INF5620 Second Project
Wave Equation - Finte elements

Nina Kristine Kylstad (`ninakky@student.matnat.uio.no`)
INF5620 - Numerical methods for partial differential equations

November 08, 2012

Abstract

This report investigates the one dimensional wave equation, and using finite element methods for solving a wave equation.

Contents

1 Mathematical problem

We consider the initial- and boundary-value problem given by

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

$$u(0,t) = U_0(t), \tag{2}$$

$$\frac{\partial u}{\partial x}(L) = 0, \tag{3}$$

$$u(x,0) = I(x), \tag{4}$$

$$\frac{\partial u}{\partial t}(x,0) = V(x), \tag{5}$$

where (1) is the 1D wave equation, (2) and (3) are boundary condition and (4) and (5) are initial conditions.

2 Numerical solution

2.1 Finite differences

We use finite differences to discretize the equation in time. We use a centered difference to approximate the time derivative:

$$\frac{\partial^2 u}{\partial t^2} \approx \frac{u^{n+1} - 2u^n + u^{n-1}}{\Delta t^2}$$

This gives (in compact form):

$$\left[D_t D_t u = c^2 u_{xx}\right]^n \tag{6}$$

Expanded and rearranged, this gives the scheme

$$u^{n+1} = 2u^n - u^{n-1} + \Delta t^2 c^2 u^n_{xx} \tag{7}$$

We then use the following approximations for $u(x)$:

$$u^n(x) \approx \sum_{j=0}^{N} c_j^n \phi_j(x), \tag{8}$$

$$u^{n-1}(x) \approx \sum_{j=0}^{N} c_j^{n-1} \phi_j(x), \tag{9}$$

$$u^{n+1}(x) \approx \sum_{j=0}^{N} c_j^{n+1} \phi_j(x), \tag{10}$$

2.2 Using a Galerkin method to derive variational form

We continue by using a Galerkin method with $v(x)$ as a test function, in order to obtain variational form:

$$\int_0^L u^{n+1} v dx = \int_0^L 2u^n v dx + \int_0^L -u^{n-1} v dx + \int_0^L \Delta t^2 c^2 u^n_{xx} v dx \tag{11}$$

We use integration by parts for the final term:

$$\int_0^L \Delta t^2 c^2 u^n_{xx} v dx = -\Delta t^2 c^2 \int_0^L u^n_x v_x dx + \Delta t^2 c^2 \left[u^n_x v\right]_0^L$$

$$\left[u^n_x v\right]_0^L = u^n_x(L) v(L) - u^n_x(0) v(0) = 0$$

This result comes from using the boundary conditions (2) and (3). Inserting this back into (11) gives

$$\int_0^L u^{n+1} v dx = 2 \int_0^L u^n v dx - \int_0^L u^{n-1} v dx - \Delta t^2 c^2 \int_0^L u^n_x v_x dx \tag{12}$$

2.3 Deriving a linear system

We now replace the u-terms in (12) with the approximations (8), (9) and (10). We also replace the test function v with $v = \phi_i(x)$. This gives

$$\sum_{j=0}^{N}\left(\int_0^L \phi_i\phi_j dx\right) c_j^{n+1} = 2\sum_{j=0}^{N}\left(\int_0^L \phi_i\phi_j dx\right) c_j^{n} - \sum_{j=0}^{N}\left(\int_0^L \phi_i\phi_j dx\right) c_j^{n-1} - \Delta t^2 c^2 \sum_{j=0}^{N}\left(\int_0^L \phi_i'\phi_j' dx\right) c_j^{n} \quad (13)$$

This gives the linear system

$$\mathbf{Mc^{n+1}} = 2\mathbf{Mc^{n}} - \mathbf{Mc^{n-1}} - C^2\mathbf{Kc^{n}} \quad (14)$$

where $C = \Delta tc$.

2.4 Interpreting as finite differences

We begin by looking at the matrices M and K. The entries in M are given by $M_{i,j} = \int_0^L \phi_i\phi_j$, and the entries in K are given by $K_{i,j} = \int_0^L \phi_i'\phi_j'$. Using P1 elements, we get the matrices

$$M = \frac{h}{6}\begin{bmatrix} 2 & 1 & 0 & \dots & 0 \\ 1 & 4 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & 4 & 1 \\ 0 & \dots & 0 & 1 & 2 \end{bmatrix} \quad (15)$$

and

$$K = \frac{1}{h}\begin{bmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \dots & 0 & -1 & 2 \end{bmatrix} \quad (16)$$

We now look at equation number i in the linear system, which becomes

$$\begin{aligned} \frac{h}{6}c_{i-1}^{n+1} + \frac{4h}{6}c_i^{n+1} + \frac{h}{6}c_{i+1}^{n+1} =& 2\left(\frac{h}{6}c_{i-1}^{n} + \frac{4h}{6}c_i^{n} + \frac{h}{6}c_{i+1}^{n}\right) \\ &- \left(\frac{h}{6}c_{i-1}^{n-1} + \frac{4h}{6}c_i^{n-1} + \frac{h}{6}c_{i+1}^{n-1}\right) \\ &- C^2\left(\frac{-1}{h}c_{i-1}^{n} + \frac{2}{h}c_i^{n} + \frac{-1}{h}c_{i+1}^{n}\right) \end{aligned}$$

We can replace c_i^n with u_i^n because the coefficients c_i are the values of u at spacial point i. h is the space between the nodes, also known as Δx. This gives

$$\begin{aligned}\frac{1}{6}u_{i-1}^{n+1}+\frac{4}{6}u_i^{n+1}+\frac{1}{6}u_{i+1}^{n+1} =& \frac{2}{6}u_{i-1}^n+\frac{8}{6}u_i^n+\frac{2}{6}u_{i+1}^n\\ &-\frac{1}{6}u_{i-1}^{n-1}-\frac{4}{6}u_i^{n-1}-\frac{1}{6}u_{i+1}^{n-1}\\ &+\frac{\Delta t^2c^2}{\Delta x^2}\left(u_{i-1}^n-2u_i^n+u_{i+1}^n\right)\end{aligned}$$

We move some terms around, and get

$$\begin{aligned}&\frac{u_i^{n+1}-2u_i^n+u_i^{n-1}}{\Delta t^2}+\frac{1}{6\Delta t^2}\left(u_{i-1}^{n+1}-2u_i^{n+1}+u_{i+1}^{n+1}-2u_{i-1}^n+4u_i^n-2u_{i+1}^n+u_{i-1}^{n-1}-2u_i^{n-1}+u_{i+1}^{n-1}\right)\\ &=c^2\frac{u_{i-1}^n-2u_i^n+u_{i+1}^n}{\Delta x^2}\end{aligned}$$

We see that the first term is the finite difference (FD) approximation for the 2nd derivative of u in time, $[D_tD_tu]_i^n$, and the last term is the FD approximation for the 2nd derivative of u in space, $[D_xD_xu]_i^n$. The term in the middle is a combination of the two, and can be written as $[D_tD_t(\frac{1}{6}\Delta x^2D_xD_xu)]_i^n$. (Note: We have to multiply by Δx^2 in the middle term to be able to write it as the approximation to the 2nd derivative in space.)

Combining the three terms, we get,

$$[D_tD_t(u+\frac{1}{6}\Delta x^2D_xD_xu)=c^2D_xD_xu]_i^n \tag{17}$$

2.5 Analysing the scheme in (17)

To perform an analysis of the scheme (17), we introduce a Fourier component, $u=e^{i(kx-\tilde{\omega}t)}$

We then get

$$\begin{aligned}[D_tD_tu]_p^n &= -\frac{4}{\Delta t^2}sin^2\left(\frac{\tilde{\omega}\Delta t}{2}\right)e^{-i\tilde{\omega}n\Delta t}e^{ikp\Delta x},\\ [D_xD_xu]_p^n &= -\frac{4}{\Delta x^2}sin^2\left(\frac{k\Delta x}{2}\right)e^{-i\tilde{\omega}n\Delta t}e^{ikp\Delta x},\\ [D_tD_t(\frac{1}{6}\Delta x^2D_xD_xu)]_p^n &= \frac{1}{6}\Delta x^2[D_tD_te^{-i\tilde{\omega}t}]^n[D_xD_xe^{ikx}]_p\\ &= \frac{8}{3\Delta t^2}e^{-i\tilde{\omega}n\Delta t}e^{ikp\Delta x}sin^2\left(\frac{\tilde{\omega}\Delta t}{2}\right)sin^2\left(\frac{k\Delta x}{2}\right)\end{aligned}$$

Putting this into the scheme (17), we get

$$\begin{aligned}&-\frac{4}{\Delta t^2}sin^2\left(\frac{\tilde{\omega}\Delta t}{2}\right)e^{-i\tilde{\omega}n\Delta t}e^{ikp\Delta x}+\frac{8}{3\Delta t^2}e^{-i\tilde{\omega}n\Delta t}e^{ikp\Delta x}sin^2\left(\frac{\tilde{\omega}\Delta t}{2}\right)sin^2\left(\frac{k\Delta x}{2}\right)\\ &=-\frac{4}{\Delta x^2}sin^2\left(\frac{k\Delta x}{2}\right)e^{-i\tilde{\omega}n\Delta t}e^{ikp\Delta x}\end{aligned}$$

After some simplification, we are left with

$$\begin{aligned}
sin^2\left(\frac{\tilde{\omega}\Delta t}{2}\right) &= C^2\frac{sin^2\left(\frac{k\Delta x}{2}\right)}{1-\frac{2}{3}sin^2\left(\frac{k\Delta x}{2}\right)} \\
sin\left(\frac{\tilde{\omega}\Delta t}{2}\right) &= C\frac{sin\left(\frac{k\Delta x}{2}\right)}{\sqrt{1-\frac{2}{3}sin^2\left(\frac{k\Delta x}{2}\right)}}
\end{aligned} \tag{18}$$

Since ω is real, we are also looking for a real $\tilde{\omega}$. This means that the left-hand-side of (18) must be in the interval $[-1, 1]$, and therefore som must the right-hand-side. The *sin* part of the right-hand-side only varies between -1 and 1. We therefore choose one of the extremes to find a stability limit for C. We choose $sin(k\Delta x/2) = 1$, which means that the left-hand-side must be smaller than, or equal to, 1 (the upper boundary of the left-hand-side).

$$\begin{aligned}
C\frac{1}{1-\frac{2}{3}} &\leq 1 \\
C &\leq \sqrt{1-\frac{2}{3}} = \sqrt{\frac{1}{3}} \\
\Rightarrow C &\leq \frac{1}{\sqrt{3}}
\end{aligned}$$

This obviously holds for the other boundary for the left-hand-side as well.

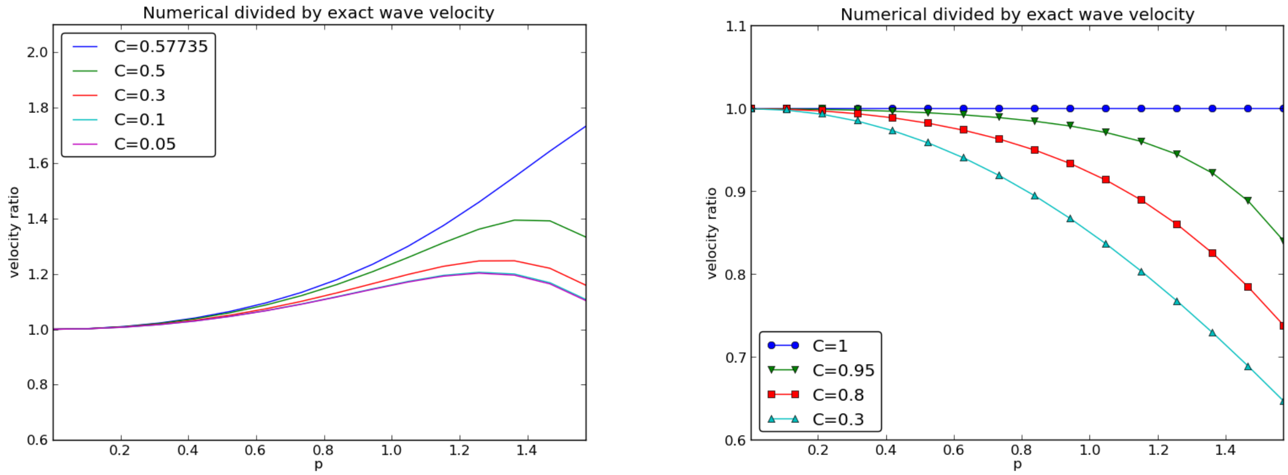
2.6 Numerical dispersion relation

We can now find the numerical dispersion relation, $\tilde{\omega}$ expressed by other parameters:

$$\begin{aligned}
sin\left(\frac{\tilde{\omega}\Delta t}{2}\right) &= C\frac{sin\left(\frac{k\Delta x}{2}\right)}{\sqrt{1-\frac{2}{3}sin^2\left(\frac{k\Delta x}{2}\right)}} \\
\frac{\tilde{\omega}\Delta t}{2} &= sin^{-1}\left(\frac{Csin\left(\frac{k\Delta x}{2}\right)}{\sqrt{1-\frac{2}{3}sin^2\left(\frac{k\Delta x}{2}\right)}}\right) \\
\Rightarrow \tilde{\omega} &= \frac{2}{\Delta t}sin^{-1}\left(\frac{Csin\left(\frac{k\Delta x}{2}\right)}{\sqrt{1-\frac{2}{3}sin^2\left(\frac{k\Delta x}{2}\right)}}\right)
\end{aligned}$$

The analytical dispersion relation is defined as $\omega = kc$. We can now compare ω and $\tilde{\omega}$. We do this by plotting $\tilde{c}/c$ as a function of $k\Delta x$ ($\tilde{c} = \tilde{\omega}/k, c = \omega/k$). We let $k\Delta x$ go from 0 to $\pi/2$. We the program `dispersion_relation_1D.py` to plot $\tilde{c}/c$ for various values of C $(=\frac{c\Delta t}{\Delta x})$. The resultant plot can be seen in figure 1a.

The plot in figure 1a looks a little strange, as we would expect that the ratio is closest to 1 for the value of C that is closest to the stability criterion. In figure 1a however, we see that the ratio is closer to 1 for smaller and smaller C.



(a) $\tilde{c}/c$ Finite Elements

(b) $\tilde{c}/c$ Finite Difference

Figure 1: Plots comparing the ratio $\tilde{c}/c$ for FEM and FD

In comparison, we can look at the corresponding plot that is made from solving (1) using finite elements and performing a similar analysis, which is shown in figure 1b. This plot shows the expected behavior.

The plot was generated using the program `dispersion_relation.py`, with

```
def r(C, p):
    return 1/(C*p)*asin(C*sin(p)/sqrt(1 - 2./3*(sin(p)**2)))
```

2.7 A diagonal M matrix

It can be very useful to have a diagonal M matrix, as this gives an explicit formula for the coefficients c_j^n, so a set of linear equations does not have to be solved at each time level. This decreases the amount of computing significantly. We can use the trapezoidal rule to produce a diagonal M matrix. Using the program `fe_approx_1D_numint.py`, and $M \in \mathbb{R}^{5\times 5}$, we get the matrix

$$M = \begin{pmatrix} \frac{1}{2}h & 0 & 0 & 0 & 0 \\ 0 & h & 0 & 0 & 0 \\ 0 & 0 & h & 0 & 0 \\ 0 & 0 & 0 & h & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2}h \end{pmatrix} \tag{19}$$

We can also find a diagonal matrix by multiplying the matrix M with $e = (1, 1, \ldots, 1)$. This gives a vector with elements that equal the sum of each row from M. In general,

$$\begin{bmatrix} a_{0,0} & a_{0,1} & a_{0,2} \\ a_{1,0} & a_{1,1} & a_{1,2} \\ a_{2,0} & a_{2,1} & a_{2,2} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{0,0} + a_{0,1} + a_{0,2} \\ a_{1,0} + a_{1,1} + a_{1,2} \\ a_{2,0} + a_{2,1} + a_{2,2} \end{bmatrix}$$

We then use the matrix created by $diag(w)$, which is the matrix with the elements of w on the diagonal. For $M \in \mathbb{R}^{5\times 5}$, we get the matrix

$$Me = \frac{h}{6}\begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 4 & 1 & 0 & 0 \\ 0 & 1 & 4 & 1 & 0 \\ 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{h}{6}(2+1) \\ \frac{h}{6}(1+4+1) \\ \frac{h}{6}(1+4+1) \\ \frac{h}{6}(1+4+1) \\ \frac{h}{6}(2+1) \end{bmatrix} = \begin{bmatrix} \frac{1}{2}h \\ h \\ h \\ h \\ \frac{1}{2}h \end{bmatrix}$$

$$diag(Me) = \begin{bmatrix} \frac{1}{2}h & 0 & 0 & 0 & 0 \\ 0 & h & 0 & 0 & 0 \\ 0 & 0 & h & 0 & 0 \\ 0 & 0 & 0 & h & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2}h \end{bmatrix}$$

We see that this gives the same result as using the trapezoidal method. By definition (see (2.7), this is also the same as replacing each row in the element matrices associated with M by the row sum on the diagonal.