COMPUTATIONAL MODELLING SAMMARY 1

BASED ON COMPUTATIONAL MODELLING BY S J HULSHOFF

SAM VAN ELSLOO

APRIL - JUNE 2017

ナマケモノは愛

Preface - 2018-2019 edition

I included examples from the maple TA quiz and the other document with examples that's available on blackboard. Other than that, there's not a whole lot to practice with, unfortunately.

Furthermore, the summary is rather long, but this is simply because we have to study quite a lot, and a lot of it is also rather complicated (so I could have made it shorter, but then you wouldn't have understood it). Furthermore, if you really want to, you skip sections 2.5.4 and 3.1.3. Also, in the third chapter, there are lot of very long matrices which take up a lot of space, so there's that.

Also, just in case you're wondering:

- Yes, I designed this front page myself, how else do you think it could have been so beautiful?
- Yes, the characters on the right hand side *are* supposed to mean something. Each front page has a different saying written on it, in fact (no I didn't secretly put any inflammatory texts on it, don't worry).
- No, I do not know any Japanese myself so don't ask me how to pronounce the characters cause genuinely, I'm clueless.

Finally, don't be a cheap-ass when printing, but please print the front page as well when you print the summary.

Contents

1 Introduction

In this course, computational modelling is the representation of complex physical systems by the numerical solution of partial differential equations. This, for example, entails numerically finding the pressure distribution across a wing, based on the Navier-Stokes equations, given that we know the wing geometry and the boundary conditions (e.g. the freestream conditions).

1.1 The computational modelling process

The complete computational modelling process is shown in figure 1.1.

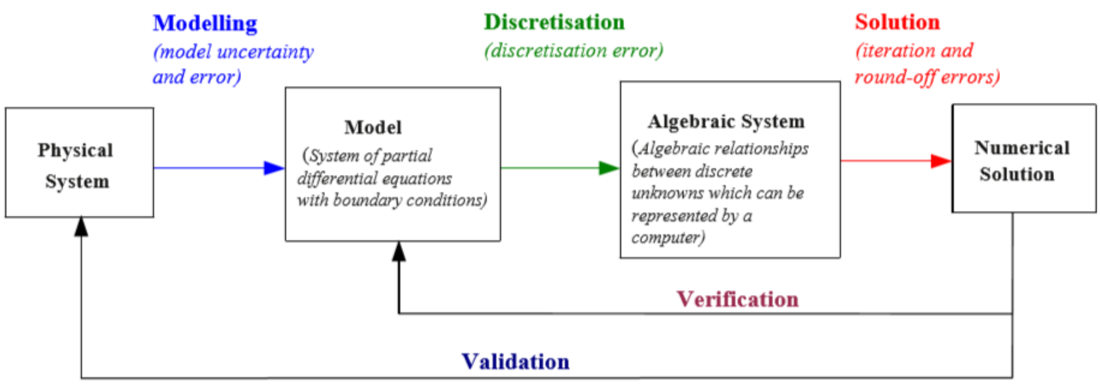


Figure 1.1: The computational modelling process.

As shown, it consists of multiple steps:

1. **Modelling**: we represent the real physical system using one or more PDEs and boundary conditions (i.e., this step involves finding the correct equations to use in your solution, and what boundary conditions need to be fulfilled). This introduces both uncertainties and model errors (more on those later). Uncertainties arise if there are no PDEs known which sufficiently describe the process under consideration, or if it is not possible to consider all deviations in problem input parameters. Model errors, on the other hand, result from using approximate (simplified) governing equations. Model errors also arise when using approximate boundary conditions, either to limit complexity, or to limit the extent of the physical domain which is considered.
2. **Discretisation**: now, the previously chosen PDEs and their boundary conditions describe a continuous solution for all possible points in the space/time domain of interest. However, for numerical methods, this domain will be discretised, meaning it will be replaced with a finite number of nodes, also called a **mesh**. The values at the nodes are called **unknowns** or **degrees of freedom**, and an **algebraic system** is derived which approximates the action of the PDEs and the boundary conditions. Discretisation clearly introduces **discretisation errors**, which arise e.g. from the inability of a mesh to represent the exact geometry (if have a finite number of nodes, it can be difficult to describe complex shapes) or from insufficient refinement in regions where the physical solution has large gradients (if you have very few nodes, it makes sense that it becomes inaccurate).
3. **Solution**: many problems of interest have non-linear governing equations or boundary conditions, which lead to non-linear systems of algebraic equations. These need to be solved by iterative methods (e.g. Newton's method). This introduces **nonlinear iteration errors**. However, iterative methods for nonlinear

problems typically make use of a local linearisation of the nonlinear algebraic system, leading to a linear system of equations to be solved each iteration. In principle, this linear system of equations can be solved using matrix algebra you already know. However, there are much more efficient methods of solving large systems of equations quickly[1], which basically come down to making a guess for the solution, and then using the linear algebraic system to estimate a correction. The error that remains due to the halting of the iteration process is called **linear iteration error**. To be clear: it thus often happens that you have both an outer non-linear iteration error and an inner linear iteration error. The original nonlinear problem is solved by making an initial guess for the solution, then estimating corrections for it (which introduces nonlinear iteration errors). However, these corrections are estimated using linearised problems, where again the corrections are estimated using an initial guess and then estimating corrections for the corrections, using an iterative method. Finally, note that this step also includes round-off errors.

4. **Verification** and **validation**: verification is the procedure where the consistency of the method is confirmed (e.g. by comparing it with an exact solution), and the magnitude of the errors is estimated. In validation, you actually check how close the numerical solution comes to the actual solution (the one that results from doing the experiment). Note that it is quintessential to do validation after verification, otherwise you don't know which errors are caused by what.

1.2 Uncertainties and errors

To recap:

UNCERTAINTIES

Uncertainties are deficiencies which can be attributed to lack of information about the system:
1. **Epistemic (systematic) uncertainties**: those associated with deviations of the model from the real system due to things you didn't expect to happen because you were either stupid or dumb.
2. **Aleatory (statistical) uncertainties**: those which are impractical to measure and therefore essentially irreducible. Examples are material imperfections.

ERRORS

Errors are deficiencies in either the modelling or numerical solution procedures which are not due to lack of knowledge:
1. Model errors: those arising from the use of approximate governing equations to describe the system
2. Discretisation errors: those associated with representing the infinitedimensional solution of the governing equations with a finite number of algebraic relations
3. Iteration error: those due to the termination of iterative procedures used to determine the numerical solution
4. Round-off errors: those arising from the finite precision with which computers carry out arithmetic.

Quiz 1: Q1

When performing structural computations, not being able to include variations in the microscale of a material leads to a deficiency best described as a
- Model error
- Discretisation error
- Aleatory uncertainty
- Epistemic uncertainty

Correct is **Aleatory uncertainty**. In case you're interested, aleatory means dependent on chance, luck or uncertain outcome.

[1]Honestly, just imagine how long it takes to row reduce to echelon form a $10^6 \times 10^6$ matrix. Taking the inverse is an even worse idea than that, and Cramer's rule is also pretty bad (if you even remember what Cramer's rule was).

Quiz 1: Q2

Errors obtained by representing a continuous physical system using a finite number of unkown values are known as:

- Model errors
- Discretisation errors
- Iteration errors
- Round-off errors

Correct is **Discretisation errors**. Pretty much textbook definition this.

Quiz 1: Q3

During verification one estimates:

- solution errors, provided model uncertainties are small.
- discretisation errors, provided model errors are small.
- discretisation errors, provided solution errors are small.
- model errors, provided discretisation errors are small.
- model errors, provided solution errors are small
- solution errors, provided model errors are small.

Correct is **discretisation errors, provided solution errors are small.** Assuming your method is correct (i.e. solution errors are small), you estimate the discretisation errors by comparing it with an exact solution (for example). Model errors are found during validation.

1.3 *What do we require from our model?*

WELL-POSED PROBLEM

A problem is **well-posed** if its solution exists, is unique, and depends continuously on the input data.

What do we mean by this? Well, that a solution exists is kind of a logical requirement. With 'depends continuously on the input data' means that if you change the input parameters slightly, your solution doesn't suddenly change wildly. Furthermore, the solution needs to be unique: this is typically a problem for inverse problems (where you have your final data and then want to reconstruct the initial data). For example, if you have a certain mixture, then it's easy to imagine that there are multiple solutions that describe how this mixture came to be, meaning the solution is not unique.

1.4 *What do we require from our discretisation?*

As explained before, discretisation is the process of replacing your domain of interest by a finite number of nodes and finding algebraic systems for them. What do we require from these algebraic systems?

CONVERGENE

A discretisation is **convergent** if its numerical solution approaches the exact model solution as its numbers of degrees of freedom is increased (i.e. you add more nodes to your mesh).

Above definition makes a lot of sense. We break convergence usually up in two parts:

CONSISTENCE AND STABILITY

- A discretisation is **consistent** if when substituting an exact solution into the discrete equations, the only terms which remain are those which tend to zero as the number of degrees of freedom is increased.
- A discretisation is **stable** if the numerical solution for a given number of degrees of freedom is unique, and small changes to the input data produce only small changes in the numerical solution.

Now, what is exactly meant with the one for consistency? Well, think back to Applied Numerical Analysis: the forward Euler-Cauchy scheme,

$$y_{n+1} = y_n + hf\left(y_n\right)$$

was consistent, whereas

$$y_{n+1} = 2y_n + hf\left(y_n\right)$$

isn't. Why, you wonder? Well, let's see what happens if we Taylor expand:

$$\begin{aligned} y_{n+1} &= y_n + hy_n' + \frac{h^2}{2}y_n' + ... \\ y_n &= y_n \\ f\left(y_n\right) &= y_n' \end{aligned}$$

If we'd substitute this into the first one, the forward Euler-Cauchy scheme, we get:

$$\begin{aligned} y_n + hy_n' + \frac{h^2}{2}y_n' + ... &= y_n + hy_n' \\ \frac{h^2}{2}y_n' + ... &= 0 \end{aligned}$$

So: if we substitute the exact solution into the discrete equations, the only terms that remain are those that tend to zero as the number of degrees of freedom is increased (which reduces h). Therefore, forward Euler-Cauchy is indeed consistent. On the other hand, if you plug it into the second one, we'd get

$$\begin{aligned} y_n + hy_n' + \frac{h^2}{2}y_n' + ... &= 2y_n + hy_n' \\ -y_n + \frac{h^2}{2}y_n' + ... &= 0 \end{aligned}$$

So: if substitute the exact solution into the discrete equations, there are terms that remain that do not tend to zero as the number of degrees of freedom is increased, thus this method is inconsistent.

Furthermore, how can a discretisation be consistent but not convergent? Again, think back to applied numerical analysis: we saw in the last module that although forward Euler-Cauchy method was consistent, you were only allowed to pick certain values for h depending on the considered ODE, otherwise the method was unstable, which meant that it was unable to converge. Indeed, we have

LAX EQUIVALENCE THEOREM

For consistent discretisations of well-posed linear problems, stability implies convergence.

Quiz 1: Q4

A numerical method developer substitutes an exact solution into the algebraic system arising from a discretisation, and finds a remainder vector. The developer repeats this exercise for several numbers of degrees of freedom and finds the magnitude of the values in the remainder vector decrease with increasing numbers of degrees of freedom. The developer has verified:

- Consistency
- Convergence
- Stability

Correct is **Consistency**. This is basically just checking whether the error goes to zero when $h \to 0$ as done in the above example of the forward-Euler scheme. The developer doesn't check anything for stability (which would involve checking a slightly disturbed solution), though, so you can't say it's convergent.

Quiz 1: Q5

Which of the below is the most logical statement

- A stable discretisation is consistent
- A stable discretisation is convergent
- A convergent discretisation is stable
- A consistent discretisation is stable

Correct is **A convergent discretisation is stable**. A convergent discretisation is both stable and consistent, but a consistent discretisation is not necessarily stable or convergent, and a stable discretisation is not necessarily consistent or convergent either.

Quiz 1: Q6

Verification involves

- Estimating model and discretisation errors
- Estimating model errors
- Estimating model errors and confirming consistency
- Estimating discretisation errors and confirming consistency

Correct is **Estimating discretisation errors and confirming consistency**. The fact that validation and not verification estimates model errors kinda gives it away.

1.5 *What do we require from our solution procedure?*

We like our solution procedure to produce numerical solutions with minimum computational effort. There are fields where this is more important then others, e.g. for the ADCS of your spacecraft it's important to do your computations quickly for obvious reasons, in F1 you are only entitled to a certain number of computer computations per month (so you shouldn't be wasting many computations on solving simple equations etc.), and in high-frequency trading[2], they actually place computers closer to the stock market so that the wires can be shorter and the data does not have to travel more distance than required to save some time, so it's kinda nice if your model is also able to do calculations quickly. Obviously you don't have to know this for the quiz but I thought it was interesting to know.

1.6 *Final remarks*

You will find out in this course that computational modelling is interesting and fun (apparently).

[2]Trading stocks on the stock market at very high frequencies (you already sell it miniseconds (even microseconds sometimes) after you bought it).

2 Modelling with PDEs

Unfortunately, this chapter is rather theoretical,

2.1 Why PDEs?

Because they are appear quite often in physics and engineering. In case you forgot what a partial differential equation was: an ordinary differential equation was an equation where the function u was only dependent on one variable, e.g. t, so that for example

$$\frac{\mathrm{d}u(t)}{\mathrm{d}t} + 2u(t) = 4t$$

where $u(t)$ is the function we're interested in in the end. Partial differential equations were equations where u dependent on multiple variables, e.g. x and t:

$$\frac{\partial u(x,t)}{\partial x} + 2\frac{\partial u(x,t)}{\partial t} = 4e^{xt}$$

It is not difficult to think of stuff that requires at least a 2D analysis: not only stuff like plate bending or buckling, flow around an airfoil etc. is clearly 2D (as it happens in the 2D plane); even when the spatial domain is effectively 1D[1], we are often interested in processes which evolve in time direction (e.g. how the temperature distribution changes over time).

2.2 Common PDEs in Science and Engineering

2.2.1 The linear diffusion equation

The linear diffusion equation is shown in figure 2.1, and the associated equation is

The **linear diffusion equation** is given by

$$\frac{\partial u}{\partial t} = v\frac{\partial^2 u}{\partial x^2} \tag{2.1}$$

LINEAR DIFFUSION EQUATION

This is actually the heat equation that you already came across during differential equations. It describes the progression of the temperature distribution in a bar of length L over time. This is depicted in figure 2.1 as well: we start with an initial temperature distribution where the middle of the bar has been heated up significantly, and we see that indeed the temperature sort of spreads to the sides[2].

Note that figure 2.1 is simply just an example of a solution to the linear diffusion equation. The exact solution for the problem you're considering depends on an initial condition (at $t = 0$) and two boundary conditions (at $x = 0$ and $x = 1$).

2.2.2 The linear advection equation

The linear advection equation describes the unattenuated propagation of information, as shown in figure 2.2. The associated equation is

[1]E.g. the temperature distribution in 1D: this is essentially only 1D, as we'd probably neglect the thickness of the bar.

[2]In case you don't really get how this graph should be interpreted: we start at a time $t = 0$, at which the temperature distribution is given by that sharp peak in the middle. As time progresses, the temperature becomes as it is shown, further 'back' in the graph.

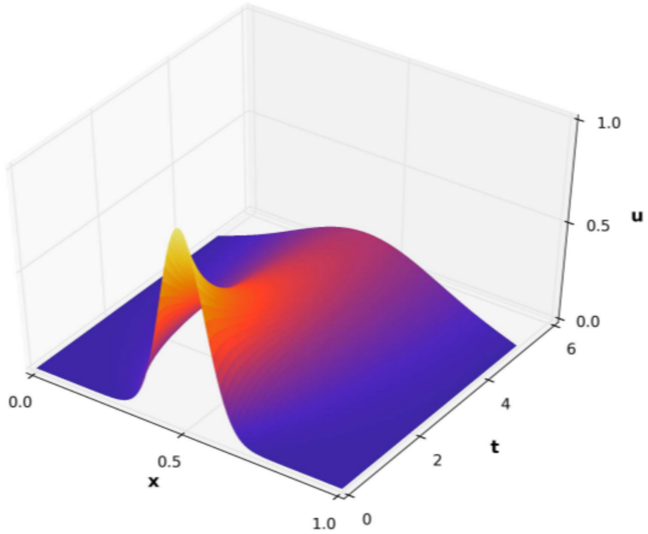


Figure 2.1: A solution for the linear diffusion (heat) equation. Boundary conditions are applied at $x = 0$ and $x = 1$, namely $u(0,t) = u(1,t) = 0$; additionally, an initial condition is applied at $t = 0$, of the form $u(x,0) = F(x)$.

LINEAR ADVECTION EQUATION

The **linear advection equation** is given by

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0 \tag{2.2}$$

What is c exactly? It's the velocity at which the wave travels to the right[3]. How do we know that c is the velocity at which the wave travels? Well, you can rather easily prove that a valid solution to this problem would be, given a lack of boundary and initial conditions:

$$u(x,t) = ct - x$$

because then

$$\frac{\partial u}{\partial t} = c, \qquad \frac{\partial u}{\partial x} = -1, \qquad \frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = c + c\cdot -1 = c - c = 0$$

so indeed the partial differential equation is satisfied. Now, if we have $u = ct - x$, if we then look at lines where u is constant (which is for example the case if we follow the peak of the wave), then we should have

$$ct - x = \text{constant}$$

or

$$c\,dt - dx = 0$$

and thus

$$\frac{dx}{dt} = c$$

So, each point on the wave moves to the right with a velocity c.

Terms similar to those in the advective equations are common in PDEs for fluids, where they represent the transport of quantities with the fluid velocity. They can also be used to represent the transmission of waves through gases and solids.

[3] Again, if you don't fully understand how the graph works: initially, at $t = 0$, the vertical displacement of the wave is as shown at the front of the graph (with its peak at $x \approx 0.3$). However, as time progresses, we move to the back of the graph; e.g. t $t = 0.4$, the peak has shifted to $x \approx 0.7$, so you kinda see the future in the back.

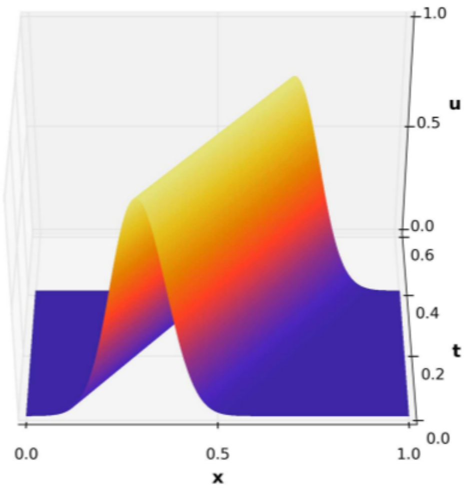


Figure 2.2: A solution for the linear advection equation. A boundary condition is applied at $x = 0$, namely $u(0,t) = 0$. Note that the boundary $x = 1$ is not specified (we'll see at the end of this chapter why). Furthermore, an initial condition at $t = 0$ is applied, namely $u(x, 0) = F(x)$.

2.2.3 *The linear advection-diffusion equation*

The linear advection-diffusion equation is simply a combination of above two equations:

LINEAR ADVECTION-DIFFUSION EQUATION

The **linear advection-diffusion equation** is given by

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = \upsilon\frac{\partial^2 u}{\partial x^2} \tag{2.3}$$

which produces solutions looking like figure 2.3; it describes the simultaneous transport and spreading out of the initial condition. The linear advection-diffusion equation is typically used as a testbed for the development of numerical methods for more complex PDEs.

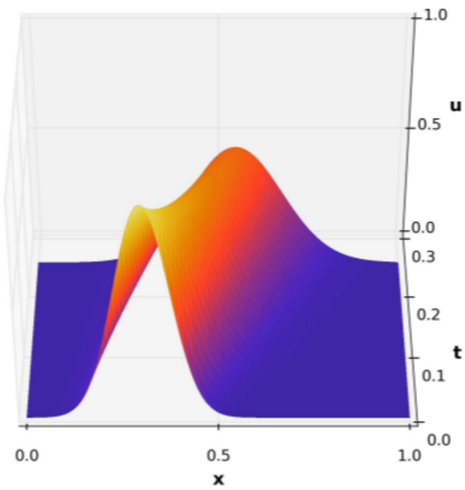


Figure 2.3: A solution for the linear advection-diffusion equation.

2.2.4 *The second-order wave equation*

SECOND-ORDER WAVE EQUATION

The **second-order wave equation** is given by

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \tag{2.4}$$

This describes advection in two directions, as shown in figure 2.4. Note that it is essentially the summation of two linear advection equations.

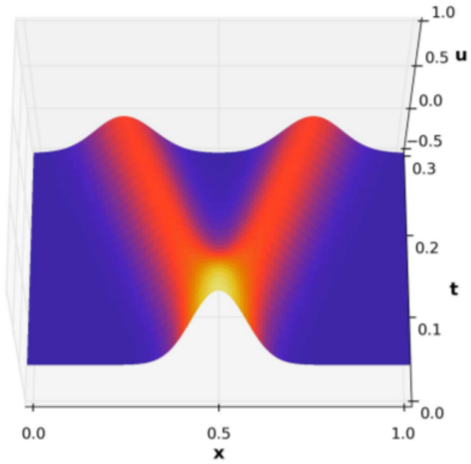


Figure 2.4: A solution for the second-order wave equation.

The second-order wave equation is common to describe electromagnetism, supersonic flow and elastic wave propagation.

2.2.5 *The non-linear advection (Burgers) equation*

NON-LINEAR ADVECTION (BURGERS) EQUATION

The **non-linear advection (Burgers) equation** is given by

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0 \tag{2.5}$$

A solution for this (with a pulse initial condition) is shown in figure 2.5. Note that this equation is pretty similar to the linear advection equation (which was simply wave travelling to the right at velocity c). However, this time, the speed is varying: the speed is equal to u. This means that in figure 2.5, points with large values of u (i.e. yellow parts) move to the right at a faster pace than points with small values of u (i.e. blue regions); this means that the peak catches up with it, and this is known as wave steepening, and is part of the process of shock formation.

The Burgers equation is used to study traffic flow (i.e. literally traffic, with bikes, cars and buses and all that), and as prototypical conservation law in theoretical and computational gasdynamics.

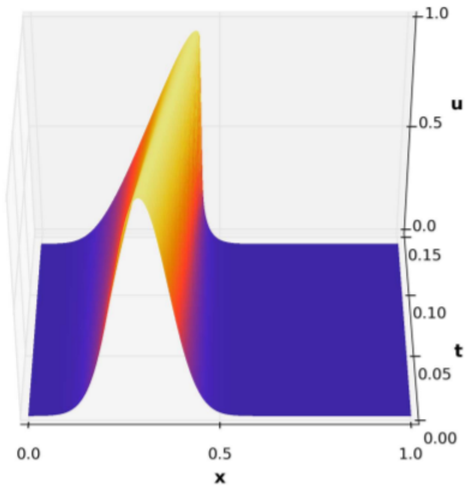


Figure 2.5: A solution for the non-linear advection (Burgers) equation.

2.2.6 The Laplace equation

LAPLACE'S EQUATION

The **Laplace equation** is given by

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \nabla^2 u = 0 \tag{2.6}$$

A solution is shown in figure 2.6. Note that Laplace acts in x and y rather than t and x.

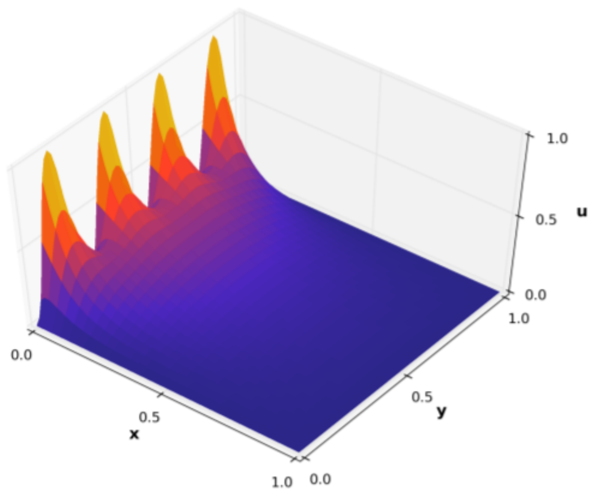


Figure 2.6: A solution for Laplace's equation with $\sin^2(y)$ boundary condition.

2.2.7 The Poisson equation

THE POISSON EQUATION

The **Poisson equation** is given by

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = S(x, y) \tag{2.7}$$

So it's simply Laplace with a source term, $S(x, y)$. A solution is shown in figure 2.7, with $S(x, y) = -1$. Here the values of u have been set to 0 on the boundaries of the domain. Note that this implies that the curvature of the solution is negative when proceeding radially, as is visible. On the other hand, if it'd have been $S(x, y) = 1$, you'd have gotten a nice smooth valley.

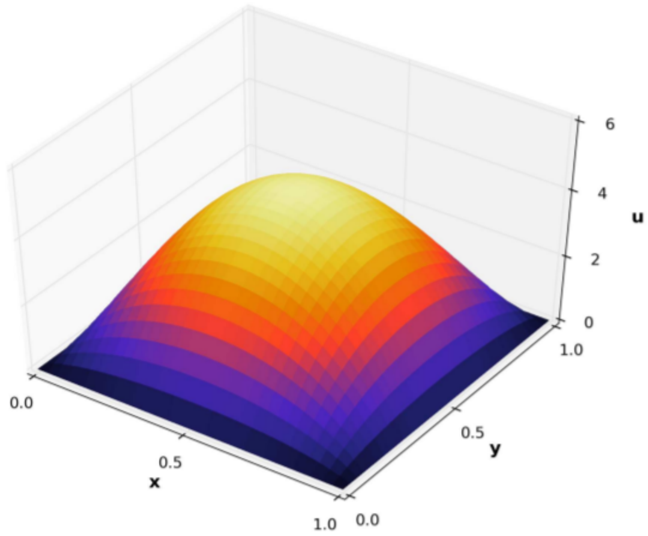


Figure 2.7: A solution for the Poisson equation.

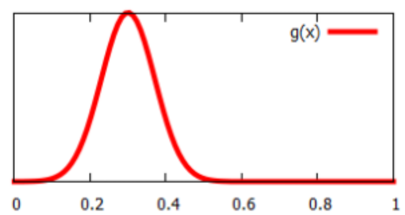
The Poisson equation $\nabla^2 u = -\rho$ appears in electrostatics, where u is the electric potential and ρ is the charge distribution. It is also used to model gravitational potential in astrodynamics.

Quiz 1: Q7

Consider the solution to the following problem:

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0$$

on $0 < x < 1$ with $u(0, t) = 0$ and $c = c(x) = x$.



If the initial condition $u(x, 0) = g(x)$ is in the form of the peak shown in the figure above, the following solution behaviour should be observed:

- The peak will not move but its magnitude will decrease in time
- The peak will split into to left-going and right-going components.
- The peak will contract as it convects to the right
- The peak will expand as it convects to the right
- The peak will contract as it convects to the left

- The peak will not move but its magnitude will increase in time
- The peak will expand as it convects to the left

Correct is **The peak will expand as it convects to the right**. Remember that this is the linear advection equation, as shown in figure 2.2, with c being the velocity at which it moves to the right. In this PDE, c is a function of x; $c = x$ to be precise. This means that the right part of the wave, where x is comparatively large, moves to the right faster than the left part of the wave, where x (and thus c) is comparatively small. Thus, the wave moves to the right (as c is positive for all values of x (as $0 < x < 1$), and expands while doing so.

2.3 *From PDEs to a computational model*

Suppose we are modelling the flow around a wing. Of course, we could, theoretically speaking, include the entire universe in our calculations, but it's obvious that this is ludicrous. Instead, we focus on a smaller numerical domain as shown in figure 2.8, where we have also shown the numerical domain for a wind tunnel flow.

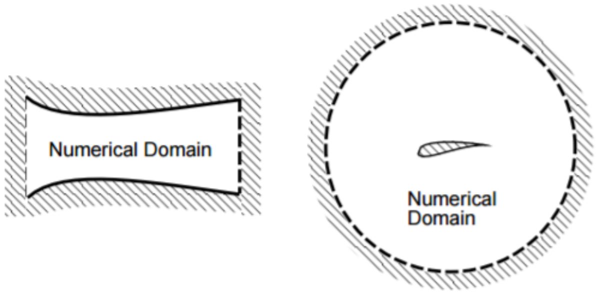


Figure 2.8: Numerical domains for channel flow (left) and an airfoil problem (right) with true boundaries (solid lines) and artificial boundaries (dashed lines).

From figure 2.8, it is clear there are two types of boundaries on our numerical domain:

- **True boundaries**: these boundaries that are actual, physical boundaries as well. In figure 2.8, this corresponds to the walls of the wind tunnel, and the surface of the airfoil. These are physically present in the real world, and thus are true boundaries.
- **Artificial boundaries**: these boundaries are boundaries imposed by humans, that don't truly exist but are introduced to reduce the size of the numerical domain to an acceptable size.

The values that are specified on true boundaries usually follow easily; for example, for an airfoil, a straightforward boundary condition would be that the flow is tangent at the surface of the airfoil (and for the wind tunnel, that the flow is tangent to the walls). On the other hand, artificial boundaries are more arbitrary. Looking at the airfoil of figure 2.8, one option would be to place the artificial boundary far away enough from the airfoil so that the perturbations are small at the boundary (meaning that the conditions at the artificial boundary are equal to the freestream conditions). However, this can lead to a very large domain, which is expensive to compute. Instead, we could (as an example) place the artificial boundary closer to the airfoil, and estimate the perturbations on the artificial boundary using the velocities induced by a vortex with a circulation equivalent to that determined for the airfoil by the computation. This can significantly reduce the size of the domain, reducing the cost of the computations.

However, do note that since the model for the processes occuring outside the computational domain will be an approximation, introducing an artificial boundary introduces an additional source of model error.

Now, to close off this section, consider again the case of a subsonic flow around an airfoil: what you have to realize is, no matter how far away you place your outer boundary, the flow at the boundary *will* be disturbed at the boundary (although it'll only be very slightly disturbed, it will be disturbed nonetheless). This can be problematic, because this means you can't simply set the boundary conditions equal to the freestream conditions. There are two ways of dealing with this problem:

- You directly change the boundary conditions, which is for example what you'd do when you'd estimate the induced velocities caused by the vortex as I described above. This is called a **physical condition**. A precise mathematical definition will follow in the next section.
- You change the PDEs themselves which you use in your system. Don't worry about how you'd go about altering them, but realize that of course, you can change your PDEs near the boundaries to take this into account. This is called a **numerical condition**. A precise mathematical definition will follow in the next section.

Do note: it is absolutely not always the case that your boundary conditions are complicated by disturbances: for example, if you have the subsonic flow over the airfoil, than the flow-tangency condition does can straightforwardly be applied. Also, if we'd have a supersonic flow over an airfoil, then the free-stream conditions are, in fact, equal to the free-stream conditions: disturbances in a supersonic flow don't propagate upstream, so even if you'd place a Christmas tree in your flow, the flow in front of the body would be exactly the undisturbed freestream flow.

Quiz 1: Q8

Moving an artificial boundary closer to a region of interest is likely to increase
- iteration error
- stability
- consistency
- convergence
- discretisation error
- model error

Correct is **model error**. Moving an artificial boundary closer to a region of interest simply makes your model less accurate. It does not have anything to do with any of the other stuff listed.

2.4 *The classification of PDEs*

2.4.1 *Nomenclature for a general problem*

As mentoned, it is normally desired to simulate some phenomena over a limited region of space. We will refer to this region as the problem domain, Ω, with boundary $\partial\Omega$, as shown in figure 2.9.

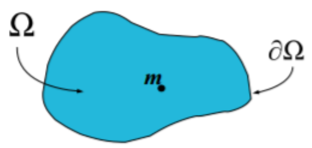


Figure 2.9: Notation for the computational domain.

Then, the solution of a PDE at all points m in Ω might be determined by a problem of the form:

$$\begin{aligned} \mathcal{L}(u) &= f \quad \forall m \in \Omega \\ u &= g \quad \forall m \in \partial\Omega \end{aligned}$$

In case you forgot, $\forall$ means 'for all'. In this case:

- $\mathcal{L}(u)$ is the differential operator of the PDE. For example, looking at the wave equation,

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

then the differential operator would be

$$\mathcal{L} = \frac{\partial^2}{\partial t^2} - c\frac{\partial^2}{\partial x^2}$$

Similarly, if we have the Laplace equation $\nabla^2 u = 0$, then $\mathcal{L} = \nabla^2$.

- f is a (possibly variable in space and time) source time. For example, in the Poisson equation, we literally had a source term $S(x, y)$ there.
- g is a (possibly variable in space and time) boundary value, i.e. the collection of boundary and initial conditions.

Now, there are three types of boundary conditions possible:

- **Dirichlet conditions** specify u directly on the boundary, i.e. $u = g$.
- **Neumann conditions** specify the derivative (or more accurately, gradient) of u on the boundary, i.e. $\frac{\partial u}{\partial n} = q$.
- **Robin conditions** mix above conditions, i.e. $\frac{\partial u}{\partial n} + u = r$.

Finally, the numerical conditions can be written as $\mathcal{L}(u) = f$. Note: in section 3.1.5, we'll better see what exactly the difference is between physical and numerical conditions.

2.4.2 *PDE order and homogeneity*

The order of a PDE is simply equal to the order of the term within $\mathcal{L}(u)$ with the highest order derivative. Homogeneous PDE are those for which $\mathcal{L}(u) = 0$.

2.4.3 *Linear, quasi-linear, and non-linear PDEs*

Linear A linear PDE is one for which the following is true:

$$\mathcal{L}\left(u_1 + u_2\right) = \mathcal{L}\left(u_1\right) + \mathcal{L}\left(u_2\right)$$

For example, Laplace equation satisfies this:

$$\mathcal{L}\left(u_1 + u_2\right) = \frac{\partial^2\left(u_1 + u_2\right)}{\partial x^2} + \frac{\partial^2\left(u_1 + u_2\right)}{\partial y^2} = \frac{\partial^2 u_1}{\partial x^2} + \frac{\partial^2 u_2}{\partial x^2} + \frac{\partial^2 u_1}{\partial y^2} + \frac{\partial^2 u_2}{\partial y^2}$$

which is indeed equal to

$$\mathcal{L}\left(u_1\right) + \mathcal{L}\left(u_2\right) = \frac{\partial^2 u_1}{\partial x^2} + \frac{\partial^2 u_1}{\partial y^2} + \frac{\partial^2 u_2}{\partial x^2} + \frac{\partial^2 u_2}{\partial y^2}$$

and so Laplace is linear. However, the non-linear advection equation is indeed non-linear:

$$\mathcal{L}\left(u\right) = \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0$$

so that

$$\mathcal{L}\left(u_1 + u_2\right) \quad = \quad \frac{\partial\left(u_1 + u_2\right)}{\partial t} + \left(u_1 + u_2\right)\frac{\partial\left(u_1 + u_2\right)}{\partial x} = \frac{\partial u_1}{\partial t} + \frac{\partial u_2}{\partial t} + \left(u_1 + u_2\right)\left(\frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial x}\right)$$

which is different from

$$\mathcal{L}\left(u_1\right) + \mathcal{L}\left(u_2\right) = \frac{\partial u_1}{\partial t} + u_1\frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial t} + u_2\frac{\partial u_2}{\partial x}$$

(you can't make those equal by getting rids of the brackets in the first equation either). Therefore, this equation is not linear.

Linear PDEs have great advantages, as solutions could simply be superimposed; we exploited this in aero I if you remember. You'll probably recall figure 2.10: we had some basic solutions which satisfied the Laplace equation[4] and since they were individually valid solutions, the summation of them was still a valid solution (since $\mathcal{L}\left(u_1+u_2\right)=\mathcal{L}\left(u_1\right)+\mathcal{L}\left(u_2\right)$ for Laplace). The idea was then to pick points on the boundary of the airfoil, and place sources at each of those boundary points; at each of these points, the flow must be tangent to the airfoil. This results in as many equations as unknowns (the unknowns being the strength of each source), and these equations are all linear, meaning it's just a matter of matrix algebra to solve the resultant system.

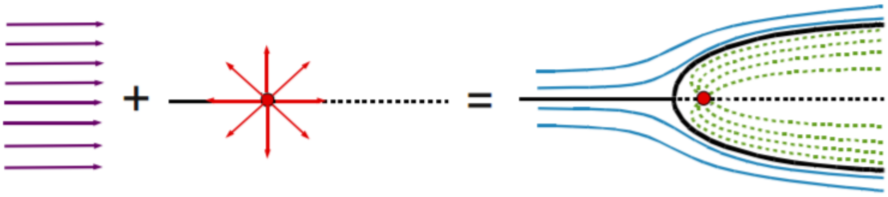


Figure 2.10: Potential flow solution for a Rankine half body obtained by the superposition of a freestream flow with a source.

This method, where you use the boundary points to determine your solution, is called **boundary-element method**. It has advantages over finite-element methods, where you analyse the points in the domain itself: you typically need significantly less boundary points for BEM than you'd need points in the domain space when using FEM. This is obviously beneficial[5].

Quasi-linear A quasi-linear PDE is one which does not have non-linearities occuring due to the highest-order derivatives and cross derivatives of its unknowns. A simple approach to determining if a PDE is quasi-linear is:

1. Underline the highest-order derivatives in all directions (i.e. the highest derivative with respect to x; the highest derivative with respect to y and the highest derivative with respect to z). Furthermore, underline all cross-derivatives.
2. Replace all other stuff by constants.
3. If the remaining operator is linear, the original operator is quasi-linear.

Let's do two examples to show exactly what I mean: suppose we have

$$\mathcal{L}(u)=x^2\frac{\partial^2 u}{\partial x\partial y}+\ln(u)\frac{\partial^2 u}{\partial x^2}+\left(\frac{\partial u}{\partial x}\right)^2\frac{\partial^2 u}{\partial y^2}$$

Then, the highest-order derivatives are shown in neonpink:

$$\mathcal{L}(u)=x^2\frac{\partial^2 u}{\partial x\partial y}+\ln(u)\frac{\partial^2 u}{\partial x^2}+\left(\frac{\partial u}{\partial x}\right)^2\frac{\partial^2 u}{\partial y^2}$$

So, substituting with constants, we get

$$\mathcal{L}(u)=a\frac{\partial^2 u}{\partial x\partial y}+b\frac{\partial^2 u}{\partial x^2}+c\frac{\partial^2 u}{\partial y^2}$$

which is clearly linear. On the other hand if we have

$$\mathcal{L}(u)=\left(\frac{\partial u}{\partial x}\right)^2+\frac{\partial^2 u}{\partial y^2}$$

[4]The formula for the potential of a uniform flow (on the left) for example equalled

$$u=V_\infty x$$

and the solution for a source equalled

$$u=\frac{\Lambda}{2\pi}\ln\sqrt{x^2+y^2}$$

[5]The only drawback is that BEM results in full matrices, whereas FEM results in matrices that are full of zeros, reducing computational times.

then the highest-order derivatives are

$$\mathcal{L}(u) = \left(\frac{\partial u}{\partial x}\right)^2 + \frac{\partial^2 u}{\partial y^2}$$

so that

$$\mathcal{L}(u) = \left(\frac{\partial u}{\partial x}\right)^2 + \frac{\partial^2 u}{\partial y^2}$$

which is not linear in $\frac{\partial u}{\partial x}$ (which is the highest order derivative in x-direction.

Quasi-linear PDEs are considerably more fucked up than linear PDEs, including development of discontinuities and extreme sensitivity to initial and boundary conditions.

Non-linear If you're not linear and not quasi-linear, you're non-linear and you can basically fuck off to Narnia because there are typically speaking no exact solutions for you.

In the remaining chapters, we'll focus on linear PDEs because they're easy.

Quiz 1: Q9

The equation

$$\frac{\partial u}{\partial t} + \sin(u)\frac{\partial^2 u}{\partial x \partial y}\left(\frac{\partial u}{\partial x}\right)^2 - 2\left(\frac{\partial u}{\partial y}\right)^2 + \frac{\partial^2 u}{\partial y^2} = 0$$

is most precisely characterised as:

- Non-Linear
- Quasi-Linear
- Linear

Correct is **Non-linear**. Indicating highest-order derivatives:

$$\mathcal{L}(u) = \frac{\partial u}{\partial t} + \sin(u)\frac{\partial^2 u}{\partial x \partial y}\left(\frac{\partial u}{\partial x}\right)^2 - 2\left(\frac{\partial u}{\partial y}\right)^2 + \frac{\partial^2 u}{\partial y^2}$$

Note that $\frac{\partial u}{\partial x}$ is indicated as it is the highest-order derivative in x-direction. $\frac{\partial u}{\partial y}$ is not indicated as it is not the highest-order derivative in y-direction. We still see one of the neonpink terms appearing squared, so it's non-linear.

2.4.4 *Hyperbolic, parabolic and elliptic PDEs*

You may recall from aero II that somewhere in the book, it was stated that the linearised flow potential equation (which we came across in chapters 11 and 12, regarding supersonic flow),

$$\left(1 - M_\infty^2\right)\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0$$

was ecliptic when $M_\infty < 1$ and hyperbolic when $M_\infty > 1$. Now, Anderson said that the difference wasn't really important, but it was important to remember that there *was* a difference. We'll now see what this difference exactly is (and I'll also explain how this difference physically relates to aerodynamics).

Now, elliptic and hyperbolic (and parabolic in between these two) have absolutely nothing to do with the shape of the PDE: a PDE that's elliptic does not necessarily produce a solution that looks elliptic when you plot it. The nomenclature arises from the following: in general, we have for the second-order linear PDE:

$$a\frac{\partial^2 u}{\partial x^2} + b\frac{\partial^2 u}{\partial x \partial y} + c\frac{\partial^2 u}{\partial y^2} = f$$

Now[6]:

HYPERBOLIC, PARABOLIC, ELLIPTIC PDEs

A second-order linear PDE is **hyperbolic** when

$$b^2 - 4ac > 0; \tag{2.8}$$

is **parabolic** when

$$b^2 - 4ac = 0; \tag{2.9}$$

is **elliptic** when

$$b^2 - 4ac < 0 \tag{2.10}$$

Now, what's the significance of this all? Some PDEs (the hyperbolic, and to a lesser extent, the parabolic ones) have **characteristics**, which are lines (or surfaces) where the solution can be described by an ODE rather than a PDE, which can make our lives much easier (we'll soon see how to identify those lines and the corresponding ODE, don't worry yet). This is an important concept so it's vital that you understand how that'll affect stuff: points on this line are *only* affected by other points on this line. What do I mean by this?

Well, think back to what Anderson said:

$$\left(1 - M_\infty^2\right)\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = 0$$

was ecliptic when $M_\infty < 1$ and hyperbolic when $M_\infty > 1$. Now, you have applied this equation twice, actually: once in aero I, where we set $M_\infty = 0$, so that we got the (elliptic) Laplace equation

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = 0$$

and of course in aero II, where we had $M_\infty > 1$ (so that it became hyperbolic). In both courses, we tried to find solutions to find the pressure distribution around the airfoil. Now, in aero II, this was very easy: if you wanted to find the C_p on the lower side of the airfoil, you just checked how much the flow was deflected and whether it was inward or outward (which determined your sign), and you were pretty much done already. For the calculation of the C_p on the lower side of the airfoil, you didn't actually give a flying fuck about what was on the upper side of the airfoil; there could have been a Christmas tree on top of it and it wouldn't have affected anything below the airfoil.

On the other hand, in aero I, your life would have been ruined if someone put a Christmas tree on top of the airfoil: this meant that you'd have to place sources and doublets fucking everywhere so that the flow didn't go through the Christmas tree: this means that even if you put it on the upper side of the airfoil, the pressure distribution of the lower airfoil is affected as well, contrary to what happens in the supersonic case.

So, the point I'm trying to make: for an elliptic PDE (the one of aero I), literally every point in the solution affects literally every other point; even points downstream from you (essentially in the future) affect you. However, for the hyperbolic PDE (where we do have characteristics), only points that are on the same characteristic line as you are important to you: if your characteristic went through the Christmas tree, then yes your pressure coefficient would have been different (and later I'll tell you what exactly the characteristic lines are for supersonic flow, in case you're interested). If it didn't go through it, it wouldn't have changed anything for you[7]. We'll soon see what I exactly mean by influencing other points because although you probably get what I'm trying to say, it's not really the best way of mathematically saying stuff.

[6] If it's not really clear why these names comes from, remember that the general equation for an eclipse is

$$\frac{x^2}{a} + \frac{y^2}{c} = 1$$

where the width will be $\sqrt{a}$ and the height $\sqrt{c}$. Now, you could rotate the eclipse a bit by adding a term in between:

$$\frac{x^2}{a} + \frac{xy}{b} + \frac{y^2}{c} = 1$$

However, if you play a bit with the values of a, b and c, you can get parabolas or hyperbolas. This is where the nomenclature refers to as well, as the similarity between above equation and the general second-order linear PDE should be quite clear. Other than that, there's no real meaning behind a 'hyperbolic' PDE.

[7] Although it may have got you wondering who the fuck still has his Christmas decorations up in May.

Now, hyperbolic PDEs have enough characteristics to describe the *entire* solution, meaning we can replace the entire PDE by some ODE(s). Parabolic PDEs also have characteristics, but not sufficiently many to describe the entire solution unfortunately. Elliptic PDEs have none. Let's now discuss exactly what characteristics are.

2.5 Hyperbolic PDEs

Yeah so I personally think the explanation in the book is rather shitty so I'll just explain it the way I interpret it (my explanation is also a bit more complete), I'm pretty sure it's correct but please correct me if you think I'm doing something wrong.

2.5.1 Example 1

Let's first cover a first-order hyperbolic example, and let's use the linear advection equation for this:

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0$$

Looking at figure 2.2, it kinda starts making sense what is meant with characteristics: we clearly see the graph uniformly 'moving' in North-West direction. So, would a characteristic in this case for example be lines that also point in North-West direction? The short answer is yes. The long answer is as follows (just follow my lead because I don't think you'd come up with most of the steps yourself):

As I said before, characteristic lines are lines along which the PDE can be replaced with an ODE. That implies, if we let s be the distance travelled along the characteristic line[8], that

$$\frac{du}{ds} = \frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0$$

which should make sense. Now, from the definition of the chain rule, we should also have

$$\frac{du}{ds} = \frac{\partial u}{\partial t}\frac{dt}{ds} + \frac{\partial u}{\partial x}\frac{dx}{ds} = 0$$

Now, comparing these two equations with each other, we see that

$$\frac{dt}{ds} = 1, \qquad \frac{dx}{ds} = c$$

Integration simply gives

$$\begin{aligned} t &= s + C_1 \\ x &= cs + C_2 \end{aligned}$$

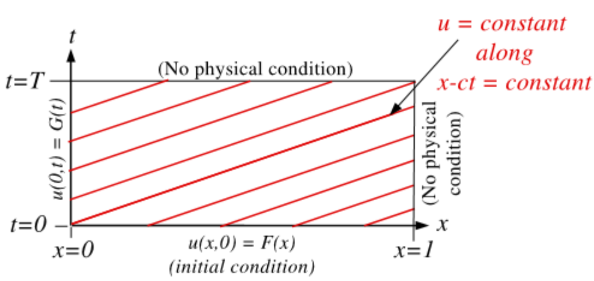


Figure 2.11: Characteristic lines for the linear advection equation.

[8] To clarify: we don't know yet what the direction is of this line, we only define s to be the distance covered along that line. Along this line, we don't measure in x and t, but simply by s distance travelled.

Now, what should we pick for C_1 and C_2? Well actually, let's take a look at figure 2.11. The red lines are the characteristic lines (we'll see really quickly that they indeed look like this, hang on for just a bit). If we take an arbitrary characteristic line, then you'll probably agree with me that it makes sense to count starting s such that when $t = 0$ (i.e. you are at the x-axis), that $s = 0$ as well (so that you basically start measuring s from the x-axis). Furthermore, let's denote the x-coordinate of the point where it then crosses the x-axis by x_0 (to be clear, each characteristic line will thus have its own x_0, corresponding to wherever it crosses the x-axis). That means that we get

$$\begin{aligned} t &= s \\ x &= cs + x_0 \end{aligned}$$

Substituting the first one into the second leads to

$$\begin{aligned} x &= ct + x_0 \\ x - ct &= x_0 \end{aligned}$$

where x_0 is a constant, belonging to the characteristic line. What I mean is, suppose a certain characteristic line has $x_0 = 1/4$, then all points (x, t) satisfying $x - ct = 1/4$ lay on this characteristic line. Similarly, if we have a point $(1/2, 2)$, then $x_0 = 1/2 - 2c$ and all points (x, t) satisfying $x - ct = 1/2 - 2c$ lay on the same characteristic line! This also means that $x - ct$ is constant, as is indicated in figure 2.11.

Now, what's the use of all this? Remember that we originally wrote

$$\frac{du}{ds} = \frac{\partial u}{\partial t}\frac{dt}{ds} + \frac{\partial u}{\partial x}\frac{dx}{ds} = 0$$

or simply

$$\begin{aligned} \frac{du}{ds} &= 0 \\ u &= \text{constant} \end{aligned}$$

In other words, u is constant *along characteristic lines*[9]. So, if we know the value of u at any point along the characteristic line, we know the value of u along the entire characteristic line. How can we make use of this? Looking at figure 2.11, we can keep in mind that the initial condition, $u(x, 0) = F(x)$ will always be known (for example, it could be $u(x, 0) = \sin(\pi x)$). The value of u of the characteristic line as it crosses the boundary, i.e. $u(x_0, 0)$ will thus be known, and thus, we have

$$u = u\left(x_0, 0\right) = F\left(x_0\right) = \sin\left(\pi x_0\right)$$

if we'd have $F\left(x_0\right) = \sin\left(\pi x_0\right)$, for example. Now, remember that we wrote somewhere that

$$x_0 = x - ct$$

and thus the solution can actually be written as

$$u(x, t) = \sin(\pi(x - ct))$$

Above stuff makes sense if we once again look at figure 2.2: the characteristic lines of the linear advection equation are parallel to the direction of movement, and indeed we see the constant values of u along those lines. This is exactly what we've been using here.

2.5.2 Boundary vs. initial conditions

Now, figure 2.11 mentions some boundary and initial conditions, so let me explain what's meant with those things. Remember that every characteristic line had only one, constant value. Now, note that if we impose an initial condition $u(x, 0) = F(x)$ on the x-axis, then we already specify the values of the characteristic lines that pass through the x-axis. Thus, we don't need to specify a condition on the right vertical axis (associated

[9] We integrate over s after all, so it should be constant in the direction of s.

with $x = 1$), because the characteristic lines passing through there already have values (imposing a boundary condition there would only lead to an overdetermined system). Now, a problem arises on the left-side: there are also lines that enter the domain by going through the vertical axis, associated with $x = 0$; these do *not* go through the x-axis and are thus not determined by the imposed initial condition. This means that we also need to impose a boundary condition on the left side, namely $u(0,t) = G(t)$. If we don't do this, the values of the characteristic lines are undetermined which is obviously bad. Now, how does that change the above solution? Well actually, it means we have to append the solution a bit.

The formula

$$u(x,t) = \sin(\pi(x - ct))$$

only holds if the characteristic line was passing through the x-axis. This is the case if $x \leq ct \leq 1$ (this is the area below the line that passes through the origin, the most extreme case of a characteristic line passing through the x-axis. If it passes through the t-axis, we can solve it in a very similar way, however. For this part of the solution, we let $x = 0$ when $s = 0$ and $t = t_0$ when $s = 0$, leading to

$$\begin{aligned} t &= s + t_0 \\ x &= cs \end{aligned}$$

so that

$$\begin{aligned} t &= \frac{x}{c} + t_0 \\ t - \frac{x}{c} &= t_0 \end{aligned}$$

Then, again, the solution is given by (if, for example, $u(0,t) = t^2$)

$$u(x,t) = u\left(0,t_0\right) = u\left(0, t - \frac{x}{c}\right) = \left(t - \frac{x}{c}\right)^2$$

which holds for $x/c \leq t \leq T$.

Also, to clarify a bit in case you're confused regarding the following:

- Initial conditions are called initial conditions cause they describe the initial condition of u, i.e. at $t = 0$.
- Boundary conditions are called boundary conditions cause they describe the condition of u at the boundaries, i.e. at $x = 0$ and $x = 1$.
- We can write $u(x,0) = F(x)$ because $u(x,0)$ is only a function of x any more. Similarly we can write $u(0,t) = G(t)$.

Furthermore, why don't we specify physical conditions at of the time period (i.e. at $t = T$) or at the right vertical axis (i.e. when $x = 1$)? Well, in that case, you'd be going back in time, which often leads to ill-posed problems and that's bad (to get philosophical, the future depends on the past but the past does not depend on the future, so it makes sense to only go forward in time).

Now, as a final note: I can explain better now with what I mean that for hyperbolic equations, points on characteristic lines are only influenced by other points on that same characteristic line: think back to all of what we did just before. The value of u along an arbitrary characteristic line was uniquely determined by the value of the physical condition (either the initial condition on the x-axis or the boundary condition on the t-axis) at the point where the characteristic line entered the domain[10]. Literally only this point mattered. If the initial conditions for the point next to it would change, then it wouldn't affect the value of this characteristic line in any way! This is what I mean: the points on characteristic lines are only affected by the initial/boundary conditions imposed on points on the characteristic lines. If the conditions change for points that are not the same characteristic as you, you simply don't give a fuck.

[10]Look back at for example $u = \sin \pi x_0$: basically, what we do is: for the desired value of u at a point with coordinates (x,t), we calculate which characteristic this point is on, and specifically, which value of x_0 is associated with this characteristic. We then simply computed what the value of the initial condition was at this specific point, and this was the value of u along the entire characteristic line.

2.5.3 Example 2

Now, let's consider a second order linear PDE, specifically the second-order wave equation,

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

Now, again, just basically follow my lead, I'll give a problem solving guide afterwards (do note that my method is totally different from what the book does , but I think my method is better).

We first rewrite this to

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

Then, you have to factorize it into[11]

$$\left(\frac{\partial}{\partial t} + c\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial t} - c\frac{\partial}{\partial x}\right) u = 0$$

This means that we essentially have two equations to satisfy:

$$\begin{aligned} \frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} &= 0 \\ \frac{\partial u}{\partial t} - c\frac{\partial u}{\partial x} &= 0 \end{aligned}$$

This means we basically have two first order PDEs which need to be replaced with ODEs. In fact, this means that we'll end up with two sets of characteristic lines. One of them follows from

$$\frac{du}{ds_1} = \frac{\partial u}{\partial t}\frac{dt}{ds_1} + \frac{\partial u}{\partial x}\frac{dx}{ds_1} = 0$$

whereas the other will follow from

$$\frac{du}{ds_2} = \frac{\partial u}{\partial t}\frac{dt}{ds_2} + \frac{\partial u}{\partial x}\frac{dx}{ds_2} = 0$$

This means that we will look for *two* solutions of u and that the total solution for u will be the summation of these two solutions. Anyways, for du/ds_1, we get that

$$\frac{dt}{ds_1} = 1, \qquad \frac{dx}{ds_1} = c$$

so that

$$\begin{aligned} t &= s + C_1 \\ x &= cs + C_2 \end{aligned}$$

Using the same logic I explained before, we set $C_1 = 0$ and $C_2 = x_0$, so that we get

$$\begin{aligned} t &= s \\ x &= cs + x_0 \\ x &= ct + x_0 \\ x - ct &= x_0 \end{aligned}$$

[11] Why don't I write

$$\left(\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x}\right)\left(\frac{\partial u}{\partial t} - c\frac{\partial u}{\partial x}\right) = 0$$

instead? Well, if you then work it out, you'd get

$$\left(\frac{\partial u}{\partial t}\right)^2 - c^2\left(\frac{\partial u}{\partial x}\right)^2 = 0$$

which is different from what we want (because now we have the first derivatives squared, but we want the second derivatives). By writing it like I do you get the *operators* squared (i.e. the second derivatives), whereas you'd otherwise get the *derivatives* squared. Fortunately for you, you'd end up at exactly the same result, but I think it's important to remember that there is a fundamental mathematical difference between them.

In other words, the first set of characteristics are lines along which $x - ct$ is constant. Along these lines, the solution is easily found:

$$\begin{aligned} \frac{du_1}{ds_1} &= 0 \\ u_1 = \text{constant} = u_1\left(x_0, 0\right) = F_1\left(x_0\right) &= F\left(x - ct\right) \end{aligned}$$

In other words: the value of u_1 (we later need to add u_2 to this) along a characteristic is equal to the value of the initial condition *corresponding to the right-running characteristics* (right-running cause lines along which $x - ct$ move to the right as time progresses). To be clear (and I'll address this point later again), we must specify *two* initial conditions: one for the right-running characteristics (call it $u_1(x,0) = F_1(x)$, one for the left-running characteristics (which we'll derive in just a second), and one for the left-running characteristic (call it $u_2(x,0) = F_2(x)$).

We can do the exact same mathematics for the second solution: for du/ds_2, we get

$$\frac{dt}{ds_2} = 1, \qquad \frac{dx}{ds_2} = -c$$

so that

$$\begin{aligned} t &= s + C_1 \\ x &= -cs + C_2 \end{aligned}$$

Using the same logic I explained before, we set $C_1 = 0$ and $C_2 = x_0$, so that we get

$$\begin{aligned} t &= s \\ x &= -cs + x_0 \\ x &= -ct + x_0 \\ x + ct &= x_0 \end{aligned}$$

In other words, the second set of characteristics are lines along which $x + ct$ is constant. Along these lines, the solution is easily found:

$$\begin{aligned} \frac{du_2}{ds_2} &= 0 \\ u_2 &= \text{constant} = u_2\left(x_0, 0\right) = F_2\left(x_0\right) = F_2\left(x + ct\right) \end{aligned}$$

In other words, the total solution becomes

$$u\left(x, t\right) = F_1\left(x - ct\right) + F_2\left(x + ct\right)$$

If you don't fully understand how these functions and their arguments work; suppose we have the initial conditions

$$\begin{aligned} F_1\left(x\right) &= \cos\left(\pi x\right) \\ F_2\left(x\right) &= \frac{\sin\left(\pi x\right)}{x} \end{aligned}$$

then

$$u\left(x, t\right) = F_1\left(x - ct\right) + F_2\left(x + ct\right) = \cos\left(\pi\left(x - ct\right)\right) + \frac{\sin\left(\pi\left(x + ct\right)\right)}{x + ct}$$

In other words, $F_1\left(x - ct\right)$ means simply that you plug in $x - ct$ for every instant of x in F_1, and $F_2\left(x + ct\right)$ means that you simply plug in $x + ct$ for every instant of x in F_2.

Note that we can beautifully draw the characteristic lines as done in figure 2.12. Note that we need to specify *four* physical conditions, as we now basically have two waves. For the right-running lines, you need to specify one initial condition on the x-axis and one boundary condition on the vertical axis where $x = 0$, as right-running waves can enter the domain via one of these boundaries. Similarly, for the right-running lines, you need to specify another initial condition on the x-axis and one boundary condition on the vertical axis where $x = 1$, as

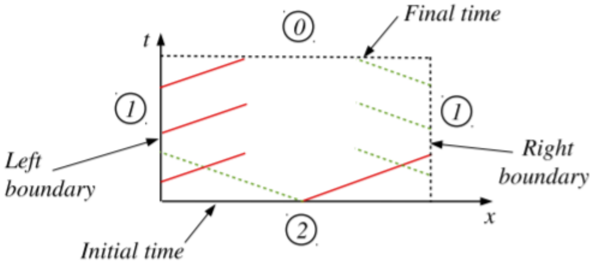


Figure 2.12: Boundary conditions for the wave equation. The number of physical conditions which can be specified on each boundary is indicated by the numbers within the circles.

left-running waves can enter the domain via one of these boundaries. Note that the solution for points where the characteristic lines enters via the boundary conditions is again different compared to the solution for points where the characteristic lines enter via the initial condition, and you should take the same care when determining exactly what solution applies for an arbitrary point (x, t) somewhere in the domain (but rest assured, I'm pretty sure they won't ask that complicated questions).

FINDING THE SOLUTION FOR A SECOND-ORDER, HYPERBOLIC PDE

Starting with a PDE of the form (note that in the former, I used t and x where I now use x and y (so x is first replaced with y, and then t is replaced with x), in case you didn't notice):

$$a\frac{\partial^2 u}{\partial x^2} + b\frac{\partial^2 u}{\partial x \partial y} + c\frac{\partial^2 u}{\partial y^2} = 0$$

1. Divide by a to get
$$\frac{\partial^2 u}{\partial x^2} + \alpha\frac{\partial^2 u}{\partial x \partial y} + \beta\frac{\partial^2 u}{\partial y^2} = 0$$
where $\alpha = b/a$ and $\beta = c/a$.
2. Factorize this as
$$\left(\frac{\partial}{\partial x} + r_1\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial x} + r_2\frac{\partial}{\partial y}\right)u = 0$$
where
$$\begin{aligned} r_1 &= \frac{\alpha + \sqrt{\alpha^2 - 4\beta}}{2} \\ r_2 &= \frac{\alpha - \sqrt{\alpha^2 - 4\beta}}{2} \end{aligned}$$
Note that this is very similar to the ABC-formula, but slightly different (but trust me, this is correct).
3. Note that this means that
$$\begin{aligned} \frac{du_1}{ds_1} &= \frac{\partial u}{\partial x} + r_1\frac{\partial u}{\partial y} = 0 \\ \frac{du_2}{ds_2} &= \frac{\partial u}{\partial x} + r_2\frac{\partial u}{\partial y} = 0 \end{aligned}$$
whilst from the chain rule, we have
$$\begin{aligned} \frac{du_1}{ds_1} &= \frac{\partial u_1}{\partial x}\frac{dx}{ds_1} + \frac{\partial u_1}{\partial y}\frac{dy}{ds_1} \\ \frac{du_2}{ds_2} &= \frac{\partial u_2}{\partial x}\frac{dx}{ds_2} + \frac{\partial u_2}{\partial y}\frac{dy}{ds_2} \end{aligned}$$

4. First focussing on the first set characteristics: from comparison, conclude that

$$\begin{aligned}\frac{dx}{ds_1} &= 1\\ \frac{dy}{ds_1} &= r_1\\ \frac{du_1}{ds_1} &= 0\end{aligned}$$

5. Integrate these equations to get

$$\begin{aligned}x &= s_1 + C_1\\ y &= r_1 s_1 + C_2\end{aligned}$$

6. Choose C_1 and C_2 such that $x = 0$ when $s_1 = 0$ and $y = y_0$ when $s = 0$, i.e.

$$\begin{aligned}x &= s_1\\ y &= r_1 s_1 + y_0\end{aligned}$$

7. Combine these equations to get the equation that describes the shape of the characteristics:

$$y - r_1 x = y_0$$

8. Integrate $du_1/ds_1 = 0$ to get

$$u_1 = \text{constant} = u_1\left(0, y_0\right) = F_1\left(y_0\right) = F_1\left(y - r_1 x\right)$$

   where $F_1(y)$ is the boundary condition imposed on the y-axis.
9. Repeat steps 4-8 for the second set of characteristics.
10. Find the total solution by

$$u(y,t) = F_1\left(y - r_1 x\right) + F_2\left(y + r_1 x\right)$$

It seems like a lot, but most of the steps are really easy (and it's literally the same every time, the only thing you actually have to do which is slightly more difficult is the factorisation).

Quiz 1: Q11

For the following equation

$$\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) + 3\frac{\partial^2 u}{\partial x \partial y} = 0$$

the characteristic directions can be defined by:

$$\frac{d}{ds_1} \equiv \frac{\partial}{\partial x} + A\frac{\partial}{\partial y}$$

The value for A which is largest in magnitude is: ...

Hint: Start by defining:

$$p = \frac{\partial u}{\partial x} \quad \text{and} \quad q = \frac{\partial u}{\partial y}$$

Then write a linear combination of the PDE and the continuity requirement as:

$$\left(\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} + 3\frac{\partial p}{\partial y}\right) + \alpha\left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) = 0$$

Yeah so we are going to ignore that hint and do it our own way. Let me write it as

$$\frac{\partial^2 u}{\partial x^2} + 3\frac{\partial^2 u}{\partial x \partial y} + \frac{\partial^2 u}{\partial y^2} = 0$$

This means that $\beta = 3$ and $\alpha = 1$, and thus

$$\begin{aligned} \left(\frac{\partial}{\partial x} + \frac{3+\sqrt{3^2-4\cdot 1}}{2\cdot 1}\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial x} + \frac{3-\sqrt{3^2-4\cdot 1}}{2\cdot 1}\frac{\partial}{\partial y}\right)u &= 0 \\ \left(\frac{\partial}{\partial x} + \frac{3+\sqrt{5}}{2}\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial x} + \frac{3-\sqrt{5}}{2}\frac{\partial}{\partial y}\right) &= 0 \end{aligned}$$

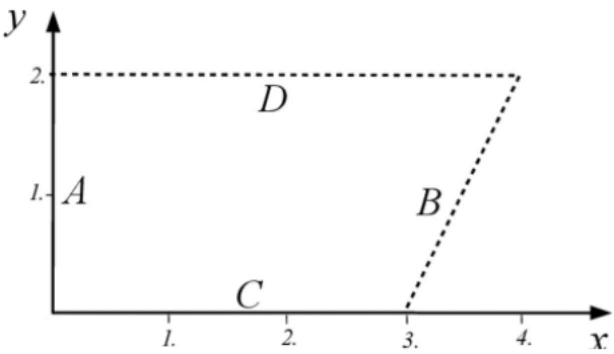
Clearly, $A = \frac{3+\sqrt{5}}{2}$. The easiest way to input this is by clicking on the button with the Σ in it, it should then speak for itself. Note: if you verify your answer, it shows a different result, but the only difference is how it's written (the expressions are exactly equivalent otherwise).

Quiz 1: Q13

A second-order hyperbolic PDE has characteristic directions defined by

$$\begin{aligned} y &= x + \text{constant} \\ y &= -x + \text{constant} \end{aligned}$$

on the domain enclosed by the segments A, C, B and D, as shown in the figure below.



Assume that when boundary conditions are applied, they specify the value of one invariant for an entire segment (e.g. one variable for all of C).

If two boundary conditions are imposed on segment B, what are the required numbers of boundary conditions which should be specified on segments A, C, and D respectively?

- 1,1,1
- 0,1,1
- 2,1,1
- 1,2,2
- 0,2,2
- 2,1,2

Correct is 0,1,1. Why? We have a second-order PDE of two dependent variables, x and y. Thus, we need $2 \cdot 2 = 4$ boundary conditions in total. With two already imposed on B, this means we only need two others. The option 0,1,1 is the only one that uses two extra boundary conditions, the rest all uses too many. Do note: as we have x and y coordinates, we are free to place our boundary conditions wherever we want, although we can't place more than two on the same boundary. You don't have the problem of not being allowed to place it at certain boundaries because then you'd be going back in time, for example (which was a problem in the PDEs we saw previously).

2.5.4 More complicated examples (skip if you're not interested)

Pretty sure this gets really beyond the scope of this course, so if you want, you can just skip till the next section, but if you want to see more complicated examples regarding all of this, you can read the following. Suppose we have the equation

$$\frac{\partial u}{\partial t} + x\frac{\partial u}{\partial x} = \sin(t)$$

Can we do apply aforementioned techniques for this? Yes, we can, and rather easily in fact. Note that this is already first-order, so we don't need to factorize it. We can compare

$$\frac{du}{ds} = \frac{\partial u}{\partial t}\frac{dt}{ds} + \frac{\partial u}{\partial x}\frac{dx}{ds} = \sin(t)$$

Note that now, $du/ds = \sin(t)$ instead of $du/ds = 0$ (so now you clearly see that not always we end up with constant values for u). We now have

$$\begin{aligned} \frac{dt}{ds} &= 1 \\ \frac{dx}{ds} &= x \end{aligned}$$

Upon integrating, this becomes

$$\begin{aligned} t &= s + C_1 \\ x &= C_2 e^s \end{aligned}$$

Again, let $t = 0$ when $s = 0$ and $x = x_0$ when $s = 0$; this straightforwardly leads to $C_1 = 0$ and $C_2 = x_0$. Thus,

$$\begin{aligned} t &= s \\ x &= x_0 e^s = x_0 e^t \\ xe^{-t} &= x_0 \end{aligned}$$

In other words, in this case, the characteristics are lines along which the product xe^{-t} is constant! These are clearly not simple straight lines as you've seen before; the presence of x in front of $\partial u/\partial x$ fucks everything up. Now, how is the solution for u along these characteristics? Well, we had

$$\frac{du}{ds} = \sin(t)$$

but it's not easy to integrate that directly over s as t is a function of s (so it does not become $u = s \cdot \sin(t)$ for example). However, we already found $t = s$, and thus

$$\begin{aligned} \frac{du}{ds} &= \sin(s) \\ u &= -\cos(s) + C_3 \end{aligned}$$

Now, if we have an initial condition $u(x, 0) = F(x)$, then for $s = 0$, u should be equal to $F(x_0)$, and thus

$$\begin{aligned} -\cos(0) + C_3 = -1 + C_3 &= F\left(x_0\right) \\ C_3 &= 1 + F\left(x_0\right) \end{aligned}$$

Thus,

$$u(x,t) = -\cos(s) + 1 + F\left(x_0\right) = -\cos(t) + 1 + F\left(xe^{-t}\right)$$

In other words, given an initial condition $F(x)$, we just substitute xe^{-t} everywhere where we see x in that formula, and that'll be the value for $F\left(xe^{-t}\right)$.

Let's do another example. Suppose we have

$$\frac{\partial u}{\partial t} + g(u)\frac{\partial u}{\partial x} = 0$$

where $g(u)$ is a function of u (e.g. it could be $g(u) = \sin(u)$, $g(u) = 3u^2$, etc., it could be anything), with the initial condition $u(x,0) = F(x)$. Again, we can apply characteristics to all of this fun. Comparing with

$$\frac{du}{ds} = \frac{\partial u}{\partial t}\frac{dt}{ds} + \frac{\partial u}{\partial x}\frac{dx}{ds} = 0$$

we easily see that

$$\begin{aligned} \frac{dt}{ds} &= 1 \\ \frac{dx}{ds} &= g(u) \end{aligned}$$

Now, the second is very difficult to integrate at first because we don't know how u relates to s. Fortunately, we can rather easily find this, however. We also have

$$\begin{aligned} \frac{du}{ds} &= 0 \\ u &= \text{constant} = F\left(x_0\right) \end{aligned}$$

In other words, u is simply a constant along a characteristic, and thus $g(u)$ is also constant along a characteristic. This means that our integration is actually very easy, and applying our conditions of $t = 0$ when $s = 0$ and $x = x_0$ when $s = 0$, this leads to

$$\begin{aligned} t &= s \\ x &= g(u)s + x_0 = g\left(F\left(x_0\right)\right) + x_0 \end{aligned}$$

so that

$$x_0 = x - g\left(F\left(x_0\right)\right)$$

and thus

$$u = F\left(x_0\right) = F\left(x - g\left(F\left(x_0\right)\right)\right) = F(x - g(u))$$

which is an *implicit* expression for u (but an expression nonetheless).

One final remark, then we'll continue: suppose we'd have

$$a\frac{\partial^2 u}{\partial x^2} + b\frac{\partial^2 u}{\partial x \partial y} + c\frac{\partial^2 u}{\partial y^2} + d\frac{\partial u}{\partial x} + e\frac{\partial u}{\partial y} = f$$

Would you still be able to apply (at least part of) the knowledge you have gained so far? Well, to some extent, yes you can. Given that above equation is hyperbolic (so you need to manually verify that $b^2 - 4ac > 0$), then to find the equations describing the shape of the characteristics, you can still divide by a and factorise as I explained in the problem solving guide. The presence of the other three terms absolutely do not matter for that!

2.6 Parabolic PDEs

We can apply what I just said to the linear diffusion equation:

$$\frac{\partial u}{\partial t} = v\frac{\partial^2 u}{\partial x^2}$$

First, we rewrite it to[12]

$$v\frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial t} = 0$$

[12] And note that this is indeed clearly parabolic:

$$b^2 - 4ac = 0^2 - 4 \cdot 0 \cdot v = 0$$

and thus it's parabolic.

and as I said, $\frac{\partial u}{\partial t}$ absolutely does not matter to determine the characteristics now! We can factorize it as

$$\left(\sqrt{v}\frac{\partial}{\partial x}\right)\left(\sqrt{v}\frac{\partial}{\partial x}\right)u + \text{lower order terms} = 0$$

This basically means that we have (we ignore all of the lower order terms, trust me, that is correct, and you don't wanna know why)

$$\begin{aligned}\frac{du_1}{ds_1} &= \sqrt{v}\frac{\partial u_1}{\partial x}\\ \frac{du_2}{ds_2} &= \sqrt{v}\frac{\partial u_2}{\partial x}\end{aligned}$$

but these are exactly the same obviously (i.e. they'd lead to exactly the same characteristic). So, we can just focus on solving stuff for

$$\frac{du}{ds} = \sqrt{v}\frac{\partial u}{\partial x} = 0$$

Comparing with

$$\frac{du}{ds} = \frac{\partial u}{\partial t}\frac{dt}{ds} + \frac{\partial u}{\partial x}\frac{dx}{ds}$$

we quickly see that

$$\begin{aligned}\frac{dt}{ds} &= 0\\ \frac{dx}{ds} &= \sqrt{v}\end{aligned}$$

This means that t is not a function of s, and thus that s is parallel to x, or in other words, the characteristics are parallel to x. Why is this important?

Well first of all, the very fact that we have only one characteristic is already a sign of trouble: since the equation is second order, you'd actually need two sets of characteristics to be able to fully describe the solution using characteristics (compare the difference between the first and second example of the hyperbolic PDEs). Therefore, although we can use characteristics to supply us with *some* information about the solution at other points[13], this information is not complete (in the second-order wave equation, we saw that one characteristic actually only gave half of the solution and you needed the other characteristic to find the other half).

Regarding boundaries, looking at figure 2.13, we need to specify an initial condition at $t = 0$, and boundary conditions at $x = 0$ and $x = 1$ to get a well-posed problem. You obviously don't do it at $t = t_{\text{final}}$; this would mean that your solution would start at the future and be going back in time and as explained before, this is generally a bad idea.

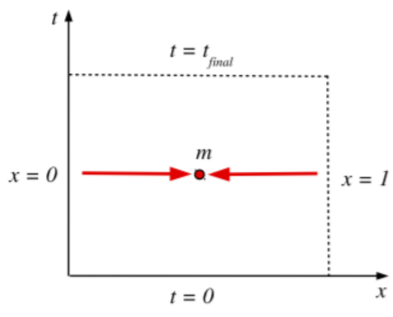


Figure 2.13: Space-time domain for the heat equation showing the domain of dependence of point m implied by the characteristic found by the analysis within the circles.

[13]To be precise, if we know the value of u (the temperature) at a x-coordinate for a fixed time t, we can say stuff about the temperature at different points along the beam for that same time t.

2.7 Elliptic PDEs

The Laplace equation is

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

and this is clearly elliptic:

$$b^2 - 4ac = 0^2 - 4 \cdot 1 \cdot 1 = -4 < 0$$

However, we could also deduce this from the fact there are no characteristics: if we'd try to factorize it, we'd end up at

$$\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) u = 0$$

and yeah you can basically already gave up because there are imaginary numbers and they're not pissing off unfortunately. Thus, we conclude, there are indeed no characteristics. This means that the solution at all points in the domain depends continuously on that of *all* other points. A well-posed problem for Laplace's equation thus requires information to be specified at *every* point on the boundary (so on all of the four boundaries). Please note that looking at for example figure 2.13, this means that you'd also be specifying a boundary condition on the far end (where $t = t_{\text{final}}$); however, bear in mind that for Laplace, we use x and y rather than t and x, so there's not really a 'future' so it does not mean you're going back in time.

Quiz 1: Q10

An analysis carried out for a PDE shows that some characteristic directions can be defined, but not enough to solve the problem using a set of ODEs. The PDE can be classified as

- elliptic
- linear
- parabolic
- non-linear
- quasi-linear
- hyperbolic

The correct answer is **parabolic**.

Quiz 1: Q12

Elliptic PDEs require boundary conditions

- on certain points depending on the interior data.
- which are a function of the applied source term.
- which respond to interior data.
- for all boundary points.

Correct is **for all boundary points**.

3 Discretisation with the finite-difference method

In this chapter, we'll actually learn how to numerically solve PDEs, using the finite-difference method.

3.1 From PDE to algebraic system

As said before, if we have the domain as shown in figure 3.1, then we discretizise it in a collection of pre-specified points in the domain. The lines connecting such points are collectively referred to as a **mesh**.

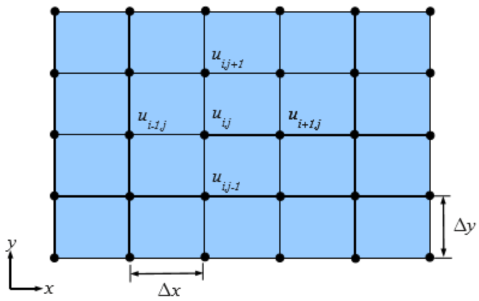


Figure 3.1: A uniform structured mesh for a rectangular domain.

Do note, we could also have periodic domains, as shown in figure 3.2. These are domains which re-connect to themselves so that boundaries are avoided and therefore boundary conditions are also not required.

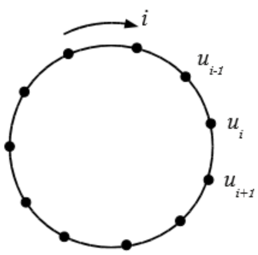


Figure 3.2: A one-dimensional periodic domain.

3.1.1 Example: linear advection equation

Suppose we have the linear advection equation:

$$\begin{aligned} \frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} &= 0 \\ u(x,0) &= u_0(x) \\ u(0,t) &= u(L,t) \end{aligned}$$

How would we go on about solving this? Well, actually, it's more straightforward than you'd think. We divide the domain up as shown in figure 3.1 (although we'll be using t instead of y). Then look at figure 3.3: we'll be using the points u^n_{i-1}, u^n_i and u^n_{i+1} to estimate u^{n+1}_i.[1] Why does this make sense? Well, our initial condition is known, thus we know the values of u at the very first row. Using the points shown in figure 3.3, this means that

[1] To be clear, these are respectively the points $(x_i - \Delta x, t_n)$, (x_i, t_n), $(x_i + \Delta x, t_n)$, $(x_i, t_n + \Delta t)$.

we use the points on the first row to estimate the values of the points on the second row. We can then again estimate the third row, based on the second row, etc.

This collection of points is collectively referred to as a **stencil** (to be clear: the stencil only refers to the points that are used in the calculation of another point, it does not refer to all the points on the domain). Now, how exactly would our method look like?

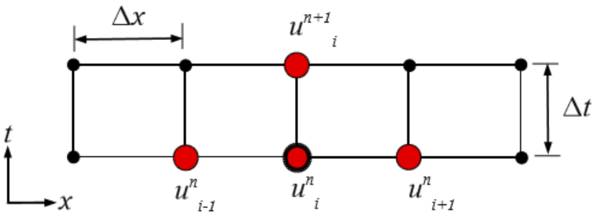


Figure 3.3: Finite-difference stencil for an explicit method. The circle with a bold outline indicates the expansion point for the Taylor series.

Well, actually, if you studied my summary for the second numerical analysis quiz well, it's actually really straightforward. $\frac{\partial u}{\partial t}$ and $\frac{\partial u}{\partial x}$ are both uni-directional derivatives (for $\frac{\partial u}{\partial x \partial y}$, the story is slightly more complicated, but we'll get to that later). I included a section in the summary that wasn't in the book, which was about finding the optimum differentiation scheme utilizing certain points. In case you forgot, this was basically what I instructed you to do (I have made slight notation changes to make it in line with this course).

METHOD OF UNDETERMINED COEFFICIENTS TO FIND OPTIMAL DIFFERENTIATION SCHEME FOR UNI-DIRECTIONAL DERIVATIVES OF ANY ORDER

If one wants to know the partial derivative $\frac{\partial^p u}{\partial t^p}$ at the point (x_i, t_n) (i.e. u_i^n), and the points $(x_i, t_n + k_j \Delta t)$, $j = 1, ..., J$ are used to determine this partial derivative, then the entries of matrix $[A]$ given by

$$[A]_{zj} = \frac{k_j^{z-1}}{(z-1)!}$$

where z denotes the zth row, and $1 \leq z \leq J$. Bear in mind that $0! = 1$ and $0^0 = 1$ as well. The undetermined coefficients of the linear combination

$$\frac{\partial^p u}{\partial t^p} = \sum_{j=1}^{J} a_j u(x_i, t_n + k_j \Delta t)$$

are determined by solving

$$\mathbf{a} = [A]^{-1}\mathbf{f}$$

where $\mathbf{f}$ has zeros everywhere, except for the $p + 1$st row, where the entry is equal to $1/\Delta t^p$.

Because I think you probably have no clue any more what the fuck I mean with this, let's do an example. In figure 3.3, we'll be looking for the partial derivatives at the circle with the bold outline. Then, to find $\frac{\partial u}{\partial t}$, we thus use the points $(x_i, t_n + \Delta t)$ and (x_i, t_n). Our final equation will look like

$$\frac{\partial u}{\partial t} = a_1 u(x_i, t_n + \Delta t) + a_2 u(x_i, t_n)$$

where a_1 and a_2 are to be determined. This means that $k_1 = 1$ and $k_2 = 0$, and $1 \leq j \leq 2$ (i.e. $J = 2$)[2]. Thus, our matrix looks like[3]

$$[A] = \begin{bmatrix} \frac{1^0}{(1-1)!} & \frac{0^0}{(1-1)!} \\ \frac{1^1}{(2-1)!} & \frac{0^1}{(2-1)!} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

[2] As we're using the points $(x_i, t_n + \Delta t)$ and (x_i, t_n).

[3] The factorials are $(z - 1)!$, z being the zth row you're on.

Furthermore, we're interested in the $p = 1$st derivative, thus

$$\mathbf{f} = \begin{bmatrix} 0 \\ \frac{1}{\Delta t} \end{bmatrix}$$

Thus,

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = [A]^{-1}\mathbf{f} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \frac{1}{\Delta t} \end{bmatrix} = \begin{bmatrix} \frac{1}{\Delta t} \\ -\frac{1}{\Delta t} \end{bmatrix}$$

and thus we have

$$\frac{\partial u}{\partial t} = \frac{u(x_i, t_n + \Delta t) - u(x_i, t_n)}{\Delta t} = \frac{u_i^{n+1} - u_i^n}{\Delta t}$$

In case you have difficulties manually computing $\mathbf{a}$; you can leave out the $/h$ and use your calculator to compute

$$\mathbf{a} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

and then add the $/\Delta t$ now.

Let's do $\frac{\partial u}{\partial x}$ as well: we use the points $(x_i - \Delta x, t_n)$, (x_i, t_n), $(x_i + \Delta x, t_n)$ for this. In other words, our final equation looks like

$$\frac{\partial u}{\partial x} = a_1 u(x_i - \Delta x, t_n) + a_2 u(x_i, t_n) + a_3 u(x_i + \Delta x, t_n) = a_1 u_{i-1}^n + a_2 u_i^n + a_3 u_{i+1}^n$$

where a_1, a_2 and a_3 need to be determined. We have $k_1 = -1$, $k_2 = 0$ and $k_3 = 1$. Thus, we have

$$[A] = \begin{bmatrix} \frac{(-1)^0}{(1-1)!} & \frac{0^0}{(1-1)!} & \frac{1^0}{(1-1)!} \\ \frac{(-1)^1}{(2-1)!} & \frac{0^1}{(2-1)!} & \frac{1^1}{(2-1)!} \\ \frac{(-1)^2}{(3-1)!} & \frac{0^2}{(3-1)!} & \frac{1^2}{(3-1)!} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

so that

$$\mathbf{a} = [A]^{-1}\mathbf{f} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \frac{1}{\Delta x} \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2\Delta x} \\ 0 \\ \frac{1}{2\Delta x} \end{bmatrix}$$

and thus we have

$$\frac{\partial u}{\partial x} = \frac{-u_{i-1}^n + u_{i+1}^n}{2\Delta x}$$

So, now we have

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = \frac{u_i^{n+1} - u_i^n}{\Delta t} + c\frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x} = 0$$

This can be rewritten to get an explicit expression for u_i^{n+1}:

$$\begin{aligned} \frac{u_i^{n+1} - u_i^n}{\Delta t} + c\frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x} &= 0 \\ u_i^{n+1} - u_i^n + \frac{c\Delta t}{\Delta x}\left(u_{i+1}^n - u_{i-1}^n\right) &= 0 \\ u_i^{n+1} &= u_i^n + \frac{c\Delta t}{2\Delta x}\left(u_{i-1}^n - u_{i+1}^n\right) \end{aligned}$$

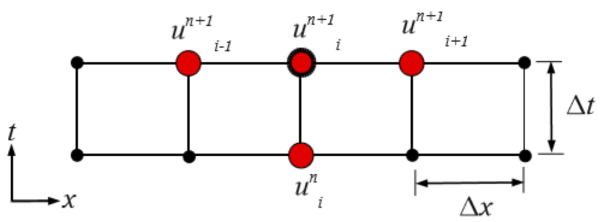


Figure 3.4: Finite-difference stencil for an implicit method.

and thus we are able to compute u_i^{n+1} if we know the values on the previous row! Absolutely wonderful stuff. Note that this is an **explicit method**, as we are able to obtain an explicit expression. This particular first-order accurate time-integration method with all spatial derivatives (i.e. the one in x-direction) at time level n is referred to as the **explicit Euler method**. However, you can just as easily determine an implicit method: look at figure 3.4.

We'll be using the points u_{i-1}^{n+1}, u_{i+1}^{n+1}, u_i^n to determine u_i^{n+1}; we'll be Taylor-expanding about this point as well. You can do this yourself, and this'll lead to

$$u_i^{n+1} + \frac{c\Delta t}{2\Delta x}\left(u_{i+1}^{n+1} - u_{i-1}^{n+1}\right) = u_i^n$$

Note that this is an **implicit method**: u_{i+1}^{n+1} and u_{i-1}^{n+1} are not known either, so it's implicit. Instead, we must solve the following matrix:

$$([I] + [D])\,\mathbf{u}^{n+1} = \mathbf{u}^n$$

with $[I]$ the identity matrix,

$$\mathbf{u}^{n+1} = \begin{bmatrix} u_1^{n+1} \\ u_2^{n+1} \\ u_3^{n+1} \\ \vdots \end{bmatrix}, \qquad \mathbf{u}^n = \begin{bmatrix} u_1^n \\ u_2^n \\ u_3^n \\ \vdots \end{bmatrix}$$

and $[D]$ given by

$$[D] = \frac{c\Delta t}{2\Delta x}\begin{bmatrix} 0 & 1 & & & & & -1 \\ -1 & 0 & 1 & & & & \\ & -1 & 0 & 1 & & & \\ & & -1 & 0 & 1 & & \\ & & & -1 & 0 & 1 & \\ & & & & -1 & 0 & 1 \\ 1 & & & & & -1 & 0 \end{bmatrix}$$

Why? The implicit equation for example leads to

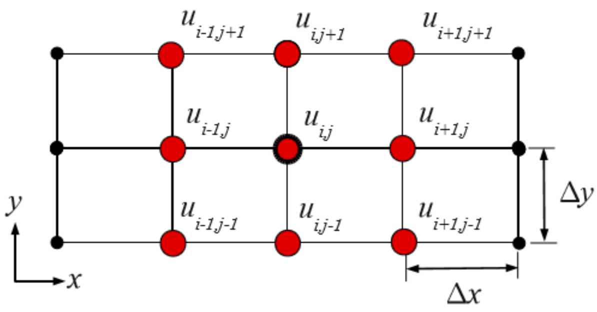
$$u_4^{n+1} + \frac{c\Delta t}{2\Delta x}\left(u_5^{n+1} - u_3^{n+1}\right) = u_i^n$$

Then, from comparison, we see that $[D]$ indeed should be like this. Note that now, we assume that u_1^{n+1} is next to u_7^{n+1} (assuming we have 7 points in x-direction). This means that we assume the domain is periodic; what we should do if it's not the case we'll see later.

3.1.2 *Example: Laplace's equation*

Suppose we have the Laplace equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$



f

Figure 3.5: Finite-difference stencil for an 2D domain.

Can you do the same analysis for that as well? Yes you can.

Consider the stencil of figure 3.5. We'll be using

$$u_{i-1,j}, \quad u_{i,j-1}, \quad u_{i,j+1}, \quad ui+1,j$$

to estimate $u_{i,j}$ (yes this method will be implicit). Let's find the two derivatives at $u_{i,j}$: for $\frac{\partial^2 u}{\partial x^2}$, we'll use the points that lay on the same horizontal line as $u_{i,j}$, i.e. $u_{i-1,j}$, $u_{i,j}$ and $u_{i+1,j}$. We then have that we eventually want to get something like

$$\frac{\partial^2 u}{\partial x^2} = a_1 u_{i-1,j} + a_2 u_{i,j} + a_3 u_{i+1,j}$$

Our matrix looks like (since we have $k_1 = -1$, $k_2 = 0$ and $k_3 = 1$):

$$\begin{bmatrix} \frac{(-1)^0}{(1-1)!} & \frac{0^0}{(1-1)!} & \frac{1^0}{(1-1)!} \\ \frac{(-1)^1}{(2-1)!} & \frac{0^1}{(2-1)!} & \frac{1^1}{(2-1)!} \\ \frac{(-1)^2}{(3-1)!} & \frac{0^2}{(3-1)!} & \frac{1^2}{(3-1)!} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix}$$

and

$$\mathbf{f} = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{\Delta x^2} \end{bmatrix}$$

so that

$$\mathbf{a} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ \frac{1}{\Delta x^2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\Delta x^2} \\ -\frac{2}{\Delta x^2} \\ \frac{1}{\Delta x^2} \end{bmatrix}$$

Thus, we get

$$\frac{\partial^2 u}{\partial x^2} = \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{\Delta x^2}$$

In an exactly similar derivation, we get

$$\frac{\partial^2 u}{\partial y^2} = \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{\Delta y^2}$$

Thus, we have

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{\Delta x^2} + \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{\Delta y^2} = 0$$

So, how would this lead to an algebraic system (i.e. some matrices and that kind of fun stuff)? Well, for that, let us consider the vector **u**, which consists of *all* points in the mesh, ordered by first moving in the i and then in the j direction, as shown in figure 3.6, i.e.

$$\mathbf{u} = \begin{bmatrix} u_{1,1} \\ u_{2,1} \\ u_{3,1} \\ \vdots \\ u_{1,2} \\ u_{2,2} \\ u_{3,2} \\ \vdots \\ u_{I,J} \end{bmatrix}$$

where I is the total number of points in i direction and J the total number of points in j direction.

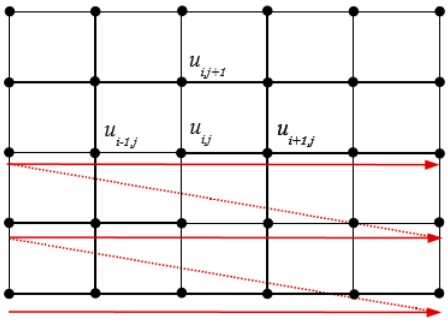


Figure 3.6: Solution vector ordering for a 2D domain.

Then, we'll write the resultant algebraic system as

$$[A]\mathbf{u} = \left(\left[A_1\right] + \left[A_2\right]\right)\mathbf{u} = 0$$

where $[A_1]$ is the matrix corresponding to $\frac{\partial^2 u}{\partial x^2}$ and $[A_2]$ the matrix corresponding to $\frac{\partial^2 u}{\partial y^2}$. How do these matrices look like? To make it a bit easier to understand, allow me to constrict ourselves to a 3×3 domain (i.e. 3 points in i direction and 3 in j direction). Then,

$$\mathbf{u} = \begin{bmatrix} u_{1,1} \\ u_{2,1} \\ u_{3,1} \\ u_{1,2} \\ u_{2,2} \\ u_{3,2} \\ u_{1,3} \\ u_{2,3} \\ u_{3,3} \end{bmatrix}$$

Let us analyse the point in the middle, i.e. $u_{2,2}$. To estimate $u_{2,2}$, we'd apply

$$\underbrace{\frac{u_{1,2} - 2u_{2,2} + u_{3,2}}{\Delta x^2}}_{\text{Contribution of } [A_1]} + \underbrace{\frac{u_{2,1} - 2u_{2,2} + u_{2,3}}{\Delta y^2}}_{\text{Contribution of } [A_2]} = 0$$

Let us consider the contribution of $[A_1]$, i.e.

$$\frac{u_{1,2} - 2u_{2,2} + u_{3,2}}{\Delta x^2}$$

$[A_1]$ will be a 9×9 matrix (as $\mathbf{u} = 9 \times 1$):

$$[A_1] = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} & A_{16} & A_{17} & A_{18} & A_{19} \\ A_{21} & A_{22} & A_{23} & A_{24} & A_{25} & A_{26} & A_{27} & A_{28} & A_{29} \\ A_{31} & A_{32} & A_{33} & A_{34} & A_{35} & A_{36} & A_{37} & A_{38} & A_{39} \\ A_{41} & A_{42} & A_{43} & A_{44} & A_{45} & A_{46} & A_{47} & A_{48} & A_{49} \\ A_{51} & A_{52} & A_{53} & A_{54} & A_{55} & A_{56} & A_{57} & A_{58} & A_{59} \\ A_{61} & A_{62} & A_{63} & A_{64} & A_{65} & A_{66} & A_{67} & A_{68} & A_{69} \\ A_{71} & A_{72} & A_{73} & A_{74} & A_{75} & A_{76} & A_{77} & A_{78} & A_{79} \\ A_{81} & A_{82} & A_{83} & A_{84} & A_{85} & A_{86} & A_{87} & A_{88} & A_{89} \\ A_{91} & A_{92} & A_{93} & A_{94} & A_{95} & A_{96} & A_{97} & A_{98} & A_{99} \end{bmatrix}$$

and thus we need to analyse

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} & A_{16} & A_{17} & A_{18} & A_{19} \\ A_{21} & A_{22} & A_{23} & A_{24} & A_{25} & A_{26} & A_{27} & A_{28} & A_{29} \\ A_{31} & A_{32} & A_{33} & A_{34} & A_{35} & A_{36} & A_{37} & A_{38} & A_{39} \\ A_{41} & A_{42} & A_{43} & A_{44} & A_{45} & A_{46} & A_{47} & A_{48} & A_{49} \\ A_{51} & A_{52} & A_{53} & A_{54} & A_{55} & A_{56} & A_{57} & A_{58} & A_{59} \\ A_{61} & A_{62} & A_{63} & A_{64} & A_{65} & A_{66} & A_{67} & A_{68} & A_{69} \\ A_{71} & A_{72} & A_{73} & A_{74} & A_{75} & A_{76} & A_{77} & A_{78} & A_{79} \\ A_{81} & A_{82} & A_{83} & A_{84} & A_{85} & A_{86} & A_{87} & A_{88} & A_{89} \\ A_{91} & A_{92} & A_{93} & A_{94} & A_{95} & A_{96} & A_{97} & A_{98} & A_{99} \end{bmatrix} \begin{bmatrix} u_{1,1} \\ u_{2,1} \\ u_{3,1} \\ u_{1,2} \\ u_{2,2} \\ u_{3,2} \\ u_{1,3} \\ u_{2,3} \\ u_{3,3} \end{bmatrix}$$

$u_{2,2}$ is in the fifth row, thus you need to analyse the fifth row of $[A_1]$ as well. If we work out the dot product between the fifth row of $[A_1]$ and $\mathbf{u}$, we get

$$A_{51}u_{1,1} + A_{52}u_{2,1} + A_{53}u_{3,1} + A_{54}u_{1,2} + A_{55}u_{2,2} + A_{56}u_{3,2} + A_{57}u_{1,3} + A_{58}u_{2,3} + A_{59}u_{3,3}$$

Now, compare this with

$$\frac{u_{1,2} - 2u_{2,2} + u_{3,2}}{\Delta x^2}$$

Clearly, we have

$$A_{54} = \frac{1}{\Delta x^2}, \quad A_{55} = \frac{-2}{\Delta x^2}, \quad A_{56} = \frac{1}{\Delta x^2}$$

and all others equal to zero, and thus we'd get the matrix

$$[A_1] = \frac{1}{\Delta x^2} \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} & A_{16} & A_{17} & A_{18} & A_{19} \\ A_{21} & A_{22} & A_{23} & A_{24} & A_{25} & A_{26} & A_{27} & A_{28} & A_{29} \\ A_{31} & A_{32} & A_{33} & A_{34} & A_{35} & A_{36} & A_{37} & A_{38} & A_{39} \\ A_{41} & A_{42} & A_{43} & A_{44} & A_{45} & A_{46} & A_{47} & A_{48} & A_{49} \\ 0 & 0 & 0 & 1 & -2 & 1 & 0 & 0 & 0 \\ A_{61} & A_{62} & A_{63} & A_{64} & A_{65} & A_{66} & A_{67} & A_{68} & A_{69} \\ A_{71} & A_{72} & A_{73} & A_{74} & A_{75} & A_{76} & A_{77} & A_{78} & A_{79} \\ A_{81} & A_{82} & A_{83} & A_{84} & A_{85} & A_{86} & A_{87} & A_{88} & A_{89} \\ A_{91} & A_{92} & A_{93} & A_{94} & A_{95} & A_{96} & A_{97} & A_{98} & A_{99} \end{bmatrix}$$

and if you'd consider other points as well, you get (assuming the domain is periodic in both x and y, so that we don't have to worry yet about boundaries)

$$[A_1] = \frac{1}{\Delta x^2} \begin{bmatrix} -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 \end{bmatrix}$$

Beautiful as a .jpeg file at 0.0 cyan, 0.97 magenta, 0.14 yellow and 0.0 black. In general for this method, we'd end up at

$$[A_1] = \frac{1}{\Delta x^2}\begin{bmatrix} \ddots & \ddots & \ddots & & & & \\ & 1 & -2 & 1 & & & \\ & & 1 & -2 & 1 & & \\ & & & 1 & -2 & 1 & \\ & & & & \ddots & \ddots & \ddots \end{bmatrix}$$

where the -2s are on the diagonal of the matrix.

Now, how would $[A_2]$ look like? Again, we'd get

$$[A_2]\mathbf{u} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} & A_{16} & A_{17} & A_{18} & A_{19} \\ A_{21} & A_{22} & A_{23} & A_{24} & A_{25} & A_{26} & A_{27} & A_{28} & A_{29} \\ A_{31} & A_{32} & A_{33} & A_{34} & A_{35} & A_{36} & A_{37} & A_{38} & A_{39} \\ A_{41} & A_{42} & A_{43} & A_{44} & A_{45} & A_{46} & A_{47} & A_{48} & A_{49} \\ A_{51} & A_{52} & A_{53} & A_{54} & A_{55} & A_{56} & A_{57} & A_{58} & A_{59} \\ A_{61} & A_{62} & A_{63} & A_{64} & A_{65} & A_{66} & A_{67} & A_{68} & A_{69} \\ A_{71} & A_{72} & A_{73} & A_{74} & A_{75} & A_{76} & A_{77} & A_{78} & A_{79} \\ A_{81} & A_{82} & A_{83} & A_{84} & A_{85} & A_{86} & A_{87} & A_{88} & A_{89} \\ A_{91} & A_{92} & A_{93} & A_{94} & A_{95} & A_{96} & A_{97} & A_{98} & A_{99} \end{bmatrix}\begin{bmatrix} u_{1,1} \\ u_{2,1} \\ u_{3,1} \\ u_{1,2} \\ u_{2,2} \\ u_{3,2} \\ u_{1,3} \\ u_{2,3} \\ u_{3,3} \end{bmatrix}$$

$u_{2,2}$ is in the fifth row, thus you need to analyse the fifth row of $[A_1]$ as well. If we work out the dot product between the fifth row of $[A_1]$ and $\mathbf{u}$, we get

$$A_{51}u_{1,1} + A_{52}u_{2,1} + A_{53}u_{3,1} + A_{54}u_{1,2} + A_{55}u_{2,2} + A_{56}u_{3,2} + A_{57}u_{1,3} + A_{58}u_{2,3} + A_{59}u_{3,3}$$

Now, in the approximation, we have

$$\frac{u_{2,1} - 2u_{2,2} + u_{2,3}}{\Delta y^2}$$

and thus from comparison, we have

$$A_{52} = \frac{1}{\Delta y^2}, \quad A_{55} = \frac{-2}{\Delta y^2}, \quad A_{58} = \frac{1}{\Delta y^2}$$

and thus we'd get the matrix

$$[A_2] = \frac{1}{\Delta x^2}\begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} & A_{16} & A_{17} & A_{18} & A_{19} \\ A_{21} & A_{22} & A_{23} & A_{24} & A_{25} & A_{26} & A_{27} & A_{28} & A_{29} \\ A_{31} & A_{32} & A_{33} & A_{34} & A_{35} & A_{36} & A_{37} & A_{38} & A_{39} \\ A_{41} & A_{42} & A_{43} & A_{44} & A_{45} & A_{46} & A_{47} & A_{48} & A_{49} \\ 0 & 1 & 0 & 0 & -2 & 0 & 0 & 1 & 0 \\ A_{61} & A_{62} & A_{63} & A_{64} & A_{65} & A_{66} & A_{67} & A_{68} & A_{69} \\ A_{71} & A_{72} & A_{73} & A_{74} & A_{75} & A_{76} & A_{77} & A_{78} & A_{79} \\ A_{81} & A_{82} & A_{83} & A_{84} & A_{85} & A_{86} & A_{87} & A_{88} & A_{89} \\ A_{91} & A_{92} & A_{93} & A_{94} & A_{95} & A_{96} & A_{97} & A_{98} & A_{99} \end{bmatrix}$$

Again, disregarding the boundaries, this leads to

$$[A_2] = \begin{bmatrix} -2 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & -2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & -2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & -2 \end{bmatrix}$$

or in general,

$$[A_2] = \begin{bmatrix} \ddots & & \ddots & & \ddots & & & & \\ & 1 & \cdots & -2 & \cdots & 1 & & & \\ & & 1 & \cdots & -2 & \cdots & 1 & & \\ & & & 1 & \cdots & -2 & \cdots & 1 & \\ & & & & \ddots & & \ddots & & \ddots \end{bmatrix}$$

Now, there are some concepts to introduce here: notice that $[A_1]$ only had three adjacent columns in each row that had nonzero entries. This maximum number of columns between coefficients on a given is known as **bandwidth**. $[A_2]$ has a bandwidth of seven, for example (in general, for $[A_2]$, using this particular method[4], has a bandwidth that is twice number of mesh nodes in i (x) direction, plus one (we have three nodes in i direction[5] here, thus $2 \cdot 3 + 1$ which is indeed seven). Note that both $[A_1]$ and $[A_2]$ are very sparse (meaning that they contain a whole lot of zeroes). In the last chapter, we'll discover methods to quickly solve these systems (because honestly, row reduction would simply take too much time).

SETTING UP MATRICES OF ALGEBRAIC SYSTEMS, DISREGARDING BOUNDARY CONDITIONS

Given some PDE:
1. Find finite difference formulas for the partial derivatives.
2. Set up a vector **u** containing the values of u at each data point, ordering them in a logical manner, i.e. go through the mesh first in horizontal direction, then in vertical direction.
3. Set up as many equations as you need to get a feel of what the pattern of the matrix equation will be.
4. Set up the matrix equation.

3.1.3 *Example: general second-order PDE (skip if you're not interested)*

This section is in the book, but it's actually not part of the stuff you have to know for the first quiz (and also not the other quizzes): we don't need to know how to construct approximations for cross-derivatives (i.e. $\frac{\partial^2 u}{\partial x \partial y}$). Therefore, only read it if you're interested, because approximations for cross-derivatives suck more than a vacuum cleaner.

Suppose we'd extend the Laplace equation to

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial x \partial y} = 0$$

And let's use the same stencil as shown in figure 3.5. We already found the two first derivatives, but the cross-derivative needs to be determined. To be precise, the Taylor-series expansion for two variables is

$$\begin{aligned} u(x+\Delta x, t+\Delta t) \quad = \quad & u(x,t) + \frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial t}\Delta t + \frac{1}{2!}\cdot\frac{\partial^2 u}{\partial x^2}\Delta x^2 + \frac{1}{1!1!}\frac{\partial u}{\partial x \partial t}\Delta x \Delta t + \frac{1}{2!}\frac{\partial^2 u}{\partial t^2}\Delta t^2 \\ & + \frac{1}{3!}\frac{\partial^3 u}{\partial x^3}\Delta x^3 + \frac{1}{2!1!}\frac{\partial^3 u}{\partial x^2 \partial t}\Delta x^2 \Delta t + \frac{1}{1!2!}\frac{\partial^3 u}{\partial x \partial t^2}\Delta x \Delta t^2 + \frac{1}{3!}\frac{\partial^3 u}{\partial t^3}\Delta t^3 + \cdots \end{aligned}$$

and so it goes on (note the factorials for the cross-derivatives: you divide by the product of the factorials; the first factorial corresponds to how many times you differentiated in x-direction, and the second factorial to how many times you differentiated in y-direction. Now, what do you have to do, exactly, to find $\frac{\partial^2 u}{\partial x \partial t}$? Please note that the following box is probably a bit overwhelming, but don't worry, I'll apply it directly afterwards and you'll see it not as bad as it seems.

METHOD OF UNDETERMINED COEFFICIENTS TO FIND OPTIMAL DIFFERENTIATION SCHEME FOR CROSS-DIRECTIONAL DERIVATIVES

If one wants to know the cross-derivative $\frac{\partial^2 u}{\partial x \partial y}$ at the point (x_i, t_n) (i.e. u_i^n), and the following set of

[4] That means, this particular stencil.

[5] To be clear: three nodes in i direction in the total domain: if we'd had used a domain that's 10 nodes wide, but we still had used this stencil, than we'd have had a bandwidth of $2 \cdot 10 + 1$.

points (10 in total) is used:

$$\left(x_i + a_0\Delta x, t_n + a_1\Delta t\right), \quad \left(x_i + b_0\Delta x, t_n + b_1\Delta t\right), \ldots, \left(x_i + j_0\Delta x_0, t_n + j_1\Delta t\right)$$

then the matrix $[A]$ is given by

$$[A] = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ a_0 & b_0 & c_0 & d_0 & e_0 & f_0 & g_0 & h_0 & i_0 & j_0 \\ a_1 & b_1 & c_1 & d_1 & e_1 & f_1 & g_1 & h_1 & i_1 & j_1 \\ \frac{a_0^2}{2} & \frac{b_0^2}{2} & \frac{c_0^2}{2} & \frac{d_0^2}{2} & \frac{e_0^2}{2} & \frac{f_0^2}{2} & \frac{g_0^2}{2} & \frac{h_0^2}{2} & \frac{i_0^2}{2} & \frac{j_0^2}{2} \\ \frac{a_1^2}{2} & \frac{b_1^2}{2} & \frac{c_1^2}{2} & \frac{d_1^2}{2} & \frac{e_1^2}{2} & \frac{f_1^2}{2} & \frac{g_1^2}{2} & \frac{h_1^2}{2} & \frac{i_1^2}{2} & \frac{j_1^2}{2} \\ a_0a_1 & b_0b_1 & c_0c_1 & d_0d_1 & e_0e_1 & f_0f_1 & g_0g_1 & h_0h_1 & i_0i_1 & j_0j_1 \\ \frac{a_0^3}{3!} & \frac{b_0^3}{3!} & \frac{c_0^3}{3!} & \frac{d_0^3}{3!} & \frac{e_0^3}{3!} & \frac{f_0^3}{3!} & \frac{g_0^3}{3!} & \frac{h_0^3}{3!} & \frac{i_0^3}{3!} & \frac{j_0^3}{3!} \\ \frac{a_0^2a_1}{2} & \frac{b_0^2b_1}{2} & \frac{c_0^2c_1}{2} & \frac{d_0^2d_1}{2} & \frac{e_0^2e_1}{2} & \frac{f_0^2f_1}{2} & \frac{g_0^2g_1}{2} & \frac{h_0^2h_1}{2} & \frac{i_0^2i_1}{2} & \frac{j_0^2j_1}{2} \\ \frac{a_0a_1^2}{2} & \frac{b_0b_1^2}{2} & \frac{c_0c_1^2}{2} & \frac{d_0d_1^2}{2} & \frac{e_0e_1^2}{2} & \frac{f_0f_1^2}{2} & \frac{g_0g_1^2}{2} & \frac{h_0h_1^2}{2} & \frac{i_0i_1^2}{2} & \frac{j_0j_1^2}{2} \\ \frac{a_1^3}{3!} & \frac{b_1^3}{3!} & \frac{c_1^3}{3!} & \frac{d_1^3}{3!} & \frac{e_1^3}{3!} & \frac{f_1^3}{3!} & \frac{g_1^3}{3!} & \frac{h_1^3}{3!} & \frac{i_1^3}{3!} & \frac{j_1^3}{3!} \end{bmatrix}$$

If more points are used, then for each extra value of u you add, you need to add an extra column, and the next few rows would look like (using the first column):

$$\begin{bmatrix} \frac{a_0^4}{4!} \\ \frac{a_0^3a_1}{3!1!} \\ \frac{a_0^2a_1^2}{2!2!} \\ \frac{a_0a_1^3}{1!3!} \\ \frac{a_1^4}{4!} \end{bmatrix}$$

Then, the cross-derivative will be the linear combination

$$\frac{\partial^2 u}{\partial x \partial y} = k_1 u\left(x_i + a_0\Delta x, t_n + a_1\Delta t\right) + \ldots + k_{10} u\left(x_i + j_0\Delta x_0, t_n + j_1\Delta t\right)$$

where $k_1, ..., k_{10}$ are determined by row reducing the matrix equation

$$A\mathbf{k} = \mathbf{f}$$

where $\mathbf{f}$ has zeros everywhere except for the fifth row, where the entry is $1/(\Delta x \Delta y)$.

Now, there are important remarks to be made here:

- I honestly don't reckon they'll ask you for a stencil that consists of 10 points to be considered for your partial derivative, so don't start killing yourself just yet.
- The matrix $[A]$ can either be have 6 rows, 10 rows, 15 rows, 21 rows, 28 rows, etc. Even if you only get only four points, you still need to include 6 rows into A.
- This is also the precise reason why you can't always say that $\mathbf{k} = [A]^{-1}\mathbf{f}$; it may be that $[A]$ is not square meaning it's not invertible.

Now, let us apply this to the stencil of figure 3.5. We will be using the points on the diagonals to compute $\frac{\partial^2 u}{\partial x \partial y}$ at $u_{i,j}$. In other words, we will be using the points

$$\left(x_i + \Delta x, y_j + \Delta y\right), \quad \left(x_i + \Delta x, y_j 0\Delta y\right), \quad \left(x_i - \Delta x, y_j - \Delta y\right), \quad \left(x_i - \Delta x, y_j + \Delta y\right)$$

meaning that $a_0 = 1$, $a_1 = 1$, $b_0 = 1$, $b_1 = -1$, $c_0 = -1$, $c_1 = -1$, $d_0 = -1$ and $d_1 = 1$. This means that our matrix becomes

$$[A] = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ \frac{1^2}{2} & \frac{1^2}{2} & \frac{(-1)^2}{2} & \frac{(-1)^2}{2} \\ 1\cdot 1 & 1\cdot -1 & -1\cdot -1 & -1\cdot 1 \\ \frac{1^2}{2} & \frac{(-1)^2}{2} & \frac{(-1)^2}{2} & \frac{1^2}{2} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 1 & -1 & 1 & -1 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

so that we get

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 1 & -1 & 1 & -1 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Unfortunately, my TI-84+ is unable to row reduce this matrix equation as there are more rows than columns. Instead, what you need to see is that the first, fourth and sixth row are row equivalent: the fourth and sixth row are simply the first row divided by 2. Thus, we can delete them, to get the matrix equation

$$\begin{aligned} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{\Delta x \Delta y} \end{bmatrix} \\ \mathbf{k} &= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{\Delta x \Delta y} \end{bmatrix} = \begin{bmatrix} \frac{1}{4\Delta x \Delta y} \\ -\frac{1}{4\Delta x \Delta y} \\ \frac{1}{4\Delta x \Delta y} \\ -\frac{1}{4\Delta x \Delta y} \end{bmatrix} \end{aligned}$$

In other words, we have

$$\frac{\partial^2 u}{\partial x \partial y} = \frac{u_{i+1,j+1} - u_{i+1,j-1} + u_{i-1,j-1} - u_{i-1,j+1}}{4\Delta x \Delta y}$$

This means that

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial x \partial y} = \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{\Delta x^2} + \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{\Delta y^2} + \frac{u_{i+1,j+1} - u_{i+1,j-1} + u_{i-1,j-1} - u_{i-1,j+1}}{4\Delta x \Delta y} = 0$$

The matrix equation will look like

$$\left([A_1] + [A_2] + [A_3]\right) \mathbf{u} = \mathbf{0}$$

Note that we already determine $[A_1]$ and $[A_2]$ in the previous section. For $[A_3]$, the derivation is exactly analogous; suppose we again merely have a 3×3 domain, than we'd get for $u_{2,2}$:

$$A_{51}u_{1,1} + A_{52}u_{2,1} + A_{53}u_{3,1} + A_{54}u_{1,2} + A_{55}u_{2,2} + A_{56}u_{3,2} + A_{57}u_{1,3} + A_{58}u_{2,3} + A_{59}u_{3,3}$$

Now, in the approximation, we have

$$\frac{u_{3,3} - u_{3,1} + u_{1,1} - u_{1,3}}{4\Delta x \Delta y}$$

Thus clearly,

$$A_{51} = \frac{1}{4\Delta x \Delta y}, \quad A_{53} = -\frac{1}{4\Delta x \Delta y}, \quad A_{57} = -\frac{1}{\Delta x \Delta y}, \quad A_{59} = \frac{1}{4\Delta x \Delta y}$$

and all other entries equal to zero, meaning we get

$$[A_3] = \frac{1}{4\Delta x \Delta y}\begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} & A_{16} & A_{17} & A_{18} & A_{19} \\ A_{21} & A_{22} & A_{23} & A_{24} & A_{25} & A_{26} & A_{27} & A_{28} & A_{29} \\ A_{31} & A_{32} & A_{33} & A_{34} & A_{35} & A_{36} & A_{37} & A_{38} & A_{39} \\ A_{41} & A_{42} & A_{43} & A_{44} & A_{45} & A_{46} & A_{47} & A_{48} & A_{49} \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ A_{61} & A_{62} & A_{63} & A_{64} & A_{65} & A_{66} & A_{67} & A_{68} & A_{69} \\ A_{71} & A_{72} & A_{73} & A_{74} & A_{75} & A_{76} & A_{77} & A_{78} & A_{79} \\ A_{81} & A_{82} & A_{83} & A_{84} & A_{85} & A_{86} & A_{87} & A_{88} & A_{89} \\ A_{91} & A_{92} & A_{93} & A_{94} & A_{95} & A_{96} & A_{97} & A_{98} & A_{99} \end{bmatrix}$$

You could again extend this to the other rows as well, yielding

$$[A_3] = \frac{1}{4\Delta x \Delta y}\begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

which you can generalize yourself. Note that the bandwidth is equal to 9 in this case (equal to 2 times the total number of points in i direction plus three).

3.1.4 Appendum

Now, the method I came up with, with the matrix and all that, is actually also the method the reader uses. However, the lecturer wants you to write down slightly more. For example, consider the following problem: we have a boundary point, u_i. We want to know $\frac{\partial u}{\partial x}$ at this boundary point, by using the boundary point u_i and the two points to the left of it, as shown in figure 3.7.

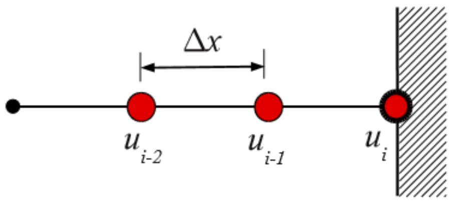


Figure 3.7: A finite-difference stencil for a boundary solution value.

We'd be ending up at (using a, b and c rather than a_1, a_2 and a_3)

$$\frac{\partial u}{\partial x} = a u_{i-2,j} + b u_{i-1,j} + c u_{i,j}$$

i.e. $k_1 = -2$, $k_2 = -1$ and $k_3 = 0$. We'd thus get

$$A = \begin{bmatrix} \frac{(-2)^0}{0!} & \frac{(-1)^0}{0!} & \frac{0^0}{0!} \\ \frac{(-2)^1}{1!} & \frac{(-1)^1}{1!} & \frac{0^1}{1!} \\ \frac{(-2)^2}{2!} & \frac{(-1)^2}{2!} & \frac{0^2}{2!} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -2 & -1 & 0 \\ 2 & \frac{1}{2} & 0 \end{bmatrix}$$

and then you'd have your other vector $\mathbf{f}$ equal to

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

However, the lecturer wants you to write this in a Taylor table, as shown in table 3.1.

Table 3.1: Taylor table for computing a one-sided 3-point expression for the first derivative. The interior columns (between the double vertical bars) contain the numerical coefficients for three Taylor expansions about the point i.

	$a \cdot u_{i-2}$	$b \cdot u_{i-1}$	$c \cdot u_i$	
u_i	a	b	c	0
$\Delta x \left.\frac{\partial u}{\partial x}\right\rvert_i$	$a(-2)$	$b(-1)$	0	1
$\Delta x^2 \left.\frac{\partial^2 u}{\partial x^2}\right\rvert_i$	$a\frac{(-2)^2}{2!}$	$b\frac{(-1)^2}{2!}$	0	0
$\Delta x^3 \left.\frac{\partial^3 u}{\partial x^3}\right\rvert_i$	$a\frac{(-2)^3}{3!}$	$b\frac{(-1)^3}{3!}$	0	

Please note the differences and similarities to what I've showed you so far:

EXPANDING THE MATRIX TO TAYLOR TABLE

- You have to add a first column, which contains a rather logical sequence for uni-directional derivatives. but I don't reckon they'll ask for cross-derivatives to be quite honest.
- You have to add a row on top with $a \cdot u_{i-2}$ etc.
- For the matrix I previously did not include a, b and c directly, you have to do that now.
- You have to add the vector $\mathbf{f}$ at the right-side of the table.
- You have to add an extra row (which follows the same pattern as the previous rows); this row will allow us to determine the leading term in the truncation error (and thus the order of the truncation error). Note that we write nothing in the vector $\mathbf{f}$ for this extra row.

NOTE REGARDING THE ORDER OF ACCURACY

If n nodes are used, then the highest order derivative that can be approximated is of order $n-1$.

If n nodes are used and the pth derivative is to be approximated, the order of accuracy will be $n-p$, except when the nodes are symmetrically distributed around the node at which the derivative is approximated; in that case:

- If p is even, then the order of accuracy will be $n-p+1$.
- If p is odd, then the order of accuracy remains $n-p$.

This shouldn't be all too difficult once you have done one or two examples with it. Now, how does one exactly go about determining the leading error in the truncation error? Well, one first solves for a, b and c, using the first three rows:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -2 & 1 & 0 \\ 2 & \frac{1}{2} & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -2 \\ \frac{3}{2} \end{bmatrix}$$

Now, to find the leading term in the truncation error, substitute these values into the last row, i.e. you get

$$\frac{1}{2} \cdot \frac{(-2)^3}{3!} + 2 \cdot \frac{(-1)^3}{3!} = \frac{-1}{3}$$

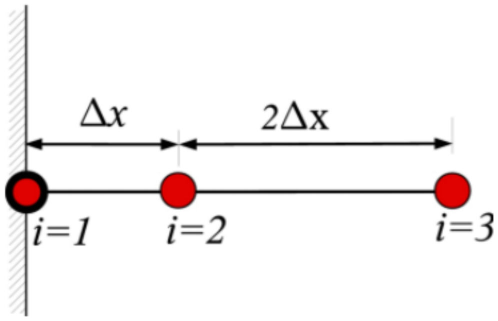
Then, multiply this with the term that's in front of this row (i.e. $\Delta x^3 \frac{\partial^3 u}{\partial x^3}$) to get $-\Delta x^3 \frac{\partial^3 u}{\partial x^3}$ and divide by the term that's in front of the derivative you were interested in (we were interested in $\frac{\partial u}{\partial x}$, and in the left column, there's a Δx in front of it, so we need to divide by Δx). This means that we end up at a leading term in the truncation error of

$$-\Delta x^2 \frac{\partial^3 u}{\partial x^3}$$

In other words, the truncation error is of order of magnitude $\mathcal{O}\left(\Delta x^2\right)$.

Quiz 1: Question 14

A Taylor table with entries labelled as shown below is to be used to construct an approximation for the second derivative $\frac{\partial^2 u}{\partial x^2}$ at the boundary point $i = 1$.



The approximation is to make use of the data at points $i = 1$, $i = 2$ and $i = 3$, and be as accurate as possible.

	$a \cdot u_i$	$b \cdot u_2$	$c \cdot u_3$	
u_i	A_2	A_3	A_4	A_5
B_1	B_2	B_3	B_4	B_5
C_1	C_2	C_3	C_4	C_5
D_1	D_2	D_3	D_4	

a. The entry in location B_5 of the table should be:
   - 1
   - 0
   - Δx
   - $2\Delta x$
   - Δx^2
   - Δx^3

b. The entry in location D_1 of the table shold be:
   - $\Delta x^3 \left(\frac{\partial^3 u}{\partial x^3}\right)_1$
   - $\Delta x^2 \left(\frac{\partial^2 u}{\partial x^2}\right)_1$
   - $\frac{\Delta x^2}{2} \left(\frac{\partial^2 u}{\partial x^2}\right)_1$
   - $\frac{2\Delta x^2}{3!} \left(\frac{\partial^2 u}{\partial x^2}\right)_1$
   - u_3
   - u_4

- The entry in location C_4 of the table should be:

Let's just construct the entire table ourselves first. The matrix I suggested would have $k_1 = 0$, $k_2 = 1$ and $k_3 = 3$, and thus we'd get, including the extra row which we'll be needing later:

$$\begin{bmatrix} \frac{0^0}{0!} & \frac{1^0}{0!} & \frac{3^0}{0!} \\ \frac{0^1}{1!} & \frac{1^1}{1!} & \frac{3^1}{1!} \\ \frac{0^2}{2!} & \frac{1^2}{2!} & \frac{3^2}{2!} \\ \frac{0^3}{3!} & \frac{1^3}{3!} & \frac{3^3}{3!} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & \frac{1}{2} & \frac{9}{2} \\ 0 & \frac{1}{6} & \frac{9}{2} \end{bmatrix}$$

Now, adding the stuff that the lecturer wants you to add, we get the beautiful table

	$a \cdot u_i$	$b \cdot u_2$	$c \cdot u_3$	
u_i	$a \cdot 1$	$b \cdot 1$	$c \cdot 1$	0
$\Delta x \left(\frac{\partial u}{\partial x}\right)_i$	$a \cdot 0$	$b \cdot 1$	$c \cdot 3$	0
$\Delta x^2 \left(\frac{\partial^2 u}{\partial x^2}\right)_i$	$a \cdot 0$	$b \cdot \frac{1}{2}$	$c \cdot \frac{9}{2}$	1
$\Delta x^3 \left(\frac{\partial^3 u}{\partial x^3}\right)_i$	$a \cdot 0$	$b \cdot \frac{1}{6}$	$c \cdot \frac{9}{2}$	

or simply

	$a \cdot u_i$	$b \cdot u_2$	$c \cdot u_3$	
u_i	a	b	c	0
$\Delta x \left(\frac{\partial u}{\partial x}\right)_i$	0	b	$3c$	0
$\Delta x^2 \left(\frac{\partial^2 u}{\partial x^2}\right)_i$	0	$\frac{b}{2}$	$\frac{9c}{2}$	1
$\Delta x^3 \left(\frac{\partial^3 u}{\partial x^3}\right)_i$	0	$\frac{b}{6}$	$\frac{9c}{2}$	

In other words:

a. The correct answer is 0.
b. The correct answer is $\Delta x^3 \left(\frac{\partial^3 u}{\partial x^3}\right)_i$.
c. The correct answer is $\frac{9c}{2}$. Note that you should not enter $\frac{3^3}{3!} \cdot c$; you need to work it out as much as possible. $4.5c$ is also accepted, but I don't recommend using rounded numbers in Maple (just make it a habit that it's always a fraction).

Examples 1: Q1

Derive an approximation for the third derivative at i of the form:

$$\frac{\left(au_{i-2} + bu_{i-1} + cu_i + du_{i+1} + eu_{i+2}\right)}{\Delta x^3} = \left.\frac{\partial^3 u}{\partial x^3}\right|_i + TE$$

with maximum accuracy. What is the leading term of the truncation error (TE)?

We have $k_1 = -2$, $k_2 = -1$, $k_3 = 0$, $k_4 = 1$ and $k_5 = 2$, meaning our matrix, with the extra row, becomes

$$\begin{bmatrix} \frac{(-2)^0}{0!} & \frac{(-1)^0}{0!} & \frac{0^0}{0!} & \frac{1^0}{0!} & \frac{2^0}{0!} \\ \frac{(-2)^1}{1!} & \frac{(-1)^1}{1!} & \frac{0^1}{1!} & \frac{1^1}{1!} & \frac{2^1}{1!} \\ \frac{(-2)^2}{2!} & \frac{(-1)^2}{2!} & \frac{0^2}{2!} & \frac{1^2}{2!} & \frac{2^2}{2!} \\ \frac{(-2)^3}{3!} & \frac{(-1)^3}{3!} & \frac{0^3}{3!} & \frac{1^3}{3!} & \frac{2^3}{3!} \\ \frac{(-2)^4}{4!} & \frac{(-1)^4}{4!} & \frac{0^4}{4!} & \frac{1^4}{4!} & \frac{2^4}{4!} \\ \frac{(-2)^5}{5!} & \frac{(-1)^5}{5!} & \frac{0^5}{5!} & \frac{1^5}{5!} & \frac{2^5}{5!} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ 2 & \frac{1}{2} & 0 & \frac{1}{2} & 2 \\ -\frac{4}{3} & -\frac{1}{6} & 0 & \frac{1}{6} & \frac{4}{3} \\ \frac{2}{3} & \frac{1}{24} & 0 & \frac{1}{24} & \frac{2}{3} \\ -\frac{4}{15} & -\frac{1}{120} & 0 & \frac{1}{120} & \frac{4}{15} \end{bmatrix}$$

so that upon extending this to a Taylor table, we get

	$a \cdot u_{i-2}$	$b \cdot u_{i-1}$	$c \cdot u_i$	$d \cdot u_{i+1}$	$e \cdot u_{i+2}$	
u_i	a	b	c	d	e	0
$\Delta x \left.\frac{\partial u}{\partial x}\right\|_i$	$-2a$	$-b$	0	d	e	0
$\Delta x^2 \left.\frac{\partial^2 u}{\partial x^2}\right\|_i$	$2a$	$\frac{b}{2}$	0	$\frac{d}{2}$	$2e$	0
$\Delta x^3 \left.\frac{\partial^3 u}{\partial x^3}\right\|_i$	$-\frac{4a}{3}$	$-\frac{b}{6}$	0	$\frac{d}{6}$	$\frac{4e}{3}$	1
$\Delta x^4 \left.\frac{\partial^4 u}{\partial x^4}\right\|_i$	$\frac{2a}{3}$	$\frac{b}{24}$	0	$\frac{d}{24}$	$\frac{2e}{3}$	0
$\Delta x^5 \left.\frac{\partial^5 u}{\partial x^5}\right\|_i$	$-\frac{4a}{15}$	$-\frac{b}{120}$	0	$\frac{d}{120}$	$\frac{4e}{15}$	

Yeah I honestly no idea how they can expect us to solve the system by hand, so I'll just use my calculator for this:

$$\begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ 2 & \frac{1}{2} & 0 & \frac{1}{2} & 2 \\ -\frac{4}{3} & -\frac{1}{6} & 0 & \frac{1}{6} & \frac{4}{3} \\ \frac{2}{3} & \frac{1}{24} & 0 & \frac{1}{24} & \frac{2}{3} \\ -\frac{4}{15} & -\frac{1}{120} & 0 & \frac{1}{120} & \frac{4}{15} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{-1}{2} \\ 1 \\ 0 \\ -1 \\ \frac{1}{2} \end{bmatrix}$$

Thus, we have

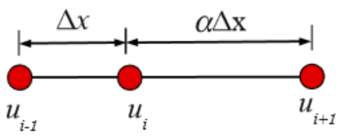
$$\left.\frac{\partial^3 u}{\partial x^3}\right|_i \approx \frac{-u_{i-2} + 2u_{i-1} - 2u_{i+1} + u_{i+1}}{2\Delta x^3}$$

Furthermore, the leading term in the truncation error will be

$$\left(-\frac{4a}{15} - \frac{b}{120} + 0 + \frac{d}{120} + \frac{4e}{15}\right) \cdot \frac{\Delta x^5 \left.\frac{\partial^5 u}{\partial x^5}\right|_i}{\Delta x^3} = \frac{\Delta x^2}{4} \left.\frac{\partial^5 u}{\partial x^5}\right|_i$$

Examples 1: Q2

Derive an approximation for the second derivative $\frac{\partial^2 u}{\partial x^2}$ at i on the uneven mesh shown below. What is its order of accuracy?



We have $k_1 = -1$, $k_2 = 0$ and $k_3 = \alpha$. Thus, we get the matrix (including the extra row)

$$\begin{bmatrix} \frac{(-1)^0}{0!} & \frac{0^0}{0!} & \frac{\alpha^0}{0!} \\ \frac{(-1)^1}{1!} & \frac{0^1}{1!} & \frac{\alpha^1}{1!} \\ \frac{(-1)^2}{2!} & \frac{0^2}{2!} & \frac{\alpha^2}{2!} \\ \frac{(-1)^3}{3!} & \frac{0^3}{3!} & \frac{\alpha^3}{3!} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & \alpha \\ \frac{1}{2} & 0 & \frac{\alpha^2}{2} \\ \frac{-1}{6} & 0 & \frac{\alpha^3}{6} \end{bmatrix}$$

In other words, we get the expanded table

	$a \cdot u_{i-1}$	$b \cdot u_i$	$c \cdot u_{i+1}$	
u_i	a	b	c	0
$\Delta x \left.\frac{\partial u}{\partial x}\right\|_i$	$-a$	0	αc	0
$\Delta x^2 \left.\frac{\partial^2 u}{\partial x^2}\right\|_i$	$\frac{a}{2}$	0	$\frac{\alpha^2 c}{2}$	1
$\Delta x^3 \left.\frac{\partial^3 u}{\partial x^3}\right\|_i$	$-\frac{a}{6}$	0	$\frac{\alpha^3 c}{6}$	

Now, finding a, b and c is slightly fucked up: we have to solve

$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & \alpha \\ \frac{1}{2} & 0 & \frac{\alpha^2}{2} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Now, α is not known, so we have to do it manually. Key is that you first only analyse

$$\begin{bmatrix} -1 & \alpha \\ \frac{1}{2} & \frac{\alpha^2}{2} \end{bmatrix} \begin{bmatrix} a \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

In other words: we leave out the middle column (and thus b) and then we can also leave out a row, namely the first row. We then have

$$\left[\begin{array}{cc|c} -1 & \alpha & 0 \\ 0 & \frac{\alpha^2+\alpha}{2} & 1 \end{array}\right]$$

Thus,

$$c = \frac{1}{\frac{\alpha^2+\alpha}{2}} = \frac{2}{\alpha^2+\alpha}$$

Furthermore,

$$\begin{aligned} -a + \alpha c &= 0 \\ -a + \alpha \cdot \frac{2}{\alpha^2+\alpha} &= 0 \\ a &= \frac{2}{\alpha+1} \end{aligned}$$

Then, from the first row of the 3×3 matrix, we have

$$\begin{aligned} a + b + c &= 0 \\ \frac{2}{\alpha+1} + b + \frac{2}{\alpha^2+\alpha} &= 0 \\ b &= -\frac{2}{\alpha^2+\alpha} - \frac{2}{\alpha+1} = -\frac{2}{\alpha^2+\alpha} - \frac{2\alpha}{\alpha^2+\alpha} = \frac{-2(\alpha+1)}{\alpha(\alpha+1)} \\ &= \frac{-2}{\alpha} \end{aligned}$$

Thus, we have

$$\frac{\partial^2 u}{\partial x^2} \approx \frac{\frac{2}{\alpha+1}u_{i-1} - \frac{2}{\alpha}u_i + \frac{2}{\alpha^2+\alpha}u_{i+1}}{\Delta x^2} = \frac{2\alpha u_{i-1} - 2(1+\alpha)u_i + 2u_{i+1}}{\alpha(1+\alpha)\Delta x^2}$$

The leading term of the truncation error is them found by

$$\begin{aligned} TE &= \left(-\frac{a}{6} + \frac{\alpha^3 c}{6}\right) \cdot \frac{\Delta x^3 \left.\frac{\partial^3 u}{\partial x^3}\right|_i}{\Delta x^2} \\ &= \left(-\frac{2}{6(\alpha+1)} + \frac{\alpha^3 \cdot 2}{6\alpha(\alpha+1)}\right) \Delta x \left.\frac{\partial^3 u}{\partial x^3}\right|_i \\ &= \left(\frac{-2\alpha + 2\alpha^3}{6\alpha(\alpha+1)}\right) \Delta x \left.\frac{\partial^3 u}{\partial x^3}\right|_i \\ &= \frac{\alpha^2+1}{3(\alpha+1)} \Delta x \left.\frac{\partial^3 u}{\partial x^3}\right|_i = \frac{(\alpha+1)(\alpha-1)}{3(\alpha+1)} \Delta x \left.\frac{\partial^3 u}{\partial x^3}\right|_i \\ &= \frac{\Delta x(\alpha-1)}{3} \left.\frac{\partial^3 u}{\partial x^3}\right|_i \end{aligned}$$

which means that the accuracy is only first order when $\gamma \neq 1$.

3.2 Boundaries

Now, I've said a few times, ignore boundaries, so let's now talk about it.

Consider, for instance, the implicit discretisation for the linear advection equation as described in section 3.1.1,

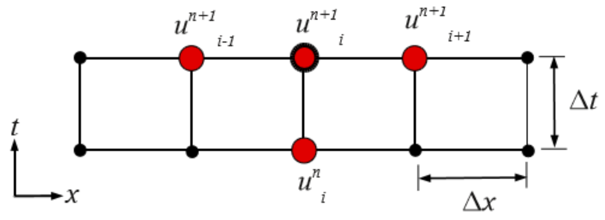


Figure 3.8: Finite-difference stencil for the implicit Euler-method.

which was derived using the stencil shown in figure 3.8. Remember that we were able to write (for a domain which consisted of 7 nodes in x direction, so that the first and seven nodes are boundaries)

$$([I] + [D])\,\mathbf{u}^{n+1} = \mathbf{u}^n$$

with $[I]$ the identity matrix,

$$\mathbf{u}^{n+1} = \begin{bmatrix} u_1^{n+1} \\ u_2^{n+1} \\ u_3^{n+1} \\ \vdots \end{bmatrix}, \qquad \mathbf{u}^n = \begin{bmatrix} u_1^n \\ u_2^n \\ u_3^n \\ \vdots \end{bmatrix}$$

and $[D]$ given by

$$[D] = \frac{c\Delta t}{2\Delta x}\begin{bmatrix} 0 & 1 & & & & & -1 \\ -1 & 0 & 1 & & & & \\ & -1 & 0 & 1 & & & \\ & & -1 & 0 & 1 & & \\ & & & -1 & 0 & 1 & \\ & & & & -1 & 0 & 1 \\ 1 & & & & & -1 & 0 \end{bmatrix}$$

or,

$$\begin{bmatrix} 1 & \mu & & & & & -\mu \\ -\mu & 1 & \mu & & & & \\ & -\mu & 1 & \mu & & & \\ & & -\mu & 1 & \mu & & \\ & & & -\mu & 1 & \mu & \\ & & & & -\mu & 1 & \mu \\ \mu & & & & & -\mu & 1 \end{bmatrix}\begin{bmatrix} u_1^{n+1} \\ u_2^{n+1} \\ u_3^{n+1} \\ u_4^{n+1} \\ u_5^{n+1} \\ u_6^{n+1} \\ u_7^{n+1} \end{bmatrix} = \begin{bmatrix} u_1^n \\ u_2^n \\ u_3^n \\ u_4^n \\ u_5^n \\ u_6^n \\ u_7^n \end{bmatrix}$$

with $\mu = \frac{c\Delta t}{2\Delta x}$. Now, the first and final row need to be adjusted so that you don't use values on the other side of the domain (because now, u_1^{n+1} uses u_7^n in its calculation, which is obviously bullshit if the domain is not periodic). Now, remember the following boundary conditions we could use:

3.2.1 Physical: Dirichlet

Dirichlet boundaries were easy: you just specified the values of u on the entire boundary, using a function $g(x)$. Thus, that'd mean that at the boundary,

$$u_1^{n+1} = g^{n+1}$$

or in other words, the matrix would become

$$\begin{bmatrix} 1 & & & & & & \\ -\mu & 1 & \mu & & & & \\ & -\mu & 1 & \mu & & & \\ & & -\mu & 1 & \mu & & \\ & & & -\mu & 1 & \mu & \\ & & & & -\mu & 1 & \mu \\ \mu & & & & & -\mu & 1 \end{bmatrix} \begin{bmatrix} u_1^{n+1} \\ u_2^{n+1} \\ u_3^{n+1} \\ u_4^{n+1} \\ u_5^{n+1} \\ u_6^{n+1} \\ u_7^{n+1} \end{bmatrix} = \begin{bmatrix} g^{n+1} \\ u_2^n \\ u_3^n \\ u_4^n \\ u_5^n \\ u_6^n \\ u_7^n \end{bmatrix}$$

Note that I did not change the other boundary, at u_7^{n+1} just yet: the reason for this is simple. Remember the analysis of the characteristics of a linear advection equation: we only were allowed to specify one physical condition on the left boundary; we weren't allowed to do so on the right boundary. Instead, we must do something different there, which we'll see after the next subsection.

3.2.2 Physical: Neumann

Neumann conditions specified the derivative u at the boundary, i.e. $\frac{\partial u}{\partial x} = q$. How would we implement this? Well, we simply find a new difference scheme to come up with an approximation of $\frac{\partial u}{\partial x}$; for example, using a very similar approach as in section 3.1.4, one could find

$$\left.\frac{\partial u}{\partial x}\right|_1 = \frac{(-3u_1 + 4u_2 - u_3)}{2\Delta x} = q$$

using the two points to the right of this boundary. We could implement this simply in our matrix equation:

$$u_1^{n+1} = g^{n+1}$$

or in other words, the matrix would become

$$\begin{bmatrix} -\frac{3}{2\Delta x} & \frac{2}{\Delta x} & \frac{-1}{2\Delta x} & & & & \\ -\mu & 1 & \mu & & & & \\ & -\mu & 1 & \mu & & & \\ & & -\mu & 1 & \mu & & \\ & & & -\mu & 1 & \mu & \\ & & & & -\mu & 1 & \mu \\ \mu & & & & & -\mu & 1 \end{bmatrix} \begin{bmatrix} u_1^{n+1} \\ u_2^{n+1} \\ u_3^{n+1} \\ u_4^{n+1} \\ u_5^{n+1} \\ u_6^{n+1} \\ u_7^{n+1} \end{bmatrix} = \begin{bmatrix} q^{n+1} \\ u_2^n \\ u_3^n \\ u_4^n \\ u_5^n \\ u_6^n \\ u_7^n \end{bmatrix}$$

3.2.3 Numerical

Now, this still leaves us with a problem what to do with the right boundary. We can't apply one of the physical conditions here, so we must do something else. Fortunately, it's much easier than you think: the problem for u_7^{n+1} was that in the stencil, we used a point that was to the right of. Remember what we were doing the entire time: for the equation

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0$$

we tried to find finite difference schemes for the partial derivatives. For $\frac{\partial u}{\partial t}$, we used the points u_i^n and u_i^{n+1} to estimate $\frac{\partial u}{\partial t}$ at u_i^{n+1}, leading to

$$\frac{\partial u}{\partial t} = \frac{u_i^{n+1} - u_i^n}{\Delta t}$$

Note that clearly, this expression is still allowed to be used, as neither of the points lay outside the domain. However, for $\frac{\partial u}{\partial x}$, we used the points u_{i-1}^{n+1}, u_i^{n+1} and u_{i+1}^{n+1}, leading to

$$\frac{\partial u}{\partial x} = \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x}$$

Now, this one is problematic for the boundary point u_7^{n+1}: this would mean u_8^{n+1} is used, which doesn't exist. Howevr, we can just use a different stencil, specifically for this boundary point, only utilizing points to the left of u_7^{n+1}; one could use the points u_5^{n+1}, u_6^{n+1} and u_7^{n+1} to derive

$$\frac{\partial u}{\partial x} = \frac{u_5^{n+1} - 4u_6^{n+1} + 3u_7^{n+1}}{2\Delta x}$$

This would mean that the PDE could be written as

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = \frac{u_7^{n+1} - u_7^n}{\Delta t} + c\frac{u_5^{n+1} - 4u_6^{n+1} + 3u_7^{n+1}}{2\Delta x} = 0$$

We can rewrite this to

$$u_7^{n+1} + \frac{c\Delta t}{2\Delta x}u_5^{n+1} - \frac{2c\Delta t}{\Delta x}u_6^{n+1} + \frac{3c\Delta t}{2\Delta x}u_7^{n+1} = u_7^n$$

This means that the final row of the matrix becomes

$$\begin{bmatrix} -\frac{3}{2\Delta x} & \frac{2}{\Delta x} & \frac{-1}{2\Delta x} & & & & \\ -\mu & 1 & \mu & & & & \\ & -\mu & 1 & \mu & & & \\ & & -\mu & 1 & \mu & & \\ & & & -\mu & 1 & \mu & \\ & & & & -\mu & 1 & \mu \\ & & & & \frac{c\Delta t}{2\Delta x} & \frac{-2c\Delta t}{\Delta x} & \left(1 + \frac{3c\Delta t}{2\Delta x}\right) \end{bmatrix} \begin{bmatrix} u_1^{n+1} \\ u_2^{n+1} \\ u_3^{n+1} \\ u_4^{n+1} \\ u_5^{n+1} \\ u_6^{n+1} \\ u_7^{n+1} \end{bmatrix} = \begin{bmatrix} q^{n+1} \\ u_2^n \\ u_3^n \\ u_4^n \\ u_5^n \\ u_6^n \\ u_7^n \end{bmatrix}$$

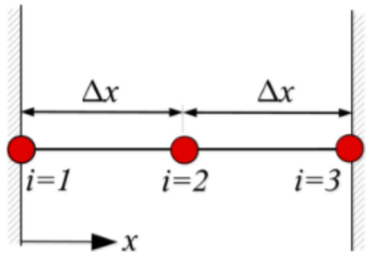
Please understand what we've done in this section:

- Dirichlet boundary conditions are really straightforward and really shouldn't be too difficult to understand.
- Neumann boundary conditions use the *derivative* of u to set the boundary condition. The actual partial differential equation is totally ignored.
- Numerical boundary conditions use the exact same approach as all other points in the domain, however, you just simply use a different stencil to find your finite difference scheme. That's the only thing it means.

DEALING WITH BOUNDARY CONDITIONS

- Physical boundaries will be given to you, and it will be stated which nodes in the mesh need to obey these conditions:
  - For Dirichlet conditions, replace the row corresponding corresponding to the node at which the boundary $u = g$ is applied with the equation corresponding to $u = g$ (this will be extremely straightforward).
  - For Neumann conditions, find a finite difference approximation for $\frac{\partial u}{\partial x}$, and replace the row corresponding to the node at which the boundary is applied with the corresponding finite difference approximation.
- Numerical boundaries may or may not be given to you. However, whenever a node would use points that lay outside of its domain (i.e., it'd be implicitly assuming that the domain is periodic even though it isn't), you need to come up with a new finite difference method that would only use nodes within the boundaries. Then, replace the corresponding row with this new finite difference method. Note that you now have to use the *entire* PDE, and not just $\frac{\partial u}{\partial x}$.

Quiz 1: Q17



The linear convection equation $\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0$ is to be approximated on the domain shown above using the following finite-difference expression:

$$\left(u_i^{n+1} - u_i^n\right) + p\left(u_i^{n+1} - u_{i-1}^{n+1}\right) = 0, \qquad \text{with } p = \frac{c\Delta t}{\Delta x}$$

On the left boundary the condition $u(0) = \sin(t)$ is to be applied. On the right boundary, a first-order accurate numerical boundary condition is to be used. This results in an algebraic system with entries labelled as shown below.

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} u_1^{n+1} \\ u_2^{n+1} \\ u_3^{n+1} \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix}$$

To answer the following questions, you will need to find expressions for the entries in the matrix A and right-hand-side vector B in terms of p.

Note, however, that the rows of the algebraic system are still valid when multiplied by a constant. Therefore, if you construct your matrix in an unusual way, a slight adjustment might be required to fit the answers. As a check, for each of the rows 1, 2 and 3 in matrix A, adding the column values should produce 1.

a. The entry in location B_1 of the right-hand side vector should be:
   - 0
   - 1
   - $\sin(t^n)$
   - $\sin\left(t^{n+1}\right)$
   - u_1^n
   - pu_2^n
   - u_2^{n+1}
   - u_3^n
   - p
b. The entry in location B_3 of the right-hand side vector should be:
   - 1
   - 0
   - u_2^n
   - u_3^n
   - pu_2^n
   - pu_3^n
   - $\sin(t^n)$
   - $\sin\left(t^{n+1}\right)$
c. The entry in location A_{22} of the matrix should be:
d. The entry in location A_{32} of the matrix should be:

Let's just construct the entire matrix equation on our own. First, we rewrite the finite-difference expression to

$$u_i^{n+1} + p\left(u_i^{n+1} - u_{i-1}^{n+1}\right) = u_i^n$$

Now, for u_1^{n+1}, there is the condition $u(0) = \sin(t)$. In other words,

$$u_1^{n+1} = \sin\left(t^{n+1}\right)$$

and this will be our first equation. For the other two, we can simply apply our finite-difference formula:

$$\begin{aligned} u_2^{n+1} + p\left(u_2^{n+1} - u_1^{n+1}\right) &= u_2^n \\ u_3^{n+1} + p\left(u_3^{n+1} - u_2^{n+1}\right) &= u_3^n \end{aligned}$$

Note that this poses no problem of the right boundary, u_3^{n+1}, as we don't use any points outside of the boundary. This means that we have the set of equations

$$\begin{aligned} u_1^{n+1} &= \sin\left(t^{n+1}\right) \\ -pu_1^{n+1} + (p+1)\,u_2^{n+1} &= u_2^n \\ -pu_2^{n+1} + (p+1)\,u_3^{n+1} &= u_3^n \end{aligned}$$

or

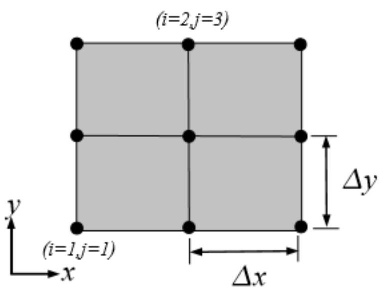
$$\begin{bmatrix} 1 & 0 & 0 \\ -p & p+1 & 0 \\ 0 & -p & p+1 \end{bmatrix}\begin{bmatrix} u_1^{n+1} \\ u_2^{n+1} \\ u_3^{n+1} \end{bmatrix} = \begin{bmatrix} \sin\left(t^{n+1}\right) \\ u_2^n \\ u_3^n \end{bmatrix}$$

Thus, the correct answers are:

a. $B_1 = \sin\left(t^{n+1}\right)$.
b. $B_3 = u_3^n$.
c. $A_{22} = 1 + p$.
d. $A_{32} = -p$.

Examples 1: Q3

Laplace's equation $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$ is to be approximated on the rectangular 3×3-node mesh shown below $\Delta x = \Delta y = h =$ constant.

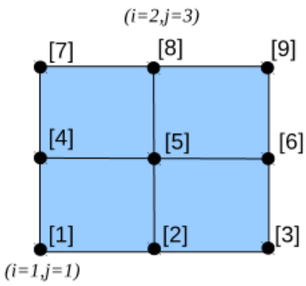


The following finite-difference approximations are to be used:

$$\begin{aligned} \left.\frac{\partial^2 u}{\partial x^2}\right|_{i,j} &= \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} \\ \left.\frac{\partial^2 u}{\partial y^2}\right|_{i,j} &= \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{h^2} \\ \left.\frac{\partial u}{\partial x}\right|_{i=3,j} &= \frac{u_{3,j} - u_{2,j}}{h} \end{aligned}$$

On the lower boundary ($j = 1$) u is set to one, on the left boundary ($i = 1, j = 2$), u is set to be two, and on the upper boundary ($j = 3$), u is set to be three. On the right boundary ($i = 3, j = 2$), the gradient of the solution in the x-direction is to be set to zero. What is the resulting algebraic system?

First, let's use the numbering system shown below.



Note that the algebraic system will have the form

$$\left(\left[A_1\right]+\left[A_2\right]\right)\mathbf{u}=0$$

where $[A_1]$ and $[A_2]$ follow from the approximation formulas. Now, before we do anything else, note that from the boundary conditions alone, we can already obtain a shitload of equations. From the fact that on the lower boundary u is set to one, we obtain

$$\begin{aligned} u_1 &= 1 \\ u_2 &= 1 \\ u_3 &= 1 \end{aligned}$$

From the fact that at $(i = 1, j = 2)$, $u = 2$, we get

$$u_4 = 2$$

From the fact that at the upper boundary, u is set to three, we get

$$\begin{aligned} u_7 &= 3 \\ u_8 &= 3 \\ u_9 &= 3 \end{aligned}$$

So, we only need equations for u_5 and u_6, really. For u_5, we simply use the approximation formulas given to write

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=\frac{u_6-2u_5+u_4}{h^2}+\frac{u_8-2u_5+u_4}{h^2}=0$$

For u_6, we have to apply the given Neumann condition:

$$\left.\frac{\partial u}{\partial x}\right|_{i=3,j=2}=0$$

Using the given approximation formula, we have

$$\left.\frac{\partial u}{\partial x}\right|_{i=3,j=2}=\frac{u_6-u_5}{h}=0$$

Thus, using these 9 equations, we get

$$\begin{bmatrix} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & 1 & & & & & & \\ & & & 1 & & & & & \\ & 1 & & 1 & -4 & 1 & & 1 & \\ & & & & -1 & 1 & & & \\ & & & & & & 1 & & \\ & & & & & & & 1 & \\ & & & & & & & & 1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \\ u_7 \\ u_8 \\ u_9 \end{bmatrix}=\begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 0 \\ 0 \\ 3 \\ 3 \\ 3 \end{bmatrix}$$

3.3 *Artificial dissipation and upwinding*

Now, suppose we look at the linear advection equation (i.e. the simple transportation of a wave in one direction). Suppose we have the square wave initial condition as denoted by the blue line in figure 3.9. The implicit Euler method we derived before (using the stencil shown in figure 3.10),

$$\frac{u_i^{n+1}-u_i^n}{\Delta t}+c\frac{u_{i+1}^n-u_{i-1}^n}{2\Delta x}=0$$

has a small problem with this: in a perfect world, after a certain amount of time, we would expect the wave to be shifted to the right a bit, but that it still looks exactly the same otherwise. However, although the implicit Euler method we derived does shit it to the right, it creates some rather obvious errors, apparent from the green line in figure 3.9.

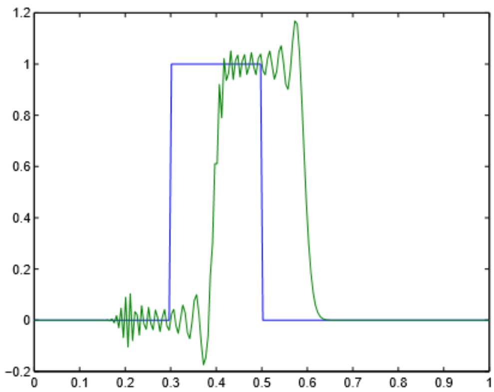


Figure 3.9: Finite-difference stencil for an implicit method.

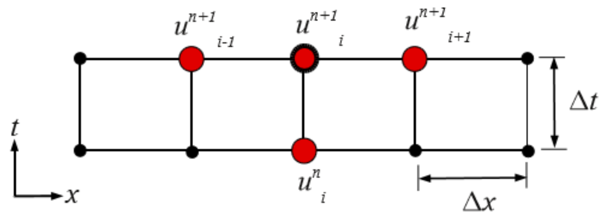


Figure 3.10: Finite-difference stencil for the implicit Euler-method.

Now, why is this the case? Well, let's first establish the fact that the implicit Euler method we derived is *central in space*. What do I mean by this? The *spatial* derivative (i.e. the derivative in x (or possibly y) direction) is evaluated in a 'central' way: you could have evaluated the spatial derivative using only nodes that are on one side of the central point (in this case, u_i^{n+1}), e.g. only the points that are on the right-side of it (so you'd only use u_{i+1}^{n+1} and u_{i+1}^n to derive an approximation for $\frac{\partial u}{\partial x}$, you wouldn't be using u_{i-1}^{n+1}). However, in our previous derivation, we used a nice symmetric distribution so that the node around which we are expanding is indeed the center. Therefore, the method we derived was central in space. Soon, we'll see methods that are not central in space (Q14 of Quiz 1 is also not central in space, fyi). Now, why is the fact that it is central in space such a problem? Well, suppose you are a point that lays very close to the cliff formed by the square wave: if you look to the right of you, you see a point which has $u = 0$, whereas to the left of you, there'll be a point where $u = 1$. What value do you take? Considering you are a point, you don't have a lot of brain cells, so you don't know what to do, and you just panic and just take some value (between approximately 0 and 1 in ofc, you're not that stupid that you suddenly think your value should be $u = 2248373$). This is basically what creates the weird behaviour near the sudden jumps of the wave: the points don't really know what they're supposed to do, so you get some weird oscillations.

Now, that kind of error as shown in figure 3.9, where the graph oscillates wildly, are called **dispersive errors**. How do we deal with them? There are two methods for that.

3.3.1 *Artificial dissipation*

Well, first of all, you could expand the scheme to

$$\frac{u_i^{n+1}-u_i^n}{\Delta t}+c\frac{u_{i+1}^n-u_{i-1}^n}{2\Delta x}-k\Delta x^2\frac{\left(u_{i-1}^{n+1}-2u_i^{n+1}+u_{i+1}^{n+1}\right)}{\Delta x^2} = 0$$

where k is a user-specified constant. Why exactly are you allowed to add this term, and why would you include it in the first place? Well, first of all, note that

$$\frac{\left(u_{i-1}^{n+1}-2u_i^{n+1}+u_{i+1}^{n+1}\right)}{\Delta x^2}$$

is actually the formula we derived in section 3.1.2 for $\frac{\partial^2 u}{\partial x^2}$, where we had

$$\frac{\partial^2 u}{\partial x^2}=\frac{u_{i-1,j}-2u_{i,j}+u_{i+1,j}}{\Delta x^2}$$

So, what we're essentially just doing is that we include the curvature of u in x direction, and we can include this (with the addition of $k\Delta x^2$ in front of it) for the following reason: generally speaking, Δx and k will both be small, meaning that this extra term will be very small if the curvature is small (so that it's influence is basically negligible), but if the curvature is very large (which is the case at those cliffs of the square wave), then this term will actually have an influence: your point, which lays very near this cliff, now has more information about what to do, so it won't have a total panic attack but actually will behave more sensibly. How will it act more sensibly?

Well, note that we actually now have an approximation for

$$\frac{\partial u}{\partial t}+c\frac{\partial u}{\partial x}-k\Delta x^2\frac{\partial^2 u}{\partial x^2}=0$$

where k and Δx are some user-specified constants. Note that this is very similar to the linear advection-diffusion equation:

$$\frac{\partial u}{\partial t}+c\frac{\partial u}{\partial x}-v\frac{\partial^2 u}{\partial x^2}=0$$

And yes indeed, our solution is basically a linear advection-difussion solution, meaning that our graph after a certain time will look like as what is shown in figure 3.11: the square wave is both transported to the right, but it's also diffused: however, as $k\Delta x^2$ is very small, this diffusion is very small, meaning it's still very similar to a square wave.

So: yes, by adding this term, we are essentially solving a linear advection-diffusion equation instead of a linear advection equation: however, because our viscosity v is very, very small, the diffusion does not have a very pronounced effect. However, it is sufficient to make sure that the dispersive errors don't occur.

Now, to conclude this subsection, let me introduce two terms: first of all, the term we added is called the **artificial dissipation** term. Furthermore, the error caused by the smoothing of the wave is called the **dissipation error**.

3.3.2 *Upwind*

Now, there's another method of dealing with this issue (not necessarily better, btw). Remember what the problem was basically: because the spatial derivative we used was central in space, your point had to comprehend very conflicting information: one coming from the dude to the left of him that was on the wave, and one to the right of him that was not on the wave, so should the point in the middle be on the wave or not?

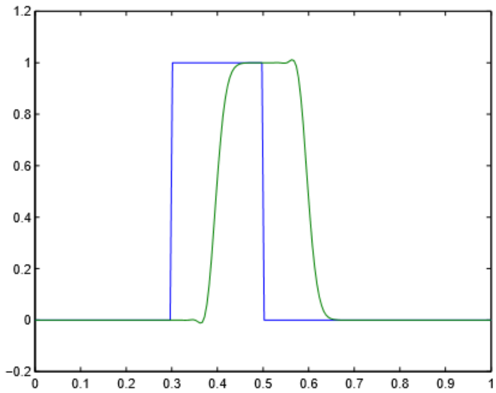


Figure 3.11: Numerical solution for the linear advection equation with a square wave initial condition produced by an implicit Euler in time, central in space discretisation with artificial dissipation.

Well, one way to solve this is to use a spatial derivative that is not central in space, but is **biased**. If we use the following approximation:

$$u_i^{n+1} + \frac{c\Delta t}{\Delta x}\left(u_i^{n+1} - u_{i-1}^{n+1}\right) = u_i^n$$

then we are only looking at the point to the left of u_i^{n+1} to evaluate the spatial derivative, and thus it is said to be biased (in that direction). Why is this helpful? Well, your point now only looks to the left, and he only sees, hey that point is already on the wave, I should probably join in as well, and the behaviour becomes much smoother. It doesn't care any more what the point to the right of him does. Now, why do we use the point to the left of u_i^{n+1} and not the one to the right of it? Well, remember that the wave is travelling to the right, in other words, your future is approaching from the left. It makes sense to look to the future to compute what will happen to you in the future (god this is such a philosophical course).

The result is shown in figure 3.12. Note that the result is even more diffusive than the one of artificial dissipation. Why is this? Well, basically, we can rewrite the spatial derivative term (you can verify yourself that these are equal):

$$\underbrace{c\frac{\left(u_i^{n+1} - u_{i-1}^{n+1}\right)}{\Delta x}}_{\text{upwind}} = \underbrace{c\frac{\left(u_{i+1}^{n+1} - u_{i-1}^{n+1}\right)}{2\Delta x}}_{\text{central}} - \underbrace{c\frac{\Delta x}{2}\frac{\left(u_{i-1}^{n+1} - 2u_i^{n+1} + u_{i+1}^{n+1}\right)}{\Delta x^2}}_{\text{artificial dissipation}}$$

What does this tell us? Well, it tells us that we basically added an artificial dissipation term just like we did in the previous subsection; however, this time, the viscosity is[6]

$$\upsilon = \frac{\Delta t}{\Delta x} \cdot c\frac{\Delta x}{2} = \frac{c\Delta t}{2}$$

where c is not necessarily large. Comparing with $k\Delta x^2$, $c\Delta t$ will generally speaking be larger thus the diffusive effect is much larger as well.

However, do note, this analogy cannot always be made (using upwind schemes (i.e. schemes that only use points upstream) does not always indirectly result in adding artificial dissipation terms). Furthermore, there's another nice benefit regarding upwind schemes: you can also apply it on the right-boundary of the domain, meaning you don't have to apply a numerical boundary there (which was a small problem in section 3.2). When you're using an upwind scheme, at the right boundary, you'll still be using points that lay inside the domain so you don't need to change stuff.

[6]Where does the $\Delta t/\Delta x$ come from? We only replaced $c\left(u_i^{n+1} - u_{i-1}^{n+1}\right)$, so we still need to multiply with $\Delta t/\Delta x$ to get the viscosity.

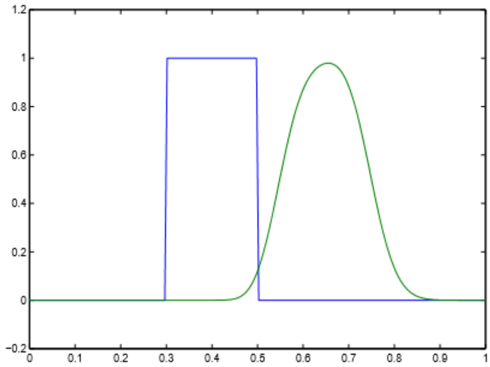


Figure 3.12: Numerical solution for the linear advection equation with a square wave initial condition produced by an implicit Euler in time, upwind in space discretisation.

Quiz 1: Q15

The most suitable artificial viscosity operator for an explicit discretisation of the linear convection equation is:

- $k\frac{\Delta x\left(u_{i-1}^n - u_{i+1}^n\right)}{2\Delta x}$
- $k\frac{\Delta x^2\left(u_{i-1}^n - 2u_i^n + u_{i+1}^n\right)}{\Delta x^2}$
- $k\frac{\left(u_{i-1}^n - 2u_i^n + u_{i+1}^n\right)}{\Delta x^2}$
- $k\frac{\left(u_{i-1}^n - u_{i+1}^n\right)}{2\Delta x}$

Correct is $k\frac{\Delta x^2\left(u_{i-1}^n - 2u_i^n + u_{i+1}^n\right)}{\Delta x^2}$. The first and fourth equation are discretisations of $\frac{\partial u}{\partial x}$, not of $\frac{\partial^2 u}{\partial x^2}$. Furthermore, you want Δx^2 multiplying the fraction to make it as small as possible.

Quiz 1: Q16

All upwind finite-difference operators for the linear convection equation

- are less accurate than central difference operators.
- are dissipative.
- can be used for numerical boundary conditions.

Correct is **can be used for numerical boundary conditions.** The first one is not always the case (if you simply include more terms in your upwind operator than it'll automatically become more accurate eventually); above derivation about the dissipation is only true for this particular upwind operator, it's not always true that you can write it like this. However, you can always use it for numerical boundary conditions.

3.4 *Dealing with irregular meshes*

As you've probably noticed, not everything is a rectangle in real life. For example, one could use the mesh shown in figure 3.13: clearly not rectangular.

In this section, we'll learn how to deal with them?

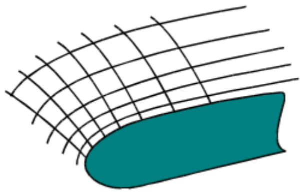


Figure 3.13: Mesh refinement near a curved boundary.

3.4.1 *Operators for unequal mesh spacing*

If the problem is merely that the mesh is unequally spaced, then we can simply apply the same method as done in Examples 1: Q2. Yes, solving the system sucks, but in real life, we can usually use computer programs for that.

3.4.2 *The generalised transformation*

We can also come up with a more fancy, general transformation procedure. Let us start by analysing the discretisation of

$$\frac{\partial u}{\partial x}+\frac{\partial u}{\partial y}=0$$

on a curvilinear mesh, as shown in figure 3.14: we know the x and y coordinates of each point in the mesh, but we need to transform these curvilinear lines to Cartesian grid lines in the (ξ,η) plane. We thus need to find transformations of the form

$$\begin{aligned}\xi &= \xi(x,y)\\ \eta &= \eta(x,y)\end{aligned}$$

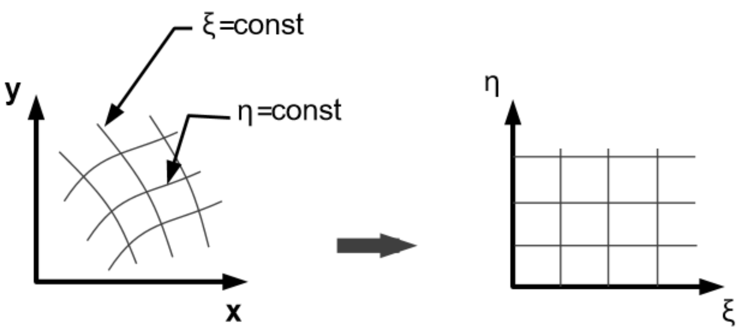


Figure 3.14: Physical and transformed coordinates.

The PDE can then be simply solved on this Cartesian domain, using methods discussed before. How exactly would our PDE now look like, if we're using ξ and η instead of x and y? Well, from the chain rule,

$$\begin{aligned}\frac{\partial u}{\partial x} &= \frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial x}+\frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial x}\\ \frac{\partial u}{\partial y} &= \frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial y}+\frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial y}\end{aligned}$$

and let me define

METRICS

The **metrics** of a transformation are denoted by

$$\xi_x = \frac{\partial \xi}{\partial x}, \quad \xi_y = \frac{\partial \xi}{\partial y}, \quad \eta_x = \frac{\partial \eta}{\partial x}, \quad \eta_y = \frac{\partial \eta}{\partial y} \tag{3.1}$$

Note that we also have

$$x_\xi = \frac{\partial x}{\partial \xi}, \quad x_\eta = \frac{\partial x}{\partial \eta}, \quad y_\xi = \frac{\partial y}{\partial \xi}, \quad y_\eta = \frac{\partial y}{\partial eta} \tag{3.2}$$

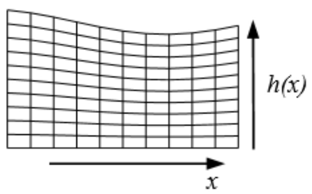
Then we can write

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = \frac{\partial u}{\partial \xi}\xi_x + \frac{\partial u}{\partial \eta}\eta_x + \frac{\partial u}{\partial \xi}\xi_y + \frac{\partial u}{\partial \eta}\eta_y = 0$$

Just to clarify: $\frac{\partial u}{\partial \xi}$ and $\frac{\partial u}{\partial \eta}$ are definitely not known beforehand, typically speaking! These are the things you'd find an approximation scheme for (which you could easily do, using the methods I explained before in this chapter (after all, in the (ξ, η) plane, it's a simple rectangular mesh). If we know functions for the transformations, we can easily compute ξ_x, η_x, ξ_y and η_y, however.

Examples 1: Q4

The physical domain shown below:



is transformed to a computational (Cartesian) domain by:

$$\xi = x \qquad \eta = \frac{y}{h(x)}$$

where $h(x)$ is the known distance between the upper and lower boundary. Derive an expression to compute the first derivative of a solution $\frac{\partial u}{\partial x}$ on the computational domain. Assume that values of $\frac{\partial u}{\partial \xi}$ and $\frac{\partial u}{\partial \eta}$ are obtained through finite-difference approximations and are given values.

We can use the chain rule to write

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial x} + \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial x}$$

Now, $\frac{\partial u}{\partial \xi}$ and $\frac{\partial u}{\partial \eta}$ are given values, so we don't need to obtain it ourselves. On the other hand, from the given transformations, we clearly have (for the second one, it's simply the quotient rule)

$$\begin{aligned} \frac{\partial \xi}{\partial x} &= 1 \\ \frac{\partial \eta}{\partial x} &= y \cdot \frac{-h'(x)}{(h(x))^2} \end{aligned}$$

and thus

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial \xi} - \frac{\partial u}{\partial \eta} y \frac{h'(x)}{h^2(x)}$$

However, we have $y = \eta \cdot h(x)$ (from the given transformation), thus this can be written as

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial \xi} - \frac{\partial u}{\partial \eta}\eta\frac{h'(x)}{h(x)}\frac{\partial u}{\partial \eta}$$

To be clear: $\frac{\partial u}{\partial \xi}$ and $\frac{\partial u}{\partial \eta}$ can be found by applying a finite-difference scheme: literally the only difference with the methods described in section 3.1 would be that you now continuously have $\Delta\xi$ and $\Delta\eta$ instead of Δx and Δy.

Now, the problem is, usually, we don't know the exact transformation function most of the time, so it's hard to come up with ξ_x, η_x, etc. So, how do we come up with them? Well, consider the following. Using the chain rule once more, we can write

$$\begin{aligned} dx &= x_\xi d\xi + x_\eta d\eta \\ dy &= y_\xi d\xi + y_\eta d\eta \end{aligned}$$

or, in matrix form:

$$\begin{aligned} \mathbf{x} &= A\boldsymbol{\xi} \\ \begin{bmatrix} dx \\ dy \end{bmatrix} &= \begin{bmatrix} x_\xi & x_\eta \\ y_\xi & y_\eta \end{bmatrix} \begin{bmatrix} d\xi \\ d\eta \end{bmatrix} \end{aligned}$$

In an exactly similar way, we can derive

$$\begin{aligned} d\xi &= \xi_x dx + \xi_y dy \\ d\eta &= \eta_x dx + \eta_y dy \end{aligned}$$

or, in matrix form:

$$\begin{aligned} \boldsymbol{\xi} &= B\mathbf{x} \\ \begin{bmatrix} d\xi \\ d\eta \end{bmatrix} &= \begin{bmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix} \end{aligned}$$

Carefully compare these matrix equations: if we have $x = A\boldsymbol{\xi}$ and $\boldsymbol{\xi} = B\mathbf{x}$, then it must be that $B = A^{-1}$. In other words

$$\begin{aligned} B &= A^{-1} \\ \begin{bmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{bmatrix} &= \begin{bmatrix} x_\xi & x_\eta \\ y_\xi & y_\eta \end{bmatrix}^{-1} = J^{-1} \begin{bmatrix} y_\eta & -x_\eta \\ -y_\xi & x_\xi \end{bmatrix} \end{aligned}$$

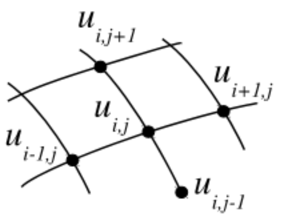
where $J = x_\xi y_\eta - y_\xi x_\eta$ is called the Jacobian of the transformation. Why did I show you all of this? Well, clearly, if we know x_ξ, x_η, y_ξ and y_η, ξ_x, ξ_y etc. are really easy to compute; it's simply calculating the inverse. Now, how do we find x_ξ etc.? Well, you can just use expressions such as

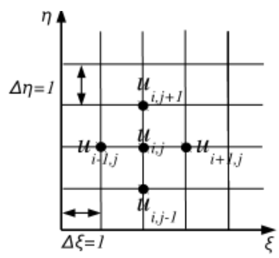
$$x_\xi = \frac{x_{i+1,j} - x_{i-1,j}}{2\Delta\xi}$$

for that. Typically, we'll choose $\Delta\xi = 1$ (and the same goes for $\Delta\eta$. Not convinced? Let's do an example:

Examples 1: Q5

Using the concept of a generalised transformation and central finite differences, estimate a numerical value for $\left.\frac{\partial u}{\partial x}\right|_{(i,j)}$ with the data given below.





node	x	y	u
(i,j)	10	9	7
$(i+1,j)$	15	8	10
$(i-1,j)$	6	11	3
$(i,j+1)$	8	13	8
$(i,j-1)$	12	4	4

Note that

$$\begin{bmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{bmatrix} = J^{-1} \begin{bmatrix} y_\eta & -x_\eta \\ -y_\xi & x_\xi \end{bmatrix} \qquad \text{where } J = \left(x_\xi y_\eta - y_\xi x_\eta\right)$$

First of, again we write

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial x} + \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial x} = u_\xi \xi_x + u_\eta \eta_x$$

Now, we can use the table to compute u_ξ and u_η:

$$\begin{aligned} u_\xi &= \frac{u_{i+1,j} - u_{i-1,j}}{2\Delta\xi} = \frac{10-3}{2\cdot 1} = \frac{7}{2} \\ u_\eta &= \frac{u_{i,j+1} - u_{i,j-1}}{2\Delta\eta} = \frac{8-4}{2\cdot 1} = 2 \end{aligned}$$

To compute ξ_x and η_x, we have to apply that matrix shit. We get:

$$\begin{aligned} x_\xi &= \frac{x_{i+1,j} - x_{i-1,j}}{2\Delta\xi} = \frac{15-6}{2\cdot 1} = \frac{9}{2} \\ x_\eta &= \frac{x_{i,j+1} - x_{i,j-1}}{2\Delta\eta} = \frac{8-12}{2\cdot 1} = -2 \\ y_\xi &= \frac{y_{i+1,j} - y_{i-1,j}}{2\Delta\xi} = \frac{8-11}{2} = -\frac{3}{2} \\ y_\eta &= \frac{y_{i,j+1} - y_{i,j-1}}{2\Delta\eta} = \frac{13-4}{2} = \frac{9}{2} \end{aligned}$$

Thus,

$$J = \left(x_\xi y_\eta - y_\xi x_\eta\right) = \frac{9}{2}\cdot\frac{9}{2} - -2\cdot -\frac{3}{2} = \frac{69}{4}$$

$$\begin{bmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{bmatrix} = J^{-1} \cdot \begin{bmatrix} y_\eta & -x_\eta \\ -y_\xi & x_\xi \end{bmatrix} = \left(\frac{69}{4}\right)^{-1} \begin{bmatrix} \frac{9}{2} & 2 \\ \frac{3}{2} & \frac{9}{2} \end{bmatrix}$$

Then, from comparison, we see that

$$\begin{aligned} \xi_x &= \frac{4}{69}\cdot\frac{9}{2} = \frac{18}{69} \\ \eta_x &= \frac{4}{69}\cdot\frac{3}{2} = \frac{6}{69} \end{aligned}$$

Thus,

$$\frac{\partial u}{\partial x} = u_\xi \xi_x + u_\eta \eta_x = \frac{7}{2}\cdot\frac{18}{69} + 2\cdot\frac{6}{69} = \frac{25}{23}$$

Absolutely beautiful stuff. Do note: normally, u_ξ and u_η would be approximated by their own finite-difference schemes, using methods described in section 3.1.