

Differential Equations *

Zhiyuan Bai

Compiled on October 2, 2020

This document serves as a set of revision materials for the Cambridge Mathematical Tripos Part IA course *Differential Equations* in Michaelmas 2019. However, despite its primary focus, readers should note that it is NOT a verbatim recall of the lectures, since the author might have made further amendments in the content. Therefore, there should always be provisions for errors and typos while this material is being used.

Contents

0	Basic Calculus	2
1	Order of Magnitude and Taylor's Theorem	3
1.1	Asymptotic Behaviour	3
1.2	Taylor Series	4
1.3	L'Hopital's Rule	5
2	Fundamental Theorem of Calculus	5
2.1	Integration	5
2.2	Some Integration Techniques	7
3	Multivariate Calculus	7
3.1	Partial Derivative	7
3.2	The Chain Rule on Higher Dimensions	8
4	First-Order ODEs	11
4.1	First-Order Linear ODEs	11
4.1.1	Prelude: Exponential Function	11
4.1.2	Rules of Linear ODEs	11
4.1.3	Forced (inhomogeneous) First Order ODEs with Constant Coefficient	11
4.1.4	First Order ODEs with Non-constant Coefficient	12
4.2	Intermezzo: Discrete Equations	12
4.2.1	Numerical Integration	13
4.2.2	Series Solution	13
4.3	Non-linear First Order ODEs	13
4.3.1	Special Types of Non-linear First Order ODEs	14

*Based on the lectures under the same name taught by Dr. J. R. Taylor in Michaelmas 2019.

4.3.2	Isoclines and Solution Curves	14
4.3.3	Autonomous DEs	15
4.3.4	Phase Portraits	15
4.3.5	Fixed Points in Discrete Equations	16
5	Higher Order Linear ODEs	16
5.1	Second Order Linear ODEs with Constant Coefficients	16
5.2	Second Order Linear ODEs with Non-constant Coefficients	18
5.3	Phase Space	18
5.4	Special Types of ODEs	20
5.5	Dirac Delta (sorry; not sorry)	22
5.6	Higher-Order Discrete/Difference Equations	24
5.7	Series Solutions	25
6	Multivariable Equations	27
6.1	Functions of Multiple Independent Variables	27
6.2	Taylor Series for Multivariate Functions	28
6.3	Classification of Stationary Points	28
6.4	System of Linear ODEs	30
6.5	Nonlinear Systems	31
6.6	Partial Differential Equations	32
6.6.1	First Order Wave Equation	32
6.6.2	Second Order Wave Equation	33
6.6.3	Diffusion Equation	34

0 Basic Calculus

Definition 0.1. Let $f : U \rightarrow \mathbb{R}$ where $U \subset \mathbb{R}$, then if $\forall \epsilon > 0, \exists \delta > 0$ such that

$$|x - x_0| < \delta \implies |f(x) - A| < \epsilon$$

for some $x_0, A \in \mathbb{R}$, we say

$$\lim_{x \rightarrow x_0} f(x) = A$$

Theorem 0.1. Assume that

$$\lim_{x \rightarrow x_0} f(x) = A, \lim_{x \rightarrow x_0} g(x) = B$$

Then the followings hold:

$$\lim_{x \rightarrow x_0} cf(x) = cA$$

$$\lim_{x \rightarrow x_0} f(x) + g(x) = A + B$$

$$\lim_{x \rightarrow x_0} f(x)g(x) = AB$$

$$B \neq 0 \implies \lim_{x \rightarrow x_0} f(x)/g(x) = A/B$$

Proof. Check definitions. □

Definition 0.2 (One-sided limits). If $\forall \epsilon > 0, \exists \delta > 0$ such that $0 < x - x_0 < \delta \implies |f(x) - A| < \epsilon$ for some $x_0, A \in \mathbb{R}$, we say

$$\lim_{x \rightarrow x_0^+} f(x) = A$$

Similarly, if it is $0 < x_0 - x < \delta \implies |f(x) - A| < \epsilon$, then

$$\lim_{x \rightarrow x_0^-} f(x) = A$$

Definition 0.3. If the limit

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

exists for all x in a given domain, we say that f is differentiable in this domain and its derivative is

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

For sufficiently smooth functions, we can differentiate it recursively. We denote the n^{th} derivative of f as

$$\frac{d^n f}{dx^n} \text{ or } f^{(n)}(x)$$

Several immediate facts are available from here:

- Theorem 0.2.**
1. *The differential operator is linear.*
 2. *(Chain Rule) Suppose both F and g are differentiable and $f(x) = F(g(x))$, then f is differentiable and $f'(x) = F'(g(x))g'(x)$.*
 3. *(Leibniz's Rule) Suppose both u and v are differentiable and $f(x) = u(x)v(x)$, then*

$$f^{(n)}(x) = \sum_{k=0}^n \binom{n}{k} u^{(k)}(x)v^{(n-k)}(x)$$

in particular $f'(x) = u(x)v'(x) + u'(x)v(x)$.

Proof. Trivial. □

1 Order of Magnitude and Taylor's Theorem

1.1 Asymptotic Behaviour

When we want to analyze the difference between functions, apart from their analytical properties, we would also be interested in the difference between their magnitudes. The following notion arrives to serve this purpose

Definition 1.1 (Little- o notation). Let f, g be real functions and $x_0 \in \bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$, we say that $f(x) = o(g(x))$ as $x \rightarrow x_0$ if

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = 0$$

Definition 1.2 (Big- O notation). Let f, g be real functions and $x_0 \in \mathbb{R}$, we say that $f(x) = O(g(x))$ as $x \rightarrow x_0$ if $\exists \delta, M > 0$ such that

$$|x - x_0| < \delta \implies |f(x)| \leq M|g(x)|$$

We say $f(x) = O(g(x))$ as $x \rightarrow \infty$ if $\exists x_1, M > 0$ such that

$$x > x_1 \implies |f(x)| \leq M|g(x)|$$

We say $f(x) = O(g(x))$ as $x \rightarrow -\infty$ if $\exists x_1 < 0, M > 0$ such that

$$x < x_1 \implies |f(x)| \leq M|g(x)|$$

Remark. Remember that the equality sign of $f(x) = o(g(x))$ or $f(x) = O(g(x))$ is not really the usual equality sign we use. It is more like $f(x) \in o(g(x))$ or $f(x) \in O(g(x))$, meaning that f is one of those functions having this property on its magnitude. The reason why we use the equality sign here is that we sometimes use the notations to denote *some* functions having this property, which we do not (need to) know any detail except its magnitude.

Example 1.1. We have $x^2 = o(x)$ as $x \rightarrow 0$ and $x^2 = O(x)$ as $x \rightarrow 0$. In fact, whenever $f(x) = o(g(x))$ as $x \rightarrow x_0$ we have $f(x) = O(g(x))$ as $x \rightarrow x_0$ as well. The proof again is just checking definitions.

Now, one of the most significant usage of the measurement of magnitude is that we can use it to approximate the rest of the terms in a series. For example, if we take the series $1+x+x^2+x^3+\dots$, we can replace it by $1+x+x^2+x^3+o(x^3)$, in which way one can include information about the magnitude of the error term of the series.

1.2 Taylor Series

The idea of the Taylor series is to locally approximate a smooth enough function by polynomials. Surely, most functions have much worse analytical properties than polynomials, making it slightly problematic to analyze some of their properties. Taylor series provides a solution. Basically, we start by assuming a local approximation of the function by a polynomial. Say $f(x) \approx a_0 + a_1x + a_2x^2 + \dots + a_nx^n$, then, by differentiating both sides recursively, we immediately have $a_n = f^{(n)}(0)/n!$. The polynomial, which we will call $P_{n,0}(x)$, is called the Taylor polynomial. In general, if we shift the polynomial by x_0 , we have the general form:

Definition 1.3. The Taylor polynomial $P_{n,x_0}(x)$ of a function f around a point x_0 is the polynomial

$$\sum_{k=0}^n \frac{f^{(k)}(x_0)(x - x_0)^k}{k!}$$

Now the big question is, we have (quite vaguely) obtained the form of the sequence of polynomials that looks as if it can approximate f as n is big enough. But does it? Obviously unless f is a polynomial as well it has no chance that the polynomial will be equal to f , but what can we say about the magnitude of the error term? Taylor's theorem saves the day.

Theorem 1.1 (Taylor's theorem). Write $h = x - x_0$. Provided that $f^{(n+1)}$ exists, then

$$E_{n,x_0}(x) = f(x) - P_{n,x_0}(x) = O(h^{n+1})$$

as $h \rightarrow 0$.

Actually $E_{n,x_0} = o(h^n)$ as $h \rightarrow 0$ as well but the big- O here is stronger.

1.3 L'Hopital's Rule

Theorem 1.2. If $f(x)$ and $g(x)$ are both differentiable at $x = x_0 \in \bar{\mathbb{R}}$, and that f, g are both continuous at x_0 and $f(x_0) = g(x_0) = 0$, then

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

Proof. The little- o notations below are taken as $x \rightarrow x_0$.

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + o(x - x_0), g(x) = g(x_0) + (x - x_0)g'(x_0) + o(x - x_0)$$

by Taylor's theorem.

$$\frac{f(x)}{g(x)} = \frac{f'(x_0) + o(x - x_0)/(x - x_0)}{g'(x_0) + o(x - x_0)/(x - x_0)} \rightarrow \frac{f'(x_0)}{g'(x_0)}$$

as $x \rightarrow x_0$. □

Note that we can use L'Hopital's rule recursively given that the conditions still hold.

2 Fundamental Theorem of Calculus

2.1 Integration

Assume below that all functions under consideration are nice enough to let the integral exist.

Consider the following sum

$$\sum_{n=0}^{N-1} f(x_n) \Delta x$$

where $\Delta x = (b - a)/N$, $x_n = a + n\Delta x$.

The question is, how close is the finite sum above to the area under the curve of f , when N is large? How big is the difference between the area and the discrete area in the sum?

Theorem 2.1 (Mean-value Theorem on Definite Integrals). For a continuous function $f(x)$,

$$\int_{x_n}^{x_{n+1}} f(x) dx = f(x_c)(x_{n+1} - x_n)$$

for some $x_c \in (x_n, x_{n+1})$.

Expand $f(x)$ about $x \rightarrow x_n$ and evaluate it at x_c . As $\Delta x \rightarrow 0$,

$$f(x_c) = f(x_n) + O(x_c - x_n) = f(x_n) + O(x_{n+1} - x_n)$$

as $|x_c - x_n| < |x_{n+1} - x_n|$. Hence by the mean-value theorem:

$$\begin{aligned} \int_{x_n}^{x_{n+1}} f(x) dx &= f(x_n)(x_{n+1} - x_n) + O(x_{n+1} - x_n)(x_{n+1} - x_n) \\ &= \Delta x f(x_n) + O(\Delta x^2) \end{aligned}$$

So $\epsilon = O(\Delta x^2)$. It follows that

$$\int_a^b f(x) dx = \lim_{N \rightarrow \infty} \sum_{n=0}^{N-1} f(x_n) \Delta x + \epsilon_n$$

By our bounds above, the error terms $\sum_n \epsilon_n = O(N \Delta x^2) = O((b-a)^2/N)$ vanish as $N \rightarrow \infty$. Therefore,

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

Theorem 2.2 (Fundamental Theorem of Calculus). *Let*

$$F(x) = \int_a^x f(t) dt$$

then $dF/dx = f(x)$.

Proof. We try to evaluate the derivative of F ,

$$\begin{aligned} \frac{dF}{dx} &= \lim_{h \rightarrow 0} \frac{1}{h} \int_x^{x+h} f(t) dt \\ &= \lim_{h \rightarrow 0} \frac{1}{h} (f(x)h + O(h^2)) \\ &= f(x) \end{aligned}$$

So $dF/dx = f(x)$. □

Corollary 2.3.

$$\frac{d}{dx} \int_x^b f(t) dt = -f(x)$$

Corollary 2.4.

$$\int_a^{g(x)} f(t) dt = f(g(x))g'(x)$$

Corollary 2.5. *Let*

$$F(x) = \int f(x) dx$$

be the indefinite integral (or anti-derivative) of f , then

$$\int_a^b f(t) dt = F(b) - F(a)$$

2.2 Some Integration Techniques

The first one is integration by substitution.

Example 2.1.

$$\int \frac{1-2x}{\sqrt{x-x^2}} dx = \int \frac{du}{\sqrt{u}} = 2\sqrt{u} + C$$

Trigonometric substitution

Example 2.2. If we see something like $\sqrt{a^2 - x^2}$, then we can substitute $x = a \sin \theta$.

If we see something like $x^2 + a^2$, we can use $x = a \tan \theta$.

If we see $\sqrt{x^2 - a^2}$, we can use $x = a \cosh \theta$ or $x = a \sec \theta$.

If we see $\sqrt{x^2 + a^2}$, we can use $x = a \sinh \theta$ or $x = a \tan \theta$.

If we see $a^2 - x^2$, we can use $x = a \tanh \theta$

And, of course, we have integration by part:

$$\int uv' = uv - \int u'v$$

from product rule.

3 Multivariate Calculus

In many real world applications, functions we might be interested in can involve more than one independent variable.

Example 3.1. Waves along a string. Let $f(x, t)$ be the displacement where x be the position and t be the time. The shape of the wave (represented by $f(x, t_0)$ where t_0 is fixed) is dependent on time.

How do we define the derivative when a function depends on more than one variable? Suppose that $f(x, y)$ is the elevation of the terrain at the specific location x, y . We can draw a contour map of its projection on a plane. If we are interested in the steepness at a point a on the surface, one should note that different trails going through a may have different steepness. So the general point is that the slope of a function at a given point depends on the direction.

3.1 Partial Derivative

We want to find the derivative of a multivariate function with respect to one variable while keeping others fixed.

Definition 3.1. Mathematically speaking, we define partial derivative $f(x, y)$ with respect to x fixing y is the limit

$$\left. \frac{\partial f}{\partial x} \right|_y = \lim_{h \rightarrow 0} \frac{f(x+h, y) - f(x, y)}{h}$$

provided that it exists. We can define the partial derivative with respect to y fixing x similarly.

Example 3.2. Let $F(x, y) = x^2 + y^3 + e^{xy^2}$, so

$$\left. \frac{\partial f}{\partial x} \right|_y = 2x + y^2 e^{xy^2}, \quad \left. \frac{\partial f}{\partial y} \right|_x = 3y^2 + 2xy e^{xy^2}$$

We can do partial derivatives recursively as well.

$$\left. \frac{\partial^2 f}{\partial x^2} \right|_y = 2 + y^4 e^{xy^2}$$

Now we can define cross-derivative as well,

$$\left. \frac{\partial}{\partial y} \left(\left. \frac{\partial f}{\partial x} \right|_y \right) \right|_x = 2y e^{xy^2} + 2xy^3 e^{xy^2}$$

Since the notation is cumbersome, we sometimes omit the symbol $|_y$.

There is a symmetry involved in mixed partial derivatives. By that we mean

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

given that all these partial derivatives exist. Some properties are required for this equality to hold, but they are out of the scope of this course.

On higher dimensions (where we can define partial derivatives analogously), for example $f(x, y, z)$, when we sometimes say

$$\left. \frac{\partial f}{\partial x} \right|_y$$

its value would depend on the path it takes in the $x - z$ plane.

We sometimes use the shorthand notation f_x, f_{xy}, f_{xx} for the partial derivatives.

3.2 The Chain Rule on Higher Dimensions

Consider $f(x(t), y(t))$, we first want to have the concept of a differential of a function.

$$\delta f = f(x + \delta x, y + \delta y) - f(x, y)$$

So we have

$$\delta f = f(x + \delta x, y + \delta y) - f(x + \delta x, y) + f(x + \delta x, y) - f(x, y)$$

When y is held constant, $f(x + \delta x, y) = f(x, y) + \delta x (\partial f / \partial x)(x, y) + o(\delta x)$ as $\delta x \rightarrow 0$. Similarly $f(x + \delta x, y + \delta y) = f(x, y) + \delta y (\partial f / \partial y)(x + \delta x, y) + o(\delta y)$ as $\delta y \rightarrow 0$. Therefore as $\delta x \rightarrow 0, \delta y \rightarrow 0$

$$\begin{aligned} \delta f &= f(x + \delta x, y) + \delta y \frac{\partial f}{\partial y}(x + \delta x, y) + o(\delta y) - f(x + \delta x, y) \\ &\quad + f(x, y) + \delta x \frac{\partial f}{\partial x}(x, y) + o(\delta x) - f(x, y) \\ &= \delta y \frac{\partial f}{\partial y}(x + \delta x, y) + \delta x \frac{\partial f}{\partial x}(x, y) + o(\delta x) + o(\delta y) \\ &= \delta y \frac{\partial f}{\partial y}(x, y) + \delta x \frac{\partial f}{\partial x}(x, y) \\ &\quad + \delta x \delta y \frac{\partial}{\partial x} \frac{\partial f}{\partial y}(x, y) + o(\delta x) + o(\delta y) + o(\delta x \delta y) \end{aligned}$$

Note that $o(\delta x \delta y) + o(\delta x) = o(\delta x)$ as $\delta y \rightarrow 0$, so we can ignore that term. We let $\delta x, \delta y \rightarrow 0$, so

$$\begin{aligned}\delta f &= \delta y \frac{\partial f}{\partial y} + \delta x \frac{\partial f}{\partial x} + \delta x \delta y \frac{\partial}{\partial x} \frac{\partial f}{\partial y} + o(\delta x) + o(\delta y) \\ df &= \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial x} dx\end{aligned}$$

This is called the chain rule in differential form. We can obtain the chain rule by dividing by another differential dt before applying the limit.

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

This is called the multivariate chain rule.

Suppose $f(x, y(x))$, then

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

We can integrate it back

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

Note that we need to integrate the above equation along a given path, but if the function is nice enough, only the endpoints matter.

Example 3.3. We choose the paths

$$(x_1, y_1) \rightarrow (x_2, y_1) \rightarrow (x_2, y_2)$$

and

$$(x_1, y_1) \rightarrow (x_1, y_2) \rightarrow (x_2, y_2)$$

then

$$\begin{aligned}f(x_2, y_2) - f(x_1, y_1) &= \int_{x_1}^{x_2} \frac{\partial f}{\partial x}(x, y_1) dx + \int_{y_1}^{y_2} \frac{\partial f}{\partial y}(x_2, y) dy \\ &= \int_{x_1}^{x_2} \frac{\partial f}{\partial x}(x, y_2) dx + \int_{y_1}^{y_2} \frac{\partial f}{\partial y}(x_1, y) dy\end{aligned}$$

An application of the multivariate chain rule is the change of variables. It is often useful to write a differential equation in a different coordinate system before solving it. To do this, we need to transform the derivatives from one to the other.

Example 3.4. We try to transform from Cartesian to polar coordinate, so $x = r \cos \theta, y = r \sin \theta$, so we can write $f(x(r, \theta), y(r, \theta))$, so

$$\left. \frac{\partial f}{\partial r} \right|_{\theta} = \left. \frac{\partial f}{\partial x} \right|_y \left. \frac{\partial x}{\partial r} \right|_{\theta} + \left. \frac{\partial f}{\partial y} \right|_x \left. \frac{\partial y}{\partial r} \right|_{\theta} = \left. \frac{\partial f}{\partial x} \right|_y \cos \theta + \left. \frac{\partial f}{\partial y} \right|_x \sin \theta$$

We can also apply it to implicit differentiation. Consider $f(x, y, z) = c$ where c is a constant. It implicitly defined $z(x, y), x(y, z), y(z, x)$. For example, we can take $xy + y^2z + z^5 = 1$, so $x = (1 - z^5 - y^2z)/y$ and we can find y by quadratic

formula, but we cannot do it easily with z since it's quintic. However, we can find the derivative $\partial z/\partial x$ fixing y by observing

$$0 = \left. \frac{\partial f}{\partial x} \right|_y = y + y^2 \left. \frac{\partial z}{\partial x} \right|_y + 5z^4 \left. \frac{\partial z}{\partial x} \right|_y$$

which we can solve for the desired derivative.

Now consider $f(x, y, z(x, y)) = 0$,

$$df = \left. \frac{\partial f}{\partial x} \right|_{y,z} dx + \left. \frac{\partial f}{\partial y} \right|_{x,z} dy + \left. \frac{\partial f}{\partial z} \right|_{x,y} dz$$

We want the derivative of z wrt x with y fixed, so

$$\left. \frac{\partial f}{\partial x} \right|_y = \left. \frac{\partial f}{\partial x} \right|_{y,z} + \left. \frac{\partial f}{\partial z} \right|_{x,y} \left. \frac{\partial z}{\partial x} \right|_y$$

In fact, $(\partial f/\partial x)|_y = 0$.

The chain rule allows us to differentiate an integral. Consider function $f(x, c)$ where each value of c gives a different function f . We want to find the (partial) derivative of the integral

$$\int_0^b f(x, c) dx$$

Then

$$\begin{aligned} \left. \frac{\partial I(b, c)}{\partial b} \right|_c &= \lim_{h \rightarrow 0} \frac{1}{h} \int_b^{b+h} f(x, c) dx \\ &= f(b, c) \end{aligned}$$

Similarly,

$$\begin{aligned} \left. \frac{\partial I(b, c)}{\partial c} \right|_b &= \lim_{h \rightarrow 0} \frac{1}{h} \int_0^b f(x, c+h) - f(x, c) dx \\ &= \int_0^b \left. \frac{\partial I(x, c)}{\partial c} \right|_x dx \end{aligned}$$

Suppose that b, c depends on t , so $I(b(t), c(t))$, so

$$\begin{aligned} \frac{dI}{dt} &= \frac{\partial I}{\partial b} \frac{db}{dt} + \frac{\partial I}{\partial c} \frac{dc}{dt} \\ &= f(b, c) \dot{b} + \dot{c} \int_0^b \left. \frac{\partial I(x, c)}{\partial c} \right|_x dx \end{aligned}$$

In general,

$$\frac{d}{dt} \int_{a(t)}^{b(t)} f(x, c(t)) dx = \dot{c} \int_{a(t)}^{b(t)} \frac{\partial f}{\partial c} dx + f(b, c) \dot{b} - f(a, c) \dot{a}$$

Note that the reciprocal rule also apply. The same rule applies to partial derivatives given that the same parameter is kept constant.

$$\left. \frac{\partial r}{\partial x} \right|_y = \frac{1}{\left. \frac{\partial x}{\partial r} \right|_y}$$

4 First-Order ODEs

Definition 4.1. An Ordinary Differential Equation (ODE) involves differentiable functions of 1 variable, while Partial Differential Equations (PDEs) involve higher dimensional functions.

The order of a differential equations is the highest order of derivative in the equation.

4.1 First-Order Linear ODEs

An linear ODE is, obviously, a linear ODE.

Example 4.1. $x^3y + y' = 0$ is a first-order linear ODE.

4.1.1 Prelude: Exponential Function

Consider the function $f(x) = a^x$, $a > 1$, then $f'(x) = a^x \lambda$ for some $\lambda > 0$.

Definition 4.2. We define $\exp(x)$ to be the solution to the differential equation $f' = f$ with $f(0) = 1$.

By definition we have $(e^h - 1)/h \rightarrow 1$ as $h \rightarrow 0$.

Definition 4.3. One can show that \exp is strictly increasing, therefore injective. So we can define \ln to be the inverse function of \exp .

So we have $\lambda = \ln a$.

The exponential function plays a central role in differential equations because it is the eigenfunction (i.e. a function that is only scaled under the operator) of the differential operator.

We obviously have $de^{\lambda x}/dx = \lambda e^{\lambda x}$, so it is indeed an eigenfunction. Actually, all the eigenfunctions of the differential operator is of the form $Ce^{\lambda x}$ for some constant C and λ . So the behaviour of a differential equation can be somehow characterized by exponential functions.

4.1.2 Rules of Linear ODEs

Firstly, any homogeneous linear ODEs with constant coefficients (e.g. $af' + bf = 0$, $a, b \in \mathbb{R}$) have solutions of the form $Ce^{\lambda x}$ where λ can be complex.

Secondly, for linear homogeneous ODEs, any constant multiple of a solution is also a solution.

Thirdly, an n^{th} order linear ODE has only n linearly independent solutions.

Lastly, an n^{th} order ODE requires n initial/boundary conditions to uniquely determine it.

4.1.3 Forced (inhomogeneous) First Order ODEs with Constant Coefficient

Case 1: Constant forcing. For example, $5y' - 3y = 10$. We will follow the following steps:

1. Find steady (equilibrium) solution where $y' = 0$. In this case, it is $y = y_p = -10/3$.

2. Then, the general solution is in the form $y = y_p + y_c$ where y_c is a complementary solution, any solution to the differential equation removing the inhomogeneous part. in this case, $5y' - 3y = 0$.

3. Solve for y_c . In this case, $y_c = Ae^{3x/5}$

4. Plug it back: $y = Ae^{3x/5} - 10/3$

Here A is any constant. This way works since linearity of the equation granted that any two solutions to it must differ by a complementary solution.

Case 2: Eigenfunction forcing.

Suppose an isotope A decays into isotope B in a way that it is proportional to a , the number of nuclei of that isotope A. While B decays into isotope C, its rate is proportional to b , the numbers of nuclei of isotope B. So $\dot{a} = -k_a a \implies a = a_0 e^{-k_a t}$ for some constants k_a, a_0 . Now $\dot{b} = -k_b b + k_a a$ for some constant k_b , so

$$\dot{b} + k_b b = k_a a_0 e^{-k_a t}$$

This is an example where the forcing term is the eigenfunction of d/dt . We can guess a particular solution

$$b_p = \frac{k_a}{k_b - k_a} e^{-k_a t}$$

if $k_b \neq k_a$. So the general solution is

$$b = \frac{k_a}{k_b - k_a} a_0 e^{-k_a t} + D e^{-k_b t}$$

If $b(t) = 0$, then

$$b = \frac{k_a}{k_b - k_a} a_0 (e^{-k_a t} - e^{-k_b t})$$

So

$$\frac{b}{a} = \frac{k_a}{k_b - k_a} (1 - e^{(k_a - k_b)t})$$

we can solve it for t and this solution allows us to date rocks etc by measuring the ratio of isotopes.

4.1.4 First Order ODEs with Non-constant Coefficient

The general form of these sort of equations are in the form $a(x)y' + b(x)y = c(x)$ or the form $y' + p(x)y = q(x)$ (given that $\forall x, a(x) \neq 0$). We can solve it by integrating factor

$$\mu(x) = \exp\left(\int p \, dx\right)$$

So

$$(\mu y)' = \mu q \implies \mu y = \int \mu q \, dx \implies y = \frac{1}{\mu} \int \mu q \, dx$$

4.2 Intermezzo: Discrete Equations

Sometimes it is useful to consider functions evaluated at a discrete set of points, which could be useful to numerical integration and series solution.

4.2.1 Numerical Integration

One approximation to dy/dx at $x = x_n$ can be written as

$$\left. \frac{dy}{dx} \right|_{y_n} \approx \frac{y_{n+1} - y_n}{h}$$

This is called the forward Euler approximation.

Example 4.2. $5y' - 3y = 0$ again. Then it is approximately equal to the discrete equation

$$5y_{n+1} - 5y_n - 3ny_n = 0 \implies y_{n+1} = \left(1 + \frac{3h}{5}\right) y_n$$

We can iterate it to approximate the solution given an initial value. This is the example of a recurrence relation. We can actually solve this by

$$y_n = \left(1 + \frac{3h}{5}\right)^n y_0$$

Which can approximate the true solution pretty well if we let $h \rightarrow 0, n \rightarrow \infty$. Note that in this case, for finite value of n , y_n is always less than the actual value of solution at that point.

4.2.2 Series Solution

A powerful way to solve DEs is to seek for solutions in the form of an infinite power series. Let

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

we can plug it in the differential equations to solve for a_n .

Example 4.3. $5y' - 3y = 0$ yet again. Plug it in and we get

$$5 \sum_{n=0}^{\infty} (5(n+1)a_{n+1} - 3a_n)x^n = 0$$

Hence

$$5(n+1)a_n - 3a_n = 0 \implies a_{n+1} = \frac{3a_n}{5n+5} \implies a_n = \left(\frac{3}{5}\right)^n \frac{1}{n!} a_0$$

Thus

$$y(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 \sum_{n=0}^{\infty} \left(\frac{3}{5}\right)^n \frac{1}{n!} x^n = a_0 e^{3x/5}$$

for some constant a_0 .

4.3 Non-linear First Order ODEs

The general form of these sort of equations can be written in the following way:

$$Q(x, y) \frac{dy}{dx} + P(x, y) = 0$$

We cannot guarantee that a nonlinear ODE can be solved in closed form. But sometimes we can.

4.3.1 Special Types of Non-linear First Order ODEs

An ODE is said to be separable if it can be written in the form $q(y)dy = p(x)dx$, then we can solve for it by integrating both sides.

An ODE in the above form is called *exact* if and only if $Q(x, y) dy + P(x, y) dx$ is an exact differential of some function $f(x, y)$, that is $df = Q dy + P dx$. If it is the case, our general form then give $df = 0$, so $f(x, y) = 0$ is the solution.

To find it, we can make use of the multivariate chain rule to get

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \frac{dy}{dx} = 0$$

$P = f_x, Q = f_y$, hence

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial P}{\partial y}, \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial Q}{\partial x}$$

So $P_y = Q_x$. Conversely, if it is true in a simply connected domain, then $P dx + Q dy$ is an exact differential. Therefore we can use it to test whether the given ODE is exact. If so, we can find f (hence a possibly implicit expression of y) back by integration.

Example 4.4. We want to solve $6y(y-x)y' + (2x-3y^2) = 0$. So $P = 2x - 3y^2, Q = 6y(y-x)$, we can check that our preceding condition hold (i.e. $P_y = Q_x$), hence it is exact.

To find $f(x, y)$, we notice

$$\begin{cases} \partial f / \partial x|_y = 2x - 3y^2 \\ \partial f / \partial y|_x = 6y(y-x) \end{cases}$$

Integrating the first equation we have $f(x, y) = x^2 - 3xy^2 + h(y)$ where h is differentiable. Now plug it in the other equation we get $6y(y-x) = -6xy + h'(y)$, thus $h'(y) = 6y^2 \implies h(y) = 2y^3 - C$ where C is a constant. So the general solution is

$$x^2 - 3xy^2 + 2y^3 = C$$

4.3.2 Isoclines and Solution Curves

Even if (most of the time) we cannot really solve the equation, we can still analyze its behaviour. Now we consider an ODE of form $\dot{y} = f(y, t)$, each initial condition $y(t_0) = y_0$ will give a different solution curve (given existence).

Example 4.5. Suppose $\dot{y} = t(1 - y^2)$, of course this is separable and we can solve it, but without solving this, we can sketch solution curves as well.

Definition 4.4. An isocline is a curve given by \dot{y} being constant, that is, $t(1 - y^2) = c$ for some constant c .

Note that if $f(y, t)$ is single-valued (i.e. actually a function), then the solution curves cannot cross (unless as tangents of each other). To sketch the solutions, we can first sketch all the Isoclines. Note that along any isocline, the corresponding \dot{y} is constant, so we can easily draw the vector field, so we could follow the directions with initial condition given to approximate the curve.

We can analysis the stability of a fixed point.

Definition 4.5. A fixed point is a constant solution of y . That is, we have $\dot{y} = f(y, t) = 0$.

Definition 4.6. A fixed point is called stable if the solution curve in a small neighbourhood of the fixed point converge to it.

We now try to analyze the stability of fixed points using perturbation analysis.

Let $y = a$ be a fixed point of some DE $\dot{y} = f(y, t)$. Consider a small perturbation of the fixed point $y = a + \epsilon(t)$, so $\dot{\epsilon} = f(a + \epsilon, t) = f(a, t) + \epsilon f_y(a, t) + O(\epsilon^2)$, so $\dot{\epsilon} \approx \epsilon f_y(a, t)$ which is linear. The behaviour of ϵ obtained from this differential equation helps us to classify the fixed points.

If $\lim_{t \rightarrow \infty} \epsilon(t) = 0$ then we call it a stable fixed point, if $\lim_{t \rightarrow \infty} \epsilon(t) = \pm\infty$, then we say it is unstable, otherwise we say it is neutral. If $f_y(a, t) = 0$, then we need higher order terms in that Taylor series in order to determine its behaviour.

Example 4.6. $\dot{y} = t(1-y^2)$, so the fixed points are $y = \pm 1$. We have $f_y = -2ty$. For $y = 1$, then $\dot{\epsilon} = -2t\epsilon \implies \epsilon_0 e^{-t^2}$, so $\epsilon \rightarrow 0$ when $t \rightarrow \infty$. So it is stable. For $y = -1$, then $\dot{\epsilon} = 2t\epsilon \implies \epsilon_0 e^{t^2}$, so $\epsilon \rightarrow \pm\infty$ (for $\epsilon_0 \neq 0$) when $t \rightarrow \infty$. So it is unstable.

4.3.3 Autonomous DEs

Definition 4.7. A DE is called autonomous if f does not depend on t , that is, it is of the form $\dot{y} = f(y)$.

So in this case, we apply the perturbation analysis to get $\dot{\epsilon} = f'(a)\epsilon$, thus $\epsilon = \epsilon_0 e^{kt}$ where $k = f'(a)$, so if $\epsilon_0 \neq 0$,

$$\epsilon \rightarrow \begin{cases} 0, & \text{if } f'(a) < 0, \text{ so it is stable} \\ \pm\infty, & \text{if } f'(a) > 0, \text{ so it is unstable} \\ \text{Others,} & \text{if } f'(a) = 0, \text{ so it is neutral} \end{cases}$$

4.3.4 Phase Portraits

Another way to analyze the behaviour of a given DE is from a geometric perspective represented by something called phase portrait.

Example 4.7. Chemical kinetics is an important example why it is useful. Consider neutralization reaction $\text{NaOH} + \text{HCl} = \text{H}_2\text{O} + \text{NaCl}$. Let $a(t), b(t)$ be the numbers of NaOH and HCl molecules at times t , and $c(t)$ be the number of H_2O which is equal to that of NaCl.

So the initial condition is $a(0) = a_0, b(0) = b_0$, and the model is $\dot{c} = \lambda ab$, also $a = a_0 - c, b = b_0 - c$, so

$$\frac{dc}{dt} = \lambda(a_0 - c)(b_0 - c)$$

which is autonomous, so a_0, b_0 are the fixed points of the DE. We can sketch the graph of \dot{c} against c , which is called the 2D phase portrait of the DE, where we can analyze the attraction vectors near the phase portrait by the trends of it to decide the stability.

Example 4.8. Let $y(t)$ be the population at time T and birth rate would be αy for some α and the death rate be βy for some β .

Case 1: Linear model. So $\dot{y} = \alpha y - \beta y \implies y = y_0 e^{(\alpha - \beta)t}$, so $y \rightarrow \infty$ if $\alpha > \beta$.

Case 2: Nonlinear model. So $\dot{y} = (\alpha - \beta)y - \gamma y^2$. the parameter γ here could be due to the death due to overcrowdedness, etc. Equivalently, we can write it as $\dot{y} = ry(1 - y/\lambda)$ where $r = \alpha - \beta$, $\lambda = (\alpha - \beta)/\gamma$, so we can sketch the phase portrait again. Note that the fixed points are 0 and λ , the former is unstable but the latter is stable.

4.3.5 Fixed Points in Discrete Equations

We can introduce fixed points in discrete equations as well. Consider a first order discrete equation (aka difference equation) of the form $x_{n+1} = f(x_n)$.

Definition 4.8. The fixed point is a discrete equation is a fixed point of the function f .

We can analyze its stability again by perturbation analysis. Let x_f be a fixed point and we perturb it by a small ϵ , so we can expand f in terms of Taylor series

$$f(x_f + \epsilon) = x_f + \epsilon f'(x_f) + O(\epsilon^2), \epsilon \rightarrow 0$$

So if $x_n = x_f + \epsilon$, $x_{n+1} = x_f + \epsilon f'(x_f)$. So x_f is stable if $|f'(x_f)| < 1$, neutral if $|f'(x_f)| = 1$, unstable if $|f'(x_f)| > 1$.

Example 4.9 (Discrete Logistic Equation). Some populations are born in distinct generations (e.g. lambs born in spring). A nonlinear model of this is

$$\frac{x_{n+1} - x_n}{\Delta t} = \lambda x_n - \gamma x_n^2$$

This is the discrete version of Example 4.8, so

$$x_{n+1} = (1 + \lambda \Delta t)x_n - \gamma \Delta t x_n^2$$

We can write alternatively $x_{n+1} = rx_n(1 - x_n) =: f(x_n)$ The function f is called the logistic map.

Note that the fixed point occurs at 0 or $1 - 1/r$. We analyze its stability by perturbation. At 0, $f' = r$, so it is stable if $0 < r < 1$ and unstable if $r > 1$. At $1 - 1/r$, $f' = 2 - r$. We assume $r < 0$ or $r > 1$ since the other cases are not physical. ¹ Then it is stable for $1 < r < 3$ and unstable for $r > 3$.

5 Higher Order Linear ODEs

5.1 Second Order Linear ODEs with Constant Coefficients

Definition 5.1. A second order linear ODE with constant coefficient is an ODE of the form

$$a \frac{d^2 y}{dx^2} + b \frac{dy}{dx} + cy = f(x)$$

where a, b, c are constants with $a \neq 0$.

¹It's an applied course, so what the hell.

Definition 5.2. A linear differential operator \mathcal{D} is a linear combination of (different orders of) differentiation operators.

We know that the (arbitrary order) differentiation operator is linear, hence any linear differential operator is linear, which gives rise to the principle of superposition.

Proposition 5.1. If y_α, y_β are solutions to $ay'' + by' + cy = f$, then $y_\alpha - y_\beta$ is a solution to $ay'' + by' + cy = 0$.

Proof. Consider the linear operator

$$\mathcal{D} = a \frac{d^2}{dx^2} + b \frac{d}{dx} + c$$

then $\mathcal{D}(y_\alpha - y_\beta) = \mathcal{D}(y_\alpha) - \mathcal{D}(y_\beta) = f(x) - f(x) = 0$. □

We can extend the above proposition to any order of linear ODEs with constant coefficient in the obvious way.

So we can solve these solutions in the following steps, assuming \mathcal{D} is defined as in the above proof:

First, we find complementary (linearly independent, defined below) solutions y_1, y_2 to $\mathcal{D}(y) = 0$.

Then we find a particular solution y_p to $\mathcal{D}(y) = f(x)$.

The general form of the solutions to $\mathcal{D}(y) = f(x)$ is $y_p + Ay_1 + By_2$ where A, B are constants.

Definition 5.3. A set of functions $(f_i)_{i \in I}$ are linearly dependent if $\sum_i c_i f_i(x) = 0$ for some constants c_i that are not all zero. The sum here is taken over some finite set of indices.

Otherwise, they are linearly independent.

Equivalently, if a function in the set of functions can be written as a linear combination of others, then the set is linearly dependent.

Consider a second-order linear differential operator. We know that the a first order one has eigenfunction to be the exponential function. Note that it is also the eigenfunction of a second-order one. In fact, the exponential is the eigenfunction of any linear differential operator. Consider the homogeneous equation $\mathcal{D}y = 0$ where

$$\mathcal{D} = a \frac{d^2}{dx^2} + b \frac{d}{dx} + c$$

Plug in $y = e^{\lambda x}$ we have $a\lambda^2 + b\lambda + c = 0$ which we call *characteristic equation* or *auxiliary equation*. From FTA, we have at least 1 (complex) solutions. Let λ_1, λ_2 be two solutions, then

Case 1: $\lambda_1 \neq \lambda_2$.

Then $y_1 = Ae^{\lambda_1 x}$, $y_2 = Be^{\lambda_2 x}$ are both solutions for each A, B constants. It is easy to see that these two are linearly independent for $AB \neq 0$. We can show that they form a basis for solution space,² and any other solutions must be of the form $Ae^{\lambda_1 x} + Be^{\lambda_2 x}$ for A, B constants.

Case 2: $\lambda_1 = \lambda_2$.

In this case, y_1, y_2 cannot span the solution space, but there is some workaround.

²This will be discussed later.

Example 5.1. $y'' - 4y' + 4y = 0$, then $(y' - 2y)' - 2(y' - 2y) = 0$ can give the solution.

Or we can consider the slightly modifies equation $y'' - 4y' + (4 - \epsilon^2)y = 0$ for some small enough ϵ . This modified equation has the general solution $y_\epsilon = Ae^{\lambda_1 x} + Be^{\lambda_2 x}$ where $\lambda_{1,2} = 2 \pm \epsilon$, so

$$y_\epsilon = e^{2x}(Ae^{\epsilon x} + Be^{-\epsilon x}) = e^{2x}((A + B) + \epsilon x(A - B) + O(\epsilon^2)) \rightarrow Ce^{2x} + De^{2x}$$

for some constants C, D by clever (or not) choices of A, B . This gives a pair of linearly independent solutions.

5.2 Second Order Linear ODEs with Non-constant Coefficients

Again we are interested in the homogeneous ones due to the superposition principle. Consider the equations in the form

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

We shall use the method of reduction of order. Given one solution y_1 to the equation, we shall find a second solution by looking for solutions of the form $y_2 = vy_1$.

First, note that $y_2' = v'y_1 + vy_1'$, $y_2'' = v''y_1 + 2v'y_1' + vy_1''$, so plugging it in we have

$$v'(2y_1' + py_1) + v''y_1 = 0$$

So it is a separable first order equation in v' which we know how to solve, plugging it back gives the solution.

5.3 Phase Space

Consider the ODE $p_i(x)y^{(i)} = f(x)$ where the summation convention is used and i is summed over $0, 1, \dots, n$. So we can write $y^{(n)}$ as a combination of $y^{(i)}$ for $i \in \{0, 1, 2, \dots, n-1\}$ and f and p_i 's.

Example 5.2. The damped oscillator has the DE

$$m\ddot{y} = -ky - L\dot{y}$$

The state of the system can be described by an n -dimensional solution vector

$$\underline{y} = \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix}$$

Going back to an undamped oscillator $y'' + 4y = 0$ which has the general solution spanned by $y_1(x) = \cos 2x$, $y_2(x) = \sin 2x$, so the solution vectors are

$$\underline{y}_1 = \begin{pmatrix} \cos 2x \\ -2 \sin 2x \end{pmatrix}, \underline{y}_2 = \begin{pmatrix} \sin 2x \\ 2 \cos 2x \end{pmatrix}$$

Thus we can do 2D phase portrait of the solutions to observe the two vectors, where we can find that the trajectories coincide. Since y_1, y_2 are linearly independent, any point in phase space can be obtained from a linear combination of them. In general, y_1, y_2, \dots, y_n are linearly independent if their solution vectors are linearly independent in the phase space.

n linearly independent solution vectors form a basis for the phase space of an n^{th} order ODE. Consider the initial conditions for a second order homogeneous ODE $y(0) = a, y'(0) = b$. If the general solution is formed by the linear combination of linearly independent functions y_1, y_2 , then in order to find a solution that complies with the initial condition, we will be solving the linear system in A, B .

$$\begin{cases} Ay_1(0) + By_2(0) = a \\ Ay_1'(0) + By_2'(0) = b \end{cases}$$

So we obtain unique solutions if and only if

$$y_1(0)y_2'(0) \neq y_2(0)y_1'(0)$$

which is true if $\underline{y_1}(0), \underline{y_2}(0)$ are linearly independent.

Definition 5.4. The Wronskian $W(x)$ is defined as

$$W(x) = \begin{vmatrix} y_1 & y_2 & \dots & y_n \\ y_1' & y_2' & \dots & y_n' \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \dots & y_n^{(n-1)} \end{vmatrix}$$

So the solutions are linearly independent if $W(x) \neq 0$. But does $W(x) = 0$ necessarily imply linear dependence?

Theorem 5.2 (Abel's Theorem). *Consider a second order linear ODE $y'' + py' + qy = 0$. If $p(x), q(x)$ are continuous on an interval I , then either $\forall x \in I, W(x) \neq 0$ or $\forall x \in I, W(x) = 0$*

Sketch of proof. Let y_1, y_2 be solutions to the ODE, and \mathcal{D} be the differential operator in the left hand side of the equation, so

$$y_2\mathcal{D}(y_1) - y_1\mathcal{D}(y_2) = 0 \implies y_2y_1'' - y_1y_2'' + (y_2y_1' - y_1y_2')p \implies W' + pW = 0$$

So we could integrate it back from x_0 to x to get

$$W(x) = W(x_0) \exp\left(-\int_{x_0}^x p(u) du\right)$$

which is called the Abel identity. Since p is continuous on a closed interval, it is bounded and integrable, therefore the exponential function part is always defined and nonzero, so $W(x_0) = 0$ if and only if $W(0) = 0$. \square

Corollary 5.3. *If $p(x) = 0$, then the Wronskian is constant.*

One can generalize the theorem above to n^{th} order linear ODEs. Indeed, for $\underline{y}' + A(x)\underline{y} = \underline{0}$, we have $W' + \text{tr}(A(x))W = 0$.

One practical application of Abel's identity can be used to find a second solution y_2 given one solution y_1 . This can be done by observing that

$$y_1 y_2' - y_2 y_1' = W(x) = W(x_0) \exp\left(-\int_{x_0}^x p(u) du\right)$$

is both linear and of first order in y_2 .

We cannot, of course, (analytically) solve all ODEs, but for some special types of them, it is sometimes possible.

5.4 Special Types of ODEs

Definition 5.5. An ODE is called equidimensional if the differential operator is unaffected by a multiplicative rescaling. So $\mathcal{D}_x = \mathcal{D}_{\bar{x}=\alpha x}$ where α is a constant. So the general form of a second order linear equidimensional is the following:

$$ax^2 \frac{d^2 y}{dx^2} + bx \frac{dy}{dx} + cy = f(x)$$

There are two methods to find y_c .

Method 1: Note that $y = x^k$ is an eigenfunction of the eigenvector of $x(d/dx)$. To solve the homogeneous equation, we can plug the eigenfunction in and find out that k satisfies

$$ak(k-1) + bk + c = 0$$

If there are two roots k_1, k_2 , then the general solution is $y_c = Ax^{k_1} + Bx^{k_2}$ where A, B are constants. Otherwise, we have at least one y_c , so we could use the method in previous sections to find the other.

Method 2: Use the substitution $z = \ln x$, so we will have

$$a \frac{d^2 y}{dz^2} + (b-a) \frac{dy}{dz} + cy = 0$$

which we know how to solve.

For the forced type of ODE, due to the superposition principle, we just (and will) need a particular solution y_p . There are two methods to do it.

Method 1: Guesswork.

Form of $f(x)$	Educated guess
e^{kx}	Ae^{kx}
$\sin(kx), \cos(kx)$	$A \sin kx + B \cos kx$
Polynomial	Polynomials

We could plug the educated guesses in and solve for the coefficients.

Method 2: Method of parameters. Given complementary functions y_1, y_2 and solution vectors $\underline{y}_1, \underline{y}_2$. Suppose that the solution vector \underline{y}_p for y_p satisfies

$$\underline{y}_p = u(x)\underline{y}_1 + v(x)\underline{y}_2$$

So what we now try to do is to define two equations for u', v'

$$\begin{cases} y_p = uy_1 + vy_2 \\ y_p' = uy_1' + vy_2' \end{cases}$$

So by differentiating the first equation and comparing it with the other, we have $u'y_1 + v'y_2 = 0$. Now differentiate the second equation and plug back to the differential equation, we know that $u'y'_1 + v'y'_2 = f(x)$. Since the Wronskian is not zero due to definitions of y_1, y_2 , we know that we can solve for u', v' and a unique solution is guaranteed. Indeed, $u' = -fy_2/W, v' = fy_1/W$. So

$$y_p = y_2 \int_{x_0}^x \frac{y_1(t)f(t)}{W(t)} dt - y_1 \int_{x_0}^x \frac{y_2(t)f(t)}{W(t)} dt$$

Definition 5.6. A forced oscillating ODE is a linear second order ODE forced by oscillating force.

This arises as many physical systems have a restoring force and damping (e.g. friction).

Example 5.3. A wheel of mass M is connected to a spring and a damper from above and a force $F(t)$ is applied from below, so $M\ddot{y} = F(t) - ky - L\dot{y}$, written in standard form,

$$\ddot{y} + \frac{L}{M}\dot{y} + \frac{k}{M}y = \frac{F(t)}{M}$$

we redefine time $\tau = \sqrt{k/M}t$ and get

$$y'' + 2Ky' + y = f(\tau), K = \frac{L}{2\sqrt{kM}}$$

We can evaluate the unforced (free, homogeneous) solution where $f \equiv 0$, which we know how to solve in the general form $Ay_1 + By_2$ where A, B are constants.

Case 1: $K < 1$, so we have complex roots of auxiliary equation, so we say the system is underdamped, so

$$y = e^{-K\tau} [A \sin(\sqrt{1 - K^2}\tau) + B \cos(\sqrt{1 - K^2}\tau)]$$

Case 2: $K = 1$, which system we call it is a critically damped.

$$y = (A + B\tau)e^{-K\tau}$$

Case 3: $K > 1$, where the system is overdamped.

$$y = Ae^{\lambda_1\tau} + Be^{\lambda_2\tau}, \lambda_{1,2} = -K \pm \sqrt{K^2 - 1}$$

So the unforced response always decays exponentially as time goes to infinity. As for forced response, if we have

$$\ddot{y} + \mu\dot{y} + \omega_0^2 y = \sin \omega t$$

which by guessing we have the particular solution

$$y_p = \frac{\omega_0^2 - \omega^2}{\omega_0^2 - \omega^2 + \mu^2\omega^2} \sin \omega t + \frac{-\mu\omega}{\omega_0^2 - \omega^2 + \mu^2\omega^2} \cos \omega t$$

For $\mu \neq 0$, we have finite amplitude oscillations matching the forcing frequency. Even when $\omega = \omega_0$, then taking the limit to get $-(\cos \omega t)/(\mu\omega)$, in which case

we still have finite amplitude oscillations.

So in general, in a damped system, the unforced part gives the short time response whilst the particular solution gives the long time behaviour.

But if $\mu = 0, \omega_0 = \omega$, we call this is a resonance. The forcing then matches the unforced response, where the equation turns to

$$\ddot{y} + \omega_0 y = \sin \omega_0 t$$

We shall use detuning to solve the equation. Consider the equation

$$\ddot{y} + \omega_0 y = \sin \omega t$$

So by guessing, we have $y_p = \sin \omega t / (\omega_0^2 - \omega^2)$. Due to linearity, $y_p + Ay_c$ works for any constant A , so

$$y_p = \frac{\sin \omega t - \sin \omega_0 t}{\omega_0^2 - \omega^2} = \frac{2}{\omega_0^2 - \omega^2} \cos\left(\frac{\omega_0 + \omega}{2} t\right) \sin\left(\frac{\omega - \omega_0}{2} t\right)$$

also solves the equation. By letting $\omega_0 \rightarrow \omega$, we have

$$y_p = -\frac{t \cos \omega_0 t}{2\omega_0}$$

which indeed solves the problem.

5.5 Dirac Delta (sorry; not sorry)

Definition 5.7. An impulse forcing is a type of forcing by a sudden change.

Example 5.4. A mountainbike riding from the road onto a side block. When they hits, then there is a sudden increase in altitude, which might be followed by some oscillation. When the time for this goes to zero, then it is considered as an (infinite) impulse. So the forced, damped oscillator has the equation

$$M\ddot{y} = F(t) - ky - L\dot{y}$$

Now if we add the impulse, and we integrate both sides and let the time elapsed during the hitting to the road block tend to 0, we will want to define the impulse by

$$I = \lim_{\epsilon \rightarrow 0} \int_{T-\epsilon}^{T+\epsilon} F(t) dt = \lim_{\epsilon \rightarrow 0} M[\dot{y}]_{T-\epsilon}^{T+\epsilon}$$

We then introduce the Dirac delta function. ³

Definition 5.8. Fix a family of functions $D(t; \epsilon)$ indexed by ϵ with the properties:

1. For any $t \neq 0$, we have $\lim_{\epsilon \rightarrow 0} D(t; \epsilon) = 0$ for $t \neq 0$.
- 2.

$$\int_{\mathbb{R}} D(t; \epsilon) dt = 1$$

We “define” the Dirac delta function by $\delta(t) = \lim_{\epsilon \rightarrow 0} D(t; \epsilon)$.

³And this is the end of the world.

Proposition 5.4. 1. $\forall t \neq 0, \delta(t) = 0$.

2. For any $a < 0 < b$,

$$\int_a^b \delta(t) dt = 1$$

3. We have the sampling property. For all functions $g(x)$ which are continuous at $x = 0$, then

$$\int_{\mathbb{R}} g(x)\delta(x) dx = g(0)$$

In general

$$\int_a^b g(x)\delta(x - x_0) dx = \begin{cases} g(x_0), & \text{if } x_0 \in [a, b] \\ 0, & \text{otherwise} \end{cases}$$

Proof. Ahem. □

Definition 5.9. The Heaviside step function $H(x)$ is defined by

$$H(x) = \int_{-\infty}^x \delta(x) dx = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{if } x < 0 \end{cases}$$

And $H(0)$ is not defined.

It is decreed that $H'(x) = \delta(x)$ “by FTC”. ⁴

Definition 5.10. The ramp function $r(x)$ is defined by

$$r(x) = \int_{-\infty}^x H(x) dx = \begin{cases} x, & \text{when } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

By FTC, $r' = H$. ⁵

Consider

$$y'' + py' + qy = \delta(x)$$

Since $\delta(x) = 0$ for all $x \neq 0$, so $y'' + py' + qy = 0$ for $x < 0$ and $x > 0$. The highest order derivative inherits the discontinuity from the forcing. But we would want y to be continuous at $x = 0$, so $\lim_{\epsilon \rightarrow 0} [y]_{-\epsilon}^{\epsilon} = 0$. And y' would have a jump near 0, so

$$\lim_{\epsilon \rightarrow 0} [y']_{-\epsilon}^{\epsilon} = \lim_{\epsilon \rightarrow 0} \int_{-\epsilon}^{\epsilon} y'' + py' + qy dx = 1$$

We first solve the equations for $x < 0$ and $x > 0$, so we will have 4 undetermined constants. and 2 initial conditions, so we need 2 more equations, which we can find from the two jump conditions as stated above.

Example 5.5. Consider $y'' - y = 3\delta(x - \pi/2)$ with $y = 0$ at $x = 0, \pi$. Note that $y'' - y = 0 \implies y = A \sinh x + B \cosh x$ Our initial conditions then implies

$$y = \begin{cases} A \sinh x, & \text{when } x < \pi/2. \\ C \sinh(\pi - x), & \text{when } x > \pi/2. \end{cases}$$

⁴I know, I know, stay calm, it is an applied course.

⁵Lol I guess?

For constants A, C . So we put in the jump condition to have $3 = \int_{\pi/2-\epsilon}^{\pi/2+\epsilon} y'' - y \, dx = [y']_{(\pi/2)^-}^{(\pi/2)^+}$. Putting in the definitions to have

$$-A \cosh(\pi/2) - c \cosh(\pi/2) = 3$$

Since we also have $0 = [y]_{(\pi/2)^-}^{(\pi/2)^+}$, we can solve to get $A = C$, therefore

$$A = C = \frac{-3}{2 \cosh(\pi/2)}$$

Thus

$$y = \begin{cases} \frac{-3 \sinh x}{2 \cosh(\pi/2)}, & \text{when } x \leq \pi/2. \\ \frac{-3 \sinh(\pi-x)}{2 \cosh(\pi/2)}, & \text{when } x > \pi/2. \end{cases}$$

We can also have the forcing by Heaviside function. Consider

$$y'' + py' + qy = H(x - x_0)$$

For p, q continuous. $y(x)$ satisfies $y'' + py' + qy = 0$ for $x < x_0$ and $y'' + py' + qy = 1$ for $x > x_0$. We can evaluate the equation on either sides of x_0 , thus

$$[y'']_{x_0^-}^{x_0^+} + p[y']_{x_0^-}^{x_0^+} + q[y]_{x_0^-}^{x_0^+} = 1$$

If y'' behaves like the Heaviside function (then y' behaves like the ramp function), then y', y are continuous, thus $[y']_{x_0^-}^{x_0^+} = [y]_{x_0^-}^{x_0^+} = 0$ and $[y'']_{x_0^-}^{x_0^+} = 1$, which is our jump conditions, which would be enough to find out the constants with the initial conditions.

5.6 Higher-Order Discrete/Difference Equations

Definition 5.11. The general form of an m^{th} order linear discrete equation with constant coefficient is

$$a_m y_{n+m} + a_{m-1} y_{n+m-1} + \cdots + a_0 y_n = f(n)$$

Turns out that they are closely related to higher order DEs, and we can solve them using the same principles.

Definition 5.12. A difference operator \mathcal{D} is such that $\mathcal{D}(y_n) = y_{n+1}$. It has eigenfunctions in the form $y_n = k^n$ for constant k as $\mathcal{D}(k^n) = k(k^n) = ky_n$.

Note that our difference equation is linear in y , thus we can dissolve y into sum of particular and complementary solutions $y_n = y_n^{(c)} + y_n^{(p)}$.

Example 5.6. We want to solve $a_2 y_{n+2} + a_1 y_{n+1} + a_0 y_n = f_n$. Consider the homogeneous equation with $f = 0$, then we can put in $y_n = k^n$ to get $a_2 k^2 + a_1 k + a_0 = 0$, so its solutions $k_{1,2}$ gives the general form of the complementary solution

$$y_n^{(c)} = \begin{cases} Ak_1^n + Bk_2^n, & \text{if } k_1 \neq k_2 \\ Ak_1^n + Bnk_1^n, & \text{if } k_1 = k_2 \end{cases}$$

Note that these are all the complementary solutions since it has 2 degrees of freedom and the sequence would be completely determined by its value at the first two initial values.

We can use guessing works for particular solution

Form of f_n	Form of $y_n^{(p)}$
$k^n, k \neq k_{1,2}$	Ak^n
$k_{1,2}^n$	$Ank_1^n + Bnk_2^n$
Polynomial	Polynomials

The Fibonacci numbers are defined as $y_0 = y_1 = 1, y_{n+1} = y_n + y_{n-1}$ for $n \geq 1$. Note that it has the auxiliary equation $k^2 - k - 1 = 0$, which has roots $k_{1,2} = (1 \pm \sqrt{5})/2$. So by plugging in our initial conditions, we obtain

$$y_n = \frac{1}{\sqrt{5}}(\phi^{n+1} - (-\phi^{-1})^{n+1})$$

where $\phi = (\sqrt{5}+1)/2$ is the golden ratio. In particular, $y_{n+1}/y_n \rightarrow \phi$ as $n \rightarrow \infty$.

5.7 Series Solutions

Often, there are no analytic solutions to some particular ODE or it is very hard to obtain one. In this case, we can try to solve the equation in the form of an infinite power series. We can use the method of Frobenius. Consider the ODE $py'' + qy' + ry = 0$. We will seek a series expansion at $x = x_0$ of a (local) solution around some point. There are many choices of x_0 . If the series expansions of q/p and r/p converge locally at x_0 , we say x_0 is an ordinary point. Otherwise, we say it is a singular point. There are two types of singular point: If x_0 is a singular point but the equation can be written in the following way:

$$P(x - x_0)^2 y'' + Q(x - x_0) y' + Ry = 0$$

and $Q/P, R/P$ are analytic, then we say x_0 is a regular singular point. Note that $Q/P = (x - x_0)q/p, R/P = (x - x_0)^2 r/p$. Otherwise it is called an irregular singular point.

Example 5.7. 1. We want to solve

$$(1 - x^2)y'' - 2xy' + 2y = 0$$

then $q/p = -2x/(1 - x^2), r/p = 2/(1 - x^2)$, so $x = \pm 1$ are singularities. But $Q/P = (x - x_0)q/p = -2x/(1 + x)$, so $x = 1$ is regular. Similarly $x = -1$ is regular as well.

2. Consider

$$y'' \sin x + y' \cos x + 2y = 0$$

Then the singularities are $n\pi, n \in \mathbb{Z}$, but since $(x - n\pi)/(\sin x)$ as $x \rightarrow n\pi$ tends to a limit, every of them is regular.

3. We look into

$$(1 + \sqrt{x})y'' - 2xy' + 2y = 0$$

So $q/p = -2x/(1 + \sqrt{x})$, one can find that the Taylor series at 0 is undefined. Indeed, 0 is an irregular singular point here.

Theorem 5.5 (Fuch's Theorem). 1. If $x = x_0$ is an ordinary point, then there are two linearly independent power series solutions of the form

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

locally near x_0 .

2. If $x = x_0$ is a regular singular point, then there is at least 1 solution of the form

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

where σ is real and $a_0 \neq 0$.

Example 5.8. 1. The equation $(1 - x^2)y'' - 2xy' + 2y = 0$ has singular points ± 1 and they are both regular. We first find series solution about an ordinary point, say $x = 0$. We try

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

So by plugging in,

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

From which we have $a_n n(n-1) - a_{n-2}(n-2)(n-3) - 2a_{n-2}(n-2) + 2a_{n-2} = 0$ for $n \geq 2$. Just simplify to get $n(n-1)a_n = (n^2 - 3n)a_{n-2}$, so

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

Consequently, a_0, a_1 , which can be arbitrary constants, could be our initial condition. Note that $a_3 = 0$, hence $a_k = 0$ for any odd $k \geq 3$. for even values of n , we have

$$a_n = \frac{n-3}{n-1} a_{n-2} = \frac{n-3}{n-1} \frac{n-5}{n-3} a_{n-4} = \frac{n-5}{n-1} a_{n-4} = \dots = \frac{n-2k-1}{n-1} a_{n-2k}$$

Thus $a_{2k} = a_0/(1-n)$, so

$$y = a_1 x + a_0 \left(1 - x^2 - \frac{x^4}{3} - \frac{x^6}{5} - \dots \right) = a_1 x + a_0 \left(1 - \frac{x}{2} \ln \frac{1+x}{1-x} \right)$$

2. Consider $4xy'' + 2(1 - x^2)y' - xy = 0$, which has a regular singular point at $x = 0$. We now try to expand the solution near it. We try $y = \sum_{n=0}^{\infty} a_n x^{n+\sigma}$ for $a_0 \neq 0$ by Fuch's Theorem. So we can plug it in our equation to try and get a recurrence for the coefficients. Note first that, by multiplying x to both sides

$$\sum_{n=0}^{\infty} a_n x^{n+\sigma} (4(n+\sigma)(n+\sigma-1) + 2(1-x^2)(n+\sigma) - x^2) = 0$$

Thus by comparing coefficients,

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

which is our equivalence relation.

To find σ , we will equate coefficient of lowest power of x . Set $n = 0$, then we can equate the coefficient of x^σ , so $a_0(4\sigma(\sigma-1)) + a_0 2\sigma = 0 \implies 2\sigma(2\sigma-1)a_0 = 0$. This is called the indicial equation. So we get $\sigma = 0$ or $\sigma = 1/2$.

For $\sigma = 0$, we again set $n = 0$ to find that a_0 is arbitrary. Then consider $n = 1$, we have $2a_1 = 0 \implies a_1 = 0$. Our equivalence relation reduced to $2n(2n-1)a_n = (2n-3)a_{n-2}$, which means that $a_k = 0$ for any odd k . For even $n = 2k$, we can calculate a few values to get

$$y = a_0 \left(1 + \frac{x^2}{4 \cdot 3} + \frac{5x^4}{8 \cdot 7 \cdot 4 \cdot 3} + \dots \right)$$

For $\sigma = 1/2$, our recurrence relation reduced to $(2n+1)(2n)b_n = (2n-2)b_{n-2}$ for $n \geq 2$ where $b_n = a_n$ to avoid ambiguity. We can equate the coefficient in the lowest power to get b_0 being arbitrary, and by considering $n = 1$ we have $b_1 = 0$, so $b_k = 0$ for any odd k as well. We can calculate b_n for n even and get

$$y = b_0 x^{1/2} \left(1 + \frac{x^2}{2 \cdot 5} + \frac{5x^4}{2 \cdot 5 \cdot 4 \cdot 9} + \dots \right)$$

There are some special cases of the indicial equation. Let x_0 be a regular singular point, and suppose $\text{Re}(\sigma_1) \leq \text{Re}(\sigma_2)$ where $\sigma_{1,2}$ are the roots of the indicial equation.

Case 1: $\sigma_2 - \sigma_1$ is a non-integer, then the two solutions are linearly independent.

Case 2: It is a nonzero integer. In this case, it is possible, but no guarantee, that the solutions y_1, y_2 are linearly dependent, so we might need an extra term in the form $cy_1 \ln(x - x_0)$ in y_2 , where c is a constant.

Case 3: It is 0. Hence $c \neq 0$, so we can set $c = 1$, so we may add $y_1 \ln(x - x_0)$ to y_2 in order to yield two linearly independent solutions.

6 Multivariable Equations

6.1 Functions of Multiple Independent Variables

Consider $f(x, y)$ and a small vector displacement $\underline{ds} = (dx, dy)$, so the change along \underline{ds} is

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

by the multivariate chain rule. Here, $\nabla f = (f_x, f_y)$ is called the gradient of f . If we write $\underline{ds} = ds \hat{s}$ with $|\hat{s}| = 1$, so $df = ds(\hat{s} \cdot \nabla f)$

Definition 6.1. The directional derivative is defined as

$$\frac{df}{ds} = \hat{s} \cdot \nabla f$$

Proposition 6.1. 1. The magnitude of ∇f is the maximum rate of change of f , that is

$$|\nabla f| = \sup_{\hat{s}, |\hat{s}|=1} \frac{df}{ds}$$

And the supremum can be attained.

2. The direction of ∇f is the direction where f increases most rapidly.

3. If \hat{s} is parallel to the contour of f , then $df/ds = 0$.

Proof. 1. Cauchy-Schwarz and take $\hat{s} = \nabla f / |\nabla f|$.

2,3. Obvious. □

Corollary 6.2. *There is always one direction where $df/ds = 0$.*

Proof. Immediate. □

Proposition 6.3. *Local extrema of f have $df/ds = 0$ for any direction. Hence $\nabla f = \underline{0}$.*

Proof. Trivial. □

However, if $\nabla f = \underline{0}$ at some point, it does *not* follow that the point is an local extremum, since it could be a saddle point. Near local extrema, the contour of f is elliptical while it is hyperboly near saddle points. Also, contours of f can only cross at saddle point.

6.2 Taylor Series for Multivariate Functions

Consider $f(x, y)$ near a point $\underline{s}_0 \in \mathbb{R}^2$ and a displacement δs along the line $\underline{\delta s}$, so

$$\begin{aligned} f(s_0 + \delta s) &= f(s_0) + \delta s \frac{df}{ds} + \frac{(\delta s)^2}{2} \frac{d^2 f}{ds^2} + \dots \\ &= f(s_0) + \delta s (\hat{\underline{s}} \cdot \nabla f) + \frac{(\delta s)^2}{2} (\hat{\underline{s}} \cdot \nabla)(\hat{\underline{s}} \cdot \nabla) f + \dots \end{aligned}$$

To write it out in coordinate form,

Definition 6.2. The Hessian matrix is defined by

$$H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \nabla(\nabla f)$$

Note that H is symmetric whenever f is nice enough (e.g. C^2) to have $f_{yx} = f_{xy}$.

We have, by this notation,

$$\begin{aligned} f(\underline{x}) &= f(x, y) \\ &= f(x_0, y_0) + (x - x_0)f_x + (y - y_0)f_y \\ &\quad + \frac{1}{2}((x - x_0)^2 f_{xx} + 2(x - x_0)(y - y_0)f_{xy} + (y - y_0)^2 f_{yy}) + \dots \\ &= f(\underline{x}_0) + \nabla f(\underline{x}_0)(\underline{x} - \underline{x}_0)^\top + \frac{1}{2}(\underline{x} - \underline{x}_0)H(\underline{x}_0)(\underline{x} - \underline{x}_0)^\top + \dots \end{aligned}$$

6.3 Classification of Stationary Points

When $\nabla f = \underline{0}$ at some \underline{x}_0 , we have

$$f(\underline{x}_0 + \underline{\delta x}) \approx f(\underline{x}_0) + \frac{1}{2} \underline{\delta x} H(\underline{x}_0) \underline{\delta x}^\top$$

Note that this can extend analogously to n dimensions.

Definition 6.3. The Hessian in n dimensions is defined by

$$H = \begin{pmatrix} f_{x_1 x_1} & f_{x_1 x_2} & \dots & f_{x_1 x_n} \\ f_{x_2 x_1} & f_{x_2 x_2} & \dots & f_{x_2 x_n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_n x_1} & f_{x_n x_2} & \dots & f_{x_n x_n} \end{pmatrix}$$

If f is nice enough to allow change of order of partial derivatives, then $H(x_0)$ is real and symmetric, hence diagonalizable by Spectral Theorem. So we can choose the basis to be the eigenbasis (which can be chosen to be orthonormal), therefore

$$\underline{\delta x}H(x_0)\underline{\delta x}^\top = \sum_{i=1}^n \lambda_i(\delta x_i)^2$$

With this, we can classify the stationary points in 3 cases.

Case 1: $\forall \underline{\delta x} \in \mathbb{R}^n \setminus \{0\}, \underline{\delta x}H(x_0)\underline{\delta x}^\top > 0$. This happens iff $\lambda_i > 0$ for each i , that is, H is positive definite.

Case 2: $\forall \underline{\delta x} \in \mathbb{R}^n \setminus \{0\}, \underline{\delta x}H(x_0)\underline{\delta x}^\top < 0$. This happens iff $\lambda_i < 0$ for each i , that is, H is negative definite.

Case 3: Otherwise, H is indefinite.

Definition 6.4. The signature of H is the pattern of signs of its subdeterminant.

For example, for $f(x_1, x_2, \dots, x_n)$, the subdeterminants are

$$H_k = \begin{pmatrix} f_{x_1x_1} & f_{x_1x_2} & \cdots & f_{x_1x_k} \\ f_{x_2x_1} & f_{x_2x_2} & \cdots & f_{x_2x_k} \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_kx_1} & f_{x_kx_2} & \cdots & f_{x_kx_k} \end{pmatrix}$$

Then the sign is the signs of $|H_1|, |H_2|, \dots, |H_n|$.

Theorem 6.4. If H is positive (negative) definite, so is all of H_i .

Therefore a minimum (maximum) point of a real function in \mathbb{R}^n is also a minimum (maximum) point in any subspace of \mathbb{R}^n that includes the point.

Type of S.P.	Signature
Minimum	+, +, +, +, ...
Maximum	-, +, -, +, ...

Sometimes $|H| = 0$, in which case this stationary point is called degenerate, so we need to look at higher order terms in the Taylor series. This helps us in sketching the contour of a two dimensional function. Consider the coordinate system aligned with the (orthonormal) eigenbasis of H , so $\underline{\delta x} = \underline{x} - \underline{x}_0 = (\xi_1, \xi_2)$, where \underline{x}_0 is a stationary point. In a small region near x_0 , contour of f satisfies

$$\text{const} = f \approx f(\underline{x}_0) + \frac{1}{2}\underline{\delta x}H\underline{\delta x}^\top \implies \lambda_1\xi_1^2 + \lambda_2\xi_2^2 = \text{const}$$

Near min/max, λ_1, λ_2 have the same sign, so the contour looks like an ellipse. Near saddles, they have the different sign and the contour looks like a hyperbola.

Example 6.1. We want to find and classify the stationary points of

$$f(x, y) = 4x^3 - 12xy + y^2 + 10y + 6$$

and sketch its contour.

We have $f_x = 12x^2 - 12y, f_y = -12x + 2y + 10$, so at the stationary point, we

have $(x, y) = (1, 1), (5, 25)$. Now $f_{xx} = 24x, f_{xy} = f_{yx} = -12, f_{yy} = 2$. So at $(1, 1)$,

$$H = \begin{pmatrix} 24 & -12 \\ -12 & 2 \end{pmatrix} \implies |H_1| > 0, |H_2| < 0$$

So it is neither a maximum nor a minimum, hence it is a saddle point.

As for $(5, 25)$,

$$H = \begin{pmatrix} 120 & -12 \\ -12 & 2 \end{pmatrix} \implies |H_1| > 0, |H_2| > 0$$

so it is a minimum. We can then sketch the contour by drawing ellipses near $(5, 25)$ and hyperbolic curves near $(1, 1)$.

6.4 System of Linear ODEs

Consider a few dependent variables $y_1(t), y_2(t), \dots$ which satisfies system of coupled ODEs.

Example 6.2. Consider

$$\begin{cases} \dot{y}_1 = ay_1 + by_2 + f_1(t) \\ \dot{y}_2 = cy_1 + dy_2 + f_2(t) \end{cases}$$

So in vector form

$$\underline{\dot{y}} = M\underline{y} + \underline{f}$$

Any n^{th} order ODE can be written as a system of n first-order ODEs.

Example 6.3. Consider $\ddot{y} + ay + by' = f$, so let $y_1 = y, y_2 = y'$, then we have

$$\frac{d}{dt} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} 0 \\ f \end{pmatrix}$$

One way to solve it is to use matrix methods. Consider

$$\underline{\dot{y}} = M\underline{y} + \underline{f}$$

First, we find the general solution \underline{y}_c to

$$\underline{\dot{y}}_c - M\underline{y}_c = 0$$

Then we find a particular solution \underline{y}_p to the system, so by the superposition principle the general solution is $\underline{y}_p + \underline{y}_c$.

To find solution to the homogeneous equation, we try $\underline{y}_c = \underline{v}e^{\lambda t}$ which leads us to the conclusion that λ must be an eigenvalue of M with eigenvector \underline{v} .

Example 6.4. Consider

$$\underline{\dot{y}} - \begin{pmatrix} -4 & 24 \\ 1 & -2 \end{pmatrix} \underline{y} = \begin{pmatrix} 4 \\ 1 \end{pmatrix} e^t$$

We can find the eigenvalues and eigenvectors of M and find that $\lambda_1 = 2, \underline{v}_1 = (4, 1)^T, \lambda_2 = -8, \underline{v}_2 = (-6, 1)^T$ So we have

$$\underline{y}_c(t) = A \begin{pmatrix} 4 \\ 1 \end{pmatrix} e^{2t} + B \begin{pmatrix} -6 \\ 1 \end{pmatrix} e^{-8t}$$

We can sketch the phase portrait of y_2 against y_1 to find a hyperbola-like path of (y_1, y_2) .

To find the particular integral, we can try

$$\underline{y}_p = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} e^t$$

and find that $(u_1, u_2) = (-4, -1)$ solves the system, therefore the general solution is

$$\underline{y} = \begin{pmatrix} -4 \\ -1 \end{pmatrix} e^t + A \begin{pmatrix} 4 \\ 1 \end{pmatrix} e^{2t} + B \begin{pmatrix} -6 \\ 1 \end{pmatrix} e^{-8t}$$

where A, B are constants.

If the forcing matches the eigenvalue and eigenvector, we can try multiplying a polynomial in t (mostly just t^k will work).

From a linear system of n first-order ODEs, we can construct n uncoupled n^{th} order ODEs.

Example 6.5. Consider the same equation

$$\underline{\dot{y}} = \begin{pmatrix} -4 & 24 \\ 1 & -2 \end{pmatrix} \underline{y} + \begin{pmatrix} 4 \\ 1 \end{pmatrix} e^t$$

We can differentiate the first component to get $\ddot{y}_1 = -4y_1 + 24y_2 + 4e^t = -6y_1 + 16y_1 + 36e^t$, which we know how to solve. Similar way works for y_2 . One can check that this essentially gives the same solution.

We can discuss the concept of Phase Portrait in a more general way. For complementary function \underline{y}_c satisfying $\underline{\dot{y}}_c = M\underline{y}_c$, then $\underline{y}_c = \underline{v}_1 e^{\lambda_1 t} + \underline{v}_2 e^{\lambda_2 t}$
 Case 1: λ_1, λ_2 are real and $\lambda_1 \lambda_2 < 0$, WLOG $\lambda_1 > 0 > \lambda_2$, then the Phase Portrait is hyperbolic and converging towards the direction of \underline{v}_1 .

Case 2: λ_1, λ_2 are real and $\lambda_1 \lambda_2 > 0$. If $\lambda_1 < \lambda_2 < 0$, the phase portrait converge to 0. This is called the stable node. If $\lambda_1 > \lambda_2 > 0$, the phase portrait diverge from 0. this is called the unstable node.

Case 3: They are both complex.

Case 3(a): The real parts are both negative, then the amplitude will decrease in time, so it produces a spiral-like phase portrait converging to 0 (stable spiral).

Case 3(b): They are both positive, so the amplitude grows in time, so it produces a spiral-like curve spiraling out of 0, then it is an unstable spiral.

Case 3(c): They are both zero, then the amplitudes does not change and it is purely oscillating (i.e. circles centering at 0). We can find the direction of the oscillation by evaluating the equation at a given point and find the sign of \dot{y}_2 .

6.5 Nonlinear Systems

Consider a nonlinear autonomous system

$$\begin{cases} \dot{x} = f(x, y) \\ \dot{y} = g(x, y) \end{cases}$$

We want to find equilibrium (fixed) point, so for $\dot{x} = \dot{y} = 0$, we have $f(x_0, y_0) = 0 = g(x_0, y_0)$ We can do perturbation analysis by a small displacement $(x, y) = (x_0 + \xi(t), y_0 + \eta(t))$ around the fixed point, so

$$\dot{\xi} = f(x_0 + \xi, y_0 + \eta) = f(x_0, y_0) + \xi f_x(x_0, y_0) + \eta f_y(x_0, y_0) + \dots$$

$$\dot{\eta} = g(x_0 + \xi, y_0 + \eta) = g(x_0, y_0) + \xi g_x(x_0, y_0) + \eta g_y(x_0, y_0) + \dots$$

But f, g are 0 at (x_0, y_0) , so

$$\begin{pmatrix} \dot{\xi} \\ \dot{\eta} \end{pmatrix} \approx \begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix} \Big|_{(x_0, y_0)} \begin{pmatrix} \xi \\ \eta \end{pmatrix}$$

Example 6.6. Lotka-Volterra Model of predator and prey.

$$\begin{cases} \dot{x} = \alpha x - \beta xy = f(x, y) \\ \dot{y} = \delta xy - \gamma y = g(x, y) \end{cases}$$

where $\alpha, \beta, \gamma, \delta > 0$. The fixed point is $(0, 0)$ and $(\gamma/\delta, \alpha/\beta)$. Note that we have

$$\begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix} = \begin{pmatrix} \alpha - \beta y & -\beta x \\ \delta y & \delta x - \gamma \end{pmatrix}$$

At $(0, 0)$, we have

$$\begin{pmatrix} \dot{\xi} \\ \dot{\eta} \end{pmatrix} \approx \begin{pmatrix} \alpha & 0 \\ 0 & -\gamma \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}$$

So the eigenvalues are $\alpha, -\gamma$, thus it is a saddle point. For the other fixed point,

$$\begin{pmatrix} \dot{\xi} \\ \dot{\eta} \end{pmatrix} \approx \begin{pmatrix} 0 & -\beta\gamma \\ \frac{\alpha\delta}{\beta} & 0 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}$$

So it has a pair of purely imaginary eigenvalues, so it is a center. The direction of rotation is counterclockwise. Indeed, we have

$$\dot{\xi} = \frac{-\beta\gamma}{\delta} \eta < 0$$

for $\eta > 0$. We can sketch the solutions.

6.6 Partial Differential Equations

Definition 6.5. A PDE is a DE with partial derivatives.

Here, we will only consider three examples.

6.6.1 First Order Wave Equation

This is the PDE

$$\frac{\partial y}{\partial t} - c \frac{\partial y}{\partial x} = 0$$

where $y = y(x, t)$ and c is a constant. We can solve it with the method of characteristics. Imagine the contour of y in the $x - t$ plane, then we can start at some point and move it back and forth along a path. So along a path $x(t)$, if we have $y(x(t), t)$, then plugging it in our equation by the multivariate chain rule,

$$\frac{dy}{dt} = \frac{\partial y}{\partial t} + \frac{\partial y}{\partial x} \frac{dx}{dt}$$

So we can take the path where $\dot{x} = -c$ thus $\dot{y} = 0$ (so $x = x_0 - ct$), hence y will be constant along that line. These paths are called characteristics. If

$y(x, t = 0) = f(x)$, then $y = f(x_0)$ along the characteristic. Therefore the general solution is

$$y = f(x + ct)$$

for some differentiable f .

Example 6.7. With the initial condition $y(x, 0) = x^2 - 3$, since we have $y = f(x + ct)$, $y(x, t) = (x + ct)^2 - 3$.

If we add some forcing,

Example 6.8. Consider

$$\frac{\partial y}{\partial t} + 5 \frac{\partial y}{\partial x} = e^{-t}$$

with $y(x, 0) = e^{-x^2}$. So dy/dt along the path with $dx/dt = 5$, so $y = A - e^{-t}$ along a path. Note that A depends on $x(0) = x_0$, the initial point of the path. We know $y(x, 0) = A - 1 = e^{-x_0^2}$, so $A = 1 + e^{-x_0^2}$. Hence $y = 1 + e^{-(x-5t)^2} - e^{-t}$.

6.6.2 Second Order Wave Equation

We want to solve

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

So we can “factorize” it to have

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

Hence solutions can have $y_t \pm cy_x = 0$, so the solution is in the form $f(x + ct) + g(x - ct)$.

Example 6.9. Suppose we want to solve $y_{tt} - c^2 y_{xx} = 0$ and $y = 1/(1+x^2)$, $y_t = 0$ at $t = 0$ and $y \rightarrow 0$ as $x \rightarrow \infty$. So

$$\begin{cases} f(x) + g(x) = 1/(1+x^2) \\ cf'(x) - cg'(x) = 0 \implies f = g + A \end{cases}$$

for some constant A . So we can solve to get

$$g(x) = \frac{1}{2(1+x^2)} - \frac{A}{2}, f(x) = \frac{1}{2(1+x^2)} + \frac{A}{2}$$

So

$$y = \frac{1}{2(1+(x+ct)^2)} + \frac{1}{2(1+(x-ct)^2)}$$

We can sketch the solution to find that this gives two waves moving to the two ends.

6.6.3 Diffusion Equation

We consider

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

where κ is a constant. Typical cases where diffusion occurs are pollution transport, heat conduction and movement of microbes. Integrate the equation over \mathbb{R} to get

$$\frac{\partial}{\partial t} \int_{-\infty}^{\infty} y \, dx = \kappa [y_x]_{-\infty}^{\infty}$$

So if $y_x \rightarrow 0$ as $x \rightarrow \pm\infty$, the integral of y over \mathbb{R} will be constant. We can solve by the use of similarity variable.

Example 6.10. Consider $y_t = \kappa y_{xx}$ where $y(x, 0) = \delta(x)$ and $y \rightarrow 0$ when $x \rightarrow \pm\infty$. Define $\eta = x^2/(4\kappa t)$.⁶ We can try solutions of the form $y = t^{-\alpha} f(\eta)$ to get

$$-\frac{\alpha}{t} + f' \eta_t = \kappa f'' (\eta_x)^2 + \kappa f' \eta_{xx}$$

But note that $\eta_t = -\eta/t$, $\eta_x = \eta/(\kappa t)$, $\eta_{xx} = 2/(4\kappa t)$. All of the terms then have a factor of $1/t$, so we can remove the time dependence and get

$$\alpha f + f' \eta + f'' \eta + f'/2 = 0$$

Let $\alpha = 1/2$, we have $\eta F' + F/2 = 0$ where $F = f + f'$. If $F = 0$ (which is a solution) then $f(\eta) = Ae^{-\eta}$. Then $y = At^{-1/2} e^{-x^2/(4\kappa t)}$, then from the delta function condition we have $A = 1/\sqrt{4\pi\kappa}$, hence

$$y(x, t) = \frac{1}{\sqrt{4\pi\kappa}} t^{-1/2} e^{-x^2/(4\kappa t)}$$

is a solution.

⁶Obtained from dimensional analysis