

# MATH 590: Meshfree Methods

## Alternate Bases

Greg Fasshauer

Department of Applied Mathematics  
Illinois Institute of Technology

Fall 2014



# Outline

- 1 Data-dependent Basis Functions
- 2 Data-Independent Basis Functions



Up until now we have mostly focused on the **positive definite kernel**  $K$  and the generally **infinite-dimensional Hilbert space**  $\mathcal{H}_K(\Omega)$  associated with it.

We now **focus on a specific scattered data fitting problem**, i.e., we fix a **finite set (of data sites)**  $\mathcal{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\} \subset \mathbb{R}^d$  and an **associated data-dependent linear function space**

$$\mathcal{H}_K(\mathcal{X}) = \text{span}\{K(\cdot, \mathbf{x}_1), \dots, K(\cdot, \mathbf{x}_N)\}$$

as **suggested by the Haar–Mairhuber–Curtis theorem** (see Chapter 1).

Then it makes sense to **consider different bases for the finite-dimensional kernel space**  $\mathcal{H}_K(\mathcal{X})$ .

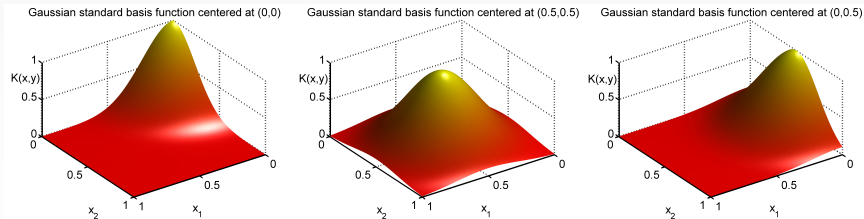
We will also consider **data-independent** (approximate) bases given in terms of the first  $N$  eigenfunctions of the Hilbert–Schmidt integral operator associated with  $K$ .



# “Standard” Basis Functions

$$\{K(\cdot, \mathbf{x}_1), \dots, K(\cdot, \mathbf{x}_N)\}$$

Corresponding system matrix often ill-conditioned



# Matrix-free Methods

Kernel interpolation leads to linear system  $K\mathbf{c} = \mathbf{y}$  with matrix

$$K_{ij} = K(\mathbf{x}_i, \mathbf{x}_j), \quad i, j = 1, \dots, N$$

Goal: **Avoid solution of linear systems**

Use **cardinal functions** in  $\text{span}\{K(\cdot, \mathbf{x}_1), \dots, K(\cdot, \mathbf{x}_N)\}$  such that

$$\hat{u}_j(\mathbf{x}_i) = \delta_{ij}, \quad i, j, \dots, N$$

Then

$$s(\mathbf{x}) = \sum_{j=1}^N y_j \hat{u}_j(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d$$



# Cardinal Basis Functions

Satisfy the **Lagrange property**

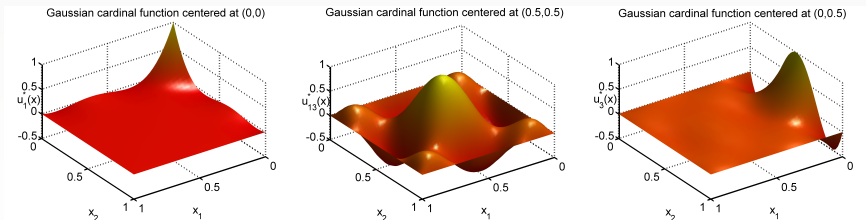
$$\hat{u}_j(\mathbf{x}_i) = \delta_{ij}$$

so that we can find them via (**hard/expensive!**)

$$\mathbf{K} \hat{\mathbf{u}}(\mathbf{x}) = \mathbf{k}(\mathbf{x}),$$

where  $K_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$  and  $\mathbf{k} = (K(\cdot, \mathbf{x}_1), \dots, K(\cdot, \mathbf{x}_N))^T$ .

**System matrix for interpolation would be identity!**



## Remark

- On an *infinite grid* it is possible to use *Fourier transform techniques* such as the Poisson summation formula to obtain closed-form expressions of the cardinal functions.
  - This was done, e.g., in [Buh90, MN90] for *multiquadrics* and *polyharmonic splines*.
  - For the *Gaussian kernel* an infinite cardinal basis was found in [HMNW12], but there is also earlier work in, e.g., [BS96].
- In the more general — *scattered* — *setting* I am aware of only one result for *univariate Gaussian cardinal functions* on  $[-1, 1]$  from [PD05]:

$$\hat{u}_j(\mathbf{x}) = e^{-\epsilon^2((x+1)^2 - (x_j+1)^2)} \prod_{\substack{i=0 \\ i \neq j}}^N \frac{e^{\beta x} - e^{\beta x_i}}{e^{\beta x_j} - e^{\beta x_i}}, \quad j = 0, 1, \dots, N,$$

where  $\beta = \frac{4\epsilon^2}{N}$  and the notation is based on the use of  $N + 1$  data points  $x_0, \dots, x_N$ .

# Newton Kernels [MS09]

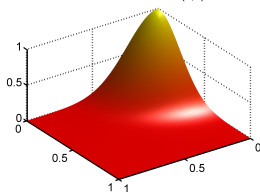
Satisfy the **Newton property**

$$\hat{v}_j(\mathbf{x}_i) = \delta_{ij}, \quad 0 \leq i \leq j \leq N$$

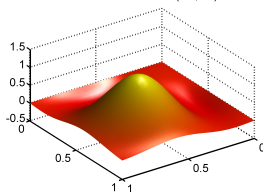
Compute via LU-decomposition of  $K$  [PS11].

Provide **orthogonal basis for native space**. **System matrix is triangular**.

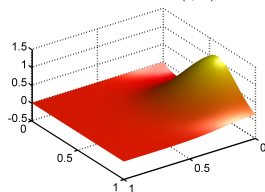
Newton basis function centered at (0,0) for Gaussian



Newton basis function centered at (0.5,0.5) for Gaussian



Newton basis function centered at (0,0.5) for Gaussian





# SVD and Weighted SVD Bases

## Remark

*By computing an SVD of the kernel matrix  $K$  one can obtain a so-called (weighted) SVD basis for the finite-dimensional kernel space  $\mathcal{H}_K(\mathcal{X})$ . This is described in [PS11].*

*A related method — based on a discretization of the Hilbert–Schmidt integral eigenvalue problem is presented in [DMS13].*

*The advantage of these methods (over the Hilbert–Schmidt SVD to be discussed soon) is that they are generic and can be applied to any kind of kernel. However, the resulting bases are usually not as robust (for small  $\varepsilon$ ) as those resulting from the Hilbert–Schmidt SVD.*



# Gaussian Eigenfunctions

Eigenfunctions for the **1D Gaussian kernel**

$$K(x, z) = e^{-\varepsilon^2|x-z|^2}, \quad x, z \in \mathbb{R},$$

were discussed in [ZWRM98] and [RW06] (including online errata).

The **general  $d$ -dimensional case follows immediately from the univariate one via the tensor product form of the Gaussian kernel**, i.e.,

$$K(\mathbf{x}, \mathbf{z}) = e^{-\varepsilon^2\|\mathbf{x}-\mathbf{z}\|_2^2} = e^{-\sum_{\ell=1}^d \varepsilon^2(x_\ell - z_\ell)^2} = \prod_{\ell=1}^d e^{-\varepsilon^2(x_\ell - z_\ell)^2},$$

where  $\mathbf{x} = (x_1, \dots, x_d)$ ,  $\mathbf{z} = (z_1, \dots, z_d) \in \mathbb{R}^d$ , so that

$$K(\mathbf{x}, \mathbf{z}) = \sum_{n \in \mathbb{N}^d} \lambda_n \varphi_n(\mathbf{x}) \varphi_n(\mathbf{z})$$

with

$$\lambda_n = \prod_{\ell=1}^d \lambda_{n_\ell}, \quad \varphi_n(\mathbf{x}) = \prod_{\ell=1}^d \varphi_{n_\ell}(x_\ell).$$



Therefore we concentrate on the univariate eigenfunctions and eigenvalues indexed by  $n = 1, 2, \dots$ . They are given by

$$\varphi_n(x) = \gamma_n e^{-\delta^2 x^2} H_{n-1}(\alpha\beta x), \quad (1)$$

$$\lambda_n = \sqrt{\frac{\alpha^2}{\alpha^2 + \delta^2 + \varepsilon^2}} \left( \frac{\varepsilon^2}{\alpha^2 + \delta^2 + \varepsilon^2} \right)^{n-1}, \quad (2)$$

where the  $H_n$  are Hermite polynomials of degree  $n$ , and

$$\beta = \left( 1 + \left( \frac{2\varepsilon}{\alpha} \right)^2 \right)^{\frac{1}{4}}, \quad \gamma_n = \sqrt{\frac{\beta}{2^{n-1} \Gamma(n)}}, \quad \delta^2 = \frac{\alpha^2}{2} (\beta^2 - 1)$$

are constants defined in terms of  $\varepsilon$  and  $\alpha$ .

The parameter  $\alpha$  determines the weight function

$$\rho(x) = \frac{\alpha}{\sqrt{\pi}} e^{-\alpha^2 x^2} \quad (3)$$

used in the Hilbert–Schmidt integral operator and the associated inner product.



## Remark

*It should be noted that the **eigenfunctions**  $\varphi_n$  are **not** the same as the well-known classical **Hermite functions**, even though there is some similarity.*

*The relative scaling of the arguments of the exponential function and the Hermite polynomials are different in the two cases.*



We now **verify** that

- the **eigenfunctions** are orthonormal with respect to the  $\rho$ -weighted  $L_2$  inner product, and
- the **Hilbert–Schmidt series** sums to the Gaussian kernel.



We need the orthogonality of Hermite polynomials (see, e.g., [AS65, Eqn. (22.2.14)]), i.e.,

$$\int_{-\infty}^{\infty} H_m(x)H_n(x)e^{-x^2} dx = \sqrt{\pi}2^n\Gamma(n+1)\delta_{m,n}.$$

Using the **definition of the eigenfunctions** (1) and of the **weight function** (3), a **substitution**  $t = \alpha\beta x$  gives us

$$\begin{aligned} \int_{-\infty}^{\infty} \varphi_m(x)\varphi_n(x)\rho(x)dx &= \gamma_m\gamma_n \int_{-\infty}^{\infty} H_{m-1}(\alpha\beta x)H_{n-1}(\alpha\beta x)e^{-2\delta^2 x^2} \frac{\alpha}{\sqrt{\pi}}e^{-\alpha^2 x^2} dx \\ &= \frac{\beta}{\sqrt{2^{m-1}\Gamma(m)2^{n-1}\Gamma(n)}} \int_{-\infty}^{\infty} H_{m-1}(\alpha\beta x)H_{n-1}(\alpha\beta x)e^{\alpha^2(1-\beta^2)x^2} \frac{\alpha}{\sqrt{\pi}}e^{-\alpha^2 x^2} dx \\ &= \frac{1}{\sqrt{\pi}\sqrt{2^{m-1}\Gamma(m)2^{n-1}\Gamma(n)}} \int_{-\infty}^{\infty} H_{m-1}(\alpha\beta x)H_{n-1}(\alpha\beta x)e^{-\alpha^2\beta^2 x^2} \alpha\beta dx \\ &= \frac{1}{\sqrt{\pi}\sqrt{2^{m-1}\Gamma(m)2^{n-1}\Gamma(n)}} \int_{-\infty}^{\infty} H_{m-1}(t)H_{n-1}(t)e^{-t^2} dt = \delta_{m,n}, \end{aligned}$$

where the last step uses the orthogonality of the Hermite polynomials.



Verification of the sum of the Mercer series is a bit more involved. The classical result needed here is *Mehler's formula* (see [DLMF12, Eqn. (18.18.28)])

$$\sum_{n=0}^{\infty} \frac{H_n(x)H_n(z)}{2^n \Gamma(n+1)} t^n = \frac{1}{\sqrt{1-t^2}} e^{\frac{2xzt - (x^2+z^2)t^2}{1-t^2}}, \quad |t| < 1. \quad (4)$$

Inserting the eigenfunctions (1) and eigenvalues (2) into the Mercer series we have

$$\begin{aligned} \sum_{n=1}^{\infty} \lambda_n \varphi_n(x) \varphi_n(z) &= \sum_{n=0}^{\infty} \lambda_{n+1} \varphi_{n+1}(x) \varphi_{n+1}(z) \\ &= \sum_{n=0}^{\infty} \lambda_{n+1} \gamma_{n+1}^2 e^{-\delta^2(x^2+z^2)} H_n(\alpha\beta x) H_n(\alpha\beta z) \\ &= e^{-\delta^2(x^2+z^2)} \sum_{n=0}^{\infty} \sqrt{\frac{\alpha^2}{\alpha^2 + \delta^2 + \varepsilon^2}} \left( \frac{\varepsilon^2}{\alpha^2 + \delta^2 + \varepsilon^2} \right)^n \frac{\beta}{2^n \Gamma(n+1)} H_n(\alpha\beta x) H_n(\alpha\beta z) \\ &= e^{-\delta^2(x^2+z^2)} \frac{\alpha\beta}{\varepsilon} \sqrt{\frac{\varepsilon^2}{\alpha^2 + \delta^2 + \varepsilon^2}} \sum_{n=0}^{\infty} \frac{H_n(\alpha\beta x) H_n(\alpha\beta z)}{2^n \Gamma(n+1)} \left( \frac{\varepsilon^2}{\alpha^2 + \delta^2 + \varepsilon^2} \right)^n. \end{aligned}$$



Now we let  $t = \frac{\varepsilon^2}{\alpha^2 + \delta^2 + \varepsilon^2}$  and apply Mehler's formula with  $x \rightarrow \alpha\beta x$  and  $z \rightarrow \alpha\beta z$  to get

$$\begin{aligned} \sum_{n=1}^{\infty} \lambda_n \varphi_n(x) \varphi_n(z) &= e^{-\delta^2(x^2+z^2)} \frac{\alpha\beta}{\varepsilon} \sqrt{\frac{t}{1-t^2}} e^{\frac{2\alpha^2\beta^2 xzt - \alpha^2\beta^2(x^2+z^2)t^2}{1-t^2}} \\ &= \frac{\alpha\beta}{\varepsilon} \sqrt{\frac{t}{1-t^2}} e^{-\frac{\alpha^2\beta^2 t}{1-t^2}(x-z)^2} e^{\frac{\alpha^2\beta^2 t - \alpha^2\beta^2 t^2 - \delta^2(1-t^2)}{1-t^2}(x^2+z^2)} \\ &= e^{-\varepsilon^2(x-z)^2}, \end{aligned}$$

where we have

- combined all of the exponential functions,
- replaced  $2\alpha^2\beta^2 xzt$  by  $-\alpha^2\beta^2(x-z)^2 t + \alpha^2\beta^2(x^2+z^2)t$ , and
- separated into two exponential functions in terms of  $(x-z)^2$  and  $x^2+z^2$ , respectively.





The remaining details for the last step are left open and can be verified (if necessary with a computer algebra system).

They are

$$\frac{\alpha\beta}{\varepsilon} \sqrt{\frac{t}{1-t^2}} = 1,$$

which takes care of both the factor multiplying the exponential functions as well as the exponent of the first exponential function.

The other fact is

$$\alpha^2\beta^2t - \alpha^2\beta^2t^2 - \delta^2(1-t^2) = 0.$$



We now have established that

$$K(x, z) = \sum_{n=1}^{\infty} \lambda_n \varphi_n(x) \varphi_n(z)$$

and

$$\int_{-\infty}^{\infty} \varphi_m(x) \varphi_n(x) \rho(x) dx = \delta_{m,n}.$$

Finally, we verify that the Hilbert–Schmidt eigenvalue problem is satisfied, i.e.,

$$\begin{aligned} \int_{-\infty}^{\infty} K(x, z) \varphi_n(x) \rho(x) dx &= \int_{-\infty}^{\infty} \sum_{m=1}^{\infty} \lambda_m \varphi_m(x) \varphi_m(z) \varphi_n(x) \rho(x) dx \\ &= \sum_{m=1}^{\infty} \lambda_m \varphi_m(z) \int_{-\infty}^{\infty} \varphi_m(x) \varphi_n(x) \rho(x) dx \\ &= \sum_{m=1}^{\infty} \lambda_m \varphi_m(z) \delta_{m,n} \\ &= \lambda_n \varphi_n(z). \end{aligned}$$



## Remark

This *argument holds as soon as we*

- *have the Hilbert–Schmidt series of an arbitrary kernel and*
- *know that the eigenfunctions are  $L_2$ -orthonormal with respect to some weight function  $\rho$ .*

Moreover, the argument *carries over to the multivariate case with arbitrary domains  $\Omega \subseteq \mathbb{R}^d$ .*



# References I

- [AS65] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*, Dover Publications, 1965.
- [BS96] B. J. C. Baxter and N. Sivakumar, *On shifted cardinal interpolation by Gaussians and multiquadrics*, *Journal of Approximation Theory* **87** (1996), no. 1, 36–59.
- [Buh90] M. D. Buhmann, *Multivariate cardinal interpolation with radial-basis functions*, *Constructive Approximation* **6** (1990), no. 3, 225–255.
- [DLMF12] *NIST Digital Library of Mathematical Functions*, <http://dlmf.nist.gov/>, Release 1.0.5 of 2012-10-01, 2012, Online companion to [OLBC10].
- [DMS13] Stefano De Marchi and Gabriele Santin, *A new stable basis for radial basis function interpolation*, *Journal of Computational and Applied Mathematics* **253** (2013), 1–13.



## References II

- [HMNW12] T. Hangelbroek, W. Madych, F. Narcowich, and J. D. Ward, *Cardinal interpolation with Gaussian kernels*, Journal of Fourier Analysis and Applications **18** (2012), no. 1, 67–86.
- [MN90] W. R Madych and S. A Nelson, *Polyharmonic cardinal splines*, Journal of Approximation Theory **60** (1990), no. 2, 141–156.
- [MS09] Stefan Müller and Robert Schaback, *A Newton basis for kernel spaces*, Journal of Approximation Theory **161** (2009), no. 2, 645–655.
- [OLBC10] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark (eds.), *NIST Handbook of Mathematical Functions*, Cambridge University Press, New York, NY, 2010, Print companion to [DLMF12].
- [PD05] Rodrigo B. Platte and Tobin A. Driscoll, *Polynomials and potential theory for Gaussian radial basis function interpolation*, SIAM Journal on Numerical Analysis **43** (2005), no. 2, 750–766.
- [PS11] M. Pazouki and R. Schaback, *Bases for kernel-based spaces*, Journal of Computational and Applied Mathematics **236** (2011), no. 4, 575–588.



## References III

- [RW06] C. E. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, Massachusetts, 2006.
- [ZWRM98] H. Zhu, C. K. Williams, R. J. Rohwer, and M. Morciniec, *Gaussian regression and optimal finite dimensional linear models*, Neural Networks and Machine Learning (C. M. Bishop, ed.), Springer-Verlag, Berlin, 1998.

