Calculus and Applications - Part II

Vahid Shahrezaei

2021-03-24

Contents

Welcome

These are lecture notes for the second part of Calculus and Applications first year module at the Department of Mathematics, Imperial College London. The notes are split into three parts on Fourier Transform, Ordinary Differential Equations and Introduction to Multivariate Calculus. Please refer to course Blackboard for additional materials recommended text books for further reading.

These lecture notes are adobted from existing courses in our department. Part I of the course is based on the old M2AA2 course (Andrew Walton) and Part II and III are based on the old M1M2 course (Frank Berkshire, Mauricio Barahona, Andrew Parry). Some examples and ideas from the old mechanics course M1A1 is included as well.

These notes are produced with accessibility in mind and I hope it meets your requirements. You can experiment with the controls in the toolbar at the top of the html version of the notes. You can search for a word, adjust typeface, font size, font and background color. You can also download a copy of these notes in differnt formats, if you wish for offline use. I hope you enjoy this course and let me know if you have any comments or questions by email.

Part I: Fourier Transforms

Chapter 1

Fourier Transforms

Last term, we saw that *Fourier series* allows us to represent a given function, defined over a finite range of the independent variable, in terms of sine and cosine waves of different amplitudes and frequencies. *Fourier Transforms* are the natural extension of Fourier series for functions defined over $\mathbb{R}$. A key reason for studying Fourier transforms (and series) is that we can use these ideas to help us solve differential equations as seen in this course regarding ordinary differential equations and more extensively next year in relation to partial differential equations. There are also many other applications for Fourier transforms in science and engineering, particularly in the context of signal processing.

1.1 Fourier's integral formula

We can represent a function $f(x)$ defined over the interval $[-L, L]$ using the Fourier series

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}\{a_n \cos\left(\frac{n\pi x}{L}\right) + b_n \sin\left(\frac{n\pi x}{L}\right)\}.$$

where the corresponding Fourier coefficients are given by

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos\left(\frac{n\pi x}{L}\right)dx, \quad n = 0, 1, 2, \cdots,$$

$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin\left(\frac{n\pi x}{L}\right)dx, \quad n = 1, 2, \cdots.$$

Expressed in the exponential form the Fourier series can be represented as

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{in\pi x/L}, \quad |x| < L,$$

$$c_n = \frac{1}{2L}\int_{-L}^{L} f(x)e^{-in\pi x/L}\,dx, \quad n = 0, \pm 1, \pm 2, \cdots.$$

By defining *angular frequency* as $\omega_n = n\pi/L$ and *frequency difference* as

$$\delta\omega = \omega_{n+1} - \omega_n,$$

we can rewrite the Fourier series in the new notation as

$$f(x) = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty}\left[\int_{-L}^{L} f(s)e^{-i\omega_n s}ds\right]e^{-i\omega_n x}\delta\omega.$$

This result can be extended for a function $f(x)$ defined on $\mathbb{R}$ by taking the limit of $L \to \infty$ from the Fourier series. Using the angular frequency notation from above and replacing sum with integral using the Riemann sum, noting that $\delta\omega \to 0$ as $L \to \infty$, we obtain

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} f(s)e^{-i\omega s}\,ds\right\}e^{i\omega x}\,d\omega.$$

We therefore have shown that for a function $f(x)$ defined over $-\infty < x < \infty$ we have the following

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{f}(\omega)e^{i\omega x}\,d\omega,$$

where

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(x)e^{-i\omega x}\,dx.$$

The function $\hat{f}(\omega)$ (also denoted as $\mathcal{F}\{f(x)\}$) is known as Fourier transform of $f(x)$, which is analogous to the Fourier coefficients in a Fourier series. The relation above between $f(x)$ and $\hat{f}(\omega)$ is also known as inverse Fourier transform. Note that some books use slightly different definitions of Fourier transform with different normalisation. In order to evaluate the integrals above, a necessary condition is that $f(x)$ and its transform decay at $\pm\infty$. Using the Dirac delta function this restriction can be overcome as seen later.

Proof of Fourier's integral formula

In the previous section in a non-rigorous way we arrived at the result

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} f(s)e^{-i\omega s}\,ds\right\}e^{i\omega x}\,d\omega.$$

To prove this more formally, we need to assume that $f(x)$ is such that

$$\int_{-\infty}^{\infty} |f(x)|\, dx$$

converges. We will also assume that $f(x)$ and $f'(x)$ are continuous for all x (this can be relaxed as discussed at the end). We start by writing the RHS above in the form

$$\lim_{L\to\infty} \frac{1}{2\pi} \int_{-L}^{L} \left\{ \int_{-\infty}^{\infty} f(s) e^{-i\omega(s-x)}\, ds \right\} d\omega =$$
$$\lim_{L\to\infty} \frac{1}{2\pi} \int_{-L}^{L} \left\{ \int_{-\infty}^{\infty} f(s) \cos\left[\omega(s-x)\right] ds - i \int_{-\infty}^{\infty} f(s) \sin\left[\omega(s-x)\right] ds \right\} d\omega.$$

The first integral in curly brackets is even about $\omega = 0$, while the second is odd. Also, because of the absolute convergence of the inner integral, we can interchange the order of integration. Therefore, the expression simplifies to

$$\begin{aligned}
&\lim_{L\to\infty} \frac{1}{\pi} \int_{0}^{L} \left\{ \int_{-\infty}^{\infty} f(s) \cos\left[\omega(s-x)\right] ds \right\} d\omega \\
&= \lim_{L\to\infty} \frac{1}{\pi} \int_{-\infty}^{\infty} f(s) \left\{ \int_{0}^{L} \cos\left[\omega(s-x)\right] d\omega \right\} ds \\
&= \lim_{L\to\infty} \frac{1}{\pi} \int_{-\infty}^{\infty} f(s) \frac{\sin\left[L(s-x)\right]}{s-x}\, ds \\
&= \lim_{L\to\infty} \frac{1}{\pi} \int_{-\infty}^{\infty} f(x+u) \frac{\sin\left(Lu\right)}{u}\, du,
\end{aligned}$$

using the substitution $u = s - x$. We now split the integral into two parts in the following form

$$\lim_{L\to\infty} \frac{1}{\pi} \left\{ \int_{-\infty}^{\infty} \frac{f(x+u) - f(x)}{u} \sin\left(Lu\right) du + f(x) \int_{-\infty}^{\infty} \frac{\sin\left(Lu\right)}{u}\, du \right\}.$$

The first integral tends to zero as $L \to 0$ using Riemann-Lebesgue Lemma (seen last term in this course). We then use the substitution $p = Lu$ in the second integral to leave

$$\lim_{L\to\infty} \frac{f(x)}{\pi} \int_{-\infty}^{\infty} \frac{\sin\left(p\right)}{p}\, dp = f(x),$$

using the fact that $\int_{-\infty}^{\infty} (\sin p)/p\, dp = \pi$ (seen last term and also in the problem sheet 1 this term).

We have therefore proved Fourier's integral formula. As remarked earlier, we have assumed here that $f(x)$ is continuous at all x. If there is a discontinuity at

x_0 (with finite left and right hand derivatives there), the LHS of the formula is replaced by $[f(x_0+) + f(x_0-)]/2$ (analogous to the Fourier series convergence we investigated earlier).

Example 1.1. Find the Fourier transform of the rectangular wave

$$f(x) = \begin{cases} 1, & \text{if } |x| < d, \\ 0, & \text{if } |x| > d. \end{cases}$$

Using the Fourier transform formula we have

$$\hat{f}(\omega) = \int_{-d}^{d} 1.e^{-iwx}\,dx = \left[\frac{e^{-i\omega x}}{-i\omega}\right]_{-d}^{d} = -\frac{1}{i\omega}(e^{-i\omega d} - e^{i\omega d}) = \frac{2}{\omega}\sin\omega d.$$

See Figure 1.1 for a graph of $\hat{f}(\omega)$ for different values of d. Note that as d gets larger, $\hat{f}$ becomes more concentrated in the vicinity of $\omega = 0$. This is a general property of Fourier transforms and its inverse and relates to uncertainty principle. A function which is more localised around zero has a wider inverse Fourier transform. See unseen question 1 for the derivation and further discussion of the uncertainty principle.

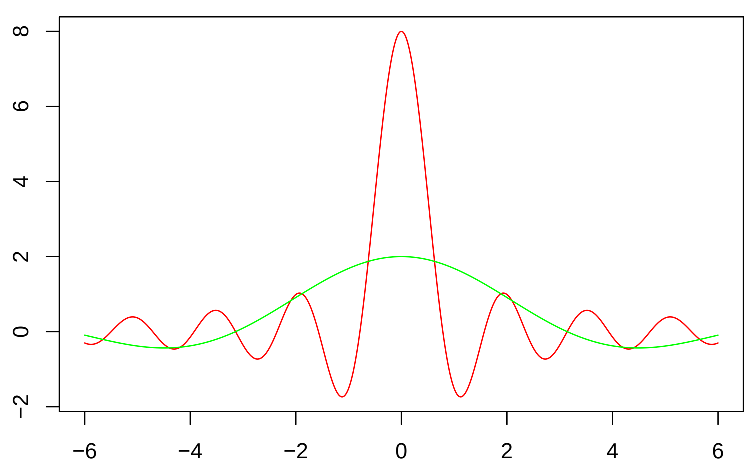

Figure 1.1: Graph of the Fourier transform for $d = 1$ (green) and $d = 4$ (red)

1.2 Fourier cosine and sine transforms

We can exploit the symmetry to define transforms over the range $[0, \infty)$. First, if we suppose that $f(x)$ is even about $x = 0$, we have

$$\begin{aligned}\hat{f}(\omega) &= \int_{-\infty}^{\infty} f(x)e^{-i\omega x}\,dx = \int_{-\infty}^{\infty} f(x)(\cos\omega x - i\sin\omega x)\,dx \\ &= 2\int_{0}^{\infty} f(x)\cos\omega x\,dx.\end{aligned}$$

We define *Fourier cosine transform* of $f(x)$ to be

$$\hat{f}_c(\omega) = \int_0^{\infty} f(x)\cos\omega x\,dx.$$

Thus, for an even function $f(x)$ we have $\hat{f}(\omega) = 2\hat{f}_c(\omega)$.

Using the inversion formula for the regular transform and exploiting the evenness of $\hat{f}_c(\omega)$, we can obtain the inversion formula for the Fourier cosine transform:

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\omega)e^{i\omega x}\ d\omega = \frac{2}{\pi}\int_0^{\infty} \hat{f}_c(\omega)\cos\omega x\,d\omega.$$

In a similar way, by considering $f(x)$ to be odd about $x = 0$, we can define a *Fourier sine transform* and derive the corresponding inversion formula. We obtain the pair of expressions:

$$\begin{aligned}\hat{f}_s(\omega) &= \int_0^{\infty} f(x)\sin\omega x\,dx \\ f(x) &= \frac{2}{\pi}\int_0^{\infty} \hat{f}_s(\omega)\sin\omega x\,dw.\end{aligned}$$

For an odd function $f(x)$, we have $\hat{f}(\omega) = -2i\hat{f}_s(\omega)$.

Example 1.2. Find the Fourier cosine transform of the rectangular wave

$$f(x) = \begin{cases} 1, & \text{if } |x| < d, \\ 0, & \text{if } |x| > d. \end{cases}$$

We can use the definition of Fourier cosine transform directly noting that the function is even. But also, as we have already obtained the Fourier transform of this function in the last example, we simply have:

$$\hat{f}_c(\omega) = \frac{1}{2}\hat{f}(\omega) = \frac{1}{\omega}\sin\omega d.$$

Chapter 2

Properties of Fourier Transforms

In the following we present some important properties of Fourier transforms. These results will be helpful in deriving Fourier and inverse Fourier transform of different functions. After discussing some basic properties, we will discuss, convolution theorem and energy theorem. Finally, we introduc Dirac delta function.

2.1 Basic Properties

- (**i**) The Fourier and inverse Fourier transforms are linear, and so

$$\mathcal{F}\{af(x)+bg(x)\} = a\hat{f}(\omega)+b\hat{g}(\omega),$$

$$\mathcal{F}^{-1}\{a\hat{f}(\omega)+b\hat{g}(\omega)\} = af(x)+bg(x),$$

where a and b are constants and $\mathcal{F}^{-1}$ denotes the inverse Fourier transform.

-(**ii**) If $a > 0$:

$$\mathcal{F}\{f(ax)\} = \frac{1}{a}\hat{f}(\frac{\omega}{a}).$$

Proof Starting on the LHS, and making the substitution $s = ax$:

$$\mathcal{F}\{f(ax)\} = \int_{-\infty}^{\infty} f(ax)e^{-i\omega x}\,dx = \frac{1}{a}\int_{-\infty}^{\infty} f(s)e^{-i(\omega/a)s}\,ds = \frac{1}{a}\hat{f}(\frac{\omega}{a}).$$

-(**iii**) In a similar way we can establish that

$$\mathcal{F}\{f(-x)\} = \hat{f}(-\omega).$$

-(**iv**) The transform of a shifted function can be calculated as follows (using $s = x - x_0$):

$$\mathcal{F}\{f(x-x_0)\} = \int_{-\infty}^{\infty} f(x-x_0)e^{-i\omega x}\,dx = \int_{-\infty}^{\infty} f(s)e^{-i\omega(s+x_0)}\,ds = e^{-i\omega x_0}\hat{f}(\omega).$$

-(**v**) A similar result, but this time involving a shift in transform space:

$$\mathcal{F}\{e^{i\omega_0 x}f(x)\} = \int_{-\infty}^{\infty} f(x)e^{-i(\omega-\omega_0)x}\,dx = \hat{f}(\omega-\omega_0).$$

-(**vi**) Symmetry formula The following result is very useful. Suppose the Fourier transform of $f(x)$ is $\hat{f}(\omega)$; change the variable ω to x; then

$$\mathcal{F}\{\hat{f}(x)\} = 2\pi f(-\omega).$$

Proof Starting with the inversion formula and changing variables from ω to s, we have

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\omega)e^{i\omega x}\,d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(s)e^{isx}\,ds.$$

If we now let $x = -\omega$ and then $s = x$, we get:

$$f(-\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(s)e^{-i\omega s}\,ds = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(x)e^{-i\omega x}\,dx = \frac{1}{2\pi}\;\mathcal{F}\{\hat{f}(x)\},$$

as required.

The following results are particularly useful when applying Fourier transforms to differential equations (as seen later this term and next year in the context of partial differential equations).

-(**vii**)

$$\mathcal{F}\{\frac{d^n f}{dx^n}\} = (i\omega)^n\hat{f}(\omega).$$

Proof This can be established by integration by parts. We assume that all derivatives of f tend to zero as $x \to \pm\infty$.

$$\begin{aligned}
\mathcal{F}\{d^n f/dx^n\} &= \int_{-\infty}^{\infty} (d^n f/dx^n)e^{-i\omega x}\,dx \\
&= \left[(d^{n-1}f/dx^{n-1})e^{-i\omega x}\right]_{-\infty}^{\infty} + i\omega\int_{-\infty}^{\infty} (d^{n-1}f/dx^{n-1})e^{-i\omega x}\,dx \\
&= i\omega\mathcal{F}\{d^{n-1}f/dx^{n-1}\} \\
&= \cdots \\
&= (i\omega)^n\hat{f}(\omega).
\end{aligned}$$

-(**viii**)

$$\mathcal{F}\{xf(x)\} = i\hat{f}'(\omega).$$

Proof Considering the LHS:

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)xe^{-i\omega x}\,dx &= \int_{-\infty}^{\infty} f(x)\frac{d}{d\omega}(ie^{-i\omega x})\,dx \\ &= i\frac{d}{d\omega}\int_{-\infty}^{\infty} f(x)e^{-i\omega x}\,dx \\ &= i\frac{d}{d\omega}\hat{f}(\omega).\end{aligned}$$

-(**ix**)

$$\begin{aligned}&\text{(a) } \mathcal{F}_c\{f'(x)\} = -f(0) + \omega\hat{f}_s(\omega), \\ &\text{(b) } \mathcal{F}_s\{f'(x)\} = -\omega\hat{f}_c(\omega), \\ &\text{(c) } \mathcal{F}_c\{f''(x)\} = -f'(0) - \omega^2\hat{f}_c(\omega), \\ &\text{(d) } \mathcal{F}_s\{f''(x)\} = \omega f(0) - \omega^2\hat{f}_s(\omega).\end{aligned}$$

Proof We prove (a) and (c) and leave the others as exercises. For (a) we have, integrating by parts:

$$\begin{aligned}\mathcal{F}_c\{f'(x)\} &= \int_0^{\infty} f'(x)\cos\omega x\,dx \\ &= [f(x)\cos\omega x]_0^{\infty} + \omega\int_0^{\infty} f(x)\sin\omega x\,dx \\ &= -f(0) + \omega\hat{f}_s(\omega).\end{aligned}$$

And we prove (c) with the use of (b):

$$\begin{aligned}\mathcal{F}_c\{f''(x)\} &= \int_0^{\infty} f''(x)\cos\omega x\,dx \\ &= [f'(x)\cos\omega x]_0^{\infty} + \omega\int_0^{\infty} f'(x)\sin\omega x\,dx \\ &= -f'(0) + \omega\mathcal{F}_s\{f'(x)\} \\ &= -f'(0) - \omega^2\hat{f}_c(\omega).\end{aligned}$$

-(**x**) If $f(x)$ is a complex-valued function and $[f(x)]^*$ is its complex conjugate, then

$$\mathcal{F}\{[f(x)]^*\} = [\hat{f}(-\omega)]^*.$$

Proof We have that

$$\hat{f}(-\omega) = \int_{-\infty}^{\infty} f(x)e^{i\omega x}\,dx$$

and so by taking complex conjugate from both sides, it follows that

$$[\hat{f}(-\omega)]^* = \int_{-\infty}^{\infty} [f(X)]^* e^{-i\omega x}\,dx = \mathcal{F}\{[f(x)]^*\}.$$

2.2 Convolution theorem for Fourier transforms

We define the convolution of two functions $f(x)$ and $g(x)$, defined over $(-\infty, \infty)$, as

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(x-u)g(u)\,du.$$

An important result is the so-called convolution theorem:

Theorem 2.1 (Convolution theorem)**.** *Suppose $f(x)$ and $g(x)$ are two functions defined over $\mathbb{R}$ with Fourier transforms given as $\hat{f}(\omega)$ and $\hat{g}(\omega)$, we have:*

$$\mathcal{F}\{f * g\} = \hat{f}(\omega)\hat{g}(\omega).$$

Proof We start on the LHS, change the order of integration and then use the substitution $s = x - u$ at fixed u:

$$\begin{aligned}
&\int_{x=-\infty}^{\infty} \left\{ \int_{u=-\infty}^{\infty} f(x-u)g(u)\,du \right\} e^{-i\omega x}\,dx \\
&= \int_{u=-\infty}^{\infty} g(u) \left\{ \int_{x=-\infty}^{\infty} f(x-u)e^{-i\omega x}\,dx \right\} du \\
&= \int_{u=-\infty}^{\infty} g(u) \left\{ \int_{s=-\infty}^{\infty} f(s)e^{-i\omega(s+u)}\,ds \right\} du \\
&= \left(\int_{-\infty}^{\infty} g(u)e^{-i\omega u}\,du \right) \left(\int_{-\infty}^{\infty} f(s)e^{-i\omega s}\,ds \right) = \hat{g}(\omega)\hat{f}(\omega),
\end{aligned}$$

as required.

The convolution theorem suggests that convolution is commutative. This can also be shown easily from the definition by using a change of variable in the integration.

A similar convolution theorem holds for the inverse functions.

$$\mathcal{F}\{f(x)g(x)\} = \frac{1}{2\pi}\hat{f}(\omega) * \hat{g}(\omega).$$

The proof of this result, using Dirac delta function is discussed as a quiz in the lectures and using symmetry formula is seen in the problem sheet.

Example 2.1. Find the inverse Fourier transform of the function

$$\frac{1}{(4+\omega^2)(9+\omega^2)}.$$

By setting

$$\hat{f}(\omega) = 1/(4+\omega^2),\ \hat{g}(\omega) = 1/(9+\omega^2),$$

we have (from the quiz in the lectures) that $\mathcal{F}\{e^{-a|x|}\} = \frac{2a}{a^2+\omega^2}$ for $a > 0$. Therefore

$$f(x) = (1/4)e^{-2|x|}, \quad g(x) = (1/6)e^{-3|x|}.$$

Thus, by the convolution theorem:

$$\begin{aligned}
\mathcal{F}^{-1}\left\{\frac{1}{(4+\omega^2)(9+\omega^2)}\right\} &= f(x) * g(x) \\
&= \frac{1}{24}\int_{-\infty}^{\infty} e^{-2|x-u|}e^{-3|u|}\,du \\
&= \cdots \\
&= \frac{1}{20}e^{-2|x|} - \frac{1}{30}e^{-3|x|}.
\end{aligned}$$

Note that there are other ways to compute the inverse, for example, we could decompose the original function into partial fractions and invert term-by-term.

2.3 Energy theorem for Fourier transforms

This is the analogous result to Parseval's theorem for Fourier series.

Theorem 2.2 (Energy theorem)**.** *Suppose $f(x)$ s real valued function defined over $\mathbb{R}$ with Fourier transform given as $\hat{f}(\omega)$, we have:*

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\left|\hat{f}(\omega)\right|^2\,d\omega = \int_{-\infty}^{\infty}[f(x)]^2\,dx.$$

Proof Properties (iii) and (x) of the Fourier transforms give

$$\mathcal{F}\{[f(-x)]^*\} = [\hat{f}(\omega)]^*.$$

Since we are assuming f to be real, this simplifies to

$$\mathcal{F}\{f(-x)\} = [\hat{f}(\omega)]^*.$$

If we now use the convolution theorem with $\hat{g}(\omega) = [\hat{f}(\omega)]^*$, we have

$$\mathcal{F}\{f(x) * f(-x)\} = \hat{f}(\omega)[\hat{f}(\omega)]^* = \left|\hat{f}(\omega)\right|^2.$$

Using the definition of convolution and the inverse transform we have

$$f(x) * f(-x) = \int_{-\infty}^{\infty} f(u+x)f(u)du = \frac{1}{2\pi}\int_{-\infty}^{\infty} \left|\hat{f}(\omega)\right|^2 e^{i\omega x}\, d\omega.$$

In particular, setting $x = 0$, we obtain the required result:

$$\int_{-\infty}^{\infty} [f(u)]^2 du = \frac{1}{2\pi}\int_{-\infty}^{\infty} \left|\hat{f}(\omega)\right|^2\, d\omega.$$

2.4 The Dirac delta-function

Before we define the Dirac delta-function, we need to be aware of the following theorem.

Theorem 2.3 (Mean-value theorem for integrals)**.** *If $g(x)$ is continuous on $[a,b]$ then*

$$\int_a^b g(x)\, dx = (b-a)g(\bar{x}),$$

for at least one $\bar{x}$ with $a \leq \bar{x} \leq b$.

The proof follows from the regular mean-value theorem for G say, by defining $g = G'$. Geometrically this means that the area under the curve is equivalent to that of a rectangle with length equal to the interval of integration.

Definition of the Dirac delta-function (impulse function)

Consider the following step-function:

$$f_k(x) = \begin{cases} k/2, & \text{if } |x| < 1/k, \\ 0, & \text{if } |x| > 1/k. \end{cases}$$

Clearly we can see that an important property of this function is that

$$\int_{-\infty}^{\infty} f_k(x)dx = 1.$$

As k increases, $f_k(x)$ gets taller and thinner (see Figure 2.1). We define the Dirac delta function to be

$$\delta(x) = \lim_{k\to\infty} f_k(x),$$

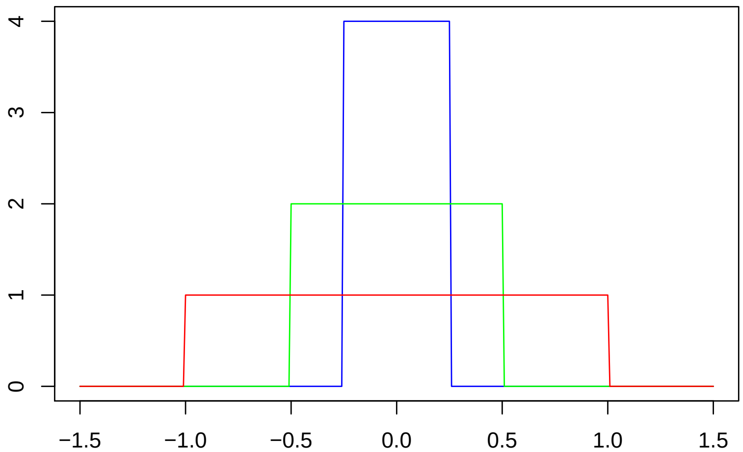

Figure 2.1: Graph of $f_k(x)$ for $k = 1$ (green), $k = 2$ (red) and $k = 4$ (blue).

although, of course, this limit doesn't exist in the usual mathematical sense. Effectively $\delta(x)$ is infinite at $x = 0$ and zero at all other values of x. The key property however, is that its integral (area under the curve) is one.

Sifting property of the delta function The delta function is most useful in how it interacts with other functions. Consider

$$\int_{-\infty}^{\infty} g(x)\delta(x)\,dx,$$

where $g(x)$ is a continuous function defined over $(-\infty, \infty)$. Using our definition of the delta-function we can rewrite this as

$$\begin{aligned}\lim_{k\to\infty}\int_{-\infty}^{\infty} g(x)f_k(x)\,dx &= \lim_{k\to\infty}\int_{-1/k}^{1/k}\frac{k}{2}g(x)\,dx \\ &= \lim_{k\to\infty}\frac{k}{2}g(\bar{x})\left(\frac{1}{k}-\left(-\frac{1}{k}\right)\right),\end{aligned}$$

for some $\bar{x}$ in $[-1/k, 1/k]$, using the mean-value theorem for integrals. Clearly, as $k \to \infty$, we must have $\bar{x} \to 0$. The expression above simplifies to

$$g(0)\frac{k}{2}\frac{2}{k} = g(0).$$

We have therefore established that for any continuous function g:

$$\int_{-\infty}^{\infty} g(x)\delta(x)\,dx = g(0).$$

This result can easily be generalized to

$$\int_{-\infty}^{\infty} g(x)\delta(x-a)\,dx = g(a).$$

Example 2.2. Find the Fourier transform of $\delta(x)$.

We have using the sifting property

$$\mathcal{F}\{\delta(x)\} = \int_{-\infty}^{\infty} \delta(x)e^{-i\omega x}\,dx = e^{-i\omega 0} = 1.$$

From this we can deduce that the inverse Fourier transform of 1 is $\delta(x)$. From this last result, and using the inversion formula, we see that an alternative representation of the delta function is

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\pm i\omega x}\,d\omega,$$

with the $\pm$ arising from the observation that $\delta(x)$ is an even function of x about $x = 0$. If we are prepared to work in terms of delta-functions, we can now take the Fourier transforms of functions that do not decay as $x \to \pm\infty$.

Example 2.3. Find the Fourier transform of $\cos \omega_0 x$.

Using the definition of $\cos \omega_0 x$ in terms of exponentials we have:

$$\begin{aligned}\mathcal{F}\{\cos \omega_0 x\} &= \int_{-\infty}^{\infty} \frac{1}{2}(e^{i\omega_0 x} + e^{-i\omega_0 x})e^{-i\omega x}\,dx \\ &= \frac{1}{2}\int_{-\infty}^{\infty} e^{-i(\omega-\omega_0)x}\,dx + \frac{1}{2}\int_{-\infty}^{\infty} e^{-i(\omega+\omega_0)x}\,dx \\ &= \pi\delta(\omega-\omega_0) + \pi\delta(\omega+\omega_0),\end{aligned}$$

which is a two-spiked 'function'.

We finish by recommending this video on a very intuitive visual introduction to Fourier transform from the popular 3Blue1Brown YouTube channel in mathematics education. Do check it out and also the additional videos on related topics such as uncertainty principle.

Part II: Ordinary Differential Equations

Chapter 3

Introduction to ordinary differential equations

Differential equations are very important in science and engineering. In this course, we focus on a specific class of differential equations called ordinary differential equations (ODEs). Ordinary refers to dealing with functions of one independent variable. We initially focus on scaler functions of the form $f : \mathbb{R} \to \mathbb{R}$. Later in the course, we discuss systems of ODEs and consider vector functions of the form $f : \mathbb{R} \to \mathbb{R}^n$. We assume $f(x)$ is differentiable upto order k. An ODE is an equation for the function $f(x)$ that involves the function, its derivatives and the independent variable. A general form for an ODE of *order k* is

$$G(x, f(x), \frac{df}{dx}, ..., \frac{d^k f}{dx^k}) = 0,$$

where the highest derivative present in the equation G is of the order k. The *degree* of the ODE is the power of highest derivative (when fractional powers have been removed). The ODE is called *linear* if G is a linear function of $f(x)$ and its derivatives. This form for the ODE is the so called *implicit* form. In an *explicit* form for an ODE, the highest order derivative is given as function of the lower derivatives:

$$\frac{d^k f}{dx^k} = F(x, f(x), \frac{df}{dx}, ..., \frac{d^{k-1} f}{dx^{k-1}}) = 0.$$

Example 3.1. Consider the following ODE for the function $f(x)$:

$$\frac{d^2 f}{dx^2} = 5 \left[1 + \left(\frac{df}{dx} \right)^2 \right]^{\frac{1}{3}}$$

This is a nonlinear explicit ODE of degree 3 and order 2.

Solving an ODE is the task of finding $f(x)$ such that the ODE is satisfied over the domain of x (*e.g.* $\mathbb{R}$).

ODEs appear naturally in many areas of sciences and humanities. In the following we provide some examples.

Example 3.2 (Second Newton Law)**.**

Mechanics: A very short introduction

Kinematics is a branch of mechanics that describes the motion of points (objects) without considering the forces that cause them to move. In one dimension $x(t)$ denotes the position of a particle at time t. Then $\frac{dx}{dt} = \dot{x} = v$ is defined as the velocity of the particle and $\frac{d^2x}{dt^2} = \ddot{x} = a$ is defined as acceleration. This can be generalised to higher dimensions using vectors of location, velocity and acceleration.

Dynamics is the branch of mechanics concerned with the study of *forces* and their effects on motion. *Isaac Newton* came up with the fundamental physical laws, which govern dynamics in physics:

-**First law** an object not acted upon by any force either remains at rest or continues to move at a constant velocity

-**Second law** the vector sum of the forces F on an object is equal to the mass m of that object multiplied by its acceleration a: $F = ma$

$$m\frac{d^2x}{dt^2} = F(t, x, \frac{dx}{dt})$$

This is a second order ODE for the position of the object $x(t)$.

-**Third law** when one body exerts a force on a second body, the second body exerts a force equal in magnitude and opposite in direction on the first body.

Mechanics used to be thaught until recently in our first year Mathematics course as it provides many links to different areas of mathemaitcs. If you have any doubts, look at this video, for a very cool counting problem for colliding particles and its very unexpected solution and link to mathematics.

Example 3.3 (Population dynamics: Malthus (1798))**.**

Consider $P(t)$ denotes the population of certain species at time t. Malthus proposed the following simple ODE:

$$\frac{dP}{dt} = kP,$$

with $k > 0$, this ODE results in an exponential increase in the population in time.

Example 3.4 (Population dynamics: Logistic Growth (Verhulst, 1845))**.**

Verhgulst proposed a modification to Malthus law, creating a carrying capacity C for the population:

$$\frac{dP}{dt} = kP\left(1 - \frac{P}{C}\right).$$

Example 3.5 (Radius of curvature)**.**

In geometry, given radius of curvature $R(x, y)$, we can find equation for the curve $y(x)$ using the following ODE that definition for the radius of curvature.

$$R(x, y) = \frac{\left(1 + (\frac{dy}{dx})^2\right)^{3/2}}{\frac{d^2y}{dx^2}}.$$

3.1 Particular and General Solutions

$f_{PI}(x)$ is called a *Particular Integral* or *Particular Solution* of an ODE such that

$$G(x, f(x), \frac{df}{dx}, ..., \frac{d^k f}{dx^k}) = 0,$$

is satisfied over the domain $x \in \mathbb{R}$.

f_{GS} is called a *General Solution* of an ODE of the order k, if

$$f_{GS} = f_{GS}(x; c_1, c_2, \cdots, c_k)$$

is a general family of solutions that fulfil the ODE. The parameters $\{c_i\}_{i=1}^k$ are the constants of integration and are usually fixed by initial or boundary conditions. In this course we, concern ourselves with methods that allow us to obtain such solutions. Rigourous mathematical results on existance and uniqueness of such solutions are discussed in the second year.

Example 3.6 (from kinematics)**.** Object moving with constant speed v:

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

One particular solution is $x_{PI} = vt$ and another one is $x_{PI} = vt+1$. The general solution is $x_{GS} = vt + c_1$, where c_1 is the constant of integration.

If we are also told that $x(t = 0) = x_0$, we have $x(t) = vt + x_0$, which is the solution to the initial value problem.

Chapter 4

First and second order ODEs

Not all ODEs are analytically solvable. In this section we discuss some types of first and second order ODEs that are analytically solvable and see some examples.

4.1 First order ODEs

A first order ODE has only the first derivative represented. The general implicit form for a first order ODE for the function $x(t)$ is:

$$G(t, x, \frac{dx}{dt}) = 0,$$

and its explicit form is:

$$\frac{dx}{dt} = F(x, t).$$

In the following we discuss some classes of first order ODEs and describe methods of obtaining a solution for them. These inculde separable and linear first order ODEs. The other types of First Order ODEs that can be solved are based on transformations or change of variables. We see two examples of this in the following. Another important class of first order ODEs that can be solved are Exact ODEs that will be discussed in part III of the course after introducing partial and total differentiation.

4.1.1 Separable First Order ODEs

A separable first order ODE can be written in the following form:

$$\frac{dx}{dt} = F_1(x)F_2(t).$$

Solution Rearranging and integrating both sides we get:

$$\int \frac{dx}{F_1(x)} = \int F_2(t)dt + c_1.$$

4.1.2 Linear First Order ODEs

First order linear ODEs have the following general form:

$$\frac{dy}{dx} + p(x)y = q(x).$$

Solution This is solved by finding an integrating factor (IF). We look for $I(x)$ such that:

$$I(x)\left[\frac{dy}{dx} + p(x)y\right] = \frac{d[I(x)y]}{dx},$$

Then, we have

$$\begin{aligned}
\frac{d[I(x)y]}{dx} &= I(x)q(x), \\
\int d[I(x)y] &= \int q(x)I(x)\,dx + c_1, \\
y(x) &= \frac{1}{I(x)}\left[\int q(x)I(x)\,dx + c_1\right].
\end{aligned}$$

Integrating factors must fulfil:

$$\begin{aligned}
\frac{d(Iy)}{dx} &= I\frac{dy}{dx} + Ipy, \\
I\frac{dy}{dx} + y\frac{dI}{dx} &= I\frac{dy}{dx} + Ipy, \\
\int \frac{dI}{I} &= \int p(x)\,dx + c'.
\end{aligned}$$

So we have:

$$I(x) = Ae^{\int p(x)\,dx},$$

where A is a new arbitrary constant (of integration).

So, we have the following for the general solution:

$$y(x) = e^{-\int p(x)\,dx}\left[\int e^{\int p(x)\,dx} q(x)\,dx + c\right],$$

where $c = c_1/A$ is a new arbitrary constant of integration.

4.1.3 Dimensionally Homogeneous

The dimensionally homogeneous have the following general form:

$$\frac{dy}{dx} = F\left(\frac{y}{x}\right).$$

Solution Let $u = y/x$ we obtain:

$$\frac{dy}{dx} = u + x\frac{du}{dx}$$

The ODE in terms of $u(x)$, which is separabale is

$$u + x\frac{du}{dx} = F(u),$$

Finding general solution $u_{GS}(x)$ for this ODE then we find the general solution for the original ODE as $y_{GS}(x) = u_{GS}(x)x$.

4.1.4 Bernoulli ODEs

There are other examples of transformations can turn specific ODEs into separable or linear. Some such as Bernoulli are classic:

$$\frac{dy}{dx} + p(x)y = q(x)y^n,$$

where $n \in \mathbb{R}$.

Solution We use the change of variable $u = y^{1-n}$. We obtain:

$$\frac{du}{dx} = (1-n)y^{-n}\frac{dy}{dx}.$$

Writing the original ODE in terms of u we have:

$$\frac{du}{dx} + (1-n)p(x)u = (1-n)q(x),$$

which is a linear ODE for $u(x)$, so we obtain $u_{GS}(x)$ and then we have

$$y_{GS} = u_{GS}^{\frac{1}{1-n}}.$$

4.2 Second Order ODEs

The general implicit form is:

$$G(x, y, \frac{dy}{dx}, \frac{d^2y}{dx^2}) = 0,$$

and the general explicit form is:

$$\frac{d^2y}{dx^2} = F(x, y, \frac{dy}{dx}).$$

The second order ODEs are common in Mechanics as Newton's second law is such ODE with independent variable as time t. They are difficult to solve for general F but there are some special cases that can be solved as described in the following. Also, the linear case is discussed in the next chapter in detail.

4.2.1 F only depends on x

$$\frac{d^2y}{dx^2} = F(x)$$

Solution Let $u = \frac{dy}{dx}$ then we have $\frac{du}{dx} = F(x)$. A first integration gives us:

$$u = \int F(x)dx + c_1$$

A second integration then gives us y_{GS}:

$$y_{GS} = \int \left[\int F(x)dx\right] dx + c_1 x + c_2.$$

4.2.2 F only depends on x and $\frac{dy}{dx}$

$$\frac{d^2y}{dx^2} = F(x, \frac{dy}{dx})$$

Solution Let $u = \frac{dy}{dx}$ then we have $\frac{du}{dx} = F(x, u)$, which is a first order ODE. If we could obtain the general solution $u_{GS}(x; c_1)$ then we have:

$$y_{GS}(x) = \int u_{GS}(x; c_1)\, dx + c_2.$$

4.2.3 F only depends on y

$$\frac{d^2y}{dx^2} = F(y)$$

Solution We let $u = \frac{dy}{dx}$ then $\frac{du}{dx} = F(y)$. Then we have:

$$\frac{du}{dx} = \frac{du}{dy}\frac{dy}{dx} = u\frac{du}{dy} = \frac{d}{dy}\left(\frac{1}{2}u^2\right) = F(y),$$

which is a first order separable ODE for $u(y)$. We have:

$$\frac{1}{2}u^2 = \int F(y)dy + c_1 = G(y) + c_1.$$

So we have:

$$\frac{dy}{dx} = u = \pm\sqrt{2G(y) + 2c_1},$$

which is a first order separable ODE for $y(x)$ and can be integrated to obtain $y_{GS}(x; c_1, c_2)$ as seen in the following example.

Example 4.1 (Mechanics Harmonic Oscillator)**.** Hooke's law states if $x(t)$ is displacement relative to an ideal spring relaxed position, the spring force is: $F = -kx$ Using second Newton Law we have: $ma = F \quad \Longrightarrow \quad m\frac{d^2x}{dt^2} = -kx$

Let velocity to be $u = \frac{dx}{dt}$, then we have:

$$a = \frac{du}{dt} = \frac{d}{dx}\left[\frac{1}{2}u^2\right] = -\frac{kx}{m}.$$

Integrating both sides we obtain:

$$\frac{u^2}{2} = -\frac{k}{2m}x^2 + c_1.$$

This equation gives us a constant of motion ($E = c_1 m$), which is known as total energy, the sum of *kinetic energy* ($1/2mu^2$) and *potential energy* ($1/2kx^2$).

$$u = \frac{dx}{dt} = \pm\sqrt{\frac{2E-kx^2}{m}} \quad \Longrightarrow \quad \int \frac{dx}{\pm\sqrt{\frac{2E-kx^2}{m}}} = \int dt$$

Sticking with the postive sign on the LHS we have:

$$\frac{1}{\sqrt{2E/m}}\int \frac{dx}{\sqrt{1-\frac{k}{2E}x^2}} = \sqrt{\frac{m}{k}}\sin^{-1}\left(\sqrt{\frac{k}{2E}}x\right) = t + c_2$$

Rearranging the solution we obtain:

$$x_{GS} = A\sin(\omega t + \phi),$$

where, $\omega = \sqrt{k/m}$ is the frequency of oscillations and $A = \sqrt{2E/k}$ and $\phi = \sqrt{k/m}c_2$ are new constants of integration. We note that, if we had chosen to use the minus sign above, we would have obtained the same family of solutions but the constants of integrations would be differently defined.

4.2.4 F only depends on y and $\frac{dy}{dx}$

$$\frac{d^2y}{dx^2} = F(y, \frac{dy}{dx})$$

let $u = \frac{dy}{dx} \quad \Longrightarrow \quad \frac{du}{dx} = F(y, u)$. So we have

$$\frac{du}{dx} = \frac{du}{dy}\frac{dy}{dx} = u\frac{du}{dy} = \frac{d}{dy}\left(\frac{1}{2}u^2\right).$$

Therefore we have the following first order ODE for $u(y)$ to solve

$$\frac{d}{dy}\left(\frac{1}{2}u^2\right) = F(y,u).$$

Given $u_{GS}(y;c_1)$ being a general solution for the above ODE, we have the following first order ODE for $y(x)$:

$$\frac{dy}{dx} = u_{GS}(y;c_1).$$

Chapter 5

Linear ODEs

The general form of linear ODEs of order k is

$$\alpha_k(x)\frac{d^k y}{dx^k} + \alpha_{k-1}(x)\frac{d^{k-1}y}{dx^{k-1}} + ... + \alpha_1(x)\frac{dy}{dx} + \alpha_0(x)y = f(x),$$

where $\alpha_k(x)$,.., $\alpha_0(x)$ and $f(x)$ are functions of only the independent variable x. The ODE is called *homogeneous* if $f(x) = 0$ and *inhomogeneous* otherwise.

Some examples of linear ODEs:

- First order ODE
$$\frac{dy}{dx} + p(x)y = q(x).$$
- Bessel's equation
$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + (x^2 - n^2)y = 0.$$
- Legendre's equation
$$(1 - x^2)\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + n(n+1)y = 0.$$

Linear Operators

We define the differential operator as $\mathcal{D}[f] \equiv \frac{d}{dx}[f]$.

Differential operator is a linear operator since we have:

$$\mathcal{D}[\lambda_1 f_1 + \lambda_2 f_2] = \lambda_1\mathcal{D}[f_1] + \lambda_2\mathcal{D}[f_2].$$

Defining differential operator of order k as $\mathcal{D}^k[f] \equiv \frac{d^k}{dx^k}[f]$, which is also a linear operator as we have:

$$\mathcal{D}^k[\lambda_1 f_1 + \lambda_2 f_2] = \lambda_1\mathcal{D}^k[f_1] + \lambda_2\mathcal{D}^k[f_2].$$

Linear ODEs are associated with a linear operator defined using the differential operators: $\mathcal{L}[y] \equiv \sum_{i=0}^{k} \alpha_i(x)\mathcal{D}^i[y]$ since:

$$\mathcal{L}[\lambda_1 f_1 + \lambda_2 f_2] = \lambda_1 \mathcal{L}[f_1] + \lambda_2 \mathcal{L}[f_2].$$

A linear ODE can thus be simply written as $\mathcal{L}[y] = f(x)$ and a homogenous ODE as $\mathcal{L}[y] = 0$. Linearity of $\mathcal{L}$ has an important consequence. If we have two solutions y_1 and y_2 of a homogenous linear ODE $\mathcal{L}[y] = 0$, then any linear combinations of these solutions are also solutions for this ODE, since:

$$\mathcal{L}[\lambda_1 f_1 + \lambda_2 f_2] = \lambda_1 \mathcal{L}[f_1] + \lambda_2 \mathcal{L}[f_2] = 0.$$

Linear independence

A set of functions $\{f_i(x)\}_{i=1}^{k}$ is said to be linearly independent if f_i's satisfy the following condition:

$$c_1 f_1(x) + c_2 f_2(x) + \cdots + c_k f_k(x) = 0,$$

if and only if $c_1 = c_2 = \cdots = c_k = 0$.

Linear ODEs are easier to solve because of the following important property of their solutions. This is a basic consequence of linearity of differential operators.

Proposition 5.1. *The solutions of the homogeneous linear ODE $\mathcal{L}[y] = 0$ form a vector space (see MATH40003: Linear Algebra and Groups) of dimension k, where k is the order of the ODE. Therefore, the general solution of a linear homogeneous ODE can be written as*

$$y_{GS}^{H}(x; c_1, ..., c_k) = c_1 y_1 + c_2 y_2 + ... + c_k y_k,$$

where $B = \{y_i(x)\}_{i=1}^{k}$ is a set of linearly independent solutions forming a basis for the linear homogeneous ODE's solution vector space.

Proposition 5.2. *To test the linear independence of a set of functions $\{y_i(x)\}_{i=1}^{k}$, we calculate the* Wronskian*, which is the determinant of the Wronskian matrix ($\mathbb{W}_{k\times k}$):*

$$W[\{y_i(x)\}_{i=1}^{k}] = det\mathbb{W} = det \begin{bmatrix} y_1(x) & y_2(x) & \cdots & y_k(x) \\ \frac{dy_1}{dx}(x) & \frac{dy_2}{dx}(x) & \cdots & \frac{dy_k}{dx}(x) \\ \vdots & \vdots & & \vdots \\ \frac{d^{k-1}y_1}{dx^{k-1}}(x) & \frac{d^{k-1}y_2}{dx^{k-1}}(x) & \cdots & \frac{d^{k-1}y_k}{dx^{k-1}}(x) \end{bmatrix}$$

The set $\{y_i(x)\}_{i=1}^{k}$ is linearly independent if

$$W[\{y_i(x)\}_{i=1}^{k}] \neq 0.$$

Proof We prove this by contradiction; If we assume $W(x) \neq 0$ but $\{y_i(x)\}_{i=1}^k$ are linearly dependent:

$$\exists i \in \{1, \cdots, k\}, \quad c_i \neq 0, \quad \sum_{i=1}^{k} c_i y_i(x) = 0.$$

By taking derivatives repeatedly with respect to x from the above equation we obtain:

$$\sum_{i=1}^{k} c_i \frac{dy_i}{dx} = 0$$
$$\vdots$$
$$\sum_{i=1}^{k} c_i \frac{d^{k-1} y_i}{dx^{k-1}} = 0,$$

which can be written as $\mathbb{W}.\vec{c} = 0$, with $\vec{c}$ defined as

$$\vec{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix},$$

since $\vec{c} \neq 0$ then $\mathbb{W}.\vec{c} = 0$ implies $W[\{y_i(x)\}_{i=1}^k] = \det\mathbb{W} = 0$, which is a contradiction. So, if $W(x) \neq 0$ then $\{y_i(x)\}_{i=1}^k$ are linearly independent, as required.

Example 5.1. Show that $\sin(x)$ and $\cos(x)$ are linearly independent.

$$\mathbb{W}_{2\times 2} = \begin{bmatrix} \sin(x) & \cos(x) \\ \cos(x) & \sin(x) \end{bmatrix},$$

then

$$W(x) = \det\mathbb{W} = -\sin^2(x) - \cos^2(x) = -1 \neq 0,$$

therefore $\sin(x)$ and $\cos(x)$ are linearly independent.

Note There exists examples in which the Wronskian vanishes without the functions being linearly dependent. An example is given in a quiz in the lectures.

5.1 General solution of the non-homogeneous linear ODE

To obtain the general solution of the non-homogeneous linear ODE

$$\mathcal{L}[y] = f(x),$$

we split the problem into two simpler steps:

1. We consider the corresponding homogeneous linear ODE $\mathcal{L}[y] = 0$. We obtain the general solution, which is also known as *complementary function* (y_{CF}):

$$y_{CF} = y^H_{GS}(x; c_1, \cdots, c_k) = \sum_{i=1}^{k} c_i y_i(x),$$

where, $\{y_i\}_{i=1}^k$ are the basis of the solution vector space (a set of linearly independent solutions of the homogeneous linear ODE).

2. We obtain any/one solution of the full non-homogeneous ODE, which is also known as *particular integral* (y_{PI}):

$$\mathcal{L}[y_{PI}] = f(x).$$

Then for the solution to the full problem by combining the results above and due to linearity, we have:

$$\mathcal{L}[y_{GS}(x; c_1, \cdots, c_k)] = \mathcal{L}[y_{CF} + y_{PI}] = \mathcal{L}[y^H_{GS}] + \mathcal{L}[y_{PI}] = f(x).$$

So the general solution of the non-homogeneous linear ODE is the sum of the complementary function and a particular integral. As seen in a quiz in the lectures different choices of particular integrals results in the same family of general solutions.

One useful consequence of the linearity is that if the RHS of the ODE is sum of two functions:

$$\mathcal{L}[y] = f_1(x) + f_2(x).$$

We can break the second step of finding particular integral into additional steps.

1. Find $y_{CF} = y^H_{GS}(x; c_1, \cdots, c_k)$ such that $\mathcal{L}[y_{CF}] = 0$.
2. Find any solution to $\mathcal{L}^1_{\mathcal{PI}}[y] = f_1(x)$.
3. Find any solution to $\mathcal{L}^2_{\mathcal{PI}}[y] = f_2(x)$.

Then, we have $y_{GS} = y_{CF} + y^1_{PI} + y^2_{PI}$.

Linear ODEs with constant coefficients

The general linear ODE is not always analytically solvable. Next year, you will see approximative and numerical methods to solve this kind of ODEs. In the rest of this course, we will focus on the case of linear ODEs with constant coefficients (α_is not depending on independent variable x):

$$\mathcal{L}[y] = \sum_{i=0}^{k} \alpha_i \mathcal{D}^i[y] = f(x)$$

5.2 First order linear ODEs with constant coefficients

The general form of the first order linear ODEs with constant coefficients is:

$$\mathcal{L}[y] = \alpha_1 \frac{dy}{dx} + \alpha_0 y = f(x).$$

As seen in chapter 4, rewriting this ODE as

$$\frac{dy}{dx} + \frac{\alpha_0}{\alpha_1} y = \frac{f(x)}{\alpha_1},$$

we can obtain the general solution using the integrating factor $I(x) = e^{\frac{\alpha_0}{\alpha_1}x}$. We obtain

$$y_{GS} = c_1 e^{-\frac{\alpha_0}{\alpha_1}x} + e^{-\frac{\alpha_0}{\alpha_1}x} \int e^{\frac{\alpha_0}{\alpha_1}x} \frac{f(x)}{\alpha_1}\, dx.$$

Example 5.2. Solve $f(x) = x$.

Using the general solution above, and by integration by parts, we obtain:

$$y_{GS} = c_1 e^{-\frac{\alpha_0}{\alpha_1}x} + \left[\frac{x}{\alpha_0} - \frac{\alpha_1}{\alpha_0^2}\right].$$

Alternative method

1. Solve the corresponding homogeneous ODE:

$$\mathcal{L}[y_{CF}] = \alpha_1 \frac{dy}{dx} + \alpha_0 y = 0.$$

   This is a separable ODE and by integration we obtain:

$$y_{CF} = y_{GS}^{H}(x; c_1) = c_1 e^{-\frac{\alpha_0}{\alpha_1}x}.$$

2. Find a particular integral for the full ODE: $\mathcal{L}[y_{PI}] = f(x) = x$.

This is done by using *ansatz*, which is an educated guess using the *method of undetermined coefficients*. In this case as $f(x)$ is polynomial, we could try a polynomial ansatz:

$$y_{PI} = Ax^2 + Bx + C,$$

where, A, B and C are constants to be determined. By plugging this ansatz in to the ODE, we check if here are suitable values for these constants that makes our ansatz a particlular solution for the ODE:

$$\mathcal{L}[y_{PI}] = \alpha_1(2Ax + B) + \alpha_0(Ax^2 + Bx + C) = x.$$

This equation should be satisfied for all $x \in \mathbb{R}$, so we equate the coefficients of the powers of x:

$$\begin{aligned} x^2: &\quad \alpha_0 A = 0 \quad \Rightarrow \quad A = 0; \\ x^1: &\quad 2\alpha_1 A + \alpha_0 B = 1 \quad \Rightarrow \quad B = \frac{1}{\alpha_0}; \\ x^0: &\quad \alpha_1 B + \alpha_0 C = 0 \quad \Rightarrow \quad C = -\frac{\alpha_1}{\alpha_0^2}, \end{aligned}$$

which gives us the same general solution obtained using the first method:

$$y_{GS} = y_{CF} + y_{PI} = c_1 e^{-\frac{\alpha_0}{\alpha_1}x} + \left[\frac{x}{\alpha_0} - \frac{\alpha_1}{\alpha_0^2}\right].$$

Example 5.3. Solve $f(x) = e^{bx}; b \neq -\frac{\alpha_0}{\alpha_1}$.

The ODE is the following:

$$\mathcal{L}[y] = \alpha_1 \frac{dy}{dx} + \alpha_0 y = e^{bx}.$$

Using the two step method, we have as before:

1. $y_{CF} = y^H_{GS}(x; c_1) = c_1 e^{-\frac{\alpha_0}{\alpha_1}x}$.
2. We try ansatz $y_{PI} = Ae^{bx}$, plugging this into the ODE, we obtain:

$$\alpha_1 Abe^{bx} + \alpha_0 Ae^{bx} = e^{bx}.$$

Solving this we obtain $A = \frac{1}{\alpha_1 b + \alpha_0}$. So we obtain:

$$y_{GS} = y_{CF} + y_{PI} = c_1 e^{-\frac{\alpha_0}{\alpha_1}x} + \frac{1}{\alpha_1 b + \alpha_0} e^{bx}.$$

What about the case $b = -\frac{\alpha_0}{\alpha_1}$? Naive ansatz $y_{PI} = Ae^{bx}$ does not work, since $\mathcal{L}[y_{PI}] = 0$. A more general ansatz is:

$$y_{PI} = A(x)e^{bx}.$$

Here we are looking for an unknown function $A(x)$, so we will obtain an ODE. This is called the *method of variation of parameters*, developed by Euler and Lagrange. Plugging this ansatz into the ODE we obtain the following simple ODE:

$$\alpha_1 \frac{dA}{dx} = 1,$$

which has the following general solution:

$$A(x) = \frac{x}{\alpha_1} + c_2.$$

So, we obtain for the general solution:

$$y_{GS} = y_{CF} + y_{PI} = c_1 e^{-\frac{\alpha_0}{\alpha_1} x} + (\frac{x}{\alpha_1} + c_2) e^{bx} = c' e^{-\frac{\alpha_0}{\alpha_1} x} + \frac{x}{\alpha_1} e^{bx},$$

where in the last step we have renamed $c_1 + c_2$ as c' a new constant of integration.

5.3 Second order linear ODEs with constant coefficients

$$\mathcal{L}[y] = \alpha_2 \frac{d^2y}{dx^2} + \alpha_1 \frac{dy}{dx} + \alpha_0 y = f(x),$$

$$y_{GS}(x; c_1, c_2) = y_{CF}(x; c_1, c_2) + y_{PI} = y^H_{GS}(x; c_1, c_2) + y_{PI}.$$

If $B = \{y_1(x), y_2(x)\}$ is a basis for the solution vector space of the homogeneous ODE: $\mathcal{L}[y^H] = 0$. Then, we have:

$$y^H_{GS}(x; c_1, c_2) = c_1 y_1(x) + c_2 y_2(x).$$

Solving the homogeneous second order linear ODE

We need to obtain two linearly independent solutions to the following ODE:

$$\mathcal{L}[y] = \alpha_2 \frac{d^2y}{dx^2} + \alpha_1 \frac{dy}{dx} + \alpha_0 y = 0.$$

We can try the ansatz: $y^H = e^{\lambda x}$,

$$\mathcal{L}[y^H] = \alpha_2 \lambda^2 e^{\lambda x} + \alpha_1 \lambda e^{\lambda x} + \alpha_0 e^{\lambda x} = 0 \quad \Rightarrow \quad \alpha_2 \lambda^2 + \alpha_1 \lambda + \alpha_0 = 0.$$

This quadratic equation is called the *characteristic equation of the linear ODE*, which has the following solutions:

$$\lambda_1, \lambda_2 = \frac{-\alpha_1 \pm \sqrt{\alpha_1^2 - 4\alpha_0\alpha_2}}{2\alpha_2}.$$

So, we have the following two candidate solutions $y^H_1 = e^{\lambda_1 x}$ and $y^H_2 = e^{\lambda_2 x}$. For these solutions to form a basis for the solution space of the homogeneous linear ODE, they should be linear independence. We evaluate the Wronskian:

$$W(x) = \det \begin{bmatrix} e^{\lambda_1 x} & e^{\lambda_2 x} \\ \lambda_1 e^{\lambda_1 x} & \lambda_2 e^{\lambda_2 x} \end{bmatrix} = e^{(\lambda_1 + \lambda_2)x}(\lambda_2 - \lambda_1).$$

So, if the roots of the characteristics equation are distinct ($\lambda_1 \neq \lambda_2$), then $W(x) \neq 0$ and the solutions form a linearly independent set. So we have:

$$y_{CF} = y^H_{GS}(x; c_1, c_2) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}.$$

For the case of $\lambda_1 = \lambda_2 = -\frac{\alpha_1}{2\alpha_2}$, we have $y_1 = e^{\lambda_1 x}$, what about the second solution y_2? We can try the ansatz $y_2 = A(x)y_1(x) = A(x)e^{\lambda_1 x}$. This is similar to the method of variation of parameters. In the context of 2nd order linear ODEs, when we have one of the solutions and looking for the second solution, this method is called the method of *reduction of order*. Plugging this ansatz into the ODE we obtain:

$$\alpha_0 \left[Ay_1\right] + \alpha_1 \left[\frac{dA}{dx}y_1 + A\frac{dy_1}{dx}\right] + \alpha_2 \left[\frac{d^2A}{dx^2}y_1 + 2\frac{dA}{dx}\frac{dy_1}{dx} + A\frac{d^2y_1}{dx^2}\right] = 0.$$

This result in the following simple ODE and solution for $A(x)$:

$$\frac{d^2A}{dx^2} = 0 \quad \Rightarrow \quad A(x) = B_1x + B_2 \quad \Rightarrow \quad y_2 = (B_1x + B_2)e^{\lambda_1 x}.$$

We note that y_2 we have obtained here contains y_1, so we can choose $y_2 = xe^{\lambda_1 x}$. Testing the linear independence of these solutions, we should evaluate the Wronskian:

$$W(x) = \det \begin{bmatrix} e^{\lambda_1 x} & xe^{\lambda_1 x} \\ \lambda_1 e^{\lambda_1 x} & e^{\lambda_1 x} + \lambda_1 xe^{\lambda_1 x} \end{bmatrix} = e^{2\lambda_1 x} \neq 0.$$

So y_1 and y_2 are linearly independent and can span the solution space. So we have the following general solution for the case characteristic equation has the repeated root λ_1:

$$y_{CF} = y^H_{GS}(x; c_1, c_2) = c_1e^{\lambda_1 x} + c_2xe^{\lambda_1 x}.$$

Possible behaviours of the 2nd order linear homogeneous ODE

If $\lambda_1 \neq \lambda_2$ then $y_{CF} = c_1e^{\lambda_1 x} + c_2e^{\lambda_2 x}$.

$$\lambda_{1,2} = -\frac{\alpha_1}{2\alpha_2} \pm \sqrt{\frac{\alpha_1^2 - 4\alpha_0\alpha_2}{4\alpha_2^2}}.$$

$\alpha_1^2 - 4\alpha_0\alpha_2 > 0 \quad \Longrightarrow \quad \lambda_{1,2} \in \mathbb{R}$

1. $\lambda_{1,2}$ can be both positive, both negative or one positive/one negative.

- If $\lambda_1 > 0$ and $\lambda_1 > \lambda_2$

$$\text{as} \quad x \to \infty, \qquad y_{CF} \to e^{\lambda_1 x} \to \infty.$$

- If $\lambda_2 < \lambda_1 < 0$

$$\text{as} \quad x \to \infty, \qquad y_{CF} \to e^{\lambda_1 x} \to 0.$$

2. $\alpha_1^2 - 4\alpha_0\alpha_2 < 0 \quad \Longrightarrow \quad \lambda_{1,2} \in \mathbb{C}$

$$\left|\frac{\alpha_1^2 - 4\alpha_0\alpha_2}{4\alpha_2^2}\right| = \omega^2 \quad \Longrightarrow \quad \lambda_{1,2} = -\frac{\alpha_1}{2\alpha_2} \pm i\omega$$

So, we have for the general solution:

$$y_{CF} = e^{-\frac{\alpha_1}{2\alpha_2}x}\left[c_1 e^{i\omega x} + c_2 e^{-i\omega x}\right] = e^{-\frac{\alpha_1}{2\alpha_2}x}\left[(c_1 + c_2)\cos\omega x + i(c_1 - c_2)\sin\omega x\right].$$

If the ODE has real coefficients the solution $y_{CF} \in \mathbb{R}$. Therefore, choosing c_1 and c_2 to be complex conjugate, we obtain $c_1' = c_1 + c_2 \in \mathbb{R}$ and $c_2' = i(c_1 - c_2) \in \mathbb{R}$. So the we can write the general solution of the homogeneous ODE with complex roots in following forms:

$$y_{CF} = e^{-\frac{\alpha_1}{2\alpha_2}x}\left[c_1' \cos\omega x + c_2' \sin\omega x\right] = e^{-\frac{\alpha_1}{2\alpha_2}x} A\cos(\omega x - \phi),$$

where in the later, we have used the following change of constants of integrations $c_1' = A\cos\phi$ and $c_2' = A\sin\phi$. Figure 5.1 shows possible behaviors of y_{CF} depending on the value of the parameter $d = \frac{\alpha_1}{2\alpha_2}$.

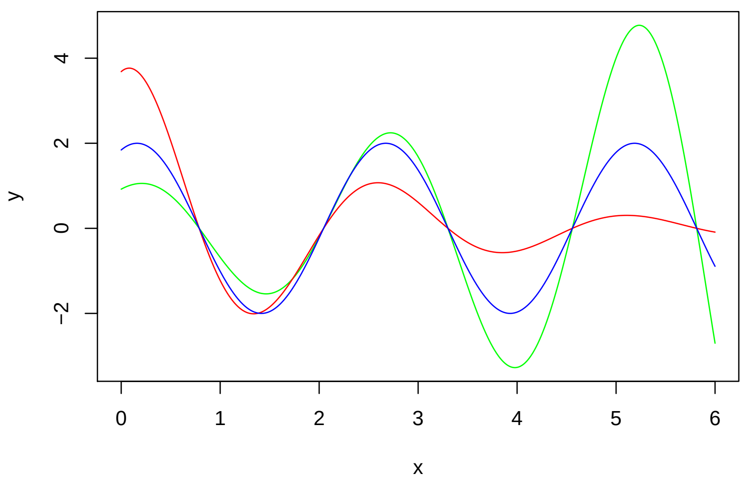

Figure 5.1: y_{CF} for $d > 0$ (red), $d < 0$ (green) and $d = 0$ (blue); all three solutions have the same phase but for clarity of visualisation different amplitudes (A) are used in each case.

Example 5.4. Find the general solution of

$$\mathcal{L}[y] = \frac{d^2y}{dx^2} - 3\frac{dy}{dx} + 2y = e^{8x}.$$

- First step: The characteristic equation is $\lambda^2 - 3\lambda + 2 = 0$, so we have $\lambda_1 = 2$ and $\lambda_2 = 1$, so
$$y_{CF} = c_1 e^{2x} + c_2 e^{x}.$$

- Second step: we try ansatz $y_{PI} = Ae^{8x}$.

$$\mathcal{L}[y_{PI}] = Ae^{8x}[64 - 24 + 2] = e^{8x} \quad \Rightarrow \quad A = \frac{1}{42},$$

So, we have:

$$y_{GS} = y_{CF} + y_{PI} = c_1 e^{2x} + c_2 e^x + \frac{1}{42} e^{8x}.$$

Example 5.5. Find the general solution of

$$\mathcal{L}[y] = \frac{d^2y}{dx^2} + 4\frac{dy}{dx} + 4y = e^{-2x}.$$

- First step: This case has a charecteristic equation with repeated root $\lambda = -2$, so we have
$$y_{CF} = c_1 e^{-2x} + c_2 x e^{-2x}.$$

- Second step: Finding a particular integral $\mathcal{L}[y_{PI}] = f(x)$. 1st try ansatz $y_{PI} = Ae^{-2x}$, which does not work. 2nd try ansatz $y_{PI} = Axe^{-2x}$, which also does not work. Let's try ansatz: $y_{PI} = A(x)e^{-2x}$ using the method of variation of parameters. By plugging into the ODE we obtain:

$$\frac{d^2A}{dx^2} = 1 \quad \Rightarrow \quad A = \frac{x^2}{2} + B_1 x + B_2.$$

So, we obtain the general solution:

$$y_{GS} = y_{CF} + y_{PI} = c_1 e^{-2x} + c_2 x e^{-2x} + \frac{x^2}{2} e^{-2x}.$$

Instead of using the method of variation of parameters, we could have guessed the ansatz $y_{PI} = Bx^2 e^{-2x}$ directly and obtaining value of B using the method of undetermined coefficients.

Example 5.6. Find the general solution of

$$\mathcal{L}[y] = \frac{d^2y}{dx^2} - 2\frac{dy}{dx} + 2y = e^x \sin(x)$$

- First step: Solving the Homogeneous problem $\mathcal{L}[y^H] = 0$. In this case roots are complex, so we have:

$$y_{CF} = c_1 e^{(1+i)x} + c_2 e^{(1-i)x}.$$

- Second step: Finding a particular integral $\mathcal{L}[y_{PI}] = f(x)$.

Here $f(x)$ can be written as the sum of two functions

$$f(x) = e^x \sin x = \frac{e^{(1+i)x}}{2i} - \frac{e^{(1-i)x}}{2i}.$$

So, we can use the following two ansatz to find particular integrals using the method of undetermined coefficients: $y_{PI}^1 = Axe^{(1+i)x}$ and $y_{PI}^2 = Axe^{(1-i)x}$. By plugging into the ODE using the first and second part of $f(x)$, one finds values for A and B respectively (left as an exercise for you) Then the general solution is:

$$y_{GS} = y_{CF} + y_{PI}^1 + y_{PI}^2.$$

5.4 kth order Linear ODEs with constant coefficients

$$\mathcal{L}[y] = \sum_{i=0}^{k} \alpha_i \mathcal{D}^i[y] = f(x); \quad \alpha_i \in \mathbb{R}$$

$$y_{GS}(x; c_1, \cdots, c_k) = y_{CF} + y_{PI} = y_{GS}^H(x; c_1, \cdots, c_k) + y_{PI}(x)$$

First step: Solving the Homogeneous problem $\mathcal{L}[y^H] = 0$

We can try the ansatz $y^H = e^{\lambda x}$:

$$\mathcal{L}[e^{\lambda x}] = e^{\lambda x} \sum_{i=0}^{k} \alpha_i \lambda^i = 0 \quad \Rightarrow \quad \sum_{i=0}^{k} \alpha_i \lambda^i = 0.$$

This is the characterstic equation of the kth order linear ODE. It has k roots that can be always obtained numerically (in the absence of an analytical solution).

- Case 1: k roots of the characteristic polynomial are distinct:

The solutions $B = \{e^{\lambda_i x}\}_{i=1}^k$ can be shown to be linearly independent using the Wronskian:

$$\mathbb{W}(x) = \begin{bmatrix} e^{\lambda_1 x} & e^{\lambda_2 x} & \cdots & e^{\lambda_k x} \\ \lambda_1 e^{\lambda_1 x} & \lambda_2 e^{\lambda_2 x} & \cdots & \lambda_k e^{\lambda_k x} \\ \vdots & \vdots & & \vdots \\ \lambda_1^{k-1} e^{\lambda_1 x} & \lambda_2^{k-1} e^{\lambda_2 x} & \cdots & \lambda_k^{k-1} e^{\lambda_k x} \end{bmatrix}$$

$$W(x) = \det \mathbb{W}(x) = e^{\sum_{i=1}^k \lambda_i x} \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_k \\ \vdots & \vdots & & \vdots \\ \lambda_1^{k-1} & \lambda_2^{k-1} & \cdots & \lambda_k^{k-1} \end{vmatrix} =$$

$$e^{\sum_{i=1}^k \lambda_i x} \prod_{1 \le i < j \le k} (\lambda_i - \lambda_j) \neq 0; \quad \text{(Vandermonde determinant)}$$

The determinant of the Vandermonde matrix (the matrix obtained above) is a well known result in linear algebra and can be proven using induction.

$$B = \{e^{\lambda_1 x}, e^{\lambda_2 x}, \cdots, e^{\lambda_k x}\} \quad \Rightarrow \quad y_{CF} = \sum_{i=1}^{k} c_i e^{\lambda_i x}.$$

- Case 2: Not all of the k roots are distinct. Below, we consider the particular case of having d repeated roots and $k-d$ distinct roots.

$$B = \{e^{\lambda_1 x}, e^{\lambda_2 x}, \cdots, e^{\lambda_r x}, xe^{\lambda_r x}, \cdots, x^{d-1}e^{\lambda_r x}, e^{\lambda_{r+1} x}, \cdots, e^{\lambda_{k-d+1} x}\} \quad \Rightarrow$$
$$y_{CF} = c_1 e^{\lambda_1 x}, c_2 e^{\lambda_2 x}, \cdots, c_r e^{\lambda_r x}, c_{r+1} x e^{\lambda_r x}, \cdots, c_{r+d-1} x^{d-1} e^{\lambda_r x}, c_{r+d} e^{\lambda_{r+1} x}, \cdots, c_k e^{\lambda_{k-d+1} x}.$$

Second step: Finding a particular integral for example for: $\mathcal{L}[y_{PI}] = e^{bx}$, for the case 2 above, we use the following ansatz, using the method of undetermined coefficients:

- if $b \neq \lambda_i$ for $\forall i$ then $y_{PI} = Ae^{bx}$.
- if $b = \lambda_i$ for $i \neq r$ then $y_{PI} = Axe^{bx}$.
- if $b = \lambda_r$ then $y_{PI} = Ax^d e^{bx}$.

5.5 Euler-Cauchy equation

A (rare) example of a linear ODE with non-constant coefficients that we can solve analytically is the Euler-Cauchy ODE:

$$\mathcal{L}[y] = \beta_k x^k \frac{d^k y}{dx^k} + \beta_{k-1} x^{k-1} \frac{d^{k-1} y}{dx^{k-1}} + \cdots + \beta_1 x \frac{dy}{dx} + \beta_0 y = f(x).$$

Using the change of variable $x = e^z$, the Euler-Cauchy equation can be transformed into a linear ODE with constant coefficients.

Example 5.7. Solve the following ODE: $x^2 \frac{d^2y}{dx^2} + 3x\frac{dy}{dx} + y = x^3$.

Using the change of variable $x = e^z$ we have $z = \ln x$ and so:

$$\frac{dy}{dx} = \frac{dy}{dz}\frac{dz}{dx} = \frac{1}{x}\frac{dy}{dz} \quad \Rightarrow \quad x\frac{dy}{dx} = \frac{dy}{dz},$$

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{d}{dz}\left(\frac{dy}{dx}\right)\frac{dz}{dx} = \frac{1}{x^2}\left[\frac{d^2y}{dz^2} - \frac{dy}{dz}\right].$$

So, in terms of the new independent variable we have the following linear ODE with constant coefficients.

$$\frac{d^2y}{dz^2} + 2\frac{dy}{dz} + y = e^{3z}.$$

By obtaining the complementary function and particular integral, we have the following general solution:

$$y_{GS}(z; c_1, c_2) = y_{CF} + y_{PI} = c_1 e^{-z} + c_2 z e^{-z} + \frac{1}{16} e^{3z}.$$

So, in the general solution in terms of x is:

$$y_{GS}(x; c_1, c_2) = \frac{c_1}{x} + c_2 \frac{\ln x}{x} + \frac{1}{16} x^3.$$

5.6 Using Fourier Transforms to solve linear ODEs

As Fourier transform is a linear operation, and given the properties we had for Fourier transforms of derivatives of a function seen in Section 2.1, one can use Fourier transforms to solve linear ODEs or find particular integrals. This is particularly relevant for solving partial differential equations as discussed in the second year. Here we discuss an example.

Example 5.8. Find a solution for the following ODE, known as the *Airy* equation or the *Stokes* equation, which arises in different areas of physics. Assume $\lim_{x\to\pm\infty} y(x) = 0$.

$$\frac{d^2 y}{dx^2} - xy = 0.$$

This is a linear 2nd order ODE with non-constant coefficients and so far we have not seen a method of solving it. Note that this ODE is not also one of the types that are discussed in Section 4.2. Our strategy is to take Fourier transform from this ODE and see if we can solve for the Fourier transform. Using the properties in Section 2.1, we obtain:

$$-\omega^2 \hat{y}(\omega) - i \frac{d\hat{y}(\omega)}{d\omega} = 0$$

This is a first order ODE for $\hat{y}(\omega)$, by solving it we obtain:

$$\hat{y}(\omega) = c e^{\frac{i\omega^3}{3}},$$

where c is an arbitrary constant of integration. Using the inverse transform we obtain:

$$y(x) = \frac{c}{2\pi} \int_{-\infty}^{\infty} e^{i(\omega x + \omega^3/3)}\, d\omega = \frac{c}{\pi} \int_0^{\infty} \cos(\omega x + \frac{\omega^3}{3})\, d\omega,$$

where, in the last step we have used the evenness of cosine and oddness of the sine function. For $c = 1$ the function $y(x)$ is known as the *Airy function* (of the first kind) and is denoted by $Ai(x)$. It is defined as the above integral and cannot be reduced further and is a solution of the Airy ODE.

Chapter 6

Introduction to Systems of ODEs

So far we have discussed ordinary differential equations where the function we have been looking for was a scaler function ($y(x) : \mathbb{R} \to \mathbb{R}$). Where the unknown function is a vector ($\vec{y}(x) : \mathbb{R} \to \mathbb{R}^n$), we are dealing with systems of ODEs.
Definition *Systems of Ordinary Differential Equations* have the following general form:

$$G_1(x, y_1, y_2, \cdots, y_n, \frac{dy_1}{dx}, \cdots, \frac{dy_n}{dx}, \cdots, \frac{d^{k_1}y_1}{dx^{k_1}}, \cdots, \frac{d^{k_n}y_n}{dx^{k_n}}) = 0,$$

$$\vdots$$

$$G_n(x, y_1, y_2, \cdots, y_n, \frac{dy_1}{dx}, \cdots, \frac{dy_n}{dx}, \cdots, \frac{d^{k_1}y_1}{dx^{k_1}}, \cdots, \frac{d^{k_n}y_n}{dx^{k_n}}) = 0.$$

The system is *ordinary* as we still have one independent variable x, but now in contrast to single ODEs, we have n functions of independent variables $y_1(x), y_2(x), \cdots, y_n(x)$ to solve for. This is the implicit form but the systems of ODEs can be written in explicit from as well. Many problems in physics and biology give rise to systems of ODEs. Here are few examples:

Example 6.1 (Predator-prey systems)**.**

These models can be used to predict the dynamics of predator and prey systems such as rabbits ($x(t)$) and foxes ($y(t)$). A classic model is Lotka-Volterra model (1925/26) that can exhibit a periodic solution of the population of predator and preys as one goes up and the other goes down.

$$\frac{dx}{dt} = ax - bxy,$$
$$\frac{dy}{dt} = dxy - cy.$$

Example 6.2 (Chemistry and biochemistry)**.**

The chemical rate equations for a set of chemical reactions. For example, consider the reversible binary reaction $A + B \rightleftharpoons C$ with forward rate of k_1 and backward rate of k_2. We have the following rate equations for the concentrations $[A]$, $[B]$ and $[C]$.

$$\frac{d[A]}{dt} = \frac{d[B]}{dt} = k_2[C] - k_1[A][B],$$
$$\frac{d[C]}{dt} = -\frac{d[A]}{dt} = -k_2[C] + k_1[A][B].$$

These kind of equations are used in mathematical modelling of biochemical reaction networks in systems and synthetic biology.

Example 6.3 (Coupled spring-mass systems)**.**

This is an example from mechanics. Consider the system of 3 masses and 4 springs that are fixed between two walls. We can write equations of motions that describe the position of the 3 masses as a function of time ($x_1(t)$, $x_2(t)$ and $x_3(t)$) as seen in Figure 6.1.

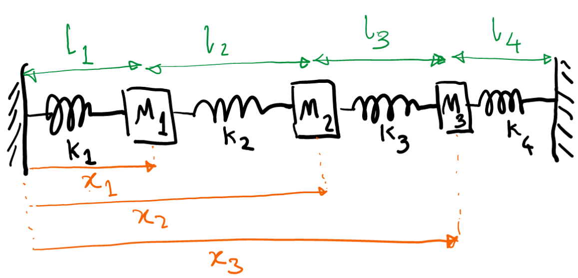

Figure 6.1: Diagram of the mass and spring system

Using the second Newton law for each mass ($F = ma$) and the Hook's law ($F = -k\Delta x$) for the springs and given the relaxed lengths of each spring (l_1 to

l_4), and assuming the distance between the walls is the sum of the the relaxed lengths of the springs, we have the following system of ODEs:

$$\frac{d^2x_1}{dt^2} = -k_1(x_1 - l_1) + k_2(x_2 - x_1 - l_2),$$
$$\frac{d^2x_2}{dt^2} = -k_2(x_2 - x_1 - l_2) + k_3(x_2 - x_2 - l_3),$$
$$\frac{d^2x_3}{dt^2} = -k_3(x_3 - x_2 - l_3) + k_4(l_1 + l_2 + l_3 - x_3).$$

Example 6.4 (SIR model of an epidemic)**.**

One of the simplest but influential models of an epidemic is the compartmental SIR model. In this model, disease propagates between the population compartments through susceptible individuals (S) becoming infected (I) after encountering other infected individuals with rate β and finally moving to a recovered population (R) with rate γ. The following simple system of ODEs describes the dynamics of this compartmental model.

$$\frac{dS}{dt} = -\beta SI,$$
$$\frac{dI}{dt} = \beta SI - \gamma I,$$
$$\frac{dR}{dt} = \gamma I.$$

This model and its variations are the basis of a lot of mathematical modelling that has been performed for predicting the effect of different interventions in the Covid-19 world-wide pandemic. If you like to learn a bit more about the SIR ODE system above check out this video. Also, the SIR model can be modeled using a so-called *agent-based* stochastic approach, where the individuals and their random interactions are specifically followed. You can check out this interesting video from the 3blue1brown series to see some cool exploration of this approach to SIR models.

6.1 Systems of first order ODEs

Systems of ODEs of general order can be rewritten in terms of systems of first order ODEs, so the following system written in explicit form is more general than it seems.

$$\frac{dy_1}{dx} = F_1(x, y_1, y_2, \cdots, y_n),$$

$$\vdots$$

$$\frac{dy_n}{dx} = F_n(x, y_1, y_2, \cdots, y_n).$$

Example 6.5 (Turning a higher order ODE into systems of 1st order ODEs)**.**

The following second order ODE is the equation of motion for damped harmonic oscillator for a mass m that is attached to an ideal spring which follows the Hook's law ($F_S = -kx$, where x is position of the mass measured from the spring relaxed position) and has a damping friction force that is propotional and opposite in direction to its velocity $F_D = -\eta\frac{dx}{dt}$ and is acted on by a deriving force $F(t)$.

$$m\frac{d^2x}{dt^2} + \eta\frac{dx}{dt} + kx = F(t).$$

By defining new variable $u = dx/dt$ as the velocity of the mass m, we can turn this second order ODE to a system of two first order ODEs:

$$\frac{dx}{dt} = u,$$
$$\frac{du}{dt} = \frac{1}{m}[F(t) - \eta u - kx].$$

We can use a general vector notation to write systems of 1st order ODEs as

$$\frac{d\vec{y}_{n\times 1}}{dt} = \vec{F}_{n\times 1}(t, \vec{y}_{n\times 1}).$$

Here n is the number of equations, t is the independent variable and $\vec{y}$ is the function we are looking for. In the next section, we discuss an important subclass of these systems.

6.2 Systems of linear 1st order ODEs with constant coefficients

Systems of linear 1st order ODEs with constant coefficients is an important class that we will discuss their solutions in detail. They have the following general form.

$$\frac{dy_1}{dt} = \sum_{i=1}^{n} \alpha_{1i} y_i + g_1(t),$$
$$\vdots$$
$$\frac{dy_n}{dt} = \sum_{i=1}^{n} \alpha_{ni} y_i + g_n(t),$$

where $\alpha_{ij} \in \mathbb{R}$ are the constant coefficients forming the matrix $A_{n\times n}$. We can write the system of linear ODEs in matrix form:

$$\begin{bmatrix} \frac{dy_1}{dt} \\ \frac{dy_2}{dt} \\ \vdots \\ \frac{dy_n}{dt} \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} + \begin{bmatrix} g_1(t) \\ g_2(t) \\ \vdots \\ g_n(t) \end{bmatrix}.$$

We can write this briefly as

$$\frac{d\vec{y}}{dt} = A\vec{y} + \vec{g}(t).$$

Now if we define the linear operator $\mathcal{L}[\vec{y}] = [\frac{d}{dt} - A]\vec{y}$ then we can write the system of linear first order ODEs as

$$\mathcal{L}[\vec{y}] = \vec{g}(t).$$

Since both matrix A and derivative $\frac{d}{dt}$ are linear operators, the operator associated to systems of linear ODEs $\mathcal{L}$ is also a linear operator. By linearity we have:

$$\mathcal{L}[\lambda_1\vec{y_1} + \lambda_2\vec{y_2}] = \lambda_1\mathcal{L}[\vec{y_1}] + \lambda_2\mathcal{L}[\vec{y_2}].$$

Therefore, the solutions of the homogenous systems of n linear ODEs $\mathcal{L}[\vec{y}_H] = 0$ forms a vector space of dimension n. So a set of linearly independent solutions $B = \{\vec{y}_i\}_{i=1}^n$ form a basis for this space. Therefore, similar to linear ODEs, the general solution can be written as

$$\vec{y}_{GS}^H = \sum_{i=1}^n c_i\vec{y}_i,$$

where c_is are n arbitrary constants of integration.

The general solution of non-homogenous systems of 1st order linear ODEs

Similar to the case of linear ODEs, here also we find the general solution in two steps

1. Obtain complimentary function $\vec{y}_{CF}$ by solving the corresponding homogenous systems of ODEs ($\mathcal{L}[\vec{y}_{CF}] = 0$)
2. Find a particular integral $\vec{y}_{PI}$ that satisfies the full non-homogenous systems of ODEs ($\mathcal{L}[\vec{y}_{PI}] = \vec{g}(t)$).

Then, for the general solution $\vec{y}_{GS}$ we have:

$$\vec{y}_{GS}(t; c_1, c_2, \cdots, c_n) = \vec{y}_{CF} + \vec{y}_{PI}.$$

Solving the homogenous problem

$$\mathcal{L}[\vec{y}_H] = 0 \quad \Longrightarrow \quad \frac{d\vec{y}_H}{dt} = A\vec{y}_H$$

First, we consider the case where matrix A has n distinct roots and therefore is diagonalizable.

That means that there exists a matrix V where, we have:

$$V^{-1}AV = \Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}.$$

where, ith column of the matrix V is the eigenvector of matrix A corresponding to the eigenvalue λ_i:

$$A\vec{v}_i = \lambda_i \vec{v}_i.$$

We use the eigenvectors matrix V to obtain the solution of the homogenous system.

$$\frac{d\vec{y}}{dt} = A\vec{y} \quad \Longrightarrow \quad V^{-1}\frac{d\vec{y}}{dt} = V^{-1}AVV^{-1}\vec{y}.$$

Letting $\vec{z} = V^{-1}\vec{y}$, we can write the system of ODEs as

$$\frac{d\vec{z}}{dt} = \Lambda\vec{z}.$$

The ith row of this equation gives us $\frac{dz_i}{dt} = \lambda_i z_i$, which can be solved, so we obtain:

$$\vec{Z} = \begin{bmatrix} c_1 e^{\lambda_1 t} \\ c_2 e^{\lambda_2 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{bmatrix} \quad \Longrightarrow \quad \vec{y}_H = V\vec{Z} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ \vdots & \vdots & & \vdots \end{bmatrix} \begin{bmatrix} c_1 e^{\lambda_1 t} \\ c_2 e^{\lambda_2 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{bmatrix}.$$

Therefore, we obtain the general solution of the homogenous system of first order linear ODEs to be:

$$\vec{y}_{CF} = \vec{y}^H_{GS} = c_1 e^{\lambda_1 t}\vec{v}_1 + c_2 e^{\lambda_2 t}\vec{v}_2 + \cdots + c_n e^{\lambda_n t}\vec{v}_n.$$

Finding particular integrals

$$\mathcal{L}[\vec{y}_{PI}] = \vec{g}(t).$$

We will use Ansatz and use the methods of undetermined coefficients and variation of parameters as done for the linear ODEs.

Example 6.6. Solve the system of ODEs for $\{x(t), y(t)\}$.

$$\frac{dx}{dt} = -4x - 3y - 5,$$
$$\frac{dy}{dt} = 2x + 3y - 2.$$

We can write this ODE in vector form as:

$$\frac{d\vec{y}}{dt} = A\vec{y} + \vec{g}(t) = \begin{bmatrix} -4 & -3 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} -5 \\ -2 \end{bmatrix}.$$

First step is to obtain $\vec{y}_{CF}$. We obtain the eigenvalues λ_1 and λ_2 and eigenvectors $\vec{v}_1$ and $\vec{v}_2$.

$$\lambda^2 + \lambda - 6 = 0 \quad \implies \quad \lambda_1 = 2,\, \lambda_2 = -3.$$

We have the corresponding eigenvectors:

$$\vec{v_1} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad \vec{v_2} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

$$\vec{y}_{CF} = \vec{y}^H_{GS}(t; c_1, c_2) = c_1 e^{2t} \begin{bmatrix} 1 \\ -2 \end{bmatrix} + c_2 e^{-3t} \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

2nd step is to find any particular integral $\vec{y}_{PI}(t)\}$ that satisfies:

$$\mathcal{L}[\vec{y}_{PI}] = \vec{g}(t) = \begin{bmatrix} -5 \\ -2 \end{bmatrix}.$$

We use the ansatz:

$$\vec{y}_{PI} = \begin{bmatrix} a \\ b \end{bmatrix}.$$

By plugging the ansatz into the ODE we obtain the undetermined coefficients a and b:

$$\vec{y}_{PI} = \begin{bmatrix} a \\ b \end{bmatrix} = A^{-1} \begin{bmatrix} 5 \\ 2 \end{bmatrix} = -\frac{1}{6} \begin{bmatrix} 3 & 3 \\ -2 & -4 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} -7/2 \\ 3 \end{bmatrix}.$$

So we have:

$$\vec{y}_{GS}(t; c_1, c_2) = \vec{y}_{CF} + \vec{y}_{PI} = c_1 e^{2t} \begin{bmatrix} 1 \\ -2 \end{bmatrix} + c_2 e^{-3t} \begin{bmatrix} 3 \\ -1 \end{bmatrix} + \begin{bmatrix} -7/2 \\ 3 \end{bmatrix}.$$

Example 6.7. Solve the problem of damped Harmonic spring with zero forcing

$$m\frac{d^2x}{dt^2} + \eta\frac{dx}{dt} + kx = 0$$

This is a second order linear ODE and we can solve it using the methods discussed in the last chapter. Using the ansatz $e^{\lambda t}$, we obtain the following characteristic equation:

$$m\lambda^2 + \eta\lambda + k = 0 \quad \Rightarrow \quad \lambda_{1,2} = \frac{-\eta \pm \sqrt{\eta^2 - 4km}}{2m}.$$

If the roots are distinct we obtain:

$$x_{GS} = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}.$$

Alternatively, we can transform the second order ODE to systems of first order ODEs as seen before:

$$\frac{dx}{dt} = u,$$
$$\frac{du}{dt} = -\frac{\eta}{m}u - \frac{k}{m}x.$$

We obtain the same $\lambda_{1,2}$ as above for the eigenvalues of corresponding matrix A for this system of ODEs.

For the eigenvectors ($\vec{v}_1$ and $\vec{v}_2$) of A, we have:

$$\vec{v}_1: \quad A\vec{v}_1 = \lambda_1\vec{v}_1 \quad \Longrightarrow \quad \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{\eta}{m} \end{bmatrix} \begin{bmatrix} v_{1x} \\ v_{1u} \end{bmatrix} = \lambda_1 \begin{bmatrix} v_{1x} \\ v_{1u} \end{bmatrix},$$

$$v_{1u} = \lambda_1 v_{1x} \quad \Longrightarrow \quad \vec{v}_1 = \begin{bmatrix} 1 \\ \lambda_1 \end{bmatrix}.$$

Similarly, we obtain for the second eigenvector:

$$\vec{v}_2: \quad A\vec{v}_2 = \lambda_2\vec{v}_2 \quad \Longrightarrow \quad \vec{v}_2 = \begin{bmatrix} 1 \\ \lambda_2 \end{bmatrix}.$$

This gives us the following general solution:

$$\vec{y}_{GS} = \begin{bmatrix} x_{GS} \\ u_{GS} \end{bmatrix} = c_1 e^{\lambda_1 t} \begin{bmatrix} 1 \\ \lambda_1 \end{bmatrix} + c_2 e^{\lambda_2 t} \begin{bmatrix} 1 \\ \lambda_2 \end{bmatrix},$$

which gives us the same general solution for x_{GS} as obtained using the previous method.

When A has repeated eigenvalues

Case 1: A is still diagonalizable (it has n linearly independent eigenvectors). Then we can still use the method described. For example for $n = 2$ we have:

$$A = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \quad \Rightarrow \quad \vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

So, we have

$$\vec{y}_{CF} = c_1 e^{\lambda t} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 e^{\lambda t} \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Case 2: A is not diagonalizable (it has less than n linearly independent eigenvectors). Then we will use the *Jordan normal form.* We first discuss an example and then see the general case.

Example 6.8.

$$\frac{d\vec{y}}{dt} = A\vec{y}; \quad A = \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix}.$$

Matrix A in this case has repeated roots:

$$(1-\lambda)(3-\lambda) + 1 = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = 2.$$

We next find the Eigenvector(s):

$$A\vec{v}_1 = \lambda_1 \vec{v}_1 \quad \Longrightarrow \quad \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} v_{1x} \\ v_{1y} \end{bmatrix} = 2 \begin{bmatrix} v_{1x} \\ v_{1y} \end{bmatrix}.$$

Which gives us:

$$v_{1x} = -v_{1y} \quad \Rightarrow \quad \vec{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

So, this matrix only has one eigenvector and the matrix is not diagonalizable.

We look for similarity transformation to a Jordan normal form (J; an almost diagonal form as defined below). We look for a matrix of the form:

$$W = \begin{bmatrix} 1 & \alpha \\ -1 & \beta \end{bmatrix}.$$

So that:

$$W^{-1}AW = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} = J,$$

where, J is the Jordan normal form for a 2 by 2 matrix. We have:

$$Aw = w \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & \alpha \\ -1 & \beta \end{bmatrix} = \begin{bmatrix} 1 & \alpha \\ -1 & \beta \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

So, we obtain the following equations for α and β.

$$\begin{aligned} \alpha - \beta &= 1 + 2\alpha \\ \alpha + 3\beta &= -1 + 2\beta \end{aligned} \quad \Longrightarrow \quad \alpha + \beta = -1 \quad \Longrightarrow \quad \vec{w}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix},$$

giving us

$$W = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix},$$

and we can check that $W^{-1}AW = J$. The Jordan normal form allows us to solve the non-diagonalizable systems of ODEs:

$$\frac{d\vec{y}}{dt} = A\vec{y} \quad \Longrightarrow \quad W^{-1}\frac{d\vec{y}}{dt} = [W^{-1}AW]\, W^{-1}\vec{y}.$$

Letting $\vec{z} = W^{-1}\vec{y}$, we obtain $\frac{d\vec{z}}{dt} = J\vec{z}$, so we have:

$$\begin{aligned} \frac{dz_1}{dt} &= 2z_1 + z_2, \\ \frac{dz_2}{dt} &= 2z_2. \end{aligned}$$

We can solve the second ODE to obtain $z_2 = c_2e^{2t}$ and we can then use this solution to solve for z_1 from the first equation above:

$$z_1 = c_1e^{2t} + c_2te^{2t}.$$

So in vector form:

$$\vec{z}_{GS} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} c_1e^{2t} + c_2te^{2t} \\ c_2e^{2t} \end{bmatrix}.$$

Now we can obtain the general solution for $\vec{y}_{GS}$.

$$\vec{y}_{GS} = W\vec{z}_{GS} = [\vec{v}_1 \quad \vec{w}_2] \begin{bmatrix} c_1e^{2t} + c_2te^{2t} \\ c_2e^{2t} \end{bmatrix} = (c_1e^{2t} + c_2te^{2t})\vec{v}_1 + c_2e^{2t}\vec{w}_2.$$

The case of non-diagonalizable $A_{n\times n}$ with one repeated eigenvalue λ

Assume λ is associated with only a single eigenvector. We can use the Jordan normal form (J) to obtain a solution to the associated systems of linear ODEs. We look for a similarity transformation W to transform A to J:

$$W^{-1}AW = J = \begin{bmatrix} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & \ddots & 1 & 0 \\ 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & \lambda \end{bmatrix}.$$

Letting $\vec{z} = W^{-1}\vec{y}$, we obtain $\frac{d\vec{z}}{dt} = J\vec{z}$, so we have:

$$\begin{aligned} \frac{dz_n}{dt} &= \lambda z_n & &\Rightarrow & z_n &= c_ne^{\lambda t} \\ \frac{dz_{n-1}}{dt} &= \lambda z_{n-1} + z_n & &\Rightarrow & z_{n-1} &= c_{n-1}e^{\lambda t} + c_nte^{\lambda t} \\ \frac{dz_{n-2}}{dt} &= \lambda z_{n-2} + z_{n-1} & &\Rightarrow & z_{n-2} &= c_{n-2}e^{\lambda t} + c_{n-1}te^{\lambda t} + c_n\tfrac{t^2}{2}e^{\lambda t} \\ &\vdots \\ \frac{dz_1}{dt} &= \lambda z_1 + z_2 & &\Rightarrow & z_1 &= c_1e^{\lambda t} + c_2te^{\lambda t} + \cdots + c_n\tfrac{t^{n-1}}{(n-1)!}e^{\lambda t} \end{aligned}$$

So, we can obtain $\vec{y}_{GS} = W\vec{z}_{GS}$.

There could be situations where the matrix has some distinct eigenvalues and some repeated eigenvalues, which will result in different Jordan normal forms. For example, consider a matrix $A_{3\times3}$ with two distinct eigenvalues one repeated. The suitable Jordan normal form would have the following form:

$$J = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix}.$$

Chapter 7

Qualitative analysis of ODEs

So far we have focused to obtain analytical solution to ODEs, but this is not always possible. Even, when it is possible, it is not always very insightful. In this section, we will focus on qualitative analysis of ODEs and as before we mostly focus on linear ODEs. We'll discuss asymptotics behavior, fixed points (and their stability) and phase plane analysis.

7.1 asymptotic behaviour

In qualitative analysis of an ODE, *asymptotic behaviour* of solution $y(t)$ as $t \to \infty$ is one aspect of solutions we look at.

Example 7.1 (Population growth)**.**

$$\frac{dP(t)}{dt} = KP(t)$$

Here, the solution is an exponential $P(t) = P(0)e^{Kt}$. So for the asymptotic behaviour we have:

$$K > 0 \quad \Rightarrow \quad \lim_{t\to\infty} P(t) \to \infty, \tag{7.1}$$

$$K < 0 \quad \Rightarrow \quad \lim_{t\to\infty} P(t) \to 0. \tag{7.2}$$

Fixed points of systems of first order ODEs

$\vec{y}^*$ is a *fixed point* or an *equilibrium point* of a system of first order ODEs, if once $\vec{y}(t_0) = \vec{y}^*$ at some time t_0 then for all future times $t > t_0$, state vector $\vec{y}$ remains equal to $\vec{y}^*$. Thus, at fixed point we have:

$$\left[\frac{d\vec{y}}{dt}\right]_{\vec{y}=\vec{y}^*} = 0$$

Example 7.2 (Logistic growth)**.** A modified model for population growth that does not lead to exponential growth for $K > 0$ is the logistic growth model, with a carrying capacity C:

$$\frac{dP}{dt} = KP(1 - \frac{P}{C}).$$

This ODE has two fixed points:

$$\frac{dP}{dt} = 0 \quad \Rightarrow \quad P_1^* = 0, \quad P_2^* = C,$$

with the former being an unstable fixed point and the latter being an stable one as defined below.

Example 7.3 (Systems of Linear homogeneous ODEs)**.**

$$\frac{d\vec{y}}{dt} = A\vec{y}.$$

A system of linear ODEs can have one or infinitely many fixed points:

$$\frac{d\vec{y}}{dt} = 0 \quad \Rightarrow \quad A\vec{y}^* = 0 \quad \Rightarrow \quad \begin{cases} \vec{y}^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, & \text{if } \mathrm{Det}(A) \neq 0, \\ \vec{y}^* \text{could be a line or plane}, & \text{if } \mathrm{Det}(A) = 0. \end{cases}$$

Stability of fixed points

Informally, a fixed point is stable if whenever the initial state is near that point, the state remains near it, perhaps even tending toward the equilibrium point as time increases. Formally, we have two types of stability as described below.

Lyapunov stability

A fixed point $\vec{y}^*$ is said to be *Lyapunov stable*, if for every $\epsilon > 0$, there exists a $\delta > 0$ such that, if $\|\vec{y}(0) - \vec{y}^*\| < \delta$, then for $\forall t \geq 0$, we have $\|\vec{y}(t) - \vec{y}^*\| < \epsilon$.

Intuitively, it means the solution does not blow up but also does not necessarily approach to the fixed point.

Asymptotic stability

A fixed point $\vec{y}^*$ is said to be *Asymptotically stable*, if it is Lyapunov stable and there exists a $\delta > 0$ such that, if $\|\vec{y}(0) - \vec{y}^*\| < \delta$, then we have

$$\lim_{t\to\infty} \|\vec{y}(t) - \vec{y}^*\| = 0.$$

Intuitively, in this case the solution does approach to the fixed point over long times.

7.2 Phase plane analysis

The general solution of systems of ODEs is given by the family of parametric curves, specified by the initial condition:

$$\vec{y}(t; c_1, \cdots, c_n) \in \mathbb{R}^n.$$

These solutions represent trajectories in $\mathbb{R}^n$ for a system of n dimensional ODEs. For a 2 dimensional systme the family of solutions starting from different initial conditions can be represented in a so called *phase plane* as illustrated in Figure 7.1.

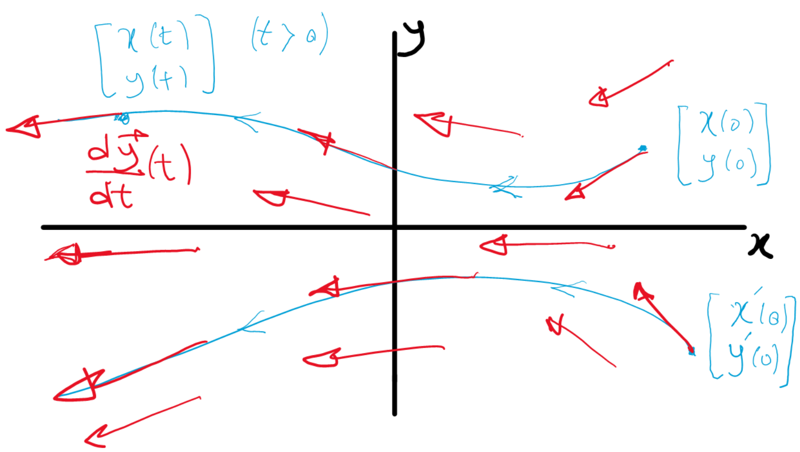

Figure 7.1: Illustration of phase plane for a 2 dimensional system. Two solutions starting at different initial conditions are shown (blue). Also, some representative velocity vectors from the vector field are drawn (red).

One can interpret the phase plane in terms of dynamics. The solution $\vec{y}(t)$ corresponds to a trajectory of a point moving on the phase plane with velocity $\frac{d\vec{y}}{dt}(t)$. For a system of first order ODEs of the form:

$$\frac{d\vec{y}}{dt} = F(\vec{y}),$$

where, there is no explicit dependence on the independent variable (time) on the right hand side, the velocity is a vector defined at every point of the phase plane and is tangent to the trajectory. This is called the *vector field*. See Figure 7.1.

Uniqueness of solutions of ODEs

Solutions of ODEs are uniquely defined by initial conditions except at some special points in the phase plane (no proof now, you will see in the second year rigorous proof). Trajectories in the phase plane cannot cross (except at some special points) as this would be equivalent of non-uniqueness of solutions. The special points are fixed points or singular points where trajectories start or end.

Phase plane analysis for the linear systems of first order ODEs

For linear systems, the vector field has some very nice properties.

$$\frac{d\vec{y}}{dt} = A\vec{y}$$

Eigenvectors define very special directions in the phase plane ($A\vec{v}_1 = \lambda_1\vec{v}_1$) in a linear system. The line defined by $\vec{v}_1$ in the phase plane is an *invariant*, meaning that if we start on $\vec{v}_1$, we will remain on it. Let $\vec{y}(0) = \alpha\vec{v}_1$ with $\alpha \in \mathbb{R}$. We have:

$$A\vec{y}(0) = \alpha A\vec{v}_1 = \alpha\lambda_1\vec{v}_1 = \frac{d\vec{y}(0)}{dt},$$

so, if $\lambda_1 > 0$, $y(t)$ grows along $\vec{v}_1$ and if $\lambda_1 < 0$, $y(t)$ decays along $\vec{v}_1$ to $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

To check this explicitly, we go back to the general solution of systems of ODE (2 dimensional case):

$$\vec{y}(t) = c_1 e^{\lambda_1 t}\vec{v}_1 + c_2 e^{\lambda_2 t}\vec{v}_2$$

Let $\vec{y}(0) = \alpha\vec{v}_1$, this gives $c_1 = \alpha$ and $c_2 = 0$. So we have:

$$\vec{y}(t) = \alpha e^{\lambda_1 t}\vec{v}_1 = e^{\lambda_1 t}\vec{y}(0),$$

so, we see explicitly, from this solution that, if $\lambda_1 > 0$, $y(t)$ grows along $\vec{v}_1$ and if $\lambda_1 < 0$, $y(t)$ decays along $\vec{v}_1$ to $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$. $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is the fixed point of the system as we have

$$\frac{d\vec{y}}{dt}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = A\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Example 7.4 (First example of qualitative and phase plane analysis of linear systems of ODE)**.**

$$\frac{d\vec{y}}{dt} = A\vec{y}; \quad A = \begin{bmatrix} -4 & -3 \\ 2 & 3 \end{bmatrix}.$$

We have the solution for this system of linear ODEs using the eigenvalues and eigenvectors for the matrix A:

$$\vec{y}_{GS}(t) = c_1 e^{2t} \begin{bmatrix} 1 \\ -2 \end{bmatrix} + c_2 e^{-3t} \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

Next we can look at the asymptotic behavior. Asymptotically, solutions blow up parallel to $\vec{v}_1$, unless we start on $\vec{v}_2$, which then we approach $\vec{y}^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. This is evident from the solution above as if $c_1 \neq 0$ as $t \to \infty$ then $\vec{y}_{GS} \to \infty$.

Aim of phase plane analysis is to obtain the *phase portrait* of the system, which is a summary of all distinct solutions, with qualitatively different trajectories in the phase plane. To do this for our linear system we draw the lines corresponding to the directions of the eigenvectors and trajectories that start on these lines. We consider the asymptotic behavior and we also compute (some examples of the) vector field at some points to draw some representative trajectories. For example, we have at $\vec{y} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$:

$$\frac{d\vec{y}}{dt}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = A \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}.$$

Figure 7.2 shows the phase portrait for this system. We can explicitly obtain the equations for the trajectories using either the general solution (as seen in the quiz in the lectures) or by obtaining an ODE for $y(x)$ by dividing the equation for $\frac{dy}{dt}$ by $\frac{dx}{dt}$:

$$\begin{aligned} \frac{dx}{dt} &= -4x - 3y \\ \frac{dy}{dt} &= 2x + 3y \end{aligned} \qquad \Longrightarrow \qquad \frac{dy}{dx} = \frac{2x + 3y}{-4x - 3y}$$

This is a homogeneous first order ODE and by using the change of variable $u = y/x$, we can obtain the following solution.

$$(x + 3y)^3 (2x + y)^2 = c,$$

where c is a constant of integration and its different values gives us the different trajectories in the phase plane as illustrated in the phase portrait in Figure 7.2. This figure and all the figures in the next section are plotted using the following online applet.

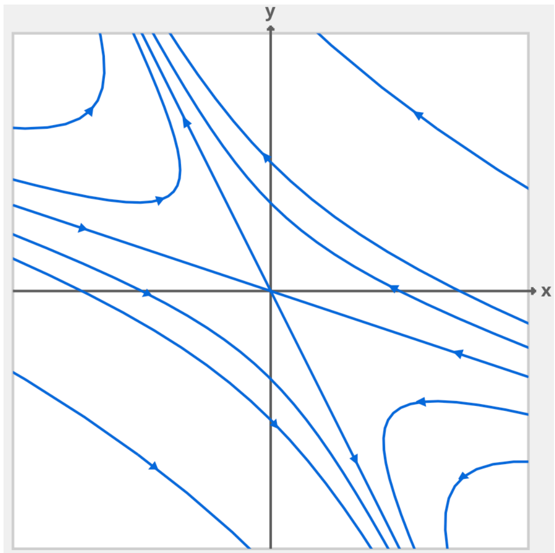

Figure 7.2: The phase portrait for the linear ODE in Example 7.4.

7.3 General system of linear ODEs in 2 dimension

In this section we present a catalogue of qualitative analysis of the general 2 dimensional system of linear ODEs:

$$\frac{d\vec{y}}{dt} = A\vec{y}; \quad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

The general solution of this system can be written in terms of the eigenvalues and eigenvectors of matrix A as seen in the last chapter. The eigenvalues can be obtained by solving the following characteristic equation:

$$\lambda^2 - \tau\lambda + \Delta = 0,$$

where $\tau = a + d$ is the trace and $\Delta = ad - bc$ is the determinant of the matrix A.

$$\tau = \text{trace}(A); \quad \Delta = \text{Det}(A); \quad \lambda_1, \lambda_2 = \frac{\tau \pm \sqrt{\tau^2 - 4\Delta}}{2}$$

The general solution is

$$y_{GS} = c_1 e^{\lambda_1 t}\vec{v}_1 + c_2 e^{\lambda_2 t}\vec{v}_2,$$

where $\vec{v}_i$ is the eigenvector corresponding to eigenvalue λ_i. In the following we consider the qualitative behaviours of this system for different values in the (τ, Δ) plane.

1. Saddle-point or Hyperbolic profile. $\Delta < 0$; Lower half of (τ, Δ).

In this case we have $\lambda_1 \in \mathbb{R}^+$ and $\lambda_2 \in \mathbb{R}^-$ since:

$$\Delta < 0 \quad \Longrightarrow \quad \tau^2 - 4\Delta > \tau^2 > 0$$

This is the case similar to Example 7.4 and we have asymptotically as $t \to \infty$, $\vec{y}(t) \to c_1 e^{\lambda_1} \vec{v}_1$, which grows exponentially as λ_1 is a real positive number. However, if we start on the line characterised by $\vec{v}_2$ direction (i.e. $c_1 = 0$), the solution goes to zero.

$$t \to \infty \quad \Longrightarrow \quad \vec{y} \to \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

An example of saddle point phase portrait can be seen in Figure 7.3.

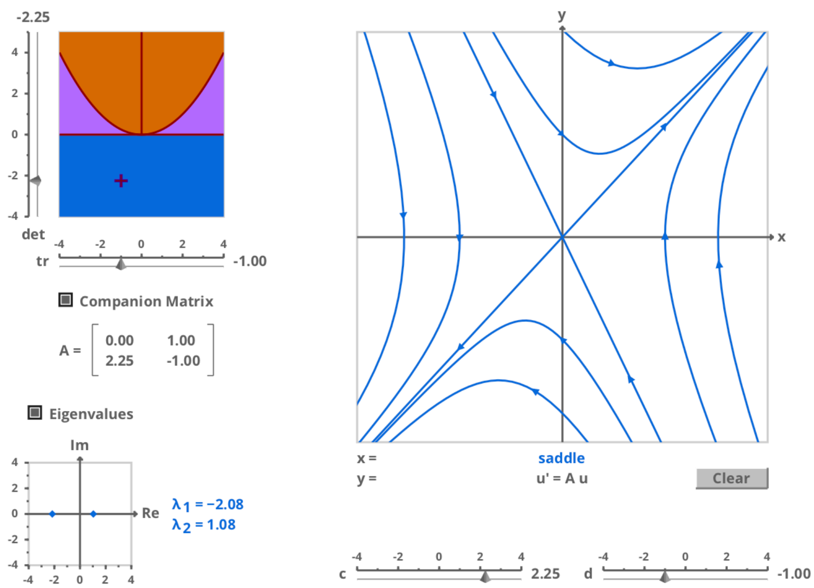

Figure 7.3: The saddle point phase portrait. This figure is plotted using this [online applet](http://mathlets.org/mathlets/linear-phase-portraits-matrix-entry/)

2.1.1: Repelling or unstable node. $0 < \Delta < \frac{\tau^2}{4}$; $\tau > 0$.

In this case we have $\lambda_1, \lambda_2 \in \mathbb{R}^+$ and $\lambda_1 > \lambda_2 > 0$. So starting on $\vec{v}_2$ blow-up along the direction of $\vec{v}_2$. Otherwise, blow up in the direction of $\vec{v}_1$.

$$t \to \infty \quad \Longrightarrow \quad \vec{y}(t) \to c_1 e^{\lambda_1} \vec{v}_1 \to \infty.$$

An example of repelling or unstable node phase portrait can be seen in Figure 7.4.

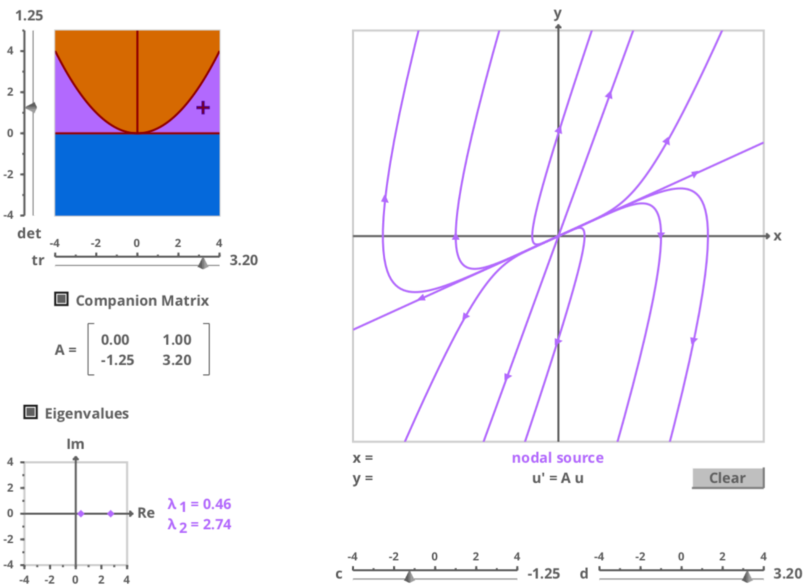

Figure 7.4: The repelling or unstable node phase portrait. This figure is plotted using this [online applet](http://mathlets.org/mathlets/linear-phase-portraits-matrix-entry/)

2.1.2: Attracting or stable node. $0 < \Delta < \frac{\tau^2}{4}; \tau < 0$.

In this case we have $\lambda_1, \lambda_2 \in \mathbb{R}^-$ and $\lambda_2 < \lambda_1 < 0$. So starting on $\vec{v}_2$ decays to $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ along the direction of $\vec{v}_2$. Otherwise, decays along the direction of $\vec{v}_1$ to $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$. So, for $c_1 \neq 0$, we have:

$$t \to \infty \quad \Longrightarrow \quad \vec{y}(t) \to c_1 e^{-|\lambda_1|} \vec{v}_1 \to \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

An example of attracting or stable node phase portrait can be seen in Figure 7.5.

2.2.1: Centre or elliptic profile. $\Delta > \frac{\tau^2}{4}$; $\tau = 0$

In this case $\lambda_{1,2} = \pm i\omega$ and the solution is periodic. Periodic behaviour corresponds to closed curves in the phase plane. For linear systems the closed curves are ellipses.

Example 7.5 (Harmonic oscillator: revisited)**.**

$$\frac{x^2}{dt^2} + \omega^2 x = 0.$$

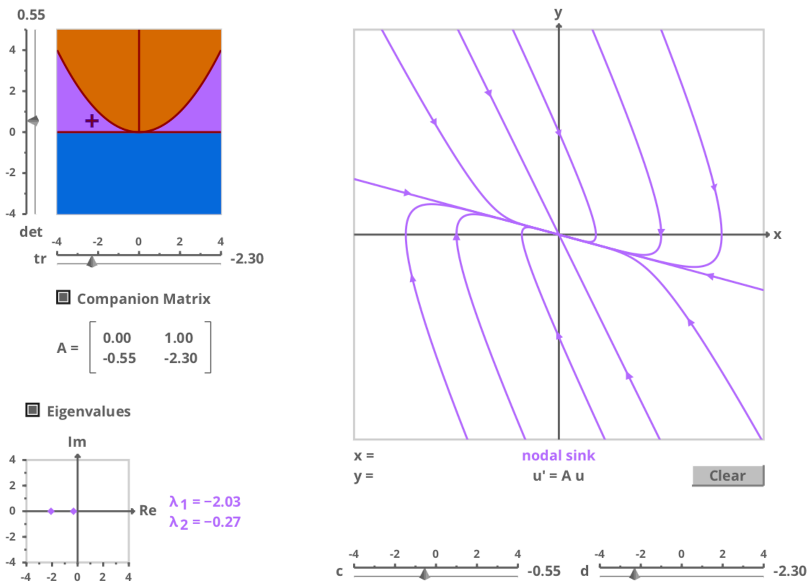

Figure 7.5: The attracting or stable node phase portrait. This figure is plotted using this [online applet](http://mathlets.org/mathlets/linear-phase-portraits-matrix-entry/)

We can write this second order linear ODE as a system of two first order linear ODEs, but defining $y = \frac{dx}{dt}$. The general solution for this system can be written as:

$$\begin{bmatrix} x_{GS} \\ y_{GS} \end{bmatrix} = \begin{bmatrix} A_0 \sin(\omega t + \phi) \\ A_0 \omega \cos(\omega t + \phi) \end{bmatrix}.$$

From this result we see that the trajectory of solutions are ellipses.

$$x^2 + \frac{y^2}{\omega^2} = A_0^2 = x_0^2 + \frac{y_0^2}{\omega^2},$$

where A_0 is a constant of integration and it depends on the initial conditions x_0 and y_0.

The phase portrait illustrating the elliptic profile can be seen in Figure 7.6. To figure out the direction of motion we evaluate the vector field at some points:

$$\vec{y} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad \Rightarrow \quad \frac{d\vec{y}}{dt} = A\vec{y} = \begin{bmatrix} y \\ 0 \end{bmatrix},$$

$$\vec{y} = \begin{bmatrix} x \\ 0 \end{bmatrix} \quad \Rightarrow \quad \frac{d\vec{y}}{dt} = A\vec{y} = \begin{bmatrix} 0 \\ -\omega^2 x \end{bmatrix}.$$

Note that in this case the point $\vec{y} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a stable fixed point, with Lyapunov stability as the trajectories around the origin do not blow up but also do not asymptotically approach the origin.

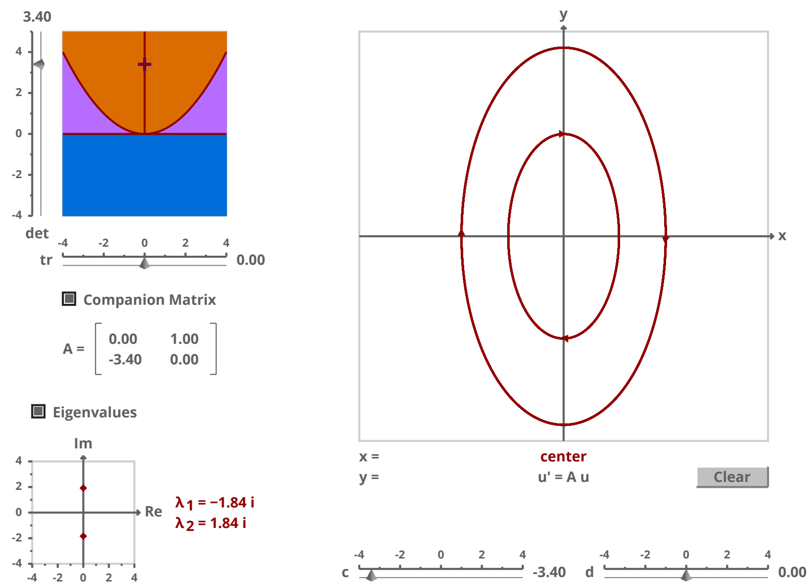

Figure 7.6: The centre or elliptic phase portrait. This figure is plotted using this [online applet](http://mathlets.org/mathlets/linear-phase-portraits-matrix-entry/)

2.2.2: Repelling or unstable spiral. $\Delta > \frac{\tau^2}{4}$; $\tau > 0$.

The eigenvalues are complex with the real part being postive. So, for the general solution we have:

$$\vec{y} = e^{\frac{\tau}{2}t}\left[c_1 e^{i\omega t}\vec{v}_1 + c_2 e^{-i\omega t}\vec{v}_2\right],$$

which, asymptotically blow up in an oscilatory fashion. The phase portrait illustrating the repelling or unstable spiral can be seen in Figure 7.7.

2.2.3: Attracting or stable spiral. $\Delta > \frac{\tau^2}{4}$; $\tau < 0$

The eigenvalues are complex with the real part being negative. So, for the general solution we have:

$$\vec{y} = e^{\frac{\tau}{2}t}\left[c_1 e^{i\omega t}\vec{v}_1 + c_2 e^{-i\omega t}\vec{v}_2\right],$$

which, asymptotically as $t \to \infty$, $\vec{y} \to \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. The phase portrait illustrating the attracting or stable spiral can be seen in Figure 7.8.

3.1: Line of repelling or unstable fixed points. $\Delta = 0$; $\quad \tau > 0$.

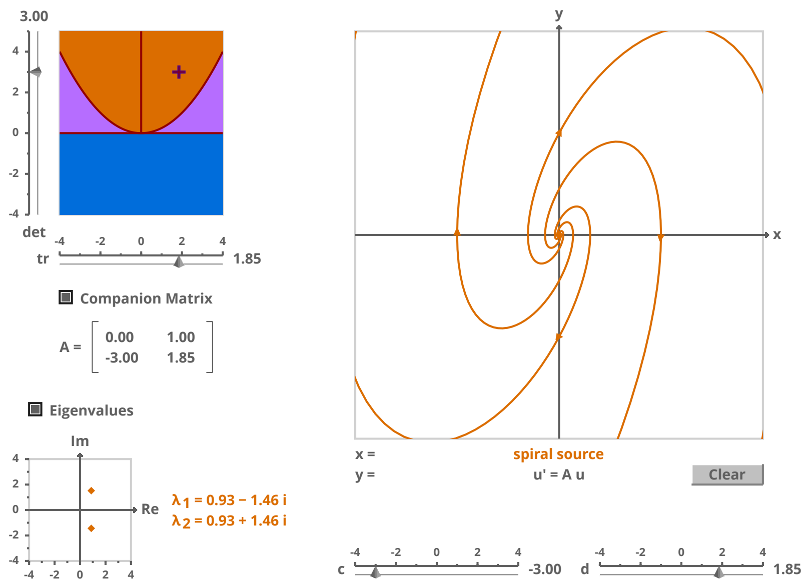

Figure 7.7: The repelling or unstable spiral phase portrait. This figure is plotted using this [online applet](http://mathlets.org/mathlets/linear-phase-portraits-matrix-entry/)

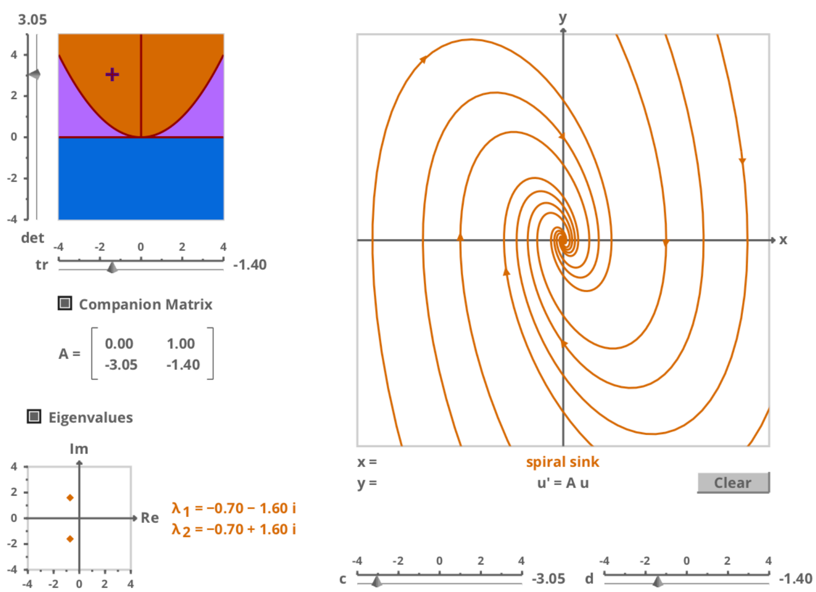

Figure 7.8: The attracting or stable spiral phase portrait. This figure is plotted using this [online applet](http://mathlets.org/mathlets/linear-phase-portraits-matrix-entry/)

In this case we have $\lambda_1 = \tau$ and $\lambda_2 = 0$, so the general solution is:

$$\vec{y} = c_1 e^{\tau t}\vec{v}_1 + c_2\vec{v}_2.$$

For the vector field we have:

$$\frac{d\vec{y}}{dt} = c_1 \tau e^{\tau t}\vec{v}_1,$$

So, $y^* = c_2\vec{v}_2$ is a line of unstable fixed points. Asymptotically as $t \to \infty$ then $\vec{y}(t) \to c_1\tau e^{\tau t}\vec{v}_1$, which blows up exponentially in the direction of $\vec{v}_1$. The phase portrait for this case can be seen in Figure 7.9.

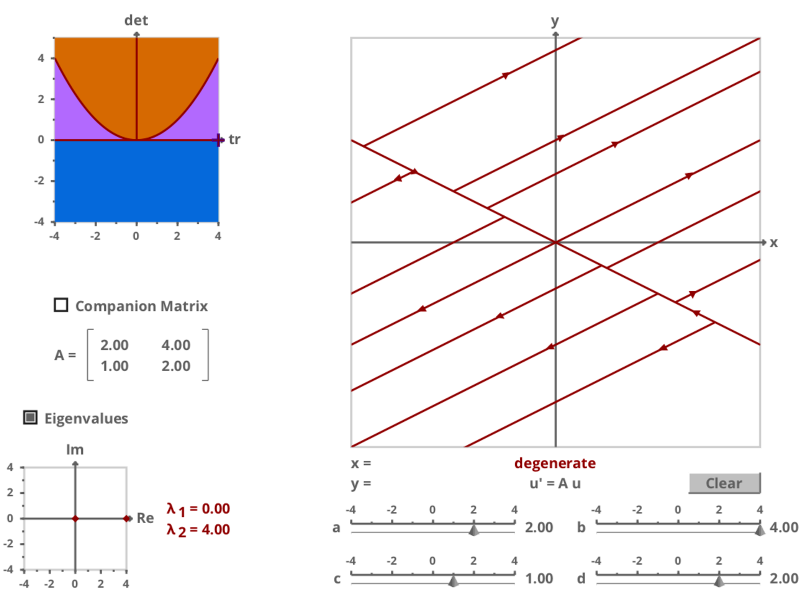

Figure 7.9: Line of repelling or unstable fixed points phase portrait. This figure is plotted using this [online applet](http://mathlets.org/mathlets/linear-phase-portraits-matrix-entry/)

3.2: Line of attracting or stable fixed points. $\Delta = 0; \quad \tau < 0.$

In this case we have $\lambda_1 = \tau$ and $\lambda_2 = 0$ again similar to last case and the general solution is

$$\vec{y} = c_1 e^{\tau t}\vec{v}_1 + c_2\vec{v}_2.$$

For the vector field we have:

$$\frac{d\vec{y}}{dt} = c_1 \tau e^{\tau t}\vec{v}_1.$$

So, $y^* = c_2\vec{v}_2$ is a line of stable fixed points. Asymptotically as $t \to \infty$ then $\vec{y}(t) \to c_1\tau e^{\tau t}\vec{v}_1$, which decays exponentially to the line of $c_2\vec{v}_2$. The phase portrait for this case can be seen in Figure 7.10.

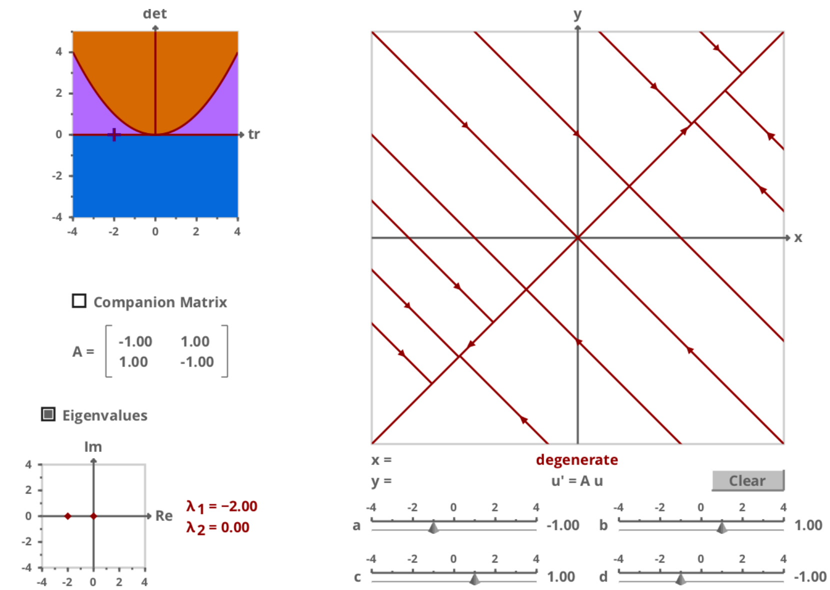

Figure 7.10: line of attracting or stable fixed points phase portrait. This figure is plotted using this [online applet](http://mathlets.org/mathlets/linear-phase-portraits-matrix-entry/)

4.1: Repelling (4.1.1) and attracting (4.1.2) star node. $\tau^2 - 4\Delta = 0$

In this case the eigenvalues are repeated $\lambda_1 = \lambda_2 = \frac{\tau}{2}$ and A is diagonalizable. We have:

$$A = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \quad \Rightarrow \quad \vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

So, we have for the general solution

$$\vec{y} = c_1 e^{\lambda t} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 e^{\lambda t} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = e^{\lambda t} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

We observe that in this case, the trajectory is always defined by $\vec{y}(0)$. The phase portrait for the repelling star node (4.1.1) can be seen in Figure 7.11. The attracting star node phase portrait is similar with an asymptotically stable fixed point at the origin $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

4.2: Unstable (4.2.1) and stable (4.2.2) improper or degenerate node. $\tau^2 - 4\Delta = 0$

In this case the eigenvalues are repeated $\lambda_1 = \lambda_2 = \frac{\tau}{2}$ but A is non-diagonalizable and $\tau > 0$ (unstable 4.2.1) or $\tau < 0$ (stable 4.2.2).

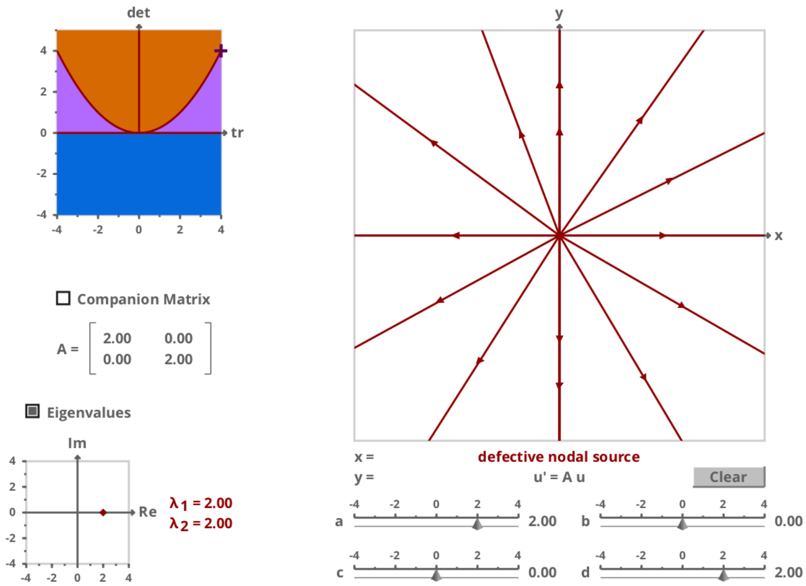

Figure 7.11: Repelling star node phase portrait. This figure is plotted using this [online applet](http://mathlets.org/mathlets/linear-phase-portraits-matrix-entry/)

Example 7.6.

$$\frac{d\vec{y}}{dt} = A\vec{y}; \quad A = \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix}.$$

In this last chapter, using the Jordan normal form, we showed the general solution of this system of ODEs to be:

$$\vec{y}_{GS} = (c_1 e^{2t} + c_2 t e^{2t})\vec{v}_1 + c_2 e^{2t}\vec{w}_2,$$

where $\vec{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is the only eigenvector, and $\vec{w}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$. We see that as $t \to \infty$, $\vec{y}(t)$ blows up in the direction of $\vec{v}_1$. We can estimate the vector field at specific points to helps draw the phase portrait. The phase portrait for this case can be seen in Figure 7.12, which is of the type unstable improper or degenerate node (4.2.1). Stable improper or degenerate node (4.2.2) have a similar phase portrait to this with an asymptotically stable fixed point at the origin $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

The catalogue of phase portraits for the 2 dimensional linear systems of ODEs can be seen in Figure 7.13 on the (τ, Δ) plane. We note that the solutions can be unstable, asymptotically stable or Lyapunov stable in the different regions

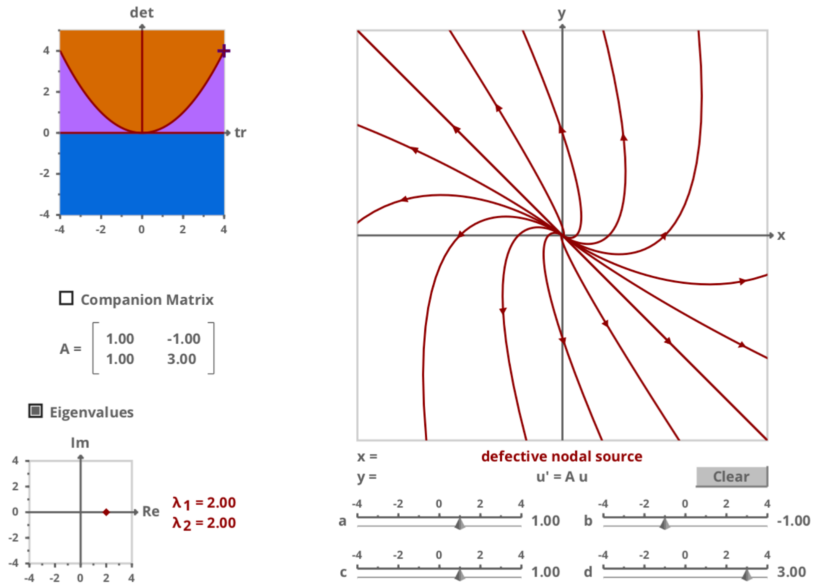

Figure 7.12: Unstable improper or degenerate node phase portrait. This figure is plotted using this [online applet](http://mathlets.org/mathlets/linear-phase-portraits-matrix-entry/)

of (τ, Δ) plane. The system can have one fixed point or infinitely many fixed points. Also, solutions could be oscillatory or non-oscillatory in different regions of the parameter space.

7.4 Phase plane analysis for 2D nonlinear systems

Phase plane analysis is a very useful tool for 2 dimensional nonlinear systems and the insights obtained from the 2D linear case is highly relevant. One obtains the special points (fixed points) and could linearise the equations in the vicinity of these points to get insight from the analysis of the corresponding linear system. Vector field and asymptotic behaviour is useful to identify the trajectories in the phase plane.

Example 7.7 (Synthetic Biology: Genetic Toggle Switch)**.**

This model of a simple synthetic genetic network was proposed in a pioneering paper by Gardener, Cantor and Colins, *Nature* 403:339-342 (2000). Consider two genes u and v, which inhibit expression of one another. The following system of nonlinear ODEs characterises the dynamics between the two genes.

Figure 7.13: The catalogue of phase portraits for the 2 dimensional linear systems of ODEs

$$\frac{du}{dt} = \frac{\alpha_1}{1+v^\beta} - u,$$
$$\frac{dv}{dt} = \frac{\alpha_2}{1+u^\gamma} - v.$$

Characterising the fixed points of this model, allowed the authors to successfully design and construct one of the first synthetic genetic networks. This system has up to 3 fixed points (two stable and one unstable ones).

Example 7.8 (Lotka-Voltera Model)**.**

Consider this classic model of predator-prey, where $x > 0$ denotes number of rabbits and $y > 0$ number of foxes in a population (with $a, b, c, d > 0$).

$$\frac{dx}{dt} = ax - bxy,$$
$$\frac{dy}{dt} = dxy - cy.$$

By setting the derivatives to zero, we obtain the following two fixed points for the system.

$$\begin{bmatrix} x^* \\ y^* \end{bmatrix}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} x^* \\ y^* \end{bmatrix}_2 = \begin{bmatrix} c/d \\ a/b \end{bmatrix}.$$

We consider dynamics near each fixed points by considering

$$x = x^* + \Delta x,$$

$$y = y^* + \Delta y,$$

where, $\Delta x, \Delta y << 1$. Linearising around the first fixed point by omitting terms of the order of $\Delta x \Delta y$, we obtain:

$$\frac{d\Delta x}{dt} = a\Delta x,$$
$$\frac{d\Delta y}{dt} = -c\Delta y.$$

This is a 2D linear system of ODEs that exhibits a saddle-point phase portrait suggesting that the first fixed point is unstable. Similarly, linearising around the second fixed point, we obtain:

$$\frac{d\Delta x}{dt} = -\frac{bc}{d}\Delta y,$$
$$\frac{d\Delta y}{dt} = \frac{da}{b}\Delta x.$$

This is a 2D linear system of ODEs that exhibits a centre phase portrait suggesting that the second fixed point has Lyapunove stability. Putting these together we get the phase portrait in Figure 7.14 for the Lotka-Voltera Model that suggests the system has periodic trajectories around the second fixed point in the phase plane.

Figure 7.14: The phase portraits for the Lotka-Voltera model

7.5 Extension of phase plane analysis to higher dimensional systems

1. $\vec{y}(t)$ can be considered a trajectory in $\mathbb{R}^n$, where n is the dimensionality of the system.
2. From each initial condition there is a unique trajectory and trajectories do not cross except at some special points.

3. We can consider asymptotic behavior to draw the trajectories.
4. We can compute vector field:

$$\frac{d\vec{y}}{dt} = \vec{F}(\vec{y}).$$

5. Fixed points ($\vec{y}^*$) are obtained by:

$$\frac{d\vec{y}}{dt}(\vec{y}^*) = \vec{0}.$$

The approach directly generalises for the linear systems as the solutions are given by the eigenvectors and eigenvalues of matrix A. As in the 2 dimensional case, solutions starting in the directions set by the eigenvectors, stay on these directions and grow or decay depending on the sign of the corresponding eigenvalue.

Stability of linear n dimensional systems For the 2 dimensional case we had stability where $\tau \leq 0$ and $\Delta \geq 0$. In terms of eigenvalue characterization this meant the real part of the eigen values are negative. Similarly, for the general n dimensional linear systems of ODEs we have stability if the real part of all the eigenvalues are negative.

Lorenz system (1933, Edward Lorenz)

More complex dynamics in phase planes are possible. Lorenz proposed a system of 3 nonlinear equations that is a model of atmospheric convection. This rather simple model for certain values of parameters (e.g. $\sigma = 10$, $\beta = \frac{8}{3}$ and $\rho = 28$), exhibits a complex non-periodic dynamics that is an example of chaos. This dynamical behavior is characterised by the divergence of trajectories in the phase plane starting from near identical initial conditions as illustrated in this video.

$$\frac{dx}{dt} = \sigma(y - x),$$
$$\frac{dy}{dt} = x(\rho - z) - y,$$
$$\frac{dz}{dt} = xy - \beta z.$$

Chapter 8

Introduction to Bifurcations

Bifurcations in a dynamical system (system of ODEs) describe the qualitative change in behavior under a variation or change of some *parameters* of the system. Parameters are constants that are tunable.

8.1 Bifurcations in linear systems

We start by looking at linear systems, first through an example:

Example 8.1 (Taking k to be the tuning parameter in the damped harmonic oscilator system)**.**

$$\frac{d^2x}{dt^2} + 2k\frac{dx}{dt} + \omega^2 x = 0$$

How does the qualitative behavior of $x(t;k)$ change when $k \in \mathbb{R}$ is varied?

We know this system is equivalent to a system of linear first order linear ODEs, letting $u = \frac{dx}{dt}$ we have:

$$\frac{d}{dt}\begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -2k \end{bmatrix}\begin{bmatrix} x \\ u \end{bmatrix}.$$

Given the catalogue of phase portraits in the (τ, Δ) plane we saw in the last chapter, we can use these to see how the qualitative behavior of $\vec{y}(t;k) = \begin{bmatrix} x \\ u \end{bmatrix}$ changes when k is varied. For this system we have $\tau = -2k$ and $\Delta = \omega^2$, so only τ depends on the tunable parameter k and Δ is always positive. As Figure 8.1 shows there are 7 different phase portraits that can be observed as k is varied. Defining a *bifurcation* as a change in stability of the system, there is only one bifurcation point at $k = 0$ for this system.

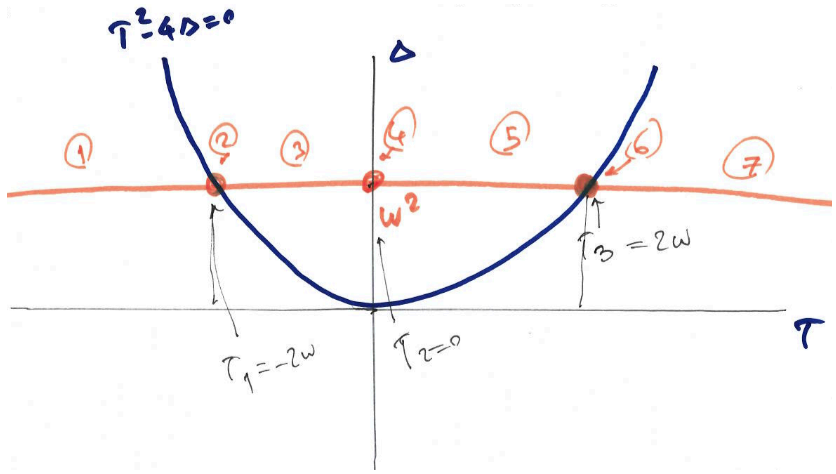

Figure 8.1: Qualitative behavior of the damped harmonic oscillator system as k is varied

In linear systems the bifurcations are related to changes in the stability of the system, as that is the main type of change that can happen in the dynamics of the system.

In non-linear systems there is a whole zoo of bifurcations and we will not cover these here but we will consider one dimensional nonlinear systems in the next section.

8.2 Qualitative behavior of non-linear 1D systems

Consider the general one dimensional nonlinear first order ODE.

$$\frac{dy}{dt} = f(y); \quad y \in \mathbb{R}^1$$

The phase plane for 1D systems can be considered. We have:

- $y(t)$ are trajectories on the real line.
- Vector field describing how we move is the velocity and is scalar in 1D case ($f(y) \in \mathbb{R}$).
- special points include
  1. Fixed points: $f(y^*) = 0$.
  2. Singularities where $f(y_{sing})$ is non-defined.

Example 8.2 (First (trivial) example of the phase plane and bifurcation analysis in one dimension)**.**

$$\frac{dy}{dt} = ky = f(y; k).$$

The fixed point for this system is $y^* = 0$ as $f(y* = 0) = 0$. The general solutio for this ODE is

$$y(t; k) = y(0)e^{kt}.$$

The general solution suggests that the fixed point is stable for $k < 0$ and is unstable for $k > 0$. Also, we can also use a plot of $\frac{dy}{dt} = f(y; k)$ vs y for different values of k to draw the vector field as illustrated in Figure 8.2 to obtain the stability of the fixed points. A stable fixed point has a flow towards it, while an unstable fixed points has an outward flow.

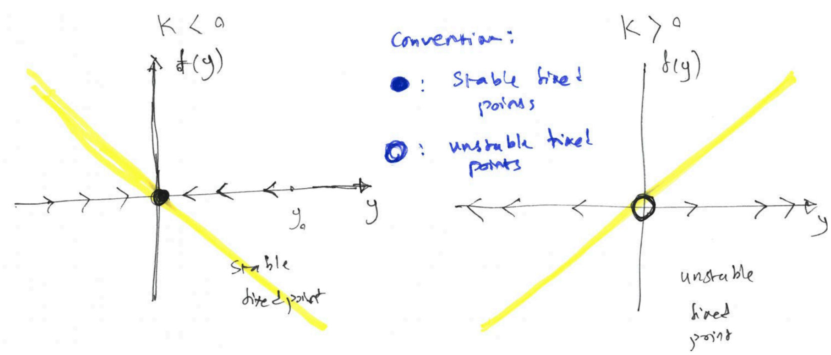

Figure 8.2: The plot of $f(y)$ vs y for different values of k, illustrates the vector field and the stability of the fixed point for Example 8.2.

A bifurcation diagram summarises all possible behaviours of the system as a parameter is varied. It represents all fixed points of the system and their stability as a function of the varying parameter. The bifurcation diagram for this example is drawn in Figure 8.3.

There are only 3 kinds of nonlinear 1D systems in terms of their bifurcation.

8.2.1 Saddle-node bifurcation

This is a basic mechanism for creation and destroying fixed points. The prototypical example of saddle-node bifurcation is given by:

$$\frac{dy}{dt} = r + y^2,$$

where $r \in \mathbb{R}$. We have $y^* = \pm\sqrt{-r}$. Figure 8.4 using the vector field, illustrates number and stability of these fixed points for different values of r.

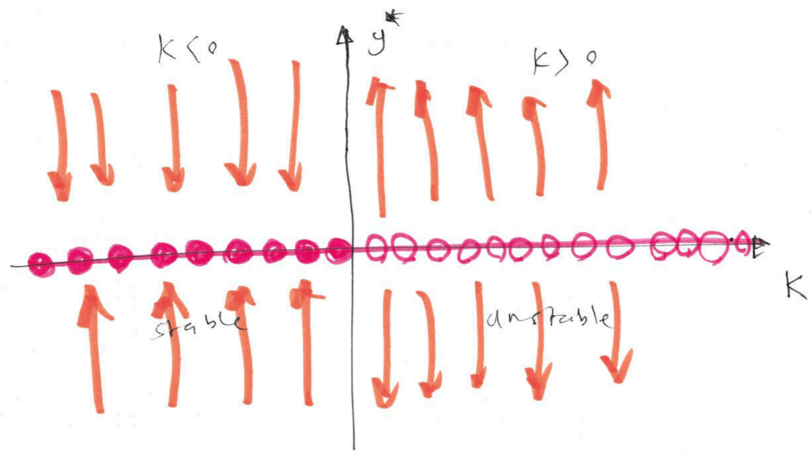

Figure 8.3: The bifurcation diagram (plot of fixed points and their stability vs the tuning parameter k) for Example 8.2.

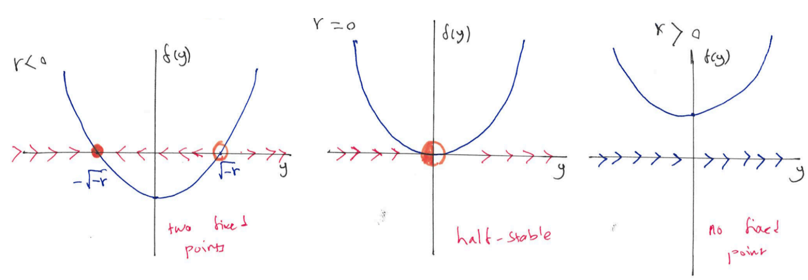

Figure 8.4: The plot of $f(y)$ vs y for different values of r, illustrates the vector field and the stability of the fixed point for saddle-node bifurcation.

For $r < 0$ we have 2 fixed points (one stable, one unstable), at $r = 0$ we have one half-stable fixed point and for $r > 0$ we have no fixed point. There is a *saddle-node bifurcation* at $r = 0$. These changes in the number and stability of the fixed points can be summarised using the bifurcation diagram, which is a plot of fixed points vs parameter r (see Figure 8.5).

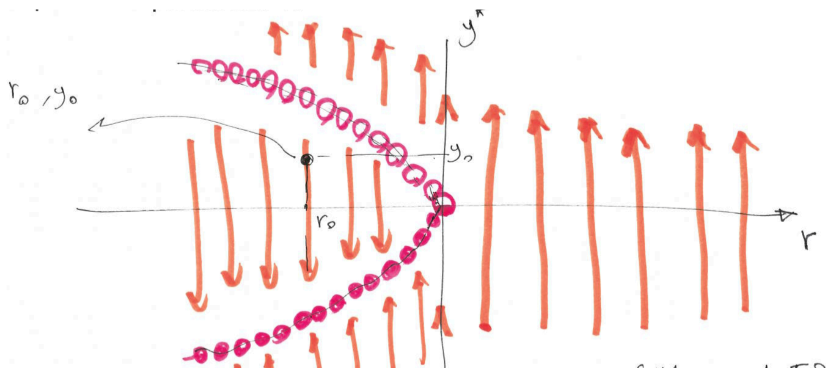

Figure 8.5: The bifurcation diagram (plot of fixed points and their stability vs the tuning parameter r) for saddle-node bifurcation.

8.2.2 Transcritical bifurcation

In certain systems a fixed point must exist for all values of a parameter. The prototypical example of this form of bifurcation is given by:

$$\frac{dy}{dt} = ry - y^2,$$

where $r \in \mathbb{R}$. This ODE has up to two fixed points $y^* = 0$ and $y^* = r$. Figure 8.6 using the vector field, illustrates number and stability of these fixed points for different values of r.

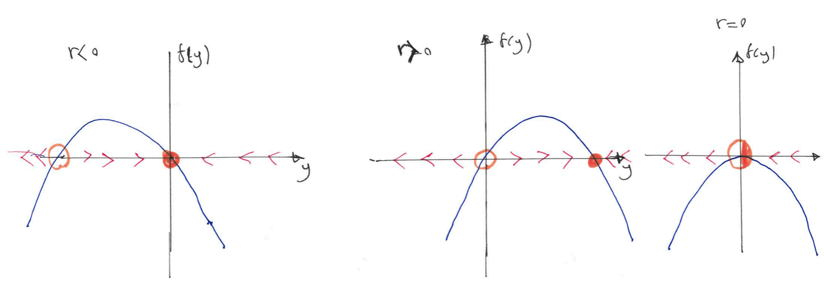

Figure 8.6: The plot of $f(y)$ vs y for different values of r, illustrates the vector field and the stability of the fixed point for transcritical bifurcation.

For $r < 0$ we have 2 fixed points (one stable, one unstable), at $r = 0$ we have one half-stable fixed point and for $r > 0$ go back to two fixed points. At the bifurcation point $r = 0$ an exchange of stabilities takes place between the two fixed points. The changes in the number and stability of the fixed points is summarised in the bifurcation diagram (see Figure 8.7).

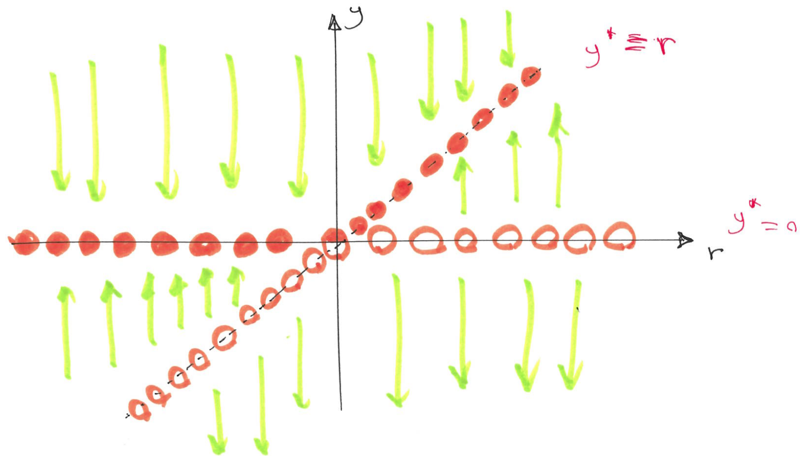

Figure 8.7: The bifurcation diagram (plot of fixed points and their stability vs the tuning parameter r) for transcritical bifurcation.

8.2.3 Pitchfork bifurcation

This kind of bifurcation is common in physical systems that have a symmetry. There are two subtypes of pitchfork bifurcation.

Supercritical pitchfork bifurcation

The prototypical example of supercritical pitchfork bifurcation is given by:

$$\frac{dy}{dt} = ry - y^3 = f(y;r)$$

Note that the equation is invariant under the change of variable $y \to -y$, which signifies the symmetry. The system has up to three fixed points $y^* = 0$ and $y^* = \pm\sqrt{r}$. Figure 8.8 using the vector field, illustrates number and stability of these fixed points for different values of r.

For $r < 0$ we have 1 stable fixed point, at $r = 0$ we have still one stable fixed point and for $r > 0$ we have three fixed points (two stable and a middle one that is unstable). At the bifurcation point $r = 0$ an exchange of stabilities takes place between the fixed points. The changes in the number and stability of the fixed points is summarised in the bifurcation diagram (see Figure 8.9).

Subcritical pitchfork bifurcation

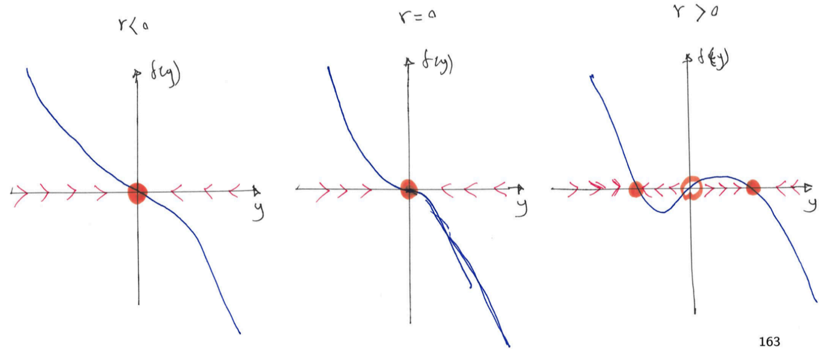

Figure 8.8: The plot of $f(y)$ vs y for different values of r, illustrates the vector field and the stability of the fixed point for supercritical pitchforkbifurcation.

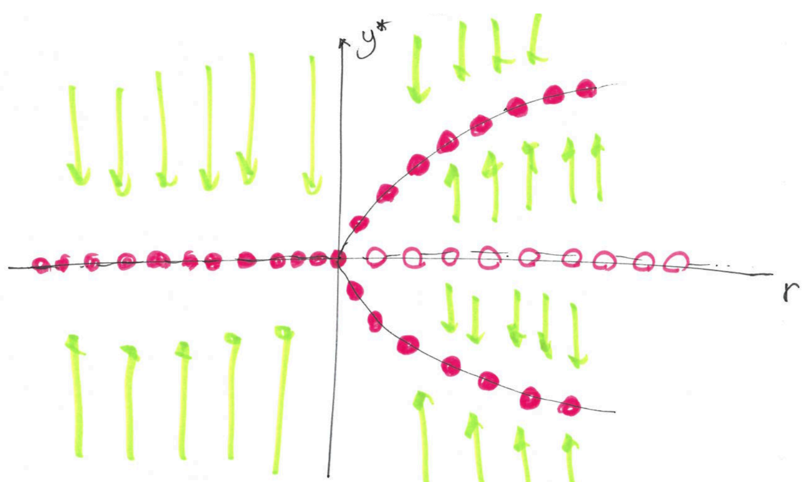

Figure 8.9: The bifurcation diagram (plot of fixed points and their stability vs the tuning parameter r) for supercritical pitchfork bifurcation.

In the subcritical pitchfork bifurcation in contrast to the supercritical case the cubic term is destabilising:

$$\frac{dy}{dt} = ry + y^3 = f(y;r)$$

The system again has up to three fixed points $y^* = 0$ and $y^* = \pm\sqrt{-r}$. Figure 8.10 using the vector field, illustrates number and stability of these fixed points for different values of r.

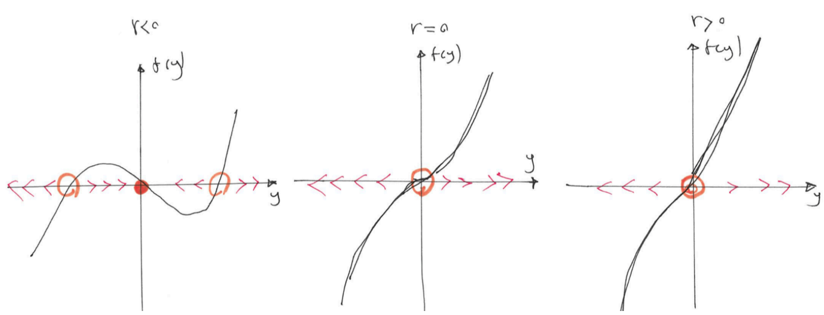

Figure 8.10: The plot of $f(y)$ vs y for different values of r, illustrates the vector field and the stability of the fixed point for subcritical pitchfork bifurcation.

For $r < 0$ we have we have three fixed points (two unstable and a middle one that is stable), at $r \geq 0$ we have one unstable fixed point. At the bifurcation point $r = 0$ an exchange of stabilities takes place between the fixed points. The changes in the number and stability of the fixed points is summarised in the bifurcation diagram (see Figure 8.11).

8.2.4 Singular points

As at the beginning of this section the special points can be the fixed points y^* or the singularities y_{sing}, where the $f(y_{sing};r)$ is not defined.

Example 8.3 (An example of an ODE with singularity)**.**

$$\frac{dy}{dt} = \frac{k}{y} = f(y,k)$$

Here, $f(y;k)$ is not defined at $y_{sing} = 0$. Figure 8.12 using the vector field, shows the *stability* of the singularity for different values of k.

For this example, we have an explicit solution that provides further insight in the behavior of the system for different values of k and initial condition.

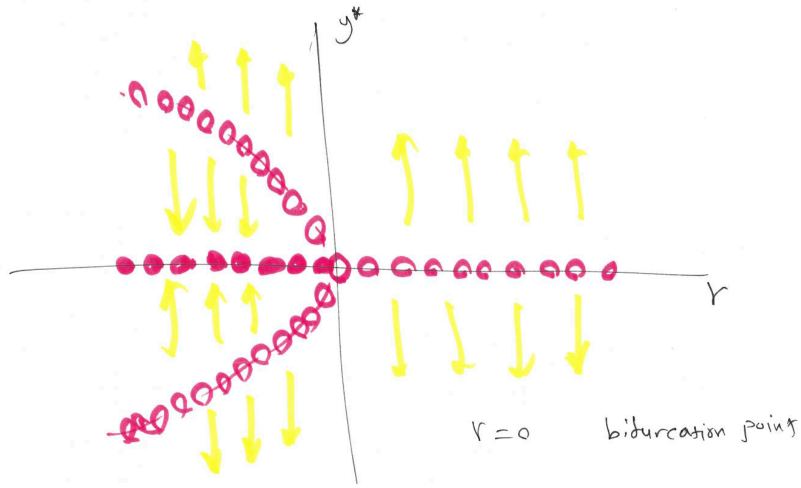

Figure 8.11: The bifurcation diagram (plot of fixed points and their stability vs the tuning parameter r) for subcritical pitchfork bifurcation.

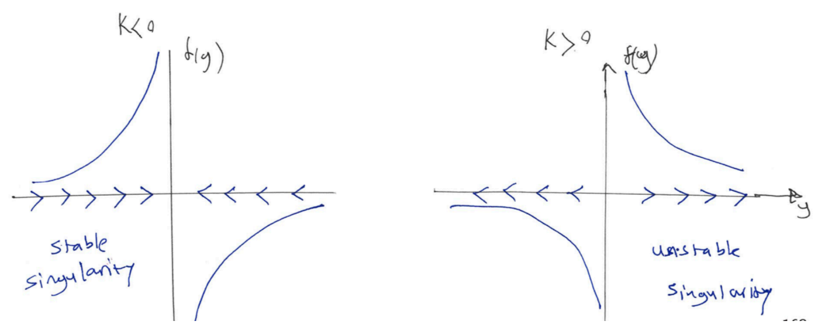

Figure 8.12: The plot of $f(y)$ vs y for different values of k, illustrates the vector field and the stability of the singularity in Example 8.3.

$$\frac{dy}{dt} = \frac{k}{y} \quad \Longrightarrow \quad \int y dy = \int k dt \quad \Longrightarrow \quad y = \pm\sqrt{2kt + y_0^2},$$

where $y_0 = y(t = 0)$ is the inital condition. For $y_0 > 0$, we have the positive solution and for $y_0 < 0$, we have the negative solution. Also, for $k > 0$ the solutions blow up to infinity and for $k < 0$ solutions approach 0. This is summarised in the bifurcation diagram (see Figure 8.13).

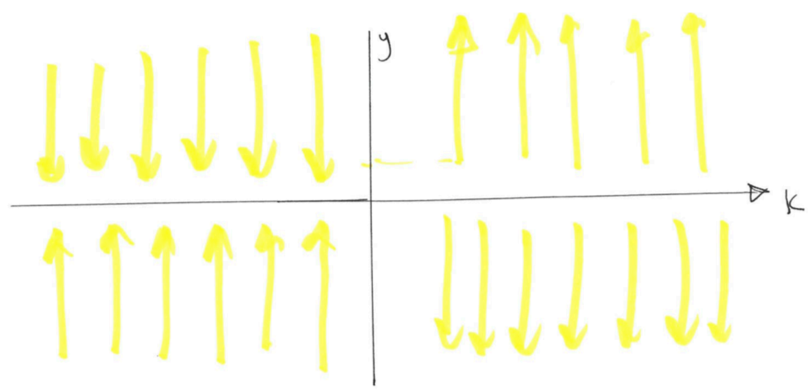

Figure 8.13: The bifurcation diagram (plot of singular points and their stability vs the tuning parameter k) for Example 8.3.

8.2.5 Impossibility of oscillations for one dimensional systems

Fixed points dominate the dynamics of first-order systems. The trajectory in one dimension phase plane never reverses direction and the approach to equilibrium is always monotonic, hence there is no over-shoot, damped oscillations or periodic solutions. This is a topological constraint, if you flow monotonically on a line, you'll never come back to your starting point. Of course, if we were moving on a circle rather than a line, we could eventually return to starting point. For example the following ODE $\frac{d\theta}{dt} = \omega$ for the angle θ has the following periodic solution $\theta = \omega t + \theta_0$.

8.2.6 Linear stability analysis

So far we have relied on graphical methods to determine stability of fixed points (using the flows of the vector field). A quantitative measure can be obtained by linearizing about the fixed point for the one dimensional systems. Let $\eta = y - y^*$ be a small perturbation away from the fixed point y^*. We have using the Taylor expansion of $f(y)$.

$$\frac{dy}{dt} = \frac{d\eta}{dt} = f(y^* + \eta) = f(y^*) + \eta\frac{df}{dy}(y = y^*) + \cdots \quad \Rightarrow \quad \frac{d\eta}{dt} \approx \eta\frac{df}{dy}(y^*).$$

Now, if $\frac{df}{dy}(y^*) > 0$ then the fixed point y^* is unstable as the perturbations around the fixed point grow. However, if $\frac{df}{dy}(y^*) < 0$ then the fixed point y^* is stable as the perturbations around the fixed point decay to zero.

Part III: Introduction to Multivariate Calculus

Chapter 9

Partial Differentiation

So far, in the course, we have considered functions of single independent variables (ordinary functions):

$$f: \quad \mathbb{R} \to \mathbb{R},$$

or in the case of systems of ODEs:

$$f: \quad \mathbb{R} \to \mathbb{R}^n.$$

In the third part of the module, we now consider functions of several variables or *Multivariable or Multivariate functions.*

$$f: \quad \mathbb{R}^n \to \mathbb{R}.$$

For every n-tuple of $\{x_i\}_{i=1}^n$, where $x_i \in \mathbb{R}$, there exists an image in $\mathbb{R}$.

$$f(x_1, \cdots, x_n) \in \mathbb{R}$$

An important example is functions of two variables:

$$f(x, y) = f(\vec{x}); \quad \vec{x} = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$$

9.1 Representation

As seen in Figure 9.1), there are the following two representations for the functions of two variables.

1. 3D representation where $f(x, y)$ is the height.
2. Level curves $\vec{x}_C = (x, y)_C$, where $f(\vec{x}_C) = C$. For each C there will be a set of points that fulfill this condition. This kind of representation is also known as a contour plot.

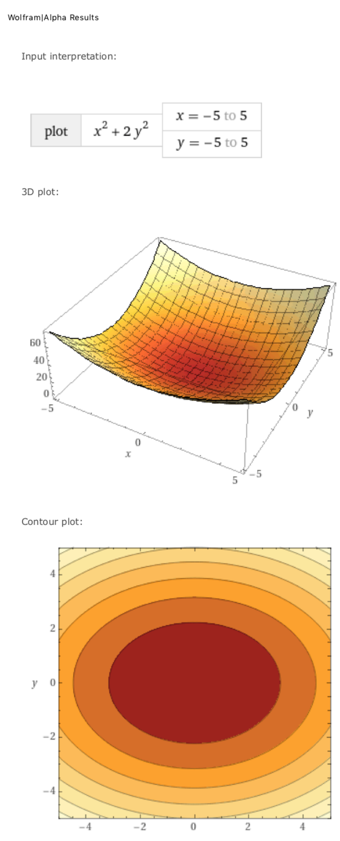

Figure 9.1: The 3D and contour plot of function $f(x, y) = x^2 + 2y^2$ drawn using [Wolfram Alpha](wolframalpha.com)

9.2 Limit and continuity

The general notions from calculus can be naturally extended to multivariate functions.

Limit of the function $f(\vec{x})$ as $\vec{x} \to \vec{x}^*$ exists and is equal C:

$$\lim_{\vec{x} \to \vec{x}^*} f(\vec{x}) = C,$$

if we have $\forall \epsilon > 0, \quad \exists \delta > 0$ so that

$$0 < |\vec{x} - \vec{x^*}| < \delta \quad \Rightarrow \quad |f(\vec{x}) - C| < \epsilon.$$

Function f is continuous at $\vec{x}^*$ if:

$$\lim_{\vec{x} \to \vec{x}^*} f(\vec{x}) = f(\vec{x}^*)$$

For example $f(x,y) = xy$ is continuous at all points in $\mathbb{R}^2$. But, the following function is not continuous at $\vec{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

$$g(x,y) = \begin{cases} xy, & \text{if } x, y \neq 0, \\ 1, & \text{if } x = y = 0. \end{cases}$$

9.3 Partial and Total Differentiation

Different derivatives are defined for functions of several variables. First we introduce partial differentiation, which is differentiation with respect to one of the variables while the other ones are held constant:

$$\frac{\partial f}{\partial x_i} = \lim_{h \to 0} \frac{f(x_1, \cdots, x_i + h, \cdots, x_n) - f(x_1, \cdots, x_n)}{h}$$

This is read as partial derivative of f with respect to x_i, it is also sometimes denoted alternatively as f_{x_i}, f'_{x_i}, $\partial_{x_i} f$ and D_{x_i}.

Higher order partial derivatives can also be defined. For example, consider a function of two variable $f(x,y)$, we denote the first partial derivatives as:

$$g_1(x,y) = \left(\frac{\partial f}{\partial x}\right)_y ; \quad g_2(x,y) = \left(\frac{\partial f}{\partial y}\right)_x$$

The subscript of y and x in each of these partial deriveatives, highlights which variable is held constant. We have the following second order partial derivatives for $f(x,y)$:

$$\left(\frac{\partial g_1}{\partial x}\right)_y = \frac{\partial^2 f}{\partial x^2}; \quad \left(\frac{\partial g_1}{\partial y}\right)_x = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$

$$\left(\frac{\partial g_2}{\partial y}\right)_x = \frac{\partial^2 f}{\partial y^2}; \quad \left(\frac{\partial g_2}{\partial x}\right)_y = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$$

Symmetry of mixed derivatives or equality of mixed derivatives: If the second partial derivatives are continuous, the order of differentiation is not important and we therefore have:

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \quad \Rightarrow \quad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}.$$

This result is known as Schwarz's theorem, Clairaut's theorem, or Young's theorem.

Operationally, calculations are simple. Partial derivatives are obtained by keeping the other variables constant, using the laws of differentiation for functions of single variables.

Example 9.1 (Obtain all the first and second partial derivatives of the following function.)**.**

$$u(x, y) = x^2 \sin y + y^3.$$

We have for the first partial derivatives:

$$\left(\frac{\partial u}{\partial x}\right)_y = 2x \sin y,$$

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

And for the second partial derivatives, we have:

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

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

$$\frac{\partial^2 u}{\partial y^2} = -x^2 \sin y + 6y,$$

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

We observe that the symmetry of the mixed derivatives in this example holds.

9.4 Total differentiation of a function of several variables

Total derivative evaluates the infinitesimal change of $f(\vec{x})$ when all the variables are allowed to change infinitesimally in contrast with partial derivatives that are about change in only one of the variables. We first consider the case of a function of two variables. We have

$$\begin{aligned}\Delta f &= f(x+\Delta x, y+\Delta y) - f(x,y)\\ &= f(x+\Delta x, y+\Delta y) - f(x,y) + f(x+\Delta x, y) - f(x+\Delta x, y)\\ &= \left[\frac{f(x+\Delta x, y) - f(x,y)}{\Delta x}\right]\Delta x + \left[\frac{f(x+\Delta x, y+\Delta y) - f(x+\delta x, y)}{\Delta y}\right]\Delta y.\end{aligned}$$

The total derivative is obtained at the limit of $\Delta x, \Delta y \to 0$:

$$df = \lim_{\Delta x, \Delta y \to 0} \Delta f = \left(\frac{\partial f}{\partial x}\right)_y dx + \left(\frac{\partial f}{\partial y}\right)_y dy.$$

For function of several variables $f(x_1, \cdots, x_n)$, total differentiation is generalized. We have:

$$df = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right) dx_i.$$

9.5 Chain rule for functions of several variables

Given ordinary functions $u(x)$ and $x(t)$, chain rule for ordinary functions is recalled to be:

$$\frac{du}{dt} = \frac{du}{dx}\frac{dx}{dt}.$$

What is the equivalent for multivariable functions? Consider a function of two variables:

$$u = u(x,y) = u(\vec{x}); \quad \vec{x} \in \mathbb{R}^2.$$

Now, if we have $x = x(t)$ and $y = y(t)$ with $t \in \mathbb{R}$. What is $\frac{du}{dt}$?

Combining total differentiation and the chain rule for ordinary functions, one can obtain:

$$\begin{aligned}du &= \left(\frac{\partial u}{\partial x}\right)_y dx + \left(\frac{\partial u}{\partial y}\right)_y dy,\\ &= \left(\frac{\partial u}{\partial x}\right)_y \left(\frac{dx}{dt}\right) dt + \left(\frac{\partial u}{\partial y}\right)_y \left(\frac{dy}{dt}\right) dt.\end{aligned}$$

So, we have:

$$\frac{du}{dt} = \left(\frac{\partial u}{\partial x}\right)_y \left(\frac{dx}{dt}\right) + \left(\frac{\partial u}{\partial y}\right)_y \left(\frac{dy}{dt}\right).$$

This generalises to a function of n variables $u(x_1, \cdots, x_n)$ with $x_i = x_i(t)$ and $t \in \mathbb{R}$:

$$\frac{du}{dt} = \sum_{i=1}^{n} \left(\frac{\partial u}{\partial x_i}\right) \frac{dx_i}{dt}.$$

Example 9.2 (Consider a cylinder that its radius and height are expanding with time)**.**

$$r(t) = 2t; \qquad h(t) = 1 + t^2.$$

Evaluate the rate of change in volume $\dfrac{dV}{dt}$.

We have $V = \pi r^2 h$, therefore

$$\frac{dV}{dt} = \left(\frac{\partial V}{\partial r}\right)_h \frac{dr}{dt} + \left(\frac{\partial V}{\partial h}\right)_r \frac{dh}{dt} = 2\pi r(2h + rt) = 8\pi t + 16\pi t^3.$$

Another example of chain rule when we have multiple dependencies. Consider

$$u = u(x, y); \quad \text{with} \quad y = y(t, x).$$

To obtain $\left(\frac{\partial u}{\partial x}\right)_t$, we combine total differentiation for $u(x, y)$ and $y(t, x)$. We get:

$$du = \left(\frac{\partial u}{\partial x}\right)_y dx + \left(\frac{\partial u}{\partial y}\right)_x dy,$$

$$dy = \left(\frac{\partial y}{\partial x}\right)_t dx + \left(\frac{\partial y}{\partial t}\right)_x dt.$$

Now, by plugging dy in the expression for du and rearranging we get:

$$du = \left[\left(\frac{\partial u}{\partial x}\right)_y + \left(\frac{\partial u}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_t\right] dx + \left(\frac{\partial u}{\partial y}\right)_x \left(\frac{\partial y}{\partial t}\right)_x dt.$$

Now, thinking of the above expression as the total derivative of $u(x, t)$, we obtain:

$$\left(\frac{\partial u}{\partial x}\right)_t = \left(\frac{\partial u}{\partial x}\right)_y + \left(\frac{\partial u}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_t,$$

$$\left(\frac{\partial u}{\partial t}\right)_x = \left(\frac{\partial u}{\partial y}\right)_x \left(\frac{\partial y}{\partial t}\right)_x.$$

Dependencies on another set of coordinates

Let

$$h = h(x, y); \quad \text{with} \quad x = x(u, v) \quad \text{and} \quad y = y(u, v).$$

Considering total derivative of $h(x, y)$ and substituting total derivatives of $x(u, v)$ and $y(u, v)$, and rearranging, we obtain:

$$dh = \left[\left(\frac{\partial h}{\partial x}\right)_y \left(\frac{\partial x}{\partial u}\right)_v + \left(\frac{\partial h}{\partial y}\right)_x \left(\frac{\partial y}{\partial u}\right)_v\right] du + \left[\left(\frac{\partial h}{\partial x}\right)_y \left(\frac{\partial x}{\partial v}\right)_u + \left(\frac{\partial h}{\partial y}\right)_x \left(\frac{\partial y}{\partial v}\right)_u\right] dv.$$

Now thinking of this expression as total derivative of $h(u, v)$, we have:

$$\left(\frac{\partial h}{\partial u}\right)_v = \left(\frac{\partial h}{\partial x}\right)_y \left(\frac{\partial x}{\partial u}\right)_v + \left(\frac{\partial h}{\partial y}\right)_x \left(\frac{\partial y}{\partial u}\right)_v,$$

$$\left(\frac{\partial h}{\partial v}\right)_u = \left(\frac{\partial h}{\partial x}\right)_y \left(\frac{\partial x}{\partial v}\right)_u + \left(\frac{\partial h}{\partial y}\right)_x \left(\frac{\partial y}{\partial v}\right)_u.$$

Note, that strictly speaking the transformed function $h(x, y)$ should be denoted as $h'(u, v)$ as the transformed function is a 'different' function of its variables. But, very commonly, the prime on $h'(u, v)$ is not used. For example $h(x, y) = x^2 + y^2$ in polar coordinates is $h'(r,' \theta) = r^2$, while common notation of $h(r, \theta)$ could imply $r^2 + \theta^2$, if one thinks of plugging r and θ in the original function $h(x, y)$. One should be aware of this notational ambiguity.

9.6 Implicit functions

First, a reminder about the explicit form for an ordinary function:

$$y = f(x); \quad x \in \mathbb{R}$$

The implicit form for an ordinary function is

$$F(x, y) = 0.$$

Trivially, if we have the explicit form we also have an implicit form:

$$F(x, y) = y - f(x) = 0.$$

For functions of two variables, we also have explicit form:

$$z = z(x, y)$$

And implicit form:

$$F(x, y, z) = 0$$

Differentiation using the Implicit form

Taking total differential from the implicit form $F(x, y, z) = 0$, we obtain:

$$dF = \left(\frac{\partial F}{\partial x}\right)_{y,z} dx + \left(\frac{\partial F}{\partial y}\right)_{x,z} dy + \left(\frac{\partial F}{\partial z}\right)_{x,y} dz = 0.$$

Taking total differential from the explicit form $z = z(x, y)$, we obtain:

$$dz = \left(\frac{\partial z}{\partial x}\right)_{y} dx + \left(\frac{\partial z}{\partial y}\right)_{x} dy.$$

Now, solving for dz in the dF equation above, we thus have the following relationship between derivatives of the implicit and explicit form:

$$\left(\frac{\partial z}{\partial x}\right)_{y} = -\frac{\left(\frac{\partial F}{\partial x}\right)_{y,z}}{\left(\frac{\partial F}{\partial z}\right)_{x,y}},$$

$$\left(\frac{\partial z}{\partial y}\right)_{x} = -\frac{\left(\frac{\partial F}{\partial y}\right)_{x,z}}{\left(\frac{\partial F}{\partial z}\right)_{x,y}}.$$

Example 9.3 (Obtain the partial derivatives of z using the explicit and implicit forms)**.**

$$\text{Let} \quad z(x, y) = x^2 + y^2 - 5.$$

We have from the explicit form:

$$dz = 2xdx + 2ydy.$$

Using the implicit form $F(x, y, z) = x^2 + y^2 - 5 - z$, we have:

$$dF = 2xdx + 2ydy - dz = 0.$$

Which results to the same total derivative for dz, that we obtained above from the explicit form.

9.7 Taylor expansion of multivariate functions

Taylor expansion for functions of one variable (reminder): Let $f(x) : \mathbb{R} \to \mathbb{R}$ and consider $x_0 \in \mathbb{R}$. We saw in the first term of the module that:

$$f(x_0 + \Delta x) = f(x_0) + \left(\frac{df}{dx}\right)_{x_0} \Delta x + \frac{1}{2}\left(\frac{d^2 f}{dx^2}\right)_{x_0} (\Delta x)^2 + \frac{1}{3!}\left(\frac{d^3 f}{dx^3}\right)_{x_0} (\Delta x)^3 + \cdots.$$

Now, let us consider $f(\vec{x})$, $\vec{x} \in \mathbb{R}^2$, we assume suitable conditions of differentiability. We can use Taylor expansion for ordinary functions first on the x direction and then y to obtain the Taylor expansion for $f(x, y)$. Up to to 3rd order we have:

$$
\begin{aligned}
&f(\vec{x_0} + \Delta\vec{x}) = f(x_0 + \Delta x, y_0 + \Delta y) \\
&= f(x_0, y_0 + \Delta y) + \left(\frac{\partial f}{\partial x}\right)_{x_0, y_0+\Delta y} \Delta x + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\right)_{x_0, y_0+\Delta y} (\Delta x)^2 + \frac{1}{3!}\left(\frac{\partial^3 f}{\partial x^3}\right)_{x_0, y_0+\Delta y} (\Delta x)^3 + \cdots \\
&= f(x_0, y_0) + \left(\frac{\partial f}{\partial y}\right)_{\vec{x}_0} \Delta y + \frac{1}{2}\left(\frac{\partial^2 f}{\partial y^2}\right)_{\vec{x}_0} (\Delta y)^2 + \frac{1}{3!}\left(\frac{\partial^3 f}{\partial y^3}\right)_{\vec{x}_0} (\Delta y)^3 + \cdots \\
&+ \Delta x \left[\left(\frac{\partial f}{\partial x}\right)_{\vec{x}_0} + \left(\frac{\partial^2 f}{\partial y \partial x}\right)_{\vec{x}_0} \Delta y + \frac{1}{2}\left(\frac{\partial^3 f}{\partial y^2 \partial x}\right)_{\vec{x}_0} (\Delta y)^2 + \cdots\right] \\
&+ \frac{1}{2}(\Delta x)^2 \left[\left(\frac{\partial^2 f}{\partial x^2}\right)_{\vec{x}_0} + \left(\frac{\partial^3 f}{\partial y \partial x^2}\right)_{\vec{x}_0} \Delta y + \cdots\right] + \frac{1}{3!}(\Delta x)^3 \left[\left(\frac{\partial^3 f}{\partial x^3}\right)_{\vec{x}_0} + \cdots\right] \\
&= f(\vec{x}_0) + \left[\left(\frac{\partial f}{\partial x}\right)_{\vec{x}_0} \Delta x + \left(\frac{\partial f}{\partial y}\right)_{\vec{x}_0} \Delta y\right] + \\
&\frac{1}{2!}\left[\left(\frac{\partial^2 f}{\partial x^2}\right)_{\vec{x}_0} (\Delta x)^2 + 2\left(\frac{\partial^2 f}{\partial x \partial y}\right)_{\vec{x}_0} \Delta x \Delta y + \left(\frac{\partial^2 f}{\partial y^2}\right)_{\vec{x}_0} (\Delta y)^2\right] \\
&+ \frac{1}{3!}\left[\left(\frac{\partial^3 f}{\partial x^3}\right)_{\vec{x}_0} (\Delta x)^3 + 3\left(\frac{\partial^3 f}{\partial x^2 \partial y}\right)_{\vec{x}_0} (\Delta x)^2 \Delta y + 3\left(\frac{\partial^3 f}{\partial x \partial y^2}\right)_{\vec{x}_0} \Delta x (\Delta y)^2 + \left(\frac{\partial^3 f}{\partial y^3}\right)_{\vec{x}_0} (\Delta y)^3\right] \\
&+ \cdots.
\end{aligned}
$$

We can write the Taylor expansion up to the second order in a vector-matrix form. We define *Gradient* of the function f evaluated at point $\vec{x}_0$ as:

$$
\vec{\nabla} f_{\vec{x}_0} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \\ \frac{\partial f}{\partial y} \end{bmatrix}_{\vec{x}_0}.
$$

Hessian Matrix associated with the function f evaluated at the point $\vec{x}_0$ is defined as:

$$
H_{ij}(\vec{x}_0) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{\vec{x}_0}
$$

We can write the Taylor expansion up to the second order in terms of the Gradient and Hessian:

$$
f(\vec{x}_0 + \Delta\vec{x}) = f(\vec{x}_0) + \vec{\nabla} f(\vec{x}_0)^T . \Delta\vec{x} + \frac{1}{2}\Delta\vec{x}^T H(\vec{x}_0)\Delta\vec{x} + \cdots
$$

This generalizes to functions of n dimensions.

Example 9.4 (Approximation)**.**

$$\text{Let} \quad A(x, y) = xy.$$

Expand A around $\vec{x}_0 = (x_0, y_0)$.

$$\begin{aligned} A(\vec{x}_0 + \vec{\Delta x}) &= A(\vec{x}_0) + \begin{bmatrix} y_o & x_0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \Delta x & \Delta y \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} + \cdots \\ &= x_0 y_0 + (y_0 \Delta x + x_0 \Delta y) + \Delta x \Delta y. \end{aligned}$$

Example 9.5 (Using Taylor Expansion for error analysis)**.** What is the maximum error in h given errors in x and θ of Δx and $\Delta\theta$, respectively:

$$h(x, \theta) = x \tan\theta.$$

We have $x = x_0 \pm \Delta x$ and $\theta = \theta_0 \pm \Delta\theta$, we are looking for Δh, using the Taylor expansion of h up to the first order we have:

$$h(x_0 \pm \Delta x, \theta_0 \pm \Delta\theta) = h(x_0, y_0) \pm \left(\frac{\partial h}{\partial x}\right)_{\vec{x}_0} \Delta x \pm \left(\frac{\partial h}{\partial \theta}\right)_{\vec{x}_0} \Delta\theta + \cdots.$$

So for maximum error we have:

$$|\Delta h| = |\tan\theta_0||\Delta x| + |x_0 \sec\theta_0{}^2||\Delta\theta|.$$

For relative error we have:

$$\left|\frac{\Delta h}{h(\vec{x}_0)}\right| = \left|\frac{\Delta x}{x_0}\right| + \left|\frac{2\Delta\theta}{\sin 2\theta_0}\right|.$$

Chapter 10

Applications of Multivariate Calculus

In Chapter 9, we introduced multivariable functions and notions of differentiation. In this chapter, we present several applications of multivariate calculus.

10.1 Change of Coordinates

In lots of situations, one may need a change of coordinates, which in general could be a nonlinear transformation. In this section, we focus on the familiar example of the change of coordinates from polar coordinates ($\vec{x} = (r, \theta)$) to cartesian ($\vec{x} = (x, y)$) and vice versa to illustrate the concepts. We generalise these results at the end of the section.

Let $x = x(r, \theta)$ and $y = y(r, \theta)$, we have:

$$x = r\cos\theta, \quad y = r\sin\theta.$$

Conversely, we have for $r = r(x, y)$ and $\theta = \theta(x, y)$:

$$r = \sqrt{x^2 + y^2}, \quad \theta = \arctan\left(\frac{y}{x}\right).$$

Using total differentiation, we can obtain a relationship between the vectors of infinitesimal change $\vec{dx}$ and $\vec{dr}$.

$$dx = \left(\frac{\partial x}{\partial r}\right)_\theta dr + \left(\frac{\partial x}{\partial \theta}\right)_r d\theta,$$
$$dy = \left(\frac{\partial y}{\partial r}\right)_\theta dr + \left(\frac{\partial y}{\partial \theta}\right)_r d\theta.$$

We can write this in vector-matrix form as

$$\begin{bmatrix} dx \\ dy \end{bmatrix} = \begin{bmatrix} \left(\frac{\partial x}{\partial r}\right)_\theta & \left(\frac{\partial x}{\partial \theta}\right)_r \\ \left(\frac{\partial y}{\partial r}\right)_\theta & \left(\frac{\partial y}{\partial \theta}\right)_r \end{bmatrix} \begin{bmatrix} dr \\ d\theta \end{bmatrix}.$$

By defining matrix J known as Jacobian of the transformation as below we can write this relationship as $\vec{dx} = J\vec{dr}$ in short.

$$J = \begin{bmatrix} \left(\frac{\partial x}{\partial r}\right)_\theta & \left(\frac{\partial x}{\partial \theta}\right)_r \\ \left(\frac{\partial y}{\partial r}\right)_\theta & \left(\frac{\partial y}{\partial \theta}\right)_r \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}.$$

Similarly for the polar to cartesian transformation we have:

$$dr = \left(\frac{\partial r}{\partial x}\right)_y dx + \left(\frac{\partial r}{\partial y}\right)_x dy$$
$$d\theta = \left(\frac{\partial \theta}{\partial x}\right)_y dx + \left(\frac{\partial \theta}{\partial y}\right)_x dy$$

In vector-matrix form we can writh this using matrix K defined below as $\vec{dr} = K\vec{dx}$:

$$K = \begin{bmatrix} \left(\frac{\partial r}{\partial x}\right)_y & \left(\frac{\partial r}{\partial y}\right)_x \\ \left(\frac{\partial \theta}{\partial x}\right)_y & \left(\frac{\partial \theta}{\partial y}\right)_x \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\frac{\sin\theta}{r} & \frac{\cos\theta}{r} \end{bmatrix}.$$

It is evident that K, the Jacobian of the transformation from polar to cartersian is equal to the inverse of J, the Jacobian of the transformation from cartesian to polar.

$$K = J^{-1}.$$

Application 1: Infinitesimal element of length

Consider the curve $y(x)$, the infinitesimal element of length ds, along this curve is

$$ds^2 = dx^2 + dy^2$$

What is the infinitesimal element of length in polar coordinates? Given we have $\vec{dx} = J\vec{dr}$, we have:

$$ds^2 = (dx)^2 + (dy)^2 = \begin{bmatrix} dx & dy \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix} = \begin{bmatrix} dr & d\theta \end{bmatrix} J^T J \begin{bmatrix} dr \\ d\theta \end{bmatrix} = (dr)^2 + r^2(d\theta)^2.$$

Application 2: Infinitesimal element of area in polar coordinate

Infinitesimal element of area is useful in taking integrals of functions of two variables ($f(x,y)$) over a domain in the (x,y) plane. In cartesian coordinates we have:

$$dA = dxdy.$$

We note that letting $\vec{dx} = dx\hat{i}$ and $\vec{dy} = dy\hat{j}$ and using the fact that the area of a parallelogram is equal to the magnitude of the cross-product of vectors spanning its edges, we can also write dA as:

$$dA = \|\vec{dx} \times \vec{dy}\| = \|\det \begin{bmatrix} \hat{i} & \hat{j} & \hat{k} \\ dx & 0 & 0 \\ 0 & dy & 0 \end{bmatrix}\|.$$

What is area element for a general transformation? Consider a general transformation in two-dimensions:

$$x = x(u,v) \quad \text{and} \quad y = y(u,v).$$

We have:

$$\begin{bmatrix} dx \\ dy \end{bmatrix} = J \begin{bmatrix} du \\ dv \end{bmatrix},$$

where J is the Jacobian of the transformation. We therefore have for $\vec{du}$ and $\vec{dv}$ in cartesian coordinates:

$$\vec{du} = J \begin{bmatrix} du \\ 0 \end{bmatrix}, \quad \vec{dv} = J \begin{bmatrix} 0 \\ dv \end{bmatrix}.$$

Using the cross-product rule for the area of the parallelogram, we have:

$$dA' = \|\vec{du} \times \vec{dv}\| = |\det J| dudv.$$

For the polar coordinate we obtain:

$$dA' = |\det J| drd\theta = rdrd\theta.$$

Note that some texts do not use prime on the transformed area element, although dA and dA' are not mathematically equal (one can easily check that $dxdy \neq rdrd\theta$). However, there is a conceptual equivalence between dA and dA'.

This result generalizes to higher dimensions for a volume element. Given a general transformation u in n dimensions we have:

$$\vec{dx}_{n\times 1} = J_{n\times n} \vec{du}_{n\times 1},$$

where J is the Jacobian of the transformation. The infinitesimal volume element in cartesian coordinates is $dV = \prod_{i=1}^{n} dx_i$. For the volume element in the transformed coordinates we have

$$dV' = |\det J| \prod_{i=1}^{n} du_i.$$

10.2 Partial Differential Equations (PDEs)

Analogous to an ordinary differential equation (ODE), one can define a partial differential equation or PDE with $f(\vec{x})$, $\vec{x} \in \mathbb{R}^n$ satisfying:

$$f(x_1, \cdots, x_n, f, \frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n}, \frac{\partial^2 f}{\partial x_i x_j}, \cdots) = 0.$$

Consider the following 2 dimensional examples. 1. Laplace Equation for $u(x, y)$ (relevant to multiple areas of physics including fluid dynamics):

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

2. Wave Equation for $u(x, t)$ (describing the dynamics of a wave with speed c in one spatial dimension and time):

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial y^2} = 0.$$

Our discussion of PDEs here will be very brief, but this topic is a major part of your multivariable calculus course and one of the applied elective courses in the second year. If you would like to have a sneak preview and for some cool connections to the Fourier series you saw last term, you can check out this video.

Transforming a PDE under a change of coordinates

We again consider the example of transformation from cartesian to polar coordinates:

$$u(x, y) \quad \longleftrightarrow \quad u(r, \theta),$$

with J being Jacobian of the transformation. Using total differentiation we have:

$$\begin{aligned}
du &= \left(\frac{\partial u}{\partial x}\right)_y dx + \left(\frac{\partial u}{\partial y}\right)_y dy \\
&= \left(\frac{\partial u}{\partial x}\right)_y \left[\left(\frac{\partial x}{\partial r}\right)_\theta dr + \left(\frac{\partial x}{\partial \theta}\right)_r d\theta\right] + \left(\frac{\partial u}{\partial y}\right)_y \left[\left(\frac{\partial y}{\partial r}\right)_\theta dr + \left(\frac{\partial y}{\partial \theta}\right)_r d\theta\right] \\
&= \left[\left(\frac{\partial u}{\partial x}\right)_y \left(\frac{\partial x}{\partial r}\right)_\theta + \left(\frac{\partial u}{\partial y}\right)_y \left(\frac{\partial y}{\partial r}\right)_\theta\right] dr + \left[\left(\frac{\partial u}{\partial x}\right)_y \left(\frac{\partial x}{\partial \theta}\right)_r + \left(\frac{\partial u}{\partial y}\right)_y \left(\frac{\partial y}{\partial \theta}\right)_r\right] d\theta.
\end{aligned}$$

Now, by equating the terms in the above and the total derivative of $u(r, \theta)$, and also using the definition of J, we obtain the following:

$$\begin{bmatrix} \left(\frac{\partial}{\partial r}\right)_\theta \\ \left(\frac{\partial}{\partial \theta}\right)_r \end{bmatrix} u(r, \theta) = J^T \begin{bmatrix} \left(\frac{\partial}{\partial x}\right)_y \\ \left(\frac{\partial}{\partial y}\right)_x \end{bmatrix} u(x, y).$$

Example 10.1 (Laplace Equation in polar coordinates)**.**

Laplace equation in Cartesian coordinates is:

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

In polar coordinates we have:

$$\begin{aligned}\frac{\partial}{\partial x}[u] &= \frac{\partial u}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial u}{\partial \theta}\frac{\partial \theta}{\partial x} \\ &= \left[\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right]u.\end{aligned}$$

Now taking second derivative we have:

$$\begin{aligned}\frac{\partial^2}{\partial x^2}[u] &= \left[\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right]\left[\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right]u \\ &= \cos^2\theta\frac{\partial^2 u}{\partial r^2} + \frac{2\cos\theta\sin\theta}{r^2}\frac{\partial u}{\partial \theta} - \frac{2\cos\theta\sin\theta}{r}\frac{\partial^2 u}{\partial r\partial\theta} + \frac{\sin^2\theta}{r}\frac{\partial u}{\partial r} + \frac{\sin^2\theta}{r^2}\frac{\partial^2 u}{\partial \theta^2}.\end{aligned}$$

Similarly, we have:

$$\frac{\partial^2}{\partial y^2}[u] = \sin^2\theta\frac{\partial^2 u}{\partial r^2} - \frac{2\cos\theta\sin\theta}{r^2}\frac{\partial u}{\partial \theta} + \frac{2\cos\theta\sin\theta}{r}\frac{\partial^2 u}{\partial r\partial\theta} + \frac{\cos^2\theta}{r}\frac{\partial u}{\partial r} + \frac{\cos^2\theta}{r^2}\frac{\partial^2 u}{\partial \theta^2}.$$

Finally, we have:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = 0.$$

For example, if one looking for function $u(r)$ (with no dependence on θ) fulfilling the Laplace equation, we could solve the following ODE:

$$\frac{d^u}{dr^2} + \frac{1}{r}\frac{du}{dr} = 0.$$

10.3 Exact ODEs

The concept of total differentiation provides an alternative method for solving first order nonlinear ODEs. Consider the given first order ODE:

$$\frac{dy}{dx} = \frac{-F(x,y)}{G(x,y)}.$$

If we have a solution of the ODE in implicit form $u(x, y) = 0$, assuming u is continuous with continuous derivatives. For total derivative of u we have:

$$du = \left(\frac{\partial u}{\partial x}\right)_y dx + \left(\frac{\partial u}{\partial y}\right)_y dy = 0 \quad \Rightarrow \quad \frac{dy}{dx} = \frac{-\left(\frac{\partial u}{\partial x}\right)_y}{\left(\frac{\partial u}{\partial y}\right)_y}.$$

For the solution $u(x, y) = 0$ to exist then we need the RHS of the above equation to be equal to the RHS of the ODE, we are trying to solve. This will be the case if

$$F = \left(\frac{\partial u}{\partial x}\right)_y \quad \text{and} \quad G = \left(\frac{\partial u}{\partial y}\right)_y.$$

But, if that is the case, due to the symmetry of the partial second mixed derivatives, we should have:

$$\frac{\partial u^2}{\partial y \partial x} = \frac{\partial u^2}{\partial x \partial y} \quad \Rightarrow \quad \frac{\partial F}{\partial y} = \frac{\partial G}{\partial x}.$$

This is known as the condition of integrability of the ODE and if it is satisfied then there exists $u(x, y)$, where

$$F = \left(\frac{\partial u}{\partial x}\right)_y \quad \text{and} \quad G = \left(\frac{\partial u}{\partial y}\right)_y,$$

and then $u(x, y) = 0$ is a solution of the first order ODE. We call this kind of ODE *exact*.

Example 10.2 (Is the following ODE exact?)**.**

$$\frac{dy}{dx} = \frac{-2xy - \cos x \cos y}{x^2 - \sin x \sin y}$$

Letting

$$F(x, y) = 2xy + \cos x \cos y \quad \text{and} \quad G(x, y) = x^2 - \sin x \sin y.$$

We can check the condition of integrability and see the ODE is exact:

$$\frac{\partial F}{\partial y} = \frac{\partial G}{\partial x} = 2x - \cos x \sin y,$$

since the ODE is exact, we can look for a solution in implicit form $u(x, y) = 0$ such that:

$$F(x, y) = \frac{\partial u}{\partial x} \quad \text{and} \quad G(x, y) = \frac{\partial u}{\partial y}.$$

We can do that in two steps. Firstly, we have:

$$\left(\frac{\partial u}{\partial x}\right)_y = 2xy + \cos x \cos y$$

Integrating with respect to x and assuming the constant of intgration could depend on y, we obtain:

$$u = yx^2 + \cos y \sin x + f(y).$$

Now in the second step, we use

$$G(x,y) = \frac{\partial u}{\partial y} = x^2 - \sin y \sin x \quad \Rightarrow \quad \frac{df}{dy} = 0.$$

This implies f is a constant, so we obtain the general solution $y(x)$ of the ODE in implicit form:

$$u(x,y) = yx^2 + \cos y \sin x + c = 0.$$

When the ODE is not exact, sometimes we can find a function (an integrating factor) that will make the equation exact. Given:

$$F(x,y)dx + G(x,y)dy = 0,$$

is not exact, we look for a function $\lambda(x)$ or $\lambda(y)$ such that:

$$\lambda F(x,y)dx + \lambda G(x,y)dy = 0,$$

is exact. Note that an integrating factor can in general be a function of both x and y, but in this case we cannot find an explicit solution for λ, and it is for this reason we can not solve very many ODEs.

Example 10.3 (Is this ODE exact? If not find an integrating factor to make it exact)**.**

$$\frac{dy}{dx} = \frac{xy-1}{x(y-x)}.$$

Letting $F = xy - 1$ and $G = x^2 - xy$, we see that the ODE is not exact as:

$$\frac{\partial F}{\partial y} \neq \frac{\partial G}{\partial x}.$$

So, we will try to find a $\lambda(x)$ (or $\lambda(y)$ that will make the ODE exact). We need to find $\lambda(x)$ such that:

$$\frac{\partial[\lambda(x)(xy-1)]}{\partial y} = \frac{\partial[\lambda(x)(x^2-xy)]}{\partial x}.$$

After simplification we obtain:

$$(x-y)\left[\frac{d\lambda}{dx}x+\lambda\right]=0.$$

As we obtain an ODE for $\lambda(x)$ (that does not depend on y), an integrating factor of the form $\lambda(x)$ exists and can be obtained by solving this ODE to be:

$$\lambda=\frac{c}{x}.$$

Now we can solve the exact ODE that is obtained (left as a quiz):

$$(y-\frac{1}{x})dx+(x-y)dy=0.$$

10.4 Sketching functions of two variables

Similar to the sketching of functions of one variable, we will use the following steps in sketching a function of two variables $f(x,y)$:

- Check continuity and find singularities.
- Find asymptotic behaviour

$$\lim_{x,y\to\pm\infty} f(x,y) \quad \text{and} \quad \lim_{\vec{x}\to\vec{x}_{sing}} f(\vec{x}).$$

- Obtain some level curves, for example $f(\vec{x})=0$.
- Find stationary points: minimum, maximum, saddle points

Stationary points for functions of two variables

Reminder that for a function of one variable $f(x)$, we find stationary points by setting the first derivative to zero:

$$f(x^*)=0.$$

Then, using Taylor expansion of $f(x)$ near x^*, we see that one can use the sign of the second derivative of $f(x)$ at x^* to decide if the stationary point is minimum (if $\frac{d^2f}{dx^2}(x^*)>0$) or maximum (if $\frac{d^2f}{dx^2}(x^*)<0$).

Using a similar approach for the functions of two variables $f(x,y)$, we have stationary points are the points where tangent plane at x^* is parallel to (x,y) plane:

$$\frac{\partial f}{\partial x}(\vec{x}^*)=\frac{\partial f}{\partial y}(\vec{x}^*)=0.$$

Then, the type of stationary point can be determined using the Taylor expansion around the stationary point $\vec{x^*}$ and by the Hessian matrix.

$$f(\vec{x}^*+\Delta\vec{x})=f(\vec{x}^*)+\begin{bmatrix}\frac{\partial f}{\partial x}(\vec{x}^*) & \frac{\partial f}{\partial y}(\vec{x}^*)\end{bmatrix}\Delta\vec{x}+\frac{1}{2}\Delta\vec{x}^T\begin{bmatrix}\frac{\partial^2 f}{\partial^2 x}(\vec{x}^*) & \frac{\partial^2 f}{\partial x\partial y}(\vec{x}^*)\\ \frac{\partial^2 f}{\partial y\partial x}(\vec{x}^*) & \frac{\partial^2 f}{\partial^2 y}(\vec{x}^*)\end{bmatrix}\Delta\vec{x}.$$

Given the fact that the first partial derivatives of $f(\vec{x})$ are zero at the stationary points, we have:

$$f(\vec{x}^* + \Delta\vec{x}) - f(\vec{x}^*) = \frac{1}{2}\Delta\vec{x}^T H(\vec{x}^*)\Delta\vec{x}.$$

Under the assumption of continuity, we have the symmetry of the second mixed partial derivatives:

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

therfore the Hessain is symmetric ($H = H^T$), which implies that H is diagonalizable and it has real eigenvalues λ_1 and λ_2. So, there exists a similarity transformation V that diagonalise the Hessian:

$$V^{-1}HV = \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

Also as H is symmetric, we have $V^{-1} = V^T$. So, we have:

$$f(\vec{x}^* + \Delta\vec{x}) - f(\vec{x}^*) = \frac{1}{2}\Delta\vec{x}^T \left[V((\vec{x}^*)\Lambda(\vec{x}^*)V^T(\vec{x}^*)\right]\Delta\vec{x}.$$

Now, if we let $\Delta\vec{z} = V^T(\vec{x}^*)\Delta\vec{x}$, we have:

$$\Delta f = f(\vec{x}^* + \Delta\vec{x}) - f(\vec{x}^*) = \frac{1}{2}\Delta\vec{z}^T\Lambda(\vec{x}^*)\Delta\vec{z} = \frac{1}{2}\left[(\Delta z_1)^2\lambda_1 + (\Delta z_2)^2\lambda_2\right].$$

Given this expression, we can use the sign of the eigenvalues to classify the stationary points.

- If $\lambda_1, \lambda_2 \in \mathbb{R}^+$, we have $\Delta f > 0$ as we move away from the stationary point, suggesting $\vec{x}^*$ is a minimum.
- If $\lambda_1, \lambda_2 \in \mathbb{R}^-$, we have $\Delta f < 0$ as we move away from the stationary point, suggesting $\vec{x}^*$ is a maximum.
- If $\lambda_1 \in \mathbb{R}^+$ and $\lambda_2 \in \mathbb{R}^-$, we classify $\vec{x}^*$ is a saddle point, since Δf can be positive or negative depending the direction of $\Delta\vec{x}$.

We note that, we could use the trace (τ) and determinant Δ of the matrix H to know the sign of the eigenvalues with explicityly calculating the eigenvalues as done in section 7.3 in the analysis of the 2D linear ODEs. In particular, $\Delta > 0, \tau > 0$ ($\Delta > 0, \tau < 0$) suggests eigenvalues are positive (negative) and we have a minima (maxima). $\Delta < 0$ indicates a saddle point stationary point.

Example 10.4 (Sketch the following function of the two variables)**.**

$$u(x, y) = (x - y)(x^2 + y^2 - 1)$$

We note the function is continuous and there are no singularities. The asymptotic behavior is that $u(x, y) \to \pm\infty$ as $x, y \to \pm\infty$.

Next, we find the level curves at zero

$$u(x, y) = 0 \quad \Rightarrow \quad x = y \quad \text{and} \quad x^2 + y^2 - 1 = 0$$

In the next step, we obtain the stationary points.

$$\frac{\partial u}{\partial x}(\vec{x}^*) = \frac{\partial u}{\partial y}(\vec{x}^*) = 0$$

$$\frac{\partial u}{\partial x} = (x^2 + y^2 - 1) + 2x(x - y) = 0,$$

$$\frac{\partial u}{\partial y} = -(x^2 + y^2 - 1) + 2y(x - y) = 0.$$

Adding these two equations we obtain $x^* = y^*$ or $x^* = -y^*$.

1. $x^* = y^* \quad \Rightarrow \quad 2{x^*}^2 - 1 = 0 \quad \Rightarrow \quad P_1 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), \quad P_2 = (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$.
2. $x^* = -y^* \quad \Rightarrow \quad 6{x^*}^2 - 1 = 0 \quad \Rightarrow \quad P_3 = (\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}), \quad P_4 = (-\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}})$.

We classify the stationary points using the Hessain:

$$H(\vec{x}) = \begin{bmatrix} 6x - 2y & 2y - 2x \\ 2y - 2x & 2x - 6y \end{bmatrix}$$

Now we use the determinant (Δ) and trace (τ) of matrix H at each stationary point to classify each stationary point:

$$P_1 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) \Rightarrow H(P_1) = \begin{bmatrix} 4\frac{1}{\sqrt{2}} & 0 \\ 0 & -4\frac{1}{\sqrt{2}} \end{bmatrix} \Rightarrow \Delta < 0 \Rightarrow P_1 \text{ is a saddle point.}$$

$$P_2 = (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}) \Rightarrow H(P_2) = \begin{bmatrix} -4\frac{1}{\sqrt{2}} & 0 \\ 0 & 4\frac{1}{\sqrt{2}} \end{bmatrix} \Rightarrow \Delta < 0 \Rightarrow P_2 \text{ is a saddle point.}$$

$$P_3 = (\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}) \Rightarrow H(P_3) = \begin{bmatrix} \frac{8}{\sqrt{6}} & -\frac{4}{\sqrt{6}} \\ -\frac{4}{\sqrt{6}} & \frac{8}{\sqrt{6}} \end{bmatrix} \Rightarrow \Delta > 0, \tau > 0 \Rightarrow P_3 \text{ is a minimum.}$$

$$P_4 = (-\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}) \Rightarrow H(P_4) = \begin{bmatrix} -\frac{8}{\sqrt{6}} & \frac{4}{\sqrt{6}} \\ \frac{4}{\sqrt{6}} & -\frac{8}{\sqrt{6}} \end{bmatrix} \Rightarrow \Delta > 0, \tau < 0 \Rightarrow P_4 \text{ is a maximum.}$$

Given the location and stability of the stationary points, we complete our sketch of

$$u(x, y) = (x - y)(x^2 + y^2 - 1),$$

as seen in Figure 10.1, by sketching some level curves, specifying the sign of the funcion and the location of the stationary points.

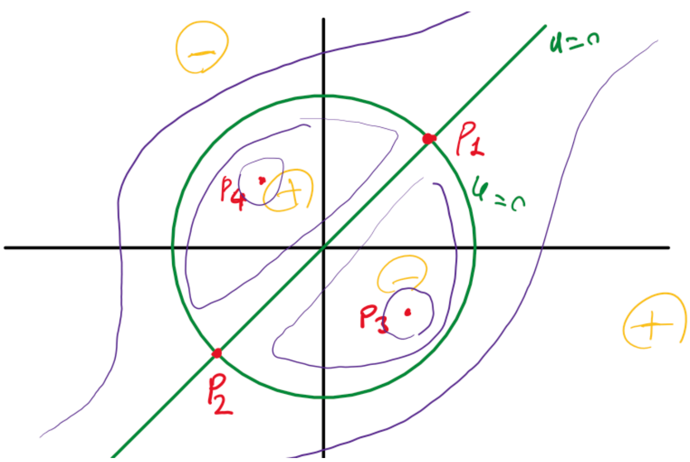

Figure 10.1: The contour plot and sketch of function $u(x, y) = (x-y)(x^2+y^2-1)$