$b(\boldsymbol{x}^*, \boldsymbol{x}_\gamma) > 0$, i.e., the matrix block A_2 in (2.9) is not the zero matrix. Then, the grid function $u(\boldsymbol{x})$ is non-positive (or non-negative, respectively) for all $\boldsymbol{x} \in \omega_h \cup \gamma_h$.

Proof. Let $L_h u(\boldsymbol{x}) \le 0$ on ω_h. Assume that there is a node $\overline{\boldsymbol{x}} \in \omega_h$ with $u(\overline{\boldsymbol{x}}) > 0$. Then, the grid function has either a positive maximum on ω_h and it is not constant, which is a contradiction to the DMP for the inner nodes, Lemma 2.19, or $u(\boldsymbol{x})$ has to be constant, i.e., $u(\boldsymbol{x}) = u(\overline{\boldsymbol{x}}) > 0$ for all $\boldsymbol{x} \in \omega_h$. For the second case, consider the boundary-connected inner node $\boldsymbol{x}^* \in \omega_h^*$. Using the same calculations as in (2.14) and taking into account that the values of u at the boundary are non-positive, one obtains

$$\begin{aligned} L_h u(\boldsymbol{x}^*) &= \underbrace{d(\boldsymbol{x}^*)}_{\ge 0}\underbrace{u(\boldsymbol{x}^*)}_{>0} - \sum_{\boldsymbol{y}\in S(\boldsymbol{x}^*), \boldsymbol{y}\notin\gamma_h} \underbrace{b(\boldsymbol{x}^*, \boldsymbol{y})}_{>0}\underbrace{(u(\boldsymbol{y}) - u(\boldsymbol{x}^*))}_{=0} \\ &\quad - \sum_{\boldsymbol{y}\in S(\boldsymbol{x}^*), \boldsymbol{y}\in\gamma_h} \underbrace{b(\boldsymbol{x}^*, \boldsymbol{y})}_{>0}\underbrace{(u(\boldsymbol{y}) - u(\boldsymbol{x}^*))}_{<0} > 0. \end{aligned} \tag{2.16}$$

In the last sum, there is at least one term since $\boldsymbol{x}_\gamma \in S(\boldsymbol{x}^*)$. Altogether, (2.16) is a contradiction to the assumption on L_h. ■

Corollary 2.22. Unique solution of the discrete Laplace equation with homogeneous right-hand side and homogeneous Dirichlet boundary conditions. *Under the assumptions of Corollary 2.21, the discrete Laplace equation $L_h u(\boldsymbol{x}) = 0$ for $\boldsymbol{x} \in \omega_h$ and $u(\boldsymbol{x}) = 0$ for $\boldsymbol{x} \in \gamma_h$ possesses only the trivial solution $u(\boldsymbol{x}) = 0$.*

Proof. The statement of the corollary follows by applying Corollary 2.21 both for $L_h u(\boldsymbol{x}) \le 0$ and $L_h u(\boldsymbol{x}) \ge 0$. ■

Theorem 2.23. Existence and uniqueness of a solution of the finite difference equation (2.6)**.** *Under the assumptions of Corollary 2.22, the finite difference equation* (2.6) *possesses a unique solution.*

Proof. Corollary 2.22 shows that the homogeneous linear system of equations (2.9) has a unique solution. Hence, the system matrix is invertible and it follows that (2.9) is uniquely solvable for all right-hand sides, where (2.9) is just the matrix-vector representation of (2.6). ■

Corollary 2.24. Comparison lemma. *Let the assumptions of Corollary 2.21 be satisfied and let*

$$\begin{aligned} L_h u(\boldsymbol{x}) &= f(\boldsymbol{x}) \text{ for } \boldsymbol{x} \in \omega_h; \quad & u(\boldsymbol{x}) &= g(\boldsymbol{x}) \text{ for } \boldsymbol{x} \in \gamma_h, \\ L_h \overline{u}(\boldsymbol{x}) &= \overline{f}(\boldsymbol{x}) \text{ for } \boldsymbol{x} \in \omega_h; \quad & \overline{u}(\boldsymbol{x}) &= \overline{g}(\boldsymbol{x}) \text{ for } \boldsymbol{x} \in \gamma_h, \end{aligned}$$

with $|f(\boldsymbol{x})| \le \overline{f}(\boldsymbol{x})$, $\boldsymbol{x} \in \omega_h$, and $|g(\boldsymbol{x})| \le \overline{g}(\boldsymbol{x})$, $\boldsymbol{x} \in \gamma_h$. Then, it is $|u(\boldsymbol{x})| \le \overline{u}(\boldsymbol{x})$ for all $\boldsymbol{x} \in \omega_h \cup \gamma_h$. The function $\overline{u}(\boldsymbol{x})$ is called majorizing function.

Proof. Exercise. ■

Remark 2.25. Remainder of this section. The remaining corollaries presented in this section will be applied in the stability proof in Section 2.4. In this proof, the homogeneous problem (right-hand side vanishes) and the problem with homogeneous Dirichlet boundary conditions will be analyzed separately. □

Corollary 2.26. Homogeneous problem. *For the solution of the problem*

$$\begin{aligned} L_h u(\boldsymbol{x}) &= 0, \quad && \boldsymbol{x} \in \omega_h, \\ u(\boldsymbol{x}) &= g(\boldsymbol{x}), && \boldsymbol{x} \in \gamma_h, \end{aligned}$$

with $d(\boldsymbol{x}) = 0$ *for all* $\boldsymbol{x} \in \omega_h^\circ$*, it holds that*

$$\|u\|_{l^\infty(\omega_h \cup \gamma_h)} \le \|g\|_{l^\infty(\gamma_h)} .$$

Proof. Consider the problem

$$\begin{aligned} L_h \overline{u}(\boldsymbol{x}) &= 0, && \boldsymbol{x} \in \omega_h, \\ \overline{u}(\boldsymbol{x}) &= \overline{g}(\boldsymbol{x}) = const = \|g\|_{l^\infty(\gamma_h)}, && \boldsymbol{x} \in \gamma_h. \end{aligned}$$

By Example 2.18, it is known that the row sums for all $\boldsymbol{x} \in \omega_h$ vanish. Hence, $\overline{u}(\boldsymbol{x}) = \|g\|_{l^\infty(\gamma_h)} = const$ is a solution of this problem.[1] By Corollary 2.22, this solution is unique.

Now, the application of Corollary 2.24 gives $\overline{u}(\boldsymbol{x}) \ge |u(\boldsymbol{x})|$ for all $\boldsymbol{x} \in \omega_h \cup \gamma_h$, so that

$$\|u\|_{l^\infty(\omega_h \cup \gamma_h)} \le \overline{u}(\boldsymbol{x}) = \|g\|_{l^\infty(\gamma_h)} ,$$

which is the statement of the corollary. ■

Corollary 2.27. Problem with homogeneous boundary condition and inhomogeneous right-hand side close to the boundary. *Consider*

$$\begin{aligned} L_h u(\boldsymbol{x}) &= f(\boldsymbol{x}), && \boldsymbol{x} \in \omega_h, \\ u(\boldsymbol{x}) &= 0, && \boldsymbol{x} \in \gamma_h, \end{aligned}$$

with $f(\boldsymbol{x}) = 0$ *for all* $\boldsymbol{x} \in \omega_h^\circ$. *Define*

$$\tilde{d}(\boldsymbol{x}) = a(\boldsymbol{x}) - \sum_{\boldsymbol{y} \in S(\boldsymbol{x}), \boldsymbol{y} \notin \gamma_h} b(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{x}) + \sum_{\boldsymbol{y} \in S(\boldsymbol{x}), \boldsymbol{y} \in \gamma_h} b(\boldsymbol{x}, \boldsymbol{y}) \quad \boldsymbol{x} \in \omega_h.$$

With respect to the finite difference scheme, it will be assumed that $\tilde{d}(\boldsymbol{x}) = 0$ *for all* $\boldsymbol{x} \in \omega_h^\circ$*, and* $\tilde{d}(\boldsymbol{x}) > 0$ *for all* $\boldsymbol{x} \in \omega_h^*$*. Then, the following estimate is valid*

$$\|u\|_{l^\infty(\omega_h \cup \gamma_h)} \le \left\| D^+ f \right\|_{l^\infty(\omega_h)}$$

with $D^+ = diag(0, \tilde{d}(\boldsymbol{x})^{-1})$*. The zero entries appear for* $\boldsymbol{x} \in \omega_h^\circ$ *and the entries* $\tilde{d}(\boldsymbol{x})^{-1}$ *for* $\boldsymbol{x} \in \omega_h^*$.

Proof. Let $\overline{f}(\boldsymbol{x}) = |f(\boldsymbol{x})|$, $\boldsymbol{x} \in \omega_h$, and $\overline{g}(\boldsymbol{x}) = 0, \boldsymbol{x} \in \gamma_h$. The corresponding solution $\overline{u}(\boldsymbol{x})$ is non-negative, $\overline{u}(\boldsymbol{x}) \ge 0$ for all $\boldsymbol{x} \in \omega_h \cup \gamma_h$, see the DMP for the boundary value

---

[1] The corresponding continuous problem is $-\Delta u = 0$ in Ω, $u = const = \|g\|_{l^\infty(\gamma_h)}$ on $\partial\Omega$. It is clear that $u = \|g\|_{l^\infty(\gamma_h)}$ is the solution of this problem. It is shown that the discrete analog holds, too.

problem, Corollary 2.21. Define $\overline{\boldsymbol{x}}$ by

$$\overline{u}(\overline{\boldsymbol{x}}) = \|\overline{u}\|_{l^\infty(\omega_h\cup\gamma_h)} .$$

One can choose $\overline{\boldsymbol{x}} \in \omega_h^*$, because if $\overline{\boldsymbol{x}} \in \omega_h^\circ$, then it holds that

$$\underbrace{d(\overline{\boldsymbol{x}})}_{=0}\,\overline{u}(\overline{\boldsymbol{x}}) - \sum_{\boldsymbol{y}\in S(\overline{\boldsymbol{x}})} \underbrace{b(\overline{\boldsymbol{x}},\boldsymbol{y})}_{>0}\underbrace{\big(\overline{u}(\boldsymbol{y}) - \overline{u}(\overline{\boldsymbol{x}})\big)}_{\le 0} = \overline{f}(\overline{\boldsymbol{x}}) = 0,$$

i.e., $\overline{u}(\overline{\boldsymbol{x}}) = \overline{u}(\boldsymbol{y})$ for all $\boldsymbol{y} \in S(\overline{\boldsymbol{x}})$. Let $\hat{\boldsymbol{x}} \in \omega_h^*$ and $\overline{\boldsymbol{x}}, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_m, \hat{\boldsymbol{x}}$ be a connection with $\boldsymbol{x}_i \notin \omega_h^*$, $i = 1, \ldots, m$. For $\boldsymbol{x}_m$, it holds analogously that

$$\overline{u}(\boldsymbol{x}_m) = \|\overline{u}\|_{l^\infty(\omega_h\cup\gamma_h)} = \overline{u}(\boldsymbol{y}) \ \forall\ \boldsymbol{y} \in S(\boldsymbol{x}_m).$$

Hence, it follows in particular that $\overline{u}(\hat{\boldsymbol{x}}) = \|\overline{u}\|_{l^\infty(\omega_h\cup\gamma_h)}$ so that one can choose $\overline{\boldsymbol{x}} = \hat{\boldsymbol{x}}$.

Using the definition of $\tilde{d}(\hat{\boldsymbol{x}})$ and the homogeneous values at the boundary yields

$$
\begin{aligned}
& d(\hat{\boldsymbol{x}})\overline{u}(\hat{\boldsymbol{x}}) - \sum_{\boldsymbol{y}\in S(\hat{\boldsymbol{x}})} b(\hat{\boldsymbol{x}},\boldsymbol{y})\big(\overline{u}(\boldsymbol{y}) - \overline{u}(\hat{\boldsymbol{x}})\big) = \overline{f}(\hat{\boldsymbol{x}}) \quad \Longleftrightarrow \\
& d(\hat{\boldsymbol{x}})\overline{u}(\hat{\boldsymbol{x}}) + \sum_{\boldsymbol{y}\in S(\hat{\boldsymbol{x}}),\boldsymbol{y}\in\gamma_h} b(\hat{\boldsymbol{x}},\boldsymbol{y})\overline{u}(\hat{\boldsymbol{x}}) \\
& - \sum_{\boldsymbol{y}\in S(\hat{\boldsymbol{x}}),\boldsymbol{y}\notin\gamma_h} b(\hat{\boldsymbol{x}},\boldsymbol{y})\big(\overline{u}(\boldsymbol{y}) - \overline{u}(\hat{\boldsymbol{x}})\big) - \sum_{\boldsymbol{y}\in S(\hat{\boldsymbol{x}}),\boldsymbol{y}\in\gamma_h} b(\hat{\boldsymbol{x}},\boldsymbol{y})\overline{u}(\hat{\boldsymbol{x}}) = \overline{f}(\hat{\boldsymbol{x}}) \quad \Longleftrightarrow \\
& \underbrace{\tilde{d}(\hat{\boldsymbol{x}})}_{>0}\ \underbrace{\overline{u}(\hat{\boldsymbol{x}})}_{=\|\overline{u}\|_{l^\infty(\omega_h\cup\gamma_h)}} - \sum_{\boldsymbol{y}\in S(\hat{\boldsymbol{x}}),\boldsymbol{y}\notin\gamma_h} \underbrace{b(\hat{\boldsymbol{x}},\boldsymbol{y})}_{>0}\underbrace{\big(\overline{u}(\boldsymbol{y}) - \overline{u}(\hat{\boldsymbol{x}})\big)}_{\le 0} = \overline{f}(\hat{\boldsymbol{x}}).
\end{aligned}
$$

It follows, using also Corollary 2.24, that

$$\|u\|_{l^\infty(\omega_h\cup\gamma_h)} \le \|\overline{u}\|_{l^\infty(\omega_h\cup\gamma_h)} \le \frac{\overline{f}(\hat{\boldsymbol{x}})}{\tilde{d}(\hat{\boldsymbol{x}})} \le \max_{\boldsymbol{x}\in\omega_h^*} \frac{\overline{f}(\boldsymbol{x})}{\tilde{d}(\boldsymbol{x})} \le \left\|D^+ f\right\|_{l^\infty(\omega_h)} .$$

■

2.4 Stability and Convergence of the Finite Difference Approximation of the Poisson Problem with Dirichlet Boundary Conditions

Remark 2.28. Decomposition of the solution. A short form to write (2.6) with $\phi(\boldsymbol{x}) = f(\boldsymbol{x})$ is

$$L_h u(\boldsymbol{x}) = f(\boldsymbol{x}),\ \boldsymbol{x} \in \omega_h, \quad u(\boldsymbol{x}) = g(\boldsymbol{x}),\ \boldsymbol{x} \in \gamma_h.$$

The solution of (2.6) can be decomposed into

$$u(\boldsymbol{x}) = u_1(\boldsymbol{x}) + u_2(\boldsymbol{x}),$$

with

$$L_h u_1(\boldsymbol{x}) = f(\boldsymbol{x}),\ \boldsymbol{x} \in \omega_h, \quad u_1(\boldsymbol{x}) = 0,\ \boldsymbol{x} \in \gamma_h \text{ (homogeneous boundary cond.)},$$
$$L_h u_2(\boldsymbol{x}) = 0,\ \boldsymbol{x} \in \omega_h, \quad u_2(\boldsymbol{x}) = g(\boldsymbol{x}),\ \boldsymbol{x} \in \gamma_h \text{ (homogeneous right-hand side)}.$$

□

Stability with Respect to the Boundary Condition

Remark 2.29. Stability with respect to the boundary condition. From Corollary 2.26, it follows that

$$\|u_2\|_{l^\infty(\omega_h)} \le \|g\|_{l^\infty(\gamma_h)}\,. \tag{2.17}$$

□

Stability with Respect to the Right-Hand Side

Remark 2.30. Decomposition of the right-hand side. The right-hand side will be decomposed into

$$f(\boldsymbol{x}) = f^\circ(\boldsymbol{x}) + f^*(\boldsymbol{x})$$

with

$$f^\circ(\boldsymbol{x}) = \begin{cases} f(\boldsymbol{x}), & \boldsymbol{x} \in \omega_h^\circ, \\ 0, & \boldsymbol{x} \in \omega_h^*, \end{cases} \qquad f^*(\boldsymbol{x}) = f(\boldsymbol{x}) - f^\circ(\boldsymbol{x}).$$

Since the considered finite difference scheme is linear, also the function $u_1(\boldsymbol{x})$ can be decomposed into

$$u_1(\boldsymbol{x}) = u_1^\circ(\boldsymbol{x}) + u_1^*(\boldsymbol{x})$$

with

$$L_h u_1^\circ(\boldsymbol{x}) = f^\circ(\boldsymbol{x}),\ \boldsymbol{x} \in \omega_h, \quad u_1^\circ(\boldsymbol{x}) = 0,\ \boldsymbol{x} \in \gamma_h,$$
$$L_h u_1^*(\boldsymbol{x}) = f^*(\boldsymbol{x}),\ \boldsymbol{x} \in \omega_h, \quad u_1^*(\boldsymbol{x}) = 0,\ \boldsymbol{x} \in \gamma_h.$$

□

Remark 2.31. Estimate for the inner nodes. Let $B((0,0),R)$ be a circle with center $(0,0)$ and radius R, which is chosen so that $R \ge \|\boldsymbol{x}\|_2$ for all $\boldsymbol{x} \in \Omega$. Consider the function

$$\overline{u}(\boldsymbol{x}) = \alpha\left(R^2 - x^2 - y^2\right) \quad \text{with } \alpha > 0,$$

that takes for $(x,y) \in \Omega$ only positive values. Applying the definition of the five point stencil, it follows that

$$\Lambda \overline{u}(\boldsymbol{x}) = -\alpha \Lambda (x^2 + y^2 - R^2)$$

$$
\begin{aligned}
&= -\alpha \left(\frac{(x+h_x)^2 - 2x^2 + (x-h_x)^2}{h_x^2} + \frac{(y+h_y)^2 - 2y^2 + (y-h_y)^2}{h_y^2} \right) \\
&= -4\alpha =: -\overline{f}(\boldsymbol{x}), \ \boldsymbol{x} \in \omega_h^\circ,
\end{aligned}
$$

and

$$
\begin{aligned}
\Lambda^* \overline{u}(\boldsymbol{x}) &= -\alpha \left[\frac{1}{\overline{h}_x} \left(\frac{(x+h_x^+)^2 - x^2}{h_x^+} - \frac{x^2 - (x-h_x^-)^2}{h_x^-} \right) \right. \\
&\quad \left. + \frac{1}{\overline{h}_y} \left(\frac{(y+h_y^+)^2 - y^2}{h_y^+} - \frac{y^2 - (y-h_y^-)^2}{h_y^-} \right) \right] \\
&= -\alpha \left(\frac{h_x^+ + h_x^-}{\overline{h}_x} + \frac{h_y^+ + h_y^-}{\overline{h}_y} \right) =: -\overline{f}(\boldsymbol{x}), \ \boldsymbol{x} \in \omega_h^*.
\end{aligned}
$$

Hence, $\overline{u}(\boldsymbol{x})$ is the solution of the problem

$$
\begin{aligned}
L_h \overline{u}(\boldsymbol{x}) &= \overline{f}(\boldsymbol{x}), && \boldsymbol{x} \in \omega_h, \\
\overline{u}(\boldsymbol{x}) &= \alpha \left(R^2 - x^2 - y^2 \right) \geq 0, && \boldsymbol{x} \in \gamma_h.
\end{aligned}
$$

It is $\overline{u}(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in \gamma_h$. Choosing $\alpha = \frac{1}{4} \left\| f^\circ \right\|_{l^\infty(\omega_h)}$, one obtains

$$
\begin{aligned}
&\overline{f}(\boldsymbol{x}) = 4\alpha = \left\| f^\circ \right\|_{l^\infty(\omega_h)} \geq \left| f^\circ(\boldsymbol{x}) \right|, \ \boldsymbol{x} \in \omega_h^\circ, \\
&\overline{f}(\boldsymbol{x}) \geq 0 = \left| f^\circ(\boldsymbol{x}) \right|, \ \boldsymbol{x} \in \omega_h^*.
\end{aligned}
$$

Now, Corollary 2.24 (Comparison Lemma) can be applied, which leads to

$$
\left\| u_1^\circ \right\|_{l^\infty(\omega_h)} \leq \left\| \overline{u} \right\|_{l^\infty(\omega_h)} \leq \alpha R^2 = \frac{R^2}{4} \left\| f^\circ \right\|_{l^\infty(\omega_h)}. \tag{2.18}
$$

One gets the last 'lower or equal' estimate because $(0,0)$ does not need to belong to Ω or ω_h. □

Remark 2.32. Estimate for the nodes that are close to the boundary. Corollary 2.27 can be applied to estimate $u_1^*(\boldsymbol{x})$. For $\boldsymbol{x} \in \omega_h^*$, one has

$$
\tilde{d}(\boldsymbol{x}) = a(\boldsymbol{x}) - \sum_{\boldsymbol{y} \in S(\boldsymbol{x}), \boldsymbol{y} \notin \gamma_h} b(\boldsymbol{x}, \boldsymbol{y}).
$$

Consider again for simplicity the one-dimensional case. With the approach from Example 2.18, one finds, using the definition of $\overline{h}_x$ and $h_x^- = h_x \geq h_x^+$ that

$$
\begin{aligned}
\tilde{d}(x) &= \frac{1}{\overline{h}_x} \left(\frac{1}{h_x^+} + \frac{1}{h_x^-} \right) - \frac{1}{\overline{h}_x h_x^-} = \frac{1}{\overline{h}_x h_x^+} = \frac{2}{h_x h_x^+ + h_x^+ h_x^+} \\
&\geq \frac{2}{h_x h_x + h_x h_x} = \frac{1}{h_x h_x} > 0.
\end{aligned}
$$

Hence, it is

$$\tilde{d}(\boldsymbol{x}) \geq \frac{1}{h^2}$$

with $h = \max\{h_x, h_y\}$. One obtains with Corollary 2.27 that

$$\|u_1^*\|_{l^\infty(\omega_h)} \leq \left\|D^+ f^*\right\|_{l^\infty(\omega_h)} \leq h^2 \|f^*\|_{l^\infty(\omega_h)} . \tag{2.19}$$

□

Lemma 2.33. Stability estimate. *The solution of the discrete Dirichlet problem* (2.6) *with* $\phi(\boldsymbol{x}) = f(\boldsymbol{x})$ *satisfies*

$$\|u\|_{l^\infty(\omega_h\cup\gamma_h)} \leq \|g\|_{l^\infty(\gamma_h)} + \frac{R^2}{4} \|f\|_{l^\infty(\omega_h^\circ)} + h^2 \|f\|_{l^\infty(\omega_h^*)} \tag{2.20}$$

with $R \geq \|\boldsymbol{x}\|_2$ *for all* $\boldsymbol{x} \in \Omega$ *and* $h = \max\{h_x, h_y\}$*, i.e., the solution* $u(\boldsymbol{x})$ *can be bounded in the norm* $\|\cdot\|_{l^\infty(\omega_h\cup\gamma_h)}$ *by the data of the problem.*

Proof. The statement of the lemma is obtained by combining the estimates (2.17), (2.18), and (2.19). ■

Convergence

Theorem 2.34. Convergence. *Let* $u(\boldsymbol{x})$ *be the solution of the Poisson equation* (2.1) *and* $u_h(\boldsymbol{x})$ *be the finite difference approximation given by the solution of* (2.6) *with* $\phi(\boldsymbol{x}) = f(\boldsymbol{x})$*. Then, it is*

$$\|u - u_h\|_{l^\infty(\omega_h\cup\gamma_h)} \leq Ch^2$$

with $h = \max\{h_x, h_y\}$.

Proof. The error in the node (x_i, y_j) is defined by $e_{ij} = u(x_i, y_j) - u_h(x_i, y_j)$. With the consistency relation $-\Lambda u(x_i, y_j) = -\Delta u(x_i, y_j) + \mathcal{O}\left(h^2\right)$, the Poisson equation (2.1) and the finite difference problem (2.6), one obtains for interior nodes

$$\begin{aligned} -\Lambda e(x_i, y_j) &= -\Lambda u(x_i, y_j) + \Lambda u_h(x_i, y_j) = -\Delta u(x_i, y_j) + \mathcal{O}\left(h^2\right) - f(x_i, y_i) \\ &= f(x_i, y_i) + \mathcal{O}\left(h^2\right) - f(x_i, y_i) = \mathcal{O}\left(h^2\right). \end{aligned}$$

Performing a similar calculation for the nodes close to the boundary leads to the following problem for the error

$$\begin{aligned} -\Lambda e(\boldsymbol{x}) &= \psi(\boldsymbol{x}),\ \boldsymbol{x} \in w_h^\circ,\ \psi(\boldsymbol{x}) = \mathcal{O}\left(h^2\right), \\ -\Lambda^* e(\boldsymbol{x}) &= \psi(\boldsymbol{x}),\ \boldsymbol{x} \in w_h^*,\ \psi(\boldsymbol{x}) = \mathcal{O}(h), \\ e(\boldsymbol{x}) &= 0, \qquad \boldsymbol{x} \in \gamma_h, \end{aligned}$$

where $\psi(\boldsymbol{x})$ is the consistency error, see Section 2.2. Applying the stability estimate (2.20) to this problem, one obtains immediately

$$\|e\|_{l^\infty(\omega_h\cup\gamma_h)} \leq \frac{R^2}{4} \|\psi\|_{l^\infty(\omega_h^\circ)} + h^2 \|\psi\|_{l^\infty(\omega_h^*)} = \mathcal{O}\left(h^2\right).$$

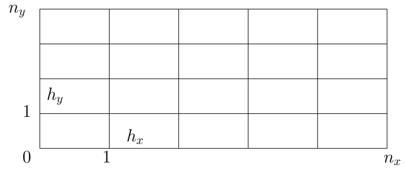

Fig. 2.6 Grid for the Dirichlet problem in the rectangular domain.

■

2.5 An Efficient Solver for the Dirichlet Problem in the Rectangle

Remark 2.35. Contents of this section. This section considers the Poisson equation (2.1) in the special case $\Omega = (0, l_x) \times (0, l_y)$. In this case, a modification of the difference stencil in a neighborhood of the boundary of the domain is not needed. The convergence of the finite difference approximation was already established in Theorem 2.34. Applying this approximation results in a large linear system of equations $A\boldsymbol{u} = \underline{f}$ which has to be solved. This section discusses some properties of the matrix A and it presents an approach for solving this system in the case of a rectangular domain in an almost optimal way.

A number of result obtained here will be needed also in Section 2.6. □

Remark 2.36. The considered problem and its approximation. The considered continuous problem consists in solving

$$\begin{aligned} -\Delta u &= f \text{ in } \Omega = (0, l_x) \times (0, l_y), \\ u &= g \text{ on } \partial\Omega, \end{aligned}$$

and the corresponding discrete problem in solving

$$\begin{aligned} -\Lambda u(\boldsymbol{x}) &= f(\boldsymbol{x}),\ \boldsymbol{x} \in \omega_h, \\ u(\boldsymbol{x}) &= g(\boldsymbol{x}),\ \boldsymbol{x} \in \gamma_h, \end{aligned}$$

where the discrete Laplacian is of the form (for simplicity of notation, the subscript h is omitted)

$$\Lambda u = \frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{h_x^2} + \frac{u_{i,j+1} - 2u_{ij} + u_{i,j-1}}{h_y^2} =: \Lambda_x u + \Lambda_y u, \quad (2.21)$$

with $h_x = l_x/n_x, h_y = l_y/n_y$, $i = 0, \ldots, n_x, j = 0, \ldots, n_y$, see Figure 2.6. □

Remark 2.37. The linear system of equations. The difference scheme (2.21) is equivalent to a linear system of equations $A\underline{u} = \underline{f}$.

For assembling the matrix and the right-hand side of the system, usually a lexicographical enumeration of the nodes of the grid is used. The nodes are called enumerated lexicographically if the node (i_1, j_1) has a smaller number than the node (i_2, j_2), if for the corresponding coordinates, it is

$$y_1 < y_2 \quad \text{or } (y_1 = y_2) \wedge (x_1 < x_2).$$

Using this lexicographical enumeration of the nodes, one obtains for the inner nodes a system of the form

$$\begin{aligned}
A &= \text{BlockTriDiag}(C, B, C) \in \mathbb{R}^{(n_x-1)(n_y-1)\times(n_x-1)(n_y-1)},\\
B &= \text{TriDiag}\left(-\frac{1}{h_x^2}, \frac{2}{h_x^2} + \frac{2}{h_y^2}, -\frac{1}{h_x^2}\right) \in \mathbb{R}^{(n_x-1)\times(n_x-1)},\\
C &= \text{Diag}\left(-\frac{1}{h_y^2}\right) \in \mathbb{R}^{(n_x-1)\times(n_x-1)},
\end{aligned}$$

$$\underline{f} = \begin{cases}
f(\boldsymbol{x}), & \boldsymbol{x} \in \omega_h^\circ,\\
f(\boldsymbol{x}) + \dfrac{g(x \pm h_x, y)}{h_x^2}, & \boldsymbol{x} \in \omega_h^*, \text{ close to right}\\
& \text{or left boundary},\\
f(\boldsymbol{x}) + \dfrac{g(x, y \pm h_y)}{h_y^2}, & \boldsymbol{x} \in \omega_h^*, \text{ close to upper}\\
& \text{or lower boundary},\\
f(\boldsymbol{x}) + \dfrac{g(x \pm h_x, y)}{h_x^2} + \dfrac{g(x, y \pm h_y)}{h_y^2}, & \boldsymbol{x} \in \omega_h^*, \text{ corner of inner nodes.}
\end{cases} \tag{2.22}$$

In this approach, the known Dirichlet boundary values are already substituted into the system and they appear in the right-hand side vector. The matrices B and C possess some modifications for nodes that have a neighbor on the boundary.

The linear system of equations has the following properties:

- high dimension: $N = (n_x - 1)(n_y - 1) \sim 10^3 \cdots 10^7$,
- sparse: per row and column of the matrix there are only 3, 4, or 5 non-zero entries,
- symmetric: hence, all eigenvalues are real,
- positive definite: all eigenvalues are positive. It holds that

$$\begin{aligned}
\lambda_{\min} &= \lambda_{(1,1)} \sim \pi^2\left(\frac{1}{l_x^2} + \frac{1}{l_y^2}\right) = \mathcal{O}(1),\\
\lambda_{\max} &= \lambda_{(n_x-1,n_y-1)} \sim \pi^2\left(\frac{1}{h_x^2} + \frac{1}{h_y^2}\right) = \mathcal{O}\left(h^{-2}\right),
\end{aligned} \tag{2.23}$$

with $h = \max\{h_x, h_y\}$, see Remark 2.38 below.

- high condition number: For the spectral condition number of a symmetric and positive definite matrix, it is

$$\kappa_2(A) = \frac{\lambda_{\max}}{\lambda_{\min}} = \mathcal{O}\left(h^{-2}\right).$$

Since the dimension of the matrix is large and the matrix is sparse, iterative solvers are an appropriate approach for solving the linear system of equations. The main costs for iterative solvers are the matrix-vector multiplications (often one per iteration). The cost of one matrix-vector multiplication is for sparse matrices proportional to the number of unknowns. Hence, an optimal solver with respect to the number of floating point operations is given if the number of operations for solving the linear system of equations is proportional to the number of unknowns. It is known that the number of iterations of many iterative solvers depends on the condition number of the matrix:

- *(damped) Jacobi method, SOR, SSOR.* The number of iteration is proportional to $\kappa_2(A)$. That means, if the grid is refined once, $h \to h/2$, then the number of unknowns is increased by around the factor 4 in two dimensions and also the number of iterations increases by a factor of around 4. Altogether, for one refinement step, the total costs increase by a factor of around 16.
- *(preconditioned) conjugate gradient (PCG) method.* The number of iterations is proportional to $\sqrt{\kappa_2(A)}$, see the corresponding theorem from the class Numerical Mathematics II. Hence, the total costs increase by a factor of around 8 if the grid is refined once.
- *multigrid methods.* For multigrid methods, the number of iterations on each grid is bounded by a constant that is independent of the grid. Hence, the total costs are proportional to the number of unknowns and these methods are optimal. However, the implementation of multigrid methods is involved.

$\square$

Remark 2.38. An eigenvalue problem. The derivation of an alternative direct solver is based on the eigenvalues and eigenvectors of the discrete Laplacian. It is possible to computed these quantities only in special situations, e.g., if the Poisson problem with Dirichlet boundary conditions is considered, the domain is rectangular, and the Laplacian is approximated with the five point stencil.

Consider the following eigenvalue problem

$$\begin{aligned} -\Lambda v(\boldsymbol{x}) &= \lambda v(\boldsymbol{x}), \ \boldsymbol{x} \in \omega_h, \\ v(\boldsymbol{x}) &= 0, \qquad\quad \boldsymbol{x} \in \gamma_h. \end{aligned}$$

Denote the node $\boldsymbol{x} = (x_i, y_j)$ by $\boldsymbol{x}_{ij}$ and grid functions in a similar way. The solution of this problem is sought in (tensor-)product form (separation of variables)

$$v_{ij}^{(\boldsymbol{k})} = v_i^{(k_x),x} v_j^{(k_y),y}, \quad \boldsymbol{k} = (k_x, k_y)^T.$$

It is

$$\Lambda v_{ij}^{(\boldsymbol{k})} = \left(\Lambda_x v_i^{(k_x),x}\right) v_j^{(k_y),y} + v_i^{(k_x),x} \left(\Lambda_y v_j^{(k_y),y}\right) = -\lambda_{\boldsymbol{k}} v_i^{(k_x),x} v_j^{(k_y),y},$$

where $i = 0, \ldots, n_x$, $j = 0, \ldots, n_y$ refers to the nodes and $k_x = 1, \ldots, n_x - 1$, $k_y = 1, \ldots, n_y - 1$ refers to the eigenvalues. Note that the number of eigenvalues is equal to the number of inner nodes, i.e., it is $(n_x - 1)(n_y - 1)$. In this ansatz, also a splitting of the eigenvalues in a contribution from the x coordinate and a contribution from the y coordinate is included. From the boundary condition, it follows that

$$v_0^{(k_x),x} = v_{n_x}^{(k_x),x} = v_0^{(k_y),y} = v_{n_y}^{(k_y),y} = 0.$$

Dividing by $v_i^{(k_x),x} v_j^{(k_y),y}$ and rearranging terms, the eigenvalue problem can be split

$$\frac{\Lambda_x v_i^{(k_x),x}}{v_i^{(k_x),x}} + \lambda_{k_x}^{(x)} = -\frac{\Lambda_y v_j^{(k_y),y}}{v_j^{(k_y),y}} - \lambda_{k_y}^{(y)}$$

with $\lambda_{\boldsymbol{k}} = \lambda_{k_x}^{(x)} + \lambda_{k_y}^{(y)}$. Both sides of this equation have to be constant since one of them depends only on i, i.e., on x, and the other one only on j, i.e., on y. The splitting of $\lambda_{\boldsymbol{k}}$ can be chosen so that the constant is zero. Then, one gets

$$\Lambda_x v_i^{(k_x),x} + \lambda_{k_x}^{(x)} v_i^{(k_x),x} = 0, \quad \Lambda_y v_j^{(k_y),y} + \lambda_{k_y}^{(y)} v_j^{(k_y),y} = 0.$$

The solution of these eigenvalue problems is known (exercise)

$$\begin{aligned} v_i^{(k_x),x} &= \sqrt{\frac{2}{l_x}} \sin\left(\frac{k_x \pi i}{n_x}\right), \quad \lambda_{k_x}^{(x)} = \frac{4}{h_x^2} \sin^2\left(\frac{k_x \pi}{2 n_x}\right), \\ v_j^{(k_y),y} &= \sqrt{\frac{2}{l_y}} \sin\left(\frac{k_y \pi j}{n_y}\right), \quad \lambda_{k_y}^{(y)} = \frac{4}{h_y^2} \sin^2\left(\frac{k_y \pi}{2 n_y}\right). \end{aligned}$$

It follows that the solution of the full eigenvalue problem is

$$v_{ij}^{(\boldsymbol{k})} = \frac{2}{\sqrt{l_x l_y}} \sin\left(\frac{k_x \pi i}{n_x}\right) \sin\left(\frac{k_y \pi j}{n_y}\right), \tag{2.24}$$

$$\lambda_{\boldsymbol{k}} = \frac{4}{h_x^2} \sin^2\left(\frac{k_x \pi}{2 n_x}\right) + \frac{4}{h_y^2} \sin^2\left(\frac{k_y \pi}{2 n_y}\right), \tag{2.25}$$

with $i = 0, \ldots, n_x, j = 0, \ldots, n_y$ and $k_x = 1, \ldots, n_x - 1, k_y = 1, \ldots, n_y - 1$. For every index $\boldsymbol{k} = (k_x, k_y)$, the eigenvalue is given by (2.25) and the entry of the corresponding eigen-grid-function $v(\boldsymbol{x})$ in the node $\boldsymbol{x}_{ij}$ is given by (2.24).

Using a Taylor series expansion, one obtains now the asymptotic behavior of the eigenvalues as given in (2.23). Note that because of the splitting of the eigenvalues into the directional contributions, the number of individual terms for computing the eigenvalues is only proportional to $(n_x + n_y)$. □

Remark 2.39. On the eigenvectors, weighted Euclidean inner product. Since the matrix corresponding to Λ is symmetric, the eigenvectors are orthogonal with respect to the Euclidean vector product. They become orthonormal with respect to the weighted Euclidean vector product

$$\langle u, v\rangle = h_x h_y \sum_{\boldsymbol{x}\in\omega_h\cup\gamma_h} u(\boldsymbol{x})v(\boldsymbol{x}) = h_x h_y \sum_{i=0}^{n_x}\sum_{j=0}^{n_y} u_{ij}v_{ij}, \tag{2.26}$$

with

$$h_x = \frac{l_x}{n_x}, h_y = \frac{l_y}{n_y},$$

i.e., then it is

$$\langle v^{(\boldsymbol{k})}, v^{(\boldsymbol{m})}\rangle = \delta_{\boldsymbol{k},\boldsymbol{m}}. \tag{2.27}$$

This property can be checked by using the relation

$$\sum_{i=0}^{n} \sin^2\left(\frac{i\pi}{n}\right) = \frac{n}{2}, \quad n > 1.$$

The norm induced by the weighted Euclidean vector product is given by

$$\|v\|_h = \langle v, v\rangle^{1/2} = \left(h_x h_y \sum_{i=0}^{n_x}\sum_{j=0}^{n_y} v_{ij}^2\right)^{1/2}. \tag{2.28}$$

The weights are such that this norm can be bounded for constant grid functions independently of the mesh, i.e.,

$$\|1\|_h = \left(h_x h_y (n_x+1)(n_y+1)\right)^{1/2} = \left(l_x l_y \frac{n_x+1}{n_x}\frac{n_y+1}{n_y}\right)^{1/2} \le 2\left(l_x l_y\right)^{1/2}. \tag{2.29}$$

□

Remark 2.40. Solver based on the eigenvalues and eigenvectors. Let $\phi(\boldsymbol{x})$ be the grid function corresponding to the right-hand side vector $\underline{f}$, see (2.22). Then, one uses the ansatz

$$\phi(\boldsymbol{x}) = \sum_{\boldsymbol{k}} \langle \phi, v^{(\boldsymbol{k})}\rangle v^{(\boldsymbol{k})}(\boldsymbol{x}) = \sum_{\boldsymbol{k}} \phi_{\boldsymbol{k}} v^{(\boldsymbol{k})}(\boldsymbol{x}) \tag{2.30}$$

with the Fourier coefficients

$$\phi_{\boldsymbol{k}} = \langle \phi, v^{(\boldsymbol{k})} \rangle = \frac{2h_x h_y}{\sqrt{l_x l_y}} \sum_{i=0}^{n_x} \sum_{j=0}^{n_y} \phi_{ij} \sin\left(\frac{k_x \pi i}{n_x}\right) \sin\left(\frac{k_y \pi j}{n_y}\right), \quad \boldsymbol{k} = (k_x, k_y),$$

with $\phi_{ij} = \phi(\boldsymbol{x}_{ij})$. The solution $u(\boldsymbol{x})$ of (2.21) is sought as a linear combination of the eigenfunctions

$$u(\boldsymbol{x}) = \sum_{\boldsymbol{k}} u_{\boldsymbol{k}} v^{(\boldsymbol{k})}(\boldsymbol{x})$$

with unknown coefficients $u_{\boldsymbol{k}}$. With this ansatz, one obtains for the finite difference operator

$$\Lambda u = \sum_{\boldsymbol{k}} u_{\boldsymbol{k}} \Lambda v^{(\boldsymbol{k})} = \sum_{\boldsymbol{k}} u_{\boldsymbol{k}} \lambda_{\boldsymbol{k}} v^{(\boldsymbol{k})}.$$

Since the eigenfunctions form a basis of the space of the grid functions, a comparison of the coefficients with the right-hand side (2.30) gives

$$-u_{\boldsymbol{k}} \lambda_{\boldsymbol{k}} = \phi_{\boldsymbol{k}} \quad \Longleftrightarrow \quad u_{\boldsymbol{k}} = -\frac{\phi_{\boldsymbol{k}}}{\lambda_{\boldsymbol{k}}}$$

or, for each component, using (2.24),

$$u_{ij} = -\sum_{\boldsymbol{k}} \frac{\phi_{\boldsymbol{k}}}{\lambda_{\boldsymbol{k}}} v_{ij}^{(\boldsymbol{k})} = -\frac{2h_x h_y}{\sqrt{l_x l_y}} \sum_{k_x=1}^{n_x-1} \sum_{k_y=1}^{n_y-1} \frac{\phi_{\boldsymbol{k}}}{\lambda_{\boldsymbol{k}}} \sin\left(\frac{k_x \pi i}{n_x}\right) \sin\left(\frac{k_y \pi j}{n_y}\right),$$

$i = 0, \ldots, n_x$, $j = 0, \ldots, n_y$.

It is possible to implement this approach with the Fast Fourier Transform (FFT) with

$$\mathcal{O}\left(n_x n_y \log_2 n_x + n_x n_y \log_2 n_y\right) = \mathcal{O}\left(N \log_2 N\right), \; N = (n_x - 1)(n_y - 1),$$

operations. Hence, this method is almost, up to a logarithmic factor, optimal. □

2.6 A Higher Order Discretization

Remark 2.41. Contents. The five point stencil is a second order discretization of the Laplacian. In this section, a discretization of higher order will be studied. In these studies, only the case of a rectangular domain $\Omega = (0, l_x) \times (0, l_y)$ and Dirichlet boundary conditions will be considered. □

Remark 2.42. Derivation of a fourth order approximation. Let $u(\boldsymbol{x})$ be the solution of the Poisson equation (2.1) and assume that $u(\boldsymbol{x})$ is sufficiently smooth. It is