IMT2112 - The finite difference method

Elwin van 't Wout*

October 3, 2019

Abstract

Many physical phenomena can be modelled with partial differential equations. In special cases, boundary value problems do have analytical solutions but for most problems of engineering interest they solution has to be approximated numerically. One of the most fundamental methods to numerically solve boundary value problems is the finite difference method. This short report[1] explains the basics of this numerical method and applies it to the Helmholtz equation for harmonic wave propagation.

1 The Helmholtz equation

Wave propagation problems are governed by the wave equation

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0 \tag{1}$$

where $u = u(x,t)$ denotes the unknown at location x and time t, and c a constant wavespeed. The wave equation allows for separable solutions, that is, $u(x,t) = v(x)w(t)$. For a given harmonic wave $w(\omega t)$ with angular frequency ω, for example $\cos(\omega t)$, the spatial unknown has to satisfy the Helmholtz equation

$$-\frac{d^2 v}{dx^2} - k^2 v = 0 \tag{2}$$

where $k = \omega/c$ denotes the wavenumber. Depending on the type of harmonic wave $w(t)$, the uknown v can be complex-valued. Furthermore, in some cases a complex-valued k is used as well.

*Instituto de Ingeniería Matemática y Computacional, Pontificia Universidad Católica de Chile, e.wout@uc.cl

[1]The purpose of these lecture notes is to introduce the methodology in undergraduate classes, not as scientific reference.

In the general case of the entire one-dimensional domain, i.e., $x \in \mathbb{R}$, the Helmholtz equation has analytical solutions, for example $v(x) = \cos(kx)$. However, when domains with boundaries are considered, analytical solutions are only available in rare cases. That is, analytical solutions of the Helmholtz equation are difficult to obtain, especially when higher dimensions are considered. For these cases, we need to use numerical methods that approximate the solution of the Helmholtz equation.

Example 1. Consider the boundary value problem

$$\begin{cases} -\frac{d^2v}{dx^2} - k^2 v = 0, \quad 0 < x < \pi; \\ v(0) = 0; \\ v(\pi) = 0. \end{cases} \tag{3}$$

One possible solution would be $v(x) = \sin(kx)$, which satisfies the Helmholtz equation and $v(0) = 0$. However, $v(\pi) = \sin(k\pi) = 0$ if and only if k is an integer. Therefore, this solution only satisfies the Helmholtz system for integer k. For non-integer k, the only solution is $v = 0$.

Remark 1. In general, the Helmholtz equation can have a non-zero right-hand side: $-\frac{d^2v}{dx^2} - k^2 v = f(x)$ for a given function f. This function typically represents sources or sinks of energy. Solving the Helmholtz equation for a general non-zero function f is complicated and analytical functions do not always exist.

2 Numerical discretisation

Because many *boundary value problems* (a differential equation with boundary conditions) do not have analytical solutions, but still have a unique solution, approximation methods need to be used. That is, instead of looking for a solution to the Helmholtz system on an interval, the idea of *discretisation methods* is to find approximations of the solution on a discrete set of points. There are two fundamental questions in the design of discretisation methods:

1. how to obtain the approximation, and
2. how many points do we need to get an accurate approximation.

Let us first look into the first question. In fact, many different approximation techniques have been proposed, of which the most common ones *finite differences*, *finite volumes*, and *finite elements*. Other common discretisation methods, with a slightly different approach, are *boundary element* and

spectral methods. Here, we will focus on the finite difference method (ES: *método de las diferencias finitas*). The second question depends on the numerical method used and the equation to be solved. More information on this for wave problems is given in Section 4.

3 The finite difference method

Let us consider the Helmholtz system

$$\begin{cases} -\frac{d^2v}{dx^2} - k^2 v = f, \quad 0 < x < \pi; \\ v(0) = 0; \\ v(\pi) = 0 \end{cases} \tag{4}$$

where $f = f(x)$ is a known function and $v = v(x)$ the unkown wave form. The function f typically models the source of the wave field and, when non-zero, prevents the trivial solution of $v = 0$ to be valid. There is no obvious analytical solution for general k and f and we therefore need an approximation method.

3.1 Computational grid

The first step for numerical methods is to define a *grid* or *mesh* (ES: *malla*), that is, the discrete points in which we are going to approximate the solution. Let us consider an equidistant grid, where the points are uniformly distributed over the interval. That is,

$$x_i = ih = i\frac{\pi}{N}, \text{ for } i = 0, 1, 2, \ldots, N \tag{5}$$

where h is called the *mesh width* or *grid spacing* (ES: *paso* or *ancho de subintervalo*) and each x_i a *node* (ES: *nodo*).

3.2 Finite difference approximation

The goal is to find a good approximation of the solution in each and every point of the computational grids.
The approximation in nodes x_0 and x_N is straightforward, namely the boundary conditions $v(x_0) = v(0) = 0$ and $v(x_N) = v(\pi) = 0$, which are called Dirichlet condition. Other boundary conditions can be used as well, for example Neumann conditions that depend on the derivative, such as $v'(0) = 0$. Notice that for Neumann boundary conditions, the value of v on the boundary is unknown and has to be approximated as well.

For the interior nodes, the only condition is that the function needs to satisfy the Helmholtz equation, that is,

$$-\frac{d^2v}{dx^2}(x_i) - k^2 v(x_i) = f(x_i), \text{ for } i = 1, 2, \ldots, N-1. \tag{6}$$

Remember that the goal is to find approximations of $v(x)$ in the nodes x_i only, so let us consider the set of unknowns v_i that approximate the solution $v(x_i)$. The aim is to find conditions for these unknowns. For this aim, we need to write each $\frac{d^2v}{dx^2}(x_i)$ as a function of v_j for $j = 1, 2, \ldots, N-1$. Notice that we can couple the approximation in one node with the approximation in neighbouring nodes. This leads us to the idea to use a Taylor series, that is,

$$v(x_{i+1}) = v(x_i) + h\frac{dv}{dx}(x_i) + \frac{1}{2}h^2\frac{d^2v}{dx^2}(x_i) + \frac{1}{6}h^3\frac{d^3v}{dx^3}(x_i) + \mathcal{O}(h^4), \tag{7}$$

$$v(x_{i-1}) = v(x_i) - h\frac{dv}{dx}(x_i) + \frac{1}{2}h^2\frac{d^2v}{dx^2}(x_i) - \frac{1}{6}h^3\frac{d^3v}{dx^3}(x_i) + \mathcal{O}(h^4). \tag{8}$$

Remember that the big-O notation means that if $f(x) = \mathcal{O}(h)$, there is a constant M and point x_0 such that $|f(x)| \leq M|h|$ for all $x \geq x_0$. Taking the difference of the two Taylor series leads to

$$\frac{v(x_{i+1}) - v(x_{i-1})}{2h} = \frac{dv}{dx}(x_i) + \mathcal{O}(h^2), \tag{9}$$

which means that $(v(x_{i+1}) - v(x_{i-1}))/(2h)$ is a good approximation of $\frac{dv}{dx}(x_i)$ for small h. This is called a *central finite difference approximation* of the first derivative. For the second derivative, we need another formula, that is,

$$\frac{v(x_{i+1}) - 2v(x_i) + v(x_{i-1})}{h^2} = \frac{d^2v}{dx^2}(x_i) + \mathcal{O}(h^2). \tag{10}$$

The $\mathcal{O}(h^2)$ term is called the *truncation error* (ES: *eror de truncamiento*).

Exercise 1. Show that the finite difference formula (10) has a truncation error of $O(h^2)$ and not just $O(h)$.

Exercise 2. Show that the forward finite difference formula

$$\frac{dv}{dx}(x_i) \approx \frac{v(x_{i+1}) - v(x_i)}{h}$$

has a truncation error of $O(h)$.

3.3 Finite difference matrix

The finite difference approximation is valid in all nodes and couples neighbouring nodes. To solve this set of equations, linear algebra is used.
First, let us substitute the finite difference approximation (10) for every internal node into the Helmholtz equation. With the observation that the truncation error is small when there are sufficient nodes in the mesh, one can use the equation

$$\frac{-v_{i+1} + 2v_i - v_{i-1}}{h^2} - k^2 v_i = f_i \text{ for } i = 1, 2, \ldots, N-1 \tag{11}$$

as approximation v_i of $v(x)$ in the node x_i. At the boundaries, $v_0 = 0$ and $v_n = 0$ because of the boundary conditions. Notice that we now have $N-1$ linear equations for $N-1$ unknowns. We can therefore use a matrix-vector notation

$$A\mathbf{u} = \mathbf{f}, \tag{12}$$

where

$$A = \frac{1}{h^2}\begin{bmatrix} 2-(kh)^2 & -1 & 0 & \ldots & \ldots & 0 \\ -1 & 2-(kh)^2 & -1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & -1 & 2-(kh)^2 & -1 \\ 0 & \ldots & \ldots & 0 & -1 & 2-(kh)^2 \end{bmatrix}$$

is the discretisation matrix, $\mathbf{u} = [u_1, u_2, \ldots, u_{N-1}]^T$ the vector of unknowns, and $\mathbf{f} = [f(x_1), f(x_2), \ldots, f(x_{N-1})]^T$ the known right-hand-side vector. Now, solving the set of linear equations gives the vector $\mathbf{u}$ and, therefore, an approximation of the solution in each node of the grid.

Exercise 3. Show that if $k = 0$, we have

$$\langle \mathbf{x}, h^2 A\mathbf{x} \rangle = x_1^2 + \sum_{m=1}^{N-2} (x_{m+1} - x_m)^2 + x_{N-1}^2$$

where the brackets denote the vector dot product. Explain from this result that the matrix A is positive definite for $k = 0$.

Exercise 4. Use the circle theorem of Gershgorin to show that the matrix A is positive definite if $\Re(k^2) < 0$ (remember that k can be complex-valued).

Exercise 5. Give a finite difference approximation and the resulting discretisation matrix for the convection-diffusion problem given by

$$\begin{cases} -a\frac{d^2v}{dx^2} + b\frac{dv}{dx} = f(x), \quad 0 < x < 1; \\ v(0) = 0; \\ v(1) = 1 \end{cases} \tag{13}$$

for positive parameters a, b.

Remark 2. For wave problems, the wavenumber k is usually real and positive, leading to an indefinite A. In general, indefinite systems have a large condition number and are, therefore, difficult to solve numerically. This is an important topic in computational sciences and the area of *numerical linear algebra* deals with the development of fast solvers for linear equations.

Remark 3. Solutions of the homogeneous Helmholtz equation $-v'' - k^2 v = 0$ that satisfy the boundary conditions are called *resonance modes* and only appear at specific values of k. These nonzero solutions are in the nullspace of the Helmholtz operator and therefore result in singularities of the discretisation matrix as well.

Remark 4. Many authors prefer to write the Helmholtz equation as $-v'' - k^2 v = 0$ instead of $v'' + k^2 v = 0$, which are equivalent. The finite difference matrix of the Laplace operator has negative values on the main diagonal, whereas the the matrix for the negative Laplace operator has positive values. Also, the finite difference matrix of $-v''$ is positive definite, instead of negative definite for v''. Since it is more intuitive to use positive definite matrices, $-v''$ is typically be used.

3.4 Neumann boundary conditions

Let us consider the Helmholtz system

$$\begin{cases} -\frac{d^2v}{dx^2} - k^2 v = f, \quad 0 < x < \pi; \\ v(0) = 0; \\ \frac{dv}{dx}(\pi) = 0 \end{cases} \tag{14}$$

where at the right boundary a Neumann condition is used.

Exercise 6. Show that $v(x) = \sin(kx)$ is a solution of this boundary value problem for $f = 0$ and specific values of k.

The finite difference method explained in Section 3.2 will not change for the left boundary condition and the Helmholtz equation. For the right boundary condition, we do not know the value $u(x_N)$ so we cannot do the same as before, and we need to treat u_N as an additional unknown. Notice that we can use the finite difference approximation

$$\frac{dv}{dx}(x_N) \approx \frac{u_N - u_{N-1}}{h} \tag{15}$$

which is first-order accurate. The resulting discretisation matrix is then given by

$$A = \frac{1}{h^2}\begin{bmatrix} 2-(kh)^2 & -1 & 0 & \dots & \dots & 0 \\ -1 & 2-(kh)^2 & -1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -1 & 2-(kh)^2 & -1 \\ 0 & \dots & \dots & 0 & -h & h \end{bmatrix}$$

for the solution vector $\mathbf{u} = [u_1, u_2, \dots, u_{N-1}, u_N]^T$.

Exercise 7. Show that if the error of the Neumann boundary condition is $\mathcal{O}(h)$ and for all other unknowns $\mathcal{O}(h^2)$, the overall error is $\mathcal{O}(h)$.

Exercise 8. Explain why the approximation $\frac{dv}{dx}(x_N) \approx \frac{u_{N+1}-u_{N-1}}{h}$, which is second-order accurate, cannot be used for the Neumann boundary condition.

Exercise 9. Given a finite-difference approximation for the Neumann boundary condition that is second-order accurate.
Hint: use the Taylor polynomial for $u(x_{N-2})$ with respect to x_N.

3.5 Two dimensional problems

Let us consider the two-dimensional Helmholtz system

$$\begin{cases} -\frac{d^2v}{dx^2} - \frac{d^2v}{dy^2} - k^2 v = f, & 0 < x, y < \pi; \\ v = 0, & \text{for } x = 0,\ x = \pi,\ y = 0, \text{ and } y = \pi \end{cases} \tag{16}$$

for the unknown $u = u(x, y)$.

Exercise 10. Show that the function $u(x, y) = \sin(k_x x)\sin(k_y y)$ is a solution of the homogeneous Helmholtz equation $-\nabla^2 u - k^2 u = 0$, but only when k_x and k_y satisfy a specific condition.

In most cases this Helmholtz problem does not have an analytical solution, so we need numerical methods to approximate the solution. In order to use the finite difference method, we first need to define the grid, which is now given by a set of points in a rectangle. Let us use an equidistant grid given by

$$\mathbf{x}_{i,j} = \begin{bmatrix} x_i \\ y_j \end{bmatrix} = \begin{bmatrix} ih_x \\ jh_y \end{bmatrix} \text{ for } i = 0, 1, 2, \ldots, N_x \text{ and } j = 0, 1, 2, \ldots, N_y \quad (17)$$

where $h_x = \pi/N_x$ and $h_y = \pi/N_y$ the mesh sizes. The finite difference approximation for interior nodes is now given by

$$\frac{-u_{i-1,j} + 2u_{i,j} - u_{i+1,j}}{h_x^2} + \frac{-u_{i,j-1} + 2u_{i,j} - u_{i,j+1}}{h_y^2} - k^2 u_{i,j} = f_{i,j}. \quad (18)$$

for the unknowns $u_{i,j}$.

Exercise 11. Show that this finite difference approximation is second-order accurate.

If we want to write two-dimensional finite differences in a matrix-vector form $A\mathbf{u} = \mathbf{b}$, we need to number the nodes in a one-dimensional way. The easiest approach is to count the nodes row-by-row or column-by-column. These are called a horizontal and vertical numbering, respectively. Both approaches are called *lexicographical* numbering. The numbering can be chosen freely, and many other numbering schemes exist.

Exercise 12. For a horizontal numbering, give the conversion formula from the double index (i, j) to a single index n for the interior nodes.

Exercise 13. For $N_x = 4$ and $N_y = 3$, give the discretisation matrix of the Helmholtz problem (16). Use a horizontal or vertical numbering.

4 The Nyquist theory on sampling

Sampling theory deals with the question how to represent continuous signals as discrete signals. This is related to the question on how to represent waves with approximations in discrete points. The Nyquist theory tells that you need to have at least two points in each wave period to represent a wave. Therefore, for a wave with wave length λ, the spacing between points should be less than $\lambda/2$. This is a lower limit and to obtain accurate results, in practice around 10 to 20 points per wave length are necessary in finite difference methods. Normally, the number of points per wave length is kept constant for different frequencies, not the total number of nodes.

Example 2. For a Helmholtz equation with a wave number of $k = 10$, we have a wave length of $\lambda = 2\pi/k = 0.2\pi$. If we want a finite difference method with 10 points per wave length, we need a spacing of $h = \lambda/10 = 0.02\pi$. If the domain is given by the interval $[0, \pi]$, this means that $\pi/h = 50$ nodes are used.

Exercise 14. How many nodes do we need if we have a Helmholtz problem on the interval $[0, \pi]$, $k = 20$ and 10 points per wavelength.

Exercise 15. We have a Helmholtz problem on the interval $[0, \pi]$ and keep 10 points per wavelength. For a problem with wave number k_1, we need n_1 nodes. Show that for a problem of wave number $k_2 = \alpha k_1$, where α is a positive constant, the number of nodes is $n_2 = \alpha n_1$.

Exercise 16. Let us consider a two-dimensional Helmholtz problem on the rectangle $[0, 1] \times [0, 1]$. For a problem with wave number k_1, we need n_1 nodes. Now, let us consider a problem with wave number $k_2 = \alpha k_1$, where α is a positive constant. How many nodes do we need in this case. Give your answer as a function of α and n_1.

Exercise 17. In the case of high-intensity focused ultrasound treatment of liver cancer, the frequency is usually kept at 1 MHz. Propagation of sound in water has a velocity of 1500 m/s. This means that the wave length is given by $\lambda = c/f = 1500/10^6 = 1.5 \cdot 10^{-3}$, so 1.5 mm. One rib is around 15 cm long and 1.2 cm thick. We use an equidistant grid with 10 nodes per wavelength. How many nodes are necessary for one rib?

References

[1] Strang, G. *Computational Science and Engineering.* Wellesley-Cambridge Press, 2007.

[2] Burden, R.L., J.D. Faires, and A.M Burden. *Numerical Analysis.* Cengage Learning, 2015.