FORSCHUNGSZENTRUM JÜLICH GmbH
Zentralinstitut für Angewandte Mathematik
D-52425 Jülich, Tel. (02461) 61-6402

Technical Report

High-Order Compact Solvers for the Three Dimensional Poisson Equation

Godehard Sutmann, Bernhard Steffen

FZJ-ZAM-IB-2003-12

December 2003
(last change: 11.12.2003)

Preprint: Submitted for publication

High-Order Compact Solvers for the Three Dimensional Poisson Equation

Godehard Sutmann[1] and Bernhard Steffen[2]

Central Institute for Applied Mathematics (ZAM) and
John von Neumann Institute for Computing (NIC)
Research Centre Jülich (FZJ)
D - 52425 Jülich, Germany

December 11, 2003

[1] g.sutmann@fz-juelich.de
[2] b.steffen@fz-juelich.de

Abstract

New compact approximation schemes for the Laplace operator of 4th- and 6th-order are proposed. The schemes are based on a Padé approximation of the Taylor expansion for the discretized Laplace operator. The new schemes are compared with other finite difference approximations in several test cases. It is found that the new schemes are superior in performance and accuracy with respect to other methods.

Keywords: Poisson equation, compact solvers, iterative solvers, Padé approximation

1 Introduction

Consider the three-dimensional Poisson equation with Dirichlet boundary conditions

$$\Delta u(x,y,z) = -f(x,y,z) \qquad x,y,z \in \Omega \tag{1}$$

$$u(x,y,z) = u_0(x,y,z) \qquad x,y,z \in \partial\Omega \tag{2}$$

In a discrete form this can be written as

$$(u_{xx})_{ijk} + (u_{yy})_{ijk} + (u_{zz})_{ijk} = -f_{ijk} \tag{3}$$

where $(u_{\alpha\alpha})_{ijk}$ denotes an approximation to the second partial derivative with respect to the coordinate direction α. The simplest approximation is obtained by

$$(u_{\alpha\alpha})_{ijk} = \frac{1}{h_\alpha^2}\delta_\alpha^2 u_{ijk} \tag{4}$$

where h_α is the grid spacing in direction α and δ_α^2 is the second-order difference operator, e.g.

$$\delta_x^2 = u_{i-1,j,k} - 2u_{i,j,k} + u_{i+1,j,k} \tag{5}$$

Higher order finite difference operators may be derived from the approximation [1]

$$\left.\frac{\partial^2 u}{\partial \alpha^2}\right|_{\alpha=ih_\alpha} = \frac{4}{h_\alpha^2}\sinh^{-1}\left(\frac{\delta_\alpha}{2}\right) \tag{6}$$

$$= \frac{1}{h_\alpha^2}\delta_\alpha^2\left\{1 - \frac{1}{12}\,\delta_\alpha^2 + \frac{1}{90}\,\delta_\alpha^4 - \frac{1}{560}\,\delta_\alpha^6 \pm \ldots\right\} u_{i,j,k} \tag{7}$$

A fourth order accurate scheme may be derived from considering only the first two terms in the expansion

$$(u_{\alpha\alpha})_{ijk} = \frac{1}{h_\alpha^2}\,\delta_\alpha^2\left(1 - \frac{1}{12}\,\delta_\alpha^2\right) \tag{8}$$

The explicit expression in terms of $u_{i,j,k}$ reads for the x-component

$$(u_{xx})_{ijk} = -\frac{1}{h_x^2}\frac{1}{12}\left(u_{i-2,j,k} - 16u_{i-1,j,k} + 30u_{i,j,k} - 16u_{i+1,j,k} + u_{i+2,j,k}\right) \tag{9}$$

If $\Delta_{i,j,k}$ is defined as the difference scheme, the resulting 13-point stencil can be written as

$$\Delta_{i,0,k} = -\frac{1}{12}\begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -16 & 0 & 0 \\ 1 & -16 & 90 & -16 & 1 \\ 0 & 0 & -16 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{10}$$

$$\Delta_{i,\pm1,k} = \frac{1}{12}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 16 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{11}$$

$$\Delta_{i,\pm2,k} = -\frac{1}{12}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{12}$$

The obvious problem with this representation is that one has to operate on the next-nearest gridpoint in the solver. This is a problem especially at the points close to the boundary. Either one has to switch to another solver of lower order, which may influence the overall accuracy of the solution. Or one needs to introduce a second boundary layer. This, however, is often impossible due to limited information. In the case of multigrid methods [2] it becomes even worse, since one should have to work consistently with two boundary layers on the coarse grid levels, which is not included in the formalism.

It is therefore of considerable interest to construct compact higher order schemes, which only need information from the next nearest grid points in space and which therefore do not get into conflict with one-layer boundary conditions.

For the case of 4th-order solvers a compact scheme was derived by Kwon and Stephenson [3], which uses a 19-point stencil for the left hand side of Eq.3. In addition, also the right hand side is modified in that it also takes into account nearest neighbor source terms

$$\begin{aligned} \frac{1}{h^2}\Big(4u_{ijk} \quad &- \quad \frac{1}{3}(u_{i-1,j,k} + u_{i+1,j,k} + u_{i,j-1,k} \\ &\quad +u_{i,j+1,k} + u_{i,j,k-1} + u_{i,j,k+1}) \\ &- \quad \frac{1}{6}(u_{i-1,j-1,k} + u_{i-1,j+1,k} + u_{i-1,j,k-1} \\ &\quad +u_{i-1,j,k+1} + u_{i,j-1,k-1} + u_{i,j-1,k+1} \\ &\quad +u_{i+1,j-1,k} + u_{i+1,j+1,k} + u_{i+1,j,k-1} + \\ &\quad u_{i+1,j,k+1} + u_{i,j+1,k-1} + u_{i,j+1,k+1})\Big) \\ &= \quad f_{i,j,k} + \frac{1}{12}\,(f_{i-1,j,k} + f_{i+1,j,k} + f_{i,j-1,k} \\ &\qquad\qquad +f_{i,j+1,k} + f_{i,j,k-1} + f_{i,j,k+1}) \end{aligned} \tag{13}$$

In Ref. [4] an approach, based on a Padé approximation was developed for two-dimensional reaction-diffusion equations. This approach was extended to three dimensions in Ref. [5]. In the present paper the Padé approximation techique will be used to derive several forms of Poisson solvers of higher order. In Sec. 2 solvers will be derived. In Sec. 3 solvers will be characterized in terms of their convergence behavior. Different test cases are used to demonstrate the

superiority with respect to other solver types. An error analysis will be given to justify our computational findings. Sec. 4 will give some conclusions and an outview for future research.

2 Theory

In this section a Padé approximation to the bracketed expression in Eq.7 will be used to derive different forms of compact stencils for the Poisson equation. The term *compact* will be used in the following for numerical schemes, which need less neighbor grid points than the straight forward expansion approach of Eq.7. For the sequel we define the $[m,n]$-Padé approximation of a function $f(x)$ as

$$\mathcal{P}_{m,n}[f(x)] = \frac{\sum_{k=0}^{m} a_k x^k}{1 + \sum_{k=1}^{m} b_k x^k} = R(x) \tag{14}$$

where a_k and b_k are chosen in a way that

$$\left.\frac{\partial^k R}{\partial x^k}\right|_{x=0} = \left.\frac{\partial^k f}{\partial x^k}\right|_{x=0} \quad , \quad k = 0, \ldots, m+n \tag{15}$$

2.1 A $\mathcal{P}_{0,2}$ compact 4th-order scheme

Eq. 8 may be approximated by an $[0,2]$-Padé approximation through

$$(u_{\alpha\alpha})_{ijk} = \frac{1}{h_\alpha^2}\,\delta_\alpha^2 \left(1 + \frac{1}{12}\,\delta_\alpha^2\right)^{-1} \tag{16}$$

$$= \frac{1}{h_\alpha^2}\,\delta_\alpha^2\,(D_{[0,2],\alpha})^{-1} \tag{17}$$

where in Eq.17 the operator $D_{[0,2],\alpha}$ was defined. This is now used in Eq.3 which can be written as

$$\left\{ \sum_{\alpha=x,y,z} \frac{1}{h_\alpha^2}\,\delta_\alpha^2 (D_{[0,2],\alpha})^{-1} \right\} u_{i,j,k} = -f_{i,j,k} \tag{18}$$

Simple algebraic manipulation gives

$$\begin{aligned} \frac{1}{h^2}\Big\{\Big[& \left(1+\delta_x^2\,(D_{[0,2],x})^{-1}\right)\left(1+\delta_y^2\,(D_{[0,2],y})^{-1}\right)\left(1+\delta_z^2\,(D_{[0,2],z})^{-1}\right) \\ - & \delta_x^2\delta_y^2 (D_{[0,2],x})^{-1}(D_{[0,2],y})^{-1} - \delta_x^2\delta_z^2 (D_{[0,2],x})^{-1}(D_{[0,2],z})^{-1} \\ - & \delta_y^2\delta_z^2 (D_{[0,2],y})^{-1}(D_{[0,2],z})^{-1} \\ - & \delta_x^2\delta_y^2\delta_z^2 (D_{[0,2],x})^{-1}(D_{[0,2],y})^{-1}(D_{[0,2],z})^{-1}\Big] - 1\Big\} u_{i,j,k} = -f_{i,j,k} \end{aligned} \tag{19}$$

In a next step, both sides of Eq.19 are multiplied by $D_{[0,2],x}D_{[0,2],y}D_{[0,2],z}$. Since the operators commute, this leads to

$$\frac{1}{h^2}\left\{\delta_x^2 D_{[0,2],y}D_{[0,2],z} + \delta_y^2 D_{[0,2],x}D_{[0,2],z} + \delta_z^2 D_{[0,2],x}D_{[0,2],y}\right\} u_{i,j,k} = g_{i,j,k} \tag{20}$$

where an effective source term $g_{i,j,k}$ was introduced

$$g_{i,j,k} = \left\{1 + \frac{1}{12}\left(\delta_x^2+\delta_y^2+\delta_z^2\right) + \frac{1}{144}\left(\delta_x^2\delta_y^2+\delta_x^2\delta_z^2+\delta_y^2\delta_z^2\right)\right\} f_{i,j,k} \tag{21}$$

Keeping terms on the *lhs* up to fourth order, Eqs.20,21 may be written finally as

$$\frac{1}{h^2}\left\{\delta_x^2+\delta_y^2+\delta_z^2+\frac{1}{6}\left(\delta_x^2\delta_y^2+\delta_x^2\delta_z^2+\delta_y^2\delta_z^2\right)\right\}u_{i,j,k} \tag{22}$$

$$= \left\{1+\frac{1}{12}\left(\delta_x^2+\delta_y^2+\delta_z^2\right)+\frac{c_1}{144}\left(\delta_x^2\delta_y^2+\delta_x^2\delta_z^2+\delta_y^2\delta_z^2\right)\right\}f_{i,j,k} \tag{23}$$

To have the *rhs* of forth order, it is only necessary to keep the first two terms. However, a higher order term has been formally kept in the expression. In the test cases it will be shown that it may improve accuracy for discontinuous sources. The factor c_1 takes values of zero or one. It was introduced in order to control the higher order approximation of the *rhs*. Note that the higher order term comes naturally from the Padé approximation and is not an artificial construction. Putting the constant c_1 to zero, reduces this scheme to the one introduced in Ref.[6] on the basis of a different approach. Here and in the following this approach is justified from the computational point of view since the modified source function is calculated only once before starting the iterative process. Therefore the computational overhead is negligible.

If a stencil notation is used and the Laplace operator is denoted as $\Delta_{i,j,k}$ whereas the source term operator, appearing on the *rhs* is denoted as $\Gamma_{i,j,k}$, the operators take the form

$$\Delta_{i,0,k} = \frac{1}{6h^2}\begin{bmatrix}1 & 2 & 1\\ 2 & -24 & 2\\ 1 & 2 & 1\end{bmatrix} \qquad \Delta_{i,\pm1,k} = \frac{1}{6h^2}\begin{bmatrix}0 & 1 & 0\\ 1 & 2 & 1\\ 0 & 1 & 0\end{bmatrix} \tag{24}$$

$$\Gamma_{i,0,k} = \frac{1}{144}\begin{bmatrix}c_1 & 12-4c_1 & c_1\\ 12-4c_1 & 72+84c_1 & 12-4c_1\\ c_1 & 12-4c_1 & c_1\end{bmatrix} \tag{25}$$

$$\Gamma_{i,\pm1,k} = \frac{1}{144}\begin{bmatrix}0 & c_1 & 0\\ c_1 & 12-4c_1 & c_1\\ 0 & c_1 & 0\end{bmatrix} \tag{26}$$

2.2 A $\mathcal{P}_{0,4}$ compact 6th-order scheme

In this section the expansion of Eq.7 is considered up to sixth order. The Padé approximation in this case is

$$(u_{\alpha\alpha})_{ijk} = \frac{1}{h_\alpha^2}\,\delta_\alpha^2\left(1+\frac{1}{12}\,\delta_\alpha^2-\frac{1}{240}\delta_\alpha^4\right)^{-1} \tag{27}$$

$$= \frac{1}{h_\alpha^2}\,D^{-1}_{[0,4],\alpha} \tag{28}$$

This leeds to the following form of the Poisson equation

$$\left\{\sum_{\alpha=x,y,z}\frac{1}{h_\alpha^2}\,\delta_\alpha^2\,D^{-1}_{[0,4],\alpha}\right\}u_{i,j,k} = -f_{i,j,k} \tag{29}$$

Similarly as before, both sides of Eq.29 may be multiplied with $\prod_{\alpha=x,y,z} D_{[0,4],\alpha}$ to give

$$\left\{\sum_\alpha \frac{1}{h_\alpha^2}\,\delta_\alpha^2\,D_{[0,4],\beta}\,D_{[0,4],\gamma}\right\} = -\left\{\prod_{\alpha=x,y,z} D_{[0,4],\alpha}\right\}\,f_{i,j,k} \tag{30}$$

which can be written as

$$\left\{\sum_\alpha \frac{1}{h_\alpha^2}\,\delta_\alpha^2\left(1+\frac{1}{12}\sum_{\beta\neq\alpha}\delta_\beta^2-\frac{1}{240}\sum_{\beta\neq\alpha}\delta_\beta^4+\frac{1}{144}\prod_{\beta\neq\alpha}\delta_\beta^2\right)\right\}\,u_{i,j,k} = -g_{i,j,k} \tag{31}$$

where an extended source term was introduced containing terms up to sixth order

$$g_{i,j,k} = \left\{1+\frac{1}{12}\sum_\alpha\delta_\alpha^2\left(1-\frac{1}{20}\,\delta_\alpha^2+\frac{1}{24}\sum_{\beta\neq\alpha}\delta_\beta^2-\frac{c_2}{240}\sum_{\beta\neq\alpha}\delta_\beta^4\right)\right\}\,f_{i,j,k} \tag{32}$$

Keeping terms up to sixth order and using the stencil notation as before, the finite difference approximation scheme for the Laplace operator then reads as

$$\Delta_{i,0,k} = \frac{1}{240\,h^2}\begin{bmatrix} 0 & -1 & 4 & -1 & 0\\ -1 & 38 & 72 & 38 & -1\\ 4 & 72 & -928 & 72 & 4\\ -1 & 38 & 72 & 38 & -1\\ 0 & -1 & 4 & -1 & 0\end{bmatrix} \tag{33}$$

$$\Delta_{i,\pm1,k} = \frac{1}{240\,h^2}\begin{bmatrix} 0 & 0 & -1 & 0 & 0\\ 0 & 5 & 38 & 5 & 0\\ -1 & 38 & 72 & 38 & -1\\ 0 & 5 & 38 & 5 & 0\\ 0 & 0 & -1 & 0 & 0\end{bmatrix} \tag{34}$$

$$\Delta_{i,\pm 2,k} = \frac{1}{240\,h^2} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 4 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{35}$$

Concerning the *rhs* of Eq. 31 it is enough to keep the first three terms in the inner brackets of Eq. 32 in order to be consistent with a 6th-order approximation. However, as before the next higher terms may be kept in order to perform a higher order approximation for the *rhs* of Eq. 31. As will be shown later a better approximation for discontinuous source distributions is obtained when keeping also the forth term. Consequently a constant c_2 is introduced which takes values zero or one. The source term operator then reads in stencil notation

$$\Gamma_{i,0,k} = \frac{1}{2880} \begin{bmatrix} 0 & -c_2 & -12+4c_2 & -c_2 & 0 \\ -c_2 & 20+8c_2 & 208-28c_2 & 20+8c_2 & -c_2 \\ -12+4c_2 & 208-28c_2 & 1464+72c_2 & 208-28c_2 & -12+4c_2 \\ -c_2 & 20+8c_2 & 208-28c_2 & 20+8c_2 & -c_2 \\ 0 & -c_2 & -12+4c_2 & -c_2 & 0 \end{bmatrix} \tag{36}$$

$$\Gamma_{i,\pm 1,k} = \frac{1}{2880} \begin{bmatrix} 0 & 0 & -c_2 & 0 & 0 \\ 0 & 0 & 20+8c_2 & 0 & 0 \\ -c_2 & 20+8c_2 & 208-28c_2 & 20+8c_2 & -c_2 \\ 0 & 0 & 20+8c_2 & 0 & 0 \\ 0 & 0 & -c_2 & 0 & 0 \end{bmatrix} \tag{37}$$

$$\Gamma_{i,\pm 2,k} = \frac{1}{2880} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -c_2 & 0 & 0 \\ 0 & -c_2 & -12+4c_2 & -c_2 & 0 \\ 0 & 0 & -c_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{38}$$

2.3 A $\mathcal{P}_{2,2}$ compact 6th-order scheme

Here the expansion of Eq.7 is considered again up to sixth order. A [2,2]-Padé approximation is used, so that Eq.7 is rewritten as

$$(u_{\alpha\alpha})_{i,j,k} = \frac{1}{h_\alpha}\ \delta_\alpha^2 D^{(1)}_{[2,2],\alpha} (D^{(2)}_{[2,2],\alpha})^{-1} \tag{39}$$

where

$$D^{(1)}_{[2,2],\alpha} = 1 + \frac{1}{20}\ \delta_\alpha^2 \tag{40}$$

$$D^{(2)}_{[2,2],\alpha} = 1 + \frac{2}{15}\ \delta_\alpha^2 \tag{41}$$

The Poisson equation, Eq.3 can then be written as

$$\left\{ \sum_{\alpha,\beta,\gamma} \delta_\alpha^2\ D^{(1)}_{[2,2],\alpha}\ D^{(2)}_{[2,2],\beta}\ D^{(2)}_{[2,2],\gamma} \right\}\ u_{i,j,k} = -\left\{ \prod_\alpha\ D^{(2)}_{[2,2],\alpha} \right\}\ f_{i,j,k} \tag{42}$$

Keeping terms up to 6th-order in δ_α^2 on the *lhs* this expression is rewritten in terms of δ_α^2 as

$$\left\{ \sum_\alpha \frac{1}{h_\alpha^2}\ \delta_\alpha^2 \left(1 + \frac{1}{20}\ \delta_\alpha^2 + \frac{2}{15} \sum_{\beta\neq\alpha} \delta_\beta^2 + \frac{1}{150} \sum_{\beta\neq\alpha}\ \delta_\alpha^2\delta_\beta^2 + \frac{4}{225} \prod_{\beta\neq\alpha} \delta_\beta^2 \right) \right\}\ u_{i,j,k} = -g_{i,j,k} \tag{43}$$

where the effective source term

$$g_{i,j,k} = \left\{ 1 + \sum_\alpha \delta_\alpha^2 \left(\frac{2}{15} + \frac{2}{225} \sum_{\beta\neq\alpha} \delta_\beta^2 + \frac{8\,c_3}{3375} \prod_{\beta\neq\alpha} \delta_\beta^2 \right) \right\}\ f_{i,j,k} \tag{44}$$

was introduced. Again a higher order approximation is kept formally on the *rhs*. Setting $c_3 = 0$ introduces the same order of accuracy on both sides of Eq. 43. The finite difference approximation scheme then reads in stencil notation as

$$\Delta_{i,0,k} = \frac{1}{300\,h^2} \begin{bmatrix} 0 & 2 & 7 & 2 & 0 \\ 2 & 32 & 40 & 32 & 2 \\ 7 & 40 & -842 & 40 & 7 \\ 2 & 32 & 40 & 32 & 2 \\ 0 & 2 & 7 & 2 & 0 \end{bmatrix} \tag{45}$$

$$\Delta_{i,\pm1,k} = \frac{1}{300\,h^2} \begin{bmatrix} 0 & 0 & 2 & 0 & 0 \\ 0 & 16 & 32 & 16 & 0 \\ 2 & 32 & 40 & 32 & 2 \\ 0 & 16 & 32 & 16 & 0 \\ 0 & 0 & 2 & 0 & 0 \end{bmatrix} \tag{46}$$

$$\Delta_{i,\pm2,k} = \frac{1}{300\,h^2} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 2 & 7 & 2 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{47}$$

$$\Gamma_{i,0,k} = \frac{1}{3375} \begin{bmatrix} 60-16c_3 & 210+32c_3 & 60-16c_3 \\ 210+32c_3 & 1395-64c_3 & 210+32c_3 \\ 60-16c_3 & 210+32c_3 & 60-16c_3 \end{bmatrix} \tag{48}$$

$$\Gamma_{i,\pm1,k} = \frac{1}{3375} \begin{bmatrix} 8c_3 & 60-16c_3 & 8c_3 \\ 60-16c_3 & 210+32c_3 & 60-16c_3 \\ 8c_3 & 60-16c_3 & 8c_3 \end{bmatrix} \tag{49}$$

2.4 A $\mathcal{P}_{2,4}$ compact 6th-order scheme

As before the expansion of Eq.7 is considered up to sixth order. The terms in brackets are rewritten in terms of a [2,4]-Padé approximation. Eq.7 may then be written as

$$(u_{\alpha\alpha})_{i,j,k} = \frac{1}{h_\alpha}\ \delta_\alpha^2 D^{(1)}_{[2,4],\alpha} (D^{(2)}_{[2,4],\alpha})^{-1} \tag{50}$$

where

$$D^{(1)}_{[2,4],\alpha} = 1 - \frac{11}{36}\,\delta^2_\alpha \tag{51}$$

$$D^{(2)}_{[2,4],\alpha} = 1 - \frac{2}{9}\,\delta^2_\alpha - \frac{4}{135}\,\delta^4_\alpha \tag{52}$$

In formal analogy to the [2,2]-Padé approximation, the Poisson equation, Eq.3 can be written as

$$\left\{ \sum_{\alpha,\beta,\gamma} \delta^2_\alpha\; D^{(1)}_{[2,4],\alpha}\; D^{(2)}_{[2,4],\beta}\; D^{(2)}_{[2,4],\gamma} \right\} u_{i,j,k} = -\left\{ \prod_\alpha\; D^{(2)}_{[2,4],\alpha} \right\} f_{i,j,k} \tag{53}$$

From this approximation it is clear that terms up to order twelve appear in the expression. Since the original approximation is of order six, also in this approximation only terms up to order six are kept. This also ensures a compact representation of the Laplace operator. Rewriting Eq.53 in terms of δ^2_α gives

$$\left\{ \sum_\alpha \frac{1}{h^2_\alpha}\;\delta^2_\alpha \left(1 - \frac{11}{36}\;\delta^2_\alpha - \frac{2}{9}\sum_{\beta\neq\alpha}\delta^2_\beta + \frac{31}{810}\sum_{\beta\neq\alpha}\;\delta^2_\alpha\delta^2_\beta + \frac{4}{27}\prod_{\beta\neq\alpha}\delta^2_\beta \right) \right\} u_{i,j,k} = -g_{i,j,k} \tag{54}$$

Again an effective source term was introduced

$$g_{i,j,k} = \left\{ 1 - \sum_\alpha \delta^2_\alpha \left(\frac{2}{9} + \frac{4}{135}\delta^2_\alpha - \frac{2}{81}\sum_{\beta\neq\alpha}\delta^2_\beta + \frac{8\,c_4}{729}\prod_{\beta\neq\alpha}\delta^2_\beta \right) \right\} f_{i,j,k} \tag{55}$$

As before, a higher order approximation is kept formally on the *rhs*. The case $c_4 = 0$ corresponds to the same order of accuracy on both sides of Eq. 54. In stencil notation the Laplace and the source term operator read

$$\Delta_{i,0,k} = \frac{1}{1620\,h^2} \begin{bmatrix} 0 & 62 & -743 & 62 & 0 \\ 62 & -1696 & 9176 & -1696 & 62 \\ -743 & 9176 & -33654 & 9176 & -743 \\ 62 & -1696 & 9176 & -1696 & 62 \\ 0 & 62 & -743 & 62 & 0 \end{bmatrix} \tag{56}$$

$$\Delta_{i,\pm1,k} = \frac{1}{1620\,h^2} \begin{bmatrix} 0 & 0 & 62 & 0 & 0 \\ 0 & 240 & -1696 & 240 & 0 \\ 62 & -1696 & 9176 & -1696 & 62 \\ 0 & 240 & -1696 & 240 & 0 \\ 0 & 0 & 62 & 0 & 0 \end{bmatrix} \tag{57}$$

$$\Delta_{i,\pm2,k} = \frac{1}{1620\,h^2} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 62 & 0 & 0 \\ 0 & 62 & -743 & 62 & 0 \\ 0 & 0 & 62 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{58}$$

$$\Gamma_{i,0,k} = \frac{1}{3645}\begin{bmatrix} 0 & 0 & -108 & 0 & 0 \\ 0 & 180+80c_4 & -1098-160c_4 & 180+80c_4 & 0 \\ -108 & -1098-160c_4 & 8721+320c_4 & -1098-160c_4 & -108 \\ 0 & 180+80c_4 & -1098-160c_4 & 180+80c_4 & 0 \\ 0 & 0 & -108 & 0 & 0 \end{bmatrix} \tag{59}$$

$$\Gamma_{i,\pm 1,k} = \frac{1}{3645}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & -40c_4 & 180+80c_4 & -40c_4 & 0 \\ 0 & 180+80c_4 & -1098-160c_4 & 180+80c_4 & 0 \\ 0 & -40c_4 & 180+80c_4 & -40c_4 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{60}$$

$$\Gamma_{i,\pm 1,k} = \frac{1}{3645}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -108 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{61}$$

3 Results

In the previous section a formal derivation of high order schemes, based on a Padé approximation was given. Although the Padé approximations should have the desired order in h, a formal proof for this fact should be given.

As a matter of fact the eigenfunctions of the exact form of the Laplace operator are known to be $\phi_{k,l,m} = \cos(kx)\cos(ly)\cos(mz)$ with corresponding eigenvalues $\lambda = -(k^2+l^2+m^2)$. For the usual stencils, the eigenvectors of the discrete problem are simply the sampled continuous eigenvectors with eigenvalues changed according to the order of accuracy. To test this feature for stencils with a modified *rhs*, a slightly more involved procedure is needed. In the derivation of the Padé approximations two operators appear in the finl results. The one acting on the field u, the other acting on the source terms f, i.e. $\tilde{\Delta}u = \tilde{\Gamma}f$, where $\tilde{\Delta}$ and $\tilde{\Gamma}$ are approximate operators. Therefore a generalized eigenvalue problem has to be solved

$$\tilde{\Delta}u = \lambda\tilde{\Gamma}u \tag{62}$$

Both $\tilde{\Delta}$ and $\tilde{\Gamma}$ have the desired property that eigenfunctions are sampled continuous eigenvectors. Consequently only the eigenvalues need to be tested. For practical purposes one may insert the eigenfunctions of the *exact* operator and controls the order in h of the approximation, i.e.

$$\left(\tilde{\Delta}\phi_{k,l,m} + (k^2+l^2+m^2)\,\tilde{\Gamma}\phi_{k,l,m}\right)\,\phi^{-1}_{k,l,m} = \mathcal{O}(h^n) \tag{63}$$

where n is the order of the approximation. Due to the linearity of the operators, this should give the order of the operator itself. For all approximations developed here, the expected order was recovered. Note, however, that this procedure is true for the Poisson equation. The derivation of the solvers introduces in a natural way a modified *rhs* of the Poisson equation. One therefore could

argue that the order for a Laplace equation will be reduced. In fact, $\tilde{\Delta}$ is only a 2nd-order approximation of the exact Laplace operator for arbitrary functions ψ. However, for those functions $\tilde{\psi}$, solving the Laplace equation $\Delta\tilde{\psi} = 0$ it has full order.

In order to validate the finite difference schemes, three different test cases have been considered. The first two test cases are the same ones as those which have already been applied by Spotz and Carey [6] and Zhang [7]. In this section we apply the lexicographic Gauss-Seidel iteration scheme and compare the discretizations, based on Padé approximations $\mathrm{D4c}_{P02}$ (Eqs.24-26), $\mathrm{D6c}_{P02}$ (Eqs.45-47), $\mathrm{D6c}_{P04}$ (Eqs.33-36) and $\mathrm{D6c}_{P24}$ (Eqs.56-58). In addition the 4th- ($\mathrm{D4}_{2bl}$) and 6th-order ($\mathrm{D6}_{3bl}$) finite difference approximations are checked which result from Eq.7 and which require two and three boundary layers for the solution respectively.

The problems are discritized on a normalized cube $(0,1)^3$, where the mesh size h is varied between $h = 1/8$ and $h = 1/64$. All calculations are carried out as long as the norm of the residuum per grid point is reduced to $h^3\|\mathbf{r}\| \leq \epsilon_{res}$ (where $\epsilon_{res} = 10^{-10}$ for the case of 64-bit precision). The norm of the residuum is defined as

$$\|\mathbf{r}^{(n)}\| = \sqrt{\sum_{i,j,k} \left(\Delta_{i,j,k} u^{(n)}_{i,j,k} + f_{i,j,k}\right)^2} \tag{64}$$

Here, $u^{(n)}_{i,j,k}$ is the field approximation after n iterations.

All calculations were performed on a LINUX IBM-T30 notebook with a Pentium-IV 2 GHz processor and the program was compiled with the Intel Fortran 90 compiler `ifc`.

As a measure of accuracy the following error norms are considered.

$$\epsilon_{max} = \frac{1}{h_x h_y h_z} \frac{\max \|\mathbf{u}^{(ex)} - \mathbf{u}\|}{\sum_{i,j,k} |u_{i,j,k}|} \tag{65}$$

$$\epsilon_{avg} = \frac{\sum_{i,j,k} |u^{(ex)}_{i,j,k} - u_{i,j,k}|}{\sum_{i,j,k} |u_{i,j,k}|} \tag{66}$$

$$\epsilon_{rel} = \frac{\|\mathbf{u}^{(ex)} - \mathbf{u}\|}{\|\mathbf{u}\|} \tag{67}$$

$$\tag{68}$$

where $u^{(ex)}_{i,j,k}$ is the exact solution of the model problem.

Before starting with test cases, the convergence characteristic of the different schemes is investigated. First of all it is clear that the schemes have to be positive definit. This is equivalent to the statement $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0,\ \forall\mathbf{x}$ or that all eigenvalues λ of $\mathbf{A}$ are greater zero. The second requirement was checked by calculating the smallest eigenvalue of the matrices $\mathbf{A}$ which were found to be positive for all cases investigated.

Another important characteristic is the convergence behavior of the finite difference schemes. In an empirical way the rate of convergence, c, is calculated as the asymtotic decay of the residuum

$$c = \lim_{n\to\infty} \frac{\|\mathbf{r}^{(n)}\|}{\|\mathbf{r}^{(n-1)}\|} \tag{69}$$

On the other hand it may be also calculated as the spectral radius of the iteration matrix $\mathbf{C}$ of the numerical scheme. For the Gauss-Seidel scheme $\mathbf{C}$ is given by

$$\mathbf{C} = (\mathbf{D} - \mathbf{L})^{-1}\mathbf{U} \tag{70}$$

where $\mathbf{D}$ is the diagonal, $-\mathbf{L}$ the lower triangular and $-\mathbf{U}$ the upper triangular part of the operator matrix $\mathbf{A}$. The spectral radius is then given as $\rho(\mathbf{C}) = \max_i \lambda_i(\mathbf{C})$, where λ_i is an eigenvalue of matrix $\mathbf{C}$. Both methods were applied to calculate the convergence rate. Table 5 shows the comparison of the different schemes. First of all it is found that the empirical determination of c is in a very good agreement with the calculation of the spectral radius. Furthermore it is found that for the different schemes a different performance has to be expected. Especially for large lattices, the schemes based on the Padé approximation perform fairly superior with respect to approximations based on the Taylor expansion, Eq.7. For the smallest lattices this seems to be different. However, it has to be noted that in the case of $\mathrm{D6}_{3bl}$ three boundary layers have to be specified, which reduces the number of unknowns in that case from $(2^3+1)^3 = 729$ to $(2^3-5)^3 = 27$! In the case, where two boundary layers have to be specified ($\mathrm{D4}_{2bl}$, $\mathrm{D6c}_{P02}$ and $\mathrm{D6c}_{P04}$) the number of unknowns on the smallest lattice is $(2^3-3)^3 = 125$, whereas for one specified boundary layer ($\mathrm{D4c}_{P02}$) there are $(2^3-1)^3 = 343$ unknowns. This imbalance is strongly reduced if one moves to larger lattices. To find the proper spectral radii the iteration matrix was constructed consequently with eliminated boundary conditions.

3.1 Test case 1

The first test case consists in solving the Poisson equation with the following source term distribution and vanishing boundary conditions

$$f_{i,j,k} = 3\pi^2 \sin(\pi i h_x)\sin(\pi j h_y)\sin(\pi k h_z) \tag{71}$$

for which the analytical solution is found to be

$$u_{i,j,k} = \sin(\pi i h_x)\sin(\pi j h_y)\sin(\pi k h_z) \tag{72}$$

In Table 1 the results are reported for the 4th- and 6th-order solvers in terms of number of iterations, CPU time and the three error norms, Eqs. 65-67.

3.2 Test case 2

The second test case consists of the solution of the Laplace equation, i.e. $f_{i,j,k} = 0$, with the following Dirichlet boundary conditions

$$\begin{array}{ll} u_{i,j,k} = \sin(\pi j h_y)\sin(\pi k h_z) & i = 0 \\ u_{i,j,k} = 2\sin(\pi j h_y)\sin(\pi k h_z) & i = nh \\ u_{i,j,k} = 0 & j,k = \{0, nh\} \end{array} \tag{73}$$

The analytical form of the field is found to be

$$u_{i,j,k} = \frac{\sin(\pi j h_y)\sin(\pi k h_z)}{\sin(\sqrt{2}\pi)} \left[2\sinh(\sqrt{2}\pi i h_x) + \sinh(\sqrt{2}\pi(1 - i h_x))\right] \tag{74}$$

3.3 Test case 3

This test case consists in solving the Poisson equation for a unit point charge, located in the centre of the cube. The source term is thereby given as

$$f_{i,j,k} = \begin{cases} \dfrac{1}{h_x h_y h_z} & : \quad i,j,k = 0.5 \\ 0 & : \quad \text{else} \end{cases} \tag{75}$$

and the analytical solution is

$$u_{i,j,k} = \frac{1}{\sqrt{(i h_x - 0.5)^2 + (j h_y - 0.5)^2 + (k h_z - 0.5)^2}} \tag{76}$$

The special feature of this test case is the singularity at $(ih_x, ih_y, ih_z) = \mathbf{0}$. This singularity cannot be accurately be described on a descrete grid. Therefore this test case will show relatively large errors for each approximation. Nevertheless, it will be interesting to see how good solutions will be obtained in the neighborhood of the singularity. Results are shown in Fig. 1.

3.4 Errors

In Tables 1-4 the error norms from Eqs. 65-67 are reported for the different approximation schemes. As expected it is found that the 6th-order schemes show a higher accuracy than the 4th-order ones. There is, however, the fact that some of the schemes (D6c$_{P04}$, D6$_{3bl}$) show an increase of the error from mesh size $h = 1/32$ to $h = 1/64$ for test case 1 and 2. An attempt to explain this phenomenon is made in terms of an error analysis.

In the following the norm of the absolute difference of the potential energy is considered, $\|\mathbf{u}^{(ex)} - \mathbf{u}\|$. This norm is connected with the residual norm according to

$$\|\mathbf{u}^{(ex)} - \mathbf{u}\| = h^2 \|\mathbf{A}^{-1}\mathbf{r}\| \tag{77}$$

$$\approx \frac{h^2}{|\lambda_{min}|}\|\mathbf{r}\| \tag{78}$$

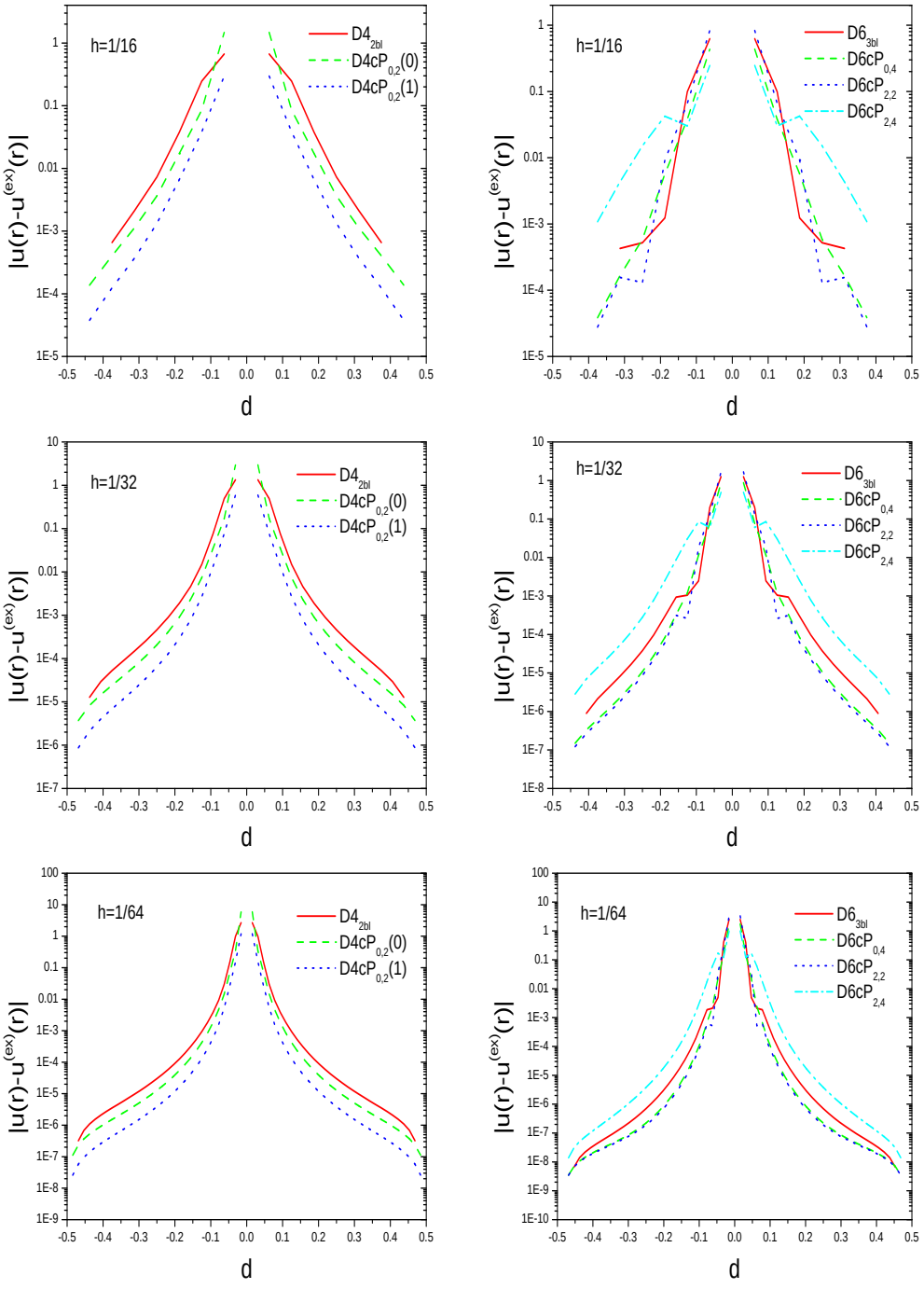

Figure 1: Absolute error $|u(r) - u^{(ex)}(r)|$ for a unit point charge in the center of the cube (test case 3) as function of distance from the charge. Shown is the behavior along the x-direction with fixed y- and z-coordinates (x,y=0,z=0). Left: 4th-order solvers $\mathrm{D4}_{2bl}$, $\mathrm{D4cP}_{0,2}(c_1 = 0)$ and $\mathrm{D4cP}_{0,2}(c_1 = 1)$; Right: 6th-order solvers $\mathrm{D6}_{3bl}$, $\mathrm{D6cP}_{0,4}$, $\mathrm{D6cP}_{2,2}$ and $\mathrm{D6cP}_{2,4}$ with $c_i = 1$ (cmp. Eqs.32,44,55).

where λ_{min} is the smallest eigenvalue of the matrix $\mathbf{A}$. As a criterion to end the iteration process for the solution of the test cases, a value of the residuum per grid point $h^3\|\mathbf{r}\| \leq \epsilon_{res}$ (for 8 bytes precision $\epsilon_{res} = 10^{-10}$) was chosen. Consequently, a natural limit for the accuracy of the potential difference is

$$\|\mathbf{u}^{(ex)} - \mathbf{u}\| = \frac{\epsilon_{res}}{h|\lambda_{min}|} \tag{79}$$

which is called the *residual error bound* henceforth.

Another limit in accuracy is connected with the condition number, κ, of the matrix $\mathbf{A}$, which gives a measure for how strong the inverse of the matrix changes, if the matrix itself is changed. Together with the machine epsilon, ε_{mach}, this defines an approximate error bound, $\epsilon_B = \varepsilon_{mach}\kappa$, where $\varepsilon_{mach} = 2.22 \times 10^{-16}$ for 8 bytes accuracy. This can be related to the potential energy difference by

$$\|\mathbf{u}^{(ex)} - \mathbf{u}\| \approx \epsilon_B\|\mathbf{u}\| \tag{80}$$

This is another criterion for a lower bound, the *machine error bound*.

The condition number of a matrix may be calculated via

$$\kappa = \frac{|\lambda_{max}|}{|\lambda_{min}|} \tag{81}$$

where λ_{max} is the biggest eigenvalue of the matrix. Since for the smallest mesh size the matrices are of size $32^3 \times 32^3$ and $64^3 \times 64^3$, methods working on the whole matrix are doomed to fail due to memory limits of the computers. Fortunately the matrices are sparse (N^*/h^3 entries, where N^* is the number of entries of the stencils) and so index field addressing methods may be used. Since $\mathbf{A}$ is symmetric, the *Rayleigh Quotient* method is used here [8], which gives a rather reliable estimate for the largest eigenvalue. The smallest eigenvalue of $\mathbf{A}$ is calculated with the help of a shifted matrix $\mathbf{B}$

$$\mathbf{B} = \mathbf{A} - \frac{2}{3}|\lambda_{max}(\mathbf{A})|\ \mathbf{I} \tag{82}$$

For this matrix the largest eigenvalue is constructed again by the *Rayleigh Quotient* method. To obtain a rather good estimate for $\lambda_{min}(\mathbf{A})$ one has to shift back [9]

$$\lambda_{min}(\mathbf{A}) = \frac{2}{3}|\lambda_{max}(\mathbf{A})| - |\lambda_{max}(\mathbf{B})| \tag{83}$$

From these calculations it is first of all found that (besides $\mathcal{P}_{2,4}$) the Padé approximation schemes have smaller condition numbers than the ones based on Eq.7 (cf. Fig. 2). This is mainly due to the fact that the largest eigenvalue is reduced. This fact was also observed for the corresponding iteration matrices which have smaller spectral radii (cf. Table 5).

Now the energy resolutions, Eq.79,80 are compared with the empirical norm of the potential difference.

As an example, Figs.3,4 show the results of test case 2 for the norms, obtained for the stencils D4c$_{P02}$, D6$_{3bl}$ and D6c$_{P04}$. It is found that in this case

Solver	Grid	Iter.	time	ϵ_{max}	ϵ_{avg}	ϵ_{rel}
$D4cP_{0,2}$	1/8	91	8.7×10^{-3}	1.07×10^{-3}	2.67×10^{-4}	2.67×10^{-4}
	1/16	365	0.189	6.48×10^{-5}	1.65×10^{-5}	1.65×10^{-5}
	1/32	1395	13.99	4.01×10^{-6}	1.03×10^{-6}	1.03×10^{-6}
	1/64	5294	1031	2.31×10^{-7}	5.96×10^{-8}	5.96×10^{-8}
$D4_{2bl}$	1/8	91	4.4×10^{-3}	7.0×10^{-4}	7.8×10^{-5}	1.21×10^{-4}
	1/16	516	0.117	5.46×10^{-5}	9.92×10^{-6}	1.20×10^{-5}
	1/32	2286	9.90	3.71×10^{-6}	8.14×10^{-7}	8.90×10^{-7}
	1/64	9298	1681	2.23×10^{-7}	5.32×10^{-8}	5.55×10^{-8}
$D6cP_{0,4}$	1/8	52	4.7×10^{-3}	1.25×10^{-5}	1.46×10^{-6}	2.24×10^{-6}
	1/16	274	0.21	2.31×10^{-7}	4.29×10^{-8}	5.12×10^{-8}
	1/32	1193	14.9	2.55×10^{-9}	4.90×10^{-10}	5.60×10^{-10}
	1/64	4819	1155	1.85×10^{-8}	4.35×10^{-9}	4.56×10^{-9}
$D6cP_{2,2}$	1/8	41	9.9×10^{-3}	4.4×10^{-4}	5.32×10^{-5}	8.11×10^{-5}
	1/16	208	0.367	3.05×10^{-5}	5.75×10^{-6}	6.82×10^{-6}
	1/32	897	22.7	2.02×10^{-6}	4.51×10^{-7}	4.88×10^{-7}
	1/64	3608	1350	1.48×10^{-7}	3.56×10^{-8}	3.70×10^{-8}
$D6cP_{2,4}$	1/8	196	1.9×10^{-2}	1.08×10^{-4}	1.13×10^{-5}	1.78×10^{-5}
	1/16	1321	0.95	2.51×10^{-6}	4.36×10^{-7}	5.35×10^{-7}
	1/32	6154	72.3	3.83×10^{-8}	8.28×10^{-9}	9.12×10^{-9}
	1/64	25447	6885	1.78×10^{-8}	4.07×10^{-9}	4.32×10^{-9}
$D6_{3bl}$	1/8	42	2.0×10^{-3}	8.22×10^{-6}	2.51×10^{-7}	7.90×10^{-7}
	1/16	413	0.116	2.70×10^{-7}	3.25×10^{-8}	4.88×10^{-8}
	1/32	2175	16.8	9.88×10^{-10}	9.62×10^{-11}	1.45×10^{-10}
	1/64	9500	1943	1.82×10^{-8}	3.83×10^{-9}	4.24×10^{-9}

Table 1

Table 1: Errors and performance for test case 1, solved with Gauss-Seidel algorithm with 4th- and 6th-order approximations for the Laplace operator as described in the text.

the lowest error bound is the *machine error bound*, which covers a range from 10^{-14} to 10^{-11} for large and small grid intervals respectively. The *residual error bound* is strictly higher, ranging between 10^{-9} to 10^{-6}. Concerning the empirical norms resulting from the iteration process, it is found that they decrease as long as they do not cross or come close to the *residual error bound*. From there the norm starts to increase again. It is found that for the 4th-order solvers the empirical norm does not cross these error bounds and consequently is decreasing with decreasing grid size h. For 6th-order solvers this is different and it is seen that the accuracy obtained is strictly limited by the *residual error bound*. For a comparison the same calculations were done for 4 Bytes accuracy. In this case, $\varepsilon = 1.19 \times 10^{-7}$, which increases strongly the *machine error bound* from 10^{-5} to 10^{-2}. For the stopping criterion in the iteration scheme a residual norm per grid point of $\epsilon_{res} \approx 10^{-5}$ was specified. Correspondingly the *residual error bound*

Solver	Grid	Iter.	time	ϵ_{max}	ϵ_{avg}	ϵ_{rel}
$D4cP_{0,2}$	1/8	87	9.3×10^{-3}	9.51×10^{-4}	2.61×10^{-4}	2.50×10^{-4}
	1/16	354	0.19	6.68×10^{-5}	1.91×10^{-5}	1.76×10^{-5}
	1/32	1361	15.2	4.50×10^{-6}	1.31×10^{-6}	1.19×10^{-6}
	1/64	5169	1009	3.51×10^{-7}	1.03×10^{-7}	9.32×10^{-8}
$D4_{2bl}$	1/8	87	4.6×10^{-3}	4.70×10^{-4}	5.62×10^{-5}	8.52×10^{-5}
	1/16	500	0.117	4.79×10^{-5}	9.62×10^{-6}	1.08×10^{-5}
	1/32	2225	13.9	3.82×10^{-6}	9.38×10^{-7}	9.37×10^{-7}
	1/64	9060	1638	3.25×10^{-7}	8.77×10^{-8}	8.32×10^{-8}
$D6cP_{0,4}$	1/8	50	4.7×10^{-3}	1.08×10^{-5}	1.35×10^{-6}	2.02×10^{-6}
	1/16	266	0.20	2.27×10^{-7}	4.66×10^{-8}	5.18×10^{-8}
	1/32	1162	14.4	4.15×10^{-8}	1.04×10^{-8}	1.04×10^{-8}
	1/64	4698	1128	5.75×10^{-8}	1.68×10^{-8}	1.59×10^{-8}
$D6cP_{2,2}$	1/8	39	8.5×10^{-3}	4.4×10^{-4}	5.32×10^{-5}	8.11×10^{-5}
	1/16	202	0.35	3.05×10^{-5}	5.75×10^{-6}	6.82×10^{-6}
	1/32	874	21.9	2.02×10^{-6}	4.51×10^{-7}	4.88×10^{-7}
	1/64	3519	1314	1.48×10^{-7}	3.56×10^{-8}	3.70×10^{-8}
$D6cP_{2,4}$	1/8	187	1.9×10^{-2}	1.09×10^{-4}	1.23×10^{-5}	1.90×10^{-5}
	1/16	1279	0.91	3.36×10^{-6}	6.47×10^{-7}	7.39×10^{-7}
	1/32	5984	67.2	1.14×10^{-7}	2.74×10^{-8}	2.77×10^{-8}
	1/64	24782	6702	6.15×10^{-8}	1.65×10^{-8}	1.59×10^{-8}
$D6_{3bl}$	1/8	40	3.1×10^{-3}	1.36×10^{-5}	4.31×10^{-7}	1.34×10^{-6}
	1/16	400	0.10	6.01×10^{-7}	7.83×10^{-8}	1.11×10^{-7}
	1/32	2112	13.7	2.51×10^{-8}	5.14×10^{-9}	5.70×10^{-9}
	1/64	9242	1898	5.71×10^{-8}	1.42×10^{-8}	1.42×10^{-8}

Table 2: Errors and performance for test case 2 (Laplace equation) solved with Gauss-Seidel algorithm with 4th- and 6th-order approximations for the Laplace operator as described in the text.

increases by factor of 10^5 with respect to the 8 bytes precision. In the present case the separation between the different bands is not well pronounced any more and since the accuracy obtained on the finest grid in the iteration scheme is already very close to the highest error limit, the accuracy decreases for most of the approximation schemes already from the finest mesh size, $h = 1/8$.

This shows that the accuracy in the present calculations is actually limitited by the criterion imposed for the residual calculation. Due to the fact that the smallest eigenvalue decreases for finer mesh sizes, the *residual error bound* increases for a fixed stopping criterion. This leads to the fact that from the point where the norm of the potential difference comes close to the *residual error bound* the accuracy starts to decrease. This can only be avoided if one introduces a renormalized stopping criterion which accounts for the increasing *residual error bound*. However, for very large matrices one is limited in any case by the *machine error bound* criterion and as a matter of fact, when reaching this

Solver	Grid	Iter.	time	ϵ_{max}	ϵ_{avg}	ϵ_{rel}
$D4cP_{0,2}$	1/8	100	7.8×10^{-3}	0.24	5.47×10^{-3}	2.85×10^{-2}
	1/16	397	0.15	0.55	1.59×10^{-3}	2.06×10^{-2}
	1/32	1527	9.03	1.16	4.28×10^{-4}	1.48×10^{-2}
	1/64	5829	1162	2.40	1.1×10^{-4}	1.05×10^{-2}
$D4_{2bl}$	1/8	100	6.1×10^{-3}	0.11	3.88×10^{-3}	1.51×10^{-2}
	1/16	561	0.12	0.25	1.30×10^{-3}	1.11×10^{-2}
	1/32	2501	14.2	0.53	3.53×10^{-4}	7.93×10^{-3}
	1/64	10229	1926	1.1	9.16×10^{-5}	5.65×10^{-3}
$D6cP_{0,4}$	1/8	57	5.1×10^{-3}	0.146	2.77×10^{-3}	1.68×10^{-2}
	1/16	298	0.22	0.328	8.04×10^{-4}	1.22×10^{-2}
	1/32	1305	14.0	0.697	2.14×10^{-4}	8.73×10^{-3}
	1/64	5302	1295	1.438	5.51×10^{-5}	6.21×10^{-3}
$D6cP_{2,2}$	1/8	43	3.9×10^{-3}	6.06×10^{-2}	2.44×10^{-3}	1.08×10^{-2}
	1/16	220	0.15	0.136	7.03×10^{-4}	7.85×10^{-3}
	1/32	954	11.39	0.289	1.87×10^{-4}	5.62×10^{-3}
	1/64	6862	1000	0.597	4.82×10^{-5}	4.0×10^{-3}
$D6cP_{2,4}$	1/8	214	1.7×10^{-2}	7.07×10^{-2}	1.52×10^{-3}	8.35×10^{-3}
	1/16	1436	1.02	0.159	5.33×10^{-4}	6.07×10^{-3}
	1/32	6730	89.6	0.338	1.43×10^{-4}	4.34×10^{-3}
	1/64	27991	7709	0.697	3.69×10^{-5}	3.09×10^{-3}
$D6_{3bl}$	1/8	45	2.2×10^{-3}	9.6×10^{-2}	1.75×10^{-3}	1.14×10^{-2}
	1/16	449	0.12	0.23	6.69×10^{-4}	8.90×10^{-3}
	1/32	2377	13.6	0.49	1.78×10^{-4}	6.37×10^{-3}
	1/64	10444	2069	1.02	4.60×10^{-5}	4.53×10^{-3}

Table 3: Errors and performance for the field distribution of a point charge (test case 3) solved with Gauss-Seidel algorithm with 4th- and 6th-order approximations for the Laplace operator as described in the text. For the cases of Padé approximations, factors c_i were set to zero.

limit, a further refinement of the mesh will not give higher but lower accuracy.

4 Conclusions

In this paper different forms of high order compact solvers for the three dimensional Poisson equation have been derived on the bases of a Padé approximation applied to the Taylor expansion of the finite difference operator, Eq.7. It was shown that both the accuracy and the convergence characteristic obtained with these new approximations is superior with respect to the *straight-forward* implementations of the Laplace stencil. The 4th-order Padé approximation led to a similar expression as found in Ref.[6]. However, we used here an extended expression, where the *rhs* of the Poisson equation was modified according to a

Solver	Grid	Iter.	time	ϵ_{max}	ϵ_{avg}	ϵ_{rel}
$D4cP_{0,2}$	1/8	100	7.9×10^{-3}	4.90×10^{-2}	2.62×10^{-3}	8.74×10^{-3}
	1/16	397	0.15	0.11	7.78×10^{-4}	6.34×10^{-3}
	1/32	1527	9.54	0.23	2.09×10^{-4}	4.54×10^{-3}
	1/64	5829	1140	0.48	5.41×10^{-5}	3.23×10^{-3}
$D6cP_{0,4}$	1/8	57	4.8×10^{-3}	7.20×10^{-2}	2.15×10^{-3}	1.02×10^{-2}
	1/16	298	0.22	0.162	6.24×10^{-4}	7.38×10^{-3}
	1/32	1305	14.6	0.344	1.66×10^{-4}	5.28×10^{-3}
	1/64	5302	1293	0.709	4.28×10^{-5}	3.76×10^{-3}
$D6cP_{2,2}$	1/8	43	3.6×10^{-3}	0.14	1.95×10^{-3}	1.56×10^{-2}
	1/16	220	0.15	0.31	5.64×10^{-4}	1.12×10^{-2}
	1/32	954	15.32	0.65	1.50×10^{-4}	8.06×10^{-3}
	1/64	3862	1001	1.35	3.86×10^{-5}	5.73×10^{-3}
$D6cP_{2,4}$	1/8	214	2.1×10^{-2}	4.1×10^{-2}	1.53×10^{-3}	5.82×10^{-3}
	1/16	1436	1.02	9.28×10^{-2}	5.46×10^{-4}	4.29×10^{-3}
	1/32	6730	84.3	0.195	1.47×10^{-4}	3.07×10^{-3}
	1/64	27991	12005	0.402	3.78×10^{-5}	2.18×10^{-3}

Table 4: Same as Table 3 but for the cases of Padé approximations where factors c_i were set to one, i.e. taking into account a higher order approximation of the *rhs*.

higher order approximation scheme. This turned out to give results of better quality in the case of discontinuous source functions, i.e. the case of a unit point charge. In the case of continuous source distributions, this higher order approximation of the *rhs* did not improve the accuracy. The same is true for the 6th-order approximations. Also here, problems with discontinuous source distributions did benefit from a higher order approximation of the *rhs*. This is understandable from the fact that the higher order approximation tends to smooth out discontinuities which become better behaved in the solution. Although the higher order approximation of the *rhs* leads to more complex expressions, this approach seems to be justified since the computational overhead is reduced to a modification of the source function, which can be done before starting the iterative process.

As a matter of fact [m,n]-Padé approximations of the bracketed expression in Eq.7 approximate the Laplace operator with order $\mathcal{O}(h^{m+n+2})$. For the 4th-order approximation, the only non-trivial choice is a [0,2]-approximation. For the case of 6th-order the two cases [0,4] and [2,2] were investigated. However, as a test case, also a [2,4]-Padé approximation was used, where in the resulting expression, terms only up to 6th-order were taken into account. This case demonstrated that with higher order Padé-approximations no better solutions may be found. It was shown that this approximation turns out to lead to a larger condition number of the operator matrix $\mathbf{A}$ and to larger spectral radii. The former property leads to a less stable solution, the latter one to a slower

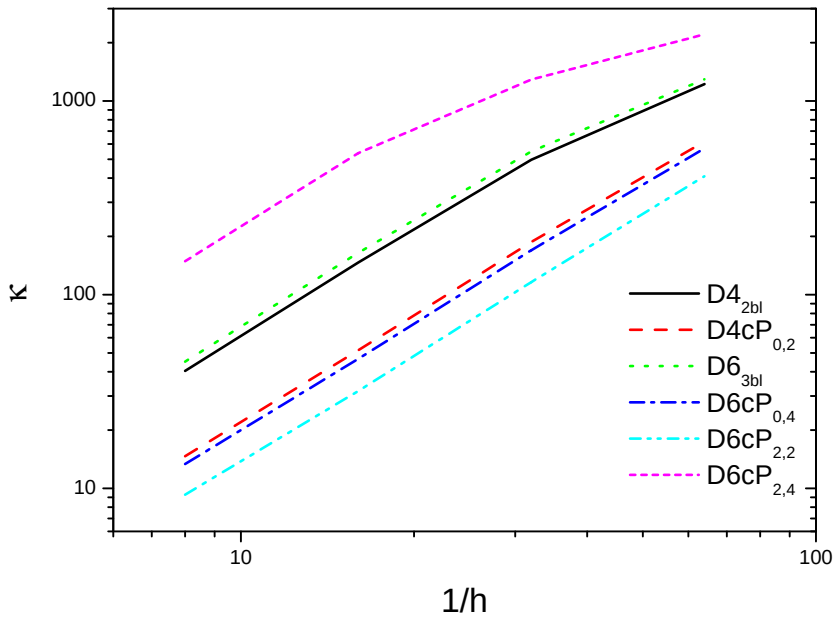

Figure 2: Condition numbers of the operator matrices **A** for Padé approximations and Taylor expansions of the Laplace operator.

convergence. It may be expected that this behavior is even more pronounced for higher order Padé approximations, where terms are kept only up to order six.

In this article the 6th-order approximations require two boundary layers for a solution of a boundary value problem. This may lead to problems where the *true* potential function is only known exactly on one boundary layer. In this case, one has to switch to lower order representations on and near the boundary. However, it is not clear how this local change of order does affect the overall solution of the problem. Investigations in this direction will be done in future. In addition the finite difference schemes of 6th-order cannot yet be used in a straight-forward way for multigrid techniques. However, double discretization techniques, i.e. different discretization of field and residuum equations, will lead to the possibility to apply also the new high order approximations to multigrid methods. Work in this direction is in progress.

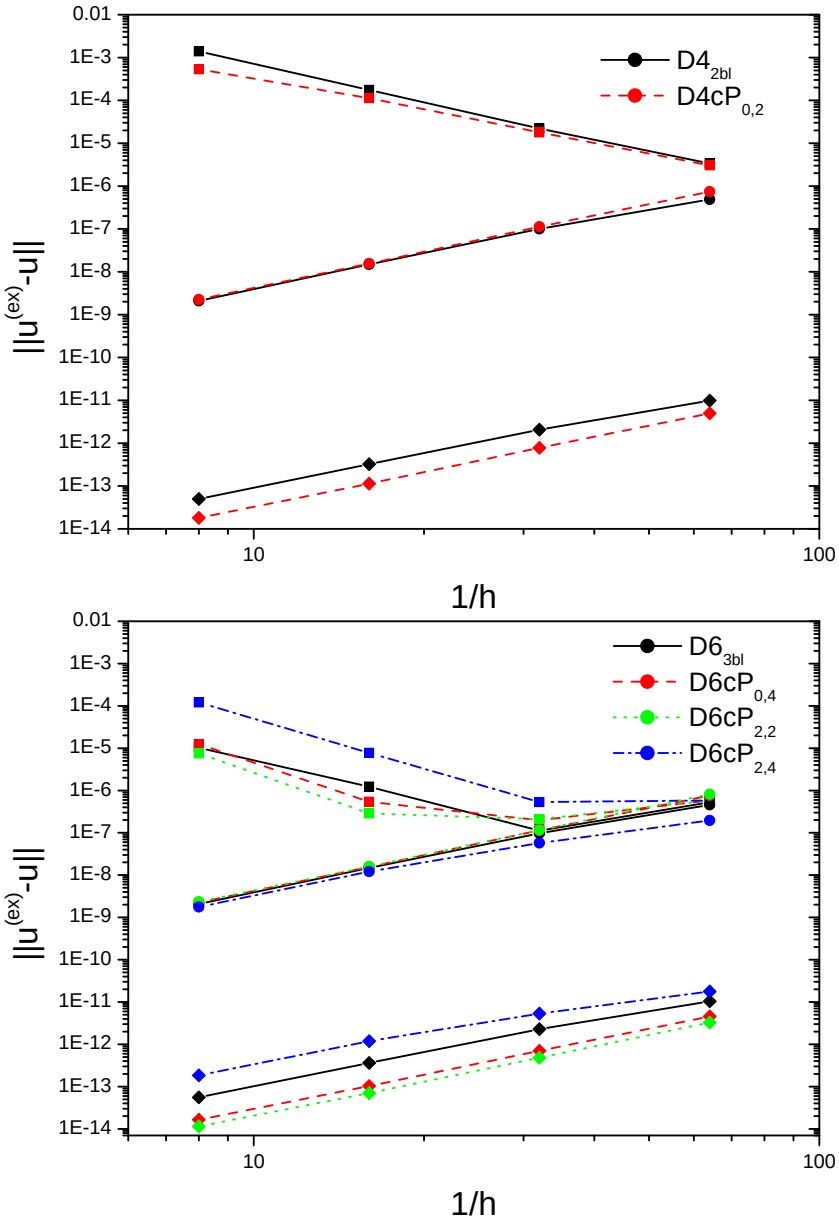

Figure 3: Comparison between absolute error norm $\|\mathbf{u}-\mathbf{u}^{(ex)}\|$ ($\square$) for test case 2 with the *machine error bound* ($\diamondsuit$) and the *residual error bound* ($\bigcirc$) for calculations with 8 bytes precision. Upper figure: 4-th order approximations $\mathrm{D4}_{2bl}$, $\mathrm{D4cP}_{0,2}(c_1 = 0)$ and $\mathrm{D4cP}_{0,2}(c_1 = 1)$; Lower figure: 6-th order approximations $\mathrm{D6}_{3bl}$, $\mathrm{D6cP}_{0,4}$, $\mathrm{D6cP}_{2,2}$ and $\mathrm{D6cP}_{2,4}$ with $c_i = 1$ (cmp. Eqs.32,44,55).

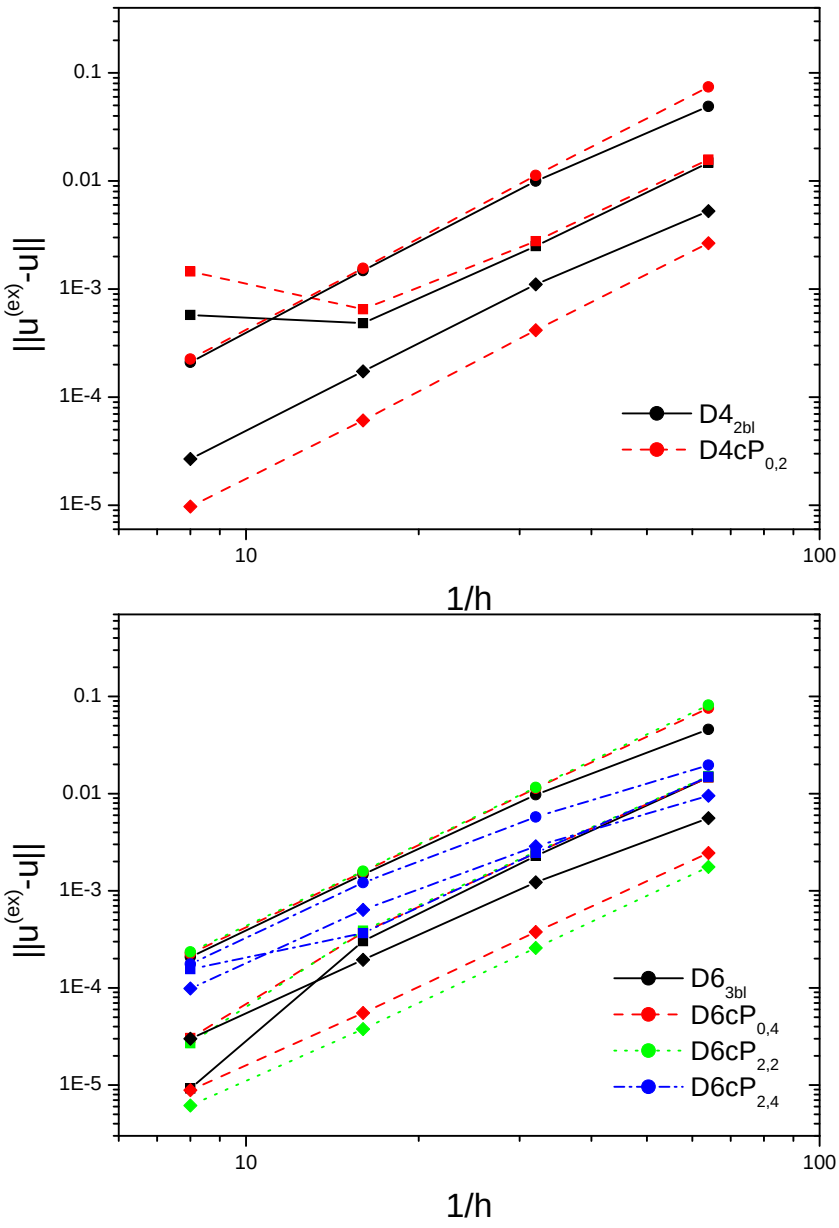

Figure 4: The same as shown in Fig.3 but for calculations with 4 bytes precision (absolute error norm $\|\mathbf{u} - \mathbf{u}^{(ex)}\|$ (□); *machine error bound* (♢); *residual error bound* (◯)).

Solver	$D4_{2bl}$		$D4cP_{0,2}$		$D6_{3bl}$	
Grid	$\frac{\|\mathbf{r}^{(n)}\|}{\|\mathbf{r}^{(n-1)}\|}$	$\rho(\mathbf{C})$	$\frac{\|\mathbf{r}^{(n)}\|}{\|\mathbf{r}^{(n-1)}\|}$	$\rho(\mathbf{C})$	$\frac{\|\mathbf{r}^{(n)}\|}{\|\mathbf{r}^{(n-1)}\|}$	$\rho(\mathbf{C})$
1/8	0.7826	0.7831	0.7921	0.7921	0.5638	0.5638
1/16	0.9592	0.9587	0.9437	0.9432	0.9486	0.9491
1/32	0.9911	–	0.9856	–	0.9906	–
1/64	0.9979	–	0.9964	–	0.9980	–

Solver	$D6cP_{0,4}$		$D6cP_{2,2}$		$D6cP_{2,4}$	
Grid	$\frac{\|\mathbf{r}^{(n)}\|}{\|\mathbf{r}^{(n-1)}\|}$	$\rho(\mathbf{C})$	$\frac{\|\mathbf{r}^{(n)}\|}{\|\mathbf{r}^{(n-1)}\|}$	$\rho(\mathbf{C})$	$\frac{\|\mathbf{r}^{(n)}\|}{\|\mathbf{r}^{(n-1)}\|}$	$\rho(\mathbf{C})$
1/8	0.6494	0.6506	0.5778	0.5772	0.8912	0.8901
1/16	0.9246	0.9265	0.9021	0.8952	0.984	0.9802
1/32	0.9831	–	0.9776	–	0.997	–
1/64	0.9960	–	0.9947	–	0.999	–

Table 5: Convergence factors for different types of finite difference schemes. Compared are factors obtained from the *empirical* residual reduction from the iterations, $\|\mathbf{r}^{(n)}\|/\|\mathbf{r}^{(n-1)}\|$, and the spectral radius $\rho(\mathbf{C})$ of the Gauss-Seidel iteration matrix, Eq. 70.

References

[1] W. F. Ames. *Numerical Methods for Partial Differential Equations.* Academic Press, New York, 1977.

[2] U. Trottenberg, C. W. Oosterlee, and A. Schüller. *Multigrid.* Academic Press, San Diego, 2001.

[3] Y. Kwon and J. W. Stephenson. Single cell finite difference approximations for Poisson's equation in three variables. *Appl. Math. Notes*, 2:13–20, 1982.

[4] W. Liao, J. Zhu, and A. Q. M. Khaliq. An efficient high order algorithm for solving systems of reaction-diffusion equations. *J. Numer. Meth. Part. Diff. Eqs.*, 18:340–354, 2002.

[5] Y. Gu, W. Liao, and J. Zhu. An efficient high order algorithm for solving systems of 3-D reaction-diffusion equations. *J. Comp. Appl. Math.*, 155:1–17, 2003.

[6] W. F. Spotz and G. F. Carey. A high-order compact formulation for the 3d Poisson equation. *Num. Meth. Part. Diff. Equat.*, 12:235–243, 1996.

[7] J. Zhang. Fast and high accuracy multigrid solution of the three dimensional Poisson equation. *J. Comp. Phys..*, 143:449–461, 1998.

[8] Y. Saad. *Numerical Methods for Large Eigenvalue Problems.* Manchester University Press, Manchester, 1992.

[9] For the smaller matrices this was checked with routines DSYTRD and DSTERF from LAPACK.