Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Approximation of $1/\|x-y\|$ by Exponentials for Wavelet Applications

(revised version: June 2005)

by

Wolfgang Hackbusch

Preprint no.: 15 2005

Approximation of $1/\|x-y\|$ by Exponentials for Wavelet Applications (Short Communication)

Wolfgang Hackbusch
Max-Planck-Institut *Mathematik in den Naturwissenschaften*
Inselstr. 22, D-04103 Leipzig
wh@mis.mpg.de

Abstract

We discuss the approximation of $1/\sqrt{t}$ by exponentials in order to apply it to the treatment of $1/\|x-y\|$. In the case of a wavelet basis, one has in addition the vanishing moment property, which allows to add polynomials without increasing the computational effort. This leads to the question whether an approximation of $1/\sqrt{t}$ by the sum of a polynomial and an exponential part yields an improvement. We show that indeed the approximation error is remarkably reduced. The improvement depends on the interval on which $1/\sqrt{t}$ is approximated.

AMS Subject Classification: 41A50, 65T60, 11L07
Key Words: approximation by exponentials, wavelets

1 Introduction

1.1 Motivation

Integrals involving $2n$-variate integrands $f(\mathbf{x},\mathbf{y}) = f(x_1,\dots,x_n,y_1,\dots,y_n)$ with $n = 3$ appear in many physical problems. In the case of Schrödinger's equation, even higher dimensions $n > 3$ must be treated (cf. [4]). Standard quadrature methods applied to f become costly for higher spatial dimensions. As a remedy, one may try to approximate $f(x_1,\dots,x_n,y_1,\dots,y_n)$ (globally or locally) by a separable sum $\sum_\nu \prod_{j=1}^n F_{\nu,j}(x_j,y_j)$. This allows to separate the spatial dimensions:

$$\iint f(x_1,\dots,x_n,y_1,\dots,y_n)\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} \approx \sum_\nu \prod_{j=1}^n \left\{ \iint F_{\nu,j}(x_j,y_j)\mathrm{d}x_j\mathrm{d}y_j \right\}.$$

An important integrand is $f(\mathbf{x},\mathbf{y}) := \frac{1}{\|\mathbf{x}-\mathbf{y}\|} = 1/\sqrt{\sum_{i=1}^d (x_i-y_i)^2}$ arising from the Newton potential. As shown in more detail in Section 1.4, the first step towards a separable sum for this function is the approximation of $1/\sqrt{t}$ by a sum of exponentials (see Section 1.2).

Certain discretisations by wavelets are in particular appropriate for high dimensional problems (see, e.g., [3] for Schrödinger's equation). Whenever wavelets are used, we recommend a modification of the approximation by a sum of exponentials which is the subject of this notice (see Section 1.3).

1.2 Formulation of the Approximation Problem

The approximation of a function like $f(t) = 1/\sqrt{t}$ by polynomials (optimal with respect to the maximum norm in $[1,R]$) is a classical subject in approximation theory, as is the approximation of f by a sum

$$E_k(t) := \sum_{\ell=1}^k \omega_\ell e^{-\alpha_\ell t} \qquad (\omega_\ell,\alpha_\ell \in \mathbb{R} \text{ parameters to be optimised}) \tag{1.1}$$

of exponentials. The set of all function E_k from (1.1) is denoted by $\mathcal{E}_k$. In the latter case, it is usual to replace $\mathcal{E}_k$ by the closure

$$\mathcal{E}_k^* = \left\{ \varphi : \varphi(t) = \sum_{\ell=1}^{k'} p_\ell(t) e^{-\alpha_\ell t},\ k' + \sum_{\ell=1}^{k'} \text{degree}(p_\ell) = 2k \right\}$$

involving polynomials as factors (cf. [1]). However, the approximation problem

$$\text{minimise} \quad \|f - (P_{q,k} + E_{q,k})\|_{\infty,[1,R]} \quad \text{over all polynomials } P_{q,k} \in \mathcal{P}_q \text{ and all } E_{q,k} \in \mathcal{E}_k, \tag{1.2}$$

where $\mathcal{P}_q$ denotes the set of polynomials of degree $\le q$, is rather unusual.

In fact, if only the number $q + 1 + 2k$ of free parameters is of interest, it is advantageous to spend all degrees of freedom on the exponential part. This is illustrated by the following table, which shows the best approximation error of $f(t) = 1/\sqrt{t}$ in $[1, R]$ for $R = 10$, 100, 1000, 10000 and different combinations of k and q with $q + 1 + 2k = 8$ degrees of freedom.

k	4	3	2	1	0
q	-1	1	3	5	7
$R = 10$	2.856E-05	3.214E-05	5.920E-05	1.956E-04	1.531E-03
$R = 100$	1.091E-03	1.237E-03	2.354E-03	8.014E-03	6.153E-02
$R = 1000$	4.016E-03	4.625E-03	9.464E-03	3.213E-02	2.169E-01
$R = 10000$	6.631E-03	7.931E-03	1.770E-02	5.846E-02	3.510E-01

The left column with $q = -1$ indicates no polynomial part (zero polynomial) and only exponential terms, while the right-most column corresponds to polynomial approximation only.

1.3 Wavelet Context

If the basis of the Galerkin method is built by a uni-variate wavelet basis $\mathcal{B}$, scalar products of the form

$$(f, \psi_\alpha)_{L^2(\mathbb{R})}$$

appear, where $\psi_\alpha \in \mathcal{B}$ is some wavelet basis function and f is some fixed function (see §1.4 for a particular application).

The number M of so-called *vanishing moments*[1] is defined by the largest M such that

$$(x^\nu, \psi_\alpha)_{L^2(\mathbb{R})} = 0 \qquad \text{for all } \nu = 0, 1, \ldots, M-1$$

and all wavelets $\psi_\alpha \in \mathcal{B}$. The multi-variate is discussed in §1.4.

Let $F := P + E$ be an approximation of f with $P \in \mathcal{P}_{M-1}$:

$$f = P + E + \varepsilon,$$

where $\varepsilon(t)$ is the approximation error. Due to the vanishing moment property, we have

$$(f, \psi_\alpha)_{L^2(\mathbb{R})} = (P, \psi_\alpha)_{L^2(\mathbb{R})} + (E, \psi_\alpha)_{L^2(\mathbb{R})} + (\varepsilon, \psi_\alpha)_{L^2(\mathbb{R})} = (E, \psi_\alpha)_{L^2(\mathbb{R})} + (\varepsilon, \psi_\alpha)_{L^2(\mathbb{R})} \approx (E, \psi_\alpha)_{L^2(\mathbb{R})}\,.$$

Since the polynomial part has vanished, the cost of the computational scheme depends only on $2k$, the number of parameters in the exponential part E. But due to the polynomial part P, the error term ε is smaller than without P.

As a consequence, the mixed approach $f = P + E + \varepsilon$ is of interest as long as $degree(P) \le M - 1$.

1.4 Particular Application

Let $P_{q,k}^*$ and $E_{q,k}^*$ be the optimal expressions in (1.2) and define the error function $\varepsilon_{q,k}$ by means of

$$1/\sqrt{t} = P_{q,k}^* + E_{q,k}^* + \varepsilon_{q,k}. \tag{1.3}$$

Replacing t by $\sum_{i=1}^d (x_i - y_i)^2 = \|\mathbf{x} - \mathbf{y}\|^2$ where $\mathbf{x} = (x_i)_{i=1}^d$ and $\mathbf{y} = (y_i)_{i=1}^d \in \mathbb{R}^d$, we get

[1] In the case of biorthogonal wavelets there are two families of wavelets with possibly different vanishing moments M, M^*.

$$\frac{1}{\|\mathbf{x}-\mathbf{y}\|} = \frac{1}{\sqrt{\sum_{i=1}^d (x_i-y_i)^2}} = \underbrace{P^*_{q,k}\left(\sum_{i=1}^d (x_i-y_i)^2\right)}_{Q_{q,k}(\mathbf{x},\mathbf{y})} + E^*_{q,k}\left(\sum_{i=1}^d (x_i-y_i)^2\right) + \varepsilon_{q,k}.$$

The hereby defined $Q_{q,k}(\mathbf{x},\mathbf{y})$ is a polynomial of total degree $\le 2q$ in $\mathbf{x},\mathbf{y}\in\mathbb{R}^d$. $Q_{q,k}$ allows the representation

$$Q_{q,k}(\mathbf{x},\mathbf{y}) = \sum_j \prod_{i=1}^d q_i^{I,j}(x_i) q_i^{II,j}(y_i) \qquad \text{with} \quad \sum_{i=1}^d degree(q_i^{I,j}) + \sum_{i=1}^d degree(q_i^{II,j}) \le 2q.$$

The standard approach to *multi-variate* wavelets are tensor-wavelets (cf. [5, §2.5.1]). For instance, in 2D the uni-variate wavelet ψ and scaling function φ are combined to the three functions $\varphi(x_1)\psi(x_2)$, $\psi(x_1)\varphi(x_2)$, $\psi(x_1)\psi(x_2)$. Together with the usual shifts and dilatations of both factors we obtain the 2D tensor-wavelets. In general, the d-dimensional tensor-wavelet contains at least one factor ψ. We abbreviate the basis functions by $b_\alpha(\mathbf{x}) = b_{\alpha_1}(x_1)\cdots b_{\alpha_d}(x_d)$ with $\alpha = (\alpha_1,\ldots,\alpha_d)$, where the uni-variate factors b_{α_j} are either ψ_{α_j} or φ_{α_j}.

The Galerkin discretisation of the Newton potential requires the coefficients

$$\iint \frac{b_\alpha(\mathbf{x}) b_\beta(\mathbf{y})}{\|\mathbf{x}-\mathbf{y}\|}\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y},$$

which in the case of $\mathrm{supp}(b_\alpha)\cap\mathrm{supp}(b_\beta)\neq\emptyset$ are approximated by the sum of $\iint Q_{q,k}(\mathbf{x},\mathbf{y}) b_\alpha(\mathbf{x}) b_\beta(\mathbf{y})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y}$ and $\iint E^*_{q,k}(\mathbf{x},\mathbf{y}) b_\alpha(\mathbf{x}) b_\beta(\mathbf{y})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y}$.

Remark 1.1 *Let $\gamma\in[2,2d]$ be the number of ψ-factors in b_α and b_β together. Then the inequality*

$$\frac{2q}{\gamma} \le M-1 \qquad (M\text{: number of vanishing moments}) \tag{1.4}$$

guarantees

$$\iint Q_{q,k}(\mathbf{x},\mathbf{y})\, b_\alpha(\mathbf{x})\, b_\beta(\mathbf{y})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} = 0. \tag{1.5}$$

Proof. Set $J^I := \{j\in\{1,\ldots,d\} : b_{\alpha_j} = \psi_{\alpha_j}\}$ and $J^{II} := \{j\in\{1,\ldots,d\} : b_{\beta_j} = \psi_{\beta_j}\}$. By definition $\#J^I + \#J^{II} = \gamma$ must be at least 2. The representation $Q_{q,k}(\mathbf{x},\mathbf{y}) = \sum_j \prod_{i=1}^d q_i^{I,j}(x_i) q_i^{II,j}(y_i)$ yields

$$\iint Q_{q,k}(\mathbf{x},\mathbf{y}) b_\alpha(\mathbf{x}) b_\beta(\mathbf{y})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} = \sum_j \prod_{i=1}^d \left(\int q_i^{I,j}(x_i) b_{\alpha_i}(x_i)\mathrm{d}x_i\right)\left(\int q_i^{II,j}(y_i) b_{\beta_i}(y_i)\mathrm{d}y_i\right).$$

Since $\sum_{i\in J^I} degree(q_i^{I,j}) + \sum_{i\in J^{II}} degree(q_i^{II,j}) \le 2q$, one of the γ polynomials must have a degree $\le 2q/\gamma$. Hence $2q/\gamma \le M-1$ proves that one of the integrals on the right-hand side vanishes, proving the remark. ∎

Furthermore,

$$E^*_{q,k}\left(\sum_{i=1}^d (x_i-y_i)^2\right) = \sum_{\ell=1}^k \omega_\ell \exp\left(-\alpha_\ell \sum_{i=1}^d (x_i-y_i)^2\right) = \sum_{\ell=1}^k \omega_\ell \prod_{i=1}^d e^{-\alpha_\ell (x_i-y_i)^2}$$

allows a simple separation of the spatial directions. It remains to approximate the one-dimensional integrals

$$\iint e^{-\alpha_\ell (x_i-y_i)^2} b_{\alpha_i}(x_i)\, b_{\beta_i}(x_i)\,\mathrm{d}x_i\,\mathrm{d}y_i \qquad (\ell=1,\ldots,k). \tag{1.6}$$

As a consequence of (1.4), two strategies are possible:

A) Given a fixed number M of vanishing moments, one provides an approximation (1.3) with $q := M-1$. Then (1.4) and thus (1.5) hold in all cases.

B) For all $q\in\left\{\left\lfloor\frac{\gamma}{2}(M-1)\right\rfloor : 2\le\gamma\le 2d\right\}$ one provides approximations (1.3), which are applied depending on the value of γ. The possibly larger q improves the accuracy; however, the exponents α_ℓ in (1.6) depend on the choice of q.

2 Behaviour of the Approximation Errors

2.1 Role of the Interval $[1, R]$

In the applications, the approximation of $\frac{1}{\|\mathbf{x}-\mathbf{y}\|}$ is needed for $\mathbf{x} \in X$ and $\mathbf{y} \in Y$, where $X, Y \subset \mathbb{R}^d$ are compact sets with $X \cap Y = \emptyset$. They define the quantities

$$d_{\min} := \operatorname{dist}(X, Y), \qquad d_{\max} := \max\{\|\mathbf{x}-\mathbf{y}\| : \mathbf{x} \in X,\ \mathbf{y} \in Y\}.$$

Therefore, the function $1/\|\mathbf{x}-\mathbf{y}\|$ is to be approximated for $d_{\min} \le \|\mathbf{x}-\mathbf{y}\| \le d_{\max}$. Correspondingly, $f(t) = 1/\sqrt{t}$ is to be approximated on the interval $\left[d_{\min}^2, d_{\max}^2\right]$. By a simple scaling argument, the latter interval may be replaced by $\left[1, d_{\max}^2/d_{\min}^2\right]$. This is, why in the following, we discuss $f(t) = 1/\sqrt{t}$ on the interval $[1, R]$, where $R = d_{\max}^2/d_{\min}^2$ gives the connection to the application problem.

2.2 Approximation for a Moderate R

Let $R = 10$ and $k = 5$ (number of terms in (1.1)). The pure exponential approximation $1/\sqrt{t} \approx E_{-1,k}^*$ uses the factors α_ℓ in the exponents of (1.1), which are given in the column $q = -1$ of the following table:

q	-1	3	7
α_1	0.028248	0.43109	1.00270
α_2	0.267270	1.04287	1.93305
α_3	0.829100	2.04404	3.29347
α_4	1.960301	3.71973	5.38334
α_5	4.329241	6.74409	8.88172

The first values $\alpha_1 = 0.028$ and $\alpha_2 = 0.267$ are comparably small. Therefore, one expects that, e.g., the first two terms $\omega_1 e^{-\alpha_1 t} + \omega_2 e^{-\alpha_2 t}$ containing four parameters can be well approximated by a polynomial of degree 3. Indeed, the column $q = 3$ shows that the exponential term in the approximation $1/\sqrt{t} \approx P_{3,k}^* + E_{3,k}^*$ starts with the larger exponential coefficient $\alpha_1 = 0.431$.

The best approximation errors $\|1/\sqrt{t} - P_{q,k}^* + E_{q,k}^*\|_{\infty,[1,R]}$ ($R = 10$, $k = 5$) for different q are given below:

q	-1	0	1	2	3	4	5	6	7	8
$k = 5$	1.96E-6	5.14E-7	1.46E-7	4.41E-8	1.41E-8	4.67E-9	1.61E-9	5.70E-10	2.08E-10	7.74E-11

Note that in the practical application, the choice of q is restricted by the number of vanishing moments (see Remark 1.1).

We add the corresponding values for $k = 1$ and $k = 2$ (and $R = 10$):

q	-1	0	1	2	3	4	5	6	7	8
$k = 1$	9.06E-2	2.35E-2	7.64E-3	2.81E-3	1.11E-3	4.58E-4	1.96E-4	8.56E-5	3.82E-5	1.73E-5
$k = 2$	6.10E-3	1.60E-3	4.86E-4	1.64E-4	5.92E-5	2.25E-5	8.85E-6	3.59E-6	1.49E-6	6.31E-7

Obviously, the accuracy can be strongly improved by the additional term $P_{q,k}^*$. For $k = 5$, the increase of the degree of $P_{q,k}^*$ by one yields a factor of about 0.32.

2.3 Case of Large R

For large R, say $R = 10\,000$, it seems questionable whether polynomials are helpful. Since $f(t) = 1/\sqrt{t}$ as well as the exponentials terms in (1.1) decrease with t, the polynomial growth of $P_{q,k}$ may be counterproductive. However, a look at the values of the exponential factors α_i in the next table shows that (for $q = -1$, i.e., no polynomial part) the value α_1 is so small that $\exp(\alpha_1 R) = \exp(-1.5) = 0.22$. Hence, this term can be approximated by a polynomial. Although the next coefficient $\alpha_2 = 0.00388$ seems to be small too, the calculation $\exp(\alpha_2 R) = 1.41_{10}\text{-}17$ shows that this term is hard to replace by polynomials. Correspondingly,

the value $\alpha_1 = 0.0038$ for $q = 3$ is close to the previous value of $\alpha_2 = 0.00388$.

q	-1	3	7
α_1	0.00015	0.0038266	0.0095261
α_2	0.00388	0.0236289	0.0443381
α_3	0.03772	0.1150225	0.1762565
α_4	0.24531	0.4779929	0.6282044
α_5	1.25838	1.7881980	2.0869439

Nevertheless, the following tables show an improvement of the approximation error beyond $q = 1$.

	$k = 1$									
R	$q = -1$	0	1	2	3	4	5	6	7	8
100	1.40E-1	6.76E-2	3.84E-2	2.44E-2	1.64E-2	1.13E-2	8.01E-3	5.76E-3	4.20E-3	3.09E-3
1000	1.40E-1	8.62E-2	6.52E-2	5.29E-2	4.41E-2	3.74E-2	3.21E-2	2.78E-2	2.43E-2	2.13E-2
10000	1.40E-1	9.25E-2	8.07E-2	7.33E-2	6.75E-2	6.26E-2	5.85E-2	5.48E-2	5.15E-2	4.86E-2

	$k = 2$									
R	$q = -1$	0	1	2	3	4	5	6	7	8
100	3.00E-2	1.31E-2	6.73E-3	3.85E-3	2.35E-3	1.50E-3	9.89E-4	6.66E-4	4.57E-4	3.18E-4
1000	4.09E-2	2.40E-2	1.63E-2	1.22E-2	9.46E-3	7.58E-3	6.18E-3	5.11E-3	4.27E-3	3.60E-3
10000	4.09E-2	2.84E-2	2.32E-2	2.00E-2	1.77E-2	1.59E-2	1.43E-2	1.31E-2	1.20E-2	1.10E-2

	$k = 5$									
R	$q = -1$	0	1	2	3	4	5	6	7	8
100	2.09E-4	9.17E-5	4.37E-5	2.22E-5	1.19E-5	6.64E-6	3.83E-6	2.28E-6	1.38E-6	8.59E-7
1000	1.22E-3	6.70E-4	4.04E-4	2.61E-4	1.77E-4	1.25E-4	9.10E-5	6.78E-5	5.15E-5	3.97E-5
10000	2.58E-3	1.62E-3	1.12E-3	8.38E-4	6.58E-4	5.34E-4	4.42E-4	3.72E-4	3.18E-4	2.74E-4

The coefficients of the polynomials $P^*_{q,k} = \sum_{i=0}^{q} \beta_i x^i$ become very small with increasing $R : \beta_i = \mathcal{O}(R^{-i})$. However, the polynomial part is never needed for the computations (except for the calculation of the optimal approximation itself).

3 Conclusion

In the wavelet context one should replace the approximation by exponentials with the combination of the same number of exponential terms plus a polynomial such that in the end this polynomial part is deleted due to the vanishing moment property. Then the amount of arithmetical work is not increased but the accuracy is improved. This improvement is substantial for moderate size of R and still remarkable for large R.

The web page `www.mis.mpg.de/scicomp/EXP_SUM/1_sqrtx/tabelle` contains the coefficients of the optimal approximations $E^*_{q,k}$ for various k, q, and R.

References

[1] D. Braess: *Nonlinear Approximation Theory*. Springer-Verlag, Berlin, 1986.

[2] D. Braess and W. Hackbusch: *Approximation of* $1/x$ *by exponential sums in* $[1, \infty)$. To appear in IMA Numer. Anal.

[3] H.-J. Flad, W, Hackbusch, D. Kolb, and R. Schneider: *Wavelet approximation of correlated wavefunctions. I. Basics*. J. Chemical Physics, 116: 9641–9657, 2002.

[4] W. Hackbusch: *The efficient computation of certain determinants arising in the treatment of Schrödinger's equation*. Computing, 67: 35–56, 2001.

[5] A.K. Louis, P. Maaß, and A. Rieder: *Wavelets*. Teubner-Verlag, Stuttgart, 1998.