F. Quevedo

DIFFERENTIAL EQUATIONS

Summary

0. Motivation

The study of Differential Equations (DE's) is arguably the area of Mathematics which has more applications to the real world. There are plenty of examples of their use in Physics, but also in Chemistry, Biology, Economics, etc. Usually differential equations are relations among quantities that change over time and/or space and therefore are relevant to the study of the evolving universe, the weather, stock market, etc. Their study has also lead to important discoveries in pure Mathematics.

N.B. This is very much a methods course. Emphasis is on understanding the main concepts and applying them to the solution of differential equations rather than on the formal aspects of the theory.

1. Basic Calculus

1.1 Differentiation and Integration

Derivatives

Define the derivative of $f(x)$ with respect to x by:

$$\frac{df}{dx} = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$$

Pictorially df/dx is the slope of the tangent to the graph of $f(x)$ at the point x. N.B. left and right-hand limits must be equal if f is differentiable (*e.g.* $f(x)=|x|$ is not differentiable at $x=0$).

Notation

$$\frac{df}{dx} \equiv \frac{d}{dx}f \equiv f'(x), \qquad \frac{d}{dx}\left(\frac{df}{dx}\right) \equiv \frac{d^2 f}{dx^2} \equiv f''(x), \qquad \text{etc.}$$

o and $\mathcal{O}$ symbols: Consider a function $H(x)$ (usually H is the difference between two other functions $H = F_1(x) - F_2(x)$ and measures the *order* of magnitude of the difference when x approaches a particular point $x \to x_0$):

$H(x) = o\,(g(x))$ as $x \to x_0$ if $\lim_{x\to x_0} \frac{H(x)}{g(x)} = 0$

$H(x) = \mathcal{O}\,(g(x))$ as $x \to x_0$ if $\lim_{x\to x_0} \frac{H(x)}{g(x)}$ is finite.

Derivatives Rules

Chain rule: If $f(x) = F\left[g(x)\right]$ then $\frac{df}{dx} = \frac{dF}{dg}\,\frac{dg}{dx}$

Product rule: If $f(x) = u(x)\,v(x)$ then $f' = \frac{df}{dx} = u'v + uv'$.

Quotient rule: If $f(x) = \frac{u(x)}{v(x)}$ then $f' = \frac{u'v - uv'}{v^2}$.

Leibnitz's rule: If $f(x) = u(x)\,v(x)$ then

$$\frac{d^n f}{dx^n} = f^{(n)}(x) = \sum_{r=0}^{n} {}^nC_r\, u^{(n-r)}\, v^{(r)}$$

where ${}^nC_r = \binom{n}{r} = \frac{n!}{(n-r)!\,r!}$.

L'Hopital's rule : Let $f(x)$ and $g(x)$ be differentiable functions at $x = x_0$ and $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = 0$ then

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f'(x)}{g'(x)}$$

provided $g'(x_0) \neq 0$. If both $f'(x_0) = g'(x_0) = 0$ then the rule can be applied again.

Taylor series

From the definitions of the derivative and the symbol $o(x)$ we can write: $f(x+h) = f(x) + hf'(x) + o(h)$. We can extend this expression to:

$$f(x+h) \;=\; f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \cdots + \frac{h^n}{n!}f^{(n)}(x) + E_n$$

where $E_n = o(h^n)$ as $h \to 0$. In fact $E_n = \mathcal{O}(h^{n+1})$ as $h \to 0$ provided $f^{(n+1)}$ exists. This is Taylor's theorem. It can be expressed in an alternative form by substituting $x \to x_0, h \to x - x_0$ in the previous expression:

$$f(x) \;=\; f(x_0) + (x-x_0)f'(x_0) + \cdots + \frac{(x-x_0)^n}{n!}f^{(n)}(x_0) + E_n = \sum_{r=0}^{n} \frac{f^{(r)}(x_0)}{r!}\,(x-x_0)^r + E_n$$

This is Taylor series of $f(x)$ about $x = x_0$. It gives a local approximation to the function $f(x)$ in the neighbourhood of $x = x_0$. If all derivatives exist, the series becomes: $f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}\,(x-x_0)^n$. Many known functions ($sinx, e^x$, etc.) can be expanded in this way. N.B. *warning*: not all functions have a well defined Taylor expansion *e.g.* $f(x) = e^{-1/x^2}$ does not have a Taylor expansion about $x = 0$ since the function and all its derivatives vanish at $x = 0$.

Integration

An (definite) integral is a sum of the form:

$$\int_a^b f(x)dx \;=\; \lim_{\Delta x\to 0} \sum_{n=0}^{N-1} f(x_n)\,\Delta x$$

where $\Delta x = \frac{b-a}{N}$ and $x_n = a + n\Delta x$. Pictorially it represents the area under the curve $f(x)$ between the points $x = a$ and $x = b$.

Fundamental Theorem of Calculus

Let $F(x) = \int_a^x f(t)dt$. Then $\frac{dF}{dx} = f(x)$.

To prove it use the definition of derivative dF/dx and $\int_x^{x+h} f(t)dt = f(x)h + \mathcal{O}(h^2)$ when h is small. Also, from the chain rule: $\frac{d}{dx}\int_a^{g(x)} f(t)dt = f\left(g(x)\right)\, g'(x)$.

Notation: $\int f(x)dx \equiv \int^x f(t)dt$. This is an *indefinite* integral. N.B. the non-appearance of a lower limit in the integral reflects the fact that an indefinite integral is defined up to an arbitrary constant.

Integration Techniques:

Integration is an art and needs practice. Examples of techniques are:

Integration by substitution : If integrand is a function of a function $\int f(u(x))u'dx$ it is helpful to recognise the form of the chain rule to change variables from x to u.

Integration by parts: Using the product rule $(uv)' = u'v + uv'$, it is often convenient to use $\int uv'dx = uv - \int vu'dx$ if $\int vu'dx$ is easier to do than $\int uv'dx$. See examples.

Derivative and integral as operators

Both derivation and integration can be seen as the action of a linear operator that takes one function $f(x)$ into a different function $g(x)$, that is $\frac{d}{dx} : f(x) \to g(x) = f'(x)$ and $\int^x : f(x) \to g(x) = \int^x f(t)dt$. Both operators are *linear*, that means that $L\left[\alpha f_1(x) + \beta f_2(x)\right] = \alpha L[f_1(x)] + \beta L[f_2(x)]$. Where L is either derivation $\frac{d}{dx}$ or integration $\int^x$ and α, β are constants.

1.2 Functions of Several Variables: Partial Differentiation

Consider a function of two variables $f(x, y)$. We are interested to find how this function changes when we move in the x or the y directions. The *partial derivative* of $f(x, y)$ w.r.t. x, keeping y constant is defined as:

$$\left(\frac{\partial f}{\partial x}\right)_y = \lim_{\delta x \to 0} \frac{f(x + \delta x, y) - f(x, y)}{\delta x}$$

Also $\left(\frac{\partial f}{\partial y}\right)_x = \lim_{\delta y \to 0} \frac{f(x, y+\delta y) - f(x,y)}{\delta y}$. This generalises to a function of many variables $f(x_1, x_2, \cdots, x_n)$. Computing partial derivatives of a given function is straightforward (just differentiate w.r.t. the corresponding variable treating the others as constants). Higher order derivatives come in several combinations: $\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial y^2}, \frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y \partial x}$. It can be shown in general that: $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$.

Notation We usually indicate which variables are being held constant but if no indication then assume everything is constant apart from the one variable w.r.t. which we are differentiating. *e.g.* for $f(x, y, z)$: $\left(\frac{\partial f}{\partial x}\right) = \left(\frac{\partial f}{\partial x}\right)_{y,z} \neq \left(\frac{\partial f}{\partial x}\right)_y$. Alternative notation $f_x = \frac{\partial f}{\partial x}, f_{xy} = \frac{\partial^2 f}{\partial x \partial y}$.

Chain rule

We define the *total differential* df of a function $f(x,y)$ as follows: Consider the variation of f from two neighbouring points: $\delta f = f(x+\delta x, y+\delta y) - f(x,y)$. In the limit when both $\delta x, \delta y \to 0$ this gives:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy.$$

If both x and y define a path parametrised by t: $x = x(t), y = y(t)$, then $f(x(t), y(t))$ is a function of t and

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

This is the chain rule for a function of several variables. Often the parametrisation may be just $x = t, y = y(t)$, that is simply $y = y(x)$. In this case the chain rule reduces to : $\frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx}$.

Change of Variables

If it is needed to change variables, *e.g.* from Cartesian to polar coordinates $x = x(r,\theta), y = y(r,\theta)$ and $f(x,y) = f\left(x(r,\theta), y(r,\theta)\right)$, then:

$$\left(\frac{\partial f}{\partial r}\right)_\theta = \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial r}\right)_\theta + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial r}\right)_\theta$$

$$\left(\frac{\partial f}{\partial \theta}\right)_r = \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial \theta}\right)_r + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial \theta}\right)_r$$

Implicit Differentiation

Consider $F(x,y,z) = \text{constant}$. This defines a surface in 3-D (*e.g.* a sphere $x^2 + y^2 + z^2 = 1$, a cone $x^2 + y^2 = z^2$, etc.). It *implicitly* defines $z = z(x,y)$ or $y = y(x,z)$ or $x = x(y,z)$. Even if it is not possible to explicitly solve for z as a function of x, y from $F =$ constant, we can use the chain rule to find $\left(\frac{\partial z}{\partial x}\right)_y$ as follows: $\left(\frac{\partial F}{\partial x}\right)_y = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial z}\left(\frac{\partial z}{\partial x}\right)_y = 0$. from which we can solve for $\left(\frac{\partial z}{\partial x}\right)_y$.

Differentiation of an Integral w.r.t. a parameter

Consider the family of functions $f(x,c)$ (parametrised by different values of c). Consider the integral $I(b,c) = \int_0^b f(x,c)dx$. From the fundamental theorem of calculus we have $\left(\frac{\partial I}{\partial b}\right)_c = f(b,c)$. Also, from direct differentiation $\left(\frac{\partial I}{\partial c}\right)_b = \int_0^b \left(\frac{\partial f}{\partial c}\right)_x dx$. Therefore, if $b = b(x), c = c(x)$, then

$$\frac{dI}{dx} = \frac{\partial I}{\partial b}\frac{db}{dx} + \frac{\partial I}{\partial c}\frac{dc}{dx} = f(b,c)\frac{db}{dx} + \frac{dc}{dx}\int_0^b \left(\frac{\partial f}{\partial c}\right)_y dy.$$

In particular, for $I(x) = \int_0^x f(x,y)dy$, $\frac{dI}{dx} = f(x,x) + \int_0^x \left(\frac{\partial f}{\partial x}\right)_y dy$.

F. Quevedo

DIFFERENTIAL EQUATIONS

Summary Chapter 2

2. First Order Ordinary Differential Equations (ODE's)

Definitions

A *Differential Equation* (DE) is an equation that contains derivatives of one or more dependent variables w.r.t. one or more independent variables.
An *Ordinary differential equation* (ODE) contains only ordinary derivatives.
A *Partial differential equation* (PDE) contains partial derivatives.
The *order* of a DE is the highest order derivative in the equation.
A DE is *linear* if the dependent variable (and its derivatives) appears only linearly.
A DE is a *DE with constant coefficients* if the independent variable does not appear explicitly.
A linear DE is *homogeneous* if $y = 0$ is a solution with y the dependent variable.

2.1 First Order Linear Ordinary Differential Equations

2.1.1 Homogeneous equations with constant coefficients

The most general first order, linear, homogeneous ODE with constant coefficients can be written as:

$$y' + \alpha y = 0$$

with $\alpha =$ constant. Recall:

$$\frac{d}{dx}\left(e^{\lambda x}\right) = \lambda \left(e^{\lambda x}\right)$$

this equation can be read as saying that $e^{\lambda x}$ is an *eigen-function* of the differential operator $L = \frac{d}{dx}$ with eigen-value λ. Similar to standard finite dimensional eigen-vectors that are transformed in a simple way by a linear operator. L keeps the functional form of the exponential function unchanged and only changes the magnitude.

Remarks

* The exponential function $y = e^{\lambda x}$ is a solution of the ODE above if $\lambda + \alpha = 0$ (*characteristic equation*).
* Since the ODE is linear and homogeneous, any multiple of a solution is a solution. Then $y = Ae^{-\alpha x}$ is a solution with A an arbitrary constant.
* $y = Ae^{-\alpha x}$ is the most general solution, as it can be found by direct integration of $\int \frac{dy}{y} = -\int \alpha dx$. So there is only one independent solution $e^{-\alpha x}$ and all other solutions can be obtained by multplication with A.
* The previous statement generalises to an *nth* order linear ODE that has n independent solutions.
* The constant A can be determined by applying *boundary conditions*, usually giving the value of y at $x = 0$. If $x = t =$ time this boundary condition is called an *initial condition*.

Series solution

A useful method to solve DE's is by series. If we assume the solution has a series expansion

$$y = \sum_{n=0}^{\infty} a_n x^n, \qquad y' = \sum_{n=1}^{\infty} n a_n x^n,$$

we can plug this *ansatz* in the ODE $y' + \alpha y = 0$ and determine the coefficients a_n by equating the coefficients of each power of x to zero. The ODE above is solved if the coefficients satisfy $a_{n+1} = -\frac{\alpha}{n+1} a_n$ which when applied iteratively leads to $a_n = \frac{(-\alpha)^n}{n!} a_0$ and therefore to $y = a_0 \sum_0^\infty \frac{(-\alpha)^n}{n!} x^n = a_0 e^{-\alpha x}$. Reproducing the most general solution we had already found. This method will be useful in more complicated ODE's.

Discrete Equation

We can approximate the ODE $y' + \alpha y = 0$ to a difference or discrete equation by just approximating $y' \sim \frac{y_{n+1} - y_n}{h}$ for a small h. The equation then becomes

$$y_{n+1} = (1 - \alpha h)\, y_n$$

which for negative α is like the equation for compound interest. To solve this equation, we can apply this *recurrence relation* repeatedly (as it was done for a_n in the series solution):

$$y_n = (1 - \alpha h) y_{n-1} = (1 - \alpha h)^2 y_{n-2} = \cdots = y_0 (1 - \alpha h)^n$$

Taking the limit $n \to \infty$, $(h \to x/n)$ gives $y(x) = y_0 \lim_{n\to\infty} (1 - \frac{\alpha x}{n})^n = y_0 e^{-\alpha x}$ reproducing the continuous result.

2.1.2 Inhomogeneous equations with constant coefficients

The inhomogeneous or forced equation with constant coefficients is $y' + \alpha y = f(x)$. Let us consider two illustrative cases.

(i) *Constant Forcing.* That is $f(x) = \beta =$ constant. A technique to solve it is to spot a particular steady (equilibrium) solution $y = y_p =$ constant. Since $y_p' = 0$ the solution is easily found to be $y_p = \beta/\alpha$. To find the most general solution write $y = y_p + y_h$ and substitute into the equation. Using the fact that y_p satisfies the inhomogeneous equation implies that y_h has to satisfy the homogeneous equation $y_h' + \alpha y_h = 0$, for which we already have the general solution $y_h = A e^{-\alpha x}$. Therefore the most general solution of the inhomogeneous equation is the sum of the particular solution y_p and y_h: $y = \beta/\alpha + A e^{-\alpha x}$. Notice that y depends on one arbitrary constant A.

(ii) *Eigenvalue forcing.* Example: In a radioactive rock, isotope A decays into isotope B at a rate proportional to the number a of remaining nuclei A and B decays to C at a rate proportional to the number b of remaining nuclei of B. The DE's are then:

$$\frac{da}{dt} = -k_a a \qquad\qquad \frac{db}{dt} = k_a a - k_b b$$

the solution of the first equation is $a = a_0 e^{-k_a t}$ which we can plug in the second equation to leave $\frac{db}{dt} + k_b b = k_a a_0 e^{-k_a t}$. Note that the forcing is an eigenfunction of the differential operator of the LHS. So try a particular integral $b_p = ce^{-k_a t}$. Plugging into the equation we find $c = \frac{k_a a_0}{k_b - k_a}$ for $k_b \neq k_a$. To find the general solution write $b = b_p + b_h$ where b_h satisfies the homogeneous equation. This leads to the solution $b = \frac{k_a a_0}{k_b - k_a} e^{-k_a t} + De^{-k_b T}$. Assuming the initial condition $b = 0$ at $t = 0$ fixes $D = -1$. Then we can easily find the ratio $b(t)/a(t)$ which allows a rock to be dated from determining the relative proportion of certain istopes.

2.1.3 The General Linear First Order ODE

The general linear first order ODE (inhomogeneous, non-constant coefficients) can be written as:

$$y' + p(x)y = f(x)$$

To solve it multiply both sides by a function (to be determined) $\mu(x)$ and determine $\mu(x)$ such that the LHS is given by the product rule $(\mu y)' = \mu y' + \mu' y$. this implies that μ satisfies $\mu' = \mu p$ which implies we can choose $\mu = e^{\int p dx}$. Plugging this into the equation leads to the general solution:

$$y(x) = \frac{1}{\mu(x)} \left[\int^x \mu(t) f(t) dt + c \right].$$

with the *integrating factor* (IF) $\mu(x) = e^{\int p dx}$. *N.B. It is better to remember the method than the solution.*

2.2 Non-linear, First Order ODE's

In general the nonlinear first order ODE is of the form:

$$Q(x,y)y' + P(x,y) = 0$$

This equation does not always have an analytic solution, so we will concentrate on classes of equations that can be solved analytically.

2.2.1 Exact Equations

The equation $Q(x,y)y' + P(x,y) = 0$ is an exact equation iff $Q(x,y)dy + P(x,y)dx$ is an exact differential $df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$ of a function $f(x,y)$. comparing bith expressions this means

$$\frac{\partial f}{\partial x} = P, \qquad \frac{\partial f}{\partial y} = Q.$$

In which case the equation reduces to $df = 0$ and the solution is $f =$ constant. From $\frac{\partial f}{\partial x} = P, \frac{\partial f}{\partial y} = Q$ we can see that $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial P}{\partial y}$ and $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial Q}{\partial x}$. This implies $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$. *N.B. reverse implication:* If $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ throughout a simply connected domain $\mathcal{D}$ then $Pdx + Qdy$ is an exact differential of a single valued function in $\mathcal{D}$. If the equation is exact then the solution can be found by integrating $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$.

2.2.2 Separable Equations

The equation $Q(x,y)y' + P(x,y) = 0$ is separable if it can be manipulated into the form $q(y)dy = p(x)dx$ in which case the solution can be found by integration.

2.2.3 Graphical Methods

We may find information about the solutions of an ODE without actually solving it. Graphical methods are very useful in this sense. Consider the equation

$$\frac{dy}{dt} \equiv \dot{y} = f(y,t)$$

* *Flow vectors* Can evaluate $f(y,t)$ at particular values of y and t and draw a small arrow (flow vector) at that point with the slope determined by the value of $f(y,t) = \dot{y}$ at that point. Then the curves joining the different arrows are solutions of the ODE.
* *Isoclines.* These are curves for which $f(y,t) =$ constant. Help to draw the flow vectors and then construct the solutions. *N.B. isoclines are usually not solutions of the ODE.*
* *Fixed points.* These are points where $\dot{y} = 0$ for all t that means they are solutions of $f(y,t) = 0$ for all t.
* *Stability* A fixed point is *stable* if solutions converge toward it and *unstable* if the solutions diverge away from it.

Perturbation Analysis To determine stability suppose that $y = a$ is a fixed point of $\dot{y} = f(y,t)$, that is $f(a,t) = 0$. To determine if $y = a$ is stable or not, write $y = a + \epsilon(t)$ ans substitute into the equation. Then $\dot{\epsilon} = f(a+\epsilon,t) = f(a,t) + \epsilon\frac{\partial f}{\partial y}(a,t) + \mathcal{O}(\epsilon^2)$. Then $\dot{\epsilon} \sim \left[\frac{\partial f}{\partial y}\right]\epsilon$ which is a linear equation for ϵ. It can be integrated. If $\epsilon \to 0$ when $t \to \infty$ then the fixed point $y = a$ is stable, if ϵ increases with t then $y = a$ is unstable.

Autonomous Systems If $\dot{y} = f(y)$ only (independent of t), then near a fixed point $y = a$ where $f(a) = 0$, we write $y = a + \epsilon(t)$ which leads to $\dot{\epsilon} = \left[\frac{df}{dy}(a)\right]\epsilon = k\epsilon$ with $k = f'(a)$ constant. The solution is then $\epsilon = \epsilon_0 e^{kt}$. So the fixed point is stable if $f'(a) < 0$ and unstable if $f'(a) > 0$. If $f'(a) = 0$ the leading $\mathcal{O}(\epsilon^2)$ terms must be considered.

Logistic Equation. A population dynamics model is described by the following (logistic) equation:

$$\dot{y} = (\alpha - \beta)y - \gamma y^2 = ry\left(1 - \frac{y}{Y}\right) = f(y)$$

Where α, β are the birth and death rates and γ measures the death rate due to fighting. Here $r = \alpha - \beta$ and $Y = r/\gamma$. Notice that there is one fixed point $y = Y$ which is stable. When population is small $\dot{y} \sim ry$ and population grows exponentially. Eventually a stable equilibrium point is reached at $y = Y$.

Discrete Logistic Equation. Approximating $y' \sim (y_{n+1} - y_n)/h$ and redefining y_n and the constants r, Y, the differential logistic equation becomes the discrete logistic equation or *logistic map* :

$$x_{n+1} = \lambda x_n \left(1 - x_n\right)$$

This is a particular case of a general class of discrete equations $x_{n+1} = f(x_n)$. The fixed points correspond to $x_{n+1} = x_n$ that is they are the solutions of $f(X) = X$. Stability is determined by writing $x_n = X + \epsilon_n$. Then $X + \epsilon_{n+1} = f(X) + \epsilon_n f'(X) + \mathcal{O}(\epsilon_n^2)$. Then $\epsilon_{n+1} \sim f'(X)\epsilon_n$. Fixed point X is stable if $|\frac{\epsilon_{n+1}}{\epsilon_n}| < 1$ which implies $|f'(X)| < 1$. For the logistic map there are two fixed points $X = 0$ and $X = 1 - 1/\lambda$. Since $f'(X) = \lambda(1 - 2X)$ we can see that $X = 0$ is stable for $\lambda < 1$ and $X = 1 - 1/\lambda$ is stable for $1 < \lambda < 3$. For $4 > \lambda > 3$ there is an extremely rich pattern in which the solutions oscillates between different points (solutions of $x_{n+2} = x_n$, etc.) then leading to *chaos*.

DIFFERENTIAL EQUATIONS

Summary Chapter 3

3. Higher Order Linear ODE's

The general *nth* order linear ODE can be written as:

$$Ly \equiv \sum_{k=0}^{n} a_k(x) y^{(k)} = f(x).$$

N.B. We will mostly concentrate on the case $n = 2$ for simplicity and the amount of applications since most equations of mathematical physics are of second order. The operator $L = \sum_{k=0}^{n} a_k(x) \frac{d^k}{dx^k}$ is linear. That is $L(\alpha y_1 + \beta y_2) = \alpha L y_1 + \beta L y_2$. This immediately implies that for $f = 0$, if y_1, y_2 are two solutions of the above ODE then $y = Ay_1 + By_2$ is also a solution. For $f \neq 0$ it implies that if y_p is a solution then $y = y_p + y_h$ is also a solution, where y_h is any solution of the homogeneous equation ($f = 0$). This suggests a method to find the general solution: first find the general solution of the homogeneous equation (the *complementary function* y_h) and then find one particular solution y_p of the inhomogeneous equation. The general solution will be the sum of both $y = y_h + y_p$.

3.1 Constant Coefficients

we will start with the simplest case of a second order DE with constant coefficients

$$ay'' + by' + cy = f$$

Recall that $e^{\lambda x}$ is an eigenfunction of the operator $\frac{d}{dx}$ and hence it is an eigenfunction of $\frac{d^2}{dx^2}$, etc. Then the complementary function can be written as $y_h = e^{\lambda x}$. Plugging this into the DE gives rise to the *characteristic equation* for λ: $a\lambda^2 + b\lambda + c = 0$. There are two solutions of the characteristic equation λ_1, λ_2 giving two complementary functions $y_1 = e^{\lambda_1 x}, y_2 = e^{\lambda_2 x}$. If λ_1, λ_2 are distinct then y_1, y_2 are *linearly independent* and *complete* (they form a basis in the space of solutions of the homogeneous equation). The general complementary function is then $y_h = Ae^{\lambda_1 x} + Be^{\lambda_2 x}$. With A, B two arbitrary constants. To determine A, B need to provide two *boundary conditions* (or *initial conditions* if the independent variable is time), say $y_h(x_0)$ and $y_h'(x_0)$.

Examples

Nature of solutions depend on the roots of the characteristic equation.

1. *Two different real roots.* e.g. $y'' - 5y' + 6y = 0$. Roots $\lambda = 2, 3$. Solution $y_h = Ae^{2x} + Be^{3x}$.
2. *Two imaginary roots.* e.g. $\ddot{y} + \omega_0^2 y = 0$. Roots $\lambda = \pm i\omega_0$. Solution $y = A\cos(\omega_0 t) + B\sin(\omega_0 t)$. This is the simple harmonic oscillator with natural frequency ω_0.
3. *Degenerate roots.* e.g. $y'' - 4y' = 4y = 0$. Roots $\lambda = 2, 2$. Then can only write one degenerate linearly independent solution $y_h = Ae^{2x}$. Need to find the second solution.

Second complementary function

(i) *De-tuning.* Perturb the previous equation slightly, e.g. $y'' - 4y' + (4 - \epsilon^2)y = 0$ for $\epsilon << 1$. Roots $\lambda = 2 \pm \epsilon$. Taking appropriately the limit $\epsilon \to 0$ gives $y_h = e^{2x}[\alpha + \beta x]$.

(ii) *A product solution.* Take $y_2(x) = v(x)y_1(x)$. For the equation above we know $y_1 = e^{2x}$ plugging y_2 in this equation gives a simple equation for $v(x)$: $v'' = 0$, implying $v = Ax + B$ and then $y_h = Ce^{2x} + Dxe^{2x}$.

This is a demonstration of a more general rule that if $y_1(x)$ is a degenerate complementary function then $y_2(x) = xy_1(x)$ is an independent complementary function, for linear ODE's with constant coefficients.

Particular Solutions

To find particular solutions y_p of the inhomogeneous equation with forcing $f \neq 0$, a simple method (*undetermined coefficients*) is to guess the form of y_p to be of the same functional form as f (e.g. $f = e^{mx}$ then take $y_p = Ae^{mx}$ plug into the DE and determine the value of A.). Remember equation is linear so we can superpose solutions and consider each forcing separately (e.g. for $f = 2x + e^{4x}$ take $y_p = ax + b + ce^{4x}$.).

Resonance

Consider $\ddot{y} + \omega_0^2 y = \sin(\omega_0 t)$. We know that $y_h = A\sin(\omega_0 t) + B\cos(\omega_0 t)$ then the standard guess $y_p = C\sin(\omega_0 t) + D\cos(\omega_0 t)$ satisfies the homogeneous equation and cannot work. A *detuning* can also be done here by writing $f = \sin\omega t$ with $\omega = \omega_0 - \epsilon$ and again $\epsilon << 1$. A solution of the type $y_p = C\sin\omega t$ can work now since $\omega \neq \omega_0$. For $\epsilon << 1$ this gives $y_p = -\frac{2}{(\omega+\omega_0)\epsilon}\left[\cos(\omega_0 - \frac{\epsilon}{2})t\ \sin(\frac{\epsilon t}{2})\right]$ For $\epsilon << 1$ this gives rise to *beating* which is an oscillation of frequency close to ω_0 with an amplitude modulated by an *envelope* given by the $\sin\epsilon t/2$ of large frequency $\mathcal{O}(1/\epsilon)$. In the limit $\epsilon \to 0$, we get $y_p = -(t/2\omega_0)\cos(\omega_0 t)$. This illustrates the phenomenon of *resonance* since the amplitude increases linearly with t. In general if the forcing is a constant linear combination of complementary functions then the particular solution is proportional to t times the standard guess.

3.2 Difference Equations

Consider a higher order difference equation e.g.

$$ay_{n+2} + by_{n+1} + cy_n = f_n$$

We can solve it in a similar way to differential equations by exploiting linearity and eigenfunctions. The difference operator $D[y_n] = y_{n+1}$ has eigen-function $y_n = k^n$ since $D[k^n] = k^{n+1} = k \cdot k^n$. To solve the difference equation above first find complementary functions satisfying the homogeneous equation ($f_n = 0$). Try $y_n = k^n$ which leads to the characteristic equation $ak^2 + bk + c = 0$ with roots k_1, k_2. The complementary function is $y_n^{(h)} = Ak_1^n + Bk_2^n$ for $k_1 \neq k_2$. If $k_1 = k_2$ the second solution is of the form nk_1^n. For particular solutions again look for $y_n^{(p)}$ similar to f_n (for $f_n = k^n$, $k \neq k_1, k_2$ try $y_n^{(p)} = Ck^n$, if $k = k_1$ try $y_n^{(p)} = Cnk_1^n$, etc.). See example for Fibonacci sequence.

3.3 General nth order Linear DE

Restrict mostly to second order, e.g.:

$$y'' + p(x)y' + q(x)y = f(x)$$

3.3.1 Phase Space and the Wronskian

A differential equation of nth order determines the nth derivative $y^{(n)}(x)$ and hence all higher derivatives in terms of $Y(x), y'(x) \cdots y^{(n-1)}(x)$. Can think in terms of a solution vector: $\mathbf{Y}(x) = (y, y', \cdots, y^{(n-1)})$ defining a point in a n-dimensional *phase space*. Two solutions of the second order ODE y_1, y_2 are independent if their corresponding vectors $\mathbf{Y_1}$ and $\mathbf{Y_2}$ are linearly independent. That means that the *Wronskian*:

$$W(x) = \det \begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix} \neq 0$$

Warning: the opposite is not always true, *i.e.* $W = 0$ does not imply linear dependence.

Abel's Theorem For the homogeneous equation $y'' + p(x)y' + q(x)y = 0$, if p and q are continuous then either $W \equiv 0$ or $W \neq 0$ for any value of x. Actually $W = W_0 \exp(-\int p dx)$.

To prove this consider W' and use the fact that y_1, y_2 satisfy the ODE which leads to $W' = -pW$ and solve for W.

Find a second solution. If we know one solution of the second order ODE and knowing the Wronskian from $W = W_0 \exp(-\int p dx)$ we can find the second solution y_2. For this just notice that $W = y_1^2 \frac{d(y_2/y_1)}{dx}$ and integrate to give $y_2 = y_1 \int (W/y_1^2)\, dx$. See examples, especially the *Cauchy-Euler equation* (homogeneous in x): $ax^2y'' + bxy' + cy = 0$ with solutions $y = x^k$.

3.3.2 Variation of Parameters

This is a second method to find particular solutions of ODE's (the first we mentioned before was the undetermined coefficients guess for y_p). Let $y_1(x), y_2(x)$ be linearly independent complementary functions of the ODE $y'' + p(x)y' + q(x)y = f(x)$. Since the corresponding solution vectors $\mathbf{Y_1}, \mathbf{Y_2}$ form a basis of the solution space, we can write a particular solution as:

$$\mathbf{Y_p} = u(x)\mathbf{Y_1}(x) + v(x)\mathbf{Y_2}(x)$$

Plugging this into the ODE leads to $y_1'u' + y_2'v' = f(x)$. Also from knowing that the second entry of $\mathbf{Y_p}$ is the derivative of the first entry we get $y_1u' + y_2v' = 0$. Therefore

$$\begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix} \begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} 0 \\ f \end{pmatrix}$$

From which we can solve for u' and v' as $u' = -\frac{y_2}{W} f$ and $v' = \frac{y_1}{W} f$. This method is more systematic than the undetermined coefficients and applies to the case of non-constant coefficients. However it is often difficult to integrate u', v' to find u, v. See examples.

3.3.3 Important Physical Systems

Let us consider three classes of physical systems that can be described by second order ODE's.

A. **Transients and damping**

In many physical systems there is a restoring force and a damping (e.g. car suspension system). Newton's second law applied to a system like this is $M\ddot{x} = F - kx - l\dot{x}$ where M is the mass of the object, F is the applied force $-kx$ is the restoring force as in a spring (Hooke's law) and $-l\dot{x}$ is a damping term (the shock absorber in the car). Another system that leads to the same equation is an RLC electric circuit in which the dependent variable is Q instead of x. By a suitable change of variables $t = \tau\sqrt{M/k}$ this equation becomes $\ddot{x} + 2\kappa\dot{x} + x = f(\tau)$ where $\kappa = l/(2\sqrt{kM})$. Then there is a single parameter κ determining the properties of the system. Consider first the free natural response $f = 0$. The characteristic equation gives $\lambda_{1,2} = -\kappa \pm \sqrt{\kappa^2 - 1}$. So we have three regimes according to κ being smaller, greater or equal to 1.

(i) For $\kappa < 1$ the solution of the homogeneous equation is $x = e^{-\kappa\tau}(A\sin\sqrt{1-\kappa^2}\tau + B\cos\sqrt{1-\kappa^2}\tau)$. This is a damped oscillator with decay time (time that takes the amplitude to decrease by $1/e$ of the original value) $\mathcal{O}(1/\kappa)$ and period $T = 2\pi/\sqrt{1-\kappa^2}$. Notice that the damping ($\kappa \neq 0$) increases the period. As $\kappa \to 1$ the oscillation period $T \to \infty$.

(ii) For $\kappa = 1$ (critically damped) the solution is $x_h = (A + B\tau)e^{-\tau}$ with the amplitude x increasing at early times before reaching a maximum value at a critical value $\tau = \tau_c$ and start decaying exponentially with decay time $\mathcal{O}(1)$.

(iii) For $\kappa > 1$ (over-damped) the solution is $x = Ae^{-(\kappa+\sqrt{\kappa^2-1})\tau} + Be^{-(\kappa-\sqrt{\kappa^2-1})\tau}$. Possible to get a large initial increase in the amplitude followed by a slow decay (decay time $\mathcal{O}(1/(\kappa - \sqrt{\kappa^2-1}))$).

For the forced system ($f \neq 0$) the complementary function determines the short time *transient* response while the particular solution determines the long-time, *asymptotic* response. e.g. for $f(\tau) = \sin\tau$ the total solution $x \to -\cos\tau/2\kappa$ as $\tau \to \infty$.

B. **Impulses and Point Forces**

Consider a ball bouncing on the ground. For a very short amount of time $t \in (T-\epsilon, T+\epsilon)$ for a time T and $\epsilon << 1$, there is a force $F(t)$ exerted by the ground on the ball. Since this force acts for a time of $\mathcal{O}(\epsilon)$ much less than the typical time of the system (between say dropping and getting back the ball), it is convenient mathematically to imagine the force acting instantaneously at $t = T$, *i.e.* consider $\epsilon \to 0$.

Newtons second law applied to this system is $m\ddot{x} = F(t) - mg$ where m is the mass of the ball and g the gravitational acceleration. Integrating this equation between $T - \epsilon$ and $T + \epsilon$ gives $\left[m\frac{dx}{dt}\right]_{T-\epsilon}^{T+\epsilon} = I + 2mg\epsilon \to \quad I$, as $\epsilon \to 0$, where the *impulse* $I = \int_{T-\epsilon}^{T+\epsilon} F(t)dt$ is the area under the force curve and is the only property of $F(t)$ that influences the macroscopic behaviour of the system. This leads us to consider mathematically a family of functions $D(t;\epsilon)$ such that $\lim_{\epsilon\to 0} D(t;\epsilon) = 0$ for all $t \neq 0$ and $\lim_{\epsilon\to 0}\int_{-\infty}^{\infty} D(t;\epsilon)dt = 1$. a typical example is $D(t;\epsilon) = \frac{1}{\epsilon\sqrt{\pi}}e^{-t^2/\epsilon^2}$. Note that as $\epsilon \to 0$, $D(0;\epsilon) \to \infty$ so $\lim_{\epsilon\to 0} D(t;\epsilon)$ is not properly defined. Nevertheless we define the *Dirac delta function* by:

$$\delta(x) = \lim_{\epsilon\to 0} D(x;\epsilon)$$

with the understanding that we can only use its integral properties. e.g.

$$\int_{-\infty}^{\infty} f(x)\delta(x)dx = \lim_{\epsilon\to 0}\int_{-\infty}^{\infty} f(x)D(x;\epsilon)dx = f(0)\lim_{\epsilon\to 0}\int_{-\infty}^{\infty} D(x;\epsilon)dx = f(0)$$

provided $f(x)$ is continuous at $x = 0$. This gives a convenient way of representing and making calculations involving impulsive or point forces. In the bouncing ball example Newton's equation can be written as $m\ddot{x} = -mg + I\delta(t - T)$. See examples.

C. **Switching on and/or off**

Define the *Heaviside step function* by $H(x) = \int_{-\infty}^{x} \delta(u)du$ which gives $H = 0$ for $x < 0$ and $H = 1$ for $x > 0$. Then $\frac{dH}{dx} = \delta(x)$. We can independently verify this by:

$$\int_{-\infty}^{\infty} H' f dx = \left[Hf\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} H f' dx = f(\infty) - \int_{0}^{\infty} f' dx = f(\infty) - [f(\infty) - f(0)] = f(0)$$

Illustrating that $H'(x)$ behaves like $\delta(x)$. But remember these functions and relationships can only be used within integrals. The function $H(x)$ is useful for switching problems such as the electric circuit example.

3.4 Series Solutions

Consider the differential equation

$$p(x)y'' + q(x)y' + r(x)y = 0$$

The point $x = x_0$ is an *ordinary point* of this DE if q/p and r/p have a Taylor expansion about $x = x_0$ (in particular if these ratios are not singular at $x = x_0$). Otherwise $x = x_0$ is a *singular point.* If $x = x_0$ is a singular point but $q(x - x_0)/p$ and $r(x - x_0)^2/p$ have Taylor expansions about $x = x_0$ then x_0 is a *regular singular point* (rsp).

If x_0 is an ordinary point then the DE above has two linearly independent solutions of the form:

$$y = \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

which converges in some neighbourhood of x_0. If x_0 is a regular singular point then the DE has at least one solution of the form (*Froebenius series*):

$$y = \sum_{n=0}^{\infty} a_n (x - x_0)^{n+\sigma}, \quad a_0 \neq 0$$

for some σ to be determined. This is known as *Fuch's theorem.* To find an explicit solution means to find all coefficients a_n and σ. This is done by plugging the series in the DE.

Examples

We consider three examples that illustrate the main distinctive cases.

1. Legendre's equation is $(1-x^2)y'' - 2xy' + l(l+1)y = 0$. it is easy to see that $x = 0$ is an ordinary point an that $x = \pm 1$ are regular singular points. We can look for solutions about the $x = 0$ point for which we expect two independent solutions. Plugging the series in the DE we find: $\sum_{n=0}^{\infty} n(n-1)a_n x^{n-2} - \sum_{n=0}^{\infty}(n-1)(n+2)a_n x^n = 0$. By changing $n \to n-2$ in the second sum we get the same powers of x in both sums and then can set each coefficient of x^n to zero. This gives the *recurrence relation*: $n(n-1)a_n = (n-3)na_{n-2}$. For $n = 0$ this gives $0 \cdot a_0 = 0$ (since $a_{-1} = a_{-2} = 0$) implying a_0 is arbitrary. Similarly for $n = 1$ we have $0 \cdot a_1 = 0$ implying a_1 is also arbitrary. For $n > 1$ we have $a_n = \frac{(n-3)}{n-1} a_{n-2}$ and so the solution is $y = a_0(1 - x^2 - \cdots) + a_1 x$. N.B. a_0 and a_1 provide the two independent constants needed.

2. Expansion about a regular singular point. consider the DE: $4xy' + 2(1-x^2)y' - xy = 0$. The point $x = 0$ is a rsp. Then the solution should be of the form $y = \sum a_n x^{n+\sigma}$. Plugging this into the equation gives $(n+\sigma)(n+\sigma-2)a_n - (2n+2\sigma-3)a_{n-2} = 0$ for $n = 0$ we have the *indicial equation*: $\sigma(2\sigma - 1) = 0$ with two roots $\sigma = 0, 1/2$. For each value of σ we find that now $a_1 = 0$ and $a_{2k} = \frac{4k-3}{4k(4k-1)} a_{2k-2}$ for $\sigma = 0$ and $a_{2k} = \frac{4k-2}{4k(4k+1)} a_{2k-2}$ for $\sigma = 1/2$. The solution is $y = a_0(1 + x^2/12 + \cdots) + a_0'(1 + x^2/10 + \cdots)$ where now the two arbitrary a_0's provide the two independent constants needed.
3. Expansion about a rsp with roots of indicial equation differing by an integer. Consider $x^2 y'' - xy = 0$. In this case $x = 0$ is a rsp but the indicial equation gives $\sigma = 0, 1$. For $\sigma = 1$ we find a_0 arbitrary and $a_n = a_{n-1}/(n(n+1))$. For $\sigma = 0$ we find $0 \cdot a_0 = 0$ but $0 \cdot a_1 = a_0$. So either get a contradiction or have $a_0 = 0$, a_1 arbitrary, with $a_n = a_{n-1}/(n(n-1))$ reproducing the same solution as the $\sigma = 1$ case. To find a second solution, use the Wronskian or use the ansatz $y_2 = y_1 \ln x + \sum_n b_n x^{n+\sigma}$. See question 4 of example sheet 4 for a justification of this ansatz.

3.5 Systems of Linear Equations

Consider two dependent variables $y_1(t), y_2(t)$ with the coupled DE's: $\dot{y}_1 = ay_1 + by_2 + f_1(t)$, and $\dot{y}_2 = cy_1 + dy_2 + f_2(t)$. This system of equations can be written in a matrix form

$$\dot{\mathbf{Y}} = \mathbf{MY} + \mathbf{F}, \qquad \text{where} \quad \mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad \mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$$

differentiating the first of these equations and using both equations to eliminate $y_2, \dot{y}_2$ we can see that this system is equivalent to one second order DE of the type $\ddot{y}_1 + A\dot{y}_1 + By_1 = F$. Conversely, a second order ODE of this type $\ddot{y} + A\dot{y} + By = F$ can be reduced to a system of two first order ODE's by setting: $y_1 \equiv y, y_2 \equiv \dot{y}$ then $\dot{y}_1 = y_2$ and $\dot{y}_2 = F - Ay_2 - By_1$ is a system of two coupled first order ODE's. In general any *nth* order ODE can be written as a system of n first order ODE's.

Solving a System of n First Order ODE's

Consider the matrix equation $\dot{\mathbf{Y}} - \mathbf{MY} = \mathbf{F}$. The complementary function can be written as $\mathbf{Y_h} = \mathbf{v}e^{\lambda t}$ where $\mathbf{v}$ is a constant vector. Plugging this into the equation gives $\mathbf{Mv} = \lambda\mathbf{v}$. This means that $\mathbf{v}$ is an eigen-vector of $\mathbf{M}$ with eigen-value λ. λ is then determined by the characteristic equation $\det(\mathbf{M} - \lambda\mathbf{I}) = 0$. For a system of two equations this immediately gives the solution of the homogeneous equation as (for $\lambda_1 \neq \lambda_2$):

$$\mathbf{Y_h} = A\,\mathbf{v_1}e^{\lambda_1 t} + B\,\mathbf{v_2}e^{\lambda_2 t}$$

(N.B. For $\lambda_1 = \lambda_2 = \lambda$ and one single independent eigen-vector $\mathbf{v}$ the second complementary function can be found of the form $\mathbf{v}te^{\lambda t} + \mathbf{u}e^{\lambda t}$ where (the generalised eigen-vector) $\mathbf{u}$ satisfies $(\mathbf{M} - \lambda\mathbf{I})\mathbf{u} = \mathbf{v}$). For a particular integral of the inhomogeneous equation try as usual $\mathbf{Y_p}$ of the same form as $\mathbf{F}(t)$, unless $\mathbf{F}(t)$ is similar to the complementary function. In that case we have to proceed in a way similar to the resonance examples before but taking into account the direction of $\mathbf{F}$. For instance if $\mathbf{F} = e^{\lambda_1 t}(a\mathbf{v_1} + b\mathbf{v_2})$ the ansatz should be $\mathbf{Y_p} = e^{\lambda_1 t}(Ct\mathbf{v_1} + D\mathbf{v_2})$. See examples.

DIFFERENTIAL EQUATIONS

Summary Chapter 4

4. Partial Differential Equations (PDE's)

This chapter retakes the study of multivariable functions, first exploring the existence of stationary points and then giving the elements of partial differential equations.

4.1 Directional Derivatives, the Gradient and Stationary Points

Let us consider a function of two variables $f(x, y)$ and study its change under the displacement $\mathbf{ds} = (dx, dy)$. We know from section 1.2 that:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = (dx, dy) \cdot (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}) \equiv \mathbf{ds} \cdot \nabla f$$

Where $\nabla f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$ is the *gradient* of f, ($\nabla f \equiv grad\, f$) written in cartesian coordinates. If we write $\mathbf{ds} = ds\, \hat{\mathbf{s}}$ where $|\hat{s}| = 1$ then $\frac{df}{ds} = \hat{s} \cdot \nabla f$. This is the *directional derivative* of f in the direction of $\hat{s}$.

N.B. The operator $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$ takes a (scalar) function f into a vector ∇f. It can also act on vectors $\mathbf{v} = (v_x, v_y)$ as $\nabla \cdot \mathbf{v} \equiv div\, \mathbf{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}$ which is a scalar function, constructed in way similar to the standard scalar product of vectors, known as the divergence of $\mathbf{v}$. It may also act as $\nabla \times \mathbf{v} \equiv curl\, \mathbf{v}$ which (in three dimensions) is a vector with components $(\nabla \times \mathbf{v})_x = \frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z}$, constructed in a way similar to the standard cross product.

In general, the gradient can be defined by the expression $\frac{df}{ds} = \hat{s} \cdot \nabla f = |\nabla f| \cos\theta$ where θ is the angle between ∇f and $\hat{\mathbf{s}}$. From this simple expression we can extract three important properties of the gradient:

(i) ∇f has magnitude equal to the maximum rate of change of $f(x, y)$ with distance in the (x, y) plane ($\cos\theta = 1$).

(ii) ∇f determines the direction in which f increases most rapidly ($\theta = 0$).

(iii) If $\mathbf{ds}$ is a displacement along a contour of f then $\frac{df}{ds} = 0$. This implies $\hat{\mathbf{s}} \cdot \nabla f = 0$ and then ∇f is the direction orthogonal to the contour.

Stationary points

There is always one direction in which $\frac{df}{ds} = 0$ namely parallel to the contours of f. But local maxima or minima have $\frac{df}{ds} = 0$ for *all* directions. This implies $\hat{\mathbf{s}} \cdot \nabla f = 0$ for all $\hat{\mathbf{s}}$ and then $\nabla f = 0$. In Cartesian coordinates this is $\frac{\partial f}{\partial x} = 0$, $\frac{\partial f}{\partial y} = 0$. Points for which $\nabla f = 0$ are called *stationary points*. They can be *maxima, minima* or *saddle points* (minima in some directions and maxima in others). To determine the nature of a stationary point we need to have an explicit expression for the Taylor series expansion for a multivariable function.

Taylor series for multivariable functions

Let us consider the displacement $\mathbf{x} = \mathbf{x_0} + \delta\mathbf{s}$. In two dimensions: $(x, y) = (x_0, y_0) + (\delta x, \delta y)$. The Taylor series expansion of $f(x, y)$ about the point (x_0, y_0) can be written as:

$$f(x, y) = a_0 + a_1\,\delta x + b_1\,\delta y + a_2\,\delta x^2 + b_2\,\delta y^2 + c_2 \delta x \delta y + \cdots$$

Similar to what we did in section 1.2, we can determine the parameters $a_0, a_1, b_1, \cdots$ by evaluating f and its partial derivatives at the point (x_0, y_0). This implies:

$$f(\mathbf{x}) = f(\mathbf{x_0}) + \delta\mathbf{s} \cdot \nabla f(\mathbf{x_0}) + \frac{1}{2}\delta\mathbf{s}^T \cdot \mathbf{H}(\mathbf{x_0}) \cdot \delta\mathbf{s} + \cdots$$

Where, in two dimensions:

$$\mathbf{H} = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

is the *Hessian matrix*, with clear generalisation for higher dimensions. N.B. $\det \mathbf{H} \equiv |\mathbf{H}|$ is called *the Hessian.* Also $\mathrm{Tr}\mathbf{H} \equiv \nabla^2 f$ is *the Laplacian* of f. In 2-dimensions and Cartesian coordinates $\nabla^2 f = f_{xx} + f_{yy}$, with a clear generalisation to higher dimensions.

Classification of Stationary Points

If $\mathbf{x} = \mathbf{x_0}$ is a stationary point $\nabla f(\mathbf{x_0}) = 0$ and so: $f(\mathbf{x}) \sim f(\mathbf{x_0}) + \frac{1}{2}\delta\mathbf{s}^T \cdot \mathbf{H} \cdot \delta\mathbf{s}$. At a minimum $\delta\mathbf{s}^T \cdot \mathbf{H} \cdot \delta\mathbf{s} > 0$ for all $\delta\mathbf{s}$, we say that $\mathbf{H}$ is *positive definite.* At a maximum $\delta\mathbf{s}^T \cdot \mathbf{H} \cdot \delta\mathbf{s} < 0$ for all $\delta\mathbf{s}$, we say that $\mathbf{H}$ is *negative definite.* At a saddle point $\mathbf{H}$ is *indefinite.* Since $\mathbf{H}$ is symmetric, it can be diagonalised, so (in a suitable basis of principal axis (x_1, x_2)):

$$\delta\mathbf{s}^T \cdot \mathbf{H} \cdot \delta\mathbf{s} = (x_1, x_2) \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda_1 x_1^2 + \lambda_2 x_2^2$$

With a clear generalisation to higher dimensions. This implies that $\mathbf{H}$ is positive definite (minimum) if *all* the eigenvalues are positive and negative definite (maximum) if *all* the eigenvalues are negative. Otherwise it is undetermined (saddle).

Criterion for Definiteness

The *signature* of $\mathbf{H}$ is the pattern of signs of the sub determinants

$$H_1 = f_{xx}, \quad H_2 = \det \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}, \cdots$$

.

$\mathbf{H}$ is positive definite (minimum) $\iff$ signature of H_i's is $+, +, \cdots, +$.
$\mathbf{H}$ is negative definite (maximum) $\iff$ signature of H_i's is $-, +, - \cdots, (-)^n$.
$\mathbf{H}$ is indefinite (saddle) otherwise. See example.

Contours of $f(x, y)$

From $f(\mathbf{x}) \sim f(\mathbf{x_0}) + \frac{1}{2}\delta\mathbf{s}^T \cdot \mathbf{H} \cdot \delta\mathbf{s}$ we can see that for contours of f (f = constant) then $\delta\mathbf{s}^T \cdot \mathbf{H} \cdot \delta\mathbf{s} \sim$ constant. This implies that in the diagonal basis (x_1, x_2), $\lambda_1 x_1^2 + \lambda_2 x_2^2 \sim$ constant. For maxima or minima λ_1, λ_2 have the same sign and this equation defines ellipsis. For a saddle λ_1, λ_2 have opposite sign and then $\lambda_1 x_1^2 + \lambda_2 x_2^2 \sim$ constant define hyperbolae. Therefore close to minima and maxima the contours are eilliptical and close to saddle points they are hyperbolae, including the crossing contours corresponding to the straight lines $x_2 = \pm\sqrt{|\lambda_1/\lambda_2|}\, x_1$.

4.2 Elements of Partial Differential Equations

Let us start discussing the basic ideas to solve partial differential equations by considering simple examples.

Simple Examples

The general first order linear PDE for a function of two variables $y(x,t)$ is: $A(x,t)\frac{\partial y}{\partial x} + B(x,t)\frac{\partial y}{\partial t} + C(x,t)y = D(x,t)$

1. If $A = 0$ or $B = 0$ then this is similar to an ODE and can be solved with the techniques we have learnt for ODE's, e.g. $y_x = \alpha x$ can be integrated immediately to give $y = \alpha^2/2 + \beta(t)$. The main difference with an ODE is that instead of finding the solution up to an arbitrary constant this solution depends on an arbitrary function $\beta(t)$.
2. A less trivial example is the *first order wave equation*: $\frac{\partial y}{\partial t} = c\frac{\partial y}{\partial x}$ Recall from section 1.2 that along a path $x(t)$: $\frac{dy}{dt} = \frac{\partial y}{\partial t} + \frac{\partial y}{\partial x}\frac{dx}{dt}$ This implies that $\frac{dy}{dt} = 0$ ($y =$ constant) along paths $\frac{dx}{dt} = -c$ for which $x + ct = x_0$ (constant). Since $x + ct = x_0$ implies $y =$ constant then the solution is $y = f(x + ct)$. The contour lines $x + ct = x_0$ are the *characteristic lines* of the PDE. Again the general solution is given in terms of an arbitrary function f now as a function of $x+ct$. This function can be fixed by imposing initial condition, *e.g.* $y(x, 0) = x^2 - 3$, this implies that $f(x) = x^2 - 3$ and then the solution for all t is $f = (x+ct)^2-3$. Another way to find the solution is to perform the change of variables $u = x+ct, v = x-ct$ which implies $y_x = y_u+y_v, y_t = c(y_u-y_v)$ and then the differential equation becomes $y_v = 0$ that has a solution $y = f(u) = f(x+ct)$. In this way identifying the characteristic lines helps to determine the proper change of variables.
3. We can also consider an inhomogeneous equation like $\frac{\partial y}{\partial t} + 5\frac{\partial y}{\partial x} = e^{-t}$. Since it is linear it shares the same property as linear ODE's for which the general solution is the general solution of the homogeneous equation y_h plus a particular solution y_p. From the previous example we can solve for y_h: $y_h = f(x - 5t)$. A particular solution can be found as $y_p = y_p(t)$ which when substituted in the equation gives $y_p = -e^{-t}$, then $y = f(x - 5t) - e^{-t}$. Imposing the initial condition $y(x, 0) = e^{-x^2}$ implies $f(x) - 1 = e^{-x^2}$ and so the solution is $y(x,t) = e^{-(x-5t)^2} - e^{-t} + 1$.
4. *Second order wave equation:* $\frac{\partial^2 y}{\partial t^2} = c^2\frac{\partial^2 y}{\partial x^2}$. This can be written as $(\frac{\partial}{\partial t} + c\frac{\partial}{\partial x})(\frac{\partial}{\partial t} - c\frac{\partial}{\partial x})y = 0$. Since these two operators commute and each would give a first order wave equation, we can immediately write the two solutions for this equation $y_1 = f(x + ct), y_2 = g(x-ct)$. From linearity we find the general solution $y = f(x+ct)+g(x-ct)$. A more explicit way to arrive at the solutions is by the same change of variables we used for the first order wave equation $u = x + ct, v = x - ct$ from which we have $y_{xx} = y_{uu} + 2y_{uv} + y_{vv}$; $y_{tt} = c^2(y_{uu} - 2y_{uv} + y_{vv})$. Plugging this into the equation gives $y_{uv} = 0$ which has the general solution $y = f(u) + g(v) = f(x + ct) + g(x - ct)$. With the initial condition $y = 1/(1 + x^2), y_t = 0$ at $t = 0$ and $y \to 0$ as $x \to \pm\infty$ we find $2f = 2g = 1/(1 + x^2)$ and $2y = 1/(1 + (x + ct)^2) + 1/(1 + (x - ct)^2)$. Which correspond to two wave packets, one moving to the left and the other to the right in the x direction.

5. *Diffusion Equation:* $\frac{\partial T}{\partial t} = k\frac{\partial^2 T}{\partial x^2}$. Here T is temperature and k a constant known as *diffusivity.* Suppose that $T(x,0) = 0, T(0,t) = H(t)$ with H the Heaviside step function. Then we can find a solution of the form $T(x,t) = \Theta(\eta)$ with $\eta = x/(2\sqrt{kt})$. substituting this in the equation gives $\Theta'' + 2\eta\Theta' = 0$ with solution $\Theta' = Ae^{-\eta^2}$ and then $\Theta = A'\mathrm{erf}\eta + B$ where $\mathrm{erf}(\eta) \equiv (2/\sqrt{\pi})\int_0^\eta e^{-z^2}dz$ is the *error function.* Since $\mathrm{erf}(\eta) \to 1$ as $\eta \to \infty$ we find $T = 1 - \mathrm{erf}(x/\sqrt{4kt})$.

4.3 Some General Aspects about PDE's

Importance of Boundary Conditions

The existence and uniqueness of solutions depend very much on the boundary conditions. For instance a first order linear PDE for $y(x,t)$ will have a well defined solution in a region if proper boundary (initial) conditions are given along a curve in the $x-t$ plane as long as the curve is *not* a characteristic (contour) line. A proper study of boundary conditions is beyond the scope of these lectures and will be covered in future courses.

The Laplacian and Important PDE's in Physics

Our physical world has at least three spatial dimensions x, y, z and time t. Physical quantities vary at different points of space and time and are therefore functions $f(x,y,z,t)$ satisfying PDE's. Most of the important equations in physics are second order PDE's involving the Laplacian $\nabla^2\phi$. The importance of the Laplacian may be inferred by taking the Taylor expansion of the previous section and compute the average of the function $f(\mathbf{x})$ in a small cube of side $2a$ centred at $\mathbf{x_0}$. The average of f is $\langle f\rangle \equiv (1/(2a)^3)\int_{-a}^{a} f dxdydz$. Since $\int_{-a}^{a} xdx = 0$, it is easy to see that this gives $\langle f\rangle = f(\mathbf{x_0}) + (a^2/6)\nabla^2 f$. Therefore $\nabla^2 f \sim (6/a^2)(\langle f\rangle - f(\mathbf{x_0}))$ implying that the Laplacian measures the difference between the average value of the function and its value at a point. *Laplace equation* $\nabla^2 f = 0$ satisfied by electric and gravitational potentials (potential theory), fluid flow and stationary heat flow. It says that the potential equals its average value at a point so it cannot increase or decrease in all directions. *Poisson's equation* $\nabla^2 f = \rho$. Satisfied by electric and gravitational potentials in presence of charge or matter density ρ. It tells that the difference between the potential and its average in a neighbourhood is proportional to the charge or mass density. *The Wave equation* $c^2\nabla^2 f = \ddot{f}$ applies when the acceleration $\ddot{f}$ is proportional to the difference between the position and its average value (Hooke's law). *The Diffusion equation* $\nabla^2 f = k\dot{f}$ states that the rate of change of the temperature is proportional to the difference between the temperature and its average value (Newton's law of cooling), etc.

Connection to ODE's We may wonder how relevant after all was to study ODE's if the most relevant equations are PDE's. However there is a technique known as *separation of variables* in which the solution for a PDE is assumed to have the form $f(x,y,z,t) = X(x)Y(y)Z(z)T(t)$ when plugged back into the PDE this reduces to an ODE for each of the functions X, Y, Z, T. For instance for $c^2 f_{xx} - f_{tt} = 0$ the ansatz $f(x,t) = X(x)T(t)$ leads to $c^2X''/X = \ddot{T}/T$ since LHS is a function of only x and the RHS is only a function of t then both have to be constant, leading to two ODE's which can be solved immediately (simple harmonic motion in each case). This technique is surprisingly powerful. Applying it to the PDE's above in Cartesian, spherical or cylindrical coordinates leads to some of the famous ODE's we have seen (like the Legendre's, Bessel's and Laguerre's equations, etc.). This is a subject that will be further studied in much detail in the Methods course.