PHY2006 - Partial Differential Equations

Contents

Library Books

Linear Partial Differential Equations and Fourier Theory - Marcus Pivato, Cambridge University Press (2012)
QUB Library E-book - https://doi-org.queens.ezp1.qub.ac.uk/10.1017/CBO9780511810183

Numerical Solution of Partial Differential Equations - An Introduction, K. W. Morton, D. F. Mayers, Cambridge University Press (2005)
QUB Library E-book - https://doi-org.queens.ezp1.qub.ac.uk/10.1017/CBO9780511812248

Online notes

Numerical Solutions

http://www.damtp.cam.ac.uk/lab/people/sd/lectures/nummeth98/index.htm#L_1_Title_Page

http://wwwf.imperial.ac.uk/~mdavis/FDM11/LECTURE_SLIDES2.PDF

http://www.cs.man.ac.uk/~fumie/tmp/introductory-finite-difference-methods-for-pdes.pdf

1. Introduction/Revision

1.1. Ordinary Differential Equations (ODEs)

Although ODEs are not the direct focus of this part of the course, they are naturally related to partial differential equations and it will be assumed you have a good grasp of what ODEs are and how they are solved. Dr. Tom Field's *PHY1002 Differential Equations Synopsis*, gives an excellent, concise summary on ODEs so only some key points about ODEs are summarised here.

Consider a function $u = f(x)$, u is the dependent variable and is x the independent variable. Ordinary differential equations have a single independent variable, which the dependent variable is differentiated with respect to. Differential equations are classified with the following definitions:

Order

This is the maximum number of times a function has been differentiated in the equation

$a\frac{du}{dx} + b = c$ (1st order)

$a\frac{d^2u}{dx^2} + b\frac{du}{dx} + cu = 0$ (2nd order)

$\frac{d^3u}{dt^3} + \omega_0^2\frac{du}{dt} = 0$ (3rd order)

$\left(\frac{du}{dx}\right)^2 + cu = x$ (1st order)

Linearity

Linearity is a very important property in Maths and Physics, in particular it allows the principle of superposition to be invoked, e.g. when considering inferences of waves. In a linear differential equation the dependent variable and its differentials must only appear with a power of 1 (i.e. must not be raised to any other power). The last equation shown above is non-linear as the differential is squared. A general linear equation has the following form

$$A(x)\frac{d^2u}{dx^2} + B(x)\frac{du}{dx} + C(x)u = f(x)$$

Note that the functions of the independent variable $A(x), B(x), C(x), f(x)$ are not necessarily linear.

Linear differential equations are much easier to solve than non-linear equations and solutions obey the principle of superposition. If two solutions u_1, u_2 of the linear ODE have been found, then a linear combination of these is also a solution

$$u = au_1 + bu_2$$

where a and b are constants.

Solutions

The methods by which differential equations are solved depends on the type of the equation (order, linearity etc.). Often an analytical solution is not possible and an approximate or numerical solution is required. You covered a number of different methods in PHY1002. We are not going to go through all these again in detail but here are a few ways in which common ODEs can be solved.

1st order - variable separable

If a first order equation is in the form $\frac{du}{dx} = f(x)g(u)$, then the variable can be separated and integrated as follows

$$\int \frac{du}{g(u)} = \int f(x)dx$$

Consider the following example

$$\frac{du}{dx} = 2u(x+1)$$

In this case $f(x) = 2(x+1)$ and $g(u) = u$ hence

$$\int \frac{du}{u} = 2\int (x+1)dx$$

$$\ln u = x^2 + 2x + c$$

$$u = e^c \exp(x^2 + 2x) = A\exp\big(x(x+2)\big)$$

1st order – integrating factor

If a first order equation is in the form $\frac{du}{dx} + f(x)u = g(x)$, the integrating factor is defined as $\exp(\int f(x)dx)$. For example consider the equation

$$\frac{du}{dx} + 2xu = e^{-x^2}$$

The integrating factor is $\exp(\int 2xdx) = \exp(x^2)$. Multiplying through by this factor gives

$$e^{-x^2}\frac{du}{dx} + 2xue^{-x^2} = 1$$

which can be written

$$\frac{d}{dx}\left[ue^{-x^2}\right] = 1$$

Integrating gives

$$ue^{-x^2} = x + c$$

$$u = e^{x^2}(x+c)$$

2nd order constant coefficients

If $f(x) = 0$, the differential equation is known as ***homogeneous***. If $A(x) = a,\ B(x) = b,\ C(x) = c$ are constants then

$$a\frac{d^2u}{dx^2} + b\frac{du}{dx} + cu = 0$$

The solution to this equation is called the ***complementary function***. It can be found by letting $y = \exp(mx)$, which when substituted gives a quadratic equation known as the ***characteristic equation***

$$am^2 + bm + c = 0$$

With two roots. The solution has the general form

$$u = C_1 e^{m_1 x} + C_2 e^{m_2 x}$$

If the roots are complex then wave solutions (imaginery exponentials) are obtained while real roots give exponential functions. If $m_1 = m_2$ then a solution of the form

$$u = (C_1 x + C_2) e^{m_1 x}$$

To solve a linear, ***inhomogeneous ODE*** **($f(x) \neq 0$)**, a particular integral is obtained along with the complementary function (by setting $f(x) = 0$).

Using linearity, the superposition of these two solutions give the general solution of the ODE.

General Solution = Complementary Function + Particular Integral

Power Series Solutions

There a numerous different ways in which solutions for linear ODEs with variable coefficients $A(x), B(x), C(x)$ are found by substituting different variables or other means. These techniques merit a course in themselves and is beyond what can be covered in PHY2006 – see for example https://tutorial.math.lamar.edu/Classes/DE/DE.aspx. However, it is worth mentioning the ***power series method*** which is can be used to solve many types of differential equations (including a few non-linear ones), not least because this is how the r and θ parts of the wavefunction for the hydrogen atom are obtained when solving the Schrödinger equation.

The solution is expressed as an infinite polynomial or power series

$$u(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \cdots$$

where a_n are constant coefficients. By substituting this into the differential equation, the goal is to find these coefficients. For some situations a solution in the form of a Taylor series $u(x) = \sum_{n=0}^{\infty} \frac{a_n x^n}{n!}$ or in the form $u(x) = x^k \sum_{n=0}^{\infty} a_n x^n$ (Frobenius method) is required or needed to simplify the algebra.

Let's take the simplest form of a 2nd order differential equation and try to solve it with the power series method (even though we know what the general solution is already)

$$\frac{d^2u}{dx^2} + u = 0$$

$$\frac{du}{dx} = a_1 + 2a_2 x + 3a_3 x^2 \ldots = \sum_{n=1}^{\infty} n a_n x^{n-1}$$

$$\frac{d^2u}{dx^2} = 2a_2 + 6a_3 x + 12a_4 x^2 \ldots = \sum_{n=2}^{\infty} n(n-1) a_n x^{n-2}$$

Note the change in the starting value of n. If we now let $i = n - 2$, then

$$\frac{d^2u}{dx^2} = \sum_{i=0}^{\infty} (i+2)(i+1) a_{i+2} x^i$$

Now the letter we use for the sum index does not matter, so if we change it back from i to n and put it into the ODE

$$\sum_{n=0}^{\infty} (n+2)(n+1) a_{n+2} x^n + \sum_{n=0}^{\infty} a_n x^n = 0$$

So if we equate each power x^n, then

$$(n+2)(n+1)a_{n+2} + a_n = 0$$

$$a_{n+2} = -\frac{a_n}{(n+2)(n+1)}$$

This is known as a ***recurrence relationship***, so that if we know the first coefficient we can obtain all others in the series. In this case as a_{n+2} and a_n are related, there are solutions based on odd and even powers of x starting with the values of a_0 and a_1

$$u(x) = a_0\left(1 - \frac{x^2}{(2)(1)} + \frac{x^4}{(4)(3)(2)(1)} - \cdots\right) = a_0\sum_{m=0}^{\infty}(-1)^m\frac{x^{2m}}{(2m)!} = a_0\cos x$$

$$u(x) = a_1\left(x - \frac{x^3}{(3)(2)(1)} + \frac{x^5}{(5)(4)(3)(2)(1)} - \cdots\right) = a_1\sum_{m=0}^{\infty}(-1)^m\frac{x^{2m+1}}{(2m+1)!} = a_1\sin x$$

which is just the general solution we expect

$$u(x) = a_0\cos x + a_1\sin x$$

Consider a more challenging example

$$(x^2+1)\frac{d^2u}{dx^2} - 4x\frac{du}{dx} + 6u = 0$$

and subsituting in the power series for each derivative

$$(x^2+1)\sum_{n=2}^{\infty}n(n-1)a_nx^{n-2} - 4x\sum_{n=1}^{\infty}na_nx^{n-1} + 6\sum_{n=0}^{\infty}a_nx^n = 0$$

$$\sum_{n=2}^{\infty}n(n-1)a_nx^2 + \sum_{n=2}^{\infty}n(n-1)a_nx^{n-2} - 4\sum_{n=1}^{\infty}na_nx^n + 6\sum_{n=0}^{\infty}a_nx^n = 0$$

If we shift the index of the 2nd term by 2 then

$$\sum_{n=2}^{\infty}n(n-1)a_nx^n + \sum_{n=0}^{\infty}(n+2)(n+1)a_{n+2}x^n - \sum_{n=1}^{\infty}4na_nx^n + \sum_{n=0}^{\infty}6a_nx^n = 0$$

So all powers in x of each of these sums is the same, it is just that the starting index is different. If we consider terms for $n = 0$ then

$$2a_2 + 6a_0 = 0$$

$$a_2 = -3a_0$$

For $n = 1$

$$6a_3 - 4a_1 + 6a_1 = 0$$

$$a_3 = -\frac{a_1}{3}$$

For all higher values of n

$$n(n-1)a_n + (n+2)(n+1)a_{n+2} - 4na_n + 6a_n = 0$$

$$(n+2)(n+1)a_{n+2} + a_n(n^2 - 5n + 6) = 0$$

$$a_{n+2} = -a_n \frac{(n-2)(n-3)}{(n+2)(n+1)}$$

Therefore this gives (for $n = 2, 3$)

$a_4 = 0$ and $a_5 = 0$, meaning that all higher power coefficients are also zero (the series is said to have been truncated). This gives solutions

$$u(x) = a_0 - 3a_0x^2$$

$$u(x) = a_1x - \frac{a_1}{3}x^3$$

so the general solution is

$$u(x) = a_0(1 - 3x^2) + a_1\left(x - \frac{1}{3}x^3\right)$$

This power series technique can be applied in a similar manner to higher order differentials.

Legendre's Differential Equation

This equation is particular common when considering 2nd order differential equations in spherical polar coordinates. The equation describing the dependence on angle θ tends to have the form

$$(1 - x^2)\frac{d^2u}{dx^2} - 2x\frac{du}{dx} + \beta u = 0$$

where the substitution $x = \sin\theta$ has been made (β is a constant). For example, a form of this equation must be solved to obtain the θ dependence of the wavefunction of the hydrogen atom. A power series solution gives

$$\sum_{n=2}^{\infty} n(n-1)a_nx^{n-2} - \sum_{n=2}^{\infty} n(n-1)a_nx^n - 2\sum_{n=1}^{\infty} na_nx^n + \sum_{n=0}^{\infty} \beta a_nx^n = 0$$

Shifting the index of the first summation by 2 gives

$$\sum_{n=0}^{\infty} (n+2)(n+1)a_{n+2}x^n - \sum_{n=2}^{\infty} n(n-1)a_nx^n - 2\sum_{n=1}^{\infty} na_nx^n + \sum_{n=0}^{\infty} \beta a_nx^n = 0$$

For $n = 0$

$$2a_2 + \beta a_0 = 0$$

$$a_2 = -\frac{\beta}{2}a_0$$

For $n = 1$

$$6a_3 - 2a_1 + \beta a_1 = 0$$

$$a_3 = -a_1\frac{\beta - 2}{6}$$

For all higher values of n

$$(n+2)(n+1)a_{n+2} - n(n-1)a_n - 2na_n + \beta a_n = 0$$

$$a_{n+2} = a_n\frac{n(n+1) - \beta}{(n+2)(n+1)}$$

As in the previous example the sum is truncated if there is a value of $n = l$ where $a_l = 0$. This happens when

$\beta = l(l+1)$ where $l = 0,1,2 \ldots$ is an integer.

This is the original of the orbital angular momentum quantum number for atoms. In order for the wavefunction to be well behaved the series must be finite and so must be truncated after a finite number of terms. These finite polynomials to a power of x^l are called Legendre polynomials and are either odd or even. Each value of l gives a different possible solution. For example if $l = 4$ ($\beta = 20$)

$$a_2 = -\frac{\beta}{2}a_0 = -10a_0$$

$$a_4 = a_2\frac{2(2+1)-\beta}{(2+2)(2+1)} = -\frac{7}{6}a_2 = \frac{35}{3}a_0$$

$$a_6 = 0, a_8 = 0, \ldots$$

$$u_4(x) = a_0\left(1 - 10x^2 + \frac{35}{3}x^4\right)$$

For $l = 3$ ($\beta = 12$)

$$a_3 = -a_1\frac{\beta-2}{6} = -\frac{5}{3}a_1$$

$$a_5 = 0, a_7 = 0, \ldots$$

$$u_3(x) = a_1\left(x - \frac{5}{3}x^3\right)$$

If $x = \sin\theta$, each of these solutions determine the angular dependence of individual atomic orbitals if $l = 0,1,2,3 \ldots (s, p, d, f \ldots)$

Linearizing non-linear ODEs

Non-linear equations usually cannot be solved analytically. However, if the physical problem can be constrained so an approximation can be made, then it might be possible to ***linearize*** the differential equation, e.g. by assuming that the dependent variable only changes by a small amount around an equilibrium value. The linear equation this produces is then solved by conventional methods, but it will mean the solution is no longer 'exact'. This technique can be applied for PDEs as well as ODEs.

Example: Air resistance of a car

The forward thrust (force) produced by a car's engine is a product of the maximum thrust F_{max} and the fraction that the accelerator pedal is pressed at any time $p(t)$ (between 0 – 1). The drag resistance is roughly proportional to the velocity of the vehicle $v(t)$ and its cross sectional area A, so by Newton's 2nd law

$$m\frac{dv}{dt} = p(t)F_{max} - \frac{1}{2}\rho ACv^2$$

where ρ is the air density and C is a dimensionless constant called the drag coefficient which is determined by the shape of the object. Clearly the squared dependence on the velocity makes the equation non-linear. However, if we want to describe the car's motion on a motorway over a time period where the velocity changes by a small amount Δv around an average speed $\bar{v}$, then the equation can be written using $v = \bar{v} + \Delta v$

$$m\frac{d(\Delta v)}{dt} = p(t)F_{max} - \frac{1}{2}\rho AC(\bar{v}+\Delta v)^2 = p(t)F_{max} - \frac{1}{2}\rho AC\bar{v}^2\left(1 + 2\frac{\Delta v}{\bar{v}} + \left(\frac{\Delta v}{\bar{v}}\right)^2\right)$$

If $\Delta v \ll \bar{v}$ then the square term can be neglected thus linearizing the equation

$$\frac{d(\Delta v)}{dt} + \frac{\rho AC\bar{v}}{m}\,\Delta v = p(t)\frac{F_{max}}{m} - \frac{1}{2m}\rho AC\bar{v}^2$$

Consider the situation where a car of mass 1000 kg and maximum thrust $F_{max} = 3000$ N is travelling at $\bar{v} = 30$ m/s (67 mph) when the pedal 30% depressed an amount $p = 0.3$. In this equilibrium situation $\Delta v = 0$ and

$$p(t)F_{max} = \frac{1}{2}\rho AC\bar{v}^2$$

$$\rho AC = \frac{2p(t)F_{max}}{\bar{v}^2} = 2\ \mathrm{kg\ m^{-1}}$$

$$\frac{d(\Delta v)}{dt} + \frac{2\bar{v}}{m}\,\Delta v = p(t)\frac{F_{max}}{m} - \frac{\bar{v}^2}{m}$$

If at $t = 0$, the driver doubles the force applied to overtake another car so that $p(t) = 0.6$, how long does it take for the car to reach 33 m/s (74 mph)? Although the right hand side is not zero (i.e. inhomogeneous), it is a constant D

$$D = p\frac{F_{max}}{m} - \frac{\bar{v}^2}{m} = 0.9\ \mathrm{m^2 s^{-2} kg^{-1}}$$

$$\frac{d(\Delta v)}{dt} + \frac{2\bar{v}}{m}\,\Delta v = D$$

So the solution has a particular integral $\Delta v = \frac{Dm}{2\bar{v}}$ and complementary function $\Delta v = A\exp\left(-\frac{2\bar{v}}{m}t\right)$ giving a general solution

$$\Delta v = A\exp\left(-\frac{2\bar{v}}{m}t\right) + \frac{Dm}{2\bar{v}}$$

Applying the initial condition $t = 0,\ \Delta v = 0$

$$A = -\frac{Dm}{2\bar{v}}$$

$$\Delta v = \frac{Dm}{2\bar{v}}\left(1 - \exp\left(-\frac{2\bar{v}}{m}t\right)\right)$$

Putting the constants into this solution (with $\Delta v = +3\ m/s$)

$$3 = 15(1 - \exp(-0.06t))$$

$$\exp(-0.06t) = 0.8$$

$$t = 3.72\ \mathrm{s}$$

Example: An Object Cooling by Conduction, Convection and Radiation

Newton's law of cooling says that the rate of cooling of an object is proportional to the temperature difference with its surroundings. If U is the internal heat energy of an object, it has a heat capacity C defined as $C = \frac{dU}{dT}$, then

$$\frac{dU}{dt} = C\frac{dT}{dt} = -K(T - T_a)$$

where K is a constant (units W.K^{-1}) determined by the thermal coupling of the object to the environment via conduction and convection. This has a simple exponential solution

$$T = T_a + (T_0 - T_a)\exp\left(-\frac{K}{C}t\right)$$

In terms of the temperature difference $\Delta T = T - T_a$

$$\Delta T = \Delta T_0 \exp\left(-\frac{K}{C}t\right)$$

However, this equation does not include heat loss by emission of radiation (which usually is much smaller and can be neglected). The power P radiated by an object of surface area A at a temperature T is given by the Stefan-Boltzmann law

$$P_{out} = \varepsilon\sigma A T^4$$

where ε is the emissivity of the object ($\varepsilon = 1$ for a pure blackbody) and $\sigma = 5.67 \times 10^{-8}$ $Wm^{-2}K^{-4}$. At the same time the object absorbs heat from its surroundings at an ambient temperature T_a at a rate

$$P_{in} = \varepsilon\sigma A T_a^4$$

So including radiation the energy balance equation is

$$C\frac{dT}{dt} = -K(T - T_a) - \varepsilon\sigma A(T^4 - T_a^4)$$

creating a non-linear ODE. If the temperature difference ΔT is small compared to the ambient temperature T_a (on the absolute scale), then we can do a Taylor expansion (see later) of the function $u(T) = T^4 - T_a^4$ about T_a

$$u(T_a + \Delta T) = u(T_a) + \Delta T\frac{du}{dT}(T_a) + \frac{(\Delta T)^2}{2}\frac{d^2u}{dT^2}(T_a) \ldots$$

Since $\frac{du}{dT} = 4T^3$ and considering up to 1st order terms only

$$f(T) = f(T_a + \Delta T) \approx 4\varepsilon\sigma A T_a^3 \Delta T$$

giving the linear differential equation

$$C\frac{d(\Delta T)}{dt} = -K\Delta T - 4\varepsilon\sigma A T_a^3 \Delta T$$

$$\frac{d(\Delta T)}{dt} + \frac{\Delta T}{C}(K + 4\varepsilon\sigma A T_a^3) = 0$$

$$\Delta T = \Delta T_0 \exp\left(-\frac{K + 4\varepsilon\sigma A T_a^3}{C}t\right)$$

So the cooling is still exponential but at a faster rate.

1.2 Multivariable Calculus

A dependent variable with two or more independent variables is a multivariable function, e.g. $u(x, y)$ and the differential of such a function must be carefully defined.

Partial Derivatives

In partial differentiation, the function is differentiated with respect to one independent variable keeping all others constant. This is signified by the symbol ∂.

Consider the following function

$$u(x,y,z) = 3x^2y^3z + \frac{zy}{x} + \sin(x+yz)$$

A partial differentiation with respect to x while keeping y and z constant is

$$\left.\frac{\partial u}{\partial x}\right|_{y,z} \equiv \frac{\partial u}{\partial x} = 6xy^3z - \frac{zy}{x^2} + \cos(x+yz)$$

Similarly

$$\left.\frac{\partial u}{\partial y}\right|_{x,z} \equiv \frac{\partial u}{\partial y} = 9x^2y^2z + \frac{z}{x} + z\cos(x+yz)$$

$$\left.\frac{\partial u}{\partial z}\right|_{x,y} \equiv \frac{\partial u}{\partial z} = 3x^2y^3 + \frac{y}{x} + y\cos(x+yz)$$

The second order partial differential w.r.t x is

$$\frac{\partial^2 u}{\partial x^2} = 6y^3z + 2\frac{zy}{x^3} - \sin(x+yz)$$

Mixed 2nd order differentials are also possible and the order of the differentiation does not matter (commutative)

$$\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 u}{\partial y \partial z} = 18xy^2z - \frac{z}{x^2} - z\sin(x+yz)$$

The ***product and chain rules*** works in exactly the same way for partial derivatives.

For a function $u(x,y) = A(x,y)B(x,y)$

$$\frac{\partial u}{\partial x} = \frac{\partial A}{\partial x}B + A\frac{\partial B}{\partial x}$$

For a function $u(x,y) = A\big(B(x,y)\big)$

$$\frac{\partial u}{\partial x} = \frac{\partial A}{\partial B}\frac{\partial B}{\partial x}$$

Total Differentials

A derivative can be obtained by considering how the function varies with respect to one independent variables while all other independent variables are not fixed. This known as a ***total derivative*** and is expressed in the same way as an ordinary derivative $\frac{du}{dx}$, but now $u(x,y,z)$ is a multivariable function. It is important to recognise that

$$\frac{du}{dx} \neq \frac{\partial u}{\partial x}$$

Using the chain rule we can obtain an expression for the total derivative

$$\frac{du}{dx} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial x}$$

$$\frac{du}{dx} = \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial x}$$

Therefore, an expression for the total derivative can only be obtained if there are constraints on x, y, z, e.g. if

$$x^2 + y^2 + z^2 = 1$$

this limits the function to points on the surface of a sphere of radius 1. Hence

$$y = (1 - x^2 - z^2)^{1/2}$$

$$\frac{\partial y}{\partial x} = -x(1 - x^2 - z^2)^{-1/2}$$

If we again consider the function $u(x, y, z) = 3x^2y^3z + \frac{zy}{x} + \sin(x + yz)$

$$\frac{\partial u}{\partial x} = 6xy^3z - \frac{zy}{x^2} + \cos(x + yz)$$

$$\frac{\partial u}{\partial y}\frac{\partial y}{\partial x} = \left[9x^2y^2z + \frac{z}{x} + z\cos(x + yz)\right]\left[-x(1 - x^2 - z^2)^{-1/2}\right]$$

$$\frac{\partial u}{\partial z}\frac{\partial z}{\partial x} = \left[3x^2y^3 + \frac{y}{x} + y\cos(x + yz)\right]\left[-x(1 - x^2 - y^2)^{-1/2}\right]$$

$$\frac{du}{dx} = 6xy^3z - \frac{zy}{x^2} + \cos(x + yz) - \frac{9x^2y^2z + \frac{z}{x} + z\cos(x + yz) + 3x^2y^3 + \frac{y}{x} + y\cos(x + yz)}{x(1 - x^2 - y^2)^{1/2}}$$

Shorthand Derivative Notation

In books/texts shorthand notation is often employed to reduce the tedium of writing out derivatives in full, so just be aware of these if reading other sources of information. These notes will mostly present the derivatives fully but will have some abbreviations for convenience in places.

Ordinary Derivatives

$$\frac{du(x)}{dx} \equiv f'(x) \quad \frac{d^2u(x)}{dx^2} \equiv u''(x) \quad \frac{d^3u(x)}{dx^3} \equiv u'''(x) \quad \frac{d^nu(x)}{dx^n} \equiv u^{(n)}(x)$$

For dependent variables which are functions of time, e.g. Newton's laws of motion, then the following notation is often used

$$\frac{dx}{dt} \equiv \dot{x} \quad \frac{d^2x}{dt^2} \equiv \ddot{x} \quad \frac{d^3x}{dt^3} \equiv \dddot{x}$$

Partial Derivatives

$$\frac{\partial u}{\partial x} \equiv u'_x \equiv \partial_x u \qquad \frac{\partial^2 u}{\partial x^2} \equiv u''_x \equiv \partial_{xx} u \equiv \partial_x^2 u \equiv u_{xx} \qquad \frac{\partial u}{\partial x\,\partial y} \equiv u''_{xy} \equiv \partial_x \partial_y u \equiv u_{xy}$$

To emphasize that an ordinary or partial derivative is being evaluated at a constant value of an independent variable, the following notation is generally used

$$\left.\frac{df}{dx}\right|_a \equiv f'(a) \quad \text{derivative evaluated at } x = a$$

1.3 Differential Operators

Gradient - Grad

When working in one dimension with a single independent variable the function corresponds to a line and the derivative gives the gradient of the line at a point. With two independent variables the function now represents a surface and the gradient of this surface now has a direction as well as a magnitude which is determined by the partial derivatives of the function. For a scalar quantity such as potential energy or temperature, e.g. $T(x, y, z)$ the gradient at any point in space can be obtained by applying the Del (or Nabla) operator to T, where

$$\boldsymbol{\nabla} = \frac{\partial}{\partial x}\hat{\boldsymbol{\imath}} + \frac{\partial}{\partial y}\hat{\boldsymbol{\jmath}} + \frac{\partial}{\partial z}\hat{\boldsymbol{k}}$$

$$\boldsymbol{\nabla} T = \text{Grad}\, T = \frac{\partial T}{\partial x}\hat{\boldsymbol{\imath}} + \frac{\partial T}{\partial y}\hat{\boldsymbol{\jmath}} + \frac{\partial T}{\partial z}\hat{\boldsymbol{k}}$$

So for instance, if a room has a uniform temperature in the x, y directions but gets linearly warmer in z, then the temperature function will have the form $T(x, y, z) = A + Bz$ and the gradient of the temperature is a vector pointing upwards

$$\boldsymbol{\nabla} T = B\hat{\boldsymbol{k}}$$

Similarly the gradient can be visualised for the gravitational potential energy of a landscape. The direction of the gradient is the direction of steepest slope at a point and would be the direction a ball would roll if placed there.

The **Grad** operation $\boldsymbol{\nabla}\Phi$

- acts on a scalar quantity Φ giving a vector
- gives the direction and magnitude of greatest gradient
- is perpendicular to a surface Φ if it is a constant (flat)

Divergence - Div

There are many instance in Physics where a physical quantity is not scalar but is vector field with a spatial dependence. Examples of this are an electric field (electric flux density $\boldsymbol{E}$) or the flow of a fluid (mass flux density $\boldsymbol{F}$) which have a magnitude and direction at every point in space. The divergence of the vector field represents the compression (+) or expansion (–) of the flux or corresponds to a source (–) or sink (–) of it.

The divergence is calculated by taking the scalar product between the Del vector operator and the vector field

$$\boldsymbol{\nabla}.\boldsymbol{F} = \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix}.\begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}$$

An example is the rate of flow of a mass of water and its direction, e.g. in a river. It is described by the vector field $\boldsymbol{F}$ – the mass flux density (units kg s^{-1} m^{-2}) which will vary in magnitude and direction across the river (in the centre - generally downriver with the largest magnitude). Since water is fairly incompressible any point there is as much flux going in as is going out and $\boldsymbol{\nabla}.\boldsymbol{F} = 0$. However, if one were to put a hosepipe in the river and turn it on, at this point there is an additional source of positive flux and hence $\boldsymbol{\nabla}.\boldsymbol{F} > 0$.

The **Div** operation $\boldsymbol{\nabla}.\boldsymbol{F}$:

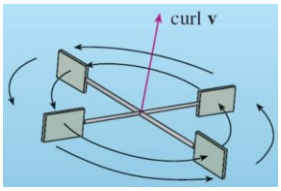
- acts on a vector field and produces a scalar value
- if $\boldsymbol{\nabla}.\boldsymbol{F} = 0$, the field is incompressible (and there are no “sources” or “sinks” of the vector field)
- if $\boldsymbol{\nabla}.\boldsymbol{F} > 0$, the field is divergent or expanding
- if $\boldsymbol{\nabla}.\boldsymbol{F} < 0$, the field is convergent or compressing

Curl

The vector product between the Del operator and another vector gives the **Curl** of the vector field.

$$\text{Curl}\,\boldsymbol{F} = \boldsymbol{\nabla} \times \boldsymbol{F} = \begin{vmatrix} \hat{\boldsymbol{\imath}} & \hat{\boldsymbol{\jmath}} & \hat{\boldsymbol{k}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix} = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right)\hat{\boldsymbol{\imath}} - \left(\frac{\partial F_z}{\partial x} - \frac{\partial F_x}{\partial z}\right)\hat{\boldsymbol{\jmath}} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)\hat{\boldsymbol{k}}$$

This vector quantity represents how much the vector field is rotating about a given point in space, with a direction normal to the plane of circulation. For the case of the vector field representing water flowing in a river, if there is an eddy then $\text{Curl}\ \boldsymbol{F} \neq 0$. If you were to put a paddle wheel into the river at that point, it would spin around.



The **Curl** operation $\boldsymbol{\nabla} \times \boldsymbol{F}$:

- acts on a vector field and produces a vector field
- $\boldsymbol{\nabla} \times \boldsymbol{F}$ is perpendicular to the direction of the field's circulation (positive for right handed circulation and vice versa)
- if $\boldsymbol{\nabla} \times \boldsymbol{F} = 0$, there is no circulation of the field at that point and the field is conservative (any line integral between two points is independent of the path taken)

Laplacian

If the gradient of a quantity is taken, followed by the divergence of this vector, this is known as a Laplacian operation denoted by the symbol ∇^2 (sometimes also by Δ)

$$\text{Div. (Grad } \Phi) = \boldsymbol{\nabla}.(\boldsymbol{\nabla}\Phi) = \nabla^2\Phi = \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2}$$

The general form of a wave equation with one spatial dimension is

$$\frac{\partial^2\Phi}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2\Phi}{\partial t^2}$$

If this is extended to 3D then it is useful to express the additional dimensions using the Laplacian

$$\nabla^2\Phi = \frac{1}{v^2}\frac{\partial^2\Phi}{\partial t^2}$$

The Laplacian operation ∇^2:

- if acting on a scalar, it generates a scalar
- can also act on a vector field generating a vector field

1.4 Taylor's Theorem

Taylor's Theorem states that if u(x) is a continuous single-valued function of x with continuous derivatives $u, u', u'', \ldots u^{(n)}$ then for a distance Δx away from a point $x = a$, the value of the function can be expressed as an infinite series

$$u(a + \Delta x) = \sum_{n=0}^{\infty} \frac{(\Delta x)^n}{n!} u^{(n)}(a) = u(a) + \Delta x\, u'(a) + \frac{\Delta x^2}{2!} u''(a) + \frac{(\Delta x)^2}{3!} u'''(a) + \cdots$$

Taylor's theorem allows the derivatives to be defined in terms of values of the function at points close to $x = a$. Considering the series expansion to 1st order, this gives a definition of the derivative of the function at $x = a$

$$u(a + \Delta x) \approx u(a) + \Delta x\, u'(a)$$

$$\left.\frac{du}{dx}\right|_a = \lim_{\Delta x \to 0} \frac{u(a+\Delta x) - u(a)}{\Delta x}$$

For the 2nd order derivative we must consider the expansion to 2nd order and the value of the function at three points $x = a - \Delta x,\ a,\ a + \Delta x$

$$u(a+\Delta x) \approx u(a) + \Delta x\, u'(a) + \frac{\Delta x^2}{2!} u''(a)$$

$$u(a-\Delta x) \approx u(a) - \Delta x\, u'(a) + \frac{\Delta x^2}{2!} u''(a)$$

Adding these together and re-arranging gives the definition

$$\left.\frac{d^2u}{dx^2}\right|_a = \lim_{\Delta x \to 0} \frac{u(a+\Delta x) - 2u(a) + u(a-\Delta x)}{(\Delta x)^2}$$

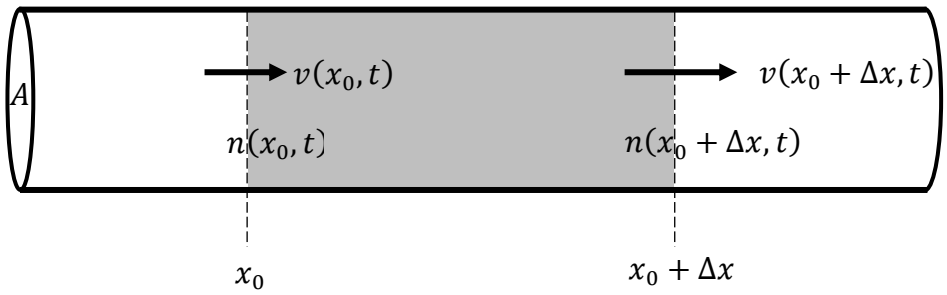
A Taylor expansion can also be written down for partial derivatives of a multivariable function.

2. Physical Examples of PDEs

PDEs are ubiquitous when the real world is modelled since any physical quantity is likely to depend on many variables, such as 3 physical dimensions and time. Different areas of physics are often described by similar PDEs, so let's look at a few key examples.

2.1. Conservation Equation

Conservation of physical quantities is the foundation of many aspects of Physics, quantities such as mass, charge and energy can be transported from one location to another but overall they are conserved. The rate of movement/transfer of a quantity in space (the flux) can be represented by a vector field, e.g. heat flux in units of Wm^{-2}. The transfer of particles/density of a fluid moving in a pipe of uniform cross sectional area A as illustrated below



The number density of fluid particles at a distance x along the pipe at a time t is $n(x,t)$ and the velocity of the fluid is $v(x,t)$. Consider the cylinder of length Δx between positions x_0 and $x_0 + \Delta x$. At the entrance (position x_0), between times t and $t + \Delta t$ a volume of fluid

$V_{in} = v(x_0,t)A\Delta t$

enters the cylinder which corresponds to a number of particles

$N_{in} = n(x_0,t)v(x_0,t)A\Delta t$

At the exit (position $x_0 + \Delta x$), in a time Δt a volume of fluid

$V_{out} = v(x_0 + \Delta x, t)\Delta t$

leaves corresponding to a number of particles

$N_{out} = n(x_0 + \Delta x, t)v(x_0 + \Delta x, t)A\Delta t$

If the total number of particles in the volume $A\Delta x$ at time t is $N(t)$ after a time Δt this changes to $N(t + \Delta t\ )$ where

$N(t + \Delta t\ ) = N(t) + N_{in} - N_{out}$

$N(t + \Delta t\ ) = N(t) + n(x_0)v(x_0)A\Delta t - n(x_0 + \Delta x)v(x_0 + \Delta x)A\Delta t$

Dividing by $A\Delta t\Delta x$

$$\frac{N(t + \Delta t\ ) - N(t)}{A\Delta t\Delta x} = \frac{n(x_0)v(x_0) - n(x_0 + \Delta x)v(x_0 + \Delta x)}{\Delta x}$$

Since $\frac{N}{A\Delta x} = n$ and using a Taylor expansion on the left hand side to first order, $n(x_0 + \Delta x) = n(x_0) + \frac{\partial n}{\partial x}\Delta x$

$$\frac{n(t + \Delta t\ ) - n(t)}{\Delta t} \approx \frac{n(x_0)v(x_0) - \left(n(x_0) + \frac{\partial n}{\partial x}\Delta x\right)\left(v(x_0) + \frac{\partial v}{\partial x}\Delta x\right)}{\Delta x}$$

$$\frac{n(t + \Delta t\ ) - n(t)}{\Delta t} \approx -\frac{v(x_0)\frac{\partial n}{\partial x}\Delta x + n(x_0)\frac{\partial v}{\partial x}\Delta x + \frac{\partial n}{\partial x}\frac{\partial v}{\partial x}(\Delta x)^2}{\Delta x}$$

As $\Delta x, \Delta t \to 0$

$$\frac{\partial n}{\partial t} + v\frac{\partial n}{\partial x} + n\frac{\partial v}{\partial x} = 0$$

$$\frac{\partial n}{\partial t} + \frac{\partial}{\partial x}(nv) = 0$$

If this is multiplied by the mass of the particles, then this can be expressed in terms of the density of the fluid ρ

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho v) = 0$$

This is for one dimensional flow, if we consider flow in three dimensions the velocity is a vector $\boldsymbol{v}$ and the net particle flux in the y, z dimensions must also be considered so that

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho v_x) + \frac{\partial}{\partial y}\left(\rho v_y\right) + \frac{\partial}{\partial z}(\rho v_z) = 0$$

The general ***conservation equation*** is therefore

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla}.(\rho \boldsymbol{v}) = \frac{\partial \rho}{\partial t} + \boldsymbol{\nabla}.\boldsymbol{F} = 0$$

where $\boldsymbol{F} = \rho \boldsymbol{v}$ is the mass flux representing the rate at which material is flowing in a particular direction per unit area (units $kg.s^{-1}.m^{-2}$)

This equation can be derived more rigourously by considering the flux density $\rho \boldsymbol{v}$ crossing an enclosed surface S and equating that with the rate of change of density in the volume V within that surface

$$\text{Rate of decrease of density in volume V} = \text{ Flux density leaving crossing the surface S}$$

$$-\frac{\partial}{\partial t}\iiint_V \rho\ dV = \oiint_S \boldsymbol{F}.d\boldsymbol{A}$$

Using the divergence theorem

$$\oiint_S (\rho \boldsymbol{v}).d\boldsymbol{A} = \iiint_V \boldsymbol{\nabla}.\boldsymbol{F} dV$$

giving

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla}.\boldsymbol{F} = 0$$

2.2. Advection Equation

Considering the 1D conservation equation again

$$\frac{\partial \rho}{\partial t} + v\frac{\partial \rho}{\partial x} + n\frac{\partial \rho}{\partial x} = 0$$

if we take the parcel of fluid in the volume $A\Delta x$, the flow is said to be isochoric (incompressible) if the density within that element remains constant as it flows down the tube. This does not mean the fluid itself is incompressible (otherwise ρ would be constant and the left hand side of the equation would just be zero), but that a group of particles maintain their density as they flow. There may be variations in density with this element but as a whole it remains constant. For this to occur, then these particles must all be travelling with the same average velocity, therefore $\frac{\partial v}{\partial x} = 0$. This produces the ***1D advection equation*** (the term advection is used into Physics for the transport of a conserved quantity by bulk motion – heat, mass, charge etc.)

$$\frac{\partial \rho}{\partial t} + v\frac{\partial \rho}{\partial x} = 0$$

In 3D the conservation equation can be expressed as (using the chain rule with the del operator)

$$\frac{\partial \rho}{\partial t} + (\nabla \rho).\boldsymbol{v} + \rho(\boldsymbol{\nabla}.\boldsymbol{v}) = 0$$

Incompressible flow is expressed as $\boldsymbol{\nabla}.\boldsymbol{v} = 0$, giving the ***3D advection equation***

$$\frac{\partial \rho}{\partial t} + (\boldsymbol{\nabla}\rho).\boldsymbol{v} = 0$$

2.3. Heat (Diffusion) Equation

The rate of heat flow H (W) in a material in one dimension is described by Fourier's law

$$H = -kA\frac{dT}{dx}$$

where A is the cross sectional area through which the heat is flowing, $\frac{dT}{dx}$ is the temperature gradient and k is the thermal conductivity of the material ($Wm^{-1}K^{-1}$). Extending this to three dimensions

$$\frac{\boldsymbol{H}}{A} = -k\,\boldsymbol{\nabla}T$$

where $\frac{\boldsymbol{H}}{A}$ is the heat flux vector. We can now use the conservation equation, replacing mass flux $\boldsymbol{F}$ with heat flux $\frac{\boldsymbol{H}}{A}$ and the rate of change of mass density $\frac{\partial \rho}{\partial t}$ with rate of change of energy density

$$\frac{\partial (U/V)}{\partial t} = \frac{C}{V}\frac{\partial T}{\partial t} = \frac{C_S M}{V}\frac{\partial T}{\partial t} = \rho C_S\frac{\partial T}{\partial t}$$

where C_S is the specific heat capacity, giving the ***Heat Equation***

$$\rho C_S\frac{\partial T}{\partial t} + \boldsymbol{\nabla}.(-k\,\boldsymbol{\nabla}T) = 0$$

$$\frac{\partial T}{\partial t} = \frac{k}{\rho C_S}\nabla^2 T = D\,\nabla^2 T$$

where $D = \frac{k}{\rho C_S}$ is the thermal diffusivity of the material (units m^2s^{-1}). If we consider a rod of material with constant cross sectional area which is insulated so that heat can only flow through its ends which are connected to a heat reservoir, then heat flow has only one direction and the heat equation reduces to

$$\frac{\partial T}{\partial t} = D\frac{\partial^2 T}{\partial x^2}$$

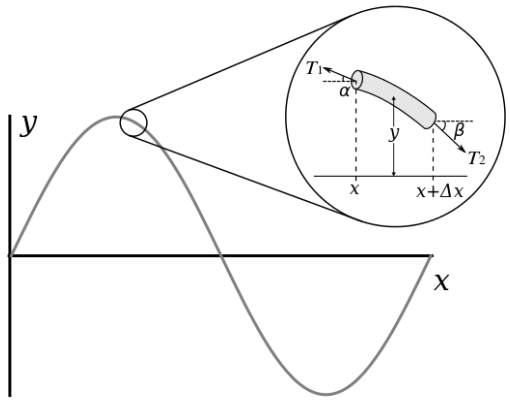
2.4. Wave Equation

Many types of waves, e.g. water waves, mechanical waves, waves on a string, sound waves, light waves…, can all be described with the same equation even though the physical underpinnings of each are different

$$\nabla^2\Phi = \frac{1}{v^2}\frac{\partial^2\Phi}{\partial t^2}$$

where v is the phase velocity of the wave and Φ is the physical quantity via which the wave is being propagated such as displacement from equilibrium. For the mechanical displacements of solids or fluids the wave equation is derived by consideration of Newton's 2nd law for the case where there is a linear restoring force following displacement from equilibrium (which applies for all stable equilibrium situations in the limit of small displacement).

Consider transverse waves on a wire under a tension T with a density per unit length μ. At a position x along the wire, it has a displacement y.



A length of wire between x and $x + \Delta x$ experiences a force in the y direction

$$\Delta F_y = T_1 - T_2 = T(\sin\alpha - \sin\beta)$$

For small displacements where $\alpha, \beta \ll 1$,

$$\Delta F_y \approx T(\tan\alpha - \tan\beta)$$

Since tan of the angle is the gradient at that point

$$\Delta F_y = T\left(\left.\frac{\partial y}{\partial x}\right|_x - \left.\frac{\partial y}{\partial x}\right|_{\substack{x+\\ \Delta x}}\right)$$

Using Newton's 2nd law

$$\mu\Delta x\frac{\partial^2 \mathrm{y}}{\partial t^2} = T\left(\left.\frac{\partial y}{\partial x}\right|_x - \left.\frac{\partial y}{\partial x}\right|_{\substack{x+\\ \Delta x}}\right)$$

$$\frac{\mu}{T}\frac{\partial^2 y}{\partial t^2} = \frac{\left(\left.\frac{\partial y}{\partial x}\right|_x - \left.\frac{\partial y}{\partial x}\right|_{x+\Delta x}\right)}{\Delta x}$$

$$\frac{\mu}{T}\frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial x^2}$$

$$\frac{1}{v^2}\frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial x^2}$$

with the wave having a phase velocity $v = \sqrt{\frac{T}{\mu}}$

Similar derivations using Newton's 2nd law can be used to obtain wave equations for material displacements. The wave equation for light (which is an electromagnetic wave) can be derived from Maxwell's equations which you will do in PHY2004.

The general solution to the wave equation has the form (https://mathworld.wolfram.com/dAlembertsSolution.html)

$$y(x,t) = f(x - vt) + g(x + vt)$$

The functions f, g are solutions to the wave equation and represents waves travelling in opposite directions with a velocity v. In x, t space $x - vt$ and $x + vt$ are known as characteristics (trajectories) along which the waves travel. The wave travels along the characteristic without changing shape so that if you were "riding" along with one of these waves it remain unchanged in your frame of reference. In most physical situations the waves will disperse or be damped as they propagate which can represented by adding terms in the wave equation, e.g. $\frac{\partial y}{\partial t}$ for damping.

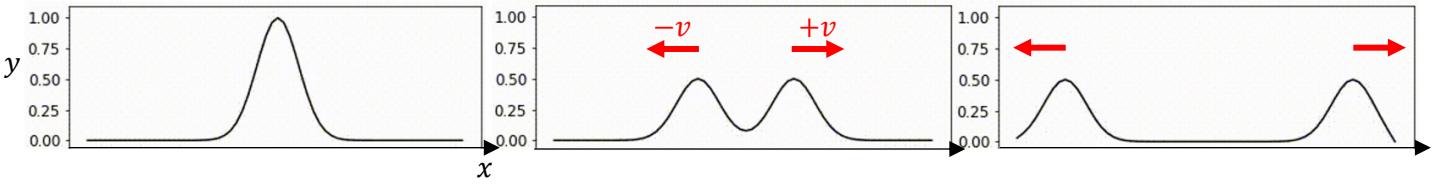
Consider a Gaussian function

$$f = \exp(-(x - vt)^2)$$

$$\frac{\partial f}{\partial x} = -2(x - vt)\exp(-(x - vt)^2) \qquad \frac{\partial f}{\partial t} = 2v(x - vt)\exp(-(x - vt)^2)$$

$$\frac{\partial^2 f}{\partial x^2} = (4(x - vt)^2 - 2)\exp(-(x - vt)^2) \qquad \frac{\partial^2 f}{\partial x^2} = (4v^2(x - vt)^2 - 2v^2)\exp(-(x - vt)^2)$$

So this satisfies the wave equation. If we have a string which initially is displaced from equilibrium with a Gaussian distribution and released, one Gaussian wave travels to the left and one to the right with a velocity $v = \sqrt{\frac{T}{\mu}}$



The solution being

$$y(x,t) = \frac{1}{2}\exp(-a(x - vt)^2) + \frac{1}{2}\exp(-a(x + vt)^2)$$

2.5. Laplace's and Poisson's Equations

Laplace's and Poisson's equation both involve the Laplacian operator acting on a scalar potential function

$$\nabla^2\Phi(x,y,z) = \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} = 0 \qquad \textbf{Laplace}$$

$$\nabla^2\Phi(x,y,z) = \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} = f(x,y,z) \quad \textbf{Poisson}$$

As there are no time dependent derivatives, these equations are must useful for describing systems which are in equilibrium, note that the heat equation reduces to Laplace's equation at equilibrium (as $\frac{\partial T}{\partial t} = 0$). These equations are particularly useful in electrostatics for calculating the electric potential in space (again you will see this in PHY2004)

2.6. Schrödinger's Wave Equation

As you have found out in PHY2001, by describing particles by a wavefunction Ψ with wavelength determined by de Broglie, it is possible to derive a wave equation for Ψ

$$i\hbar\frac{\partial\Psi(\boldsymbol{r},t)}{\partial t} = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r},t)\right]\Psi(\boldsymbol{r},t)$$

Like all the other PDEs presented in this section, the SWE is linear meaning that superpositions of particle waves cause interference in the same manner as other types of waves.

3. Definitions/Classification of PDEs

PDEs are classified in terms of order and linearity in the same way as ODEs. In all the examples given in Section 2, the PDEs are linear which makes it easier to obtain an analytical solution. Even if an equation is non-linear, by making an approximation it is often possible to linearize the equation so a solution can be obtained. For equations which can't be solved analytically, numerical solutions can be obtained by computer.

Order

The order of a PDE maximum number of times a dependent variable has been differentiated for any term in the equation.

$(x^2 + y^2)\frac{\partial u}{\partial x} + 2y\frac{\partial^2 u}{\partial x\partial y} - 3u = 0 \quad \text{2nd order}$

$u\left(\frac{\partial u}{\partial x}\right)^2 + x\frac{\partial u}{\partial y} = 0 \quad \text{1st order}$

Linearity

For a PDE to be linear, any terms with the dependent variable u or its partial derivatives must be raised to the power of 1 only. Examples of non-linear terms

$u^2 \qquad \left(\frac{\partial u}{\partial x}\right)^{1/2} \qquad u\frac{\partial^2 u}{\partial x\partial y} \qquad \left(\frac{\partial u}{\partial x}\right)\left(\frac{\partial u}{\partial y}\right)$

Examples of linear terms

$u \qquad \frac{1}{x}\frac{\partial^2 u}{\partial x^2} \qquad \frac{\partial^2 u}{\partial x\partial y} \qquad xy\left(\frac{\partial u}{\partial x}\right)$

Homogeneity

The equation is homogeneous if all the terms in the equation contain the dependent variable or its partial derivatives. For example the following equation is a 2nd order, non-linear, inhomogeneous equation

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

Types of 2nd Order PDEs

A quadratic equation in two dimensions with the general form

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

can be transformed into following form

$$\frac{X^2}{A^2} \pm \frac{Y^2}{B^2} = 1$$

Where $X = x - x_0$ and $Y = y - y_0$, and A and B are constants depending on $a - f$

For $b^2 - 4ac > 0$, the sign is negative and this describes a ***hyperbolic*** function

If $b^2 - 4ac < 0$, the sign is positive and this describes an ***elliptical*** function (a circle if $A = B$)

If $b^2 - 4ac = 0$, then $Y = AX^2$ a ***parabolic*** function is obtained

A 2nd order linear PDE with one dependent variable $u(x, y)$ and two independent variables x, y be generally written in the form

$$A\frac{\partial^2 u}{\partial x^2} + B\frac{\partial^2 f}{\partial x \partial y} + C\frac{\partial^2 u}{\partial y^2} + D\frac{\partial u}{\partial x} + E\frac{\partial u}{\partial y} + F = 0$$

Similarly the PDE can be classified in the same way where

$B^2 - 4AC > 0$ ***hyperbolic***

$B^2 - 4AC < 0$ ***elliptical***

$B^2 - 4AC = 0$ ***parabolic***

If A, B, C are not constants but functions of x, y, then the equation may be a mixed type where it will have one type in a particular region of space and another elsewhere. For example

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

Is elliptic for $x < 0$, hyperbolic for $x > 0$ and parabolic for $x = 0$

The significance of these classifications is that they often determine the way in which the PDE needs to be solved (see next section 4.1).

Examples

	PDE	**Linearity**	**Order**	**2nd order type**	**Homogeneous/ Inhomogeneous**
Advection	$\frac{\partial u}{\partial t} + v(x,t)\frac{\partial u}{\partial x} = 0$	Linear	1	-	Homogeneous
Heat	$\frac{\partial T}{\partial t} = D\frac{\partial^2 T}{\partial x^2}$	Linear	2	Parabolic	Homogeneous

Wave	$\frac{\partial^2 u}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 u}{\partial t^2}$	Linear	2	Hyperbolic	Homogeneous
Laplace	$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} = 0$	Linear	2	Elliptical	Homogeneous
Helmholtz	$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + k^2 u = 0$	Linear	2	Elliptical	Homogeneous
Kuramoto–Sivashinsky	$\frac{\partial u}{\partial t} + \frac{\partial^4 u}{\partial x^4} + \frac{\partial^2 u}{\partial x^2} + \frac{1}{2}\left(\frac{\partial u}{\partial x}\right)^2 = 0$	Non-linear	4	-	Homogeneous
Fisher	$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} + ru(1-u)$	Non-linear	2	-	Homogeneous
1st Maxwell equation	$\boldsymbol{\nabla}.\boldsymbol{E} = \frac{\rho(x,y,z)}{\varepsilon_0}$	Linear	1	-	Inhomogeneous

4. Analytical Solutions of PDEs

An analytical solution to a differential equation is where an algebraic expression can be obtained for the solution based on the initial conditions and/or boundary conditions. Analytical solutions are 'exact' and in this section we will look at different ways of obtaining analytical solutions to PDEs. However, in many cases the PDE does not have an analytical solution.

4.1. Constraints Required for a Solution

Sufficiency/Uniqueness of solutions

You have seen for ODEs that general solutions can be obtained, but to obtain a unique solution then some data is required to constrain the solution. For example to solve a 1st order ODE we only need one data point (an initial condition) while 2nd order requires two data points, e.g. the value of the dependent variable and its derivative at a certain value of the independent variable, e.g. $u(x=0)$ and $u'(x=0)$.

As PDEs have more than one dependent variable then many more solutions are possible and it is more difficult to determine what data is required to achieve a unique solution. The problem is said to be *well-posed* if the data leads to a unique solution, and the solution depends continuously on the data. If there is insufficient data, then there is not enough constraint on possible solutions and a unique solution cannot be obtained. Conversely, if the data is too restrictive then no solution is possible – these constraints would represent an unphysical situation.

Initial / Boundary conditions

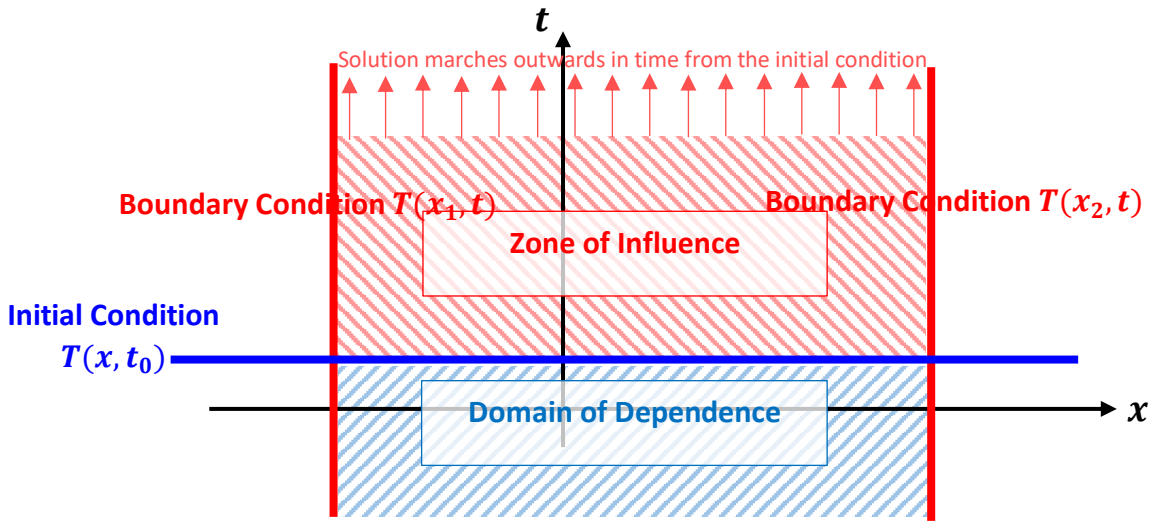
For a PDE to describe a physical system, then the independent variables will include at least one spatial dimension along with time. For example consider the one dimensional heat equation governing the flow of heat along a bar of length L. We can see that there could be an infinite number of possible temperature distributions along the bar (provided $T(x,t)$ and $\partial T(x,t)/\partial x$ are continuous) which will evolve continuously in time as the heat is redistributed. However, a unique solution can only obtained for appropriate initial conditions (ICa) and boundary conditions (BC), corresponding to the temperature distribution $T(x,t_0)$ along the bar at a time t_0 and values of T and/or $\partial T/\partial x$ at the boundaries, i.e. at $x = 0, L$.

If, for example, the ends of the bar are in perfect thermal contact with a heat reservoir at 0°C, then the solution must satisfy these boundary conditions where $T(0,t) = T(L,t) = 0$. This is an example of a ***Dirichlet Boundary***

Condition where the value of the dependent variable (temperature in this instance) is constant in time at particular spatial positions (boundaries). However, if the end at $x = L$ is removed from the reservoir and surrounded by a perfect insulator, then no heat can flow out this end of the bar, so from Fourier's law this must mean $\partial T/\partial x|_{x=L} = 0$. If the derivative of the dependent variable is constant, this is known as a ***Neumann Boundary Condition***. Either of these examples in conjunction with an initial temperature distribution of the bar is sufficient to obtain a unique solution for all x and t.

Parabolic Equations

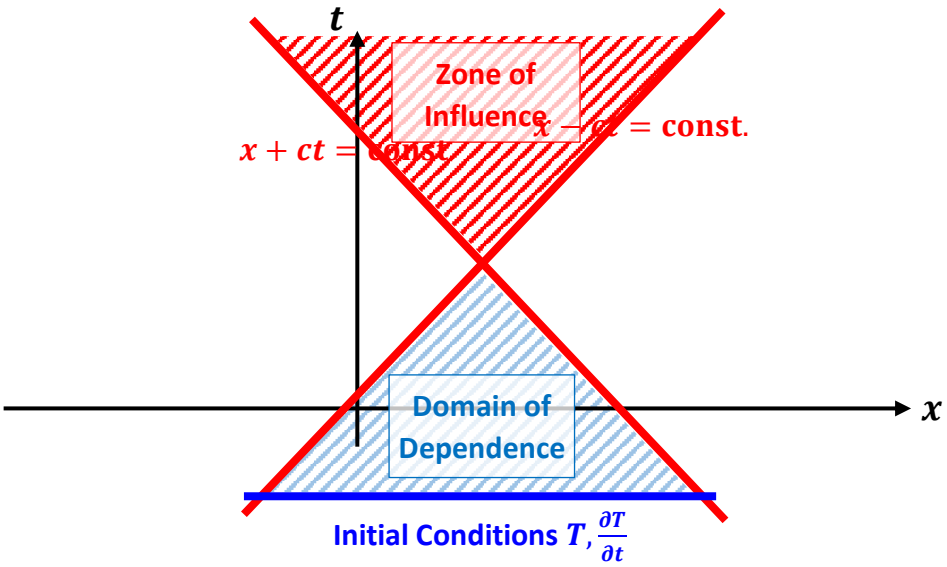
The heat equation $\frac{\partial T}{\partial t} = D\frac{\partial^2 T}{\partial x^2}$ is an example of a parabolic PDE. The category of a PDE determines the way which the IC and BC influence the solution $T(x,t)$. The figure below shows (x,t) space divided up by one initial condition $T(x,t_0)$ and two boundary conditions $T(x_1,t)$ and $T(x_2,t)$. The region $t < t_0$ has values which ultimately determine the IC - $T(x,t_0)$ and is known as the domain of dependence. The region $t > t_0$ where we are trying to find a solution is called the zone of influence (it is influenced by the value of the function in the domain of dependence). The solution is found by ***marching forward in time*** where the solution at a particular time is dependent only on the solution from the immediately preceding time, along with the boundary constraints (this will be evident when we solve this equation numerically in Section 5).



Hyperbolic Equations

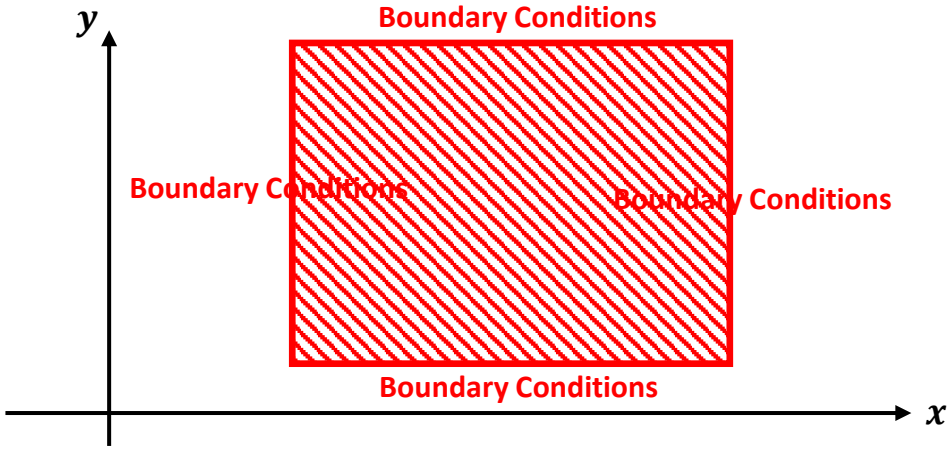
The wave equation $\frac{1}{c^2}\frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial x^2}$ is an example of a hyperbolic equation. This equation has solutions corresponding to a forward travelling wave ($x - ct = \text{const.}$) and a backward travelling wave ($x + ct = \text{const.}$). This determines boundaries in space-time as shown below, which in conjunction with initial conditions defines a triangle in (x,t) which is the domain of dependence. The values in this region determine the value at point P (x_0,t_0) at the apex of the triangle and the points in the zone of influence. As in the parabolic case, solutions can be found by *forward marching in time*.

P (x_0,t_0)



Elliptical Equations

Laplace's equation $\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} = 0$ is an example of an elliptical equation and is often used to describe systems in equilibrium (hence there is no time dependence). So in this case there is no marching forward in time, but the value of the function at a point in space is influenced by all other points (a change/disturbance at one point "instantaneously" propagates to all other points). For a solution to be obtained for this problem in a specific region of space, then boundary conditions are required (either Dirichlet, Neumann or a combination) for all points along the boundary enclosing the space. The domain of solution for an elliptic PDE is enclosed by these boundaries.



4.2. Method of Characteristics

PDEs which can be solved by *forward marching in time*, have trajectories which they follow through (x, t) space. These curves are known as ***characteristics*** along which the solution propagates. This reduces the dimensionality of the problem so that the PDE is reduced into two (or more) ordinary ODEs. We will demonstrate how this is achieved for 1st order PDEs with one dependent variable u and two independent variables x, t

$$\frac{\partial u}{\partial t} + v\frac{\partial u}{\partial x} = f$$

where v, f may be a functions of x, t, u. This is just our 1D advection equation. If we compare this for an expression for the *total derivative* $\frac{du}{dt}$

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

then we can see that we write down two ODEs

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

which is the velocity at which the solution is propagating along x

$$\frac{du}{dt} = f$$

$\frac{dx}{dt} = v$ defines the trajectory or *characteristic* of a solution through (x, t) while $\frac{du}{dt} = f$ determines how the dependent variable evolves along the trajectory. This best visualised with some simple examples for instance consider a 1st order equation which can describe a single wave travelling at a constant speed c

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

An example would be a single wavefront moving along the surface of water, with the wave shape initially described by a Gaussian function $u(x, 0) = \exp(-x^2)$.

The characteristic trajectories are given by

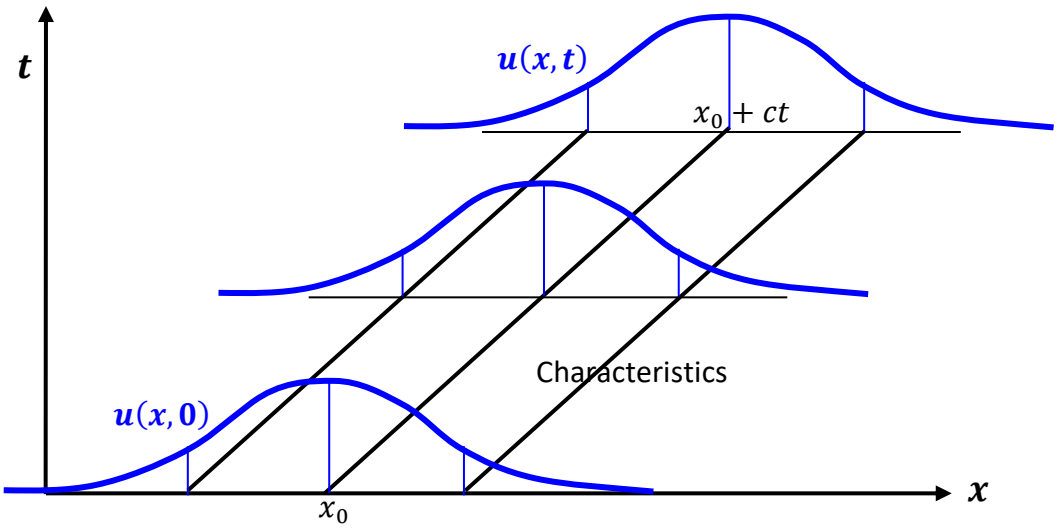
$$\frac{dx}{dt} = c$$

$$x = ct + x_0$$

which correspond to a series of parallel straight lines in (x, t) space. Each of these lines is governed by the parameter $x_0 = x - ct$ which is where the characteristic passes through the x axis. The other ODE is

$$\frac{du}{dt} = 0$$

which determines how the value of u evolves in time along a characteristic parameterised by x_0 (the initial point along the x axis). In this case u remains constant along each characteristic



Integration of the total derivative $\frac{du}{dt}$ occurs along the characteristic curves determine by x_0. As the derivative is zero, a constant is obtained which depends only on x_0.

$u(x_0, t) = C(x_0)$

For each initial position x_0 along the x axis we have a distribution

$u(x_0, 0) = \exp(-x_0^2) = C(x_0)$

$u(x_0, t) = \exp(-x_0^2)$

So $u(x_0, t)$ is a solution along a characteristic defined by x_0. To obtain the solution for all characteristics $u(x, t)$, then $x_0 = x - ct$ is substituted in

$u(x, t) = \exp(-x_0^2) = \exp(-(x - ct)^2)$

Solution for characteristics representing different velocities

Consider the same initial conditions as the previous example but with $u(x, 0) = \exp(-x^2)$ but the velocity of the wave increases linearly with time $v = \frac{dx}{dt} = t$.

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

Solving

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

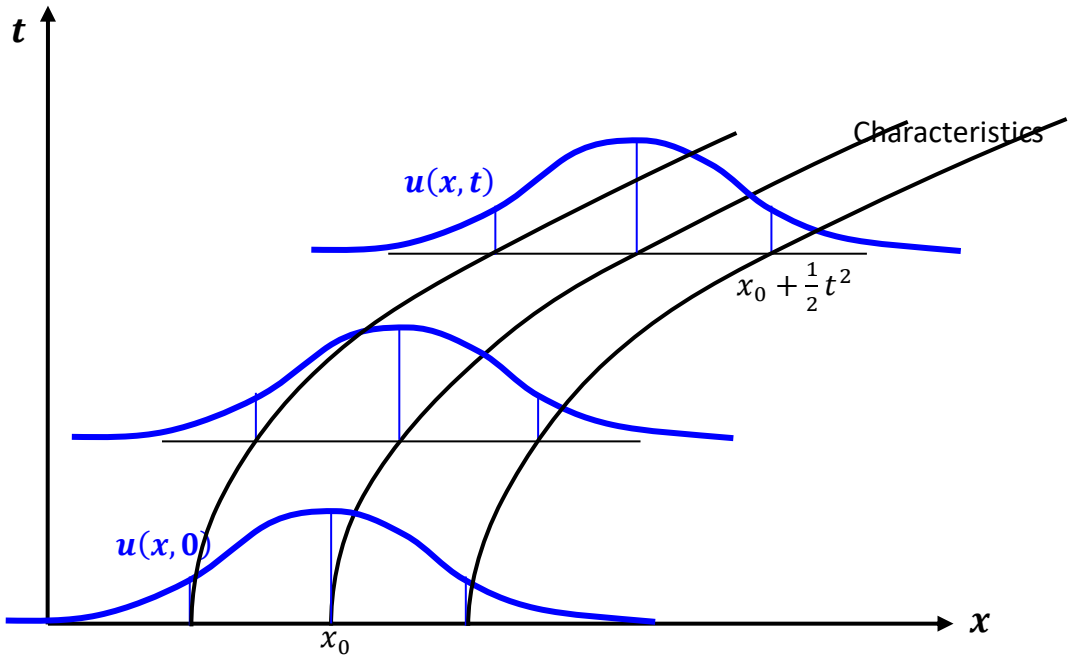
$x = \frac{1}{2}t^2 + x_0$

The characteristics are

$x_0 = x - \frac{1}{2}t^2$

As before $\frac{du}{dt} = 0$ and $u(x_0, t) = C(x_0)$, therefore

$u(x, t) = \exp(-x_0^2) = \exp\left(-\left(x - \frac{1}{2}t^2\right)^2\right)$



When the function evolves along a characteristic, i.e. $du/dt \neq 0$

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

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

$$x = ct + x_0$$

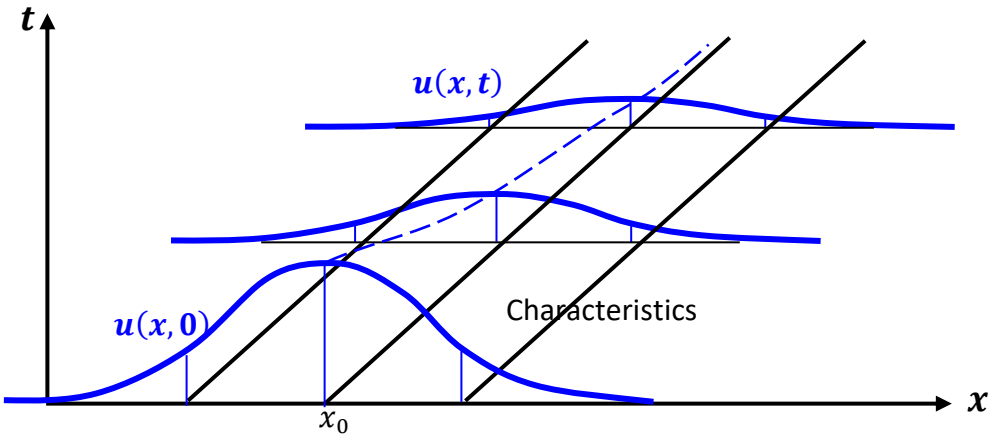
$$\frac{du}{dt} = -u$$

$$u(x_0, t) = C(x_0)\exp(-t)$$

From the initial conditions

$$u(x_0, 0) = C(x_0)\exp(0) = C(x_0) = \exp(-x_0^2)$$

$$u(x, t) = \exp(-x_0^2)\exp(-t) = \exp(-(x - ct)^2)\exp(-t)$$



Wave with amplitude dependent velocity

I am sure you have watched waves approaching a beach and notice that as the water depth reduces, that the waves break. This is because the parts of the wave with a higher amplitude travel faster so that the crest of the wave "gets ahead" of the rest of the waves until it falls due to gravity. This motion can be approximated by the following equation

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

where the velocity of a wave is dependent on the amplitude, $v = u$, so different parts of the wave travel faster than others but the amplitude is maintained (as $\frac{du}{dt} = 0$). The ODEs in this case are

$$\frac{dx}{dt} = u \qquad \frac{du}{dt} = 0$$

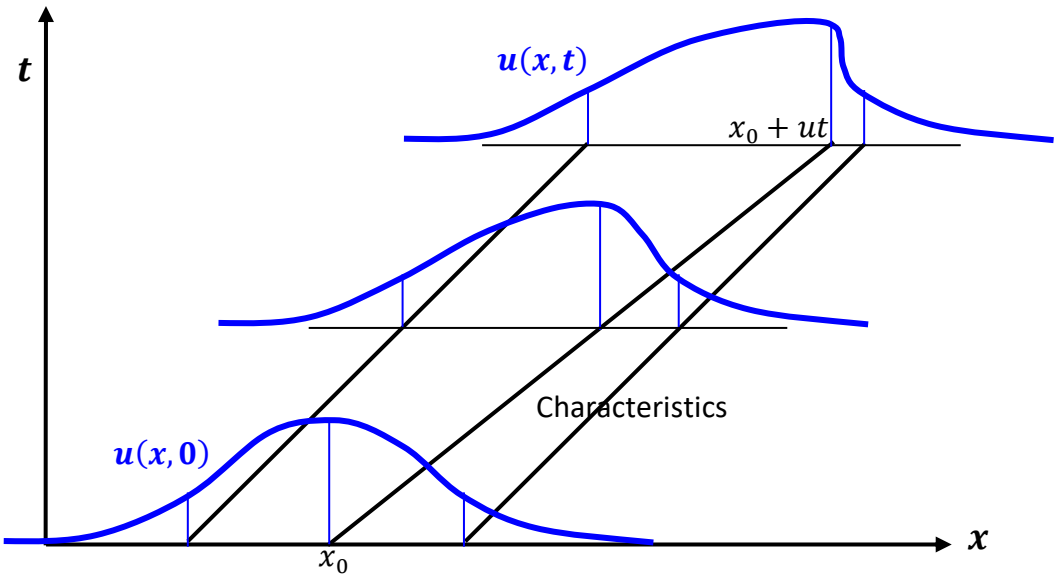
Since u is constant along a characteristic then integrating the left hand equation we have

$$x = ut + x_0$$

giving characteristics $x_0 = x - ut$ and as before $u(x_0, 0) = C(x_0) = \exp(-x_0^2)$ so that

$$u(x,t) = \exp(-(x-ut)^2)$$

Although we do not have an explicit expression for u, the initial conditions and characteristics show how it evolves in time



4.3. Separation of Variables

ODEs

You will already be aware of using separation of variables to solve ODEs. For example the 1st order equation

$$\frac{du}{dx} = -2xu$$

If we put all the y terms on the left hand side and all the x on the right hand side, then we can intergrate each side independently

$$\int \frac{du}{y} = -2\int x dx$$

$$\ln u = -x^2 + C$$

$$u(x) = \exp(C - x^2) = A\exp(-x^2)$$

PDEs

However for a PDE we have one or more additional independent variables. Consider a dependent variable which is a function of two independent variables $y(x,t)$. For some PDEs, usually linear equations, the solution can be written as a product of two functions

$$y(x,t) = f(x)g(t)$$

When this is substituted into the PDE, all the terms dependent on x are brought to one side of the equality and terms dependent on y to the right side. As an example consider the wave equation for waves on a wire fixed at two points $x = 0, L$, where $y(x,t)$ is the displacement of the wire from its equilibrium position. The wave equation when there is no energy loss in the motion (c - phase velocity of waves along the wire)

$$\frac{1}{c^2}\frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial x^2}$$

Substituting the solution

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}[f(x)g(t)] = \frac{\partial^2}{\partial x^2}[f(x)g(t)]$$

$$\frac{1}{c^2}\frac{d^2g}{dt^2}f(x) = \frac{d^2f}{dx^2}g(t)$$

These are now ordinary (total) derivatives since f and g are functions of a single variable only. Dividing by $f(x)g(t)$

$$\frac{1}{c^2}\frac{d^2g}{dt^2}\frac{1}{g(t)} = \frac{d^2f}{dx^2}\frac{1}{f(x)}$$

means we have separated the variables where the left hand side is exclusively a function of t, while the right hand is exclusively a function of x. Hence, it is not possible to consider a variation in one side of the equation without it somehow affecting the other. But both sides depend on completely independent variables. The only way this can be achieved is if both sides are in fact equal to a constant.

$$\frac{1}{c^2}\frac{d^2g}{dt^2}\frac{1}{g(t)} = \frac{d^2f}{dx^2}\frac{1}{f(x)} = K$$

K is known as a separation constant and we now have two separate ODEs

$$\frac{d^2g}{dt^2} - c^2Kg(t) = 0$$

$$\frac{d^2f}{dx^2} - Kf(x) = 0$$

We know how to solve these 2nd order, linear, homogeneous ODEs. Consider the equation in x, solutions have the form $f(x) = \exp(mx)$ giving the characteristic equation

$$m^2 - K = 0$$

$$m = \pm\sqrt{K}$$

giving a general solution

$$f(x) = C_1e^{+\sqrt{K}x} + C_2e^{-\sqrt{K}x}$$

The form of the solutions depends on whether $K < 0, K = 0, K > 0$. The solution which applies depends on the boundary conditions. Since $f(0) = 0$, then $C_1 = -C_2$

$$f(x) = C_1\left(e^{+\sqrt{K}x} - e^{-\sqrt{K}x}\right)$$

For $f(L) = 0$

$$0 = C_1\left(e^{+\sqrt{K}L} - e^{-\sqrt{K}L}\right)$$

This can be satisfied if $C_1 = 0$ or $K = 0$, this gives $f(x) = 0$ which is a mathematically acceptable solution but not very interesting for our physical problem (the wire is at equilibrium and doesn't move). Therefore

$$e^{+\sqrt{K}L} = e^{-\sqrt{K}L}$$

Which means that if $K < 0$ this gives an imaginery exponential

$$e^{+i\sqrt{|K|}L} = e^{-i\sqrt{|K|}L}$$

From Euler's relationship $e^{i\theta} = \cos\theta + i\sin\theta$

$$\cos\left(\sqrt{|K|}L\right) + i\sin\left(\sqrt{|K|}L\right) = \cos\left(\sqrt{|K|}L\right) - i\sin\left(\sqrt{|K|}L\right)$$

$$2i\sin\left(\sqrt{|K|}L\right) = 0$$

Therefore $\sqrt{|K|}L = n\pi$, where n is an integer

$$\sqrt{|K|} = \frac{n\pi}{L}$$

$$f(x) = 2iC_1 \sin\left(\frac{n\pi}{L}x\right) = 2iC_1 \sin\left(\frac{n\pi}{L}x\right)$$

Now we know the allowed values of K, we can now solve the other ODE

$$\frac{d^2g}{dt^2} + \left(\frac{n\pi}{L}c\right)^2 g(t) = 0$$

$$g(t) = C_3 \sin\left(\frac{n\pi}{L}ct\right) + C_4 \cos\left(\frac{n\pi}{L}ct\right)$$

which gives an overall solution

$$y(x,t) = f(x)g(t) = \sin\left(\frac{n\pi}{L}x\right)\left[a_n \sin\left(\frac{n\pi}{L}ct\right) + b_n \cos\left(\frac{n\pi}{L}ct\right)\right]$$

Note that the constants C_1, C_3, C_4 have been combined into two new constants a_n and b_n. To find these we need to consider the initial conditions, i.e. displacement of the wire at $t = 0$

$$y(x,0) = B_n \sin\left(\frac{n\pi}{L}x\right)$$

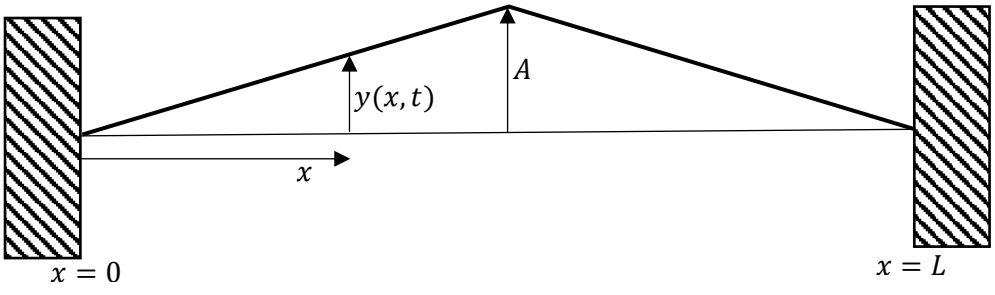
But what if the initial displacement of the wire does not have a sinusoidal shape? Note that there are an infinite number of possible solutions as n is a positive integer. Using the principal of superposition (the equation is linear), then we can combine any linear combination of these solutions and it will also be a solution. Such a sum of sines and cosines is a Fourier Series!

Obtaining a Complete Solution using the Fourier Method

The most general solution is a Fourier Series

$$y(x,t) = \sum_{n=1}^{\infty} \sin\left(\frac{n\pi}{L}x\right)\left[a_n \sin\left(\frac{n\pi}{L}ct\right) + b_n \cos\left(\frac{n\pi}{L}ct\right)\right]$$

Consider the case where the wire is pulled at its midpoint by a distance A from equilibrium and released (like plucking a stringed instrument)



The initial condition is defined as

$$y(x,0) = \frac{2Ax}{L} \text{ for } 0 < x < \frac{L}{2}$$

$$y(x,0) = 2A\left(1 - \frac{x}{L}\right) \text{ for } \frac{L}{2} < x < L$$

These coefficients a_n and b_n are found using the normal Fourier methods. For $0 < x < \frac{L}{2}$ and $\frac{L}{2} < x < L$

$$\frac{2Ax}{L} = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi}{L}x\right) \qquad 2A\left(1-\frac{x}{L}\right) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi}{L}x\right)$$

multiplying both sides by $\sin\left(\frac{m\pi}{L}x\right)$ and integrating between $x = 0, L$

$$\int_0^{L/2} \frac{2Ax}{L}\sin\left(\frac{m\pi}{L}x\right)dx + \int_{L/2}^{L} 2A\left(1-\frac{x}{L}\right)\sin\left(\frac{m\pi}{L}x\right)dx = \int_0^{L} \sin\left(\frac{m\pi}{L}x\right)\sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi}{L}x\right)dx$$

Due to the orthogonality of sine functions, in the range $x = 0 - L$, $\int_0^L \sin\left(\frac{m\pi}{L}x\right)\sin\left(\frac{n\pi}{L}x\right)dx = 0$ if $m \neq n$

$$\int_0^{L/2} \frac{x}{L}\sin\left(\frac{m\pi}{L}x\right)dx - \int_{L/2}^{L} \frac{x}{L}\sin\left(\frac{m\pi}{L}x\right)dx + \int_{L/2}^{L} \sin\left(\frac{m\pi}{L}x\right)dx = \frac{b_m}{2A}\int_0^{L} \sin^2\left(\frac{m\pi}{L}x\right)dx$$

Using integration by parts $\int \frac{x}{L}\sin\left(\frac{m\pi}{L}x\right)dx = -\frac{x}{m\pi}\cos\left(\frac{m\pi}{L}x\right) + \frac{L}{m^2\pi^2}\sin\left(\frac{m\pi}{L}x\right)$ and $2\sin^2\theta = 1 - \cos 2\theta$

$$\left[-\frac{x}{m\pi}\cos\left(\frac{m\pi}{L}x\right) + \frac{L}{m^2\pi^2}\sin\left(\frac{m\pi}{L}x\right)\right]_0^{\frac{L}{2}} - \left[-\frac{x}{m\pi}\cos\left(\frac{m\pi}{L}x\right) + \frac{L}{m^2\pi^2}\sin\left(\frac{m\pi}{L}x\right)\right]_{\frac{L}{2}}^{L} + \left[-\frac{L}{m\pi}\cos\left(\frac{m\pi}{L}x\right)\right]_{\frac{L}{2}}^{L}$$
$$= \frac{b_m}{2A}\int_0^{L} \frac{1}{2}\left(1 - \cos\left(\frac{2m\pi}{L}x\right)\right)dx$$

$$-\frac{L}{2m\pi}\cos\left(\frac{m\pi}{2}\right) + \frac{L}{m^2\pi^2}\sin\left(\frac{m\pi}{2}\right) - \left(-\frac{L}{m\pi}\cos(m\pi) + \frac{L}{m^2\pi^2}\sin(m\pi) - \left(-\frac{L}{2m\pi}\cos\left(\frac{m\pi}{2}\right) + \frac{L}{m^2\pi^2}\sin\left(\frac{m\pi}{2}\right)\right)\right)$$
$$-\frac{L}{m\pi}\cos(m\pi) + \frac{L}{m\pi}\cos\left(\frac{m\pi}{2}\right) = \frac{b_m}{4A}\left[x - \frac{L}{2m\pi}\sin\left(\frac{2m\pi}{L}x\right)\right]_0^{L}$$

$$-\frac{L}{2m\pi}\cos\left(\frac{m\pi}{2}\right) + \frac{L}{m^2\pi^2}\sin\left(\frac{m\pi}{2}\right) + \frac{L}{m\pi}\cos(m\pi) - \frac{L}{m^2\pi^2}\sin(m\pi) - \frac{L}{2m\pi}\cos\left(\frac{m\pi}{2}\right) + \frac{L}{m^2\pi^2}\sin\left(\frac{m\pi}{2}\right)$$
$$-\frac{L}{m\pi}\cos(m\pi) + \frac{L}{m\pi}\cos\left(\frac{m\pi}{2}\right) = \frac{b_m}{4A}\left(L - \frac{L}{2m\pi}\sin(2m\pi)\right)$$

$$\frac{2L}{m^2\pi^2}\sin\left(\frac{m\pi}{2}\right) = \frac{b_m L}{4A}$$

$$b_m = \frac{8A}{\pi^2 m^2}\sin\left(\frac{m\pi}{2}\right)$$

When m is even $\sin\left(\frac{m\pi}{2}\right) = 0$, $b_m = 0$

when m is odd $\sin\left(\frac{m\pi}{2}\right) = (-1)^{\frac{m-1}{2}}$

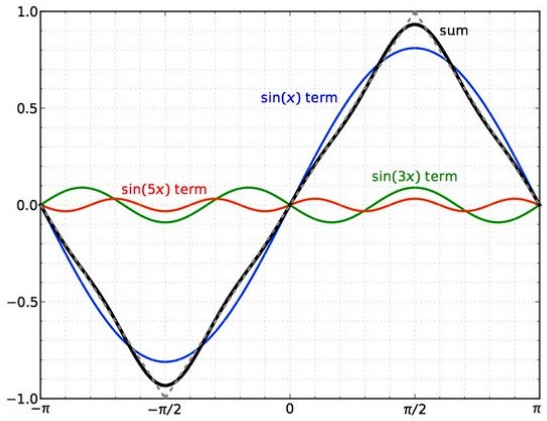
The other initial condition is that the wire is release with zero initial velocity, i.e. $\left.\frac{\partial y}{\partial t}\right|_{t=0} = 0$

$$\frac{\partial y}{\partial t} = \sum_{n=0}^{\infty} \sin\left(\frac{n\pi}{L}x\right)\left[a_n \frac{L}{n\pi c}\cos\left(\frac{n\pi}{L}ct\right) - b_n \frac{L}{n\pi c}\sin\left(\frac{n\pi}{L}ct\right)\right]$$

When $t = 0$ we see that $a_n = 0$. This gives an overall solution

$$y(x,t) = \frac{8A}{\pi^2}\sum_{n=1}^{\infty} \frac{(-1)^{\frac{n-1}{2}}}{n^2}\sin\left(\frac{n\pi}{L}x\right)\cos\left(\frac{n\pi}{L}ct\right) \quad \text{for odd } n$$

$$y(x,t) = \frac{8A}{\pi^2}\left(\sin\left(\frac{\pi}{L}x\right)\cos\left(\frac{\pi}{L}ct\right) \quad -\frac{1}{9}\sin\left(\frac{3\pi}{L}x\right)\cos\left(\frac{3\pi}{L}ct\right) \quad +\frac{1}{25}\sin\left(\frac{5\pi}{L}x\right)\cos\left(\frac{5\pi}{L}ct\right)\ldots\ldots\right)$$



This shows that when a string is plucked on a stringed musical instrument, not only the fundamental frequency is hears but many harmonics as well. If the string is plucked at a different position then there is a different distribution of harmonics as the Fourier coefficients are different, giving a different sound (even though the fundamental frequency is unchanged).

Schrödinger's Wave Equation for the H atom

You will already have seen in PHY2001 how separation of variables can be used to obtain an analytical solution for the quantum states of the hydrogen atom from the time independent Schrödinger. Generally for quantum problems in spherical polar coordinates, the variables are separable if there is a central potential, i.e. the potential energy is only a function of r only and is independent of θ, ϕ.

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V(r)\psi = E\psi$$

Expressing the Laplacian operator in spherical polar coordinates

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2} + \frac{2m}{\hbar^2}\big(E - V(r)\big)\psi = 0$$

$$\frac{1}{r^2}\left(2r\frac{\partial\psi}{\partial r} + r^2\frac{\partial^2\psi}{\partial r^2}\right) + \frac{1}{r^2\sin\theta}\left(\cos\theta\frac{\partial\psi}{\partial\theta} + \sin\theta\frac{\partial^2\psi}{\partial\theta^2}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2} + \frac{2m}{\hbar^2}\big(E - V(r)\big)\psi = 0$$

$$\frac{\partial^2\psi}{\partial r^2} + \frac{2}{r}\frac{\partial\psi}{\partial r} + \frac{\cos\theta}{r^2\sin\theta}\frac{\partial\psi}{\partial\theta} + \frac{1}{r^2}\frac{\partial^2\psi}{\partial\theta^2} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2} + \frac{2m}{\hbar^2}\big(E - V(r)\big)\psi = 0$$

Substituting $\psi(r,\theta,\phi) = R(r)\Theta(\theta)\Phi(\phi)$

$$\frac{d^2R}{dr^2}\Theta\Phi + \frac{2}{r}\frac{dR}{dr}\Theta\Phi + \frac{\cos\theta}{r^2\sin\theta}\frac{d\Theta}{d\theta}\mathrm{R}\Phi + \frac{1}{r^2}\frac{d^2\Theta}{d\theta^2}\mathrm{R}\Phi + \frac{1}{r^2\sin^2\theta}\frac{d^2\Phi}{d\phi^2}\mathrm{R}\Theta + \frac{2m}{\hbar^2}\big(E - V(r)\big)R\Theta\Phi = 0$$

Dividing by $\frac{R\Theta\Phi}{r^2\sin^2\theta}$ and putting all the ϕ terms on the RHS, this is set equal to a separation constant α

$$\sin^2\theta\frac{r^2}{R}\frac{d^2R}{dr^2} + \sin^2\theta\frac{2r}{R}\frac{dR}{dr} + \sin^2\theta\frac{d^2\Theta}{d\theta^2}\frac{1}{\Theta} + \sin\theta\cos\theta\frac{d\Theta}{d\theta}\frac{1}{\Theta} + r^2\sin^2\theta\frac{2m}{\hbar^2}\big(E - V(r)\big) = -\frac{d^2\Phi}{d\phi^2}\frac{1}{\Phi} = \alpha$$

which gives an ODE in ϕ

$$\frac{d^2\Phi}{d\phi^2} + \alpha\Phi = 0$$

This has a straightforward solution $\Phi = Ae^{m_l\phi}$ ($m_l = \sqrt{\alpha}$), where the quantum number m_l must be an integer to ensure the wavefunction is single valued.

The r, θ terms on the other side of the equation can be separated again. Dividing by $\sin^2\theta$, and setting equal to another separation constant β

$$\frac{r^2}{R}\frac{d^2R}{dr^2} + \frac{2r}{R}\frac{dR}{dr} + \frac{2mr^2}{\hbar^2}(V(r) - E) = \frac{\alpha}{\sin^2\theta} - \frac{d^2\Theta}{d\theta^2}\frac{1}{\Theta} - \frac{\cos\theta}{\sin\theta}\frac{d\Theta}{d\theta}\frac{1}{\Theta} = \beta$$

we get two more ODEs

$$\frac{d^2\Theta}{d\theta^2} + \frac{\cos\theta}{\sin\theta}\frac{d\Theta}{d\theta} + \left(\beta - \frac{m_l^2}{\sin^2\theta}\right)\Theta = 0$$

$$\frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr} + \frac{2m}{\hbar^2}\left(V(r) - E - \frac{\hbar^2}{2mr^2}\beta\right)R = 0$$

The latter 2 ODEs can be solved using appropriate substitutions and series solutions. In order for each power series (and hence the wavefunction) to be finite, the series are truncated after l and n terms ($l \geq 0,\ n \geq 1,\ l < n$) respectively, where $\beta = l(l+1)$.

5. Numerical Solutions of ODEs and PDEs

5.1. Finite Difference Methods

Solutions of ordinary and partial differential equations can be obtained when the problem is well-posed, i.e. it has sufficient and appropriate initial and boundary conditions. Analytical solutions to these problems can be limited to special cases, especially for PDEs. As a result it is necessary to use numerical techniques to obtain approximate solutions and with the use of modern computational power the errors in these approximations can minimised.

Numerical techniques used to solve differential equations are based on the principle of discretisation, where a continuous function to be solved is approximated by dividing the space up into an array of grid points (figure 5.1). For example, in a 2D space $0 \leq x \leq X$, $0 \leq y \leq Y$, this area is divided up into an evenly spaced rectangular grid $N \times M$ where

$$x = x_0, x_1, x_2, \dots x_n, \dots x_N,$$

$$y = y_0, y_1, y_2, \dots y_m, \dots y_M,$$

and a differential equation in $u(x, y)$ is solved numerically for points $u_{n,m} = u(x_n, y_m)$.

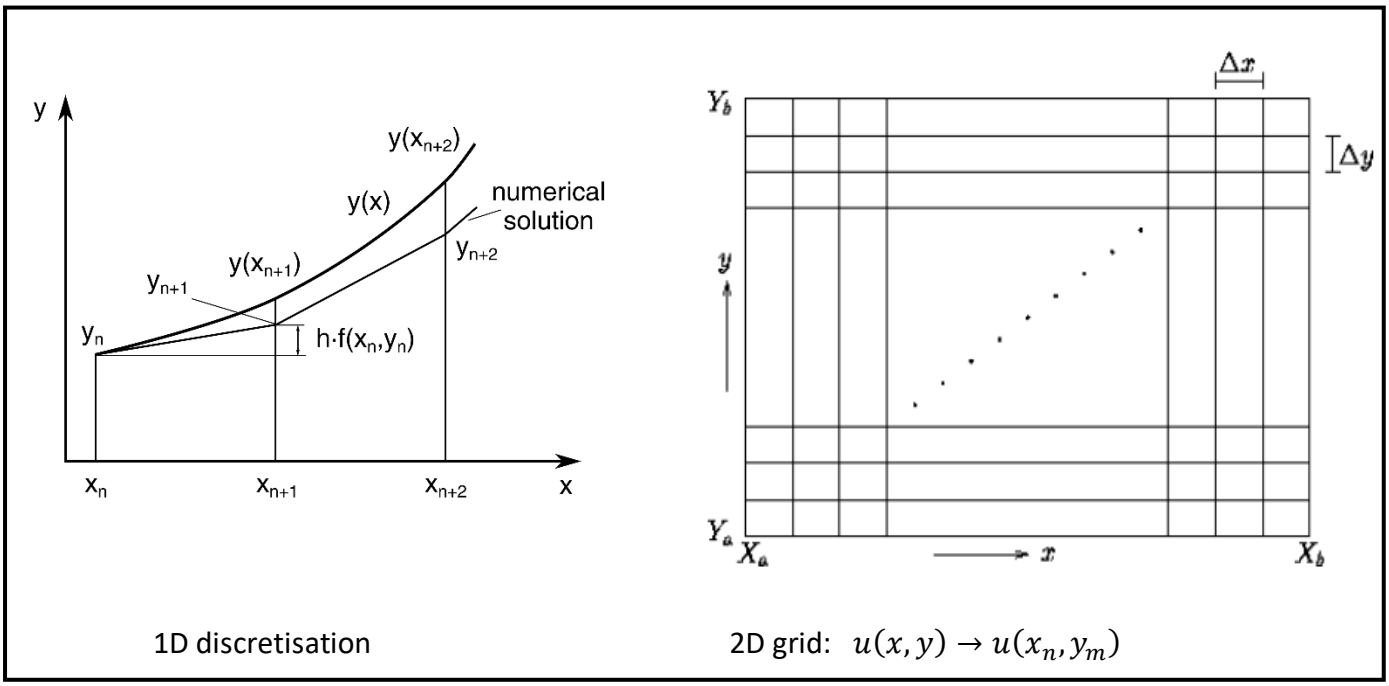


Figure 5.1: Demonstrating how a 1D and 2D function can be discretised

The value of a function at a particular grid point is related to the values at other grid points via the differentials. For a 1D function, the next point on the curve can be estimated by using the gradient at neighbouring points. Using Taylor's theorem a 1D function can be expanded around a position x as follows

$$u(x+\Delta x) = u(x) + u'(x)\Delta x + \frac{\Delta x^2}{2!}u''(x) + \frac{(\Delta x)^2}{3!}u'''(x) + \cdots$$

So to first order

$$u(x+\Delta x) = u(x) + u'(x)\Delta x + O(\Delta x^2)$$

where the term $O(\Delta x^2)$ is the error in the approximation which depends on $(\Delta x)^2$ or higher powers thereof. Finite difference methods are the simplest and most widely used technique for solving differential equations numerically. They are relatively straightforward to implement and computationally efficient. Euler's method is the simplest way of achieving this.

5.2. Ordinary Differential Equations

Euler's Method - Forward (Explicit)

This is a first order approximation where the gradient u_i' at position x_i is used to estimate the value of u_{i+1}. For a discrete point along the curve of a 1D function this is (from a 1st order Taylor expansion)

$$u_{i+1} \approx u_i + u_i' \, \Delta x$$

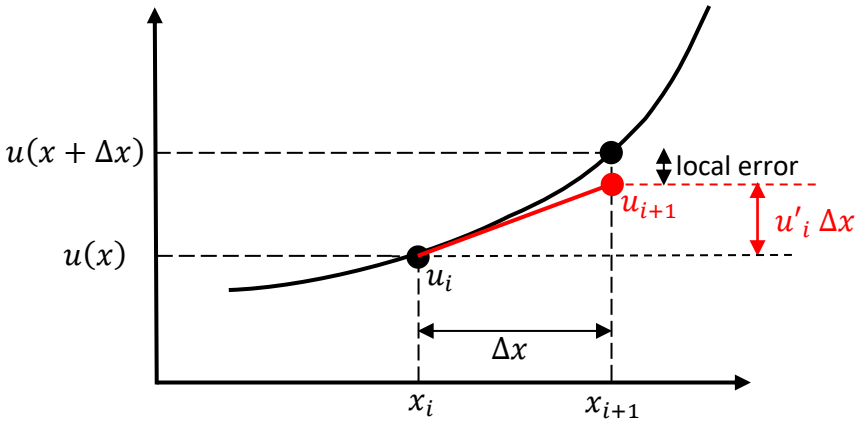


Figure 5.2: Euler's method where the gradient of a 1D function at x_i is used to a estimate the function value at $x_i + \Delta x$

Therefore, when an initial condition has been defined, a 1st order differential equation can be easily solved iteratively. This is known as the ***explicit Euler method***. For a first order ordinary differential equation:

$$\frac{du}{dx} = u' = f(x,u)$$

In discrete form to first order this is estimated as

$$u_{i+1} \approx u_i + f(x_i,u_i)\,\Delta x = u_i + f_i\,\Delta x \qquad \text{EULER FORWARD (EXPLICIT)}$$

For a single step there is an error which scales as the square of the step size $O(\Delta x^2)$. This is known as the ***local error*** $= [u_{i+1} - u(x+\Delta x)]$. If we start with an initial condition value $u(x_0)$ the value on the next grid point can be determined and if this is repeated N times, then a value $u(x_N)$ can be obtained. If $L = x_N - x_0$, where $N = L/\Delta x$ then the ***global (total accumulated) error*** $= [u_N - u(x_N)]$ is linear with the step size $O(\Delta x)$.

Consider a simple example, the 1D ODE

$$\frac{du}{dx} = u' = -u$$

It has an analytical solution

$$u(x) = A\exp(-x)$$

and if $y = 1$ at $x = 0$

$$u(x) = \exp(-x)$$

Let's say we need to find a solution in the range $x = 0$ to $x = 2$. We divide up this range into $N = 5$ sections, so $\Delta x = 0.4$ and $x_0, x_1, x_2, \ldots x_N = 0, 0.4, 0.8 \ldots 2.0$. Values of u_i can be found from

$$u_{i+1} \approx u_i + u'_i\,\Delta x$$

$$u_{i+1} \approx u_i - u_i\,\Delta x = u_i(1 - \Delta x)$$

Starting at $u_0 = 1$ at $x = 0$ and we iterate for a stepsize of $\Delta x = 0.4$

i	x_i	u_i	$u(x)$	Error $\frac{f(x)-f_i}{f(x)}$ %
0	0.00	1.000	1.000	0.0
1	0.40	0.600	0.670	10.5
2	0.80	0.360	0.449	19.9
3	1.20	0.216	0.301	28.3
4	1.60	0.130	0.202	35.8
5	2.00	0.078	0.135	42.5

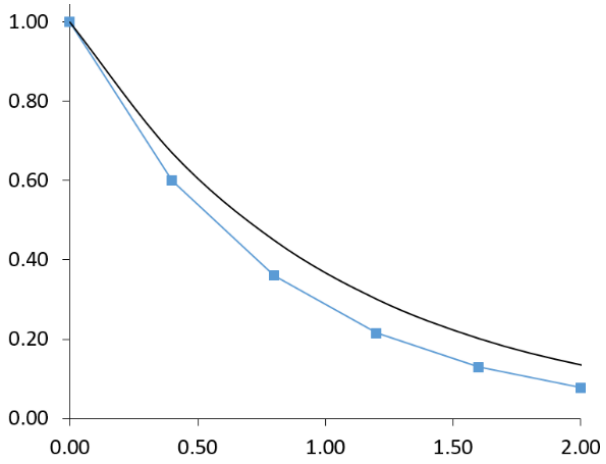


Figure 5.3: Forward (explicit) Euler solution for $\frac{du}{dx} = -u(x)$ with $N = 5$

This produces a large global error due to the large step size. If we reduce the step size to $\Delta x = 0.2$, then the error reduces by half indicating showing the linear dependence on Δx.

i	x_i	u_i	$u(x)$	Error $\frac{f(x)-f_i}{f(x)}$ %
0	0.00	1.000	1.000	0.0
1	0.20	0.800	0.819	2.3
2	0.40	0.640	0.670	4.5
3	0.60	0.512	0.549	6.7
4	0.80	0.410	0.449	8.8
5	1.00	0.328	0.368	10.9
6	1.20	0.262	0.301	13.0
7	1.40	0.210	0.247	15.0
8	1.60	0.168	0.202	16.9
9	1.80	0.134	0.165	18.8
10	2.00	0.107	0.135	20.7

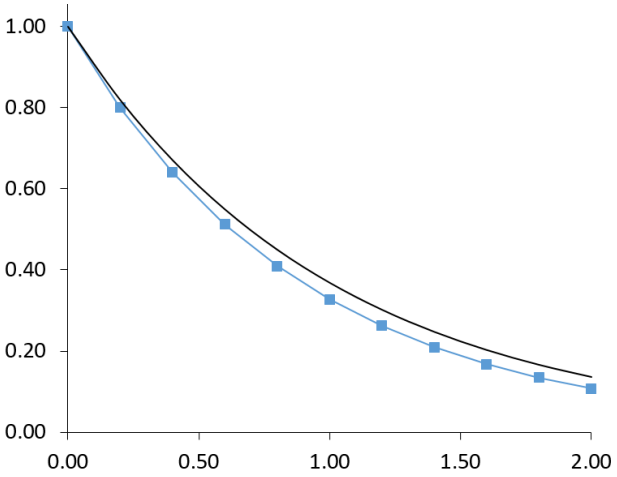


Figure 5.4: Same a Figure 3 but with $N = 10$

Euler's method features in the story of female, African American mathematicians who contributed to NASA's spacecraft trajectory calculations in the move *Hidden Figures* - https://www.youtube.com/watch?v=7UUd9t7XspE

Euler's Method - Backward (Implicit)

This Euler method is an example of a forward (or explicit) finite difference method. Working from initial conditions, the "next" value can be calculated directly from the previous values or u_{i+1} is an explicit function of the currently known values u_i and u_i'. However, the iteration process can be "traced backwards" instead. If we move a step backwards then

$$u(x - \Delta x) = u(x) - u'(x)\Delta x + O(\Delta x^2)$$

which for discrete points gives the ***backwards approximation***

$$u_{i-1} \approx u_i - u'_i \, \Delta x$$

$u_{i-1} \approx u_i - f_i \, \Delta x$ EULER BACKWARD (IMPLICIT)

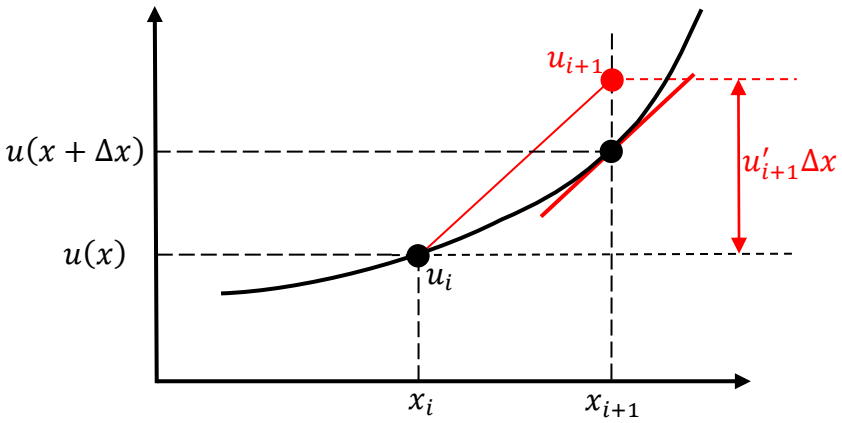


Figure 5.5: Euler's backwards method where the gradient of a 1D function at x_{i+1} is used to estimate increment the function from x_i

But we don't know the final value u_N only the initial condition u_0, so can we iterate backwards? Let's look again at the previous example where $u_i' = -u_i$, $u_0 = 1$, writing out expressions at each grid point for $N = 5$, working backwards

$u_4 \approx u_5 - u'_5\,\Delta x = u_5(1+\Delta x)$

$u_3 \approx u_4(1+\Delta x)$

$u_2 \approx u_3(1+\Delta x)$

$u_1 \approx u_2(1+\Delta x)$

$u_0 \approx u_1(1+\Delta x)$

This backwards iteration is known as an ***implicit method***. The next value of the function is not calculated directly but as there are 5 simultaneous equations and 5 unknowns (a relatively simple example in this case), this is sufficient to find each value u_i.

i	x_i	u_i	$u(x)$	$100 \times \frac{u(x)-u_i}{u(x)}$ %
0.00	0.00	1.000	1.000	0.0
1.00	0.40	0.714	0.670	-6.6
2.00	0.80	0.510	0.449	-13.5
3.00	1.20	0.364	0.301	-21.0
4.00	1.60	0.260	0.202	-28.9
5.00	2.00	0.186	0.135	-37.4

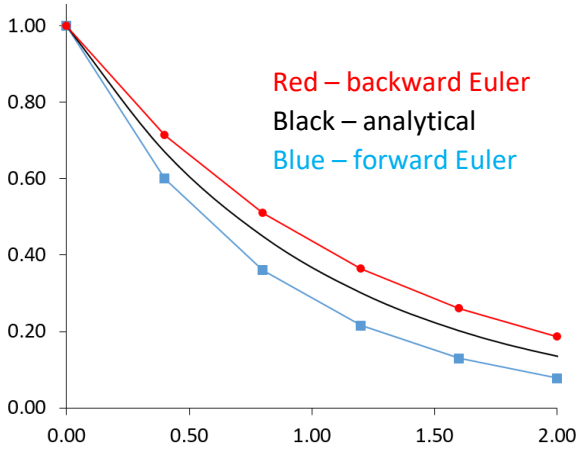


Figure 5.6: Comparison between backwards and forwards Euler solutions for $f(x,u) = -u$

Midpoint Approximation

The main limitation of the forward and backward Euler methods is that the estimate of the differential is not very good as the gradient of the curve is over- or under-estimated. A more accurate estimate of the gradient can be obtained half way between the points at $x_{i+\frac{1}{2}} = x_i + \frac{\Delta x}{2}$, i.e. at the ***mid-point***.

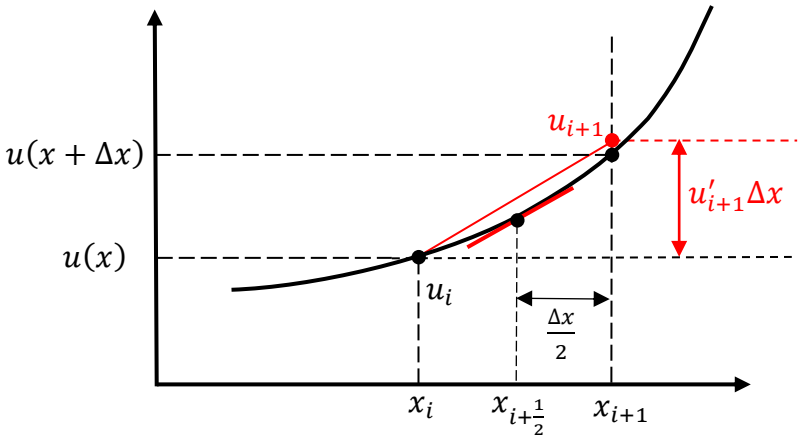


Figure 5.7: The mid-point approximation where the value of the function at x_{i+1} is calculated by using the gradient at the midpoint $x_{i+\frac{1}{2}}$

For this case the estimate is now

$$u(x+\Delta x) \approx u(x) + u'\left(x+\frac{\Delta x}{2}\right)\Delta x$$

The gradient at the midpoint is

$$u'\left(x+\frac{\Delta x}{2}\right) = f\left(x+\frac{\Delta x}{2}, u\left(x+\frac{\Delta x}{2}\right)\right)$$

In discrete terms

$$u'_{i+\frac{1}{2}} = f_{i+\frac{1}{2}}$$

$u_{i+1} \approx u_i + f_{i+\frac{1}{2}}\Delta x$	MIDPOINT (FORWARD)

We know the simple forwards Euler is not accurate but it is sufficient for estimating the gradient $f_{i+\frac{1}{2}}$ because this is a lower order correction

$$f_{i+\frac{1}{2}} \approx f_i + {f_i}'\frac{\Delta x}{2}$$

$$u_{i+1} \approx u_i + \left(f_i + {f_i}'\frac{\Delta x}{2}\right)\Delta x$$

Remembering ${u'}_i = f_i$

$$u_{i+1} \approx u_i + {u_i}'\Delta x + {f_i}'\frac{(\Delta x)^2}{2}$$

So compared to the simple Euler method, there is an additional term in $(\Delta x)^2$

For the simple function, where $f(x,u) = -u(x)$ then ${u_i}' = -u_i$ and ${f_i}' = -{u_i}' = u_i$

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

i	x_i	u_i	$u(x)$	Error $\frac{f(x)-f_i}{f(x)}$ %
0.00	0.00	1.000	1.000	0.0
1.00	0.40	0.680	0.670	-1.4
2.00	0.80	0.462	0.449	-2.9
3.00	1.20	0.314	0.301	-4.4
4.00	1.60	0.214	0.202	-5.9
5.00	2.00	0.145	0.135	-7.4

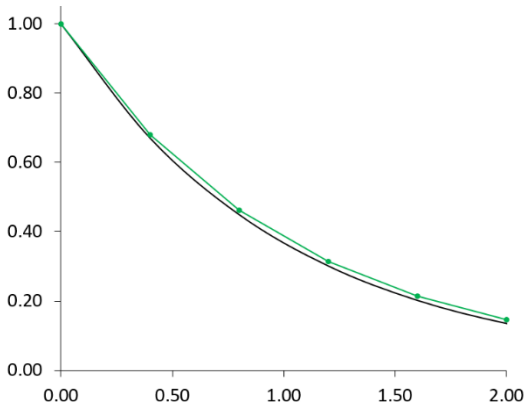


Figure 5.8: Comparison between the analytical solution and the mid-point approximation

This is a substantially better estimate than the simple Euler, even when the number of steps is doubled.

Runge-Kutta Methods

The Euler and mid-point methods are part of a family of techniques called Runge-Kutta methods. The 1st order Runge-Kutta method or RK1, where the increment k_1 is determined from the gradient at the start of the interval x_i (therefore this is just the simple forward Euler)

$$u_{i+1} \approx u_i + k_1$$

$$k_1 = \Delta x . f_i$$

The 2nd order estimate or RK2 where the increment k_2 is determined from the gradient at midpoint between x_i and x_{i+1} where the value of $u_{i+\frac{1}{2}}$ and hence $f_{i+\frac{1}{2}}$ is estimated using k_1 (therefore this is just the midpoint method which uses the Euler to estimate the gradient at the midpoint)

$$u_{i+1} \approx u_i + k_2$$

$$k_2 = \Delta x . f_{i+\frac{1}{2}}^{(k_1)}$$

The third estimate RK3 also uses the gradient at the midpoint but with the value of $u_{i+\frac{1}{2}}$ and hence $f_{i+\frac{1}{2}}$ is estimated using k_2 (this is also the midpoint method but instead uses the value obtained from RK2 to estimate the gradient at the midpoint)

$$u_{i+1} \approx u_i + k_3$$

$$k_3 = \Delta x . f_{i+\frac{1}{2}}^{(k_2)}$$

The fourth estimate RK4 the increment k_4 are determined from the gradient at end of the interval x_{i+1}. For this the value of u_{i+1} and hence f_{i+1} is estimated using k_3 (this is the backwards Euler but instead uses the value obtained from RK3 to estimate the gradient at the end point)

$$u_{i+1} \approx u_i + k_4$$

$$k_4 = \Delta x . f_{i+1}^{(k_3)}$$

These four estimates when combined they are known as the 4th order Runge-Kutta method as follows

$$u_{i+1} \approx u_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

where the two midpoint estimates k_2 and k_3 are given double the weighting of the endpoint estimates k_1 and k_4 as they are more accurate. This gives a local error of $O(\Delta x^5)$ and a global (or total accumulated) error of $O(\Delta x^4)$.

RK4 is widely used to accurately solved ordinary differential equations as it is far more accurate than the RK1 (Euler) and RK2 (Midpoint) methods while still being computationally efficient.

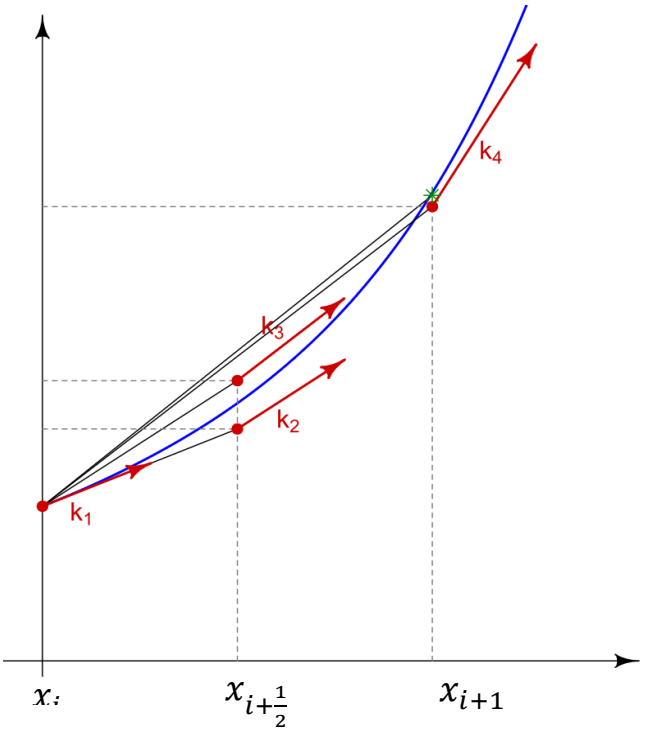


Figure 5.9: Gradients at the start k_1, mid-point k_2 and k_3, and end k_2 of the iteration increment used in a 4th order Runge-Kutta approximation.

Second Order Differentials

These methods can be applied to 2nd order differential equations as well. The 2nd order derivative can be expressed using Taylor's theorem to third order. For a forward and backward increment:

$$u(x+\Delta x) = u(x) + u'(x)\Delta x + \frac{\Delta x^2}{2!}u''(x) + \frac{\Delta x^3}{3!}u'''(x) + O(\Delta x^4)$$

$$u(x-\Delta x) = u(x) - u'(x)\Delta x + \frac{\Delta x^2}{2!}u''(x) - \frac{\Delta x^3}{3!}u'''(x) + O(\Delta x^4)$$

Adding these together

$$u''(x) = \frac{u(x+\Delta x) - 2u(x) + u(x-\Delta x)}{(\Delta x)^2} + O(\Delta x^2)$$

or for the discretised function

$$u''_m = \frac{u_{m+1} - 2u_m + u_{m-1}}{(\Delta x)^2} + O(\Delta x^2) \qquad \text{2nd ORDER DIFFERENTIAL TAYLOR APPROXIMATION}$$

Examples

Accelerating Car

Let's trying getting a numerical answer to the problem of an accelerating car we looked at previously. The equation of motion is

$$\frac{dv}{dt} + \frac{\rho AC}{2m} v^2 = p(t) \frac{F_{max}}{m}$$

Where ρ – air density, C – drag coefficient, A – cross sectional area of the car, F_{max} – maximum forward force from the engine, $p(t)$ – fraction which the accelerator pedal is pressed, m – mass of the car.

Euler

Previously a solution was obtained by linearizing the equation by assuming the car's velocity changes by small amounts around an average velocity. To solve the non-linear equation by using the forward Euler method we can rewrite the equation of motion as a finite difference equation

$$v_{i+1} = v_i + \Delta t \left(\frac{dv}{dt}\right)_i$$

$$v_{i+1} = v_i + \Delta t \left(p \frac{F_{max}}{m} - \frac{\rho AC}{2m} v_i^2\right)$$

For $m = 1000,\ \rho AC = 2, F_{max} = 3000,\ p = 0.6$

$$v_{i+1} = v_i + \Delta t\left(1.8 - 0.001 v_i^2\right)$$

We want to find how long it takes for the speed to accelerate from an initial speed of 30 m/s to 33 m/s. Let's take time steps $\Delta t = 0.5\ s$,

i	t_i	v_i
0	0	30.000
1	0.5	30.450
2	1	30.886
3	1.5	31.309
4	2	31.719
5	2.5	32.116
6	3	32.500
7	3.5	32.872
8	4	33.232

This predicts that it takes 3.5-4 s which is in good agreement with the prediction of 3.719 s from the linearized analytical solution, but how accurate is this value? If we gradually reduce Δt, the following answers are obtained (it is simple to set up in Excel)

Δt	N	t
0.2	19	3.6 – 3.8
0.1	38	3.7 – 3.8
0.05	75	3.7 – 3.75
0.02	187	3.72 – 3.74
0.01	374	3.73 – 3.74
0.005	747	3.73 – 3.735
0.002	1867	3.734

This gives a difference of 0.015 (0.4%) showing that the linearized model of the equation of motion is a pretty good approximation for this situation.

Saturn V Rocket

The Saturn V rocket which took Apollo spacecraft to the Moon is the most powerful machine built by humans. It had a total mass of $M_0 = 3{,}000{,}000$ kg just before launch and during its 1st stage burned fuel at a rate of $L = 13{,}500$ kg/s so that gas was expelled out the engine nozzles at a speed of $u = 2600$ m/s. At the end of the 1st stage, the rocket's mass had reduced to $M = 800{,}000$ kg. The evolution of the velocity v and mass m of the rocket are governed by the following differential equation

$$\frac{dv}{dm} = \frac{g}{L} - \frac{u}{m}$$

where $g = 9.8$ ms^{-2} is the acceleration due to gravity (assumed constant during the first stage). We can solve this analytically by separating the variables and integrating from the ignition point to the end of the 1st stage

$$\int_0^{v_f} dv = \int_{M_0}^{M} \left(\frac{g}{L} - \frac{u}{m}\right) dm$$

$$v_f = \left[\frac{g}{L}m - u \ln m\right]_{M_0}^{M}$$

$$v_f = -u \ln \frac{M}{M_0} + \frac{g}{L}(M - M_0)$$

$$v_f = u \ln \frac{M_0}{M} - \frac{g}{L}(M_0 - M)$$

Inputting the parameters this gives $v_f = 1837$ m/s, which is a pretty good estimate of the actual speed achieved during the launches – 2300 m/s. This simple model does not take into account the reduction in gravity at higher altitudes and the fact that the rocket did not travel vertically upwards (it had travelled about 100 km from the launch site by the end of the 1st stage).

Let's see how well the Euler method does in estimating this velocity. The differential equation becomes the following finite difference equation

$$\left(\frac{dv}{dm}\right)_i = \frac{g}{L} - \frac{u}{m_i}$$

$$v_{i+1} = v_i + \left(\frac{dv}{dm}\right)_i \Delta m$$

$$v_{i+1} = v_i + \Delta m \left(\frac{g}{L} - \frac{u}{m_i}\right)$$

If we take 5 steps then $\Delta m = -440{,}000$ kg, so the final velocity can be obtained as shown in the table below complete the table below

i	m (kg)	v_i (m/s)	$\left(\frac{dv}{dm}\right)_i$
0	3,000,000	0	-0.000140
1	2,560,000	61.6	-0.000289
2	2,120,000	188.7	-0.000500
3	1,680,000	408.6	-0.000821
4	1,240,000	769.9	-0.00137
5	800,000	1372.7	-0.00252

So the velocity at the end of the first stage is estimated to be 1373 m/s which corresponds to a global error of 464 m/s (25%). Let's see how well the midpoint method works for the same number of steps. It has a difference equation

$$v_{i+1} = v_i + \left(\frac{dv}{dm}\right)_{i+\frac{1}{2}} \Delta m$$

$$\left(\frac{dv}{dm}\right)_{i+\frac{1}{2}} = \frac{g}{L} - \frac{u}{m_{i+\frac{1}{2}}}$$

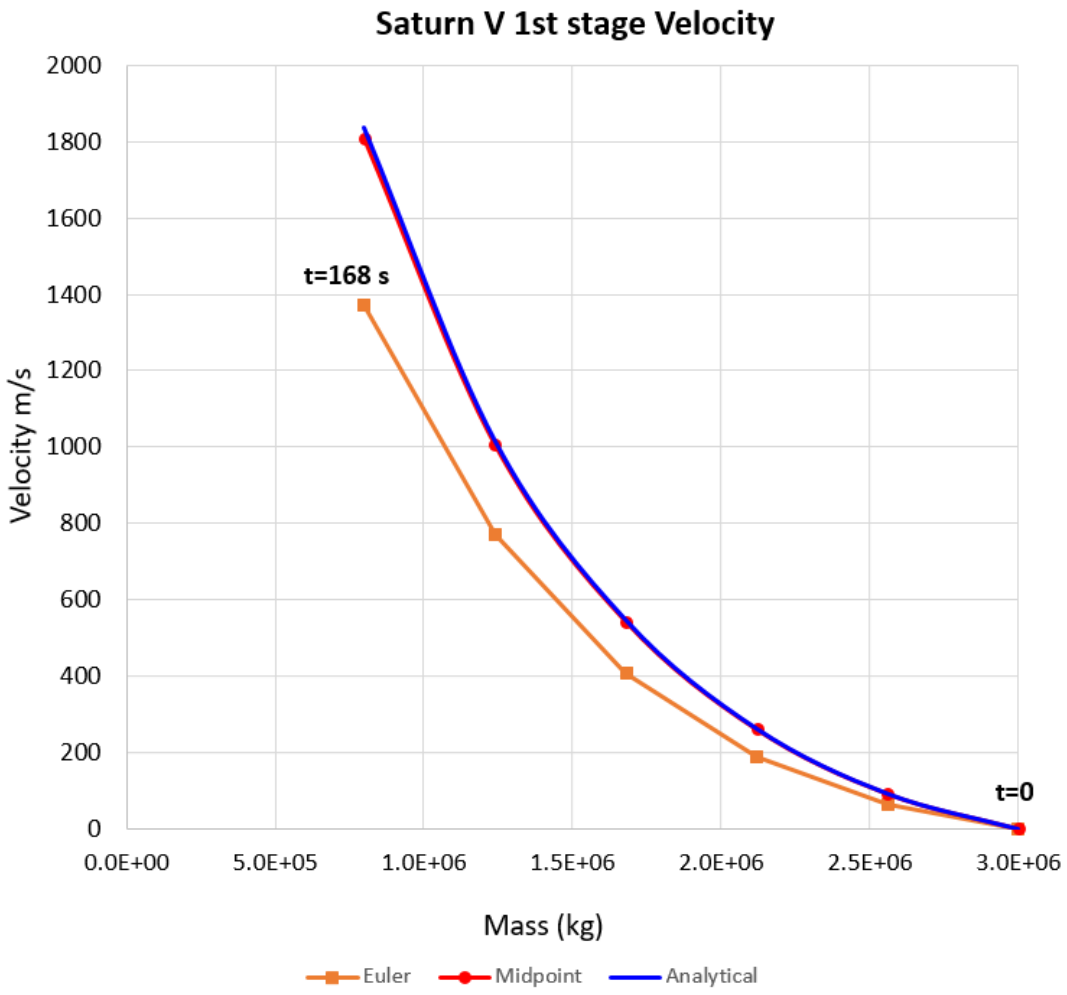
where

$$m_{i+\frac{1}{2}} = m_i + \frac{\Delta m}{2}$$

$$v_{i+1} = v_i + \left(\frac{g}{L} - \frac{u}{m_i + \frac{\Delta m}{2}}\right)\Delta m$$

i	m (kg)	v_i (m/s)	$\left(\frac{dv}{dm}\right)_{i+\frac{1}{2}}$
0	3,000,000	0	-0.00021
1	2,560,000	91.8	-0.00038
2	2,120,000	260.9	-0.00064
3	1,680,000	543.3	-0.00105
4	1,240,000	1007.1	-0.00182
5	800,000	1809.0	-0.00376

This is a big improvement compared to Euler, with the global error now 26 m/s (1.4%). Since the global error in the Euler method scales linearly with Δm, it would take approximately 18 times more steps (90 iterations) to get the same precision as the midpoint method.



5.3. PDEs – Heat Equation

The heat equation (and the Schrödinger equation) is a parabolic equation can be solved numerical by *marching forward in time*. In 1D we have

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

Explicit Finite Difference Method

For a function $T(x,t) \rightarrow T_{m,n}(x_m, t_n)$, the solution obtained by explicit FDM expresses the partial differentials as finite differences

$$\frac{\partial T}{\partial t} \approx \frac{T_{m,n+1} - T_{m,n}}{\Delta t}$$

$$\frac{\partial^2 T}{\partial x^2} \approx \frac{T_{m+1,n} - 2T_{m,n} + T_{m-1,n}}{(\Delta x)^2}$$

$$\frac{T_{m,n+1} - T_{m,n}}{\Delta t} \approx D\frac{T_{m+1,n} - 2T_{m,n} + T_{m-1,n}}{(\Delta x)^2} \qquad \text{FORWARD TIME (EXPLICIT), SPACE CENTRED}$$

This method is known as space centred as the second order differential is obtained symmetrically from the function's values at $m-1,\ m,\ m+1$.

$$T_{m,n+1} = T_{m,n} + \lambda\left(T_{m+1,n} - 2T_{m,n} + T_{m-1,n}\right)$$

$$T_{m,n+1} = \lambda T_{m+1,n} + T_{m,n}(1-2\lambda) + \lambda T_{m-1,n} \quad \text{where} \quad \lambda = D\frac{\Delta t}{(\Delta x)^2}$$

For example, consider a rod of length L which is insulated so that heat can only flow along the rod. If the initial temperature distribution $T(x,0)$ and boundary conditions at the end of the rod are defined, then the points on the grid can be obtained from using the equation above. Consider the whole rod at the same initial temperature $T(x,0) = 0$. At $t = 0$ with one end ($x = 0$) put in perfect thermal contact with a heat reservoir at a temperature of $T(0,t) = 100$ while the other end ($x = 10$) is connected to a heat reservoir $T(10,t) = 0$. For a diffusivity of $D = 1$, let's create an M=10 x N=20 grid where $\Delta x = 1$ and $\Delta t = 0.2$, giving $\lambda = 0.2$. The initial and boundary conditions are shown in bold below

		x ⟶											
t			$m=0$	1	2	3	4	5	6	7	8	9	10
▼	0	$n=0$	**100.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
	0.2	1	**100.0**										**0.0**
	0.4	2	**100.0**										**0.0**
	0.6	3	**100.0**										**0.0**
	0.8	4	**100.0**										**0.0**
	1.0	5	**100.0**										**0.0**
	1.2	6	**100.0**										**0.0**
	1.4	7	**100.0**										**0.0**
	1.6	8	**100.0**										**0.0**
	1.8	9	**100.0**										**0.0**
	2.0	10	**100.0**										**0.0**

Figure 5.10: A 10 x 20 (x,t) grid with the initial and boundary conditions for a bar initially at 0 °C connected at one end to a heat reservoir at 100 °C

Each row of the grid is filled up using the equation above to obtain the spatial distribution of the temperature for the next time step. This is then repeated for the next time step filling the grid from top to bottom so that the process is iterated in time.

	$m=0$	1	2	3	4
$n=0$	**100.0**	**0.0**	**0.0**	**0.0**	**0.0**
1	**100.0**	20.0	0.0	0.0	0.0
2	**100.0**	32.0	4.0	0.0	0.0
3	**100.0**	40.0	8.8	0.8	0.0

$\times\lambda$ $\times(1-2\lambda)$ $\lambda\times$

$\times\lambda$ $\times(1-2\lambda)$ $\lambda\times$

Figure 5.11: Demonstration of the relationship between different grid points for the explicit finite difference method

In this way the time evolution of the temperature distribution can be incremented. The figure below plots $T(x,t)$ at four later times and shows how heat flows down the rod.

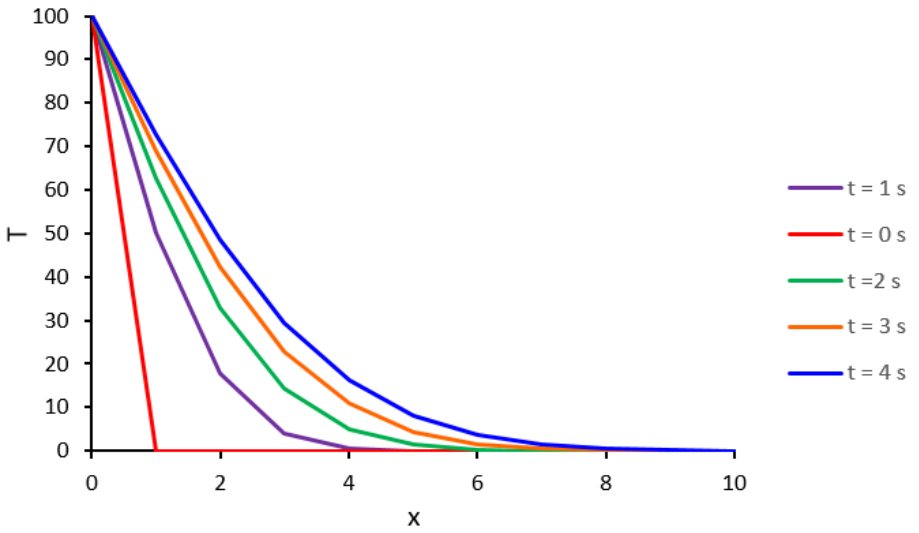


Figure 5.12: Solution to heat equation for different time intervals

The explicit method is a very simple and computationally efficient way of solving a diffusion equation. The global error in the calculation is $O(\Delta t)$ for time and $O(\Delta x^2)$ for space. However, the iteration process is only stable for $\lambda \leq 0.5$. If this condition is not obeyed, each iteration step is overestimated causing the calculation to diverge from the actual solution. This puts constraints on the values of Δt and Δx required.

The stability of numerical calculations is obviously important for a sensible solution to be obtained. The stability criteria for linear PDEs can be predicted by von Neumann stability analysis - https://en.wikipedia.org/wiki/Von_Neumann_stability_analysis, but this is beyond the scope of the current course.

Implicit Finite Difference Method (backward in time)

As with the Euler method, an implicit (backwards in time) approach can be employed to solve the heat equation. This has the advantage of being inherently stable but requires some linear algebra to solve the simultaneous equations which couple the function's values on different points of the grid. Now the heat equation on the grid is expressed as (note the 2nd order space derivative is now evaluated at n+1 rather than n)

$$\frac{T_{m,n+1} - T_{m,n}}{\Delta t} \approx D \frac{T_{m+1,n+1} - 2T_{m,n+1} + T_{m-1,n+1}}{(\Delta x)^2} \qquad \text{BACKWARD TIME (IMPLICIT), SPACE CENTRED}$$

$$T_{m,n} = -\lambda T_{m+1,n+1} + T_{m,n+1}(1 + 2\lambda) - \lambda T_{m+1,n+1} \quad \text{where} \quad \lambda = D\frac{\Delta t}{(\Delta x)^2}$$

The relationship between different grid points for this implicit scheme is now

	$m = 0$	1	2	3
$n = 0$	**100.0**	**0.0**	**0.0**	**0.0**
1	**100.0**	$\times -\lambda$	$\times (1 + 2\lambda)$	$\times -\lambda$
2	**100.0**			
3	**100.0**			

Figure 5.13: Demonstration of the relationship between different grid points for the implicit finite difference method

Therefore, if we are filling the grid from initial conditions at $t = 0$ $(n = 0)$, then we cannot directly calculate the grid points at the next time step $n = 1$. However, they can be obtained by solving the series of algebraic (simultaneous) equations. In this way the differential equation is converted into a series of linear algebraic equations which can be represented by the matrix equation shown below. Like analytical solutions, if sufficient initial and boundary conditions are specified a solution can be obtained.

$$\begin{pmatrix} (1+2\lambda) & -\lambda & 0 & 0 & 0 \\ -\lambda & (1+2\lambda) & -\lambda & 0 & 0 \\ 0 & -\lambda & (1+2\lambda) & -\lambda & 0 \\ & & \ddots & \ddots & \\ 0 & 0 & -\lambda & (1+2\lambda) & -\lambda \\ 0 & 0 & 0 & -\lambda & (1+2\lambda) \end{pmatrix} \begin{pmatrix} T_{1,n+1} \\ T_{2,n+1} \\ T_{3,n+1} \\ \vdots \\ T_{M-2,n+1} \\ T_{M-1,n+1} \end{pmatrix} = \begin{pmatrix} T_{1,n} \\ T_{2,n} \\ T_{3,n} \\ \vdots \\ T_{M-2,n} \\ T_{M-1,n} \end{pmatrix} + \lambda \begin{pmatrix} T_{0,n+1} \\ 0 \\ 0 \\ \vdots \\ 0 \\ T_{M,n+1} \end{pmatrix}$$

This type of linear algebra problem can be solved by finding the inverse to the matrix but that becomes too inefficient as the grid gets larger. A number of different techniques such as Gaussian elimination can be used to solve this computationally. For a sparsely populated matrix (tri-diagonal) such as this, Crout's Algorithm is particularly efficient. See the Appendix to see how this is implemented for this implicit scheme.

The advantage of the implicit technique is that, unlike the explicit scheme, it is inherently stable. However, the solution of the algebraic equations requires more computational steps making it slower.

Crank-Nicolson Method

The Crank-Nicolson estimates the 2nd order differential in space as an average of the explicit method at point n, m and the implicit method at point $n, m + 1$.

$$\frac{T_{m,n+1}-T_{m,n}}{\Delta t} \approx \frac{D}{2}\frac{T_{m+1,n+1}-2T_{m,n+1}+T_{m-1,n+1}}{(\Delta x)^2} + \frac{D}{2}\frac{T_{m+1,n}-2T_{m,n}+T_{m-1,n}}{(\Delta x)^2}$$

IMPLICIT EXPLICIT

using $\lambda = D\frac{\Delta t}{(\Delta x)^2}$

$$-\lambda T_{m+1,n+1} + 2(\lambda+1)T_{m,n+1} - \lambda T_{m-1,n+1} = \lambda T_{m+1,n} + 2(1-\lambda)T_{m,n} + \lambda T_{m-1,n}$$

CRANK-NICOLSON METHOD

As with the implicit stable, Crank-Nicolson is stable but improves on the global error, so that it depends on second order in both time and space $O(\Delta x^2) + O(\Delta t^2)$. So when solving parabolic PDEs which are to first order in time and second order in space, such as the heat equation and the time dependent Schrödinger equation, the Crank-Nicolson method is commonly used to obtain the most accurate solutions.

Summary of Finite Difference Methods for Solving Diffusion Equations

The algebraic connection between different points on the grid can be visualised with a "stencil". The Figure below shows this for the explicit, implicit and Crank-Nicolson finite difference solutions to the diffusion equation.

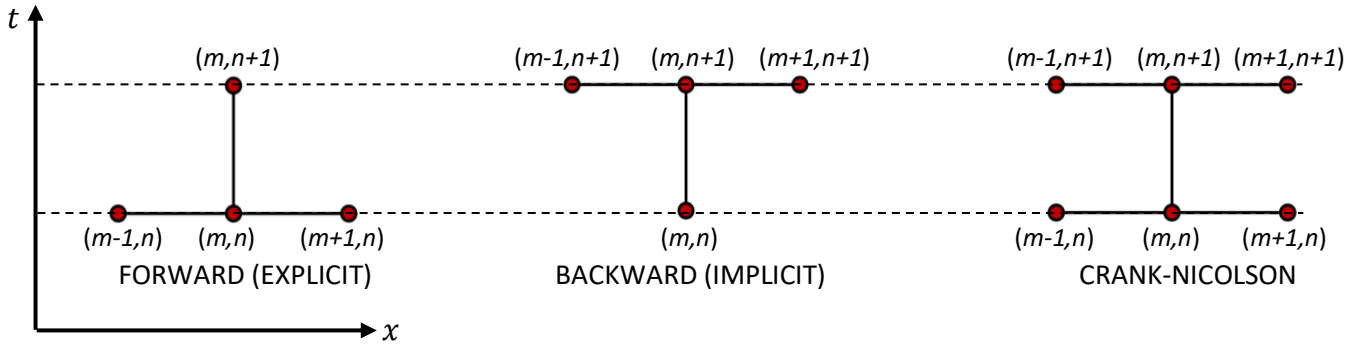


Figure 5.14: Stencil representation of different finite difference methods for solving the heat equation.

Usually the Crank–Nicolson scheme is the most accurate scheme for small time steps. The explicit scheme is the least accurate and can be unstable, but is also the easiest to implement and the least numerically intensive. The implicit scheme works the best for large time steps.

5.4. PDEs - Wave Equation

The wave equation is a 2nd order hyperbolic PDE. Like parabolic equations it can solved *marching forward in time* with suitable initial/boundary conditions. Consider the 1D wave equation

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

We can express these 2nd order differentials as finite differences
(remember m is the distance x index and n is the time t index)

$$v^2\frac{u_{m+1,n}-2u_{m,n}+u_{m-1,n}}{(\Delta x)^2} = \frac{u_{m,n+1}-2u_{m,n}+u_{m,n-1}}{(\Delta t)^2}$$

$$u_{m,n+1} = r\left(u_{m+1,n}+u_{m-1,n}\right) + 2u_{m,n}(1-r) - u_{m,n-1}$$

where $r = \left(\frac{v\Delta t}{\Delta x}\right)^2$. The stencil for this finite difference equation is

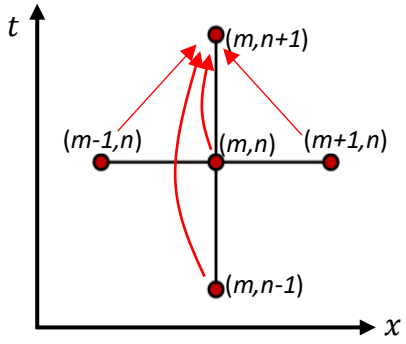


Figure 5.15: Stencil representation of for finite difference methods for solving the wave equation.

So the next time forward step in time is calculated from the current time point and the two positions adjacent to this, and the previous time point.

5.5. PDEs – Laplace/Poisson's Equation

Laplace's and Poisson's equations are 2nd order elliptic PDEs. Analytical solutions can be obtained for some configurations of the boundary conditions but numerical solutions are generally needed. If we consider the Laplace's equation it describes the electric potential V distribution in a space with no free charges.

$\nabla^2 V(x,y,z) = 0$ which in 2D is

$$\frac{\partial^2 V(x,y)}{\partial x^2} + \frac{\partial^2 V(x,y)}{\partial y^2} = 0$$

Within a region of space $0 < x < X,\ \ 0 < y < Y$, a finite difference grid of $M \times N$ where $\Delta x = \frac{X}{M},\ \ \Delta y = \frac{Y}{N}$, Laplace's equation can be written

$$\frac{V_{m,n-1} - 2V_{m,n} + V_{m,n+1}}{(\Delta x)^2} + \frac{V_{m+1,n} - 2V_{m,n} + V_{m-1,n}}{(\Delta y)^2} = 0$$

If we choose the grid so that $\Delta x = \Delta y$, then we get a simple expression for a grid point based on the average of the values of its 4 closest neighbours

$$V_{m,n} = \frac{1}{4}\left(V_{m,n-1} + V_{m,n+1} + V_{m+1,n} + V_{m-1,n}\right)$$

which can be represented by the following stencil

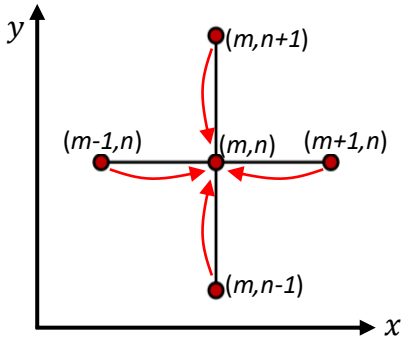


Figure 5.16: Stencil representation of for finite difference methods for solving the Laplace equation.

The finite difference equations which relate grid points can be solved with directly algebraic techniques, but it is more common to use relaxation techniques. These iterative methods use an initial guess at the solution which is

then allowed to slowly "*relax*" towards the true solution, reducing the errors as it does so. In its simplest form the following process is followed:

- An solution is "*guessed*", i.e. a value at each grid point $V_{m,n}$ is defined
- The boundary conditions are applied
- A new value $\overline{V}_{m,n}$ at each grid point is calculated using to create new values
  $$\overline{V}_{m,n} = \frac{1}{4}\left(V_{m,n-1} + V_{m,n+1} + V_{m+1,n} + V_{m-1,n}\right)$$
- The iteration is repeated until a tolerance δ is reached
  $$\left|\frac{\overline{V}_{m,n} - V_{m,n}}{V_{m,n}}\right| < \delta$$

Consider the following physical problem. There are four electrodes, two with +100 V potential and two with -100V potential as shown below (in a quadrupole formation). These act as boundary conditions for the problem. Our initial "guess" for the other grid points is simply to set them to zero – Iteration 0. Calculating the new values for the grid points yields – Iteration 1 (note that the undefined points along the edges are calculated from the average of the nearest 3 points and an average of 2 for the corners).

Iteration 0

0	0	100	100	100	100	100	100	100	0	0
0	0	0	0	0	0	0	0	0	0	0
-100	0	0	0	0	0	0	0	0	0	-100
-100	0	0	0	0	0	0	0	0	0	-100
-100	0	0	0	0	0	0	0	0	0	-100
-100	0	0	0	0	0	0	0	0	0	-100
-100	0	0	0	0	0	0	0	0	0	-100
-100	0	0	0	0	0	0	0	0	0	-100
-100	0	0	0	0	0	0	0	0	0	-100
0	0	0	0	0	0	0	0	0	0	0
0	0	100	100	100	100	100	100	100	0	0

Iteration 1

0	33.3	100	100	100	100	100	100	100	33.3	0
-33	0	25	25	25	25	25	25	25	0	-33
-100	-25	0	0	0	0	0	0	0	-25	-100
-100	-25	0	0	0	0	0	0	0	-25	-100
-100	-25	0	0	0	0	0	0	0	-25	-100
-100	-25	0	0	0	0	0	0	0	-25	-100
-100	-25	0	0	0	0	0	0	0	-25	-100
-100	-25	0	0	0	0	0	0	0	-25	-100
-100	-25	0	0	0	0	0	0	0	-25	-100
-33	0	25	25	25	25	25	25	25	0	-33
0	33.3	100	100	100	100	100	100	100	33.3	[illegible]

Iteration 2

0	33.3	100	100	100	100	100	100	100	33.3	0
-33	0	31.3	37.5	37.5	37.5	37.5	37.5	31.3	0	-33
-100	-31	0	6.25	6.25	6.25	6.25	6.25	0	-31	-100
-100	-38	-6.3	0	0	0	0	0	-6.3	-38	-100
-100	-38	-6.3	0	0	0	0	0	-6.3	-38	-100
-100	-38	-6.3	0	0	0	0	0	-6.3	-38	-100
-100	-38	-6.3	0	0	0	0	0	-6.3	-38	-100
-100	-38	-6.3	0	0	0	0	0	-6.3	-38	-100
-100	-31	0	6.25	6.25	6.25	6.25	6.25	0	-31	-100
-33	0	31.3	37.5	37.5	37.5	37.5	37.5	31.3	0	-33
0	33.3	100	100	100	100	100	100	100	33.3	0

Figure 5.17: The first few iterations of relaxation method for solving Laplace's equation for a 10x10 grid.

With each iteration the potential "progresses" across the grid. After 20 iterations the grid points have converged sufficiently so that the next iteration causes less than a 0.1% change.

Iteration 20

0.00	33.33	100	100	100	100	100	100	100	33.33	0.00
-33.33	0.00	39.12	56.51	64.33	66.61	64.33	56.51	39.12	0.00	-33.33
-100	-39.12	0.00	22.70	34.40	38.00	34.40	22.70	0.00	-39.12	-100
-100	-56.51	-22.70	0.00	12.78	16.89	12.78	0.00	-22.70	-56.51	-100
-100	-64.33	-34.40	-12.78	0.00	4.21	0.00	-12.78	-34.40	-64.33	-100
-100	-66.61	-38.00	-16.89	-4.21	0.00	-4.21	-16.89	-38.00	-66.61	-100
-100	-64.33	-34.40	-12.78	0.00	4.21	0.00	-12.78	-34.40	-64.33	-100
-100	-56.51	-22.70	0.00	12.78	16.89	12.78	0.00	-22.70	-56.51	-100
-100	-39.12	0.00	22.70	34.40	38.00	34.40	22.70	0.00	-39.12	-100
-33.33	0.00	39.12	56.51	64.33	66.61	64.33	56.51	39.12	0.00	-33.33
0.00	33.33	100	100	100	100	100	100	100	33.33	0.00

Iteration 20 – Iteration 19

0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.01	0.03	0.06	0.07	0.06	0.03	0.01	0.00	0.00
0.00	-0.01	0.00	0.04	0.08	0.09	0.08	0.04	0.00	-0.01	0.00
0.00	-0.03	-0.04	0.00	0.05	0.07	0.05	0.00	-0.04	-0.03	0.00
0.00	-0.06	-0.08	-0.05	0.00	0.02	0.00	-0.05	-0.08	-0.06	0.00
0.00	-0.07	-0.09	-0.07	-0.02	0.00	-0.02	-0.07	-0.09	-0.07	0.00
0.00	-0.06	-0.08	-0.05	0.00	0.02	0.00	-0.05	-0.08	-0.06	0.00
0.00	-0.03	-0.04	0.00	0.05	0.07	0.05	0.00	-0.04	-0.03	0.00
0.00	-0.01	0.00	0.04	0.08	0.09	0.08	0.04	0.00	-0.01	0.00
0.00	0.00	0.01	0.03	0.06	0.07	0.06	0.03	0.01	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Figure 5.18: Simple relaxation method after 20 iterations and the change in each grid point it introduces.

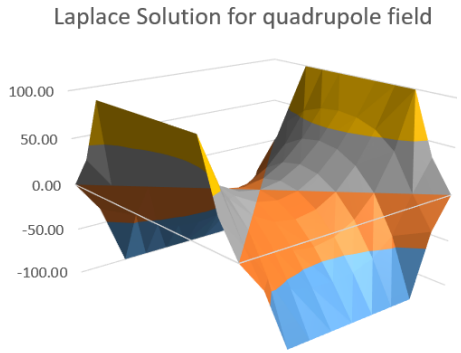


Figure 5.19: Surface plot of the solution after 20 iterations

While this convergence seems to be relatively fast, it is very slow for larger grids because for a grid of size $N \times N$ it takes $N/2$ iterations before the boundary values begin to diffuse into the centre of the mesh. This process can be made faster by updating the grid point to the value immediately after it is calculated rather than deferring to the next iteration. In the initial calculations, some points can also be skipped to get a faster adjustment of points far away from the electrodes. For instance, only every other point is adjusted in the first few "sweeps" through in a checkboard arrangement.

The quadrupolar potential can be used to trap particles in space by having the voltages on the electrodes oscillate in time. See https://www.youtube.com/watch?v=wlmJ5goHd3g to get a simulation of how this is achieved.

Successive Over-Relaxation (SOR)

The convergence can be speeded up considerably by over-estimating each iteration so it is "*over-relaxed*" by a factor σ. This calculation is

$$\overline{V}_{m,n} = (1-\sigma)\, V_{m,n} + \frac{1}{4}\sigma\left(V_{m,n-1} + V_{m,n+1} + V_{m+1,n} + V_{m-1,n}\right)$$

The optimal value of σ will depend on the problem but is typically between 1.2 and 1.4. This dramatically speeds up convergence especially for larger grids.

Appendix: Crout's Algorithm

Crout's algorithm solves a linear algebra problem of the form

$\underline{\underline{A}}\,\underline{x} = \underline{b}$

by expressing the 2D matrix $\underline{\underline{A}}$ as the product of two triangular matricies $\underline{\underline{L}}$ and $\underline{\underline{U}}$

$\underline{\underline{A}} = \underline{\underline{L}}\,\underline{\underline{U}}$

so that letting $\underline{\underline{L}}\,\underline{y} = \underline{b}$ then

$\underline{\underline{U}}\,\underline{x} = \underline{b}$

As the $\underline{\underline{L}}$ and $\underline{\underline{U}}$ are triangular, the elements of $\underline{y}$ and then $\underline{x}$ can be found

For the implicit (backwards) finite difference method, the solution to the heat equation can be found from

$$\underline{\underline{A}}\,\underline{T}_{n+1} = \underline{T}_n + \underline{b}_{n+1}$$

where $\underline{\underline{A}}$ is an $(M-1, M-1)$ matrix

$$\underline{\underline{A}} = \begin{pmatrix} (1+2\lambda) & -\lambda & 0 & 0 & 0 \\ -\lambda & (1+2\lambda) & -\lambda & 0 & 0 \\ 0 & -\lambda & (1+2\lambda) & -\lambda & 0 \\ & & \ddots & \ddots & \\ 0 & 0 & -\lambda & (1+2\lambda) & -\lambda \\ 0 & 0 & 0 & -\lambda & (1+2\lambda) \end{pmatrix}$$

This can be written as

$$\underline{\underline{L}} = \begin{pmatrix} c_1 & 0 & 0 & 0 & 0 \\ d & c_2 & 0 & 0 & 0 \\ 0 & d & c_2 & 0 & 0 \\ & & \ddots & \ddots & \\ 0 & 0 & 0 & d & c_{M-1} \end{pmatrix} \quad \underline{\underline{U}} = \begin{pmatrix} 1 & f_1 & 0 & 0 & 0 \\ 0 & 1 & f_2 & 0 & 0 \\ & & \ddots & \ddots & \\ 0 & 0 & 0 & 1 & f_{M-2} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

so that

$c_1 = 1 + 2\lambda, \;\; c_1 f_1 = -\lambda, \;\; d = -\lambda$

$df_1 + c_2 = 1 + 2\lambda$

$df_2 + c_3 = 1 + 2\lambda$

$\vdots$

$df_{M-2} + c_{M-2} = 1 + 2\lambda$

$c_{M-1} = 1 + 2\lambda$

Letting $\underline{\underline{L}}\,\underline{y} = \underline{T}_n$ where $\underline{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{M-1} \end{pmatrix}$ $\underline{b}_{n+1} = \lambda \begin{pmatrix} T_{0,n+1} \\ 0 \\ \vdots \\ T_{M,n+1} \end{pmatrix}$

so that

$c_1 y_1 = T_{1,n} + \lambda T_{0,n+1}$

$dy_1 + c_2 y_2 = T_{2,n}$

$\vdots$

$dy_{m-1} + c_m y_m = T_{m,n}$

$\vdots$

$dy_{M-2} + c_{M-1} y_{M-1} = T_{M-1,n} + \lambda T_{M,n+1}$

Then $\underline{\underline{U}}\;\underline{T}_{n+1} = \underline{y}$

and the temperature distribution solution at the next time step can be found from

$T_{M-1,n+1} = y_{M-1}$

$T_{M-2,n+1} + f_{M-2} T_{M-1,n+1} = y_{M-2}$

$\vdots$

$T_{m,n+1} + f_m T_{m+1,n+1} = y_m$

$\vdots$

$T_{1,n+1} + f_1 T_{2,n+1} = y_1$