

# Discretization error estimates in maximum norm for convergent splittings of matrices with a monotone preconditioning part

Owe Axelsson, János Karátson<sup>2,\*</sup>

March 30, 2016

<sup>1</sup> Institute of Geonics AS CR, IT4 Inovations,  
Ostrava, The Czech Republic,  
owea@it.uu.se,

<sup>2</sup> Department of Applied Analysis & MTA-ELTE Numerical Analysis  
and Large Networks Research Group, ELTE University;  
Department of Analysis, Technical University; Budapest, Hungary;  
karatson@cs.elte.hu

Keywords: finite difference method, error estimates, matrix splitting, preconditioning

## Abstract

For finite difference matrices that are monotone, a discretization error estimate in maximum norm follows from the truncation errors of the discretization. It enables also discretization error estimates for derivatives of the solution. These results are extended to convergent operator splittings of the difference matrix where the major, preconditioning part is monotone but the whole operator is not necessarily monotone.

## 1 Introduction

Finite difference methods for elliptic problems [3, 9, 11] are most suitable for regular grids. They can often be formulated so that the discrete operator is a monotone matrix. This enables a simple estimate of the discretization error in maximum norm with a best constant factor in the upper bound. In addition, one can approximate first and higher order derivatives of the solution with the same order of convergence, if the solution is sufficiently regular. The estimates are local, and therefore hold even for problems with discontinuous coefficients on macroelements, if the solution is regular in the interior of each element. One can consider here both rectangular and hexagonal meshes. The difference operator is particularly simple for hexagonal meshes. High order difference approximations can be constructed either with use of approximations on locally extended meshes of the higher order derivative terms in the truncation error, or with use of extrapolation for regularly refined meshes.

In the case when the matrix is not monotone, one can split it in a monotone and remainder term. If this, with proper scaling of the matrices, leads to a convergent splitting, one can still estimate the maximum norm of the error, but with a factor that becomes larger when the splitting leads to a larger convergence factor. This approach can be illustrated for the Helmholtz equation. Other

---

\* Supported by the Hungarian Scientific Research Fund OTKA, No. 112157

possible applications might arise for the system of elasticity equation, using a splitting into the divergence and grad div terms, or proper splittings can be based on equivalent operator pairs, see [4] for examples of equivalent operators.

In this paper, first discretization error estimates for monotone matrices are described. In Section 3 some high order difference approximations are derived. Estimates using convergent splittings are presented in Section 4, which includes the case of the Helmholtz equation, illustrated by a numerical test.

## 2 Preliminaries: discretization error estimates for monotone matrices

Recall that a discretization matrix  $\mathcal{L}_h$  on a difference mesh  $\Omega_h$ , where  $h$  denotes the mesh size, is called *monotone* if

$$\mathcal{L}_h u \geq 0 \quad \text{implies} \quad u \geq 0. \quad (2.1)$$

It is readily seen that monotone matrices are nonsingular, because if  $\mathcal{L}_h u \leq 0$  then  $\mathcal{L}_h(-u) \geq 0$  so  $-u \geq 0$ , i.e.  $u \leq 0$ . Hence, if  $\mathcal{L}_h u = 0$  then both  $u \geq 0$  and  $u \leq 0$ , that is  $u = 0$ . Further, a nonsingular discretized operator (matrix)  $\mathcal{L}_h = A$  is monotone if and only if  $A^{-1} \geq 0$ , i.e. the entries of its inverse are nonnegative. The sufficiency follows immediately. To show the necessity part, if  $A^{-1}$  contains a negative entry in position  $(i, j)$ , then  $A_h u = e_j$  (the  $j$ th unit vector) has a solution with  $i$ th component  $u_i = (A_h^{-1} e_j)_i < 0$ , so  $A_h$  cannot be monotone.

If  $A = M - R$ , where  $M$  is monotone and  $M^{-1}R \geq 0$ , and the splitting is convergent, i.e.  $\varrho(M^{-1}R) < 1$ , where  $\varrho(\cdot)$  is the spectral radius, then  $A$  is monotone. Such a splitting is called a convergent weak regular splitting [10]. This is seen simply by expanding the inverse of  $M^{-1}A = I - M^{-1}R$  in a Neumann series. In many applications one gets a difference operator  $\mathcal{L}_h = D - R$  of positive type, i.e. where  $D$  is monotone, and  $R \geq 0$ . Then this operator is monotone if  $\varrho(D^{-1}R) < 1$ . If  $D$  is diagonal, such a matrix is called a diagonally dominant  $M$ -matrix.

A major advantage of dealing with monotone matrices is that the inverse of the matrices is bounded in maximum norm,  $\|\cdot\|_\infty$ , uniformly with respect to the mesh parameter  $h$ . This leads to a simple and useful discretization error estimate. To see this, let  $v \geq 0$  be a normalized vector, i.e.  $\max_i v(x_i) = 1$  for which  $\mathcal{L}_h v \geq \alpha$ ,  $\alpha > 0$ . (Such a vector or function is called a barrier function for the operator.) Then with  $\mathcal{L}_h v = \alpha e$ ,

$$1 = \|v\|_\infty = \alpha \|\mathcal{L}_h^{-1} e\|_\infty,$$

where  $e^T = (1, 1, \dots, 1)$ . Hence

$$\|\mathcal{L}_h^{-1}\|_\infty = \|\mathcal{L}_h^{-1} e\|_\infty \leq \frac{1}{\alpha}.$$

Therefore the best constant,  $\|\mathcal{L}_h^{-1}\|_\infty$  can be computed by solving  $\mathcal{L}_h v = e$ . Now let  $\mathcal{L}_h v_h = f_h$  be the discrete equation for an elliptic differential operator  $\mathcal{L}u = f$ , where  $\mathcal{L}_h$  is monotone. Then

$$\mathcal{L}_h(u - u_h) = \mathcal{L}_h u - f_h \quad (2.2)$$

is the truncation error, and the discretization error can be estimated by

$$\|u - u_h\|_\infty \leq \frac{1}{\alpha} \|\mathcal{L}_h u - f_h\|_\infty.$$

For regular problems, i.e. with a sufficiently differentiable solution  $u$ , one can estimate the truncation error

$$\tau_h := \mathcal{L}_h u - f_h$$

from a Taylor series expansion. Note that the remainder term of  $O(h^k)$  in the Taylor expansion can be written in integral form as  $\int_x^{x+h} (x+h-s)^{k-1}/(k-1)! u^{(k)}(s)ds$ .

There are various ways one can further improve the accuracy of the discrete solution. One can estimate the lowest order derivative terms in the Taylor expansion by use of difference approximations on a locally extended mesh or one can use higher order difference approximations, see Section 3. Another way is by extrapolating the solution on a mesh and its refinement. To show this, let  $\mathcal{L}u = f$  be a second order elliptic differential operator approximated by a difference operator with second order truncation error. Assume for simplicity given Dirichlet boundary conditions and assume that the solution  $u \in C^6(\Omega)$ . Let the truncation error satisfy

$$\mathcal{L}_h(u - u_h) = \mathcal{L}_h u - f = h^2 G u + O(h^4),$$

where  $G$  is a differential operator of fourth order. As an example, let  $\mathcal{L} = -\Delta$  be the Laplacian, then for a rectangular mesh  $G u = -\frac{1}{12}(u_x^{(4)} + u_y^{(4)})$ . Further let  $\varphi$  be the solution of

$$\mathcal{L}\varphi = G u \text{ in } \Omega, \quad \varphi = 0 \text{ on } \partial\Omega.$$

Then

$$\mathcal{L}_h \varphi - \mathcal{L}\varphi = h^2 G \varphi + O(h^4)$$

and

$$\mathcal{L}_h(u - u_h - h^2 \varphi) = \mathcal{L}_h(u - u_h) - h^2 \mathcal{L}\varphi - h^4 G \varphi + O(h^4) = O(h^4).$$

By use of the boundedness of the inverse of  $\mathcal{L}_h$ , we get

$$u(x_i) - u_h(x_i) = h^2 \varphi(x_i) + O(h^4), \quad h \rightarrow 0. \quad (2.3)$$

Since  $\varphi$  does not depend on  $h$ , we can combine solutions  $u_h$  and  $u_{2h}$  on the mesh  $\Omega_{2h}$  and its refinement  $\Omega_h$ , by extrapolation to get

$$\begin{aligned} u(x_i) - u_h(x_i) &= h^2 \varphi(x_i) + O(h^4), \\ u(x_i) - u_{2h}(x_i) &= 4h^2 \varphi(x_i) + O(h^4) \end{aligned}$$

that is,

$$u(x_i) - \frac{4u_h(x_i) - u_{2h}(x_i)}{3} = O(h^4),$$

which means that the extrapolated value  $(4u_h(x_i) - u_{2h}(x_i))/3$  has a discretization error  $O(h^4)$ . Under assumptions of sufficient regularity, one can improve this further to get even higher order of approximations.

Approximations for difference methods for regular meshes can be improved not only by extrapolation, but (2.3) shows also that we can compute approximations of derivatives of the solution to the same order of accuracy as for the solution itself. It holds namely

$$\frac{u(x_{i+1}) - u(x_{i-1}))}{2h} - \frac{u_h(x_{i+1}) - u_h(x_{i-1}))}{2h} - h^2 \frac{\varphi(x_{i+1}) - \varphi(x_{i-1}))}{2h} = O(h^3),$$

that is,

$$u'_x(x_i) - \frac{u_h(x_{i+1}) - u_h(x_{i-1}))}{2h} - h^2 \varphi'_x(x_i) = \frac{h^2}{6} u_x^{(3)}(x_i) + O(h^3)$$

so

$$u'_x(x_i) = \frac{u_h(x_{i+1}) - u_h(x_{i-1}))}{2h} + O(h^2),$$

if  $u \in C^6(\Omega)$ . A corresponding expression holds for the derivative in direction  $y$ .

In a similar way, assuming a correspondingly higher order of regularity of the solution, higher order approximations can be computed by extrapolation and even higher-order derivatives can be computed with error  $O(h^2)$  or higher order. Due to the existence of an  $h$ -expansion of the errors, we do not lose any accuracy even though we divide by powers of  $h$  to compute the derivatives. This is one of the major advantages with difference methods.

### 3 High order difference approximations and monotone matrices

We show now that there exist also high order difference approximations that lead to monotone matrices. There are two ways to obtain high order difference approximations for elliptic problems:

- (i) Use of a higher order difference approximation scheme involving some more mesh points and use of difference approximations of the lowest order derivative terms in the truncation error for the basic method.
- (ii) Use of extrapolation of the approximation solutions in a given and refined mesh.

We illustrate first the methods for a second order elliptic problem and consider then application for more involved problems. Advantages and disadvantages of the different approaches are discussed briefly. In the following we use the readily understandable notations,  $u_x^{(k)}$ ,  $u_y^{(k)}$ ,  $u_{x,y}^{(k,l)}$  etc. For instance, the mixed derivative  $u_{x,y}^{(4,2)} = u_{xxxxyy}$ .

#### 3.1 Rectangular meshes for the Laplacian

For a square mesh, for the standard five point and cross direction five point differences for a sufficiently regular solution it holds

$$\begin{aligned}\Delta_h^{(5)} u &= \Delta u + \frac{2}{4!} h^2 (u_x^{(4)} + u_y^{(4)}) + \frac{2}{6!} h^4 (u_x^{(6)} + u_y^{(6)}) + O(h^6) \\ \Delta_h^{(5,\times)} u &= \Delta u + \frac{2}{4!} h^2 (u_x^{(4)} + 6u_{x,y}^{(2,2)} + u_y^{(4)}) + \frac{2}{6!} h^4 (u_x^{(6)} + 15u_{x,y}^{(4,2)} + 15u_{x,y}^{(2,4)} + u_y^{(6)}) + O(h^6).\end{aligned}$$

Let  $\Delta_h^{(9)}$  be the nine-point difference scheme defined by

$$\Delta_h^{(9)} = \frac{2}{3} \Delta_h^{(5)} + \frac{1}{3} \Delta_h^{(5,\times)}.$$

The coefficients in this stencil equation equal  $1/6$  for the corner vertex points in a square with edges  $2h$ , equal  $2/3$  for the midedge points, and equal  $-10/3$  for the center point.

A computation shows that for a uniform rectangular mesh,

$$\Delta_h^{(9)} u_h = f + \frac{1}{12} h^2 \Delta f + \frac{1}{360} h^4 (\Delta^2 f + 2f_{x,y}^{(2,2)}) + O(h^6)$$

where  $\Delta^2 = \Delta(\Delta f)$ . Using a modified right-hand side in the difference formula, it follows that the difference approximation

$$\Delta_h^{(9)} u_h = \left[ I + \frac{h^2}{12} \Delta_h^{(5)} \right] f, \quad (x, y) \in \Omega_h \quad (3.1)$$

has truncation error  $O(h^4)$ . Moreover, it follows from (3.1) that for a sufficiently smooth function  $f$ ,  $\Delta f = \Delta_h^{(9)} f - (1/12) h^2 \Delta^2 f + O(h^4)$ . A computation shows that  $h^2 f_{x,y}^{(2,2)} = 2[\Delta_h^{(5,\times)} f - \Delta_h^{(5)} f] + O(h^4) = 6(\Delta_h^{(9)} - \Delta_h^{(5)}) f + O(h^4)$  and, therefore, the nine-point stencil with the next modified right-hand side,

$$\Delta_h^{(9)} u_h = f + \frac{1}{12} h^2 \Delta_h^{(9)} f + \frac{1}{30} h^2 (\Delta_h^{(9)} - \Delta_h^{(5)}) f - \frac{1}{240} h^4 \Delta_h^{(5)} \Delta f \quad (x, y) \in \Omega_h,$$

has a truncation error  $O(h^6)$ .

The implementation of this scheme is simplified if  $f$  is given analytically so that  $\Delta f$  can be computed explicitly. If  $f \equiv 0$ , then  $\Delta_h^{(9)} u_h \equiv 0$  has an order of approximation  $O(h^6)$ . Hence, this scheme provides a very accurate approximation, for instance, for far-field equations, where frequently  $\Delta u = 0$ . The above is an example of a compact difference scheme; for further references on such schemes, see [7].

### 3.2 Difference methods for orthotropic problems and problems with mixed derivatives

Consider first the anisotropic differential equation

$$au_{xx} + bu_{yy} = f(x, y) \quad (x, y) \in \Omega,$$

where  $u = g(x, y)$ ,  $(x, y) \in \partial\Omega$  and  $f$  and  $g$  are given, sufficiently smooth functions. Let  $a > 0$  and  $b > 0$ . Here, the nine-point difference approximation has a stencil, as shown in (3.2) with  $c = 0$ . If we modify the right-hand side to be  $f + 1/12h^2(af_{xx} + bf_{yy})$ , then it can be seen (as has already been shown above for  $a = b$ ) that in this case the local truncation error becomes  $O(h^4)$ .

Consider next the differential equation with a mixed derivative

$$au_{xx} + 2cu_{xy} + bu_{yy} = f(x, y), \quad (x, y) \in \Omega$$

with given boundary conditions. We assume that  $a > 0$ ,  $b > 0$ , and  $c^2 < ab$ , which are the conditions for ellipticity of the operator. For the mixed derivative, we use the central difference approximation

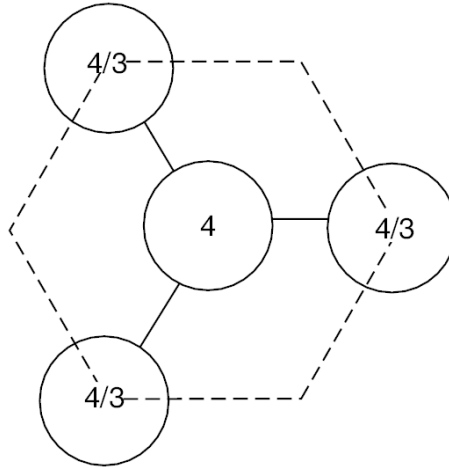
$$u_{xy} \approx \frac{1}{4h^2}[u_h(x-h, y-h) - u_h(x+h, y-h) - u_h(x-h, y+h) + u_h(x+h, y+h)].$$

Combined with the nine-point difference stencil, the stencil becomes

$$\frac{1}{6h^2} \begin{bmatrix} \frac{a+b}{2} - 3c & 5b - a & \frac{a+b}{2} + 3c \\ 5a - b & -10(a+b) & 5a - b \\ \frac{a+b}{2} + 3c & 5b - a & \frac{a+b}{2} - 3c \end{bmatrix}. \quad (3.2)$$

### 3.3 Difference schemes for other regular tessellations

Finite differences can be extended to nonrectangular meshes. For a regular (isosceles) triangular mesh, one can form the obvious seven-point difference stencil. For a hexagonal ('honeycomb') mesh, one finds a four-point stencil:



The symmetrically located nodepoints in the seven-point scheme in the hexagonal mesh allow one to readily approximate second-order cross derivatives.

The corresponding seven-point scheme takes the form

$$\Delta_h^{(7)} u_h(x, y) = \left(\frac{h}{2}\right)^2 \left[ \frac{1}{6} \sum_{i=1}^6 u_h(x_i, y_i) - u(x, y) \right],$$

where  $u(x_i, y_i)$ ,  $i = 1, 2, \dots, 6$  are the hexagonal mesh node points. A Taylor expansion shows that

$$\Delta_h^{(7)} u_h = f + \frac{1}{4} \left( \frac{h}{2} \right)^2 \Delta f + \frac{1}{360} \left( \frac{h}{2} \right)^4 (11u_x^{(6)} + 15u_{x,y}^{(4,2)} + 45u_{x,y}^{(2,4)} + 9u_y^{(6)}) + O(h^6).$$

This scheme corresponds to a horizontal ordering of the hexagonal mesh. Similarly, for a 90 degree reoriented scheme, one gets

$$\Delta_h^{(7,\times)} u_h = f + \frac{1}{4} \left( \frac{h}{2} \right)^2 \Delta f + \frac{1}{360} \left( \frac{h}{2} \right)^4 (9u_x^{(6)} + 45u_{x,y}^{(4,2)} + 15u_{x,y}^{(2,4)} + 11u_y^{(6)}) + O(h^6).$$

One can solve the difference approximations

$$\begin{aligned} \Delta_h^{(7)} u_h^{(1)} &= f + \frac{1}{4} \left( \frac{h}{2} \right)^2 \Delta f + \frac{1}{36} \left( \frac{h}{2} \right)^4 \Delta^2 f, \\ \Delta_h^{(7,\times)} u_h^{(2)} &= f + \frac{1}{4} \left( \frac{h}{2} \right)^2 \Delta f + \frac{1}{36} \left( \frac{h}{2} \right)^4 \Delta^2 f \end{aligned}$$

separately and take the average  $(u_h^{(1)} + u_h^{(2)})/2$  of the solutions or, alternatively, solve

$$\frac{1}{2} (\Delta_h^{(7)} + \Delta_h^{(7,\times)}) u_h = f + \frac{1}{4} \left( \frac{h}{2} \right)^2 \Delta f + \frac{1}{36} \left( \frac{h}{2} \right)^4 \Delta^2 f$$

which in both cases results in an  $O(h^6)$  truncation and discretization error. Here

$$\Delta^3 f = \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right)^3 u = u_x^{(6)} + 3u_{x,y}^{(4,2)} + 3u_{x,y}^{(2,4)} + u_y^{(6)}.$$

Clearly, as before, we can replace  $\Delta f$  with  $(\Delta_h^{(7)} + \Delta_h^{(7,\times)})/2$ , to enable pointwise computations of  $\Delta f$  and also avoiding the need to compute fourth order derivatives of  $f$ .

The difference schemes are of positive type and give monotone matrices. The methods can be used in splitting a given operator in a monotone operator and remainder, if the splitting is convergent.

The advantage with the above methods of high order is that one can use a fairly coarse mesh and still get sufficient accuracy. The disadvantage is that they require a uniform mesh.

The alternative method is to use the extrapolation method. This method is more flexible and can be applied locally even if the mesh is not globally uniform. A slight disadvantage is that it requires a second solution, but this can take place on a coarser mesh and is therefore less expensive.

It is possible to use extrapolation even for the high order difference schemes. This can give extremely high orders of approximation.

Besides rectangular, prism and pyramid meshes there exist also other regular meshes in 3D that lead to monotone difference schemes. However, the application in 3D will not be considered in this paper.

## 4 Error estimates using convergent operator splittings

We now describe two extensions of the estimation of discretization errors for matrices that are not necessarily monotone, but where one can split the operator in a sum of a monotone matrix and a remainder term.

## 4.1 Estimation based on eigenvalue bounds

Assume that the discrete operator has been split as

$$A = M - N,$$

where  $A$  is nonsingular,  $M$  is monotone and  $N$  is the remainder term. Further assumptions will be specified below. Assume further that eigenvalue bounds,

$$0 < \lambda_1 \leq \lambda(M^{-1}A) \leq \lambda_0$$

exist for the preconditioned matrix  $M^{-1}A$ , which do not depend on  $h$ .

In order to get a convergent splitting, we scale  $M$ , i.e. replace  $M$  with  $\frac{1}{\beta}M$ , for some scalar  $\beta > 0$ . Then the corresponding splitting becomes

$$A = \frac{1}{\beta}M + (A - \frac{1}{\beta}M).$$

Here the eigenvalues of  $(\frac{1}{\beta}M)^{-1}(A - \frac{1}{\beta}M) = \beta M^{-1}A - I$  are contained in the interval  $[\beta\lambda_1 - 1, \beta\lambda_0 - 1]$ . To get a convergent splitting we must choose  $\beta$  such that  $0 < \beta < \frac{2}{\lambda_0}$ . The optimal value of  $\beta$  equals

$$-(\beta\lambda_1 - 1) = \beta\lambda_0 - 1, \quad \text{i.e. } \beta = \frac{2}{\lambda_0 + \lambda_1},$$

which results in the spectral bound

$$c := \rho \left( \left( \frac{1}{\beta}M \right)^{-1} \left( A - \frac{1}{\beta}M \right) \right) = \frac{\lambda_0 - \lambda_1}{\lambda_0 + \lambda_1} = \frac{1 - \kappa^{-1}}{1 + \kappa^{-1}},$$

where  $\kappa = \lambda_0/\lambda_1$  is the condition number of the corresponding preconditioned matrix  $M^{-1}A$ .

Let  $\tilde{M} = \frac{1}{\beta}M$ ,  $\tilde{N} = \tilde{M} - A$ . The splitting can be applied to estimate the discretization error. From (2.2) we have

$$\mathcal{L}_h(u - u_h) = \tau_h \tag{4.1}$$

(where  $\tau_h$  is the truncation error), i.e.

$$(\tilde{M} - \tilde{N})(u - u_h) = \tau_h$$

or

$$(I - \tilde{M}^{-1}\tilde{N})(u - u_h) = \tilde{M}^{-1}\tau_h = \beta M^{-1}\tau_h.$$

Hence

$$\|u - u_h\|_p \leq \frac{\beta}{1 - \|\tilde{M}^{-1}\tilde{N}\|_p} \|M^{-1}\tau_h\|_p,$$

where we let  $p = 2$  or  $p = \infty$ . If  $M$  is normal, or even symmetric, for  $p = 2$  we have  $\|\tilde{M}^{-1}\tilde{N}\|_2 = \rho(\tilde{M}^{-1}\tilde{N}) = c$ , and  $\|M^{-1}\|_2 \leq \|M^{-1}\|_\infty$ . (Note the familiar inequality  $\|M\|_2^2 \leq \|M^T\|_\infty \|M\|_\infty$ , so  $\|M\|_2 \leq \|M\|_\infty$  for symmetric matrices.)

Since  $1 - c = \frac{2\lambda_1}{\lambda_0 + \lambda_1}$  and  $\frac{\beta}{1 - c} = \frac{1}{\lambda_1}$ , it follows that

$$\|u - u_h\|_2 \leq \frac{1}{\lambda_1} \|M^{-1}\|_2 \|\tau_h\|_2.$$

If we approximate  $\|\tilde{M}^{-1}\tilde{N}\|_\infty$  with  $\|\tilde{M}^{-1}\tilde{N}\|_2$ , we get

$$\|u - u_h\|_\infty \lesssim \frac{1}{\lambda_1} \|M^{-1}\|_\infty \|\tau_h\|_\infty. \tag{4.2}$$

## 4.2 Estimation based on a monotone part of the matrix

Let us consider again the splitting

$$A = M - N,$$

where  $M$  is monotone. One can readily reduce the error estimation with  $A^{-1}$  to the one with  $M^{-1}$  if  $N$  is properly dominated by the other matrices. First we describe this for general matrices, then in the next subsection we give analogous estimates on operator level on an example.

### 4.2.1 The case of general matrices

First, assume that  $A$  dominates  $N$  in the sense that

$$\exists \beta > 0 : \quad \|Ax\|_\infty \geq \beta \|Nx\|_\infty \quad (\forall x \in \mathbf{R}^n). \quad (4.3)$$

**Proposition 4.1.** *If assumption (4.3) holds, then*

$$\|A^{-1}\|_\infty \leq \left(1 + \frac{1}{\beta}\right) \|M^{-1}\|_\infty. \quad (4.4)$$

PROOF. Since

$$\|Mx\|_\infty \leq \|Ax\|_\infty + \|Nx\|_\infty \leq \left(1 + \frac{1}{\beta}\right) \|Ax\|_\infty, \quad (4.5)$$

we have

$$\|MA^{-1}\|_\infty = \max_{\substack{y \in \mathbf{R}^n \\ y \neq 0}} \frac{\|MA^{-1}y\|_\infty}{\|y\|_\infty} = \max_{\substack{x \in \mathbf{R}^n \\ x \neq 0}} \frac{\|Mx\|_\infty}{\|Ax\|_\infty} \leq 1 + \frac{1}{\beta} \quad (4.6)$$

and thus

$$\|A^{-1}\|_\infty \leq \|M^{-1}MA^{-1}\|_\infty \leq \left(1 + \frac{1}{\beta}\right) \|M^{-1}\|_\infty. \quad \blacksquare$$

Second, we can reduce the dominance requirement to the splitted parts  $M$  and  $N$ , but then a larger constant is needed. Assume that

$$\exists \alpha > 1 : \quad \|Mx\|_\infty \geq \alpha \|Nx\|_\infty \quad (\forall x \in \mathbf{R}^n). \quad (4.7)$$

**Proposition 4.2.** *If assumption (4.7) holds, then*

$$\|A^{-1}\|_\infty \leq \left(\frac{\alpha}{\alpha - 1}\right) \|M^{-1}\|_\infty. \quad (4.8)$$

PROOF. Now

$$\alpha \|Nx\|_\infty \leq \|Mx\|_\infty \leq \|Ax\|_\infty + \|Nx\|_\infty,$$

hence

$$\|Ax\|_\infty \geq (\alpha - 1) \|Nx\|_\infty$$

for all  $x \in \mathbf{R}^n$ . Then we can apply Proposition 4.1 with  $\beta := \alpha - 1$ , in which case

$$1 + \frac{1}{\beta} = 1 + \frac{1}{\alpha - 1} = \frac{\alpha}{\alpha - 1}$$

and we thus obtain the desired statement.  $\blacksquare$

We note that here the analogous upper estimate yields in the same way the counterpart of (4.5), namely

$$\frac{\|Ax\|_\infty}{\|Mx\|_\infty} \leq 1 + \frac{1}{\alpha}$$



i.e. altogether  $A$  and  $M$  are equivalent operators in the sense of [5]. However, in below we only need the lower estimate.

Applying Proposition 4.1, we obtain that if

$$Ae = r$$

for some vectors  $e, r \in \mathbf{R}^n$ , then

$$\|e\|_\infty \leq \left(1 + \frac{1}{\beta}\right) \|M^{-1}\|_\infty \|r\|_\infty.$$

In particular, we can consider the case of an error equation (4.1) where we split the discretization matrix  $\mathcal{L}_h$  as  $M_h - N_h$  and assume that estimate (4.3) holds uniformly w.r.t. the discretization parameter  $h$ . That is, let

$$\exists \beta > 0 : \quad \|A_h x\|_\infty \geq \beta \|N_h x\|_\infty \quad (\forall h \leq h_0, x \in \mathbf{R}^n). \quad (4.9)$$

**Corollary 4.1.** *Let us split the discretization matrix as  $\mathcal{L}_h = M_h - N_h$  and let assumption (4.9) hold. Then the error equation (4.1) satisfies*

$$\|u - u_h\|_\infty \leq \left(1 + \frac{1}{\beta}\right) \|M_h^{-1}\|_\infty \|\tau_h\|_\infty.$$

#### 4.2.2 Estimates on operator level for Helmholtz problems

Let us consider the following Helmholtz equation on a bounded domain  $\Omega \subset \mathbf{R}^d$  (where  $d = 2$  or  $3$ ):

$$\begin{cases} Lu := -\Delta u - \kappa^2 u = f \\ u|_{\partial\Omega} = 0, \end{cases} \quad (4.10)$$

and assume that  $\kappa^2$  is not an eigenvalue of  $-\Delta$ . Then the Fredholm alternative theorem implies that for all  $f \in L^2(\Omega)$  problem (4.10) has a unique weak solution  $u \in H_0^1(\Omega)$ , see, e.g., [12, Ch. 5.27]. We assume that  $\Omega$  is  $C^2$ -convex, i.e. it is  $C^2$ -diffeomorphic to a convex domain.

We consider the operator  $L$  such that its domain of definition  $D(L)$  includes the given boundary conditions, i.e.  $u \in D(L)$  implies  $u|_{\partial\Omega} = 0$ . We decompose  $L$  as

$$L = M - N, \quad \text{where} \quad Mu := -\Delta u, \quad Nu := \kappa^2 u.$$

Then  $M$  is monotone in the sense that if  $Mu \geq 0$  then  $u \geq 0$ .

Our goal is to derive an analogue of (4.3) in the sense that

$$\exists \beta > 0 : \quad \|Lu\|_\infty \geq \beta \|Nu\|_\infty \quad (\forall u \in C^2(\overline{\Omega}), u|_{\partial\Omega} = 0), \quad (4.11)$$

where  $\|\cdot\|_\infty$  is the maximum norm on  $\overline{\Omega}$ . First we need a  $H^2$ -regularity result. We introduce the space  $H^2(\Omega) \cap H_0^1(\Omega) = \{u \in H^2(\Omega) : u|_{\partial\Omega} = 0\}$ .

**Proposition 4.3.** *There exists  $c_1 > 0$  such that for all  $u \in H^2(\Omega) \cap H_0^1(\Omega)$*

$$\|Lu\|_{L^2} \geq c_1 \|u\|_{H^2}.$$

PROOF. Let us consider the Hilbert spaces  $H := H^2(\Omega) \cap H_0^1(\Omega)$ ,  $K := L^2(\Omega)$  and the operator  $L$  from  $H$  to  $K$ . Then, clearly, there exists  $K_1 > 0$  such that

$$\|Lu\|_{L^2} \leq K_1 \|u\|_{H^2} \quad (\forall u \in H)$$

since the norm  $\|u\|_{H^2}$  involves all up to second derivatives, that is,  $L : H \rightarrow K$  is a bounded linear operator. We verify that  $L$  is bijective, i.e. that for all  $f \in L^2(\Omega)$  there exists a unique  $u \in H^2(\Omega) \cap H_0^1(\Omega)$  such that  $Lu = f$ , in other words, that (4.10) has a unique solution in  $H^2(\Omega) \cap H_0^1(\Omega)$ . Let  $f \in L^2(\Omega)$ . We have seen that there is a unique weak solution  $u \in H_0^1(\Omega)$ , hence we only have to prove that  $u \in H^2(\Omega) \cap H_0^1(\Omega)$ . Let  $\hat{f} := f + ku$ , which is in  $L^2(\Omega)$ , then  $u$  is the solution of the problem

$$-\Delta u = \hat{f}, \quad u|_{\partial\Omega} = 0.$$

Since  $\Omega$  is  $C^2$ -convex, we have  $u \in H^2(\Omega) \cap H_0^1(\Omega)$  (see [8]). Altogether,  $L : H \rightarrow K$  is a bounded linear bijection. Then Banach's homeomorphism theorem asserts that  $L$  has a bounded inverse, which just means that the desired estimate holds.  $\blacksquare$

**Proposition 4.4.** *There exists  $c_2 > 0$  such that for all  $u \in C^2(\overline{\Omega})$ ,  $u|_{\partial\Omega} = 0$  we have*

$$\|Lu\|_{\infty} \geq c_2 \|u\|_{\infty}.$$

PROOF. Since  $d = 2$  or  $3$ , we have the bounded inclusion  $H^2(\Omega) \subset C(\overline{\Omega})$  (see [1]), i.e. there exists  $k_1 > 0$  such that

$$\|u\|_{\infty} \leq k_1 \|u\|_{H^2} \quad (\forall u \in H^2(\Omega)).$$

This also holds for the considered case  $u \in C^2(\overline{\Omega})$ ,  $u|_{\partial\Omega} = 0$ . On the other hand, the  $L^\infty$ -norm is stronger than the  $L^2$ -norm, i.e. there exists  $k_2 > 0$  such that

$$\|v\|_{L^2} \leq k_2 \|v\|_{\infty} \quad (\forall v \in L^\infty(\Omega)),$$

and this also holds for  $v := Lu \in C(\overline{\Omega})$ . Altogether, combining these with Proposition 4.3, we have

$$\|u\|_{\infty} \leq k_1 \|u\|_{H^2} \leq \frac{k_1}{c_1} \|Lu\|_{L^2} \leq \frac{k_1 k_2}{c_1} \|Lu\|_{\infty},$$

i.e. the desired statement holds with  $c_2 = \frac{c_1}{k_1 k_2}$ .  $\blacksquare$

It follows readily that (4.11) holds with  $\beta = \frac{c_2}{\kappa^2}$ :

**Corollary 4.2.** *For all  $u \in C^2(\overline{\Omega})$ ,  $u|_{\partial\Omega} = 0$  we have*

$$\|Lu\|_{\infty} \geq \frac{c_2}{\kappa^2} \|Nu\|_{\infty}.$$

Then we can follow (4.5) to derive

$$\|-\Delta u\|_{\infty} \leq \|Lu\|_{\infty} + \|Nu\|_{\infty} \leq \left(1 + \frac{\kappa^2}{c_2}\right) \|Lu\|_{\infty}.$$

To sum up, we have

**Corollary 4.3.** *For all  $u \in C^2(\overline{\Omega})$ ,  $u|_{\partial\Omega} = 0$  we have*

$$\frac{\|-\Delta u\|_{\infty}}{\|Lu\|_{\infty}} \leq 1 + \frac{\kappa^2}{c_2},$$

i.e. for  $v := Lu$  we have

$$\frac{\|-\Delta L^{-1}v\|_{\infty}}{\|v\|_{\infty}} \leq 1 + \frac{\kappa^2}{c_2}.$$

**Remark 4.1.** (i) Altogether, we have obtained an analogue of (4.6). Since the set of above functions  $v$  is dense, we have the operator norm estimate

$$\| -\Delta L^{-1} \|_{\infty} \leq 1 + \frac{\kappa^2}{c_2}.$$

(ii) For finite difference discretizations, under proper further regularity we have

$$L_h u_h = Lu + O(h^2), \quad -\Delta_h u_h = -\Delta u + O(h^2)$$

in the node points, hence for sufficiently small  $h$  we have

$$\frac{\| -\Delta_h u_h \|_{\infty}}{\| L_h u_h \|_{\infty}} = \frac{\| -\Delta u \|_{\infty} + O(h^2)}{\| Lu \|_{\infty} + O(h^2)} \approx \frac{\| -\Delta u \|_{\infty}}{\| Lu \|_{\infty}} \leq 1 + \frac{\kappa^2}{c_2}.$$

We are thus approximately in the situation of (4.9). In particular, following Corollary 4.1, the discretization error and truncation error satisfy

$$\| u - u_h \|_{\infty} \leq C \| -\Delta_h^{-1} \|_{\infty} \| \tau_h \|_{\infty} \quad (4.12)$$

where  $C \approx 1 + \frac{\kappa^2}{c_2}$ . A more precise formulation would involve higher regularity related to the  $O(h^2)$  estimate.

### 4.2.3 Numerical illustration

In this paper we have considered several examples of monotone difference schemes, also of higher order. However, to illustrate the main idea of the paper, namely to extend the estimates to convergent splittings of matrices, it suffices to use a standard second order difference scheme.

Let us then consider the Helmholtz equation (4.10) on the unit square  $\Omega := [0, 1]^2$ . We apply the finite difference method using the standard five-point discretization.

We let  $\kappa^2 = 32$ , which is not an eigenvalue of  $-\Delta$ . The exact solution is chosen as  $u(x, y) = \sin \pi x \cdot \sin \pi y$ . Then the right hand side is  $f(x, y) = \gamma \sin \pi x \cdot \sin \pi y$ , where  $\gamma = 2\pi^2 - 32 \approx -12.28$ .

Our goal is to illustrate (4.12). For this sake we have run the finite difference method for the above problem on different grids, and we have calculated the  $\infty$ -norms from (4.12). We denote  $C_h := \| u - u_h \|_{\infty} / \| -\Delta_h^{-1} \|_{\infty} \| \tau_h \|_{\infty}$ , i.e. we have the equality

$$\| u - u_h \|_{\infty} = C_h \| -\Delta_h^{-1} \|_{\infty} \| \tau_h \|_{\infty}$$

corresponding to (4.12). We wish to demonstrate that  $C_h \leq C$  for some constant  $C > 0$  independently of  $h$ , further, we are interested in the approximate value of  $C$ .

The results are as follows. (The  $n \times n$  grids are understood to consist of interior points, due to the homogeneous boundary condition, the letter  $h$  stands for the mesh width  $h := 1/(n + 1)$ .)

grid	$\  u - u_h \ _{\infty}$	$\  -\Delta_h^{-1} \ _{\infty}$	$\  \tau_h \ _{\infty}$	$C_h$
$3 \times 3$	0.0750	0.0703	0.9940	1.0733
$10 \times 10$	0.0106	0.0722	0.1311	1.1198
$30 \times 30$	0.0014	0.0735	0.0168	1.1337
$100 \times 100$	1.29e-04	0.0737	0.0016	1.1024

Table 1: Error bounds for the discretized Helmholtz equation.

The values of  $C_h$  indicate as expected that they are bounded independently of the grid. Altogether, we have illustrated that the operator  $-\Delta$ , when being the monotone part of the Helmholtz operator, is able to produce a similar bound for the discretization error as in classical problems where the original matrix was monotone itself.

## References

- [1] ADAMS, R.A., *Sobolev Spaces*, Academic Press, New York-London, 1975.
- [2] AXELSSON, O., *Iterative Solution Methods*, Cambridge University Press, 1994.
- [3] AXELSSON, O., *Finite Difference Methods*, in: Encyclopedia of Computational Mechanics, Vol. 1, eds. E. Stein, R. de Borst and T. J. R. Hughes, John Wiley & Sons, Ltd., Chichester, 2004; pp. 7–53.
- [4] AXELSSON, O., KARÁTON, J., Equivalent operator preconditioning for elliptic problems, *Numer. Algor.*, 50:297–380, 2009.
- [5] FABER, V., MANTEUFFEL, T., PARTER, S.V., On the theory of equivalent operators and application to the numerical solution of uniformly elliptic partial differential equations, *Adv. in Appl. Math.*, 11 (1990), 109-163.
- [6] FIEDLER, M., *Special Matrices and Their Applications in Numerical Mathematics*: Second Edition, Dover Books on Mathematics, 2008.
- [7] HOUSTIS, E. N.; RICE, J. R., High order methods for elliptic partial differential equations with singularities, *Internat. J. Numer. Methods Engrg.* 18 (1982), no. 5, 737-754.
- [8] KADLEC, J., On the regularity of the solution of the Poisson problem on a domain with boundary locally similar to the boundary of a convex open set, *Czech. Math. J.*, 14 (89), (1964), pp. 386-393.
- [9] LEVEQUE, R. J., *Finite Difference Methods for Ordinary and Partial Differential Equations. Steady-state and Time-dependent Problems*, SIAM, Philadelphia, 2007.
- [10] ORTEGA, J.M., RHEINOLDT, W.C., *Iterative Solutions for Nonlinear Equations in Several Variables*, Academic Press, 1970.
- [11] STRIKWERDA, J. C. *Finite Difference Schemes and Partial Differential Equations*, SIAM, Philadelphia, 2004.
- [12] VLADIMIROV, V. S., *Equations of Mathematical Physics*, Marcel Dekker, 1971.