



# On the expansion of solutions of Laplace-like equations into traces of separable higher dimensional functions

Harry Yserentant<sup>1</sup>

Received: 8 October 2018 / Revised: 14 November 2019 / Published online: 31 July 2020  
© The Author(s) 2020

## Abstract

This paper deals with the equation  $-\Delta u + \mu u = f$  on high-dimensional spaces  $\mathbb{R}^m$  where  $\mu$  is a positive constant. If the right-hand side  $f$  is a rapidly converging series of separable functions, the solution  $u$  can be represented in the same way. These constructions are based on approximations of the function  $1/r$  by sums of exponential functions. The aim of this paper is to prove results of similar kind for more general right-hand sides  $f(x) = F(Tx)$  that are composed of a separable function on a space of a dimension  $n$  greater than  $m$  and a linear mapping given by a matrix  $T$  of full rank. These results are based on the observation that in the high-dimensional case, for  $\omega$  in most of the  $\mathbb{R}^n$ , the euclidian norm of the vector  $T^t \omega$  in the lower dimensional space  $\mathbb{R}^m$  behaves like the euclidian norm of  $\omega$ .

**Mathematics Subject Classification** 41A25 · 41A63 · 65N99

## 1 Introduction

The approximation of high-dimensional functions, whether they be given explicitly or implicitly as solutions of differential equations, represents a grand challenge for applied mathematics. High-dimensional problems arise in many fields of application such as data analysis and statistics, but first of all in the natural sciences. The Schrödinger equation, which links chemistry to physics and describes a system of electrons and nuclei that interact by Coulomb attraction and repulsion forces, forms an important example. The present work is partly motivated by applications in the context of quantum theory and is devoted to the equation

$$-\Delta u + \mu u = f \quad (1.1)$$

---

✉ Harry Yserentant  
yserentant@math.tu-berlin.de

<sup>1</sup> Technische Universität Berlin, Institut für Mathematik, 10623 Berlin, Germany

on  $\mathbb{R}^m$  for high dimensions  $m$ , with  $\mu > 0$  a given constant. Provided the right-hand side  $f$  of this equation possesses an integrable Fourier transform,

$$u(x) = \left( \frac{1}{\sqrt{2\pi}} \right)^m \int \frac{1}{\mu + \|\omega\|^2} \widehat{f}(\omega) e^{i\omega \cdot x} d\omega \quad (1.2)$$

is a solution of this equation, and the only solution that tends uniformly to zero as  $x$  goes to infinity. If the right-hand side  $f$  of the equation is a tensor product

$$f(x) = \prod_i \phi_i(x_i) \quad (1.3)$$

of univariate functions or a rapidly converging series of such tensor products, the same holds for the Fourier transform of  $f$ . If one replaces the corresponding term in the high-dimensional integral (1.2) by an approximation

$$\frac{1}{r} \approx \sum_k a_k e^{-\beta_k r}, \quad \frac{1}{\mu + \|\omega\|^2} \approx \sum_k a_k e^{-\beta_k \mu} \prod_i e^{-\beta_k \omega_i^2}, \quad (1.4)$$

the integral collapses in this case therefore to a sum of products of one-dimensional integrals. That is, the solution of the equation can, independent of the space dimension, be approximated by a finite or infinite sum of such tensor products. Usually one starts from approximations of  $1/r$  of given absolute accuracy. Such approximations of  $1/r$  are studied in [1] and [2] and result in error estimates in terms of the right-hand side of the equation. By reasons that will become clear later, we will focus in the present paper on approximations of  $1/r$  of given relative accuracy. In the context here, they lead to error estimates in terms of the solution of the equation itself. An example of such an approximation in form of an infinite series is

$$\frac{1}{r} \approx h \sum_{k=-\infty}^{\infty} e^{kh} \exp(-e^{kh} r). \quad (1.5)$$

It arises from an integral representation of  $1/r$  that is discretized by the trapezoidal or midpoint rule. It has been analyzed in [9, Sect. 5] and is extremely accurate. The relative error tends exponentially with the distance  $h$  of the nodes to zero. It is less than  $5 \cdot 10^{-8}$  for  $h = 1/2$ , and for  $h = 1$  still less than  $7 \cdot 10^{-4}$ . The approximation properties of partial sums of this series will be studied later.

The conclusion is that a tensor product structure like (1.3) of the right-hand side of the equation is directly reflected in its solution. This effect enables the solution of truly high-dimensional equations [4,6] and probably forms one of the bases [3] for the success of modern tensor product methods [5]. The aim of the present paper is to generalize this observation to right-hand sides

$$f(x) = F(Tx) \quad (1.6)$$

that are composed of a separable function  $F$  on a space of a dimension  $n$  greater than the original dimension  $m$  and a linear mapping given by a matrix  $T$  of full rank. This covers, for example, right-hand sides  $f$  that depend explicitly on differences of the components of  $x$ . We will prove that the solution  $u$  of the Eq. (1.1) can in such cases almost always be well approximated by finite sums of functions of the same type, provided the ratio  $n/m$  of the dimensions does not become too large. Background is some kind of concentration of measure effect in high space dimensions. Our main tool is the representation  $u(x) = U(Tx)$  of the solution in terms of the solution  $U$  of a degenerate elliptic equation  $\mathcal{L}U = F$  in the higher dimensional space. Approximations to  $U$  are then iteratively generated.

## 2 Functions with integrable Fourier transform and their traces

We consider in this paper functions  $u : \mathbb{R}^d \rightarrow \mathbb{R}$ ,  $d$  a varying and potentially high dimension, that possess a then also unique representation

$$u(x) = \left( \frac{1}{\sqrt{2\pi}} \right)^d \int \widehat{u}(\omega) e^{i\omega \cdot x} d\omega \quad (2.1)$$

in terms of a function  $\widehat{u} \in L_1(\mathbb{R}^d)$ , their Fourier transform. Such functions are uniformly continuous and tend uniformly to zero as  $x$  goes to infinity, the Riemann–Lebesgue theorem. The space  $W_0(\mathbb{R}^d)$  of these functions becomes under the norm

$$\|u\| = \left( \frac{1}{\sqrt{2\pi}} \right)^d \int |\widehat{u}(\omega)| d\omega \quad (2.2)$$

a Banach space and even a Banach algebra. The norm (2.2) dominates the maximum norm of the functions in this space. If the functions  $(i\omega)^\beta \widehat{u}(\omega)$ ,  $\beta \leq \alpha$ , in multi-index notation, are integrable as well, the partial derivative  $D^\alpha u$  of  $u$  exists, is given by

$$(D^\alpha u)(x) = \left( \frac{1}{\sqrt{2\pi}} \right)^d \int (i\omega)^\alpha \widehat{u}(\omega) e^{i\omega \cdot x} d\omega, \quad (2.3)$$

and is as the function  $u$  itself uniformly continuous and vanishes as  $u$  at infinity. For partial derivatives of first order, this follows from the Fourier representation of the corresponding difference quotients and the dominated convergence theorem, and, for derivatives of higher order, it follows by induction.

Let  $T$  be a from now on fixed  $(n \times m)$ -matrix of full rank  $m < n$  and let

$$u : \mathbb{R}^m \rightarrow \mathbb{R} : x \rightarrow U(Tx) \quad (2.4)$$

be the trace of a function  $U : \mathbb{R}^n \rightarrow \mathbb{R}$  with an integrable Fourier transform. We first calculate the Fourier transform of such trace functions.

**Theorem 2.1** *Let  $U : \mathbb{R}^m \times \mathbb{R}^{n-m} \rightarrow \mathbb{R}$  be a function with an integrable Fourier transform. Its trace function (2.4) possesses then an integrable Fourier transform as well. It reads, in terms of the Fourier transform of  $U$ ,*

$$\widehat{u}(\omega) = \left(\frac{1}{\sqrt{2\pi}}\right)^{n-m} \int \frac{1}{|\det \widetilde{T}|} \widehat{U}\left(\widetilde{T}^{-t}\left(\begin{smallmatrix} \omega \\ \eta \end{smallmatrix}\right)\right) d\eta, \quad (2.5)$$

where  $\widetilde{T}$  is an arbitrary invertible matrix of dimension  $n \times n$  whose first  $m$  columns coincide with those of  $T$ . The norm (2.2) of the trace function satisfies the estimate

$$\|u\| \leq \|U\| \quad (2.6)$$

in terms of the corresponding norm of the function  $U$ .

**Proof** The scaled  $L_1$ -norm

$$\left(\frac{1}{\sqrt{2\pi}}\right)^m \int |\widehat{u}(\omega)| d\omega = \left(\frac{1}{\sqrt{2\pi}}\right)^n \int \left| \int \frac{1}{|\det \widetilde{T}|} \widehat{U}\left(\widetilde{T}^{-t}\left(\begin{smallmatrix} \omega \\ \eta \end{smallmatrix}\right)\right) d\eta \right| d\omega$$

of the function (2.5) remains, by Fubini's theorem and the transformation theorem for multivariate integrals, finite and satisfies the estimate

$$\left(\frac{1}{\sqrt{2\pi}}\right)^m \int |\widehat{u}(\omega)| d\omega \leq \left(\frac{1}{\sqrt{2\pi}}\right)^n \int \left| \widehat{U}\left(\begin{smallmatrix} \omega \\ \eta \end{smallmatrix}\right) \right| d(\omega, \eta).$$

That the function (2.5) is the Fourier transform of the trace function follows from

$$\omega \cdot x = \widetilde{T}^{-t}\left(\begin{smallmatrix} \omega \\ \eta \end{smallmatrix}\right) \cdot \widetilde{T}\left(\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right) = \widetilde{T}^{-t}\left(\begin{smallmatrix} \omega \\ \eta \end{smallmatrix}\right) \cdot T x$$

and again Fubini's theorem and the transformation theorem.  $\square$

The estimate (2.6) is sharp. Every function  $u$  in  $W_0(\mathbb{R}^m)$  is trace of a function  $U$  in  $W_0(\mathbb{R}^n)$  with norm  $\|U\| = \|u\|$ , for example of that with the Fourier transform

$$\widehat{U}\left(\begin{smallmatrix} \omega \\ \eta \end{smallmatrix}\right) = |\det \widetilde{T}| V\left(\widetilde{T}^t\left(\begin{smallmatrix} \omega \\ \eta \end{smallmatrix}\right)\right), \quad V\left(\begin{smallmatrix} \omega \\ \eta \end{smallmatrix}\right) = \widehat{u}(\omega) e^{-\frac{1}{2}\|\eta\|^2}. \quad (2.7)$$

A consequence of Theorem 2.1 is that the traces of functions with Fourier transform vanishing outside of a strip around the kernel of  $T^t$  are bandlimited.

**Lemma 2.1** *Let the Fourier transform  $\widehat{U} \in L_1$  of the function  $U$  vanish outside of the set of all  $\omega$  for which  $\|T^t \omega\| \leq \Omega$  holds in a given norm. Then the Fourier transform of its trace function vanishes outside of the ball of radius  $\Omega$  around the origin.*

**Proof** We split the vectors in  $\mathbb{R}^n$  again into the parts  $\omega \in \mathbb{R}^m$  and  $\eta \in \mathbb{R}^{n-m}$ , as in Theorem 2.1 and its proof. Because of our assumption on the support of  $\widehat{U}$ ,

$$\|\omega\| = \left\| T^t \widetilde{T}^{-t} \left(\begin{smallmatrix} \omega \\ \eta \end{smallmatrix}\right) \right\| \leq \Omega$$

holds for the  $\omega$  and  $\eta$  for which the integrand in the representation (2.5) of the Fourier transform  $\widehat{u}(\omega)$  of the trace function takes a value different from zero, which means that  $\widehat{u}(\omega)$  must vanish for arguments  $\omega$  of norm  $\|\omega\| > \Omega$ .  $\square$

### 3 Shifted Laplace equations with trace functions as right-hand sides

We now return to the equation  $-\Delta u + \mu u = f$  from Sect. 1. We show that its solution can, for a right-hand side  $f(x) = F(Tx)$  that is trace of a function  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  with Fourier transform in  $L_1$ , be written as trace  $u(x) = U(Tx)$  of a function  $U : \mathbb{R}^n \rightarrow \mathbb{R}$  that solves a degenerate elliptic equation.

**Theorem 3.1** *Let  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  be a function with Fourier transform in  $L_1$  and let  $\mu$  be a positive constant. The twice continuously differentiable function*

$$u(x) = \left( \frac{1}{\sqrt{2\pi}} \right)^m \int \frac{1}{\mu + \|\omega\|^2} \widehat{f}(\omega) e^{i\omega \cdot x} d\omega, \quad (3.1)$$

*with  $\|\omega\|$  the euclidian norm of  $\omega$ , is then the only solution of the equation*

$$-\Delta u + \mu u = f \quad (3.2)$$

*on the  $\mathbb{R}^m$  that vanishes at infinity.*

**Proof** That  $u$  is a twice continuously differentiable function that solves the equation and vanishes at infinity follows from the remarks made in the last section on functions with integrable Fourier transform. The maximum principle states that it is the only solution of the equation with this property.  $\square$

The solution (3.1) of the Eq. (3.2) can also be characterized in a different way. We call a function  $u \in W_0(\mathbb{R}^m)$  a weak solution of this equation if

$$\int u(-\Delta\varphi + \mu\varphi) dx = \int f\varphi dx \quad (3.3)$$

holds for all rapidly decreasing functions  $\varphi$ .

**Lemma 3.1** *The function (3.1) is the only weak solution of the Eq. (3.2).*

**Proof** That the function (3.1) is a weak solution of the Eq. (3.2) follows from

$$\int e^{i\omega \cdot x} (-\Delta\varphi + \mu\varphi)(x) dx = \int (\mu + \|\omega\|^2) e^{i\omega \cdot x} \varphi(x) dx$$

and Fubini's theorem, which is applied here twice, first to exchange the order of integration with respect to  $x$  and  $\omega$ , and then to revert this process. If  $u_1$  and  $u_2$  are weak solutions of the equation, we have

$$\int (u_1 - u_2)(-\Delta\varphi + \mu\varphi) dx = 0$$

for all rapidly decreasing functions  $\varphi$ . As the equation  $-\Delta\varphi + \mu\varphi = \chi$  possesses for all rapidly decreasing functions  $\chi$  a rapidly decreasing solution  $\varphi$ , with a Fourier representation as above, for all rapidly decreasing functions  $\chi$  then

$$\int (u_1 - u_2) \chi \, dx = 0$$

holds. The difference of  $u_1$  and  $u_2$  must therefore vanish.  $\square$

Let the right-hand side now be the trace  $f(x) = F(Tx)$  of a function  $F$  in  $W_0(\mathbb{R}^n)$ . As such, it is by the results of the previous section a function in  $W_0(\mathbb{R}^m)$ . The crucial observation is that we can lift the equation from  $\mathbb{R}^m$  into  $\mathbb{R}^n$ .

**Theorem 3.2** *The solution (3.1) is the trace  $u(x) = U(Tx)$  of the function*

$$U(y) = \left( \frac{1}{\sqrt{2\pi}} \right)^n \int \frac{1}{\mu + \|T^t \omega\|^2} \widehat{F}(\omega) e^{i\omega \cdot y} \, d\omega \quad (3.4)$$

*mapping the higher dimensional space to the real numbers.*

**Proof** Using again Fubini's theorem in the previously described manner and observing that for rapidly decreasing functions  $\varphi$  because of  $\omega \cdot Tx = T^t \omega \cdot x$

$$\int e^{i\omega \cdot Tx} (-\Delta \varphi + \mu \varphi)(x) \, dx = \int (\mu + \|T^t \omega\|^2) e^{i\omega \cdot Tx} \varphi(x) \, dx$$

holds, one recognizes that the trace function  $u$  is a weak solution of Eq. (3.2). As such, it coincides by Lemma 3.1 with the solution (3.1) of this equation.  $\square$

The function (3.4) is in the domain of the operator  $\mathcal{L}$  given by

$$(\mathcal{L}U)(y) = \left( \frac{1}{\sqrt{2\pi}} \right)^n \int (\mu + \|T^t \omega\|^2) \widehat{U}(\omega) e^{i\omega \cdot y} \, d\omega. \quad (3.5)$$

By definition, it solves the equation

$$\mathcal{L}U = F. \quad (3.6)$$

Because the expression  $\mu + \|T^t \omega\|^2$  is a second order polynomial in the components of  $\omega$ ,  $\mathcal{L}$  can be considered as a second order differential operator and this equation therefore as a degenerate elliptic equation. If the Fourier transform of  $F$  and then also that of the solution  $U$  have a bounded support,  $U$  is infinitely differentiable and its derivatives can be obtained by differentiation under the integral sign. In this case, the function (3.4) is a classical solution of the Eq. (3.6), which is, however, irrelevant for the following considerations.

Instead of attacking the original Eq. (3.2) directly, we will approximate the solution of the higher dimensional Eq. (3.6). We are here primarily interested in right-hand sides  $F$  that are products of lower-dimensional functions or sums or rapidly converging series of such functions. Because  $TT^t$  will only in exceptional cases be a diagonal matrix, the function

$$\frac{1}{\mu + \|T^t \omega\|^2} \quad (3.7)$$

can, however, in general not be approximated as easily by sums of separable Gauss functions as sketched in the introduction for its counterpart in the representation (3.1) of the solution of the original equation. We will show that this problem vanishes in the high-dimensional case due to a concentration of measure effect.

#### 4 The iterative solution of the degenerate elliptic equation

To start with, let  $\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$  be a measurable, bounded function for which

$$\left| 1 - \alpha(\omega)(\mu + \|T^t \omega\|^2) \right| < 1 \quad (4.1)$$

holds for all  $\omega \in \mathbb{R}^n$  and assign to it the operator  $\alpha : W_0(\mathbb{R}^n) \rightarrow W_0(\mathbb{R}^n)$  given by

$$(\alpha F)(y) = \left( \frac{1}{\sqrt{2\pi}} \right)^n \int \alpha(\omega) \widehat{F}(\omega) e^{i\omega \cdot y} d\omega. \quad (4.2)$$

The function  $\widetilde{U} = \alpha F$  can then serve as a first approximation of the solution (3.4). In general, this approximation will not be very precise but can be iteratively improved. Starting from  $U_0 = 0$  or  $U_1 = \alpha F$ , let

$$U_{k+1} = (I - \alpha \mathcal{L})U_k + \alpha F. \quad (4.3)$$

The convergence of these iterates to the solution (3.4) of the equation  $\mathcal{L}U = F$  can already be shown under these very modest assumptions.

**Theorem 4.1** *The iterates (4.3) converge in the norm (2.2) to the solution (3.4).*

**Proof** The errors possess the representation

$$U - U_k = (I - \alpha \mathcal{L})^k U.$$

The  $L_1$ -norm of the Fourier transform of the errors is therefore

$$\|\widehat{U} - \widehat{U}_k\|_{L_1} = \int \left| \left( 1 - \alpha(\omega)(\mu + \|T^t \omega\|^2) \right)^k \widehat{U}(\omega) \right| d\omega.$$

The integrands are by the assumption made above bounded by the absolute value of the integrable function  $\widehat{U}$  and tend almost everywhere to zero as  $k$  goes to infinity. The dominated convergence theorem yields therefore

$$\lim_{k \rightarrow \infty} \|\widehat{U} - \widehat{U}_k\|_{L_1} = 0,$$

which proves the proposition.  $\square$

The same kind of result obviously also holds in the norm

$$\|U\| = \left(\frac{1}{\sqrt{2\pi}}\right)^n \int (\mu + \|T^t \omega\|^2) |\widehat{U}(\omega)| d\omega, \quad (4.4)$$

by which the derivatives up to second order of the trace of  $U$  can be estimated, and in all norms of similar type, be they based on the  $L_1$ - or the  $L_2$ -norm of the Fourier transform, corresponding properties of the right-hand side provided, of course. An interesting example is the Hilbert space norm given by the expression

$$\|U\|^2 = \int |\widehat{U}(\omega)|^2 \prod_{i=1}^n \frac{1 + \omega_i^2}{2} d\omega, \quad (4.5)$$

which dominates the  $L_1$ -based norm used so far and thus also the maximum norm. It measures the size of the first order mixed derivatives and is tensor compatible.

Provided that on the support of  $\widehat{F}$  and thus also of the solution and the iterates

$$\left| 1 - \alpha(\omega)(\mu + \|T^t \omega\|^2) \right| \leq q < 1 \quad (4.6)$$

holds, one obtains from the Fourier representation of the errors the estimate

$$\|U - U_k\| \leq q^k \|U\| \quad (4.7)$$

in the norms (2.2) and (4.4). In this case, one can use polynomial acceleration to speed up the convergence of the iteration or, in other words, to improve the quality of the approximations of the solution. That is, one replaces the  $k$ -th iterate by a weighted mean of the iterate itself and all previous ones. The errors of the recombined iterates possess then again a Fourier representation

$$(U - U_k)(y) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \int P_k(\alpha(\omega)(\mu + \|T^t \omega\|^2)) \widehat{U}(\omega) e^{i\omega \cdot y} d\omega, \quad (4.8)$$

but now not with the polynomials  $P_k(\lambda) = (1 - \lambda)^k$  but polynomials

$$P_k(\lambda) = \sum_{\ell=0}^k \alpha_{k\ell} (1 - \lambda)^\ell, \quad \sum_{\ell=0}^k \alpha_{k\ell} = 1. \quad (4.9)$$

Let  $T_k$  denote the Chebyshev polynomial of degree  $k$ . Among all polynomials  $P$  of degree  $k$  that satisfy the normalization condition  $P(0) = 1$ , the polynomial

$$P_k(\lambda) = T_k\left(\frac{b+a-2\lambda}{b-a}\right) / T_k\left(\frac{b+a}{b-a}\right) \quad (4.10)$$



is then the only one that attains on a given interval  $0 < a \leq \lambda \leq b$  the smallest possible maximum absolute value, which is, in terms of the ratio  $\kappa = b/a$ , given by

$$\max_{a \leq \lambda \leq b} |P_k(\lambda)| = \frac{2r^k}{1 + r^{2k}}, \quad r = \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}. \quad (4.11)$$

As is well-known, this property plays a central role in the analysis of the conjugate gradient method. In our case we have  $a = 1 - q$  and  $b = 1 + q$ . Inserting the corresponding polynomials  $P_k$ , one obtains the error estimate

$$\|U - U_k\| \leq \frac{2r^k}{1 + r^{2k}} \|U\| \quad (4.12)$$

for the recombined iterates. This is in comparison to the convergence rate

$$q = \frac{\kappa - 1}{\kappa + 1} \quad (4.13)$$

of the original iteration a potentially big and very substantial improvement.

## 5 A particular class of iterative methods

Let  $\rho : \mathbb{R}^n \rightarrow \mathbb{R}$  be a measurable, locally bounded function for which

$$\|T^t \omega\| \leq \rho(\omega) \quad (5.1)$$

holds for all  $\omega \in \mathbb{R}^n$ . We assign to this  $\rho$  the iteration (4.3) based on the function

$$\alpha(\omega) = \frac{1}{\mu + \rho(\omega)^2}. \quad (5.2)$$

Our basic example is the function  $\rho(\omega) = \|T\| \|\omega\|$ . The function (5.2) can then in a second step, as indicated in the introduction and explained below in more detail, again be approximated by a sum of Gauss functions. In any case,

$$1 - \alpha(\omega)(\mu + \|T^t \omega\|^2) = \frac{\rho(\omega)^2 - \|T^t \omega\|^2}{\rho(\omega)^2 + \mu}. \quad (5.3)$$

That is, the condition (4.1) is because of (5.1) everywhere satisfied. The error thus tends by Theorem 4.1 in the norm (2.2) to zero. The same holds for the norm (4.4), and for any other norm like that given by (4.5), corresponding regularity properties of the solution provided. The key feature for understanding the convergence of the iteration comes from the analysis of the sets

$$S(\delta) = \{\omega \mid \|T^t \omega\| \geq \delta \rho(\omega)\}, \quad 0 \leq \delta \leq 1. \quad (5.4)$$

If  $\rho(\omega)$  is a norm or seminorm,  $S(\delta)$  is a cone, that is, with  $\omega$  every scalar multiple of  $\omega$  is contained in  $S(\delta)$ . Independent of the size of  $\mu$ , on the set  $S(\delta)$

$$\delta^2 \leq \alpha(\omega)(\mu + \|T^t \omega\|^2) \leq 1 \quad (5.5)$$

holds. If the Fourier transform of the right-hand side  $F$  of the Eq. (3.6) and with that also the Fourier transform of its solution  $U$  vanish outside of the set  $S(\delta)$ , this implies the error estimate

$$\|U - U_k\| \leq q^k \|U\|, \quad q = 1 - \delta^2. \quad (5.6)$$

The crucial point is that the regions  $S(\delta)$  fill in case of high dimensions  $m$  almost the complete space once  $\delta$  falls below a certain bound, without any sophisticated further adaption of the function  $\rho(\omega) = \|T\| \|\omega\|$  to the matrix  $T$ .

**Theorem 5.1** *Let  $\rho(\omega) = \|T\| \|\omega\|$  and let  $\kappa$  be the ratio of the maximum and the minimum singular value of the matrix  $T$ , the condition number of  $T$ . If  $\kappa\delta < 1$ , then*

$$\frac{\lambda(\{\omega \mid \|T^t \omega\| < \delta \rho(\omega), \|\omega\| \leq R\})}{\lambda(\{\omega \mid \|\omega\| \leq R\})} \leq \psi\left(\frac{\kappa\delta}{\sqrt{1 - \kappa^2\delta^2}}\right) \quad (5.7)$$

holds for all  $R > 0$ , where  $\lambda$  is the volume measure on the  $\mathbb{R}^n$  and

$$\psi(\varepsilon) = \frac{2 \Gamma(n/2)}{\Gamma(m/2)\Gamma((n-m)/2)} \int_0^\varepsilon \frac{t^{m-1}}{(1+t^2)^{n/2}} dt. \quad (5.8)$$

Equality holds if and only if  $\kappa = 1$ , that is, if all singular values of  $T$  coincide.

**Proof** Differing from the notation in the theorem but consistent within the proof, we split the vectors in  $\mathbb{R}^n$  into parts  $\omega \in \mathbb{R}^m$  and  $\eta \in \mathbb{R}^{n-m}$ . Starting point of our argumentation is a singular value decomposition  $T^t = U \Sigma V^t$  of  $T^t$ . As the multiplication with the orthogonal matrices  $U$  and  $V^t$ , respectively, does not change the euclidian length of a vector, the set whose volume has to be estimated consists of the points with components  $\omega$  and  $\eta$  in the ball of radius  $R$  around the origin for which

$$\|\Sigma V^t \begin{pmatrix} \omega \\ \eta \end{pmatrix}\| < \delta \|T\| \|V^t \begin{pmatrix} \omega \\ \eta \end{pmatrix}\|$$

holds. Because the volume is invariant under orthogonal transformations, the volume of this set coincides with the volume of the set of all points in the ball for which

$$\|\Sigma \begin{pmatrix} \omega \\ \eta \end{pmatrix}\| < \delta \|T\| \left\| \begin{pmatrix} \omega \\ \eta \end{pmatrix} \right\|$$

holds. Let  $0 < \sigma_1 \leq \dots \leq \sigma_m$  be the diagonal elements of the diagonal matrix  $\Sigma$ , the singular values of the transpose  $T^t$  of  $T$  and of  $T$  itself, and let  $\kappa = \sigma_m/\sigma_1$ . Since

$$\sigma_1 \|\omega\| \leq \left\| \Sigma \begin{pmatrix} \omega \\ \eta \end{pmatrix} \right\|, \quad \|T\| = \sigma_m,$$

the given set is a subset of the set of all those points in the ball for which

$$\|\omega\| < \kappa \delta \left\| \begin{pmatrix} \omega \\ \eta \end{pmatrix} \right\|$$

holds. If  $\kappa = 1$ , that is, if all singular values of the matrix  $T$  are equal,

$$\sigma_1 \|\omega\| = \left\| \Sigma \begin{pmatrix} \omega \\ \eta \end{pmatrix} \right\|.$$

The two sets then coincide and nothing is lost up to here. If  $\kappa > 1$ , there exists a vector with components  $\omega_0$  and  $\eta_0$  inside of the ball under consideration for which

$$\left\| \Sigma \begin{pmatrix} \omega_0 \\ \eta_0 \end{pmatrix} \right\| = \sigma_m \|\omega_0\|, \quad \frac{1 - \kappa^2 \delta^2}{\kappa^2 \delta^2} \|\omega_0\|^2 < \|\eta_0\|^2 < \frac{1 - \delta^2}{\delta^2} \|\omega_0\|^2$$

holds. For this vector and thus also for all vectors sufficiently close to it we have

$$\left\| \Sigma \begin{pmatrix} \omega \\ \eta \end{pmatrix} \right\| > \delta \|T\| \left\| \begin{pmatrix} \omega \\ \eta \end{pmatrix} \right\|, \quad \|\omega\| < \kappa \delta \left\| \begin{pmatrix} \omega \\ \eta \end{pmatrix} \right\|.$$

That means that the two sets then differ and that the second one has a greater volume. In what follows, we will calculate the volume of the latter one and compare it with the volume of the ball. We can restrict ourselves here to the radius  $R = 1$ . Let

$$\varepsilon = \frac{\kappa \delta}{\sqrt{1 - \kappa^2 \delta^2}}.$$

The set consists then of the points in the  $n$ -dimensional unit ball for which

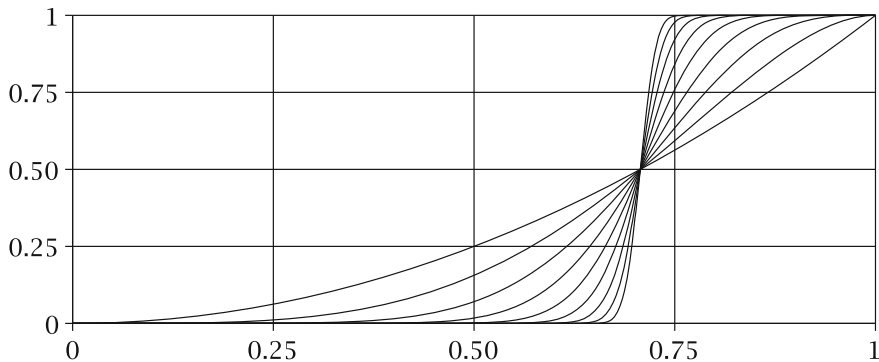
$$\|\omega\| < \varepsilon \|\eta\|$$

holds. Its volume can, by Fubini's theorem, be expressed as double integral

$$\int \left( \int H(\varepsilon \|\eta\| - \|\omega\|) \chi(\|\omega\|^2 + \|\eta\|^2) d\omega \right) d\eta,$$

where  $H(t) = 0$  for  $t \leq 0$ ,  $H(t) = 1$  for  $t > 0$ ,  $\chi(t) = 1$  for  $t \leq 1$ , and  $\chi(t) = 0$  for arguments  $t > 1$ . It tends, by the dominated convergence theorem, to the volume of the unit ball as  $\varepsilon$  goes to infinity. In terms of polar coordinates, it reads

$$(n - m) v_{n-m} \int_0^\infty \left( m v_m \int_0^{\varepsilon s} \chi(r^2 + s^2) r^{m-1} dr \right) s^{n-m-1} ds,$$



**Fig. 1** The bound (5.7) as function of  $0 < \kappa\delta < 1$  for dimensions  $m = 2, 4, 8, \dots, 512$  and  $n = 2m$

with  $v_d$  the volume of the  $d$ -dimensional unit ball. Substituting  $t = r/s$  in the inner integral and interchanging the order of integration, it attains finally the value

$$\frac{(n-m)v_{n-m}mv_m}{n} \int_0^\varepsilon \frac{t^{m-1}}{(1+t^2)^{n/2}} dt.$$

Dividing this by the volume  $v_n$  of the unit ball itself and remembering that

$$v_d = \frac{2}{d} \frac{\pi^{d/2}}{\Gamma(d/2)},$$

this completes the proof of the estimate (5.7) and shows that equality holds if and only if  $\kappa = 1$ , that is, if all singular values of  $T$  coincide. Moreover, we have shown that the function (5.8) tends to one as  $\varepsilon$  goes to infinity, and the bound for the ratio of the two volumes itself to one as  $\kappa\delta$  goes to one.  $\square$

The theorem states that ratio of the two volumes tends like  $\delta^m$  to zero as  $\delta$  goes to zero. It can also for rather large  $\delta$  still attain extremely small values, which means that in most of the frequency space  $\|T^t \omega\|$  behaves like the norm of  $\omega$ . For  $m = 128$  and  $n = 256$ , for example, the ratio is for  $\kappa\delta \leq 1/4$  less than  $1.90 \cdot 10^{-42}$ , and for  $\kappa\delta \leq 1/2$  still less than  $6.95 \cdot 10^{-10}$ . Figure 1 shows the bound (5.7) as a function of  $\kappa\delta < 1$  for  $n = 2m$ , with  $m = 2, 4, 8, \dots, 512$ . As long as the Fourier transform of the solution is not strongly concentrated around the kernel of  $T^t$ , the regions of worse convergence will in such cases hardly affect the iterates and can be ignored.

Polynomial acceleration as in Sect. 4 does not change this picture. If one inserts in (4.8) the polynomials (4.10) that attain the smallest possible maximum absolute value on the interval  $a \leq \lambda \leq b$  with  $a = \delta^2$  and  $b = 1$ , one gets the error estimate

$$\|U - U_k\| \leq \frac{2q^k}{1+q^{2k}} \|U\|, \quad q = \frac{1-\delta}{1+\delta}, \quad (5.9)$$

for the recombined iterates as long as the Fourier transform of the solution  $U$  vanishes outside of  $S(\delta)$ . In the general case, the corresponding part of the error is thus reduced

by a much larger factor than with the basic iteration. As the polynomials (4.10) take on the interval  $0 < \lambda < a$  values  $0 < P_k(\lambda) < 1$ , and satisfy there the estimate

$$0 < P_k(\lambda) < \left(1 - \frac{2\lambda}{a+b}\right)^k, \quad (5.10)$$

the remaining part of the error does not blow up and even tends to zero as  $k$  goes to infinity. The iteration error can thus also in the described cases be reduced very substantially replacing the original iterates by linear combinations of all iterates up to the given one. These arguments also apply when the function (5.2), here

$$\alpha(\omega) = \frac{1}{\mu + \|T\|^2 \|\omega\|^2}, \quad (5.11)$$

is approximated by a sum  $\tilde{\alpha}(\omega)$  of Gauss functions that satisfies an estimate

$$0 \leq \tilde{\alpha}(\omega) \leq (1 + \varepsilon)\alpha(\omega) \quad (5.12)$$

on the whole frequency space and the reverse estimate

$$(1 - \varepsilon)\alpha(\omega) \leq \tilde{\alpha}(\omega) \quad (5.13)$$

on a ball around the origin. The radius of this ball determines the spatial resolution and must therefore be chosen sufficiently large. Approximations that meet these requirements can be constructed truncating the infinite series (1.5) mentioned in the introduction; see the appendix for details. The part of the error with Fourier transform supported on the intersection of the set  $S(\delta)$  and this ball tends then, for the choice  $a = (1 - \varepsilon)\delta^2$  and  $b = 1 + \varepsilon$  in the polynomials (4.10), again rapidly to zero, from one iteration step to the next asymptotically at least by the factor

$$q' = \frac{1 - \delta'}{1 + \delta'}, \quad \delta' = \sqrt{\frac{1 - \varepsilon}{1 + \varepsilon}} \delta, \quad (5.14)$$

which differs also for comparatively crude approximations of the function (5.11) not substantially from the factor in (5.9) for the unperturbed case.

Assume that the approximation  $\tilde{\alpha}(\omega)$  of the function (5.11) is based on an approximation of  $1/r$  with relative accuracy  $\varepsilon$  for  $\mu \leq r \leq \mu R$ , where  $R$  is a large number, say  $R = 10^{12}$  or  $10^{18}$ . The estimate (5.13) holds then on the ball

$$B = \{\omega \mid \mu + \|T\|^2 \|\omega\|^2 \leq \mu R\}. \quad (5.15)$$

The Fourier transform of the solution (3.4) of the Eq. (3.6) can in this case on the complement  $S(\delta) \setminus B$  of the set  $S(\delta) \cap B$  be pointwise estimated as

$$|\widehat{U}(\omega)| \leq \frac{1}{\delta^2 \mu R} |\widehat{F}(\omega)| \quad (5.16)$$

by the Fourier transform of the right-hand side of the equation. This means that already due to the inherent smoothing properties of the equation not much is lost when also the part of the error with Fourier transform supported on  $S(\delta) \setminus B$  is ignored and (5.13) holds only on the given ball and not on the whole frequency space.

The topic of this paper is structural properties of the solutions of the differential Eq. (1.1). Our results open the possibility to apply tensor product methods to the approximation of solutions that themselves do not possess a tensor product structure and are not well separable. The practical feasibility of the approach depends on the representation of the involved tensors and of the factors of which they are composed, in particular on the access to their Fourier transform or the difficulty to calculate their convolution with a Gauss function. An interesting case is when these factors are themselves expanded into Gauss–Hermite functions. The iteration steps (4.3) then do not lead out of this class of functions. This enables a very efficient realization. What remains is the question how and to what extent the iterates can be compressed in between to keep the amount of work and storage under control without affecting the accuracy too much. We refer to [3] for such considerations.

## 6 The limit behavior for fixed dimension ratios

Figure 1 suggests that the bound (5.7) tends to zero for all  $\kappa\delta$  below some jump discontinuity and to one for the  $\kappa\delta$  above this point if the ratio of the dimensions  $m$  and  $n$  is kept fixed and  $m$  tends to infinity. This is indeed the case.

**Lemma 6.1** *If one keeps the ratio  $m/n$  of the dimensions fixed and lets  $m$  tend to infinity, the functions (5.8) tend for arguments  $\varepsilon$  left of the jump discontinuity*

$$\varepsilon_0 = \sqrt{\frac{m}{n-m}} \quad (6.1)$$

*pointwise to zero and for arguments  $\varepsilon$  right of it pointwise to one. The maximum distance of the function values  $\psi(\varepsilon)$  to zero and one, respectively, tends for all  $\varepsilon$  outside any given interval around the jump discontinuity exponentially to zero.*

**Proof** Stirling's formula states that there is a function  $0 < \mu(x) < 1/(12x)$  such that

$$\ln \Gamma(x) = \frac{2x-1}{2} \ln(x) - x + \ln(\sqrt{2\pi}) + \mu(x)$$

holds for all arguments  $x > 0$ ; see [8, Eq. 5.6.1] and [7] for a proof. If the ratio of the dimensions  $m$  and  $n$  is kept fixed and  $m$  goes to infinity, this representation leads to

$$\ln \left( \frac{2 \Gamma(n/2)}{\Gamma(m/2) \Gamma((n-m)/2)} \right) = \beta m + \frac{\ln(m)}{2} - \beta_0 - \ln(\sqrt{\pi}) + \mathcal{O}\left(\frac{1}{m}\right),$$

where the constants  $\beta$  and  $\beta_0$  depend only on the ratio  $m/n$  and are given by

$$2\beta = \frac{n}{m} \ln \left( \frac{n}{m} \right) - \left( \frac{n}{m} - 1 \right) \ln \left( \frac{n}{m} - 1 \right),$$

$$2\beta_0 = \ln\left(\frac{n}{m}\right) - \ln\left(\frac{n}{m} - 1\right).$$

We keep the integers  $m$  and  $n$  in the following fixed, assume that they are relatively prime, and study with help of this representation the limit behavior of the functions

$$\psi_k(\varepsilon) = \frac{2\Gamma(kn/2)}{\Gamma(km/2)\Gamma((kn-km)/2)} \int_0^\varepsilon \frac{t^{km-1}}{(1+t^2)^{kn/2}} dt$$

as  $k$  goes to infinity. The with the prefactor multiplied integrands can be written as

$$C(k)\sqrt{k} \left( \frac{e^{\beta m} t^m}{(1+t^2)^{n/2}} \right)^{k-1} \frac{t^{m-1}}{(1+t^2)^{n/2}},$$

where  $C(k)$  remains bounded and tends to the limit

$$\lim_{k \rightarrow \infty} C(k) = \sqrt{\frac{m}{\pi}} e^{\beta m - \beta_0}.$$

The term in the brackets attains its global maximum at the point  $t = \varepsilon_0$  specified above. As it increases strictly for  $t < \varepsilon_0$ , takes the value one at  $t = \varepsilon_0$ , and decreases strictly for  $t > \varepsilon_0$ , there exists for every open interval around  $\varepsilon_0$  a  $q < 1$  with

$$0 \leq \frac{e^{\beta m} t^m}{(1+t^2)^{n/2}} \leq q$$

for all  $t \geq 0$  outside of it. As stated in the proof of Theorem 5.1,  $\psi_k(\varepsilon)$  tends to one as  $\varepsilon$  goes to infinity. For arguments  $\varepsilon$  right of the interval therefore

$$0 < 1 - \psi_k(\varepsilon) < C(k)\sqrt{k} q^{k-1} \int_\varepsilon^\infty \frac{t^{m-1}}{(1+t^2)^{n/2}} dt$$

holds. For arguments  $\varepsilon > 0$  left of the interval one obtains

$$0 < \psi_k(\varepsilon) < C(k)\sqrt{k} q^{k-1} \int_0^\varepsilon \frac{t^{m-1}}{(1+t^2)^{n/2}} dt.$$

As the integrals are uniformly bounded in  $\varepsilon$ , this proves the proposition.  $\square$

If the ratio  $m/n$  of the dimensions  $m$  and  $n$  is kept fixed and  $m$  tends to infinity, the bound (5.7) from Theorem 5.1 for the ratio of the two volumes tends therefore for values of  $\kappa\delta$  left of the jump discontinuity

$$\xi_0 = \sqrt{\frac{m}{n}} \quad (6.2)$$

to zero and for values of  $\kappa\delta$  right of it to one, uniformly and exponentially outside every interval around  $\xi_0$  and the faster, the larger the interval is. For large dimensions  $m$ , the effective convergence rate in (5.9) thus approaches the value

$$q = \frac{1 - \delta}{1 + \delta}, \quad \delta = \frac{1}{\kappa} \sqrt{\frac{m}{n}}. \quad (6.3)$$

In a similar way, one can derive a local estimate that describes the behavior of the bound (5.7) for small values of  $\kappa\delta$  in a more explicit manner. One starts from the observation that for dimensions  $n \geq m + 2$

$$\psi\left(\frac{\delta}{\sqrt{1 - \delta^2}}\right) \leq \frac{2 \Gamma(n/2)}{\Gamma(m/2) \Gamma((n - m)/2)} \frac{\delta^m}{m} \quad (6.4)$$

holds, which is shown estimating the integrand in (5.8) by the function

$$\varphi(t)^{m-1} \varphi'(t), \quad \varphi(t) = \frac{t}{\sqrt{1 + t^2}}. \quad (6.5)$$

That is, the function on the left-hand side can be estimated by the leading term of its Taylor expansion at  $\delta = 0$ . Inserting the given representation of the prefactor,

$$\psi\left(\frac{\delta}{\sqrt{1 - \delta^2}}\right) \leq c(m) \frac{1}{\sqrt{m}} \left(\frac{\delta}{\delta_0}\right)^m \quad (6.6)$$

follows, where  $\delta_0 = e^{-\beta}$  and  $c(m)$  tends to the limit value  $e^{-\beta_0}/\sqrt{\pi}$  or

$$\sqrt{\frac{n - m}{n\pi}} \quad (6.7)$$

as  $m$  goes to infinity. The scaling factor  $\delta_0$  does not deteriorate when  $n/m$  gets large, at least not faster than the position (6.2) of the jump discontinuity.

**Lemma 6.2** *The scaling factor  $\delta_0 = e^{-\beta}$  can be written as product*

$$e^{-\beta} = \vartheta\left(\frac{m}{n}\right) \sqrt{\frac{m}{n}}, \quad (6.8)$$

where  $\vartheta(x)$  increases monotonously from  $\vartheta(0) = 1/\sqrt{e}$  to  $\vartheta(1) = 1$ .

**Proof** Let  $x = m/n$ . Then we have

$$\vartheta(x) = \exp\left(\frac{(1 - x) \ln(1 - x)}{2x}\right).$$



The exponent possesses for  $|x| < 1$  the power series expansion

$$\frac{(1-x)\ln(1-x)}{2x} = -\frac{1}{2} + \frac{1}{2} \sum_{k=1}^{\infty} \frac{1}{k(k+1)} x^k$$

and tends to the limit value zero as  $x$  goes to 1.  $\square$

## 7 A final example

Finally we return to the initially mentioned example of right-hand sides that depend explicitly on differences of the components  $x_i$  of  $x$ . Let  $\mathcal{J}_0$  be the set of the index pairs  $(i, j)$ ,  $i = 1, \dots, m-1$  and  $j = i+1, \dots, m$ , and let  $\mathcal{J}$  be the subset of  $\mathcal{J}_0$  assigned to the involved differences  $x_i - x_j$ . The total number of variables is then  $n = m + |\mathcal{J}|$ , with  $|\mathcal{J}|$  the number of the index pairs in  $\mathcal{J}$ . We label the first  $m$  components of the vectors in  $\mathbb{R}^n$  by the indices  $1, \dots, m$ , and the remaining components doubly, by the index pairs in  $\mathcal{J}$ . The first  $m$  components of  $Tx$  are in this notation

$$Tx|_i = x_i, \quad i = 1, \dots, m, \quad (7.1)$$

and the remaining components are

$$Tx|_{ij} = x_i - x_j, \quad (i, j) \in \mathcal{J}. \quad (7.2)$$

The components of  $T^t \omega$  can be calculated via the relation  $T^t \omega \cdot x = \omega \cdot Tx$ . If one sets  $\omega'_{ij} = \omega_{ij}$  for the index pairs  $(i, j) \in \mathcal{J}$  and otherwise formally  $\omega'_{ij} = 0$ , they are

$$T^t \omega|_i = \omega_i + \sum_{j=1}^m (\omega'_{ij} - \omega'_{ji}). \quad (7.3)$$

Interesting is the case when  $\omega$  is one the standard basis vectors  $e_i$  or  $e_{ij}$  for  $i = 1, \dots, m$  and  $(i, j) \in \mathcal{J}$ , respectively, pointing into the direction of the coordinate axes. Then

$$T^t e_i|_i = e_i|_i, \quad T^t e_{ij}|_i = e_{ij}|_{ij}, \quad T^t e_{ij}|_j = -e_{ij}|_{ij}, \quad (7.4)$$

and  $T^t \omega|_k = 0$  for all other components. We have

$$\|T^t e_i\| = \|e_i\|, \quad \|T^t e_{ij}\| = \sqrt{2} \|e_{ij}\|. \quad (7.5)$$

If  $\delta \|T\| < 1$ , the coordinate axes hence are contained in the set  $S(\delta)$  on which fast convergence is guaranteed, an advantageous property when the Fourier transforms of the functions under consideration are concentrated around them. This will, for example, be the case when their mixed derivatives are bounded, and in particular when the functions are tensor products of univariate functions.

Because of  $\|Tx\| \geq \|x\|$  and  $\|Te\| = \|e\|$  for  $e = (1, \dots, 1)^t$ , the minimum singular value of  $T$  is one. Since the spectral norm and the spectral condition of  $T$  therefore coincide, the estimate (5.7) reduces here to

$$\frac{\lambda(\{\omega \mid \|T^t \omega\| < \delta \|\omega\|, \|\omega\| \leq R\})}{\lambda(\{\omega \mid \|\omega\| \leq R\})} \leq \psi\left(\frac{\delta}{\sqrt{1 - \delta^2}}\right), \quad (7.6)$$

where  $\delta$  can attain values between zero and one. The norm of the matrix  $T$  depends on the involved differences  $x_i - x_j$ , that is, on the set  $\mathcal{J}$  of index pairs. Let  $q_{ij} = 1$  if either  $(i, j)$  or  $(j, i)$  belongs to  $\mathcal{J}$ , and  $q_{ij} = 0$  otherwise. Then

$$\|Tx\|^2 = \sum_{i=1}^m x_i \left( \left(1 + \sum_{j=1}^m q_{ij}\right) x_i - \sum_{j=1}^m q_{ij} x_j \right). \quad (7.7)$$

The diagonal and off-diagonal elements of  $T^t T$  are therefore

$$T^t T|_{ii} = 1 + \sum_{j=1}^m q_{ij}, \quad T^t T|_{ij} = -q_{ij}, \quad (7.8)$$

and the row-sum norm of  $T^t T$  assigned to the maximum norm is

$$\|T^t T\|_{\infty} = \max_{i=1, \dots, m} \left( 1 + 2 \sum_{j=1}^m q_{ij} \right). \quad (7.9)$$

It represents an upper bound for the eigenvalues of  $T^t T$  and its square root therefore an upper bound for the singular values of  $T$ . Like the ratio

$$\frac{n}{m} = 1 + \frac{1}{2m} \sum_{i=1}^m d_i, \quad d_i = \sum_{j=1}^m q_{ij}, \quad (7.10)$$

of the dimensions  $n$  and  $m$ , the spectral norm and the spectral condition of  $T$  can thus be bounded in terms of the degrees  $d_i$  of the vertices of the underlying graph.

**Acknowledgements** Open Access funding provided by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix. On the exponential approximation

The approximation (1.5) of  $1/r$  by exponential functions can be written in the form

$$\frac{1}{r} \approx h \sum_{k=-\infty}^{\infty} e^{kh} \exp(-e^{kh} r) = \frac{\phi(\ln r)}{r},$$

where  $\phi$  denotes the continuous,  $h$ -periodic function

$$\phi(s) = h \sum_{k=-\infty}^{\infty} \exp(-e^{kh+s} + kh + s).$$

It has been shown in [9, Sect. 5] by means of tools from Fourier analysis that this function approximates the constant 1 with an absolute error

$$\sim 4\pi h^{-1/2} e^{-\pi^2/h}$$

as  $h$  tends to zero, that is, already for rather large values of  $h$  with very high accuracy. The same kind of representation holds when the series is replaced by a finite sum

$$\frac{1}{r} \approx h \sum_{k=k_1}^{k_2} e^{kh} \exp(-e^{kh} r) = \frac{\tilde{\phi}(\ln r)}{r}.$$

To find out with which relative error  $\varepsilon$  this sum approximates the function  $1/r$  on a given interval  $1 \leq r \leq R$ , and the correspondingly rescaled sum

$$\frac{1}{r} \approx h \sum_{k=k_1}^{k_2} \frac{e^{kh}}{\mu} \exp\left(-\frac{e^{kh}}{\mu} r\right)$$

the function then on the interval  $\mu \leq r \leq R\mu$ , thus one has to study the function

$$\tilde{\phi}(s) = h \sum_{k=k_1}^{k_2} \exp(-e^{kh+s} + kh + s).$$

If the approximation  $\tilde{\alpha}(\omega)$  of the function (5.11) is based on the described approximation of  $1/r$  on the interval  $\mu \leq r \leq R\mu$ , (5.12) means that  $\tilde{\phi}(s) \leq 1 + \varepsilon$  must hold for all arguments  $s \geq 0$ . The condition (5.13) is satisfied for the  $\omega$  in the ball

$$\mu + \|T\|^2 \|\omega\|^2 \leq \mu R$$

if for all  $s$  in the interval  $0 \leq s \leq \ln(R)$  conversely  $1 - \varepsilon \leq \tilde{\phi}(s)$  holds. It does not require much effort to fulfill these conditions for moderate accuracies  $\varepsilon$  and large  $R$ ,

as needed in the present context. The relative error  $\varepsilon$  is, for instance, less than 0.01 on the interval  $\mu \leq r \leq 10^{18}\mu$  for  $h = 1.4$ ,  $k_1 = -35$ , and  $k_2 = 1$ , that is, in the one-percent range on an interval that spans eighteen orders of magnitude.

## References

1. Braess, D., Hackbusch, W.: Approximation of  $1/x$  by exponential sums in  $[1, \infty)$ . *IMA J. Numer. Anal.* **25**, 685–697 (2005)
2. Braess, D., Hackbusch, W.: On the efficient computation of high-dimensional integrals and the approximation by exponential sums. In: DeVore, R., Kunoth, A. (eds.) *Multiscale, Nonlinear and Adaptive Approximation*. Springer, Heidelberg (2009)
3. Dahmen, W., DeVore, R., Grasedyck, L., Süli, E.: Tensor-sparsity of solutions to high-dimensional elliptic partial differential equations. *Found. Comput. Math.* **16**, 813–874 (2016)
4. Grasedyck, L.: Existence and computation of low Kronecker-rank approximations for large linear systems of tensor product structure. *Computing* **72**, 247–265 (2004)
5. Hackbusch, W.: *Tensor Spaces and Numerical Tensor Calculus*. Springer, Cham (2019)
6. Khoromskij, B.: Tensor-structured preconditioners and approximate inverse of elliptic operators in  $\mathbb{R}^d$ . *Constr. Approx.* **30**, 599–620 (2009)
7. Königsberger, K.: *Analysis I*. Springer, Berlin (2004)
8. Olver, F.W., Lozier, D.W., Boisvert, R.F., Clark, C.W. (eds.): *NIST Handbook of Mathematical Functions*. Cambridge University Press, Cambridge (2010)
9. Scholz, S., Yserentant, H.: On the approximation of electronic wavefunctions by anisotropic Gauss and Gauss–Hermite functions. *Numer. Math.* **136**, 841–874 (2017)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.