**After today's class you should be able to**

- explain the principle of finite differencing
- assemble the "big linear system" corresponding to a finite-difference scheme
- define the consistency, stability and convergence of a scheme, use these terms correctly and describe the relationship between them
- determine the order of consistency of a finite-difference scheme
- prove the stability of the classical 5-point discretisation of Poisson's equation on equidistant rectangular grids
- recognise the limitations of the convergence analysis of finite-difference schemes

## 2.2 Finite Differences for Poisson's Equation

**Executive Summary** The finite-difference method yields a discrete approximation of the *strong formulation* of a PDE. A problem is discretised in three steps:

**1**. mesh $\Omega$ with a (normally Cartesian) point grid $\Omega^h$

**2**. approximate all derivatives with difference quotients

**3**. set up a system of equations $L^h u^h = f^h$ for the unknown function values $u^h$ on $\Omega^h$

Here we are going to focus on domains that are Cartesian products of intervals (intervals in 1D, rectangles in 2D, cuboids in 3D), where we usually limit our presentation to the 2D case. On such rectangular domains, finite-difference methods exhibit one of their greatest strengths: they are relatively simple and possess a lot of structure. From the didactic viewpoint, this setting makes it easier for us to discover concepts of the numerical analysis of PDEs. We will treat more general domains in the context of finite-element methods.

### Finite Differencing

We consider our well-known elliptic boundary value problem: Find $u \in C^2(\Omega) \cap C(\bar{\Omega})$ such that

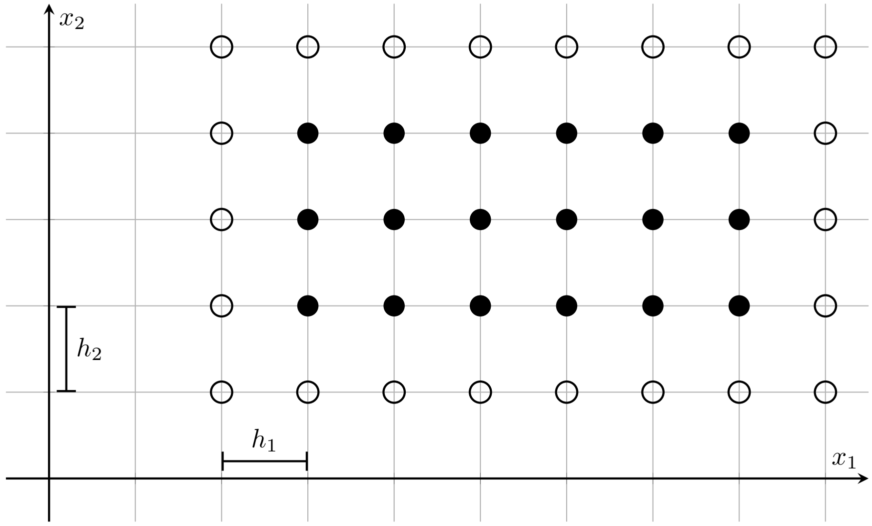
$$\begin{aligned} -\Delta u &= f && \text{in } \Omega \\ u &= g && \text{on } \partial\Omega \end{aligned} \tag{2.7}$$

on the rectangle $\Omega \subset \mathbb{R}^2 = ]a_1, b_1[ \times ]a_2, b_2[$. This domain is discretised with a Cartesian grid

$$\Omega^h = \left\{ (x_1, x_2) \in \mathbb{R}^2 \;\middle|\; x_1 = a_1 + i\frac{b_1 - a_1}{N_1},\ x_2 = b_1 + j\frac{b_2 - a_2}{N_2},\ (i = 1, \dots, N_1 - 1,\ j = 1, \dots, N_2 - 1) \right\}$$

using $N_1$ subintervals of constant grid spacing $h_1 = \frac{b_1 - a_1}{N_1} > 0$ in $x_1$-direction and $N_2$ subintervals of constant grid spacing $h_2 = \frac{b_2 - a_2}{N_2} > 0$ in $x_2$-direction. For the discrete domain with the boundary points included, we use the notation

$$\bar{\Omega}^h = \left\{ (x_1, x_2) \in \mathbb{R}^2 \;\middle|\; x_1 = a_1 + i\frac{b_1 - a_1}{N_1},\ x_2 = b_1 + j\frac{b_2 - a_2}{N_2},\ (i = 0, \dots, N_1,\ j = 0, \dots, N_2) \right\}.$$




Recall the definition of the partial derivative

$$\frac{\partial u}{\partial x_i}(x) = \lim_{h \to 0} \frac{u(x + he_i) - u(x)}{h}.$$

By dropping the limit on the right hand side, we obtain a difference quotient with finite $h > 0$ (hence the name *finite differencing*) that approximates the true derivative. Based on this principle, we introduce the following difference quotients:

**2.2.1 Definition (First-Order Difference Quotients)** Let $u \in C^1(\mathbb{R}^d)$. For $i \in \{1, \ldots, d\}$ we define

- the *forward difference quotient*
$$\partial_i^{+h} u(x) = \frac{u(x + he_i) - u(x)}{h}$$
- the *backward difference quotient*
$$\partial_i^{-h} u(x) = \frac{u(x) - u(x - he_i)}{h}$$
- the *central difference quotient*
$$\partial_i^{h} u(x) = \frac{u(x + he_i) - u(x - he_i)}{2h}$$

To approximate second partial derivatives, one can now compose two first-order difference quotients. For instance,

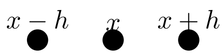
$$\partial_i^{+h}\partial_i^{-h} u(x) = \partial_i^{-h}\partial_i^{+h} u(x) = \frac{u(x-h) - 2u(x) + u(x+h)}{h^2}$$

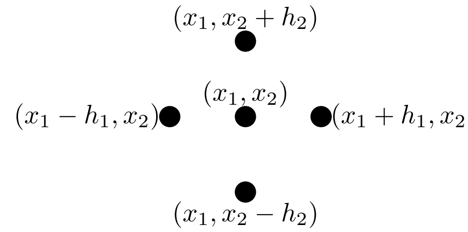
approximates the second partial derivative $\frac{\partial^2 u}{\partial x_i^2}$. Therefore, a finite-difference approximation of the negative Laplacian at $x \in \bar{\Omega}^h$ is given by

$$\begin{aligned} -\Delta^h u(x) &= -\sum_{i=1}^{d} \partial_i^{+h_i}\partial_i^{-h_i} u(x) \\ &= \begin{cases} \dfrac{-u(x-h) + 2u(x) - u(x+h)}{h^2} & \text{in 1D} \\ \dfrac{-u(x_1 - h_1, x_2) + 2u(x_1, x_2) - u(x_1 + h_1, x_2)}{h_1^2} + \dfrac{-u(x_1, x_2 - h_2) + 2u(x_1, x_2) - u(x_1, x_2 + h_2)}{h_2^2} & \text{in 2D} \end{cases} \end{aligned}$$

The last equation shows that the discrete Laplacian evaluated at the grid point $x$ depends on the function values at the point $x$ itself plus the two neighbouring grid points $x \pm h$ in 1D, or the four neighbouring grid points

$(x_1 \pm h_1, x_2 \pm h_2)$ in 2D, respectively. These dependency patterns are referred to as a *stencil*, here a 3-point stencil or a 5-point stencil, respectively.




For a computational implementation of the finite difference method, it is desirable to write the difference equations for the unknown function values of $u^h$ on the grid points of $\Omega^h$ in matrix form

$$L^h u^h = f^h.$$

The entries of the vector $f^h$ are the function values of the right hand side $f$ evaluated on each grid point, plus any source terms arising in difference quotients that refer to given boundary data.

For the Poisson-Dirichlet problem in 1D

$$L^h = \frac{1}{h^2}\left( \quad \right) \qquad f^h = \left( \quad \right).$$

In 2D, if the points in $\Omega^h$ are ordered row-wise from bottom left to top right, then $L^h = L_1^h + L_2^h$ with

$$L_1^h = \frac{1}{h_1^2}\left( \quad \right)$$

$$L_2^h = \frac{1}{h_2^2}\begin{pmatrix} \\ \end{pmatrix}$$

and

$$f^h = \begin{pmatrix} \\ \end{pmatrix} + \frac{1}{h_1^2}\begin{pmatrix} \\ \end{pmatrix} + \frac{1}{h_1^2}\begin{pmatrix} \\ \end{pmatrix} + \frac{1}{h_2^2}\begin{pmatrix} \\ \end{pmatrix} + \frac{1}{h_2^2}\begin{pmatrix} \\ \end{pmatrix}.$$

With the Kronecker product, these matrices and vectors can be written in more compact form

$$L_1^h = \frac{1}{h_1^2}\begin{pmatrix} \\ \end{pmatrix} \otimes \begin{pmatrix} \\ \end{pmatrix} \qquad L_2^h = \frac{1}{h_2^2}\begin{pmatrix} \\ \end{pmatrix} \otimes \begin{pmatrix} \\ \end{pmatrix}.$$

(and similarly for the boundary terms on the right hand side).

It is important to note that the discrete negative Laplacian $L^h$ is represented by a *sparse matrix*. If $L^h$ is an $N \times N$ matrix, it contains a total number of $N^2$ entries. The 5-point stencil of the 2D central differencing scheme reveals that out of the $N$ entries in each row, at most 5 can be non-zero. For points near the boundary, the stencil includes known boundary values. Therefore, the corresponding rows of $L^h$ have 3 or 4 non-zero entries only. Rows corresponding to grid points further away from the boundary contain 5 non-zero entries.

Consequently, only $O(N)$ out of the $N^2$ entries of $L^h$ are non-zero. For this purpose, software packages for scientific computing usually offer a special data type for sparse matrices where only the non-zero entries and their indices are stored, but not all the $O(N^2)$ zero entries. Since PDE problems lead to systems with very large $N$ ($N \sim 10^7 - 10^9$ for most industrial problems), the memory requirements for the full matrix would be excessive. Furthermore, the multiplication of a sparse matrix with a vector involves only $O(N)$ multiplications of non-zero numbers. If the sparse matrix had been stored fully, a computer would still carry out all $N^2$ multiplications of matrix entries with vector entries, despite the fact that almost all of these multiplications give zero. Sparse data formats hence allow for an economical use of both memory and CPU resources.

### Fundamental Notions of the Numerical Analysis of PDEs

Many problems in numerical analysis have the form of a root-finding problem $T(u) = 0$. This includes in particular ODEs and PDEs (e.g. $T(u) = Lu - f$), no matter whether or not they are linear. Three closely related notions describe a numerical scheme $T^h(u^h) = 0$ that discretises the original problem $T(u) = 0$, where $h > 0$ is some discretisation parameter (typically a measure of the grid spacing or a time step size):

(1) *Consistency:* Is the discretisation scheme $T^h$ a good approximation of the exact problem $T$? Or, somewhat more precisely, does the true solution $u$ only produce a small residual $T^h(u)$ in the numerical approximation of the problem?

(2) *Stability:* Is the discrete problem $T^h(u^h) = 0$ well-posed with continuous dependence on the data independently of $h$? In other words, does a small residual $T^h(v^h)$ imply a small error $u^h - v^h$?

(3) *Convergence:* Is the discrete solution $u^h$ a good approximation of the exact solution $u$?

**2.2.2 Definition (Consistency)** The numerical scheme $T^h$ is said to be

1. *consistent*, if for every solution of $T(u) = 0$

$$\|T^h(u)\| \to 0 \qquad \text{as } h \to 0.$$

2. *consistent of order* $O(h^p)$, if additionally

$$\|T^h(u)\| = O(h^p) \qquad \text{as } h \to 0.$$

The residual $T^h(u)$ is sometimes called *consistency error*, and if the inconsistency arises due to a truncated Taylor series, then it is also known as *truncation error*.

**2.2.3 Definition (Stability)** The numerical scheme $T^h$ is said to be *stable* (with respect to $h$) if there are constants $h_0 > 0$ and $C > 0$ such that for all $h \in ]0, h_0]$ the discrete problem $T^h(u^h) = 0$ has a unique solution and if additionally the stability inequality

$$\|u^h - v^h\| \leq C\|T^h(v^h)\|$$

holds for all discrete functions $v^h$ and all $h \in ]0, h_0]$.

Note that since $T^h(u^h) = 0$, we could have equally written $\|T^h(v^h)\| = \|T^h(u^h) - T^h(v^h)\|$, and then the stability inequality looks like an inverse Lipschitz condition.

**2.2.4 Definition (Convergence)** The numerical scheme $T^h$ is said to be

1. *convergent*, if there exists a constant $h_0 > 0$ such that for all $h \in ]0, h_0]$ the discrete problem $T^h(u^h) = 0$ has a unique solution and if with the solution of $T(u) = 0$

$$\|u^h - u\| \to 0 \qquad \text{as } h \to 0.$$

**2**. *convergent of order* $O(h^p)$, if additionally

$$\|u^h - u\| = O(h^p) \qquad \text{as } h \to 0.$$

The difference $u^h - u$ between the numerical and the exact solution is called the *error*.

**2.2.5 Theorem (Consistency $\wedge$ Stability $\Rightarrow$ Convergence)** If the scheme $T^h$ is

(a) consistent and stable, then $T^h$ is also convergent;

(b) consistent of order $p$ and stable, then $T^h$ is also convergent of order $p$.

*Proof.* For a scheme that is stable, we have existence and uniqueness of a discrete solution $u^h$ on sufficiently fine grids along with the estimate

$$\|u^h - u\| \le C\|T^h(u)\|$$

where the right hand side $\to 0$ (or $= O(h^p)$) if the scheme is consistent (of order $O(h^p)$). □

### Consistency of Finite-Difference Methods

So far, we have simply written down an ad hoc approximation of the Poisson-Dirichlet problem. We shall now (i) confirm that the approximations made do in fact lead to a consistent approximation and (ii) also find a more general, systematic approach for deriving a finite-difference scheme. This approach relies on Taylor expansions.

**2.2.6 Lemma (Truncation Errors of First-Order Difference Quotients)** If the line segment $\{\, x + the_i : -1 \le t \le 1 \,\}$ is fully contained in $\bar{\Omega}$ then

- provided that $u \in C^2(\bar{\Omega})$

$$\left|\partial_i^{\pm h} u(x) - \frac{\partial u}{\partial x_i}\right| \le \frac{h}{2} \max_{t\in[-1,1]} \left|\frac{\partial^2 u}{\partial x_i^2}(x + te_i)\right|$$

- provided that $u \in C^3(\bar{\Omega})$

$$\left|\partial_i^{h} u(x) - \frac{\partial u}{\partial x_i}\right| \le \frac{h^2}{6} \max_{t\in[-1,1]} \left|\frac{\partial^3 u}{\partial x_i^3}(x + te_i)\right|$$

*Proof.* Since we're just looking at differences along a (one-dimensional) line, we'll simplify our notation and consider the one-dimensional case only.

- For the one-sided difference quotients where the solution is assumed to be in $C^2(\bar{\Omega})$, we expand

$$u(x \pm h) = u(x) \pm hu'(x) + \frac{h^2}{2} u''(\xi_\pm)$$

  where the second derivative is evaluated at some point $\xi_+ \in ]x, x+h[$ or $\xi_- \in ]x-h, x[$, respectively. For the difference between the finite-difference approximation and the exact derivative, the local truncation error, we obtain

$$\begin{aligned}
\left|\partial^{\pm h} u(x) - u'(x)\right| &= \left|\frac{\pm u(x \pm h) \mp u(x)}{h} - u'(x)\right| \\
&= \left|\frac{\pm u(x) + hu'(x) \pm \frac{h^2}{2} u''(\xi_\pm) \mp u(x)}{h} - u'(x)\right| \\
&= \left|\frac{h}{2} u''(\xi_\pm)\right| \\
&\le \frac{h}{2} \max_{[x-h, x+h]} |u''|.
\end{aligned}$$

- For the central difference quotient and a solution $u \in C^3(\bar\Omega)$, we expand

$$u(x \pm h) = u(x) \pm hu'(x) + \frac{h^2}{2}u''(x) \pm \frac{h^3}{6}u'''(\xi_\pm)$$

with $\xi_+ \in ]x, x+h[$ or $\xi_- \in ]x-h, x[$, respectively. The local truncation error can be estimated as follows:

$$\begin{aligned}
\left|\partial^h u(x) - u'(x)\right| &= \left|\frac{u(x+h) - u(x-h)}{2h} - u'(x)\right| \\
&= \left|\frac{u(x) + hu'(x) + \frac{h^2}{2}u''(x) + \frac{h^3}{6}u'''(\xi_+) - u(x) + hu'(x) - \frac{h^2}{2}u''(x) + \frac{h^3}{6}u'''(\xi_-)}{2h} - u'(x)\right| \\
&= \left|\frac{h^2}{12}u''(\xi_+) + \frac{h^2}{12}u''(\xi_-)\right| \\
&\le \frac{h^2}{6}\max_{[x-h,x+h]}|u'''|.
\end{aligned}$$

□

**2.2.7 Lemma (Truncation Error of the Three-Point Second-Order Difference Quotient)** If the line segment $\{\, x + the_i : -1 \le t \le 1 \,\}$ is fully contained in $\bar\Omega$ then, provided that $u \in C^4(\bar\Omega)$,

$$\left|\partial_i^{+h}\partial_i^{-h}u(x) - \frac{\partial^2 u}{\partial x_i^2}\right| \le \frac{h^2}{12}\max_{t\in[-1,1]}\left|\frac{\partial^4 u}{\partial x_i^4}(x + te_i)\right|$$

*Proof.* We consider the one-dimensional case case again.

Taylor expansion leads to

$$u(x \pm h) = u(x) \pm hu'(x) + \frac{h^2}{2}u''(x) \pm \frac{h^3}{6}u'''(x) + \frac{h^4}{24}u^{(4)}(\xi_\pm)$$

and thus we obtain the estimate

$$\begin{aligned}
\left|\partial^{+h}\partial^{-h}u(x) - u''(x)\right| &= \left|\frac{u(x-h) - 2u(x) + u(x+h)}{h^2} - u''(x)\right| \\
&= \left|\frac{\frac{h^2}{2}u''(x) + \frac{h^4}{24}u^{(4)}(\xi_-) + \frac{h^2}{2}u''(x) + \frac{h^4}{24}u^{(4)}(\xi_+)}{h^2} - u''(x)\right| \\
&= \left|\frac{h^2}{24}u^{(4)}(\xi_-) + \frac{h^2}{24}u^{(4)}(\xi_+)\right| \\
&\le \frac{h^2}{12}\max_{[x-h,x+h]}|u^{(4)}|.
\end{aligned}$$

□

**2.2.8 Theorem (Consistency on Equidistant Grids)** If the grid $\Omega^h$ is equidistant in each direction, then the finite-difference discretisation for the Poisson-Dirichlet problem is $2^{\text{nd}}$ order consistent, provided that the solution $u$ of the continuous problem is in $C^4(\bar\Omega)$.

In 1D:

$$\max_{\bar\Omega^h}\left|-\Delta^h u - f\right| \le \frac{h^2}{12}\max_{\bar\Omega}|u^{(4)}| \tag{2.8}$$

In 2D:

$$\max_{\bar\Omega^h}\left|-\Delta^h u - f\right| \le \frac{h^2}{6}\max_{\bar\Omega}\left\{\left|\frac{\partial^4 u}{\partial x_1^4}\right|, \left|\frac{\partial^4 u}{\partial x_2^4}\right|\right\} \tag{2.9}$$

with $h = \max\{\, h_1, h_2 \,\}$.

*Proof.* According to Lemma 2.2.7, noting that $-u''(x) = f(x)$, we have

$$\max_{\bar{\Omega}^h} |T^h(u)| = \max_{x\in\bar{\Omega}^h} |-\Delta^h u(x) - f(x)| = \max_{x\in\bar{\Omega}^h} |-\partial^{+h}\partial^{-h} u(x) + u''(x)| \le \frac{h^2}{12} \max_{\bar{\Omega}^h} |u^{(4)}|.$$

For a 2D domain, we apply Lemma 2.2.7 once for the partial derivatives in $x_1$-direction and once for the partial derivatives in $x_2$-direction, then add up the two truncation errors:

$$\begin{aligned}
\max_{\bar{\Omega}^h} |T^h(u)| &= \max_{x\in\bar{\Omega}^h} |-\Delta^h u(x) - f(x)| \\
&= \max_{x\in\bar{\Omega}^h} \left| -\partial_1^{+h}\partial_1^{-h} u(x) - \partial_2^{+h}\partial_2^{-h} u(x) + \frac{\partial^2 u}{\partial x_1^2}(x) + + \frac{\partial^2 u}{\partial x_2^2}(x) \right| \\
&\le \max_{x\in\bar{\Omega}^h} \left| -\partial_1^{+h}\partial_1^{-h} u(x) + \frac{\partial^2 u}{\partial x_1^2}(x) \right| + \max_{x\in\bar{\Omega}^h} \left| -\partial_2^{+h}\partial_2^{-h} u(x) + \frac{\partial^2 u}{\partial x_2^2}(x) \right| \\
&\le \frac{h^2}{6} \max_{\bar{\Omega}} \left\{ \left| \frac{\partial^4 u}{\partial x_1^4} \right|, \left| \frac{\partial^4 u}{\partial x_2^4} \right| \right\}.
\end{aligned}$$

□

### Stability of Finite-Difference Methods

For the finite difference method on equidistant rectangular grids, the discrete matrix $L^h$ possesses a lot of structure with repeating patterns. In such special cases, its eigenvalues and eigenvectors can be determined analytically. We will exploit this for a simplified stability analysis. On more general domains, where the eigenvalues and eigenvectors are unknown, one would have to resort to different techniques that are analogous to the well-posedness proof of the continuous problem: discrete Green's functions and a discrete maximum principle. We will discuss the discrete maximum principle next time, but not bother with discrete Green's functions.

In the following analysis, we assume without loss of generality that the domain is the cube $]0,1[^d$. (If not, translate and re-scale the domain. This only changes the mesh parameter $h$ by a constant.)

Recall that the negative Laplacian on $]0,1[$ with homogeneous Dirichlet boundary conditions has the eigenfunctions

$$v_k(x) = sin(k\pi x) \qquad k = 1, 2, \dots \tag{2.10}$$

with corresponding eigenvalues

$$\lambda_k(x) = k^2\pi^2 \qquad k = 1, 2, \dots \tag{2.11}$$

Astonishingly, the eigenvectors of the discrete Laplacian $L^h$ are simply the eigenfunctions $v_k$ sampled at the grid points $h, 2h, \dots, 1-h$.

**2.2.9 Lemma (Eigenvalues and Eigenvectors of the Discrete Laplacian)** The eigenvalues of the matrix

$$L^h = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & & \ddots & \\ & & -1 & 2 \end{pmatrix} \in \mathbb{R}^{(N-1)\times(N-1)}$$

are

$$\lambda_k = \frac{2}{h^2} - \frac{2}{h^2}\cos\frac{k\pi}{N} \qquad k = 1, \dots, N-1$$

with eigenvectors

$$v_k = (\sin(k\pi h), \sin(2k\pi h), \dots, \sin((N-1)k\pi h))^\top$$

Here $N = 1/h$ is the number of subintervals.

*Proof.* Elementary calculations... □

First of all, we observe that all eigenvalues are positive. Consequently, the linear system of the discrete problem has a unique solution.

Furthermore, with any discrete function $v^h$ we derive

$$\begin{aligned} L^h(u^h - v^h) &= f^h - L^h v^h \\ \Rightarrow \qquad u^h - v^h &= (L^h)^{-1}(f^h - L^h v^h) \\ \Rightarrow \qquad |u^h - v^h| &\le \|(L^h)^{-1}\||(f^h - L^h v^h)| \le \frac{1}{\lambda_{\min}}|(f^h - L^h v^h)| \end{aligned}$$

where

$$\lambda_{\min} = \lambda_1 = \frac{2}{h^2} - \frac{2}{h^2}\cos\left(\frac{\pi}{N}\right) = \frac{2}{h^2}\left(1 - \cos(\pi h)\right). \tag{2.12}$$

On fine grids with small $h$ and large $N$,

$$\lambda_{\min} \approx \frac{2}{h^2}\pi^2 h^2 = \pi^2, \tag{2.13}$$

so the eigenvalues remain bounded away from zero and we have derived the stability inequality

$$|u^h - v^h| \le C|(f^h - L^h v^h)| = C\|T^h(v^h)\|.$$

If we use the maximum norm $|\cdot|_\infty$ instead of the Euclidean norm $|\cdot|$, the constant $C$ in this inequality may change.

On a rectangular domain in 2D, one may use separation of variables and then apply the same arguments.

**2.2.10 Theorem (Stability Inequality)** For the discrete Poisson-Dirichlet problem we have the stability inequality

$$|u^h - v^h|_\infty \le C|L^h v^h - f^h|_\infty \tag{2.14}$$

for all grid functions $v^h$ and with a constant $C$ that is independent of $h$.

### Convergence of Finite-Difference Methods

**2.2.11 Theorem (Convergence on Equidistant Grids)** Let $\Omega^h$ have constant grid spacing $h_1$ and $h_2$ in $x_1$- and $x_2$-direction, respectively. Then the finite difference discretisation for the Poisson-Dirichlet problem is $2^{\text{nd}}$ order convergent, provided that the solution $u$ to the continuous problem is in $C^4(\bar{\Omega})$.

*Proof.* Second-order consistency & stability ⇒ second-order convergence. □

**2.2.12 Remark (C⁴-Regularity up to the Boundary)** The assumption of an analytical solution in $C^4(\bar{\Omega})$ is not normally satisfied, as such a high regularity of the solution usually requires a sufficiently regular domain, with no corners. Even for the problem

$$-\Delta u = 1 \qquad \text{in } \Omega \qquad\qquad u = 0 \qquad \text{on } \partial\Omega$$

with $C^\infty$-data, $u \notin C^4(\bar{\Omega})$ if $\Omega$ is, for example, the square $]0,1[^2$: assuming that $u$ is four times continuously differentiable up to the boundary, then the PDE prescribes

$$-\Delta u|_{x=0} =$$

but the boundary condition implies

$$-\Delta u|_{x=0} =$$

Consequently, even with the perfectly smooth data in this example, the corners in the domain do not even allow for second derivatives that are continuous up to the boundary.

As we shall see, the convergence analysis of finite-element methods requires far weaker assumptions on the regularity of the solution and is therefore, in contrast to the analysis of finite-difference methods, applicable to many practical problems where solutions are not that smooth.