# Numerical Analysis - Part II

Anders C. Hansen

Lecture 2

# Solving PDEs with finite difference methods

Our goal is to solve the Poisson equation

$$\nabla^2 u = f \qquad (x, y) \in \Omega, \tag{1}$$

where  $\nabla^2 = \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$  is the Laplace operator and  $\Omega$  is an open connected domain of  $\mathbb{R}^2$  with a Jordan boundary, specified together with the *Dirichlet boundary condition* 

$$u(x,y) = \phi(x,y)$$
  $(x,y) \in \partial \Omega.$  (2)

(You may assume that  $f \in C(\Omega)$ ,  $\phi \in C^2(\partial \Omega)$ , but this can be relaxed by an approach outside the scope of this course.)

### Approximation of $\nabla^2$

We cannot solve

$$abla^2 u = f \qquad (x, y) \in \Omega,$$
(3)

directly, meaning that we typically do not have a closed form solution u, nor can we solve (3) directly on a computer.

However, we do know how to solve a linear system of equations

$$Ax = y, \qquad A \in \mathbb{R}^{N \times N}, \quad x, y \in \mathbb{R}^{N}.$$

If we can "approximate" a function u with a vector x, what should be the approximation of the operator  $\nabla^2$ ?

**Crazy Idea:** Use finite differences! After all, the derivative is the limit of differences of the function with some scaling. Indeed,

$$u'(a) = \lim_{h \to 0} \frac{u(a+h) - u(a)}{h}$$

To this end we impose on  $\Omega$  a square grid with uniform spacing of h>0 and replace

$$\nabla^2 u = f \qquad (x, y) \in \Omega,$$

by a *finite-difference* formula. For simplicity, we require for the time being that  $\partial \Omega$  'fits' into the grid: if a grid point lies inside  $\Omega$  then all its neighbours are in cl  $\Omega$ . We will discuss briefly in the sequel grids that fail this condition.

### Example of a grid on a square



Figure: A square domain  $\Omega$  with an equidistant grid.

## Example of a grid on a more complicated domain



Figure: A more complicated domain  $\Omega$  with an equidistant grid.

### **Computational stencil**

We have the five-point method

$$u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1} - 4u_{i,j} = h^2 f_{i,j},$$
 (*ih*, *jh*)  $\in \Omega$ , (4)

where  $f_{i,j} = f(ih, jh)$  are given, and  $u_{i,j} \approx u(ih, jh)$  is an approximation to the exact solution. It is usually denoted by the following *computational stencil* 



Whenever  $(ih, jh) \in \partial\Omega$ , we substitute appropriate Dirichlet boundary values. Note that the outcome of our procedure is a set of linear algebraic equations whose solution approximates the solution of the Poisson equation (1) at the grid points.

## **Finite-difference discretization**

Finite-difference discretization of  $\nabla^2 u = f$  replaces the PDE by a large system of linear equations. In the sequel we pay special attention to the *five-point formula*, which results in the approximation

$$h^{2}\nabla^{2}u(x,y) \approx u(x-h,y) + u(x+h,y) + u(x,y-h) + u(x,y+h) - 4u(x,y).$$
(5)

For the sake of simplicity, we restrict our attention to the important case of  $\Omega$  being a *unit square*, where  $h = \frac{1}{m+1}$  for some positive integer *m*. Thus, we estimate the  $m^2$  unknown function values  $u(ih, jh)_{i,j=1}^m$  (where  $(ih, jh) \in \Omega$ ) by letting the right-hand side of (5) equal  $h^2 f(ih, jh)$  at each value of *i* and *j*. This yields an  $n \times n$  system of linear equations with  $n = m^2$  unknowns  $u_{i,j}$ :

$$u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1} - 4u_{i,j} = h^2 f(ih, jh).$$
 (6)

(Note that when *i* or *j* is equal to 1 or *m*, then the values  $u_{0,j}$ ,  $u_{i,0}$  or  $u_{i,m+1}$ ,  $u_{m+1,j}$  are known boundary values and they should be moved to the right-hand side, thus leaving fewer unknowns on the left.) Having ordered grid points, we can write (6) as a linear system, say

$$A \boldsymbol{u} = \boldsymbol{b}$$
 .

Our present concern is to prove that, as  $h \rightarrow 0$ , the numerical solution (6) tends to the exact solution of the Poisson equation  $\nabla^2 u = f$  (with appropriate Dirichlet boundary conditions).

### Example of a grid on a square



Figure: A square domain  $\Omega$  with an equidistant grid.

#### Example 1 (Natural ordering)

The way the matrix A of this system looks depends of course on the way how the grid points (ih, jh) are being assembled in the one-dimensional array. In the *natural ordering*, when the grid points are arranged by columns, A is the following block tridiagonal matrix:

$$A = \begin{bmatrix} B & I & & \\ I & B & I & \\ & \ddots & \ddots & \ddots & \\ & I & B & I \\ & & & I & B \end{bmatrix}, \qquad B = \begin{bmatrix} -4 & 1 & & \\ 1 & -4 & 1 & \\ & \ddots & \ddots & \ddots & \\ & 1 & -4 & 1 \\ & & & 1 & -4 \end{bmatrix}$$

Before heading on let us prove the following simple but useful theorem whose importance will become apparent in the course of the lecture.

### Theorem 2 (Gershgorin theorem)

All eigenvalues of an  $n \times n$  matrix A are contained in the union of the Gershgorin discs in the complex plane:

$$\sigma(A) \subset \cup_{i=1}^{n} \Gamma_{i}, \qquad \Gamma_{i} := \{ z \in \mathbb{C} : |z - a_{ii}| \le r_{i} \}, \qquad r_{i} := \sum_{j \neq i} |a_{ij}|.$$

### Proof of the Gershgorin theorem

**Proof.** Let  $\lambda$  be an eigenvalue of A. Choose a corresponding eigenvector  $x = (x_j)$  so that one component  $x_i$  is equal to 1 and the others are of absolute value less than or equal to  $x_i = 1$  and  $|x_j| \le 1$   $j \ne i$ . There is always such an x, which can be obtained simply by dividing any eigenvector by its component with largest modulus. Since  $Ax = \lambda x$ , in particular

$$\sum_{j} a_{ij} x_j = \lambda x_i = \lambda.$$

So, splitting the sum and taking into account once again that  $x_i = 1$ , we get

$$\sum_{j\neq i}a_{ij}x_j+a_{ii}=\lambda.$$

Therefore, applying the triangle inequality,

$$|\lambda - a_{ii}| = \left|\sum_{j \neq i} a_{ij} x_j\right| \le \sum_{j \neq i} |a_{ij}| |x_j| \le \sum_{j \neq i} |a_{ij}| = r_i.$$

#### Lemma 3

For any ordering of the grid points, the matrix A of the system (6) is symmetric and negative definite.

### **Proof** I

**Proof.** Equation (6) implies that if  $a_{ij} \neq 0$  for  $i \neq j$ , then the *i*-th and *j*-th points of the grid (ph, qh), are nearest neighbours. Hence  $a_{ij} \neq 0$  implies  $a_{ij} = a_{ji} = 1$ , which proves the symmetry of *A*. Therefore *A* has real eigenvalues and eigenvectors.

It remains to prove that all the eigenvalues are negative. The arguments are parallel to the proof of Gershgorin theorem. Let  $A\mathbf{x} = \lambda \mathbf{x}$ , and let *i* be an integer such that  $|x_i| = \max |x_j|$ . With such an *i* we address the following identity (which is a reordering of the equation  $(A\mathbf{x})_i = \lambda x_i$ ):

$$\underbrace{\left| (\lambda - \mathbf{a}_{ii}) \mathbf{x}_i \right|}_{|\lambda + 4| |\mathbf{x}_i|} = \underbrace{\left| \sum_{j \neq i}^n \mathbf{a}_{ij} \mathbf{x}_j \right|}_{\leq 4| \mathbf{x}_i|} \,. \tag{7}$$

Here  $a_{ii} = -4$  and  $a_{ij} \in \{0, 1\}$  for  $j \neq i$ , with at most four nonzero elements on the right-hand side. It is seen that the case  $\lambda > 0$  is impossible. Assuming  $\lambda = 0$ , we obtain  $|x_j| = |x_i|$  whenever  $a_{ij} = 1$ , so we can alter the value of *i* in (7) to any of such *j* and repeat the same arguments. Thus, the modulus of every component of *x* would be  $|x_i|$ , but then the equations (7) that occur at the boundary of the grid and have fewer than four off-diagonal terms (see (6)) could not be true. Hence,  $\lambda = 0$  is impossible too, hence  $\lambda < 0$  which proves that *A* is negative definite.

# Proof II

**Proof.** Let *U* be any linear operator changing the grid ordering. Then *U* is clearly unitary  $(||Ux||_2 = ||x||_2 \text{ for any } x)$ . Note that any matrix  $\tilde{A}$  representing the the system of equations (6) can be written as  $\tilde{A} = UAU^*$  for some unitary matrix *U*, where *A* is as in Example 1. Self-adjointness is preserved by unitary operators, and so is the spectrum. Thus,  $\tilde{A}$  is self-adjoint (symmetric as it is real). Moreover,  $\sigma(A)$  does not intersect the positive half plane by the Gershgorin theorem, so we only need to show that  $0 \notin \sigma(A)$ . If Ax = 0 then, by the definition of *A*, *x* must have elements of equal modulus, however, then the definition of *B* (that gives *A*) implies that x = 0.

#### Proposition 4

The eigenvalues of the matrix A are

$$\lambda_{k,\ell} = -4\left(\sin^2rac{k\pi h}{2}+\sin^2rac{\ell\pi h}{2}
ight), \qquad h=rac{1}{m+1}\,, \qquad k,\ell=1...m.$$

### Proof

**Proof.** Let us show that, for every pair  $(k, \ell)$ , the vectors

$$v = (v_{i,j}), \quad v_{i,j} = \sin ix \sin jy, \quad \text{where} \quad x = k\pi h, \quad y = \ell \pi h,$$

are the eigenvectors of A. Indeed, for i, j = 1...m, we have

$$\begin{aligned} (Av)_{i,j} &= \sin(jy) \left[ \sin(ix - x) - 2\sin(ix) + \sin(ix + x) \right] \\ &+ \sin(ix) \left[ \sin(jy - y) - 2\sin(jy) + \sin(jy + y) \right] \\ &= \sin(jy) \sin(ix) \left[ 2\cos x - 2 \right] + \sin(ix) \sin(jy) \left[ 2\cos y - 2 \right] = \lambda v_{i,j} \,. \end{aligned}$$

Note that the terms  $u_{i\pm 1,j}$ ,  $u_{i,j\pm 1}$  do not appear in (6) for i, j=1 or i, j=m, respectively, therefore (for such i, j) we should have dropped the corresponding components from above equation, but they are equal to zero because  $\sin(i-1)x = 0$  for i = 1, while  $\sin(i+1)x = 0$  for i = m, since  $x = \frac{k\pi}{m+1}$ . Thus, the eigenvalues are

$$\lambda_{k,\ell} = \left[2\cos x - 2\right] + \left[2\cos y - 2\right] = -4\left(\sin^2 \frac{x}{2} + \sin^2 \frac{y}{2}\right) \\ = -4\left(\sin^2 \frac{k\pi h}{2} + \sin^2 \frac{\ell\pi h}{2}\right).$$
(8)

### Remark

### Remark 5

As a matter of independent mathematical interest, note that for  $1 \leq k, \ell \ll m$  we have  $\sin x \approx x$ , hence the eigenvalues for the discretized Laplacian  $\nabla_h^2$  are

$$rac{\lambda_{k,\ell}}{h^2}pprox -rac{4}{h^2}\left[rac{k^2\pi^2h^2}{4}+rac{\ell^2\pi^2h^2}{4}
ight]=-(k^2+\ell^2)\pi^2\,.$$

Now, recall (e.g. from the solution of the Poisson equation in a square by separation of variables in Maths Methods) that the *exact* eigenvalues of  $\nabla^2$  (in the unit square) are  $-(k^2 + \ell^2)\pi^2$ ,  $k, \ell \in \mathbb{N}$ , with the corresponding eigenfunctions  $V_{k,\ell}(x, y) = \sin k\pi x \sin \ell\pi y$ . So, the eigenvectors of the discretized  $\nabla_h^2$  are the values of  $V_{k,\ell}(x, y)$  on the grid-points, and the eigenvalues of  $\nabla_h^2$  approximate those for continuous case.