pi}^{\pi} Y_\ell^m(\theta,\varphi) Y_k^n(\theta,\varphi) \cos\theta \, d\varphi d\theta = \delta_{k\ell}\delta_{mn}$$

- They form a complete orthonormal basis for $L_2(\mathbb{S}^2)$.
- If $f \in L_2(\mathbb{S}^2)$ then

$$f(\mathbf{x}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{f}_\ell^m Y_\ell^m(\mathbf{x}), \text{ where } \hat{f}_\ell^m = \int_{\mathbb{S}^2} f(\mathbf{x}) Y_\ell^m(\mathbf{x}) d\mu(\mathbf{x}).$$

- There is no counter part to the fast Fourier transform (FFT) for computing the spherical harmonic coefficients $\hat{f}_\ell^m$.
  - Fast methods of similar complexity ($\mathcal{O}(N \log N)$) have been developed, but have very large constants associated with them. So an actual computational advantage does not occur until $N$ is extremely large.

# Overview of spherical harmonics



- Two useful results on spherical harmonics we will use:
- Addition theorem: Let $\mathbf{x}, \mathbf{y} \in \mathbb{S}^2$, then for $\ell = 0, 1, \ldots$

$$\frac{4\pi}{2\ell+1} \sum_{m=-\ell}^{\ell} Y_\ell^m(\mathbf{x}) Y_\ell^m(\mathbf{y}) = P_\ell(\mathbf{x}^T \mathbf{y}),$$

where $P_\ell$ is the standard Legendre polynomial of degree $\ell$.

- Funk-Hecke formula: Let $f \in L_1(-1,1)$ and have the Legendre expansion

$$f(t) = \sum_{k=0}^{\infty} a_k P_k(t), \text{ where } a_k = \frac{2k+1}{2} \int_{-1}^{1} f(t) P_k(t) dt.$$

Then for any spherical harmonic $Y_\ell^m$ the following holds:

$$\int_{\mathbb{S}^2} f(\mathbf{x} \cdot \mathbf{y}) Y_\ell^m(\mathbf{x}) d\mu(\mathbf{x}) = \frac{4\pi a_\ell}{2\ell+1} Y_\ell^m(\mathbf{y}).$$

# Theorems for positive definite zonal kernels

*Supplementary material*

**Definition.** A zonal kernel $\psi : [-1,1] \to \mathbb{R}$ is said to be a positive definite zonal kernel on $\mathbb{S}^2$ if for any distinct set of nodes $X = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{S}^2$ and $\underline{b} \in \mathbb{R}^N \backslash \{0\}$ the matrix $A = \{\psi(\mathbf{x}_i^T \mathbf{x}_j)\}$ is positive definite, i.e.

$$\sum_{i=1}^N \sum_{j=1}^N b_i \psi(\mathbf{x}_i^T \mathbf{x}_j) b_j > 0.$$

**Theorem** (Schoenberg (1942)). If a zonal kernel $\psi : [-1,1] \to \mathbb{R}$ is expressible in a Legendre series as

$$\psi(t) = \sum_{\ell=0}^{\infty} a_\ell P_\ell(t)$$

where $a_\ell > 0$ for $\ell \geq 0$ and $\sum_{\ell=0}^{\infty} a_\ell < \infty$ then $\psi$ is a positive definite zonal kernel on $\mathbb{S}^2$.

# Theorems for positive definite zonal kernels



**Theorem** (Schoenberg (1942)). If a zonal kernel $\psi : [-1,1] \to \mathbb{R}$ is expressible in a Legendre series as

$$\psi(t) = \sum_{\ell=0}^{\infty} a_\ell P_\ell(t)$$

where $a_\ell > 0$ for $\ell \geq 0$ and $\sum_{\ell=0}^{\infty} a_\ell < \infty$ then $\psi$ is a positive definite zonal kernel on $\mathbb{S}^2$.

Proof:

1. The condition $\sum_{\ell=0}^{\infty} a_\ell < \infty$ guarantees that $\psi \in C(\mathbb{S}^2)$.
2. Use the addition theorem: Let $X = \{\mathbf{x}_j\}_{j=1}^{N} \subset \mathbb{S}^2$ and $\underline{b} \in \mathbb{R}^N \backslash \{0\}$ then

$$\begin{aligned}\sum_{i=1}^{N}\sum_{j=1}^{N} b_i \psi(\mathbf{x}_i^T \mathbf{x}_j) b_j &= \sum_{i=1}^{N}\sum_{j=1}^{N} b_i b_j \sum_{\ell=0}^{\infty} a_\ell P_\ell(\mathbf{x}_i^T \mathbf{x}_j) \\ &= \sum_{\ell=0}^{\infty} \frac{4\pi a_\ell}{2\ell+1} \sum_{m=-\ell}^{\ell} \sum_{i=1}^{N}\sum_{j=1}^{N} b_i b_j Y_\ell^m(\mathbf{x}_i) Y_\ell^m(\mathbf{x}_j) \\ &= \sum_{\ell=0}^{\infty} \frac{4\pi a_\ell}{2\ell+1} \sum_{m=-\ell}^{\ell} \left| \sum_{j=1}^{N} b_j Y_\ell^m(\mathbf{x}_j) \right|^2 \geq 0\end{aligned}$$

3. Show that the quadratic form must be strictly positive.

# Theorems for positive definite zonal kernels

*Supplementary material*

**Theorem** (Schoenberg (1942)). If a zonal kernel $\psi : [-1, 1] \to \mathbb{R}$ is expressible in a Legendre series as

$$\psi(t) = \sum_{\ell=0}^{\infty} a_\ell P_\ell(t)$$

where $a_\ell > 0$ for $\ell \geq 0$ and $\sum_{\ell=0}^{\infty} a_\ell < \infty$ then $\psi$ is a positive definite zonal kernel on $\mathbb{S}^2$.

- Necessary and sufficient conditions on the Legendre coefficients $a_\ell$ were only given in 2003 by Chen, Menegatto, & Sun.
  - Their result says the set $\left\{ \ell \in \mathbb{N}_0 \middle| a_\ell > 0 \right\}$ must contain infinitely many odd and infinitely many even integers.

# Conditionally positive definite zonal kernels



- Similar to $\mathbb{R}^d$, we can define conditionally positive definite zonal kernels.

**Definition.** A continuous zonal kernel $\psi : [-1, 1] \to \mathbb{R}$ is said to be conditionally positive definite of order $k$ on $\mathbb{S}^2$ if, for any distinct $X = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{S}^2$, and all $\mathbf{b} \in \mathbb{R}^N \backslash \{\mathbf{0}\}$ satisfying

$$\sum_{j=1}^{N} b_j p(\mathbf{x}_j) = 0$$

for all spherical harmonics of degree $< k$, the following is satisfied:

$$\sum_{i=1}^{N} \sum_{j=1}^{N} b_i \psi(\mathbf{x}_i^T \mathbf{x}_j) b_j > 0.$$

**Theorem.** If the Legendre expansion coefficients of $\psi : [-1, 1] \to \mathbb{R}$ satisfy $a_\ell > 0$ for $\ell \geq k$ and $\sum_{\ell=0}^{\infty} a_\ell < \infty$.

Proof: Use same ideas as the positive definite case.

# Conditionally positive definite zonal kernels



**Definition.** Let $\psi : [-1,1] \to \mathbb{R}$ be a continuous zonal kernel and $\{p_i(\mathbf{x})\}_{i=1}^{k^2}$ be a basis for the space of all spherical harmonics of degree $k-1$. The general SBF interpolant for the distinct nodes $X = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{S}^2$ and some target, $f$, sampled on $X$, $\{f_j\}_{j=1}^N$ is

$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \psi(\mathbf{x}^T \mathbf{x}_j) + \sum_{\ell=1}^{k^2} d_\ell p_\ell(\mathbf{x}),$$

where $I_X f(\mathbf{x}_i) = f_i$, $i = 1,\ldots,N$ and $\sum_{j=1}^{N} c_j p_\ell(\mathbf{x}_j) = 0$, $\ell = 1,\ldots,k^2$.

In linear system form, these constraints are

$$\begin{bmatrix} A & P \\ P^T & 0 \end{bmatrix} \begin{bmatrix} \underline{c} \\ \underline{d} \end{bmatrix} = \begin{bmatrix} \underline{f} \\ \underline{0} \end{bmatrix}, \text{ where } a_{i,j} = \psi(\mathbf{x}_i^T \mathbf{x}_j),\; p_{i,\ell} = p_\ell(\mathbf{x}_i)$$

**Theorem.** The above linear system is invertible for any distinct $X$, provided

- $\text{rank}(P) = k^2$,
- $\psi$ is conditionally positive definite of of order $k$.

# Conditionally positive definite zonal kernels

**Definition.** Let $\psi : [-1,1] \to \mathbb{R}$ be a continuous zonal kernel and $\{p_i(\mathbf{x})\}_{i=1}^{k^2}$ be a basis for the space of all spherical harmonics of degree $k-1$. The general SBF interpolant for the distinct nodes $X = \{\mathbf{x}_j\}_{j=1}^{N} \subset \mathbb{S}^2$ and some target, $f$, sampled on $X$, $\{f_j\}_{j=1}^{N}$ is

$$I_X f(\mathbf{x}) = \sum_{j=1}^{N} c_j \psi(\mathbf{x}^T \mathbf{x}_j) + \sum_{\ell=1}^{k^2} d_\ell p_\ell(\mathbf{x}),$$

where $I_X f(\mathbf{x}_i) = f_i$, $i = 1, \ldots, N$ and $\sum_{j=1}^{N} c_j p_\ell(\mathbf{x}_j) = 0$, $\ell = 1, \ldots, k^2$.

In linear system form, these constraints are

$$\begin{bmatrix} A & P \\ P^T & 0 \end{bmatrix} \begin{bmatrix} \underline{c} \\ \underline{d} \end{bmatrix} = \begin{bmatrix} \underline{f} \\ \underline{0} \end{bmatrix}, \text{ where } a_{i,j} = \psi(\mathbf{x}_i^T \mathbf{x}_j),\ p_{i,\ell} = p_\ell(\mathbf{x}_i)$$

**Example** (Restricted thin plate spline, or surface spline). Let

- $\psi(t) = (1-t)\log(2-2t)$
- $p_1(\mathbf{x}) = 1$, $p_2(\mathbf{x}) = x$, $p_3(\mathbf{x}) = y$, and $p_4(\mathbf{x}) = z$.

The system has a unique solution provided $X$ are distinct.

# Spherical Fourier coefficients



- More useful to work with a zonal kernels spherical Fourier coefficients $\hat{\psi}_\ell$. These are related to Legendre coefficients through the Funk-Hecke formula:

$$\psi(\mathbf{x}^T\mathbf{y}) = \sum_{\ell=0}^{\infty} \hat{\psi}(\ell) \sum_{m=-\ell}^{\ell} Y_\ell^m(\mathbf{x}) Y_\ell^m(\mathbf{y}) \Longrightarrow \hat{\psi}(\ell) := \frac{4\pi a_\ell}{2\ell+1}$$

- Error estimates for SBF interpolants are governed by the asymptotic decay of $\hat{\psi}_\ell$.
- Stable algorithms (RBF-QR) also work with $\hat{\psi}_\ell$ (more on this later...)
- Baxter & Hubbert (2001) computed $\hat{\psi}_\ell$ for many standard RBFs restricted to $\mathbb{S}^2$.
- zu Castell & Filbir (2005) and Narcowich, Sun, & Ward (2007) linked the spherical Fourier coefficients of restricted RBFs to the standard Fourier coefficients in $\mathbb{R}^3$:

$$\hat{\psi}_\ell = \int_0^\infty u\hat{\phi}(u) J_{\ell+1/2}(u)du,$$

where $\hat{\phi}$ is the Hankel transform of the RBF in $\mathbb{R}^3$.

# Examples of positive definite zonal kernels



- Examples of positive definite (PD) and order $k$ conditionally positive definite (CPD($k$)) zonal kernels with their spherical Fourier coefficients.

| Name | Kernel ($r(t)=\sqrt{2-2t}$) | Fourier coefficients $\hat{\psi}_\ell$ ($0<h<1$, $\varepsilon>0$) | Type |
|---|---|---|---|
| Legendre | $\psi(t)=(1+h^2-2ht)^{-1/2}$ | $\hat{\psi}_\ell=\dfrac{2\pi h^\ell}{\ell+1/2}$ | PD |
| Poisson | $\psi(t)=(1-h^2)(1+h^2-2ht)^{-3/2}$ | $\hat{\psi}_\ell=4\pi h^\ell$ | PD |
| Spherical | $\psi(t)=1-r(t)+\dfrac{(r(t))^2}{2}\log\left(\dfrac{r(t)+2}{r(t)}\right)$ | $\hat{\psi}_\ell=\dfrac{2\pi}{(\ell+1/2)(\ell+1)(\ell+2)}$ | PD |
| Gaussian | $\psi(t)=\exp(-(\varepsilon r(t))^2)$ | $\varepsilon^{2\ell}\dfrac{4\pi^{3/2}}{\varepsilon^{2\ell+1}}e^{-2\varepsilon^2}I_{\ell+1/2}(2\varepsilon^2)$ | PD |
| IMQ | $\psi(t)=\dfrac{1}{\sqrt{1+(\varepsilon r(t))^2}}$ | $\varepsilon^{2\ell}\dfrac{4\pi}{(\ell+1/2)}\left(\dfrac{2}{1+\sqrt{4\varepsilon^2+1}}\right)^{2\ell+1}$ | PD |
| MQ | $\psi(t)=-\sqrt{1+(\varepsilon r(t))^2}$ | $\varepsilon^{2\ell}\dfrac{2\pi(2\varepsilon^2+1+(\ell+1/2)\sqrt{1+4\varepsilon^2})}{(\ell+3/2)(\ell+1/2)(\ell-1/2)}\left(\dfrac{2}{1+\sqrt{4\varepsilon^2+1}}\right)^{2\ell+1}$ | CPD(1) |
| TPS | $\psi(t)=(r(t))^2\log(r(t))$ | $\dfrac{8\pi}{(\ell+2)(\ell+1)\ell(\ell-1)}$ | CPD(2) |
| Cubic | $\psi(t)=(r(t))^3$ | $\dfrac{18\pi}{(\ell+5/2)(\ell+3/2)(\ell+1/2)(\ell-1/2)(\ell-3/2)}$ | CPD(2) |

- First three kernels are specific to $\mathbb{S}^2$, while the last 5 are RBFs restricted to $\mathbb{S}^2$.

# Error estimates



- Goal: Present some known results on error estimates for SBF interpolants for target function of various smoothness.

- We will introduce (or review) some background notation and material that is necessary for the proofs of the estimates, but will not prove them.
  - Reproducing kernel Hilbert spaces (RKHS)
  - Sobolev spaces on $\mathbb{S}^2$;
  - Native spaces;
  - Geometric properties of node sets $X \subset \mathbb{S}^2$.

- Brief historical notes regarding SBF error estimates:
  - Earliest results appear to be Freeden (1981), but do not depend on $\psi$ or target.
  - First Sobolev-type estimates were given in Jetter, Stöckler, & Ward (1999).
  - Since then many more results have appeared, e.g.
    Levesley, Light, Ragozin, & Sun (1999), v. Golitschek & Light (2001), Morton & Neamtu (2002), Narcowich & Ward (2002), Hubbert & Morton (2004,2004), Levesley & Sun (2005), Narcowich, Sun, & Ward (2007), Narcowich, Sun, Ward, & Wendland (2007), Sloan & Sommariva (2008), Sloan & Wendland (2009), Hangelbroek (2011).

# Reproducing kernel Hilbert spaces

*Supplementary material*

- Reproducing kernel Hilbert spaces (RKHS) play a key role deriving error estimates for SBF (and more generally RBF) interpolants.

- They allow one to view the interpolation problem as the solution to a particular optimization problem.

**Definition.** Let $\mathcal{F}(\Omega)$ be a Hilbert space of functions $f : \Omega \to \mathbb{R}$ with inner product $\langle \cdot, \cdot \rangle_{\mathcal{F}}$. If there exists a kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$ such that for all $\mathbf{y} \in \Omega$

$$f(\mathbf{y}) = \langle f, \Phi(\cdot, \mathbf{y}) \rangle_{\mathcal{F}} \text{ for all } f \in \mathcal{F},$$

then $\mathcal{F}$ is called a RKHS with reproducing kernel $\Phi$.

# Reproducing kernel Hilbert spaces

*Supplementary material*

- Reproducing kernel Hilbert spaces (RKHS) play a key role deriving error estimates for SBF (and more generally RBF) interpolants.

- They allow one to view the interpolation problem as the solution to a particular optimization problem.

  **Definition.** Let $\mathcal{F}(\Omega)$ be a Hilbert space of functions $f : \Omega \to \mathbb{R}$ with inner product $\langle \cdot, \cdot \rangle_{\mathcal{F}}$. If there exists a kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$ such that for all $\mathbf{y} \in \Omega$

  $$f(\mathbf{y}) = \langle f, \Phi(\cdot, \mathbf{y}) \rangle_{\mathcal{F}} \text{ for all } f \in \mathcal{F},$$

  then $\mathcal{F}$ is called a RKHS with reproducing kernel $\Phi$.

- The reproducing kernel $\Phi$ of a RKHS is unique.

- Existence of $\Phi$ is equivalent to the point evaluation functional $\delta_{\mathbf{y}} : \mathcal{F} \to \mathbb{R}$ being continuous. (Implied by Reisz representation theorem).

- $\Phi$ also satisfies the following:
  (1) $\Phi(\mathbf{x}, \mathbf{y}) = \Phi(\mathbf{y}, \mathbf{x})$ for $x, y \in \Omega$;     (2) $\Phi$ is positive semi-definite on $\Omega$.

# Reproducing kernel Hilbert spaces



**Example.** The space spanned by all spherical harmonics of degree $n$ with the standard $L_2(\mathbb{S}^2)$ inner product $\langle \cdot, \cdot \rangle_{L_2}$ is a RKHS with reproducing kernel

$$\Phi_n(\mathbf{x}, \mathbf{y}) = \sum_{k=0}^{n} \frac{2k+1}{4\pi} P_k(\mathbf{x}^T \mathbf{y}).$$

# Reproducing kernel Hilbert spaces



**Example.** The space spanned by all spherical harmonics of degree $n$ with the standard $L_2(\mathbb{S}^2)$ inner product $\langle \cdot, \cdot \rangle_{L_2}$ is a RKHS with reproducing kernel

$$\Phi_n(\mathbf{x}, \mathbf{y}) = \sum_{k=0}^{n} \frac{2k+1}{4\pi} P_k(\mathbf{x}^T \mathbf{y}).$$

Let $\mathbf{x}, \mathbf{y} \in \mathbb{S}^2$ and $f(\mathbf{x}) = \sum_{\ell=0}^{n} \sum_{m=-\ell}^{\ell} c_\ell^m Y_\ell^m(\mathbf{x})$ for some coefficients $c_\ell^m$. Then

$$\begin{aligned}
\langle f, \Phi_n(\cdot, \mathbf{y}) \rangle_{L_2} &= \int_{\mathbb{S}^2} f(\mathbf{x}) \Phi_n(\mathbf{x}, \mathbf{y}) d\mu(\mathbf{x}) \\
&= \int_{\mathbb{S}^2} \left( \sum_{\ell=0}^{n} \sum_{m=-\ell}^{\ell} c_\ell^m Y_\ell^m(\mathbf{x}) \right) \left( \sum_{k=0}^{n} \frac{2k+1}{4\pi} P_k(\mathbf{x}^T \mathbf{y}) \right) d\mu(\mathbf{x}) \\
&= \sum_{k=0}^{n} \frac{2k+1}{4\pi} \sum_{\ell=0}^{n} \sum_{m=-\ell}^{\ell} c_\ell^m \int_{\mathbb{S}^2} P_k(\mathbf{x}^T \mathbf{y}) Y_\ell^m(\mathbf{x}) d\mu(\mathbf{x}) \\
&= \sum_{k=0}^{n} \frac{2k+1}{4\pi} \sum_{m=-k}^{k} \frac{4\pi}{2k+1} c_k^m Y_k^m(\mathbf{y}) \qquad \text{(Funk-Hecke formula)} \\
&= \sum_{k=0}^{n} \sum_{m=-k}^{k} c_k^m Y_k^m(\mathbf{y}) = f(\mathbf{y})
\end{aligned}$$

# Sobolev spaces



- Sobolev spaces on $\mathbb{S}^2$ can be defined in terms of spherical Harmonics.

**Definition.** The Sobolev space of order $\tau$ on $\mathbb{S}^2$ is given by

$$H^\tau(\mathbb{S}^2) = \left\{ f \in L_2(\mathbb{S}^2) \middle| \|f\|^2_{H^\tau} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} (1+\ell(\ell+1))^\tau \left| \hat{f}_\ell^m \right|^2 < \infty \right\}.$$

Here $\|\cdot\|_{H^\tau}$ is a norm induced by the inner product

$$\langle f, g \rangle_{H^\tau} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} (1+\ell(\ell+1))^\tau \hat{f}_\ell^m \hat{g}_\ell^m,$$

where $\hat{f}_\ell^m = \langle f, Y_\ell^m \rangle_{L_2} = \int_{\mathbb{S}^2} f(\mathbf{x}) Y_\ell^m(\mathbf{x}) d\mu(\mathbf{x})$.

- Compare to Sobolev spaces on $\mathbb{R}^3$:

$$H^\beta(\mathbb{R}^3) = \left\{ f \in L_2(\mathbb{R}^3) \middle| \|f\|^2_{H^\beta} = \int_{\mathbb{R}^3} (1+\|\boldsymbol{\omega}\|^2)^\beta \left| \hat{f}(\boldsymbol{\omega}) \right|^2 d\mathbf{x} < \infty \right\}.$$

# Sobolev spaces



- Sobolev spaces on $\mathbb{S}^2$ can be defined in terms of spherical Harmonics.

  **Definition.** The Sobolev space of order $\tau$ on $\mathbb{S}^2$ is given by

$$H^\tau(\mathbb{S}^2) = \left\{ f \in L_2(\mathbb{S}^2) \,\middle|\, \|f\|^2_{H^\tau} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} (1 + \ell(\ell+1))^\tau \left| \hat{f}_\ell^m \right|^2 < \infty \right\}.$$

- Sobolev embedding theorem implies $H^\tau(\mathbb{S}^2)$ is continuously embedded in $C(\mathbb{S}^2)$ for $\tau > 1$. Thus, $H^\tau(\mathbb{S}^2)$ is a RKHS.

- Can show the reproducing kernel is $\Phi_\tau(\mathbf{x},\mathbf{y}) = \sum_{\ell=0}^{\infty} (1+\ell(\ell+1))^{-\tau} \frac{2\ell+1}{4\pi} P_\ell(\mathbf{x}\cdot\mathbf{y})$.

# Native spaces



- Each positive definite zonal kernel $\psi$ naturally gives rise to a RKHS on $\mathbb{S}^2$, which is called the native space of $\psi$.

- This is the natural space to understand approximation with shifts of $\psi$.

**Definition.** Let $\psi$ be a positive definite zonal kernel with spherical Fourier coefficients $\hat{\psi}_\ell$, $\ell = 0, 1, \ldots$. The native space $\mathcal{N}_\psi$ of $\psi$ is given by

$$\mathcal{N}_\psi = \left\{ f \in L_2(\mathbb{S}^2) \middle| \|f\|_{\mathcal{N}_\psi} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \frac{|\hat{f}_\ell^m|^2}{\hat{\psi}_\ell} < \infty \right\},$$

with inner product

$$\langle f, g \rangle_{\mathcal{N}_\psi} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \frac{\hat{f}_\ell^m \hat{g}_\ell^m}{\hat{\psi}_\ell}.$$

- A similar definition holds for conditionally positive definite kernels, but the inner product has to be slightly modified (see Hubbert, 2002).

# Native spaces



- An important "optimality" result stems from $\mathcal{N}_\psi(\mathbb{S}^2)$ being a RKHS.

- Consider the following optimization problem:

  **Problem.** Let $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ be a distinct set of nodes on $\mathbb{S}^2$ and let $\{f_1, \ldots, f_N\}$ be samples of some target function $f$ on $X$. Find $s \in \mathcal{N}_\psi(\mathbb{S}^2)$ that satisfies $s(\mathbf{x}_j) = f_j$, $j = 1, \ldots, N$ and has minimal native space norm $\|s\|_{\mathcal{N}_\psi}$, i.e.

$$\text{minimize} \left\{ \|s\|_{\mathcal{N}_\psi} \middle| s \in \mathcal{N}_\psi(\mathbb{S}^2) \text{ with } s\big|_X = f\big|_X \right\}.$$

  Solution: $s$ is the unique SBF interpolant to $f\big|_X$ using the kernel $\psi$.

- SBF interpolants also have nice properties in their respective native spaces:

  1. $\|f - I_{\psi,X} f\|_{\mathcal{N}_\psi}^2 + \|I_{\psi,X} f\|_{\mathcal{N}_\psi}^2 = \|f\|_{\mathcal{N}_\psi}^2$
  2. $\|f - I_{\psi,X} f\|_{\mathcal{N}_\psi} \leq \|f\|_{\mathcal{N}_\psi}$

# Native spaces



- Note similarity between Sobolev space $H^\tau(\mathbb{S}^2)$ and $\mathcal{N}_\psi(\mathbb{S}^2)$:

$$H^\tau(\mathbb{S}^2) = \left\{ f \in L_2(\mathbb{S}^2) \,\middle|\, \|f\|^2_{H^\tau} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} (1+\ell(\ell+1))^\tau \left|\hat{f}_\ell^m\right|^2 < \infty \right\}$$

$$\mathcal{N}_\psi(\mathbb{S}^2) = \left\{ f \in L_2(\mathbb{S}^2) \,\middle|\, \|f\|_{\mathcal{N}_\psi} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \frac{|\hat{f}_\ell^m|^2}{\hat{\psi}_\ell} < \infty \right\}$$

- If $\hat{\psi}_\ell \sim (1+\ell(\ell+1))^{-\tau}$, then it follows that $\mathcal{N}_\psi = H^\tau$, with equivalent norms.

- This is one reason we care about the asymptotic behavior of $\hat{\psi}_\ell$.

- For RBFs restricted to $\mathbb{S}^2$, we have the following nice result connecting the asymptotics of the spherical Fourier coefficients to the Fourier transform (Levesley & Hubbert (2001), zu Castell & Filbir (2005), Narcowich, Sun, & Ward (2007)):

> If $\psi$ is an SBF obtained by restricting an RBF $\phi$ to $\mathbb{S}^2$ and if $\hat{\phi}(\boldsymbol{\omega}) \sim (1+\|\boldsymbol{\omega}\|_2^2)^{-(\tau+1/2)}$ then $\hat{\psi}_\ell \sim (1+\ell(\ell+1))^{-\tau}$.

# Native spaces



- Examples of radial kernels $\phi$ and their norm-equivalent native spaces $\mathcal{N}_\psi$ when restricted to $\mathbb{S}^2$:

| Name | RBF (use $r = \sqrt{2-2t}$ to get SBF $\psi$) | $\mathcal{N}_\psi(\mathbb{S}^2)$ |
|---|---|---|
| Matern | $\phi_2(r) = e^{-\varepsilon r}$ | $H^{1.5}(\mathbb{S}^2)$ |
| TPS(1) | $\phi(r) = r^2 \log(r)$ | $H^{2}(\mathbb{S}^2)$ |
| Cubic | $\phi(r) = r^3$ | $H^{2.5}(\mathbb{S}^2)$ |
| TPS(2) | $\phi(r) = r^4 \log(r)$ | $H^{3}(\mathbb{S}^2)$ |
| Wendland | $\phi_{3,2}(r) = (1-\varepsilon r)_+^6 (3 + 18(\varepsilon r) + 15(\varepsilon r)^2)$ | $H^{3.5}(\mathbb{S}^2)$ |
| Matern | $\phi_5(r) = e^{-\varepsilon r}(15 + 15(\varepsilon r) + 6(\varepsilon r)^2 + (\varepsilon r)^3)$ | $H^{4.5}(\mathbb{S}^2)$ |

- The spherical Fourier coefficients for all these restricted kernels have algebraic decay rates.

- For kernels with spherical Fourier coefficients with exponential decay rates (*e.g.* Gaussian and multiquadric) the Native spaces are no longer equivalent to Sobolev spaces.

- These natives spaces do satisfy: $\mathcal{N}_\psi(\mathbb{S}^2) \subset H^\tau(\mathbb{S}^2)$ for all $\tau > 1$.

- Error estimates for interpolants are directly linked to the native space of $\psi$.

# Geometric properties of node sets



- The following properties for node sets on the sphere appear in the error estimates:

- Mesh norm

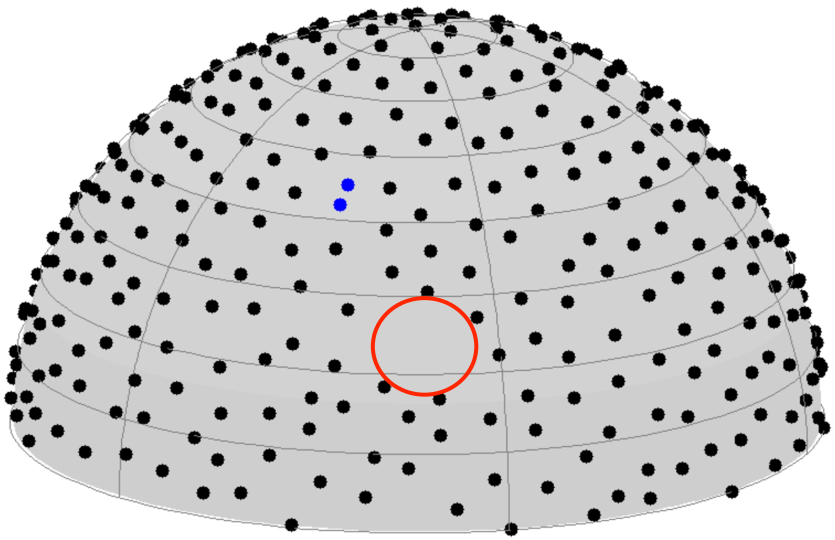
$$h_X = \sup_{\mathbf{x}\in\mathbb{S}^2} \text{dist}_{\mathbb{S}^2}(\mathbf{x}, X)$$

- Separation radius

$$q_X = \frac{1}{2}\min_{i\neq j} \text{dist}_{\mathbb{S}^2}(\mathbf{x}_i, \mathbf{x}_j)$$

- Mesh ratio

$$\rho_X = \frac{h_X}{q_X}$$



$$X = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{S}^2$$

(Only part of the sphere is shown)

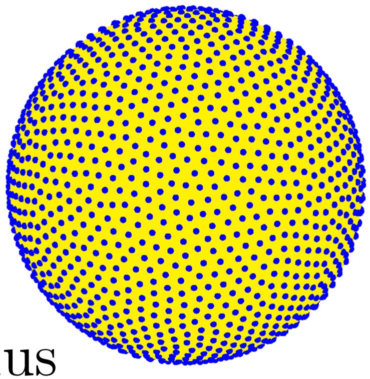
# Interpolation error estimates



- We start with known error estimates for kernels of finite smoothness.
  Jetter, Stöckler, & Ward (1999), Morton & Neamtu (2002), Hubbert & Morton (2004,2004), Narcowich, Sun, Ward, & Wendland (2007)

Notation:

- $\psi$ is the SBF
- $\hat{\psi}_\ell \sim (1+\ell(\ell+1))^{-\tau}$, $\tau > 1$
- $\mathcal{N}_\psi(\mathbb{S}^2) = H^\tau(\mathbb{S}^2)$
- $I_X f$ is SBF interpolant of $f\big|_X$
- $X = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{S}^2$
- $h_X$ = mesh-norm
- $q_X$ = separation radius
- $\rho_X = h_X/q_X$, mesh ratio



Theorem. Target functions in the native space.

If $f \in H^\tau(\mathbb{S}^2)$ then $\|f - I_X f\|_{L_p(\mathbb{S}^2)} = \mathcal{O}(h_X^{\tau-2(1/2-1/p)_+})$ for $1 \le p \le \infty$. In particular,

$$\|f - I_X f\|_{L_1(\mathbb{S}^2)} = \mathcal{O}(h_X^{\tau})$$
$$\|f - I_X f\|_{L_2(\mathbb{S}^2)} = \mathcal{O}(h_X^{\tau})$$
$$\|f - I_X f\|_{L_\infty(\mathbb{S}^2)} = \mathcal{O}(h_X^{\tau-1})$$

# Interpolation error estimates



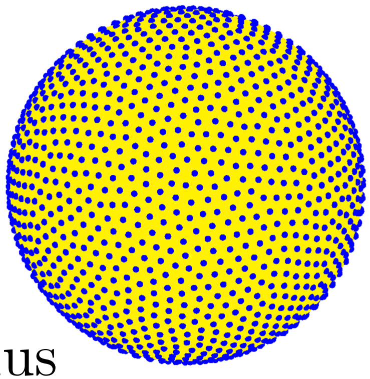
- We start with known error estimates for kernels of finite smoothness.

  Jetter, Stöckler, & Ward (1999), Morton & Neamtu (2002), Hubbert & Morton (2004,2004), Narcowich, Sun, Ward, & Wendland (2007)

Notation:

- $\psi$ is the SBF
- $\hat{\psi}_\ell \sim (1+\ell(\ell+1))^{-\tau}$, $\tau > 1$
- $\mathcal{N}_\psi(\mathbb{S}^2) = H^\tau(\mathbb{S}^2)$
- $I_X f$ is SBF interpolant of $f\big|_X$

- $X = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{S}^2$
- $h_X$ = mesh-norm
- $q_X$ = separation radius
- $\rho_X = h_X/q_X$, mesh ratio



Theorem. Target functions twice as smooth as the native space.

$$\text{If } f \in H^{2\tau}(\mathbb{S}^2) \text{ then } \|f - I_X f\|_{L_p(\mathbb{S}^2)} = \mathcal{O}(h_X^{2\tau}) \text{ for } 1 \le p \le \infty.$$

Remark. Known as the "doubling trick" from spline theory. (Schaback 1999)

# Interpolation error estimates



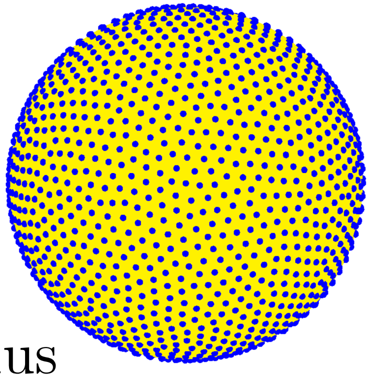
- We start with known error estimates for kernels of finite smoothness.

  Jetter, Stöckler, & Ward (1999), Morton & Neamtu (2002), Hubbert & Morton (2004,2004), Narcowich, Sun, Ward, & Wendland (2007)

Notation:

- $\psi$ is the SBF
- $\hat{\psi}_\ell \sim (1+\ell(\ell+1))^{-\tau}$, $\tau > 1$
- $\mathcal{N}_\psi(\mathbb{S}^2) = H^\tau(\mathbb{S}^2)$
- $I_X f$ is SBF interpolant of $f\big|_X$

- $X = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{S}^2$
- $h_X$ = mesh-norm
- $q_X$ = separation radius
- $\rho_X = h_X / q_X$, mesh ratio



Theorem. Target functions rougher than the native space.

If $f \in H^\beta(\mathbb{S}^2)$ for $\tau > \beta > 1$ then $\|f - I_X f\|_{L_p(\mathbb{S}^2)} = \mathcal{O}(\rho^{\tau-\beta} h_X^{\tau - 2(1/2-1/p)_+})$ for $1 \le p \le \infty$.

Remark.

(1) Referred to as "escaping the native space". (Narcowich, Ward, & Wendland (2005, 2006)).

(2) These rates are the best possible.

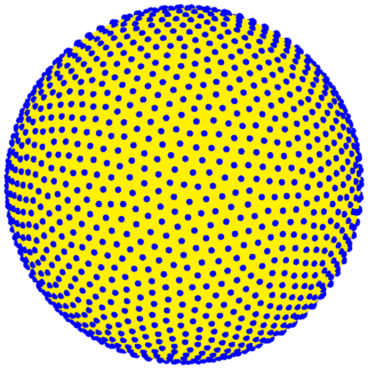
# Interpolation error estimates

*Supplementary material*

- Error estimates for infinitely smooth kernels (e.g. Gaussian, multiquadric).
  Jetter, Stöckler, & Ward (1999)

Notation:

- $\psi$ is the SBF
- $\hat{\psi}_\ell \sim \exp(-\alpha(2\ell+1)),\ \alpha > 0$
- $X = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{S}^2$
- $h_X = \text{mesh-norm}$



- $\mathcal{N}_\psi(\mathbb{S}^2) = \left\{ f \in L_2(\mathbb{S}^2) \,\middle|\, \|f\|_{\mathcal{N}_\psi} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \frac{|\hat{f}_\ell^m|^2}{\hat{\psi}_\ell} < \infty \right\}$

Theorem. Target functions in the native space.

$$\text{If } f \in \mathcal{N}_\psi(\mathbb{S}^2) \text{ then } \|f - I_X f\|_{L_\infty(\mathbb{S}^2)} = \mathcal{O}(h_X^{-1} \exp(-\alpha/2h_X)).$$

Remarks:

(1) This is called spectral (or exponential) convergence.

(2) Function space may be small, but does include all band-limited functions.

(3) Only known result I am aware of (too bad there are not more).

(4) Numerical results indicate convergence is also fine for less smooth functions.

# Optimal nodes

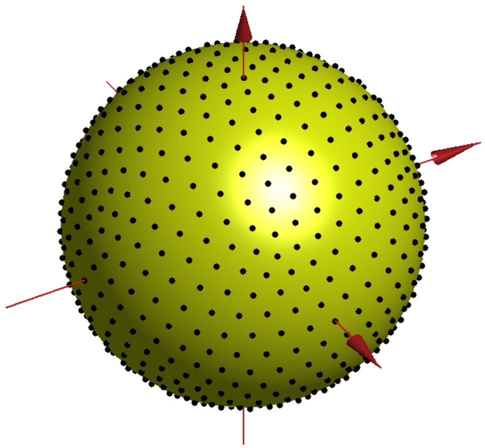
*Supplementary material*

- If one has the freedom to choose the nodes, then the error estimates indicate they should be roughly as evenly spaced as possible.

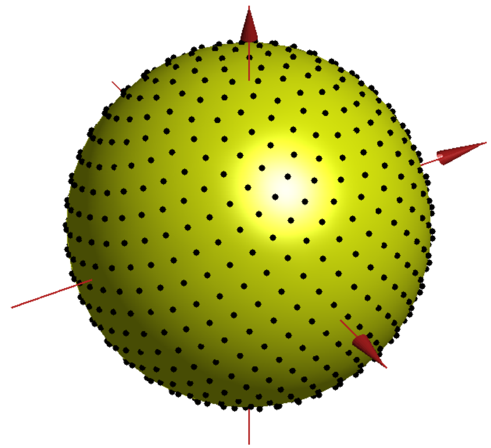
Examples:

Deterministic

Icosahedral

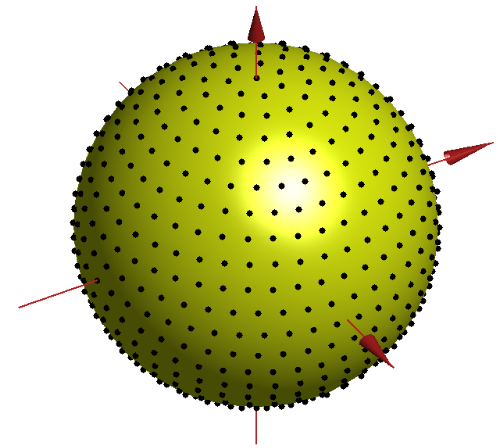


Fibonacci



Swinbank & Purser (2006)

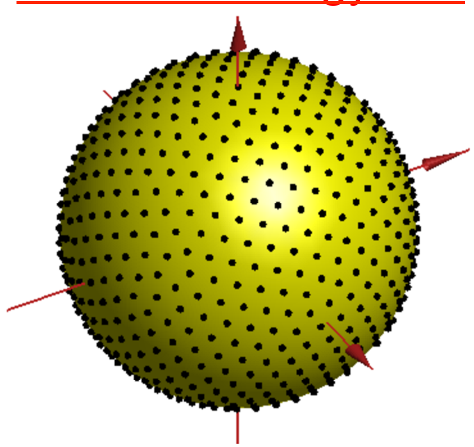
Equal area



Saff & Kuijlaars (1997)

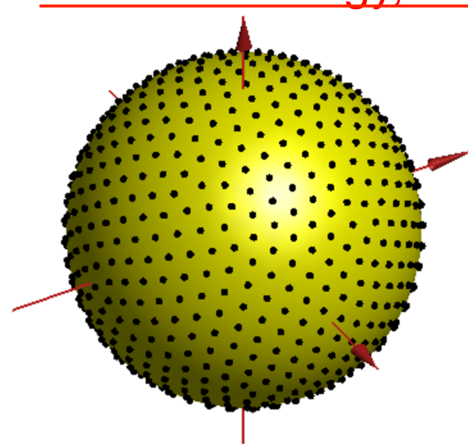
Non-deterministic

Minimum energy *s*=2



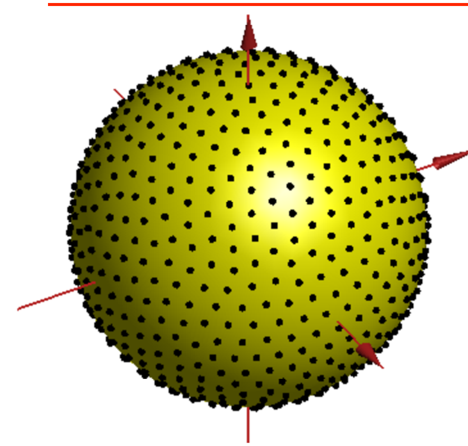
Hardin & Saff (2004)

Minimum energy, *s*=3



Riesz energy: $\|\mathbf{x}-\mathbf{y}\|_2^{-s}$

Maximal determinant



Womersley & Sloan (2001)

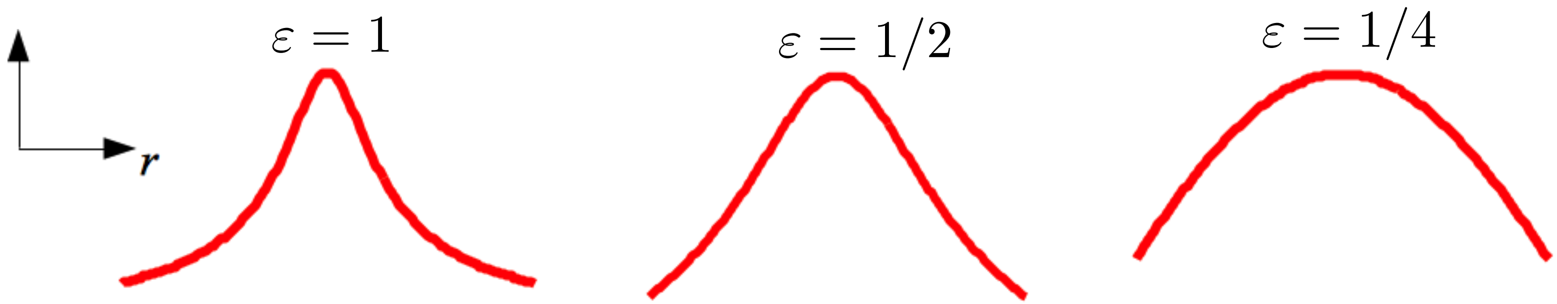
# What about the shape parameter?

*Supplementary material*

- Smooth kernels with a shape parameter.

<u>Ex:</u> $\phi(r) = \exp(-(\varepsilon r)^2)$ $\phi(r) = \dfrac{1}{\sqrt{1+(\varepsilon r)^2}}$ $\phi(r) = \sqrt{1+(\varepsilon r)^2}$

<u>Issue:</u> Effect of decreasing $\varepsilon$ leads to severe ill-conditioning of interp. matrices




Basis functions get flatter as $\varepsilon \longrightarrow 0$

<u>Linear system for determining the interpolation coefficients</u>

$$\underbrace{\begin{bmatrix} \phi(\|\mathbf{x}_1-\mathbf{x}_1\|) & \phi(\|\mathbf{x}_1-\mathbf{x}_2\|) & \cdots & \phi(\|\mathbf{x}_1-\mathbf{x}_N\|) \\ \phi(\|\mathbf{x}_2-\mathbf{x}_1\|) & \phi(\|\mathbf{x}_2-\mathbf{x}_2\|) & \cdots & \phi(\|\mathbf{x}_2-\mathbf{x}_N\|) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(\|\mathbf{x}_N-\mathbf{x}_1\|) & \phi(\|\mathbf{x}_N-\mathbf{x}_2\|) & \cdots & \phi(\|\mathbf{x}_N-\mathbf{x}_N\|) \end{bmatrix}}_{A_X} \underbrace{\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix}}_{\underline{c}} = \underbrace{\begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_N \end{bmatrix}}_{\underline{f}}$$

$A_X$ is guaranteed to be positive definite if $\phi$ is positive definite.

RBF-Direct

# RBF interpolation in the "flat" limit



RBF interpolant: $$I_{X,\varepsilon} f(\mathbf{x}) = \sum_{j=1}^{N} c_j(\varepsilon)\phi_\varepsilon(\|\mathbf{x}-\mathbf{x}_j\|)$$

**Theorem** (Driscoll & Fornberg (2002)). For $N$ nodes in 1-D, the RBF interpolant (for certain smooth kernels) converges to the standard Lagrange interpolant as $\varepsilon \longrightarrow 0$ (flat limit)

- Higher dimensions: Limit usually exits and takes the form of a multivariate polynomial as $\varepsilon \longrightarrow 0$.
  - Fornberg, W, & Larsson (2004), Larsson & Fornberg (2005), Schaback (2005,2006), Lee, Yoon, & Yoon (2007)
  - In the case of the Gaussian kernel, the interpolant always converges to the de Boor & Ron "least polynomial interpolant".
- Sphere: Limit (usually) exits and converges to a spherical harmonic interpolant (Fornberg & Piret (2007)).

# Base vs. space



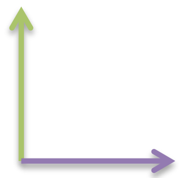
- Key observation: The space spanned by linear combinations of positive definite radial kernels (in $\mathbb{R}^d$ or $\mathbb{S}^2$) is good for approximation

  BUT, the standard basis $\{\phi(\cdot, \mathbf{x}_1), \ldots, \phi(\cdot, \mathbf{x}_N)\}$ can be problematic.

<u>Analogy:</u>
<u>(Fornberg)</u>

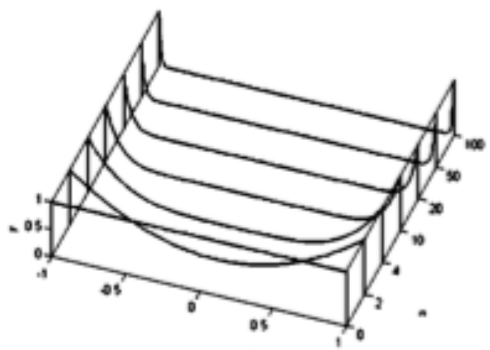
Vectors

Bad basis for $\mathbb{R}^2$

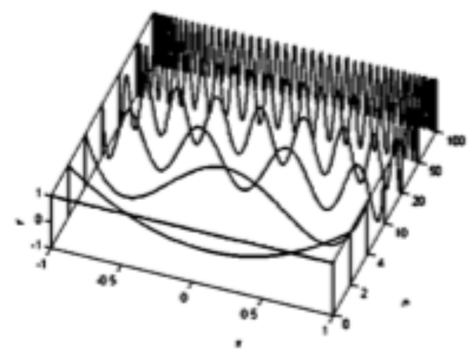


Good basis for $\mathbb{R}^2$

Polynomials

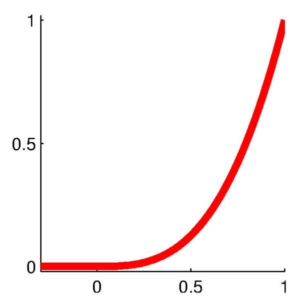


Bad basis: $x^n$, $n = 0, 1, \ldots$

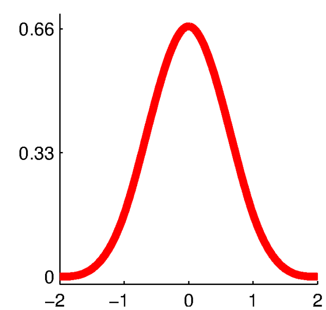


Chebyshev basis: $T_n(x)$, $n = 0, 1, \ldots$

Splines



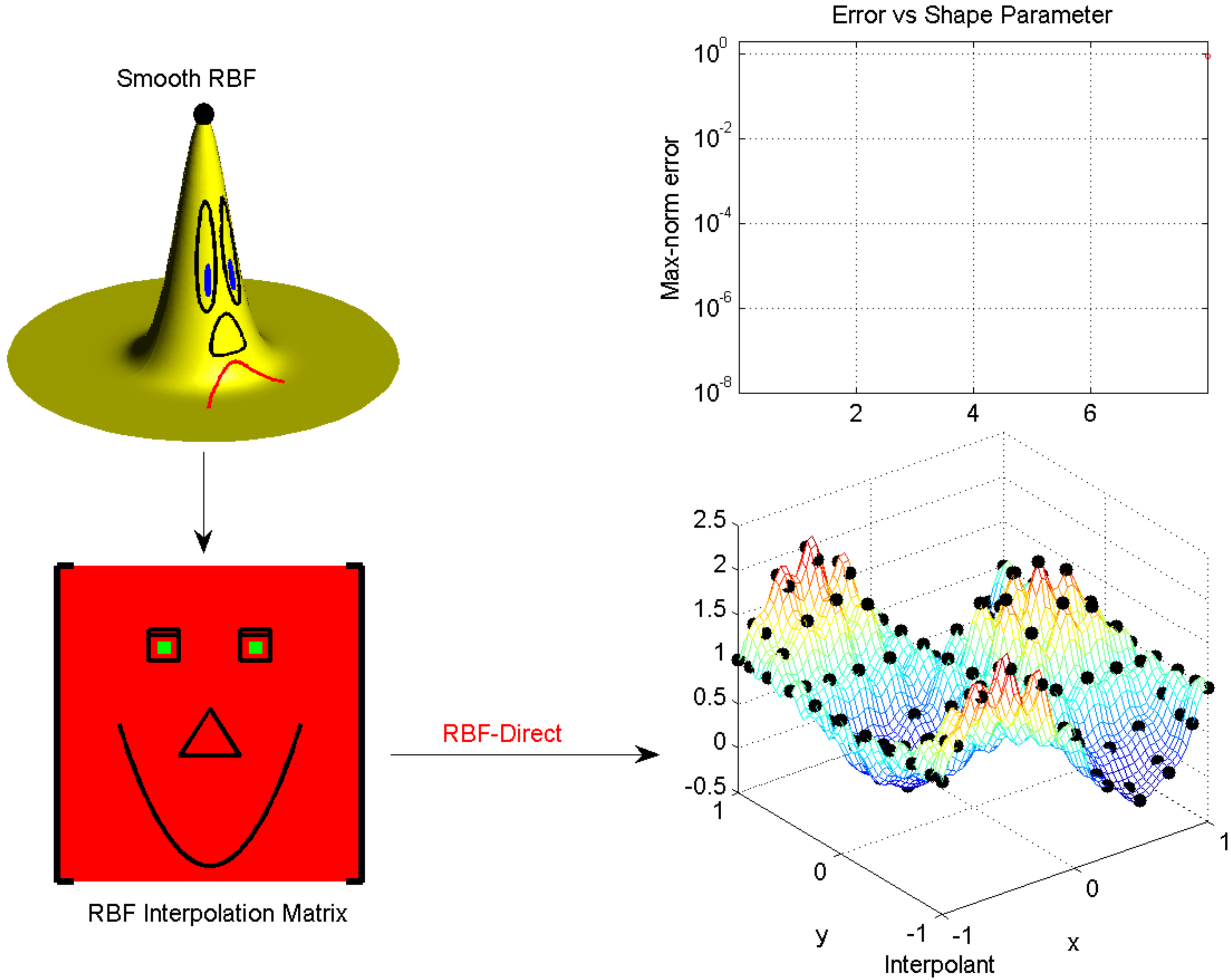

Truncated powers: $(x)^3_+$




Bspline basis: $b_3(x)$

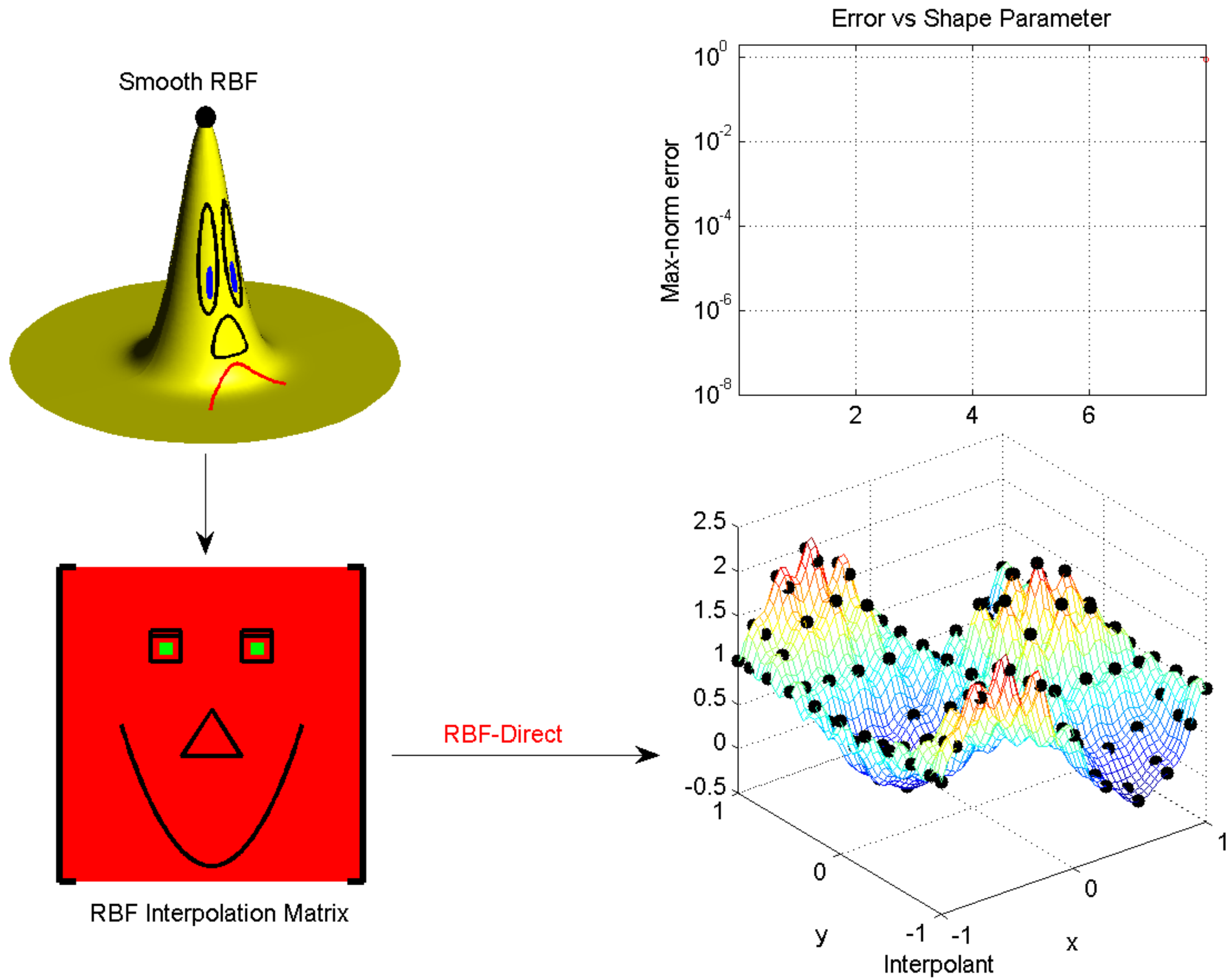
# Using a bad basis for flat kernels:





# Using a good basis for flat kernels:

*Supplementary material*



# Behavior of interpolants in the flat limit

*Supplementary material*

- For cardinal data $f(\mathbf{x}_j) = \begin{cases} 1 & \text{if } j = 1 \\ 0 & \text{if } j \neq 1 \end{cases}$ the interpolant can be written as

$$I_{X,\varepsilon} f(\mathbf{x}) = \frac{\det \begin{bmatrix} \phi_\varepsilon(\|\mathbf{x}-\mathbf{x}_1\|) & \phi_\varepsilon(\|\mathbf{x}-\mathbf{x}_2\|) & \cdots & \phi_\varepsilon(\|\mathbf{x}-\mathbf{x}_N\|) \\ \phi_\varepsilon(\|\mathbf{x}_2-\mathbf{x}_1\|) & \phi_\varepsilon(\|\mathbf{x}_2-\mathbf{x}_2\|) & \cdots & \phi_\varepsilon(\|\mathbf{x}_2-\mathbf{x}_N\|) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_\varepsilon(\|\mathbf{x}_N-\mathbf{x}_1\|) & \phi_\varepsilon(\|\mathbf{x}_N-\mathbf{x}_2\|) & \cdots & \phi_\varepsilon(\|\mathbf{x}_N-\mathbf{x}_N\|) \end{bmatrix}}{\det \begin{bmatrix} \phi_\varepsilon(\|\mathbf{x}_1-\mathbf{x}_1\|) & \phi_\varepsilon(\|\mathbf{x}_1-\mathbf{x}_2\|) & \cdots & \phi_\varepsilon(\|\mathbf{x}_1-\mathbf{x}_N\|) \\ \phi_\varepsilon(\|\mathbf{x}_2-\mathbf{x}_1\|) & \phi_\varepsilon(\|\mathbf{x}_2-\mathbf{x}_2\|) & \cdots & \phi_\varepsilon(\|\mathbf{x}_2-\mathbf{x}_N\|) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_\varepsilon(\|\mathbf{x}_N-\mathbf{x}_1\|) & \phi_\varepsilon(\|\mathbf{x}_N-\mathbf{x}_2\|) & \cdots & \phi_\varepsilon(\|\mathbf{x}_N-\mathbf{x}_N\|) \end{bmatrix}}$$

- Expand determinants using $\phi_\varepsilon(r) = a_0 + a_1\varepsilon^2 r^2 + a_2\varepsilon^4 r^4 + a_3\varepsilon^6 r^6 + \cdots$:

$$I_{X,\varepsilon} f(\mathbf{x}) = \frac{\varepsilon^{2p}\{\text{poly. in } \mathbf{x}\} + \varepsilon^{2(p+1)}\{\text{poly. in } \mathbf{x}\} + \cdots}{\varepsilon^{2q}\{\text{constant}\} + \varepsilon^{2(q+1)}\{\text{constant}\} + \cdots}$$

- In general (and always for GA) $p = q$ so that the $\lim_{\varepsilon \longrightarrow 0} s(\mathbf{x}, \varepsilon)$ exists.

# Behavior of interpolants in the flat limit



- Expand determinants using $\phi_\varepsilon(r) = a_0 + a_1\varepsilon^2 r^2 + a_2\varepsilon^4 r^4 + a_3\varepsilon^6 r^6 + \cdots$:

$$I_{X,\varepsilon} f(\mathbf{x}) = \frac{\varepsilon^{2p}\{\text{poly. in } \mathbf{x}\} + \varepsilon^{2(p+1)}\{\text{poly. in } \mathbf{x}\} + \cdots}{\varepsilon^{2q}\{\text{constant}\} + \varepsilon^{2(q+1)}\{\text{constant}\} + \cdots}$$

- In general (and always for GA) $p = q$ so that the $\lim_{\varepsilon \longrightarrow 0} s(\mathbf{x}, \varepsilon)$ exists.
- Example values of $2p$ and $2q$:

| | $d$ - dimension | $N$ - number of data points | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 2 | 3 | 5 | 10 | 20 | 50 | 100 | 200 |
| Leading power of $\epsilon$ | 1 | 2 | 6 | 20 | 90 | 380 | 2450 | 9900 | 39800 |
| in det($A$) for both | 2 | 2 | 4 | 12 | 40 | 130 | 570 | 1690 | 4940 |
| numer. and denom. | 3 | 2 | 4 | 10 | 30 | 90 | 360 | 980 | 2610 |

- High powers of $\varepsilon$ indicate extreme ill-conditioning.
- Stable algorithms (better bases) are needed to reach the flat limit.

# Uncertainty principle misconception



- Schaback's uncertainty principle:

Principle: *One cannot simultaneously achieve good conditioning and high accuracy*.

Misconception: Accuracy that can be achieved is limited by ill-conditioning.

Restatement:

*One cannot simultaneously achieve good conditioning and high accuracy when using the standard basis.*

- It's a matter of base vs. space.
- Literature for interpolation with "flat" kernels is growing:

Theory:
Driscoll & Fornberg (2002)
Larsson & Fornberg (2003; 2005)
Fornberg, Wright, & Larsson (2004)
Schaback (2005; 2008)
Platte & Driscoll (2005)
Fornberg, Larsson, & Wright (2006)
deBoor (2006)
Fornberg & Zuev (2007)
Lee, Yoon, & Yoon (2007)
Fornberg & Piret (2008)
Buhmann, Dinew, & Larsson (2010)
Platte (2011)
Song, Riddle, Fasshauer, & Hickernell (2011)

Stable algorithms:
Fornberg & Wright (2004)
Fornberg & Piret (2007)
Fornberg, Larsson, & Flyer (2011)
Fasshauer & McCourt (2011)
Gonnet, Pachon, & Trefethen (2011)
Pazouki & Schaback (2011)
De Marchi & Santin (2013)
Fornberg, Letho, Powell (2013)
Wright & Fornberg (2013)

# Better kernel bases for the sphere: RBF-QR algorithm



- Key idea behind the RBF-QR algorithm is to exploit the spherical harmonic expansion of the kernel:

$$\phi_\varepsilon(\|\mathbf{x}-\mathbf{y}\|) = \phi_\varepsilon(\sqrt{2-2\mathbf{x}^T\mathbf{y}}) = \psi_\varepsilon(\mathbf{x}^T\mathbf{y}) = \sum_{\ell=0}^{\infty} \hat{\psi}_\varepsilon(\ell) \sum_{m=-\ell}^{\ell} Y_\ell^m(\mathbf{x}) Y_\ell^m(\mathbf{y})$$

- And use the nice properties of the resulting spherical Fourier coefficients:

| Name | Kernel ($r(t)=\sqrt{2-2t}$) | Fourier coefficients $\hat{\psi}_\varepsilon(\ell)$ ($\varepsilon>0$) |
|---|---|---|
| Gaussian | $\psi(t)=\exp(-(\varepsilon r(t))^2)$ | $\varepsilon^{2\ell}\dfrac{4\pi^{3/2}}{\varepsilon^{2\ell+1}}e^{-2\varepsilon^2}I_{\ell+1/2}(2\varepsilon^2)$ |
| IMQ | $\psi(t)=\dfrac{1}{\sqrt{1+(\varepsilon r(t))^2}}$ | $\varepsilon^{2\ell}\dfrac{4\pi}{(\ell+1/2)}\left(\dfrac{2}{1+\sqrt{4\varepsilon^2+1}}\right)^{2\ell+1}$ |
| MQ | $\psi(t)=-\sqrt{1+(\varepsilon r(t))^2}$ | $\varepsilon^{2\ell}\dfrac{2\pi(2\varepsilon^2+1+(\ell+1/2)\sqrt{1+4\varepsilon^2})}{(\ell+3/2)(\ell+1/2)(\ell-1/2)}\left(\dfrac{2}{1+\sqrt{4\varepsilon^2+1}}\right)^{2\ell+1}$ |

Note how the powers of $\varepsilon$ appear in the coefficients.

- Can redefine the spherical harmonic expansion as:

$$\psi_\varepsilon(\mathbf{x}^T\mathbf{y}) = \sum_{\ell=0}^{\infty} \varepsilon^{2\ell}\tilde{\psi}_\varepsilon(\ell) \sum_{m=-\ell}^{\ell} Y_\ell^m(\mathbf{x}) Y_\ell^m(\mathbf{y}) \quad (\tilde{\psi}_\varepsilon(\ell) = \varepsilon^{-2\ell}\hat{\psi}(\ell))$$

# Better kernel bases for the sphere: RBF-QR algorithm



- For $X = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{S}^2$, we can write each basis function $\psi_\varepsilon(\mathbf{x}^T\mathbf{x}_j)$ as

$$
\begin{aligned}
\psi_\varepsilon(\mathbf{x}^T\mathbf{x}_1) =& \widetilde{\psi}_\varepsilon(0)Y_0^0(\mathbf{x}_1)Y_0^0(\mathbf{x})+ \\
& \varepsilon^2\widetilde{\psi}_\varepsilon(1)\left\{Y_1^{-1}(\mathbf{x}_1)Y_1^{-1}(\mathbf{x}) + Y_1^0(\mathbf{x}_1)Y_1^0(\mathbf{x}) + Y_1^1(\mathbf{x}_1)Y_1^1(\mathbf{x})\right\} + \\
& \varepsilon^4\widetilde{\psi}_\varepsilon(2)\left\{\ldots\ldots\right\} + \varepsilon^6\widetilde{\psi}_\varepsilon(3)\left\{\ldots\ldots\right\} + \varepsilon^8\widetilde{\psi}_\varepsilon(4)\left\{\ldots\ldots\right\} + \ldots \\
\psi_\varepsilon(\mathbf{x}^T\mathbf{x}_2) =& \widetilde{\psi}_\varepsilon(0)Y_0^0(\mathbf{x}_2)Y_0^0(\mathbf{x})+ \\
& \varepsilon^2\widetilde{\psi}_\varepsilon(1)\left\{Y_1^{-1}(\mathbf{x}_2)Y_1^{-1}(\mathbf{x}) + Y_1^0(\mathbf{x}_2)Y_1^0(\mathbf{x}) + Y_1^1(\mathbf{x}_2)Y_1^1(\mathbf{x})\right\} + \\
& \varepsilon^4\widetilde{\psi}_\varepsilon(2)\left\{\ldots\ldots\right\} + \varepsilon^6\widetilde{\psi}_\varepsilon(3)\left\{\ldots\ldots\right\} + \varepsilon^8\widetilde{\psi}_\varepsilon(4)\left\{\ldots\ldots\right\} + \ldots \\
\vdots \qquad & \qquad \vdots \\
\psi_\varepsilon(\mathbf{x}^T\mathbf{x}_N) =& \widetilde{\psi}_\varepsilon(0)Y_0^0(\mathbf{x}_N)Y_0^0(\mathbf{x})+ \\
& \varepsilon^2\widetilde{\psi}_\varepsilon(1)\left\{Y_1^{-1}(\mathbf{x}_N)Y_1^{-1}(\mathbf{x}) + Y_1^0(\mathbf{x}_N)Y_1^0(\mathbf{x}) + Y_1^1(\mathbf{x}_N)Y_1^1(\mathbf{x})\right\} + \\
& \varepsilon^4\widetilde{\psi}_\varepsilon(2)\left\{\ldots\ldots\right\} + \varepsilon^6\widetilde{\psi}_\varepsilon(3)\left\{\ldots\ldots\right\} + \varepsilon^8\widetilde{\psi}_\varepsilon(4)\left\{\ldots\ldots\right\} + \ldots
\end{aligned}
$$

- For simplicity we assume $N$ is a perfect square $(N = n+1)^2$.

# Better kernel bases for the sphere: RBF-QR algorithm

*Supplementary material*

- Or in matrix vector form as

$$\begin{bmatrix} \psi(\mathbf{x}^T\mathbf{x}_1) \\ \psi(\mathbf{x}^T\mathbf{x}_2) \\ \vdots \\ \psi(\mathbf{x}^T\mathbf{x}_N) \end{bmatrix} = BE \begin{bmatrix} Y_0^0(\mathbf{x}) \\ Y_1^{-1}(\mathbf{x}) \\ Y_1^0(\mathbf{x}) \\ Y_1^1(\mathbf{x}) \\ \vdots \end{bmatrix}$$

$$B = \begin{bmatrix} \widetilde{\psi}(0)Y_0^0(\mathbf{x}_1) & \widetilde{\psi}(1)Y_1^{-1}(\mathbf{x}_1) & \widetilde{\psi}(1)Y_1^0(\mathbf{x}_1) & \widetilde{\psi}(1)Y_1^1(\mathbf{x}_1) & \dots \\ \widetilde{\psi}(0)Y_0^0(\mathbf{x}_2) & \widetilde{\psi}(1)Y_1^{-1}(\mathbf{x}_2) & \widetilde{\psi}(1)Y_1^0(\mathbf{x}_2) & \widetilde{\psi}(1)Y_1^1(\mathbf{x}_2) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \widetilde{\psi}(0)Y_0^0(\mathbf{x}_N) & \widetilde{\psi}(1)Y_1^{-1}(\mathbf{x}_N) & \widetilde{\psi}(1)Y_1^0(\mathbf{x}_N) & \widetilde{\psi}(1)Y_1^1(\mathbf{x}_N) & \dots \end{bmatrix}$$

$$E = \begin{bmatrix} 1 & & & & & \\ & \varepsilon^2 & & & & \\ & & \varepsilon^2 & & & \\ & & & \varepsilon^2 & & \\ & & & & \varepsilon^4 & \\ & & & & & \ddots \end{bmatrix}$$

# Better kernel bases for the sphere: RBF-QR algorithm



$$\begin{bmatrix} \psi(\mathbf{x}^T\mathbf{x}_1) \\ \psi(\mathbf{x}^T\mathbf{x}_2) \\ \vdots \\ \psi(\mathbf{x}^T\mathbf{x}_N) \end{bmatrix} = BE \begin{bmatrix} Y_0^0(\mathbf{x}) \\ Y_1^{-1}(\mathbf{x}) \\ Y_1^0(\mathbf{x}) \\ Y_1^1(\mathbf{x}) \\ \vdots \end{bmatrix} \Longleftrightarrow \underline{\psi}(\mathbf{x}) = BE\underline{Y}$$

- We can't deal with infinite matrices so we truncate according to some tolerance (see Fornberg & Piret (2007)).
- Let $\mu > n$ be the truncation degree of the expansion.
- Partition the truncated matrices:

$$B = \begin{bmatrix} B_1 & B_2 \end{bmatrix}$$

where the columns of $B_1$ consist of spherical harmonics up to degree $n$ and $B_2$ consist of spherical harmonics from degree $n+1$ to $\mu$.

$$E = \begin{bmatrix} E_1 & \\ & E_2 \end{bmatrix}$$

where $E_1$ is diagonal with blocks of powers of $\varepsilon$ from 0 to $2n$ and where $E_2$ is diagonal with blocks of powers of $\varepsilon$ from $2(n+1)$ to $2\mu$.

# Better kernel bases for the sphere: RBF-QR algorithm



$$\begin{bmatrix} \psi(\mathbf{x}^T\mathbf{x}_1) \\ \psi(\mathbf{x}^T\mathbf{x}_2) \\ \vdots \\ \psi(\mathbf{x}^T\mathbf{x}_N) \end{bmatrix} = BE \begin{bmatrix} Y_0^0(\mathbf{x}) \\ Y_1^{-1}(\mathbf{x}) \\ Y_1^0(\mathbf{x}) \\ Y_1^1(\mathbf{x}) \\ \vdots \\ Y_\nu^\nu(\mathbf{x}) \end{bmatrix} \Longleftrightarrow \underline{\psi}(\mathbf{x}) = \begin{bmatrix} B_1 & B_2 \end{bmatrix} \begin{bmatrix} E_1 & \\ & E_2 \end{bmatrix} \underline{Y}$$

- Now we change basis, by $QR$ decomposition of $[B_1\ B_2]$ and manipulations with $E$.

$$\begin{aligned} \underline{\psi}(\mathbf{x}) &= QR \begin{bmatrix} E_1 & \\ & E_2 \end{bmatrix} \underline{Y} \\ &= Q \begin{bmatrix} R_1 & R_2 \end{bmatrix} \begin{bmatrix} E_1 & \\ & E_2 \end{bmatrix} \underline{Y} \quad (R_1 \text{ is } N\text{-by-}N \text{ and upper triangular}) \\ &= QR_1 \begin{bmatrix} I & R_1^{-1}R_2 \end{bmatrix} \begin{bmatrix} E_1 & \\ & E_2 \end{bmatrix} \underline{Y} \\ &= QR_1E_1 \underbrace{\begin{bmatrix} I & E_1^{-1}R_1^{-1}R_2E_2 \end{bmatrix} \underline{Y}}_{\text{New stable basis}} \end{aligned}$$

- All negative powers of $\varepsilon$ in $E_1^{-1}R_1^{-1}R_2E_2$ analytically cancel after some matrix manipulations.

# Numerical example

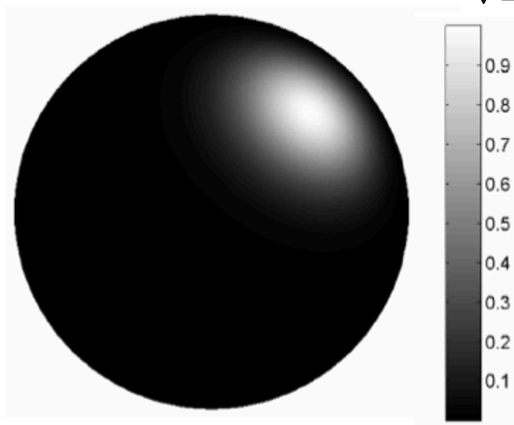


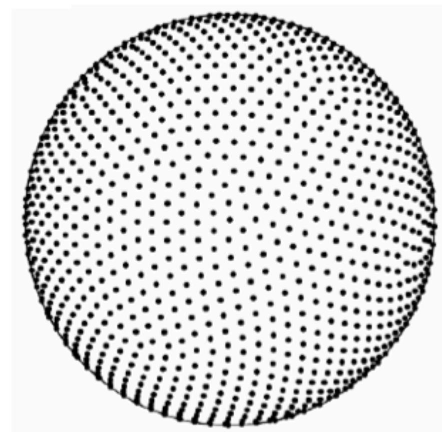
Target function:

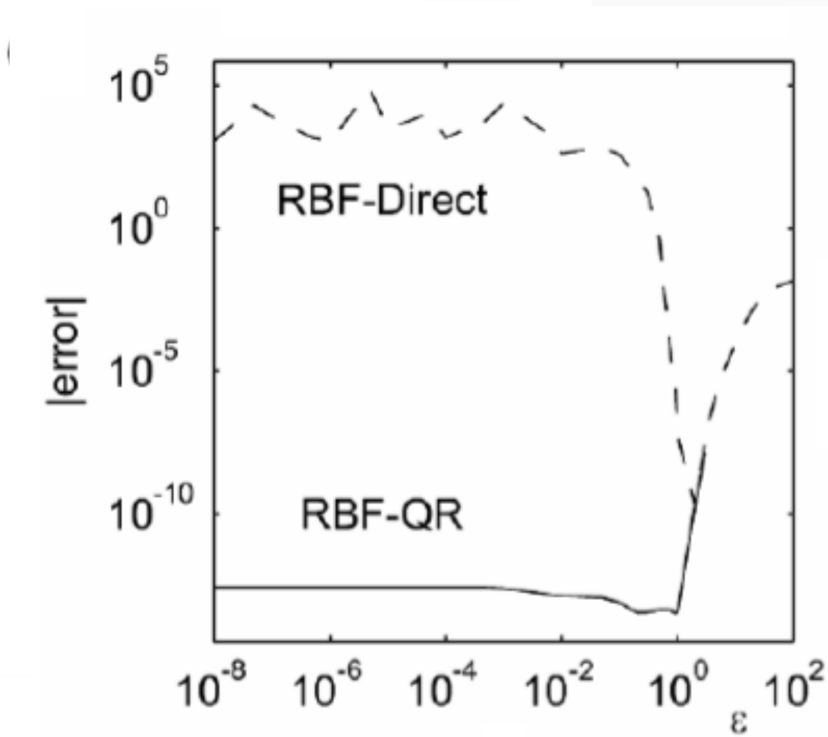
$$f(\mathbf{x}) = \exp(-7(x + \tfrac{1}{2}^{2}) - 8(y + \tfrac{1}{2})^2 - 9(z - \tfrac{1}{\sqrt{2}})^2)$$

Node Set:

$N = 1849$ Minimal energy (ME) nodes





# Remarks on RBF-QR



- RBF-QR allows one to stably compute “flat” kernel interpolants on the sphere.
- One can reach full numerical precision using this procedure (for smooth enough target functions and large enough $N$)
- It is more expensive than standard approach (RBF-Direct).
- Work has gone into extending this idea to general Euclidean space, but the procedure is much more complicated.
- Matlab Code for RBF-QR is given in Fornberg & Piret (2007) and is available in the **rbfsphere** package.

# Concluding remarks



- This was general background material for getting started in this area.
- There is still much more to learn and many interesting problems:
  - Approximation (and decomposition) of vector fields.
  - Fast algorithms for interpolation using localized bases
  - Numerical integration
  - RBF generated finite differences
  - RBF partition of unity methods
  - Numerical solution of partial differential equations on spheres.
  - Generalizations to other manifolds.

❖ If you have any questions or want to chat about research ideas, please come and talk to me.