# Course Notes for <br> MS4025: Applied Analysis <br> CHAPTER 1: Integral equations 

Dr. Sarah Mitchell

October 31, 2008

## Contents

Introductory Information ..... 1
0 About the Course ..... 2
I Integral Equations ..... 5
1 Introduction ..... 6
2 Volterra Equations ..... 12
2.1 Solution by differentiation ..... 17
2.2 Solution by Laplace Transform ..... 20
MS4025 Applied Analysis ..... 0-2
2.3 Solution by Iteration ..... 24
2.4 Exercises ..... 31
3 Fredholm Integral Equations ..... 32
3.1 Separable Kernels ..... 32
3.1.1 The Fredholm Alternative ..... 42
3.2 Symmetric Kernels ..... 54
3.2.1 Inhomogeneous Fredholm Integral Equations of the First Kind ..... 60
3.2.2 Inhomogeneous Fredholm Integral Equations of the Second Kind ..... 65
3.3 Exercises ..... 76
MS4025 Applied Analysis ..... 0-3
II Green's Functions ..... 78
4 Introduction ..... 79
4.1 Examples of Sturm-Liouville problems ..... 80
4.2 Discussion of definition of Sturm-Liouville problem ..... 82
4.3 Definition of Green's Function ..... 85
4.4 Physical Interpretation of Green's Functions ..... 100
4.5 Inhomogeneous Boundary Conditions ..... 112
4.6 L Non-invertible ..... 117
4.7 Green's functions via Eigenfunctions ..... 127
4.8 Exercises ..... 132
5 Distributions ..... 134
5.1 Definitions ..... 141
5.2 Formal Results ..... 146
5.3 Distributional Solutions to ODE's ..... 149
5.4 Partial Differential Equations ..... 163
5.4.1 Elliptic Problems ..... 163
5.5 Transforms of distributions ..... 178
5.6 Diffusion problems - problems involving time ..... 182
5.7 Exercises ..... 188
III Complex Analysis ..... 191
6 Calculus in $\mathbb{C}$ ..... 192
6.1 Brief Summary of Terms and Ideas ..... 193
6.2 Limits and Derivatives ..... 200
MS4025 Applied Analysis ..... 0-5
6.3 Cauchy-Riemann Equations ..... 202
7 Complex Integration ..... 213
7.1 Line integrals in $\mathbb{C}$ ..... 214
7.2 Integration Methods ..... 221
7.3 Cauchy's Theorem ..... 228
7.4 Consequences of Cauchy's Theorem ..... 239
7.5 Exercises ..... 253
8 The Method of Residues ..... 258
8.1 Residues ..... 263
8.2 The Residue Theorem ..... 271
8.3 Evaluating Real Integrals ..... 274
8.4 Inverting the Laplace Transform ..... 284
MS4025Applied Analysis0-6
8.4.1 Deriving the Inverse Laplace transform ..... 285
8.4.2 Examples of Inverse Laplace transform ..... 293
8.5 Exercises ..... 304
9 Conformal Mappings and their Applications ..... 308
9.1 The Conformal Property ..... 309
9.2 The Bilinear Transformation ..... 321
9.3 Constructing the "Right" Bilinear Transformation ..... 330
9.4 Conformal Mappings and Boundary Value Problems ..... 342

## Part

## Introductory Information

## 0 About the Course

- The Course is split into three Chapters (or Parts): Integral Equations, Green's Functions \& Complex Analysis.
- Lectures will be held on
- Mondays at 16:00 in CG058
- Tuesdays at 15:00 in KBG14.
- Tutorials will be held on Thursdays at 10:00 in KBG15.
- Notes available at http://www.staff.ul.ie/mitchells/MS4025.html These will be separated into individual Chapters \& made available during the course.
- The main reference text for the course is "Applied Mathematics" by J.David Logan (available in the Library at Shelfmark: 510/LOG) - especially for Chapter II and some of Chapter I.
- "Advanced Engineering Mathematics" by Kreyszig (available in the Library at Shelfmark: 510.2462/KRE) covers most of Chapter III.
- There are Exercises given during and at the end of each Chapter - you will be asked to attempt one or more before the next tutorial.
- There are also statements made in the notes that you are asked to check.
- There will be an end of semester examination for $100 \%$ of the marks for the course.


## Part I

Integral Equations

## 1 Introduction

An integral equation is an equation where the unknown function $u(x)$ occurs under an integral sign. We will restrict ourselves to two types of linear integral equation.

- Fredholm equation

$$
\begin{equation*}
\int_{a}^{b} k(x, y) u(y) d y+\alpha(x) u(x)=f(x), a \leq x \leq b \tag{1.1}
\end{equation*}
$$

- Volterra equation

$$
\begin{equation*}
\int_{a}^{x} k(x, y) u(y) d y+\alpha(x) u(x)=f(x), a \leq x \leq b \tag{1.2}
\end{equation*}
$$

Here $k(x, y)$ is the kernel - assumed continuous on the square $a \leq x, y \leq b \quad(a, b$ finite $)$.

Note the apparently minor difference between (1.1) and (1.2) the solution methods are very different; also (as we will see) (1.1) is closely related to boundary value problems while (1.2) is closely related to initial value problems.

- A solution is a function $\mathfrak{u}(x)$ that satisfies the equation.
- If $f \equiv 0$, the equation is homogeneous, otherwise it is inhomogeneous.
- If $\alpha(x) \equiv 0$ the equation is "of the first kind" otherwise "of the second kind".
- For an equation of the second kind, check that provided $\alpha$ is everywhere non-zero, we can eliminate it from the equation.
- If $k(x, y)=k(y, x)$ we say that the kernel is symmetric integral equations with symmetric kernels have nice properties that make their solution easier.

Just as we have operator notation for ode's (e.g. $\mathrm{Lu} \equiv-\left(\mathrm{p}(\mathrm{x}) \mathfrak{u}^{\prime}(\mathrm{x})\right)^{\prime}+\mathrm{q}(\mathrm{x}) \mathbf{u}(\mathrm{x})-\mathrm{a}$ Sturm Liouville operator), we can use operator notation for integral equations:

$$
\begin{gathered}
(\mathrm{Ku})(\mathrm{x})=\int_{\mathrm{a}}^{\mathrm{b}} k(x, y) \mathfrak{u}(\mathrm{y}) \mathrm{dy} \quad \text { a new function of } \mathrm{x} \\
\mathrm{~K}: \mathbf{u} \rightarrow \mathrm{Ku} \quad \text { a new function. }
\end{gathered}
$$

So our Fredholm integral equation of the second kind may be written

$$
\mathrm{Ku}+\alpha u=\mathrm{f}
$$

This general operator form for an integral equation of the second kind is often re-written as

$$
\begin{equation*}
u=f+\lambda K u \tag{1.3}
\end{equation*}
$$

setting $\alpha \equiv 1$ and introducing $\lambda$ — which as we will see is related to the eigenvalues of the operator $K$. We will refer to this as the standard form for an integral equation of the second kind.

We can consider eigenvalue problems for integral equations (just as we can for o.d.e.'s ) $\mathbf{K u}=\lambda u$ or in standard form (replacing $\lambda$ with $1 / \lambda)$ :

$$
\begin{equation*}
u=\lambda K u \tag{1.4}
\end{equation*}
$$

An eigenvalue is of course just a value of $\lambda$ that satisfies (1.4) for some function $u$ - called an eigenfunction.

Note that the standard form (1.4) for a homogenous integral equation of the second kind is the "opposite" of the analogous eigenvalue problem for the matrix operator $A-A x=\lambda x$. This choice (rather than $\mathrm{Ku}=\lambda \boldsymbol{u}$ ) is convenient when converting an integral equation into a differential equation.

The set of eigenvalues is called the spectrum of $K$ - the multiplicity is just the dimension of the function space spanned by its corresponding eigenfunctions. We will find it useful to study the spectrum of (1.4) when trying to solve (1.1) and (1.2).

Definition 1.1 (Inner Products) Some definitions that we will need:

- If $\mathbf{u}, \boldsymbol{v}$ are functions on $[\mathrm{a}, \mathrm{b}]$ then $(\mathbf{u}, \boldsymbol{v})=\int_{\mathrm{a}}^{\mathrm{b}} u(\mathrm{x}) \overline{\boldsymbol{v}}(\mathrm{x}) \mathrm{d} x$ where $\bar{v}$ is the complex conjugate of $v(x)$.
- The norm of $\mathbf{u},\|\mathbf{u}\|$ is defined by
$\|u\|^{2}=\int_{a}^{b} u \bar{u} d x=\int_{a}^{b}|u(x)|^{2} d x$.
- The norm of $\mathfrak{u}$ is zero iff $\mathfrak{u} \equiv 0$.
- Define the set of square integrable functions $\mathrm{L}^{2}(\mathrm{a}, \mathrm{b})$ to be the functions f such that $\|\mathrm{f}\|^{2}=\int_{\mathrm{a}}^{\mathrm{b}}|\mathrm{f}(\mathrm{x})|^{2} \mathrm{dx}$ is defined. It can be shown that $\mathrm{C}^{2}(\mathrm{a}, \mathrm{b}) \subset \mathrm{L}^{2}(\mathrm{a}, \mathrm{b})$.


## 2 Volterra Equations

We need to develop solution methods for Volterra integral equations - we begin with an example.

Example 2.1 (Inventory control) Suppose a shopkeeper knows that if goods are purchased at any given time then a fraction $\mathrm{k}(\mathrm{t})$ of the goods will remain t days later. At what rate should goods be purchased to keep stock constant?
Solution: Let $\mathfrak{u}(\mathrm{t})$ be the rate (goods per unit time) at which goods are to be bought. Let $A$ be the initial stock level. In the time interval $[\tau, \tau+\Delta \mathrm{t}]$ the shop will buy $\mathfrak{u}(\tau) \Delta \mathrm{t}$ quantity of goods. At the later time $\mathrm{t}(\mathrm{t}-\tau$ days later $), \mathrm{k}(\mathrm{t}-\tau) \mathfrak{u}(\tau) \Delta \tau$ of that purchase will be left.

So the amount of goods left in the shop at time t is is just the sum of these"infinitesimal" contributions plus what remains of the opening balance so:

$$
\text { Stock at time } \mathrm{t}=A k(\mathrm{t})+\int_{0}^{\mathrm{t}} \mathrm{k}(\mathrm{t}-\tau) \mathrm{u}(\tau) \mathrm{d} \tau
$$

The problem to be solved is to find a function $\mathbf{u}(\mathrm{t})$ (given $\mathrm{k}(\mathrm{t})$ ) such that:

$$
\begin{equation*}
A=A k(t)+\int_{0}^{t} k(t-\tau) u(\tau) d \tau \tag{2.1}
\end{equation*}
$$

a Volterra integral equation of the first kind. Check that this is a Volterra integral equation of the first kind. We will see later how to solve this problem.

The next Example illustrates the fact that Volterra integral equations are closely related to initial value problems.

Example 2.2 Given the o.d.e.

$$
\begin{equation*}
u^{\prime \prime}=\lambda u+g(x) ; \quad u(0)=1, u^{\prime}(0)=0 \tag{2.2}
\end{equation*}
$$

we can integrate w.r.t. x. For convenience, we write the RHS in 2.2 as $\mathrm{F}(\mathrm{x})$ :

$$
\begin{gathered}
u^{\prime}=\int_{0}^{x} F(y) d y+C_{1} \\
u=\int_{0}^{x} \int_{0}^{y} F(z) d z d y+C_{1} x+C_{2} .
\end{gathered}
$$

The following neat identity (check it) is just what we need:

$$
\begin{equation*}
\int_{0}^{x} \int_{0}^{y} F(z) d z d y=\int_{0}^{x}(x-y) F(y) d y \tag{2.3}
\end{equation*}
$$

and so

$$
u(x)=\int_{0}^{x}(x-y)(\lambda u(y)+g(y)) d y+C_{1} x+C_{2}
$$

Use the initial conditions (you will need the Leibnitz formula (2.5) below):

$$
\begin{gathered}
u(0)=1 \quad \therefore C_{2}=1 \\
u^{\prime}(0)=C_{1} \quad \therefore C_{1}=0 \\
\therefore u(x)=\lambda \int_{0}^{x}(x-y) u(y) d y \\
\quad+\int_{0}^{x}(x-y) g(y) d y \\
\quad+1
\end{gathered}
$$

This is a Volterra integral equation of the 2nd kind. (Check.)

Exercise 2.1 Consider the o.d.e.
$u^{\prime \prime}+p(x) u^{\prime}+q(x) u=f(x), x>a ; u(a)=u_{0}$ and $u^{\prime}(a)=u_{0}^{\prime}$. Use a procedure similar to the above example to transform the o.d.e. into a Volterra integral equation of the 2nd kind.

### 2.1 Solution by differentiation

We have seen that initial value problems can be reduced to Volterra integral equations. The opposite is also true. We illustrate this by re-visiting Example 2.1.

$$
A=A k(t)+\int_{0}^{t} k(t-\tau) u(\tau) d \tau
$$

Take $k(t)=1-\frac{t}{t_{0}}$, for $t<t_{0}$ and $k(t)=0$ for $t \geq t_{0}$ - i.e. the stock is "run down" at a linear rate of reduction in $t_{0}$ days after which it remains at zero. Then

$$
A=A\left(1-\frac{t}{t_{0}}\right)+\int_{0}^{t}\left(1-\frac{t-\tau}{t_{0}}\right) u(\tau) d \tau
$$

Note that $k(t-\tau)=0$ for $t_{0}<t-\tau \equiv \tau<t-t_{0}$ but $t<t_{0}$ so the lower limit in the integral is unchanged).

## Differentiate w.r.t. t:

$$
\begin{equation*}
0=-\frac{A}{t_{0}}+u(t)+\int_{0}^{t}\left(-\frac{1}{t_{0}}\right) u(\tau) d \tau \tag{2.4}
\end{equation*}
$$

Here we used the Leibniz formula

$$
\begin{equation*}
\frac{d}{d x} \int_{a(x)}^{b(x)} F(x, t) d t=F(x, b(x)) b^{\prime}(x)-F(x, a(x)) a^{\prime}(x)+\int_{a(x)}^{b(x)} F_{x}(x, t) d t \tag{2.5}
\end{equation*}
$$

You can easily check the Leibnitz formula by differentiating from first principles. Differentiating (2.4) again w.r.t. $t$,

$$
\begin{gathered}
0=u^{\prime}(t)-\frac{1}{t_{0}} u(t) \\
\therefore u^{\prime}(t)=\frac{1}{t_{0}}(u(t)) \\
\therefore u=C e^{t / t_{0}}, \quad \text { for } t<t_{0} .
\end{gathered}
$$

If we set $t=0$ in (2.4), we have $u(0)=\frac{A}{t_{0}}$ and so $C=\frac{A}{t_{0}}$ which gives us the final result $u(t)=\frac{A}{t_{0}} e^{t / t_{0}}$.
We can check this answer by substituting in (2.1). The equation to be satisfied for $t<t_{0}$ is:

$$
A=A\left(1-t / t_{0}\right)+\int_{0}^{t} \frac{\left[t_{0}-t+\tau\right]}{t_{0}} \frac{A}{t_{0}} e^{\tau / t_{0}} d \tau
$$

It is easy to check that the RHS reduces to $A$ for all $t<t_{0}$.

### 2.2 Solution by Laplace Transform

An alternative solution method for Example 2.1 is to use the fact (check) that the Laplace Transform of a convolution is just the product of the transforms:

$$
\begin{equation*}
\overline{\int_{0}^{t} a(x) b(t-x) d x}=\bar{a}(s) \bar{b}(s) \tag{2.6}
\end{equation*}
$$

When we take the Laplace Transform of both sides of (2.1) and apply 2.6 we find that $\frac{A}{s}=A \bar{k}+\bar{k} \bar{u}$. We also have $\bar{k}=\frac{1}{s}-\frac{1}{s^{2}} \cdot \frac{1}{t_{0}}$ and so $\bar{u}=\frac{A}{t_{0}\left(s-\frac{1}{t_{0}}\right)}$ which gives $u(t)=\frac{A}{t_{0}} e^{t / t_{0}}$.

Exercise 2.2 What solution do we get if $k(t)=e^{-t / t_{0}}$ for $t \geq 0$ (rapid depletion)?

In general, Volterra equations of convolution type

$$
u(t)=f(t)+\int_{0}^{t} k(t-\tau) u(\tau) d \tau(\text { second kind })
$$

and

$$
0=f(t)+\int_{0}^{t} k(t-\tau) u(\tau) d \tau(\text { first kind })
$$

can be solved most easily using the L.T. method:

- Second kind: $\overline{\mathrm{u}}=\overline{\mathrm{f}}+\overline{\mathrm{k}} \overline{\mathrm{u}} \quad \therefore \overline{\mathrm{u}}=\frac{-\overline{\mathrm{f}}}{\overline{\mathrm{k}}-1} .=\frac{\overline{\mathrm{f}}}{1-\overline{\mathrm{k}}}$
- First kind: $0=\bar{f}+\bar{k} \bar{u} \quad \therefore \bar{u}=-\bar{f} / \bar{k}$

Another example of the L.T. method:

## Example 2.3

$$
\sin x=\lambda \int_{0}^{x} e^{x-t} u(t) d t
$$

Use the L.T. method. We have $\mathcal{L}(\sin )=\frac{1}{s^{2}+1}$,
$\mathcal{L}(k(x))=\mathcal{L}\left(e^{x}\right)=\frac{1}{s-1}$ and so

$$
\begin{gathered}
\frac{1}{s^{2}+1}=\lambda \frac{1}{s-1} \cdot \bar{u} \\
\therefore \bar{u}=\frac{1}{\lambda} \frac{s-1}{s^{2}+1}=\frac{1}{\lambda}\left(\frac{s}{s^{2}+1}-\frac{1}{s^{2}+1}\right)
\end{gathered}
$$

so finally

$$
u(t)=\frac{1}{\lambda}(\cos t-\sin t)
$$

But things can go wrong when this approach is used - Volterra
Integral equations of the first kind do not always have a solution as there may not be any $u(x)$ such that $K u=f$ !

Example 2.4 If we replace $\sin x$ by 1 in the above Example and try to use the L.T. method again

$$
\begin{gathered}
\frac{1}{s}=\lambda \cdot \frac{1}{s-1} \bar{u} \\
\therefore \bar{u}=\frac{1}{\lambda} \frac{s-1}{s}=\left(1-\frac{1}{s}\right) \cdot \frac{1}{\lambda}
\end{gathered}
$$

but there is no function $\mathfrak{u}(\mathrm{t})$ that has $\overline{\mathfrak{u}}$ as its transform - check. This problem has no solution.

### 2.3 Solution by Iteration

Consider the general Volterra integral equation of the second kind - it is convenient to use the "standard form" which is more natural for iterating

$$
u(x)=f(x)+\lambda \int_{a}^{x} k(x, y) u(y) d y
$$

or just $u=f+\lambda K u$ where $(K u)(x)=\int_{a}^{x} K(x, y) u(y) d y$. Choose $u_{0}(x)=f(x)$ as our initial estimate of $\mathfrak{u}(x)$ then $u_{n+1}=f+\lambda K u_{n}$ and so

$$
\begin{gathered}
u_{1}=f+\lambda K f \\
u_{2}=f+\lambda K f+\lambda^{2} K^{2} f
\end{gathered}
$$

and

$$
u_{n}=f+\sum_{1}^{n} \lambda^{i} K^{i} f
$$

Ignoring questions of convergence we write

$$
u(x)=f(x)+\sum_{1}^{\infty} \lambda^{i} K^{i} f
$$

Or more neatly

$$
u(x)=f(x)+\lambda \int_{a}^{x} \Gamma(x, y) f(y) d y
$$

where

$$
\begin{equation*}
\Gamma(x, y)=\sum_{n=0}^{\infty} \lambda^{n} K_{n+1}(x, y) \tag{2.7}
\end{equation*}
$$

and

$$
K_{n+1}(x, y)=\int_{t}^{x} K(x, t) K_{n}(t, y) d t
$$

with

$$
K_{1}(x, y) \equiv K(x, y)
$$

To see where this comes from we derive the result for $u_{2}(x)=\int_{a}^{x} K(x, y) \int_{a}^{y} f(t) K(y, t) d t d y$. Now interchanging the order of integration we have (note the changes in the limits of integration

- this is easiest to see by drawing a sketch)

$$
u_{2}(x)=\int_{a}^{x} f(t)\left[\int_{t}^{x} K(x, y) K(y, t) d y\right] d t
$$

So we can write

$$
u_{2}(x)=\int_{a}^{x} f(t) K_{2}(x, t) d t
$$

where

$$
K_{2}(x, t)=\int_{t}^{x} K(x, y) K(y, t) d y
$$

Similarly (as the above trick with re-ordering the integral still works)

$$
K_{3}(x, t)=\int_{t}^{x} K(x, y) K_{2}(y, t) d y
$$

and in general

$$
K_{n+1}(x, t)=\int_{t}^{x} K(x, y) K_{n}(y, t) d y
$$

and as claimed

$$
u(x)=f(x)+\lambda \int_{a}^{x} \Gamma(x, y) f(y) d y
$$

We assume that the sum (2.7) for $\Gamma$ converges and call $\Gamma(x, y)$ the "resolvent kernel".

If the kernel $K$ is particularly simple we may be able to explicitly calculate the resolvent kernel $\Gamma$.
Example 2.5 Let $u=f+\lambda \int_{0}^{x} e^{x-t} u(t) d t$

$$
\begin{gathered}
K_{1}(x, t)=K(x, t)=e^{x-t} \\
K_{2}(x, t)=\int_{t}^{x} K(x, y) K(y, t) d y=\int_{t}^{x} e^{x-y} e^{y-t} d y \\
\text { so } K_{2}(x, t)=(x-t) e^{x-t}
\end{gathered}
$$

Continuing,

$$
\begin{aligned}
& K_{3}(x, t)=\int_{t}^{x} e^{x-y}(y-t) e^{y-t} d y \\
= & e^{x-t} \int_{t}^{x}(y-t) d y=e^{x-t} \frac{(x-t)^{2}}{2!} .
\end{aligned}
$$

In general, it is easy to see that $\mathrm{K}_{n}(\mathrm{x}, \mathrm{t})=\mathrm{e}^{\mathrm{x}-\mathrm{t}} \frac{(\mathrm{x}-\mathrm{t})^{n}}{\mathrm{n}!}$ and so

$$
\begin{gathered}
\Gamma(x, t)=K_{1}(x, t)+\lambda K_{2}(x, t)+\cdots+\lambda^{n} K_{n+1}(x, t) \\
=e^{x-t}\left[1+\lambda(x-t)+\cdots+\lambda^{n} \frac{(x-t)^{n}}{(n)!}+\right] \\
=e^{x-t} e^{\lambda(x-t)}=e^{(1+\lambda)(x-t)}=\Gamma(x, t ; \lambda)
\end{gathered}
$$

So the solution is

$$
u(x)=f(x)+\lambda \int_{0}^{x} e^{(1+\lambda)(x-t)} f(t) d t
$$

It is not often possible to find a closed form for $\Gamma$ - a finite sum yields a numerical approximation.

A simpler way to "approximately solve" the problem is to simply iterate

$$
u_{n+1}=f+\lambda \int_{a}^{x} k(x, y) u_{n}(x) d x
$$

where $u_{0}$ is chosen appropriately $u_{0}=0 \equiv u_{1}=f$.
Example 2.6 Find the first two terms in the series solution to:

$$
u=x-\int_{0}^{x}(x-t) u(t) d t
$$

### 2.4 Exercises

1. Check the Leibnitz Formula (2.5)
2. Solve the integral equation $u(x)=x+\lambda \int_{0}^{x}(x-y) u(y) d y$.
3. Rewrite the initial value problem $u^{\prime \prime}-\lambda u=f(x), \quad x>0$; $u(0)=1, u^{\prime}(0)=0$ as a Volterra integral equation.
4. Solve the integral equation $\int_{0}^{x} y u(y) d y-\lambda u(x)=f(x)$, $0 \leq x \leq 1$ using any method you wish. (Assume that $\lambda \neq 0$.)
5. Find the first three terms in the series solution to:

$$
u=1+\int_{0}^{x}(x+y) u(y) d y
$$

## 3 Fredholm Integral Equations

In this Chapter we consider Fredholm Integral equations of the second kind. Using operator notation:

$$
\begin{equation*}
u=f+\lambda K u \tag{3.1}
\end{equation*}
$$

We first consider an important special case.

### 3.1 Separable Kernels

Separable (sometimes called degenerate) kernels $k(x, y)$ can be written as a finite sum of terms, each of which is the product of a function of $x$ times a function of $y$ :

$$
k(x, y)=\sum_{i} X_{i}(x) Y_{i}(y)
$$

Then the general Fredholm Integral equation of the second kind (3.1) takes the special form:

$$
\begin{equation*}
\mathfrak{u}(x)=f(x)+\lambda \sum X_{j}(x) \int_{a}^{b} Y_{j}(y) \mathfrak{u}(y) d y \tag{3.2}
\end{equation*}
$$

or just

$$
\begin{equation*}
u(x)=f(x)+\lambda \sum_{j} u_{j} x_{j}(x) \tag{3.3}
\end{equation*}
$$

where the numbers $u_{i} \equiv \int_{a}^{b} Y_{\mathfrak{j}}(y) u(y) d y$ are to be determined once we have calculated all the $\mathrm{U}_{\mathrm{j}}$ we can write down the solution from (3.3).

The solution $\mathfrak{u}(x)$ is just the inhomogeneous term $f(x)$ plus a finite linear combination of the $X_{j}(x)$ so separable Fredholm integral equations are (as the word degenerate suggests) a very special case.

Now multiply (3.3) on the left by $Y_{i}(x)$ and integrate from $y=a$ to $y=b$ (change the dummy variable to $y$ for convenience). Then

$$
u_{i}=F_{i}+\lambda \sum_{j} u_{j} \int_{a}^{b} X_{j}(y) Y_{i}(y) d y \quad \text { where } F_{i}=\left(f, Y_{i}\right)
$$

Then we have (using a vector notation) $U=F+\lambda A U$ with $A_{i j}=\int X_{j}(y) Y_{i}(y) d y=\left(X_{j}, Y_{i}\right)$. We can now write

$$
\begin{equation*}
(\mathrm{I}-\lambda A) \mathrm{U}=\mathrm{F} \tag{3.4}
\end{equation*}
$$

and so $U=(I-\lambda A)^{-1} F$ if the matrix $I-\lambda A$ is invertible or equivalently provided $\operatorname{det}(I-\lambda A) \neq 0$. Once we solve (3.4) for $\mathbb{U}$, we can substitute into (3.3) and find the solution $u(x)$.

If $\operatorname{det}(I-\lambda A)=0$ then $\frac{1}{\lambda}$ must be an eigenvalue of $A$ (why?).
Then there is either no solution or infinitely many - depending on whether $F$ is not/is in the column space of $I-\lambda A$.

If $f(x) \equiv 0$ (homogeneous case) the problem reduces to $u=\lambda K u$, so the eigenvalues of $A$ are the reciprocals of the eigenvalues of $K$. The corresponding eigenvectors/eigenfunctions are found by following the above procedure with $f \equiv 0$, namely $(I-\lambda A) U=0$.

So the vector $U$ is just the eigenvector of $A$ corresponding to the eigenvalue $\frac{1}{\lambda}$.

This analysis can be summarised as what is sometimes called a "Theorem of the alternative": Given a homogeneous Fredholm Integral Equation of the second kind with separable kernel; then defining $A$ as above:

- if $\frac{1}{\lambda}$ is not an eigenvalue of $A$ then there is a unique solution
- if $\frac{1}{\lambda}$ is an eigenvalue of $\mathcal{A}$ then there is either no solution or there are infinitely many.

Example 3.1 Solve the homogeneous Fredholm Integral Equation of the second kind with separable kernel;

$$
u(x)=x+\lambda \int_{0}^{1}\left(x y^{2}+x^{2} y\right) u(y) d y
$$

We have

$$
k(x, y)=x y^{2}+x^{2} y=X_{1} Y_{1}+X_{2} Y_{2}
$$

and $\mathrm{f}(\mathrm{x}) \equiv \mathrm{x}$. We will need the table

$$
\begin{array}{ccc}
i & X_{i} & Y_{i} \\
1 & x & y^{2} \\
2 & x^{2} & y
\end{array}
$$

when we calculate the various coefficients $\mathcal{A}_{i j}$ and $\mathrm{F}_{\mathfrak{i}}$.

By definition,

$$
F=\binom{F_{1}}{F_{2}}=\binom{\left(f, Y_{1}\right)}{\left(f, Y_{2}\right)} .
$$

So

$$
F_{1}=\left(f, Y_{1}\right)=\int_{0}^{1} x \cdot x^{2} d x=\frac{1}{4}
$$

and

$$
F_{2}=\left(f, Y_{2}\right)=\int_{0}^{1} x \cdot x d x=\frac{1}{3}
$$

so

$$
F=\left[\begin{array}{c}
\frac{1}{4} \\
\frac{1}{3}
\end{array}\right]
$$

We need to form the matrix $A$, defined by $A_{i j}=\left(X_{\mathfrak{j}}, Y_{i}\right)$,
$A_{11}=\frac{1}{4} ; \quad A_{12}=\frac{1}{5} ; \quad A_{21}=\frac{1}{3} ; A_{22}=\frac{1}{4}$. So $A=\left[\begin{array}{cc}\frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4}\end{array}\right]$.

We need to solve

$$
\mathrm{U}=\left[\begin{array}{c}
\frac{1}{4} \\
\frac{1}{3}
\end{array}\right]+\lambda\left[\begin{array}{cc}
\frac{1}{4} & \frac{1}{5} \\
\frac{1}{3} & \frac{1}{4}
\end{array}\right] \mathrm{U}
$$

Rewrite as

$$
\left[\begin{array}{cc}
1-\frac{\lambda}{4} & -\frac{\lambda}{5} \\
-\frac{\lambda}{3} & 1-\frac{\lambda}{4}
\end{array}\right]\left[\begin{array}{l}
U_{1} \\
U_{2}
\end{array}\right]=\left[\begin{array}{c}
\frac{1}{4} \\
\frac{1}{3}
\end{array}\right]
$$

The determinant $|\mathrm{I}-\lambda A|$ evaluates to $\frac{1}{240}\left(240-120 \lambda-\lambda^{2}\right)$.
Provided $\operatorname{det}(\mathrm{I}-\lambda A) \neq 0(\lambda \neq-60 \pm 16 \sqrt{15})$ the problem has a unique solution

$$
U=\frac{1}{240-120 \lambda-\lambda^{2}}\binom{60+\lambda}{80}
$$

Finally,

$$
u(x)=x+\lambda\left(u_{1} x+u_{2} x^{2}\right)
$$

Now suppose that $\lambda$ is equal to (say) $-60+16 \sqrt{15} \equiv \frac{1}{1 / 4-1 / \sqrt{15}}$,
(i.e. the reciprocal of one of the two eigenvalues of $\mathcal{A}$ ). Then

$$
(I-\lambda A)=\left[\begin{array}{cc}
16-4 \sqrt{15} & 12-\frac{16}{5} \sqrt{15} \\
20-16 / 3 \sqrt{15} & 16-4 \sqrt{15}
\end{array}\right]
$$

We know that the two columns are parallel, (check why?) so for F to be in the column space of $(\mathrm{I}-\lambda \mathrm{A})$ it must be the case that F is a multiple of (say) the first column of $(\mathrm{I}-\lambda \mathcal{A})$. For this to be true we must have the ratio $\frac{F_{1}}{F_{2}}=\frac{16-4 \sqrt{15}}{20-16 / 3 \sqrt{15}} \equiv-\sqrt{3 / 5}$. But $F_{1} / F_{2}=3 / 4$. So for this value of $\lambda$ there is no solution. (You should check that this is also true when $\lambda=-60-16 \sqrt{15}$.)

If by good fortune F is a linear combination of the columns of ( $\mathrm{I}-\lambda A$ ) then we will have infinitely many solutions.

For example if $\mathrm{C}_{1}=\mathrm{c}_{1} \mathrm{~F}$ and $\mathrm{C}_{2}=\mathrm{c}_{2} \mathrm{~F}$ where $\mathrm{C}_{1}$ and $\mathrm{C}_{2}$ are the first and second columns of $(\mathrm{I}-\lambda \mathrm{A})$ respectively then we must have

$$
\begin{equation*}
\mathrm{c}_{1} \mathrm{u}_{1}+\mathrm{c}_{2} \mathrm{u}_{2}=1 \tag{3.5}
\end{equation*}
$$

and the solution is just

$$
u(x)=x+\left(u_{1} x+u_{2} x^{2}\right)
$$

where $\mathrm{U}_{1}$ and $\mathrm{U}_{2}$ are any of the infinitely many solutions to (3.5).
We will see shortly a more systematic (and simpler) way of doing this analysis.

### 3.1.1 The Fredholm Alternative

Consider the (separable) homgeneous Fredholm Integral Equation of the second kind.

$$
u=\lambda K u
$$

For degenerate/separable problems

$$
u=\lambda \sum X_{k}(x) u_{k}
$$

where

$$
\mathrm{u}_{\mathrm{k}}=\int_{\mathrm{a}}^{\mathrm{b}} \mathrm{Y}_{\mathrm{k}}(\mathrm{y}) \mathrm{u}(\mathrm{y}) \mathrm{dy}
$$

As before this reduces to $(I-\lambda A) U=0$. So if $|I-\lambda A| \neq 0$, the only solution is $u(x)=0$. Otherwise (zero determinant) there are infinitely many solutions (a linear system has 0,1 or $\infty$ solutions).

For this problem $u=\lambda K u$, we refer to $\lambda$ as the eigenvalue of $K$ (it would be more natural to call $\frac{1}{\lambda}$ the eigenvalue but it is useful to stick with $u=\lambda K u)$. The non-trivial solutions $u_{j}(x)$ corresponding to each $\lambda_{j}$ are the eigenfunctions of $K$.

Example 3.2 Given the (separable) homogeneous integral equation $u=\lambda \int_{0}^{\pi}\left(\cos ^{2} x \cos 2 y+\cos 3 x \cos ^{3} y\right) u(y) d y$, we have:
$\begin{array}{lll}i & X_{i} & Y_{i}\end{array}$
$1 \cos ^{2} x \quad \cos 2 y$
$2 \cos 3 x \cos ^{3} y$

Using the definition $\mathrm{A}_{\mathrm{ij}} \equiv\left(\mathrm{X}_{\mathrm{j}}, \mathrm{Y}_{\mathrm{i}}\right)$ or $\quad\left(\mathrm{Y}_{\mathrm{i}}, \mathrm{X}_{\mathrm{j}}\right)$;
$A_{11}=\int_{0}^{\pi} \cos ^{2} x \cos 2 x=\frac{\pi}{4}, A_{12}=\int_{0}^{\pi} \cos 3 x \cos 2 x=0$,
$A_{21}=\int_{0}^{\pi} \cos ^{3} x \cos ^{2} x=0$ and $A_{22}=\int_{0}^{\pi} \cos 3 x \cos ^{3} x=\frac{\pi}{8}$ and therefore

$$
A=\left[\begin{array}{cc}
\pi / 4 & 0 \\
0 & \pi / 8
\end{array}\right]
$$

So the matrix equation for U is $(\mathrm{I}-\lambda \mathrm{A}) \mathrm{U}=0$ :

$$
\left[\begin{array}{ccc}
1-\frac{\pi}{4} \lambda & 0 \\
0 & 1-\frac{\pi}{8} \lambda
\end{array}\right]\left[\begin{array}{l}
\mathrm{U}_{1} \\
\mathrm{U}_{2}
\end{array}\right]=0 .
$$

This has a non-trivial solution only if

$$
\operatorname{det}(I-\lambda A)=\left(1-\lambda \frac{\pi}{4}\right)\left(1-\lambda \frac{\pi}{8}\right) \text { is zero. }
$$

So if $\lambda$ takes a value other than $\frac{4}{\pi}$ or $\frac{8}{\pi}$ then the only solution is $\mathrm{U}=0$ as $(\mathrm{I}-\lambda A)$ is invertible. So $\mathrm{u}(\mathrm{x}) \equiv 0$ for all $\mathrm{x} \in[0, \pi]$.
Now consider the two special cases, $\lambda=\frac{4}{\pi}$ and $\lambda=\frac{8}{\pi}$.
(a) $\lambda=\frac{4}{\pi}$ so $0 \times \mathrm{U}_{1}=0\left(\mathrm{U}_{1}\right.$ arbitrary $)$ and $\left(1-\frac{4}{\pi} \frac{\pi}{8}\right) \mathrm{U}_{2}=0$ so
$\mathrm{U}_{2}=0$. We have

$$
\mathrm{u}^{(1)}=\binom{1}{0} \text {, say, and } \quad \mathrm{u}^{(1)}(\mathrm{x})=\frac{4}{\pi} \cos ^{2} x
$$

(b) $\lambda=\frac{8}{\pi}$. In this case $\mathrm{U}_{1}=0$ and $\mathrm{U}_{2}$ is arbitrary.

$$
\therefore \mathrm{U}^{(2)}=\binom{0}{1}, \quad \text { say, and } \quad \mathrm{u}^{(2)}(\mathrm{x})=\frac{8}{\pi} \cos 3 \mathrm{x}
$$

The insights gained from studying Fredholm Integral equations of the second kind (both inhomogeneous and homogeneous) are still valid for the much more general case of a symmetric (non-separable) kernel.

We state without proof two Theorems.
Theorem 3.1 ( First Fredholm Alternative) If the
homogeneous Fredholm integral equation of the second kind $\mathfrak{u}=\lambda \mathrm{Ku}$ has only the trivial solution $\mathbf{u}(\mathrm{x}) \equiv 0$, then the corresponding inhomogeneous equation $\mathbf{u}=\mathrm{f}+\lambda \mathrm{Ku}$ has exactly one solution for any given f
and
If the homogeneous integral equation has non-trivial solutions, then the inhomogeneous integral equation has either no solution or infinitely many.

This Theorem is often stated as:
Theorem 3.2 ( First Fredholm Alternative-rewritten)
Either the inhomogeneous Fredholm integral equation $u=f+\lambda K u$ has exactly one solution for any given f or the homogeneous integral equation $\mathfrak{u}=\lambda \mathrm{Ku}$ has non-trivial solutions (but not both).

See Exercise 1 to show that Theorems 3.1 and 3.2 are equivalent.
If the kernel is symmetric then we can say more.
Theorem 3.3 ( Second Fredholm Alternative) When the homogeneous Fredholm integral equation of the second kind with symmetric kernel $u=\lambda \mathrm{Ku}$ has a non-trivial solution (or solutions) $\mathfrak{u}_{\mathfrak{j}}(\mathrm{x})$ corresponding to $\lambda=\lambda_{\mathrm{j}}$ then the associated inhomogeneous equation (with the same value for the parameter $\lambda$ ), namely $u=\mathrm{f}+\lambda \mathrm{Ku}$, will have a solution if and only if $\left(\mathrm{f}, \mathrm{u}_{\mathfrak{j}}\right)=0$ for every eigenfunction $\mathfrak{u}_{\mathfrak{j}}(\mathrm{x})$ (corresponding to $\lambda=\lambda_{\mathfrak{j}}$ ) of the homogeneous integral equation (eigenvalue problem) $u=\lambda К u$.

Consider again Example 3.2 (non-symmetric kernel) above.
Theorem 3.2 applies. It tells us that the associated inhomogeneous Fredholm integral equation of the second kind $u=f+\lambda K u$ will have exactly one solution if $\lambda \neq \frac{4}{\pi}$ or $\frac{8}{\pi}$. For $\lambda=\frac{4}{\pi}$ or $\frac{8}{\pi}$ the associated inhomogeneous problem will have either no solution or infinitely many - we cannot say which as Theorem 3.3 only applies to symmetric kernels. Of course we could simply try to construct a solution by solving $(\mathrm{I}-\lambda A) \mathrm{U}=\mathrm{F}$ for $\lambda=\lambda_{1} \equiv \frac{4}{\pi}$ and $\lambda=\lambda_{2} \equiv \frac{8}{\pi}$ - with no way of knowing in advance whether solutions existed.

Exercise 3.1 Check whether the Fredholm integral equation $u=\sin (x)+\lambda \int_{0}^{\pi}\left(\cos ^{2} x \cos 2 y+\cos 3 x \cos ^{3} y\right) u(y) d y$ has zero, one or infinitely many solutions for $\lambda=\frac{4}{\pi}$ ?

Example 3.3 Consider the symmetric problem

$$
u=f+\lambda \int_{0}^{2 \pi} \sin (x+y) u(y) d y
$$

As the kernel is symmetric, Theorem 3.3 applies.

- Theorem 3.1 tells us that for $\lambda$ not an eigenvalue of the kernel K we will have a unique solution.
- Theorem 3.3 tells us that if $\lambda$ is an eigenvalue of $\mathrm{K}\left(\lambda_{j}\right.$, say) then the existence of (infinitely many) solutions requires that $\left(\mathrm{f}, \phi_{\mathfrak{j}}\right)=0$ for each $\phi_{\mathfrak{j}}(\mathrm{x})$ corresponding to the eigenvalue $\lambda_{j}$. To see how this works, check that $A=\left[\begin{array}{cc}0 & \pi \\ \pi & 0\end{array}\right]$ so $|I-\lambda A|=0$ when $\lambda= \pm \frac{1}{\pi}$.
So for $\lambda \neq \pm \frac{1}{\pi}$ we find the unique solution for $u(x)$ by solving $(\mathrm{I}-\lambda A) \mathrm{U}=\mathrm{F}$ for U and substituting for U in $u(x)=f(x)+\lambda \sum u_{i} X_{i}(x)$.

On the other hand when $\lambda=\lambda_{j}= \pm \frac{1}{\pi}$ we expect to have a solution to the inhomogeneous problem only if $\left(\mathrm{f}, \phi_{\mathfrak{j}}\right)=0$ for each $\phi_{\mathfrak{j}}(\mathrm{x})$ corresponding to $\lambda_{j}$.
We can find the eigenfunctions $\phi_{j}$ corresponding to $\lambda_{1}=\frac{1}{\pi}$ and $\lambda_{2}=-\frac{1}{\pi}$ as before - check

$$
\begin{aligned}
& \phi_{1}=\frac{\mathrm{U}_{1}}{\pi}(\sin x+\cos x) ; \mathrm{U}_{1} \quad \text { arbitrary } \\
& \phi_{2}=-\frac{\mathrm{U}_{2}}{\pi}(\sin x-\cos x) ; \mathrm{U}_{2} \quad \text { arbitrary }
\end{aligned}
$$

We usually normalise the eigenfunctions so that $\left\|\phi_{1}\right\|=\left\|\phi_{2}\right\|=1$ which results in check:

$$
\phi_{1}=\frac{1}{\sqrt{2 \pi}}(\sin +\cos ) ; \phi_{2}=\frac{1}{\sqrt{2 \pi}}(\sin -\cos )
$$

So consider the case $\lambda=\lambda_{1}=\frac{1}{\pi}$. Theorem 3.3 tells us that we have no solution unless $\left(\mathrm{f}, \phi_{1}\right)=0$ in which case we have infinitely many solutions.

Take $\mathrm{f}(\mathrm{x})=\mathrm{x}$ :

$$
\begin{aligned}
& \frac{1}{\sqrt{2 \pi}} \int_{0}^{2 \pi} x(\sin (x)+\cos (x)) d x= \\
& \frac{1}{\sqrt{2 \pi}}\left(\left.[-\cos (x)+\sin (x)] x\right|_{0} ^{2 \pi}-\int_{0}^{2 \pi}[-\cos (x)+\sin (x)] d x\right) \\
& =\frac{1}{\sqrt{2 \pi}} \neq 0
\end{aligned}
$$

so we expect no solution.

Let's see. Try solving $\left(\mathrm{I}-\lambda_{1} \mathrm{~A}\right) \mathrm{U}=\mathrm{F}$ We can see immediately that there is no solution as

$$
\left(\begin{array}{cc}
1 & -1 \\
-1 & 1
\end{array}\right)\binom{u_{1}}{u_{2}}=\binom{\left(f, Y_{1}\right)}{\left(f, Y_{2}\right)}=\binom{(x, \cos (x))}{(x, \sin (x))}
$$

We have $\mathrm{K}(\mathrm{x}, \mathrm{y}) \equiv \sin \mathrm{x} \cos \mathrm{y}+\cos \mathrm{x} \sin \mathrm{y}$ giving the table:

$$
\begin{array}{lll} 
& X & Y \\
1 & S & C \\
2 & C & S
\end{array}
$$

So we have have $\mathrm{U}_{1}-\mathrm{U}_{2}=0$ and $-\mathrm{U}_{1}+\mathrm{U}_{2}=-2 \pi-$ an inconsistent linear system - so no solution.

Suppose $\mathrm{f}=\sin 2 \mathrm{x}$. It is easy to check that $(\sin 2 x, \sin x+\cos x)=0$ so we expect infinitely many solutions

$$
\begin{gathered}
\left(\mathrm{I}-\lambda_{1} A\right) \mathrm{U}=\mathrm{F} \\
\mathrm{U}_{1}-\mathrm{U}_{2}=(\sin 2 x, \cos 2 x)=0 \\
-\mathrm{U}_{1}+\mathrm{U}_{2}=(\sin 2 x, \sin 2 x)=0 \\
\therefore \mathrm{U}_{1}=\mathrm{U}_{2} \quad \text { arbitrary. }
\end{gathered}
$$

Therfore the solution is $u(x)=\mathrm{U}_{1}\left[\phi_{1}(x)+\phi_{2}(x)\right]$, where $\mathrm{U}_{1}$ is arbitary.

### 3.2 Symmetric Kernels

This more general case requires extra techniques - in particular we will need to work systematically to find the eigenvalues and eigenfunctions of the operator K in the "eigenvalue equation"

$$
u=\lambda K u, \quad(K u)(x)=\int_{a}^{b} k(x, t) u(t) d t
$$

where $k(x, y)=k(y, x)$ is symmetric. First we prove two theorems on useful properties of symmetric kernels.

Theorem 3.4 If the kernel $\mathrm{k}(\mathrm{x}, \mathrm{y})$ is symmetric and real then $(K u, v)=(u, K v)$.

Proof: we take the general case where $\left.(u, v)=\int u(x) \bar{v}(x) d x\right)$. Now

$$
\begin{aligned}
(K u, v) & =\int_{a}^{b}\left(\int_{a}^{b} k(x, t) u(t) d t\right) \overline{v(x)} d x \\
& =\int_{a}^{b} \int_{a}^{b} k(x, t) u(t) \overline{v(x)} d t d x \\
& =\int_{a}^{b} \int_{a}^{b} k(t, x) u(x) \overline{v(t)} d x d t \quad \text { re-labelling variables } \\
& =\int_{a}^{b} \int_{a}^{b} k(x, t) u(x) \overline{v(t)} d t d x \quad \text { symmetric kernel } \\
& =\int_{a}^{b} u(x) \overline{\left(\int_{a}^{b} k(x, t) v(t) d t\right)} d x \quad \text { real kernel } \\
& =(u, K v)
\end{aligned}
$$

Theorem 3.5 If a kernel K is real, symmetric and continuous, then its eigenvalues are real and eigenfunctions corresponding to distinct eigenvalues are orthogonal.

Proof: First note that $(\mathbb{K u}, \mathfrak{u})=(\mathfrak{u}, \mathrm{Ku})$ by Thm. 3.4. But $\overline{(\mathbb{K u}, \mathfrak{u})}=\int_{a}^{b}(\overline{\mathrm{Ku}})(x) \mathfrak{u}(x) \mathrm{d} x=\int_{a}^{b} \mathfrak{u}(x)(\overline{\mathrm{Ku}})(x) \mathrm{d} x=(\mathfrak{u}, \mathrm{Ku})$ so $(\overline{K u, u})=(\mathbb{K u}, \mathfrak{u})$ and so $(\mathbb{K u}, \mathfrak{u})$ is real. Let $\mathfrak{u}=\lambda K \mathfrak{u}$ - then $(K \mathfrak{u}, \mathfrak{u})=\frac{1}{\lambda}(\mathfrak{u}, \mathfrak{u})$ and $\lambda$ is therefore real as $(\mathfrak{u}, \mathfrak{u}) \geq 0$.
Now we have $u=\lambda K \mathfrak{u}$; let $v=\mu K v, \lambda \neq \mu$. Then

$$
\begin{aligned}
& (\mathrm{Ku}, v)=\frac{1}{\lambda}(\mathfrak{u}, v) \quad \text { and } \\
& (\mathfrak{u}, \boldsymbol{K} v)=\frac{1}{\mu}(\mathfrak{u}, v) .
\end{aligned}
$$

But the left hand sides are equal by Thm. 3.4 so, as $\lambda \neq \mu$, we must have $(u, v)=0$.

It can be shown that, when a kernel K is real, symmetric (and non-separable), $K$ has infinitely many eigenvalues $\lambda_{1}, \lambda_{2}, \ldots$, each with finite multiplicity (i.e. only a finite number of corresponding eigenfunctions) which we can label

$$
0<\cdots \leq\left|\lambda_{n}\right| \leq \cdots \leq\left|\lambda_{2}\right| \leq\left|\lambda_{1}\right|
$$

with $\lim _{n \rightarrow \infty} \lambda_{n}=0$.
Also any square-integrable function $f \in L^{2}[a, b]$ can be written as

$$
f(x)=\sum_{n=1}^{\infty} f_{n} \phi_{n}(x), \quad f_{n}=\left(\phi_{n}, f\right)
$$

The reason why we focus on eigenfunction expansions is that they give us a general method for solving symmetric problems. Consider the homogeneous Fredholm integral equation of the first kind

$$
f=\lambda K u
$$

The eigenfunctions $\phi_{n}(x)$ satisfy

$$
\phi_{\mathfrak{n}}(x)=\lambda_{\mathfrak{n}} \int k(x, y) \phi(y) d y
$$

We know that they are orthogonal and that the eigenvalues are real. We now state without proof a useful result:

Theorem 3.6 (Hilbert-Schmidt) If $\mathrm{f}=\lambda \mathrm{Ku}$ as above where k is symmetric and both k and u are square integrable, then f can be expanded in a "Fourier Series" - i.e. $f(x)=\sum f_{k} \phi_{k}(x)$ where $f_{k}=\left(f, \phi_{k}\right)$. (The series converges to $f(x)$ in the mean and is absolutely and uniformly convergent.)

This result will allow us to develop a general solution for any (symmetric) inhomogeneous Fredholm integral equation of the second kind. Despite the Theorem not telling us anything about what we are interested in (the solution, $\mathfrak{u}(\mathrm{x})$ ) we will base our solution technique on it.

First, we briefly examine the question of the existence of solutions to an inhomogeneous Fredholm integral equation of the first kind $f=\lambda K u$.

### 3.2.1 Inhomogeneous Fredholm Integral Equations of the First Kind

A solution $u$ will not necessarily exist for all $f$ - more precisely:
Theorem 3.7 For a continuous real, symmetric kernel and continuous $\mathrm{f}(\mathrm{x})$ an inhomogeneous Fredholm integral equation of the first kind $\mathrm{f}=\lambda \mathrm{Ku}$ has a solution only if f can be expressed in a series of the eigenfunctions $\phi_{\mathrm{k}}(\mathrm{x})$ of the kernel.

Now it is easy to see that if $f(x)=\sum f_{k} \phi_{k}(x)$ where $\phi_{k}(x)=\lambda_{k} K \phi_{k}(x)$ then

$$
\begin{aligned}
\sum f_{k} \phi_{k} & =\lambda K \sum u_{k} \phi_{k} \\
& =\lambda \sum u_{k} K \phi_{k} \\
& =\sum u_{k} \frac{\lambda}{\lambda_{k}} \phi_{k}
\end{aligned}
$$

Therefore $u_{k}=\frac{f_{k} \lambda_{k}}{\lambda}$ and so $u(x)=\sum_{k} \frac{f_{k} \lambda_{k}}{\lambda} \phi_{k}$

So we are guaranteed a solution, though not a unique solution as any $\Psi(x)$ that is orthogonal to the kernel can be added to $u(x)$ without changing the output. So to ensure the solution $u(x)$ is unique we must check that there are no functions $\Psi$ such that

$$
\int_{a}^{b} K(x, y) \Psi(y) d y=0 \text { all } x
$$

Before taking an example we recall that the Fourier Sine Series for a function $f \in L^{2}(0, l)$ is as follows:

$$
\begin{aligned}
f(x) & =\sum_{0}^{\infty} f_{n} \sin \frac{n \pi x}{l} \\
f_{n} & =\frac{2}{l} \int_{0}^{l} f(x) \sin \frac{n \pi x}{l} d x
\end{aligned}
$$

Example 3.4 Let $f(x)=\frac{1}{2}\left(x-x^{2}\right), \lambda=1$. We can check that with

$$
k(x, t)= \begin{cases}x(1-t) & 0 \leq x \leq t \\ t(1-x) & t \leq x \leq 1\end{cases}
$$

the equation $\mathrm{f}(\mathrm{x})=\mathrm{Ku}$ is satisfied by $\mathfrak{u}(\mathrm{x})=1,0<\mathrm{x}<1$.

- The Fourier Sine Series for $\mathrm{f}(\mathrm{x})$ is

$$
f(x) \equiv \frac{1}{2}\left(x-x^{2}\right)=\sum_{1}^{\infty} \frac{2 \sqrt{2} \cdot \sqrt{2}}{\pi^{3}(2 k+1)^{3}} \sin (2 k+1) \pi x
$$

- The Fourier Sine Series for $\mathfrak{u}(x) \equiv 1$ is

$$
\sum_{1}^{\infty} \frac{2 \sqrt{2} \cdot \sqrt{2}}{\pi(2 k+1)} \sin (2 k+1) \pi x
$$

The condition $\mathfrak{u}_{\mathrm{k}}=\frac{\mathrm{f}_{\mathrm{k}} \lambda_{\mathrm{k}}}{\lambda}$ (with $\lambda=1$ ) is satisfied as for the above kernel the eigenvalue equation $\mathbf{u}=\lambda \mathrm{ku}$ is satisfied by eigenvalues $\lambda_{k}=\pi^{2}(2 k+1)^{2}$.

Example 3.5 On the other hand the equation $\mathrm{x}=\mathrm{Ku}$ with the same kernel $\mathrm{k}(\mathrm{x}, \mathrm{y})$ has no solution. Check: $\mathrm{f}_{\mathrm{k}}=\frac{2}{\mathrm{k} \pi}(-1)^{\mathrm{k}+1}$ so $u_{k}=\frac{2}{k \pi}(-1)^{k+1} \pi^{2}(2 \mathrm{k}+1)^{2}$. But this is not a convergent series so $u(x)$ is not defined.

### 3.2.2 Inhomogeneous Fredholm Integral Equations of the Second Kind

As in the case of a Volterra integral equation of the second kind we can form a "resolvent kernel": $\Gamma(x, y ; \lambda)$ that effectively is the inverse operator to K . We will express the result as a Theorem.

Theorem 3.8 Given a Fredholm integral equation of the second kind $u=\mathrm{f}+\lambda \mathrm{Ku}$, the solution can be expressed as

$$
\begin{equation*}
u(x)=f(x)+\lambda \int_{a}^{b} f(y) \Gamma(x, y ; \lambda) d y \tag{3.6}
\end{equation*}
$$

where

$$
\begin{equation*}
\Gamma(x, y ; \lambda)=\sum_{1}^{\infty} \frac{\phi_{k}(x) \phi_{k}(y)}{\lambda_{k}-\lambda} \quad \lambda \neq \lambda_{k} \tag{3.7}
\end{equation*}
$$

Alternatively

$$
\begin{equation*}
u(x)=f(x)+\lambda \sum \frac{f_{k} \phi_{k}(x)}{\lambda_{k}-\lambda} \tag{3.8}
\end{equation*}
$$

Proof: We rewrite our integral equation $\mathfrak{u}=\mathrm{f}+\lambda \boldsymbol{K} u$ as $\mathrm{d}(\mathrm{x}) \equiv \boldsymbol{u}-\mathrm{f}=\lambda \mathrm{K} \mathbf{u}$. Now (provided that $\mathbf{u}$ and f are in $L_{2}(\mathrm{a}, \mathrm{b})$ ) the Hilbert-Schmidt Theorem 3.6 tells us that we have $d(x)=\sum_{1}^{\infty} d_{k} \phi_{k}(x)$ where $d_{k}=u_{k}-f_{k}$.
We also have

$$
\begin{aligned}
d_{k} & =\int_{a}^{b} d(x) \phi_{k}(x) d x \\
& =\int_{a}^{b}\left(\int_{a}^{b} \lambda k(x, y) u(y) d y\right) \phi_{k}(x) d x
\end{aligned}
$$

Swapping the order of integration \& using the symmetry of the kernel, we have

$$
\begin{aligned}
d_{k} & =\lambda \int_{a}^{b} u(y)\left(\int_{a}^{b} k(y, x) \phi_{k}(x) d x\right) d y \\
& =\lambda \int_{a}^{b} u(y) \frac{\phi_{k}(y)}{\lambda_{k}} d y \\
& =\frac{\lambda}{\lambda_{k}} u_{k}
\end{aligned}
$$

So using $d_{k}=u_{k}-f_{k}$ and $d_{k}=\frac{\lambda}{\lambda_{k}} u_{k}$, we have $d_{k}=\frac{\lambda_{k}}{\lambda} d_{k}-f_{k}$ and $d_{k}=\frac{\lambda f_{k}}{\lambda_{k-\lambda}}$. So $d(x)=\lambda \sum \frac{f_{k} \phi_{k}(x)}{\lambda_{k}-\lambda}$ and therefore $u(x)=f(x)+\lambda \sum \frac{f_{k} \phi_{k}(x)}{\lambda_{k}-\lambda}$. This is Eq. 3.8. Substituting for $f_{k}$ gives us (3.7)

Example 3.6 Solve $u=f+\lambda K u$ where

$$
k(x, t)= \begin{cases}x(1-t) & 0 \leq x \leq t \\ t(1-x) & t \leq x \leq 1\end{cases}
$$

Consider the eigenvalue problem $\mathbf{u}=\lambda \mathrm{Ku}$, where
$\mathrm{Ku}=\int_{0}^{1} \mathrm{k}(\mathrm{x}, \mathrm{t}) \mathrm{u}(\mathrm{t}) \mathrm{dt}$. We can check that $\mathrm{k}(\mathrm{x}, \mathrm{t})$ above is zero at $x=0$ and $x=1$, so $u(0)=u(1)=0$ for any solution to the eigenvalue problem. We can expand $(\mathrm{Ku})(\mathrm{x})$ as

$$
(K u)(x)=\int_{0}^{x} t(1-x) u(t) d t+\int_{x}^{1} x(1-t) u(t) d t
$$

Now, we need to solve the eigenvalue problem - the standard method is to turn this integral equation into an o.d.e. by differentiating the equation $\mathrm{u}=\lambda \mathrm{Ku}$ twice w.r.t. x .

$$
\begin{aligned}
u(x) & =\lambda K u \\
\frac{d u}{d x} & =\lambda\left\{x(1-x) u(x)+\int_{0}^{x}(-t) u(t) d t\right. \\
& \left.-x(1-x) u(x)+\int_{x}^{1}(-t) u(t) d t(1-t) u(t) d t\right\} \\
\frac{d^{2} u}{d x^{2}} & =\lambda\{-x u(x)-(1-x) u(x)\}
\end{aligned}
$$

which simplifies to

$$
u^{\prime \prime}=-\lambda u ; \quad u(0)=u(1)=0
$$

For negative $\lambda$ we only get the trivial solution $u(x) \equiv 0$ so write $\lambda=c^{2} \geq 0$. This gives $\mathfrak{u}(x)=A \cos c x+B \sin c x$. As $u(0)=0$ we have $\mathrm{A}=0$ and $\mathbf{u}(1)=0$ gives us $\mathrm{B} \sin \mathrm{c}=0$ so $\mathrm{c}=\mathrm{n} \pi$.

Therefore the eigenvalues (for the o.d.e.) are $\lambda_{n}=n^{2} \pi^{2}$ and the corresponding eigenfunctions are $\phi_{n}(x)=B_{n} \sin n \pi x$. We fix $B_{n}$ by requiring that $\left\|\phi_{n}\right\|=1$ which gives us $\mathrm{B}_{\mathrm{n}}=\sqrt{2}$.
So for $\lambda \neq \lambda_{n} \equiv n^{2} \pi^{2}$ we have a solution for any $\mathrm{f}(\mathrm{x})$.
Take $\mathrm{f}(\mathrm{x})=\mathrm{e}^{\mathrm{x}}$, then $\mathrm{u}(\mathrm{x})=\mathrm{f}(\mathrm{x})+\lambda \sum_{\mathrm{n}=1}^{\infty} \frac{\mathrm{f}_{\mathrm{n}} \phi_{\mathrm{n}}(\mathrm{x})}{\lambda_{\mathrm{n}}-\lambda}$ where
$f_{n}=\left(e^{x}, \phi_{\mathfrak{n}}(x)\right)$. Use the trick of replacing $\sin \mathfrak{n} \pi x$ by $e^{i n \pi x}$ and taking the imaginary part of the answer. This allows us to perform the integral and find $f_{n}=-\frac{\sqrt{2} n \pi}{1+n^{2} \pi^{2}}\left((-1)^{n} e-1\right)$.
So the solution is

$$
u(x)=e^{x}+\lambda \sum_{n=1}^{\infty} \frac{f_{n} \phi_{n}(x)}{n^{2} \pi^{2}-\lambda}
$$

with $\mathrm{f}_{\mathrm{n}}$ and $\phi_{\mathrm{n}}(\mathrm{x})$ as above.

Finally, note that if $\lambda=\lambda_{n}$ for some $n$ then a solution exists only if the corresponding $f_{n}=0$. So require $\left(f, \phi_{n}\right)=0$ for a solution.

Example 3.7 Take $u(x)=\cos 2 x+2 \int_{0}^{\frac{\pi}{2}} k(x, t) u(t) d t$ with

$$
k(x, t)= \begin{cases}\sin x \cos t & 0 \leq x \leq t \\ \sin t \cos x & t \leq x \leq \pi / 2\end{cases}
$$

It is easy to check that $\mathrm{k}(\mathrm{x}, \mathrm{t})$ is symmetric and square integrable on $\left[0, \frac{\pi}{2}\right] \times\left[0, \frac{\pi}{2}\right]$.

We need to solve the eigenvalue problem $u=\lambda \mathrm{Ku}$ so as before we reduce $u=\lambda К u$ to an o.d.e.

$$
\begin{aligned}
u(x) & =\lambda\left[\int_{0}^{x} \cos x \sin t u(t) d t+\int_{x}^{\frac{\pi}{2}} \sin x \cos t u(t) d t\right] \\
\frac{d u}{d x} & =\lambda\left[\cos x \sin x u(x)+\int_{0}^{x}(-\sin x) \sin t u(t) d t\right. \\
& \left.-\sin x \cos x u(x)+\int_{x}^{\frac{\pi}{2}} \cos x \cos t u(t) d t\right] \\
\frac{d^{2} u}{d x^{2}} & =\lambda\left\{-\sin ^{2} x u(x)-\int_{0}^{x} \cos x \sin t u(t) d t\right. \\
& \left.-\cos ^{2} x u(x)-\int_{x}^{\frac{\pi}{2}} \sin x \cos t u(t) d t\right\} \\
& =-(1+\lambda) u(x)
\end{aligned}
$$

So our eigenvalue equation reduces to $u^{\prime \prime}=-(1+\lambda) u$. It is easy to check that the definition of the kernel implies that $u(0)=u(\pi / 2)=0$.

As usual there is no non-trivial solution for $1+\lambda<0$ or equivalently for $\lambda<-1$. So we take $1+\lambda=\mathrm{k}^{2} \geq 0$ and $\mathfrak{u}(\mathrm{x})=\mathrm{A} \cos \mathrm{kx}+\mathrm{B} \sin \mathrm{kx}$. As usual, $\mathrm{A}=0$ as $\mathfrak{u}(0)=0$ and $B \sin k \frac{\pi}{2}=0$ so $k=2 n$.
We therefore have $\lambda_{n}=4 n^{2}-1$ and $\phi_{n}=B_{n} \sin 2 n x$. If we take the usual normalisation $\left\|\phi_{\mathrm{n}}\right\|^{2}=1$ we find that $\mathrm{B}_{\mathrm{n}}=\frac{2}{\sqrt{\pi}}$ and so $\phi_{n}=\frac{2}{\sqrt{\pi}} \sin 2 n x$.
Now we note that $\lambda_{n}=4 n^{2}-1$ is always odd. So if, for example, we take $\lambda=2$ we expect a unique solution for any $\mathrm{f}(\mathrm{x})$.

To illustrate what happens, take $\mathrm{f}(\mathrm{x})=\cos 2 \mathrm{x}$. So $($ for $\lambda=2)$

$$
\begin{aligned}
u(x) & =\cos 2 x+2 \sum_{1}^{\infty} \frac{f_{n} \phi_{n}(x)}{\lambda_{n}-\lambda} \\
& =\cos 2 x+\frac{2.2}{\sqrt{\pi}} \sum \frac{f_{n}}{\left(4 n^{2}-3\right)} \sin 2 n x
\end{aligned}
$$

where

$$
\begin{aligned}
f_{n} & =\frac{2}{\sqrt{\pi}} \int_{0}^{\frac{\pi}{2}} \sin 2 n x \cos 2 x d x \\
& =\left\{\begin{array}{lll}
\frac{2}{\sqrt{\pi}} \cdot \frac{n}{n^{2}-1} & n & \text { even } \\
0 & n & \text { odd }
\end{array}\right.
\end{aligned}
$$

Suppose that instead of $\lambda=2$ we take $\lambda=3$ then $\lambda_{1}=\lambda=3$. For $a$ solution to exist we must have $\mathrm{f}_{1}=0$ - this is the case for $\mathrm{f}(\mathrm{x})=\cos 2 \mathrm{x}$. So a solution exists.

If, on the other hand, $\lambda=3$ but $f(x)=\sin 2 x$ then there is no solution as $\mathrm{f}_{1}=\frac{2}{\sqrt{\pi}} \int_{0}^{\frac{\pi}{2}} \sin ^{2} 2 x \mathrm{~d} x \neq 0$.

### 3.3 Exercises

1. Show that the First Fredholm Alternative Theorem 3.1 is equivalent to Theorem 3.2 on the succeeding slide. Hint: note that the latter is asserting $A \vee B-$ but that $A$ and $B$ cannot both be true. (Why?) Show that $(A \vee B) \wedge(A \wedge B)^{\prime}$ is equivalent to $A \Leftrightarrow B^{\prime}$ (and $B \Leftrightarrow A^{\prime}$ ). You should now be able to see that Theorem 3.1 is equivalent to Theorem 3.2.
2. Does the operator

$$
K u(x)=\int_{0}^{\pi} \sin x \sin 2 y u(y) d y
$$

have any eigenvalues?
3. Solve the integral equation

$$
\int_{0}^{1} k(x, y) u(y) d y-\lambda u(x)=x
$$

using eigenfunction expansions where

$$
k(x, y)= \begin{cases}x(1-y), & x<y \\ y(1-x), & x>y\end{cases}
$$

4. Investigate the existence of solutions to

$$
u(x)=\sin x+3 \int_{0}^{\pi}(x+y) u(y) d y
$$

5. Solve the separable integral equation

$$
\int_{0}^{1} e^{x+y} u(y) d y-\lambda u(x)=f(x)
$$

Examine any special cases carefully.

## Part II

## Green's Functions

## 4 Introduction

The most general linear second-order o.d.e. can be written

$$
a(x) u^{\prime \prime}(x)+b(x) u^{\prime}(x)+c(x) u(x)=f(x)
$$

with either initial values or boundary conditions. We will study solution methods for a particular class of o.d.e.s - the Sturm-Liouville problems.

Definition 4.1 (Sturm-Liouville problem)

$$
\begin{align*}
A u & \equiv-\left(p(x) u^{\prime}(x)\right)^{\prime}+q(x) u(x)=f(x), \quad a<x<b  \tag{4.1a}\\
B_{1} u(a) & \equiv \alpha_{1} u(a)+\alpha_{2} u^{\prime}(a)=0  \tag{4.1b}\\
B_{1} u(b) & \equiv \beta_{1} u(b)+\beta_{2} u^{\prime}(b)=0 \tag{4.1c}
\end{align*}
$$

### 4.1 Examples of Sturm-Liouville problems

Any o.d.e. of the form $a(x) u^{\prime \prime}+b(x) u^{\prime}+c(x) u=0$ can be
transformed into the form $-\left(p u^{\prime}\right)^{\prime}+q u=0$ using integrating factors - though the integration cannot always be carried out explicitly. See Q 1 for more examples.
[Bessel's Equation]

$$
\begin{array}{r}
x^{2} u^{\prime \prime}+x u^{\prime}+\left(x^{2}-n^{2}\right) u=0 \\
\left(x u^{\prime}\right)^{\prime}+\left(x-n^{2} / x\right) u=0 \\
p(x) \equiv-x \quad q(x) \equiv x-n^{2} / x \quad \text { defined on }(0, \infty)
\end{array}
$$

[Legendre's Equation]

$$
\begin{aligned}
\left(1-x^{2}\right) u^{\prime \prime}-2 x u^{\prime}+\mathfrak{n}(n+1) \mathfrak{u} & =0 \\
\left(\left(1-x^{2}\right) \mathfrak{u}^{\prime}\right)^{\prime}+\mathfrak{n}(n+1) \mathfrak{u} & =0 \\
p(x) \equiv x^{2}-1 \quad q(x) \equiv \mathfrak{n}(n+1) \quad \text { defined on }[-1,1] &
\end{aligned}
$$

### 4.2 Discussion of definition of Sturm-Liouville problem

Note that the boundary conditions (4.1b) \& (4.1c) are separable (conditions are specified separately at $a$ and $b$ ) and homogeneous.
(The latter condition may be relaxed, allowing boundary
conditions like $\mathfrak{u}(\mathfrak{a})=\mathfrak{u}_{\boldsymbol{a}}$ and $\mathfrak{u}(\mathfrak{b})=\mathfrak{u}_{\mathfrak{b}}$. We will see later how to do this.)

In (4.1a), $\mathcal{A}$ is a differential operator that maps the function $\mathfrak{u}(x)$ into the function $f(x)$ - we write $A u=f$. It is convenient to incorporate the boundary conditions $\mathrm{B}_{1} \mathrm{u}(\mathrm{a})=0$ and $\mathrm{B}_{2} \mathfrak{u}(\mathrm{~b})=0$ into a single (symbolic) operator L so the whole Sturm-Liouville problem (4.1a)-(4.1c) can be represented by the single equation $\mathrm{Lu}=\mathrm{f}$.

We think of $L$ as operating on the set of $C^{2}[a, b]$ functions that satisfy the boundary conditions (4.1b) \& (4.1c). So L incorporates the boundary conditions in its definition as well as the
Sturm-Liouville operator $A$ where $A u \equiv-\frac{d}{d x}\left(p(x) \frac{d u}{d x}\right)+q(x) u$.

The form of the equation $L u=f$ is reminiscent of the familiar linear system $A x=b$ and suggests that we try to find an "inverse" operator $L^{-1}$ so that we can write $u=L^{-1} f$ and so solve the Sturm-Liouville problem. As L is a differential operator it is plausible that the inverse operator will be an integral operator of the form

$$
\begin{equation*}
u(x)=\left(L^{-1} f\right)(x) \equiv \int_{a}^{b} g(x, \xi) f(\xi) d \xi \tag{4.2}
\end{equation*}
$$

Note that this equation is similar to the solution to a Fredholm integral equation of the second kind using a resolvent kernel $\Gamma$ in (3.7). Again drawing on the analogy with matrix algebra we expect that if $\mathrm{Lu}=0$ has non-trivial solutions then $\mathrm{L}^{-1}$ will not be defined. If, on the other hand, $\mathrm{Lu}=0$ has only the trivial solution and the kernel $g(x, \xi)$ is defined, we call $g$ the Green's Function associated with L.

### 4.3 Definition of Green's Function

We will see later that physically $g(x, \xi)$ is the response at $\chi$ of the system represented by $L u=\mathrm{f}$ when the system is acted on at the point $\xi$ by a unit point source.

First, back to the mathematics. We state and prove a theorem that gives an explicit formula for the Green's Function $g(x, \xi)$ and confirms that it defines an inverse for the Sturm-Liouville operator
$L$ and therefore a solution for the Sturm-Liouville problem Lu $=\mathrm{f}$. The formula on the next slide is complicated and we will see later that there is a simpler alternative.

Theorem 4.1 (Green's Function) Given the Sturm-Liouville problem $\mathrm{Lu}=\mathrm{f}$, if the homogeneous problem $\mathrm{Lu}=0$ has only the trivial solution then the inverse operator $\mathrm{L}^{-1}$ exists and is defined by (4.2) where

$$
g(x, y)= \begin{cases}\frac{-u_{1}(x) \mathfrak{u}_{2}(y)}{p(y) W(y)}, & x<y  \tag{4.3}\\ \frac{-u_{1}(y) \mathfrak{u}_{2}(x)}{\mathfrak{p}(y) W(y)}, & x>y\end{cases}
$$

Here $\mathfrak{u}_{1}$ and $u_{2}$ are the solutions to $\mathrm{Au}=0$ with $\mathrm{B}_{1} \mathfrak{u}(\mathrm{a})=0$ and $\mathrm{B}_{2} \mathrm{u}(\mathrm{b})=0$ repectively. (Note if both boundary conditions held then by assumption the only solution is $\mathbf{u}=0$.) The function $\mathbf{W}(\mathrm{y})$ is called the Wronskian and is defined by

$$
\begin{equation*}
W(y)=u_{1}(y) u_{2}^{\prime}(y)-u_{2}(y) u_{1}^{\prime}(y) \tag{4.4}
\end{equation*}
$$

Before proving Thm. 4.1 we remind ourselves that we can now write the solution as $u(x)=\int_{a}^{b} g(x, y) f(y) d y$. We can write $g(x, y)$ as a single function using the Heaviside function $H$ that is zero for $x<0$ and equals one for $x \geq 0$ :

$$
\begin{equation*}
g(x, y)=-\frac{1}{p(y) W(y)}\left(u_{1}(x) u_{2}(y) H(y-x)+u_{1}(y) u_{2}(x) H(x-y)\right) \tag{4.5}
\end{equation*}
$$

Now we state the defining properties of the Green's Function that follow from the definition (4.3).
(a) $\operatorname{Ag}(x, y)=0$ for all $x \neq y$ ( $y$ is treated as a fixed parameter). This follows directly from the definition (4.3).
(b) $g(x, y)$ satisfies the boundary conditions
(i) $B_{1} u(a)=0$ as $u_{1}$ satisfies this boundary condition by definition.
(ii) $\mathrm{B}_{2} u(b)=0$ as $u_{2}$ satisfies this boundary condition by definition.
(c) $g(x, y)$ is continuous on $[a, b]$ including $x=y$ (w.r.t. both $x$ and $y$ separately).
(d) $g^{\prime}(x, y)$ is not continuous at $x=y$. To see this, calculate $g^{\prime}(x, y)$ from the definition (4.3):

$$
g^{\prime}(x, y)= \begin{cases}-\frac{u_{1}^{\prime}(x) u_{2}(y)}{p(y) W(y)}, & x<y \\ \frac{-u_{1}(y) u_{2}^{\prime}(x)}{p(y) W(y)}, & x>y .\end{cases}
$$

So
$g^{\prime}(y+\epsilon, y)-g^{\prime}(y-\epsilon, y)=\frac{-u_{1}(y) u_{2}^{\prime}(y+\epsilon)+u_{1}^{\prime}(y-\epsilon) u_{2}(y)}{p(y) W(y)}$.
Taking the limit as $\epsilon \rightarrow 0^{+}$the top line is just $-W(y)$ so we find that

$$
\begin{equation*}
\left.\Delta g^{\prime}(x, y)\right|_{x=y}=-\frac{1}{p(y)} \tag{4.6}
\end{equation*}
$$

These four properties define $g(x, y)$ uniquely for any
Sturm-Liouville problem and can be used to calculate the Green's
Function directly - bypassing the definition (4.3).
Before proving the Theorem, an example.
Example 4.1 Consider the (very simple) Sturm-Liouville problem $-\mathfrak{u}^{\prime \prime}(\mathrm{x})=\mathrm{f}(\mathrm{x})$ on $0<\mathrm{x}<1$ with $\mathfrak{u}(0)=\mathfrak{u}(1)=0$. Let's use the four defining properties to construct the Green's Function for this problem. Here $\boldsymbol{A u}=-\mathfrak{u}^{\prime \prime}$. Use each property in turn:
(a) Solutions to $\mathrm{A} \mathfrak{u}=0$ are just $\mathfrak{u}=\mathrm{ax}+\mathrm{b}$.
(b) The boundary conditions require that $\mathfrak{u}_{1}(x)=a x$ and $u_{2}(x)=b(1-x)$.
(c) Continuity at $\mathrm{x}=\mathrm{y}$ implies that $\mathrm{a} \mathrm{y}=\boldsymbol{b}(1-\mathrm{y})$.
(d) $\left.\Delta \mathrm{g}\right|_{\mathrm{x}=\mathrm{y}}=-1 / \mathrm{p}(\mathrm{y}) \equiv-1$ so as $\mathrm{g}^{\prime}=\mathrm{a}$ for $\mathrm{x}<\mathrm{y}$ and $\mathrm{g}^{\prime}=-\mathrm{b}$ for $\mathrm{x}>\mathrm{y}$ we have $-\mathrm{b}-\mathrm{a}=-1$ or $\mathrm{a}+\mathrm{b}=1$.

Combining the two equations for a and b gives $\mathrm{b}=\mathrm{y}$ and $a=1-y$, so

$$
g(x, y)= \begin{cases}x(1-y), & x<y \\ y(1-x), & x>y .\end{cases}
$$

Example 4.2 We can, of course, calculate $\mathrm{g}(\mathrm{x}, \mathrm{y})$ directly from the definition (4.3). For the Sturm-Liouville problem above, we have $\mathrm{p}(\mathrm{y})=1$ and $\mathrm{W}(\mathrm{y})=\mathrm{ay}(-\mathrm{b})-\mathrm{b}(1-\mathrm{y}) \mathrm{a}$ which reduces to $\mathrm{W}(\mathrm{y})=-\mathrm{ab}$. So substituting directly into the formula for $\mathrm{g}(\mathrm{x}, \mathrm{y})$ we find as expected

$$
g(x, y)= \begin{cases}\frac{-a \times b(1-y)}{-a b}=x(1-y), & x<y \\ \frac{-b(1-x) a y}{-a b}=y(1-x), & x>y .\end{cases}
$$

It is usually easier to use the defining properties to calculate the Green's Function.

So the solution to the above simple Sturm-Liouville problem is

$$
u(x)=\int_{0}^{1} g(x, y) f(y) d y
$$

and the function $g(x, y)$ is piecewise linear - sketch it!
We now prove the Theorem.

Proof: (of Thm 4.1) We need to show that the function $u(x) \equiv \int_{a}^{b} g(x, y) f(y) d y$-with $g$ defined as in (4.3) — satisfies the o.d.e. $\mathrm{Au}=\mathrm{f}(4.1 \mathrm{a})$ together with the boundary conditions (4.1b) \& (4.1c). We can expand the integral for $u(x)$ using the two-fold definition of $g(x, y)$ to:

$$
\begin{equation*}
u(x)=-u_{2}(x) \int_{a}^{x} \frac{u_{1}(y)}{p(y) W(y)} d y-u_{1}(x) \int_{x}^{b} \frac{u_{2}(y)}{p(y) W(y)} d y \tag{4.7}
\end{equation*}
$$

- First check that $\mathfrak{u}(x)$ satisfies the boundary conditions. From (4.7) we have that $u(a)=-u_{1}(a) K_{2}$, where $K_{2}=\int_{a}^{b} \frac{u_{2}(y)}{p(y) W(y)} d y$. Also, use the Leibnitz formula (2.5) to check that $u^{\prime}(a)=-u_{1}^{\prime}(a) K_{2}$. It follows, as $B_{1} u_{1}(a)=0$, that $B_{1} u(a)=0$. A similar argument shows that $B_{2} u(b)=0$.
- Now we check that $A \mathfrak{u}=0, x \neq y$, with $\mathfrak{u}(x)$ given by (4.7). We just differentiate (4.7) w.r.t. $x$, giving (after cancellations)

$$
\begin{aligned}
u^{\prime}(x) & =-u_{2}^{\prime}(x) \int_{a}^{x} \frac{u_{1}(y) f(y) d y}{p(y) W(y)}-u_{1}^{\prime}(x) \int_{x}^{b} \frac{u_{2}(y) f(y) d y}{p(y) W(y)} \\
\left(p(x) u^{\prime}(x)\right)^{\prime} & =-\left(p u_{2}^{\prime}\right)^{\prime} \int_{a}^{x} \frac{u_{1}(y) f(y)}{p(y) W(y)} d y-\frac{p(x) \mathbf{u}_{2}^{\prime}(x) \mathbf{u}_{1}(x) f(x)}{p(x) W(x)} \\
& -\left(p u_{1}^{\prime}\right)^{\prime} \int_{x}^{b} \frac{u_{2}(y) f(y)}{p(y) W(y)} d y+\frac{p(x) u_{1}^{\prime}(x) \mathbf{u}_{2}(x) f(x)}{p(x) W(x)}
\end{aligned}
$$

So combining the second and fourth (highlighted in blue) terms in the latter equation - using the definition (4.4) of the Wronskian $\mathrm{W}(\mathrm{y})$;

$$
\begin{aligned}
\left(p u^{\prime}\right)^{\prime}(x)=-f(x)-\left(p u_{2}^{\prime}\right)^{\prime}(x) & \int_{a}^{x} \frac{u_{1}(y) f(y)}{p(y) W(y)} \\
& -\left(p u_{1}^{\prime}\right)^{\prime}(x) \int_{x}^{b} \frac{u_{2}(y) f(y)}{p(y) W(y)} d y .
\end{aligned}
$$

Finally, assembling $-\left(p u^{\prime}\right)^{\prime}(x)+q(x) u(x)$ and using $-\left(p u_{\mathfrak{i}}^{\prime}\right)^{\prime}(x)=-q u_{i}(x), \mathfrak{i}=1,2$ we have

$$
\begin{aligned}
-\left(p u^{\prime}\right)^{\prime}(x)+q(x) u(x) & =f(x)+\left(q u_{2}\right)(x) \int_{a}^{x} \frac{u_{1}(y) f(y)}{p(y) W(y)} d y \\
& +\left(q u_{1}\right)(x) \int_{x}^{b} \frac{u_{2}(y) f(y)}{p(y) W(y)} d y \\
& -q(x) u_{2}(x) \int_{a}^{x} \frac{u_{1}(y) f(y)}{p(y) W(y)} d y \\
& -q(x) u_{1}(x) \int_{x}^{b} \frac{u_{2}(y) f(y)}{p(y) W(y)} d y
\end{aligned}
$$

When we examine the RHS, all the integral terms cancel and we are left with

$$
-\left(p u^{\prime}\right)^{\prime}(x)+q(x) u(x)=f(x)
$$

Example 4.3 Solve the Sturm-Liouville problemu" $+4 u=f(x)$ with $u(0)=u(1)=0$.

Solution: The Green's Function $\mathrm{G}(\mathrm{x}, \mathrm{y})$ satisfies
$g^{\prime \prime}+4 g=0, \quad x \neq y$. So the general solution is
$\mathrm{g}=\mathrm{A} \cos 2 \mathrm{x}+\mathrm{B} \sin 2 \mathrm{x}$. At $\mathrm{x}=0$ the boundary condition requires
that $A=0$. So $\mathrm{g}(\mathrm{x}, \mathrm{y})=A \sin 2 \mathrm{x}$ for $\mathrm{x}<\mathrm{y}$.
For $\mathrm{x}>\mathrm{y}$ we have $\mathrm{g}=\mathrm{C} \cos 2 \mathrm{x}+\mathrm{D} \sin 2 \mathrm{x}$.
Important Note: use different dummy variables at the two boundaries. At $\mathrm{x}=1$, we must have $\mathrm{C} \cos 2+\mathrm{D} \sin 2=0$.

So $\mathrm{C}=-\mathrm{D} \sin 2 / \cos 2$ which gives us:

$$
\begin{aligned}
g(x, y) & =\frac{D}{\cos 2}(-\cos 2 x \sin 2+\sin 2 x \cos 2) \\
& =D \sin (2 x-2) \quad \text { for } x>y
\end{aligned}
$$

- dropping $\cos 2$ in denominator as D is arbitrary.

Now require that $\mathrm{g}(\mathrm{x}, \mathrm{y})$ be continuous at $\mathrm{x}=\mathrm{y}$. This gives us:

$$
A \sin 2 y=D \sin (2 y-2)
$$

or

$$
A=D \frac{\sin (2 y-2)}{\sin 2 y}
$$

Finally we know that the "jump" discontinuity in $\mathrm{g}^{\prime}$ at $\mathrm{x}=\mathrm{y}$ is $\Delta \mathrm{g}(\mathrm{y}, \mathrm{y})=-1 / \mathrm{p}(\mathrm{y})=1 . S o$

$$
2 D \cos (2 y-2)-2 D \frac{\sin (2 y-2)}{\sin 2 y} \cos 2 y=1
$$

Simplifying:

$$
\begin{aligned}
2 \frac{\mathrm{D}}{\sin 2 y}(\sin 2 y \cos (2 y-2)-\sin (2 y-2) \cos 2 y) & =1 \\
\text { which reduces nicely to: } 2 \mathrm{D} \frac{\sin 2}{\sin 2 y} & =1 \\
\text { or just } \mathrm{D} & =\frac{\sin 2 \mathrm{~s}}{2 \sin 2} .
\end{aligned}
$$

So the final expression for $\mathrm{g}(\mathrm{x}, \mathrm{y})$ is

$$
g(x, y)= \begin{cases}\frac{\sin (2 y-2) \sin 2 x}{2 \sin 2} & x<y \\ \frac{\sin (2 x-2) \sin 2 y}{2 \sin 2} & x>y .\end{cases}
$$

Exercise 4.1 Solve the above Sturm-Liouville problem with $f(x)=x$.

In Exercise 5 at the end of this Chapter you are asked to solve the Exercise using the original definition of the Green's Function (4.3).
[Note 1.] It is easy to get confused as to which "piece" of the Green's Function to use when calculating the solution:

$$
u(x)=\int_{a}^{b} g(x, y) f(y) d y
$$

Remember that we are integrating w.r.t. $y$ so use the $x>y$ piece when integrating over the left hand part of the interval (from 0 to $x$ in the above Example) and the $x<y$ piece when integrating over the right hand part of the interval (from $x$ to 1 in the above Example
[Note 2.] This simple Sturm-Liouville problem (with $f(x)=x$ ) can be solved much more easily by finding a particular solution to the ode, adding it to the general solution to the homogeneous equation and using the boundary conditions. But this ad-hoc approach is of little use when the inhomogeneous term is more general or the ode more complicated.

### 4.4 Physical Interpretation of Green's Functions

For the sake of definiteness we'll consider heat flow in a one-dimensional bar. The heat equation is

$$
u_{t}-u_{x x}=f(x, t)
$$

where $f$ is the heat source over the length of the bar. Suppose that the ends of the bar are kept at a constant $0^{\circ}$ then the steady state (no time dependence) temperature distribution satisfies

$$
-u_{x x}=f(x), \quad 0<x<1 ; \quad u(0)=u(1)=0
$$

This is the problem that we solved in Example 4.1 using Green's Functions.

Now suppose that $f(x)$ is a heat source of unit strength (i.e. the heat per unit time being supplied to the bar is one - in appropriate units) that acts at a single point $x=y$ in $(0,1)$.

There are two ways of looking at this scenario:
Unreal: "heat wire at the point $x=y$ to an infinite temperature" "using a unit amount of energy per unit time"
Real: "heat the section of wire near $x=y$ to high temperature" "using a unit amount of energy per unit time"

The unreal scenario is a mathematical idealisation that is easier to handle than the actual situation.

We will use the symbol $\delta(x, y)$ to denote this "unit point source at y" so we write:

$$
-u_{x x}=\delta(x, y)
$$

- This unit source or "delta function" acts only at a single point - so we must have

$$
\delta(x, y)=0, \quad x \neq y
$$

- As the source has unit strength, we must have

$$
\int_{0}^{1} \delta(x, y) d x=1
$$

There is no function that satisfies these two conditions as the first condition implies that the integral is zero! We will see later how this mathematical difficulty can be resolved. For the moment we will think of $\delta(x, y)$ as a function that is zero for $x \neq y$ and has a "spike" at $x=y$ so that it integrates to 1 .

Now, continuing the sloppy physical description, if the heat source is $f(x)=\delta(x, y)$, we have:

$$
\begin{equation*}
-u^{\prime \prime}=\delta(x, y) \tag{4.8}
\end{equation*}
$$

For $x \neq y$ we just have $u^{\prime \prime}=0$ so $u(x)=a x+b$, for $x \neq y$. For $x<y$ we have $u=a x$ as $u(0)=0$. For $x>y$ we have $u=b(1-x)$ as $u(1)=0$.

To get a second equation for $\mathfrak{a}$ and $\boldsymbol{b}$ we just integrate (4.8) over any small interval round $x=y$ so

$$
-\int_{y-\epsilon}^{y+\epsilon} u^{\prime \prime}(x) d x=\int_{y-\epsilon}^{y+\epsilon} \delta(x, y) \equiv 1
$$

So

$$
-u^{\prime}(y+\epsilon)+u^{\prime}(y-\epsilon)=1
$$

Taking the limit as $\epsilon \rightarrow 0^{+}$gives us

$$
\Delta u_{x=y}^{\prime}=-1
$$

This is just the "jump condition" we saw when we solved the same equation using Green's Functions. So (as before) $-b-a=-1$, $a=1-y, b=y$ and

$$
u(x)= \begin{cases}x(1-y), & x<y \\ y(1-x), & x>y\end{cases}
$$

We have re-constructed the Green's Function for the problem. So now we have a physical interpretation of the Green's Function : it is the response of the system to a unit point source at $x=y$.

In general (for any Sturm-Liouville problem) the Green's
Function $g(x, y)$ is the solution to the symbolic boundary value problem:

$$
A g(x, y) \equiv-\left(p(x) g^{\prime}(x, y)\right)^{\prime}+q(x, y) g(x, y)=\delta(x, y), \quad a<x<b
$$

$B_{1} g(a, y)=0$
$B_{2} g(b, y)=0$.
The solution $u(x)$ to the Sturm-Liouville problem $L u=f$, namely $u=\int_{a}^{b} g(x, y) f(y) d y$, can be interpreted as the response of the system to the superposition of point sources of magnitude $f(y)$ over the whole interval $a<y<b$.

Example 4.4 Consider the same differential operator $\mathrm{Au} \equiv-\mathbf{u}^{\prime \prime}$ on $0<x<1$ but with boundary conditions $u^{\prime}(0)=u^{\prime}(1)=0-$ perfect insulators at each end so no heat flowing into/out of wire.

Now the Green's Function does not exist as the equation $\mathrm{Lu}=0$ does have non-trivial solutions - namely $\mathfrak{u}(x)$ constant, $0<x<1$ and so the operator L cannot have an inverse.

Physically; again interpreting the problem as steady-state heat flow, the boundary conditions mean that no heat can pass through the end-points $x=0$ and $x=1$ so heat cannot escape. If we imposed a unit heat source - inserting heat energy at a constant unit rate at the point $\mathrm{x}=\mathrm{y}$ - heat would build up in the bar, preventing a steady-state solution.

Note: if we try to go ahead and construct a Green's Function using the defining formula (4.3) in terms of $\mathfrak{u}_{1}$ and $\mathfrak{u}_{2}$, we find that $\mathrm{u}_{1}(\mathrm{x})=\mathrm{K}_{1}$ and $\mathrm{u}_{2}(\mathrm{x})=\mathrm{K}_{2}$ so the Wronskian $\mathrm{W}(\mathrm{x})=0$ and so the Green's Function $\mathrm{g}(\mathrm{x}, \mathrm{y})$ is not defined as W appears in the denominator. (The Wronskian is zero when the two functions are not linearly independant - i.e. one is a multiple of the other - as here.)

We know that $g(x, y)$ is not differentiable at $x=y$ so how can it make sense to apply the differential operator
$A \mathfrak{u}=-\left(p(x) \mathfrak{u}^{\prime}(x)\right)^{\prime}+q(x) \mathfrak{u}(x)$ to it?
Intuitively; $g$ is twice differentiable except at $x=y$ and is continuous at $x=y, g^{\prime}$ is differentiable except at $y$ with a discontinuity at $y$ and $g^{\prime \prime}$ is continuous except at $y$. As $g^{\prime}$ has a jump at $x=y$ it is not surprising that
$g^{\prime \prime}(y) \equiv \lim _{\epsilon \rightarrow 0} \frac{1}{\epsilon}\left(g^{\prime}(y+\epsilon)-g^{\prime}(y-\epsilon)\right)$ is not defined but for any small non-zero value of $\epsilon$, we have $\Delta g^{\prime} \approx-\frac{1}{\mathfrak{p}(y)}$ so

$$
\frac{\left.\Delta\left(\mathrm{pg}^{\prime}\right)\right|_{x=y}}{\epsilon} \approx-\frac{1}{\epsilon},
$$

"explaining" the spike when $\epsilon \rightarrow 0$ in $\delta(x, y)$.
These ideas obviously need tightening up - but if used carefully will give correct results. We will present a precise mathematical treatment later.

Initial Value Problems The present treatment for boundary value problems needs only a small change to apply Green's functions to initial value problems.

$$
A u \equiv-\left(p u^{\prime}\right)^{\prime}+q u=f(t), \quad t>0 \quad u(0)=u^{\prime}(0)=0
$$

The "causal" Green's Function (impulse response function) is the solution to the above when $f$ is a unit impulse applied at $t=\tau$, namely $\delta(t, \tau)$. So $A g(t, \tau)=\delta(t, \tau)$

To determine the Green's Function g, we note that as the initial data is zero $\left(u(0)=u^{\prime}(0)=0\right)$ and as the impulse does not occur till time $t=\tau$, we must have $g(t, \tau)=0, t<\tau$. For $t>\tau$ we must have

$$
\operatorname{Ag}(\mathrm{t}, \tau)=0
$$

The rest of the analysis is familiar.
We require that $g$ be continuous at $t=\tau$ so $g(\tau, \tau)=0$.
At $t=\tau$ we require that $g^{\prime}(t, \tau)$ have a jump of magnitude
$\Delta g^{\prime} \equiv g^{\prime}\left(\tau^{+}, \tau\right)-0=\frac{-1}{p(\tau)}$.
The above is sufficient to determine $g(t, \tau)$ for $t>\tau$.
Example 4.5 (Spring-mass system) We have
$\mathfrak{m} \ddot{u}+\mathrm{ku}=\mathrm{f}(\mathrm{t}), \quad \mathrm{t}>0 \quad \mathbf{u}(0)=\dot{u}(0)=0$, where $\mathbf{u}$ is
displacement from equilibrium and f is the applied force.
This is a Sturm-Liouville problem with $\mathrm{p}=-\mathrm{m}$ and $\mathrm{q}=\mathrm{k}$. We have $\mathrm{g}(\mathrm{t}, \tau)=0$ for $\mathrm{t}<\tau$ and $\mathrm{m} \ddot{g}+\mathrm{kg}=0$ for $\mathrm{t}>\tau$. So $\ddot{g}(t, \tau)=-k / m g, k, m>0$. Set $k / m=\lambda^{2}$. Then

$$
g=A \cos \lambda t+B \sin \lambda t, \quad t>\tau
$$

Continuity at $\mathrm{t}=\tau$ implies that $0=A \cos \lambda \tau+\mathrm{B} \sin \lambda \tau$ so $A=-B \tan \lambda \tau$.

The condition $\Delta \mathrm{g}^{\prime}=-\frac{1}{\mathrm{p}}=\frac{1}{\mathrm{~m}}$ implies that $g^{\prime}\left(\tau^{+}, \tau\right)=\lambda(-A \sin \lambda \tau+B \cos \lambda \tau)=\frac{1}{m}$.
Doing the algebra, we find $B=\frac{1}{m \lambda} \cos \lambda \tau$ and $A=-\frac{1}{m \lambda} \sin \lambda \tau$.
Finally, $g(t, \tau)=\frac{1}{m \lambda}(-\sin \lambda \tau \cos \lambda t+\cos \lambda \tau \sin \lambda)-$ simplifying gives

$$
g=\frac{1}{\sqrt{\mathrm{mk}}} \sin \lambda(\mathrm{t}-\tau), \text { for } \mathrm{t}>\tau \text { and } 0 \text { for } \mathrm{t}<\tau
$$

Now the solution is defined for $\mathrm{t}>\tau$. So $\mathrm{u}(\mathrm{t})=\int_{0}^{\infty} \mathrm{g}(\mathrm{t}, \tau) \mathrm{f}(\tau) \mathrm{d} \tau$ - the sum over all responses. But as $\mathrm{g}=0$ for $\mathrm{t}<\tau$ we can write $u(t)=\frac{1}{\sqrt{\mathrm{~km}}} \int_{0}^{\mathrm{t}} \sin \lambda(\mathrm{t}-\tau) \mathrm{f}(\tau) \mathrm{d} \tau$. (As we would expect the solution at time t is only affected by inputs up to that time.)

### 4.5 Inhomogeneous Boundary Conditions

Up to now we have only considered problems with separable and homogeneous boundary conditions (4.1b) \& (4.1c). Allowing inhomogeneous boundary conditions means that we want to consider boundary conditions of the form

$$
\begin{aligned}
& B_{1} u(a) \equiv \alpha_{1} u(a)+\alpha_{2} u^{\prime}(a)=\alpha \\
& B_{1} u(b) \equiv \beta_{1} u(b)+\beta_{2} u^{\prime}(b)=\beta
\end{aligned}
$$

There is a systematic way to generalise the Green's Function method to problems with boundary conditions like this by extending the definition of $g(x, y)$. Here we take a simpler approach - if $u(x)$ satisfies inhomogeneous boundary conditions then define $v(x)=u(x)+C x+D$ and choose $C$ and $D$ so that $v(x)$ satisfies the corresponding homogeneous boundary conditions.

Just substitute for $u(x)=v(x)-C x-D$ in the inhomogeneous boundary conditions above:

$$
\begin{aligned}
& \alpha_{1}(v(a)-C a-D)+\alpha_{2}\left(v^{\prime}(a)-C\right)=\alpha \\
& \beta_{1}(v(b)-C b-D)+\beta_{2}\left(v^{\prime}(b)-C\right)=\beta
\end{aligned}
$$

or

$$
\begin{aligned}
& \alpha_{1} v(a)+\alpha_{2} v^{\prime}(a)=\alpha+\alpha_{1}(C a+D)+\alpha_{2} C \\
& \beta_{1} v(b)+\beta_{2} v^{\prime}(b)=\beta+\beta_{1}(C b+D)+\beta_{2} C
\end{aligned}
$$

But $v(x)$ satisfies homogeneous boundary conditions so we must have:

$$
\begin{aligned}
& \alpha+\alpha_{1}(C a+D)+\alpha_{2} C=0 \\
& \beta+\beta_{1}(C b+D)+\beta_{2} C=0
\end{aligned}
$$

two equations for the two unknowns $C$ and $D$.

We could solve giving a general formula for $C$ and $D$. It is just as easy to apply the method from scratch to the problem at hand see the Example below.

Once $C$ and $D$ have been found then we just substitute for $u(x)=v(x)-C x-D$ in the ode $A u=f$ giving a new ode $A v=g$, say - where $g$ is just $f$ augmented with all the terms involving $C$ and $D$ from $u(x)=v(x)-C x-D$. Finally, once the solution $v(x)$ is found use the equation $u(x)=v(x)-C x-D$ again to finish the calculation.

An Example should make the above clear.

Example 4.6 Solve the inhomogeneous (mixed boundary conditions) Sturm-Liouville problem:

$$
\begin{aligned}
x^{2} u^{\prime \prime}-x u^{\prime}-8 u & =f \\
u(1) & =1 \\
u(2)+2 u^{\prime}(2) & =3
\end{aligned}
$$

Set $\mathbf{u}(\mathrm{x})=v(\mathrm{x})-\mathrm{C} \mathrm{x}-\mathrm{D}$ as above. So we have:

$$
\begin{aligned}
x^{2} v^{\prime \prime}-x\left(v^{\prime}-\mathrm{C}\right)-8(v-\mathrm{C} x-\mathrm{D}) & =\mathrm{f} \\
v(1)-\mathrm{C}-\mathrm{D} & =1 \\
v(2)-2 \mathrm{C}-\mathrm{D}+2\left(v^{\prime}(2)-\mathrm{C}\right) & =3
\end{aligned}
$$

Rewrite the boundary conditions in terms of $v$ :

$$
\begin{aligned}
v(1) & =\mathrm{C}+\mathrm{D}+1 & & =0 \\
v(2)+2 v^{\prime}(2) & =2 \mathrm{C}+\mathrm{D}+2 \mathrm{C}+3 & & =0
\end{aligned}
$$

So we have two equations

$$
\begin{aligned}
C+D & =-1 \\
4 C+D & =-3
\end{aligned}
$$

for the two unknowns C and D giving $\mathrm{C}=-2 / 3$ and $\mathrm{D}=-1 / 3$.
Finally, solve the homogeneous Sturm-Liouville problem

$$
\begin{aligned}
x^{2} v^{\prime \prime}-x v^{\prime}-8 v & =f-9 C x-8 D \\
v(1) & =0 \\
v(2)+2 v^{\prime}(2) & =0
\end{aligned}
$$

and find the solution $\mathfrak{u}(\mathrm{x})$ to the original inhomogeneous problem using $u(x)=v(x)-C x-D$.

The remainder of the problem is straightforward and is left as an Exercise.

### 4.6 L Non-invertible

Suppose again that, as in Example 4.4, $\mathrm{Lu}=0$ has a non-trivial solution. Based on our experience with integral equaltions we expect that there may not be a solution to $L u=f$ - and if there is it is not unique. Of course (as in the Example) when the solution is not unique (having specified the boundary conditions) the problem is "ill-posed" - meaning that the mathematical model is un-physical.
The following Theorem ties things up:
Theorem 4.2 Suppose the Sturm-Liouville problem defined above has a non-trivial solution $\phi$ to the homogeneous problem $\mathrm{L} \phi=0$. Then the inhomogeneous problem $\mathrm{Lu}=\mathrm{f}$ has a solution if and only if

$$
(\phi, f) \equiv \int_{a}^{b} \phi(x) f(x) d x=0
$$

## Proof:

[ $\Rightarrow$ ] Assume a solution $\mathfrak{u}$ exists to $L \mathfrak{u}=\mathrm{f}$. Then

$$
\begin{aligned}
(\phi, f) & =(\phi, A u) \\
& =-\int_{a}^{b} \phi\left(p u^{\prime}\right)^{\prime}+\int_{a}^{b} \phi q u \\
& =-\left.\phi p u^{\prime}\right|_{a} ^{b}+\int_{a}^{b}\left(p u^{\prime} \phi^{\prime}+\phi q u\right) \\
& =-\left.\phi p u^{\prime}\right|_{a} ^{b}+\left.u\left(p \phi^{\prime}\right)\right|_{a} ^{b}-\int_{a}^{b} u\left(p \phi^{\prime}\right)^{\prime}+\int_{a}^{b} \phi q u .
\end{aligned}
$$

So $(\phi, f)=\left.p\left(u \phi^{\prime}-\phi u^{\prime}\right)\right|_{\mathrm{a}} ^{\mathrm{b}} \quad \mathrm{T}_{1}+\int u A \phi d x \quad \mathbf{T}_{\mathbf{2}}$.
Now, $\mathrm{T}_{1}=0$ as $u$ and $\phi$ satisfy the separable homogeneous boundary conditions at $a$ and $b$. (Exercise: check.) Also, $T_{2}=0$ as $\phi$ satisfies the equation $A \phi=0$.
$[\Leftarrow]$ Assume that $(\phi, f)=0$. Let $v$ be independent of $\phi$ $(W(v, \phi) \neq 0)$ and satisfy $A v=0$ but not the boundary conditions. Now define a "pseudo-Green's function" (the actual Green's function does not exist as $\mathrm{L} \phi=0$ has a non-trivial solution)

$$
\begin{align*}
& G(x, y)= \\
& -\frac{1}{p(y) W(y)}[\phi(x) v(y) H(y-x)+\phi(y) v(x) H(x-y)] \\
& W=\phi v^{\prime}-v \phi^{\prime} \tag{4.9}
\end{align*}
$$

We must check that $u=\mathbf{c} \phi+\int_{a}^{b} G(x, y) f(y) d y$ satisfies $L u=f$ for any constant c. We can drop the $\mathbf{c} \phi$ term as $\phi$ satisfies $\mathrm{L} \phi=0$ by assumption. RTP that the remaining terms in $u(x)$ $-\bar{u}(x)$ say - satisfy $L \bar{u}=f$.

We have

$$
\bar{u}(x)=-\int_{a}^{x} \frac{\phi(y) v(x)}{B L(y)} f(y) d y-\int_{x}^{b} \frac{\phi(x) v(y)}{B L(y)} f(y) d y .
$$

So

$$
\begin{aligned}
\bar{u}^{\prime}=-\frac{\phi(\mathbf{x}) \mathbf{v}(\mathbf{x})}{\operatorname{BL}(\mathbf{x})} \mathbf{f}(\mathrm{x}) & -\int_{\mathrm{a}}^{x} \frac{\phi(\mathrm{y}) v^{\prime}(\mathrm{x}) f(\mathrm{y})}{\operatorname{BL}(\mathrm{y})} \mathrm{d} y \\
& +\frac{\phi(\mathbf{x}) \mathbf{v}(\mathbf{x})}{\operatorname{BL}(\mathbf{x})} \mathbf{f}(\mathbf{x})-\int_{x}^{b} \frac{\phi^{\prime}(x) v(\mathrm{y})}{\operatorname{BL}(\mathrm{y})} f(\mathrm{y}) \mathrm{d} y
\end{aligned}
$$

Therefore cancelling $T_{1}$ and $T_{3}$ on RHS,

$$
p \bar{u}^{\prime}=-\int_{a}^{x} \frac{\phi(y) v^{\prime}(x) p(x)}{B L(y)} d y-\int_{x}^{b} \frac{\phi^{\prime}(x) v(y) p(x)}{B L(y)} d y
$$

and

$$
\begin{aligned}
& \frac{d}{d x}\left(p \bar{u}^{\prime}\right)=-\frac{\phi(x)}{} v^{\prime}(x) \\
& W(x) f(x)-\int_{a}^{x} \frac{\phi(y)\left(p v^{\prime}\right)^{\prime}(x)}{\operatorname{BL}(y)} f(y) \\
& \quad+\frac{\phi^{\prime}(x) v(x)}{W(x)} f(x)-\int_{x}^{b} \frac{\left(p \phi^{\prime}\right)^{\prime} v(y) f(y)}{B L(y)} d y .
\end{aligned}
$$

So, finally, using the fact that $A v=0$ and assembling the pieces we have $A \bar{u}=\mathrm{f}$ and so $A \bar{u}=\mathrm{f}$ as required.

We haven't yet used the condition $(\phi, f)=0$. But we still need to check the boundary conditions. We saw in the first half of the proof that if $A \phi=0$ and $A u=f$ (which we have just checked for the definition of $u$ above) then $(\phi, f)=\left.p\left(u \phi^{\prime}-\phi u^{\prime}\right)\right|_{\mathfrak{a}} ^{b}$. As we are given that $(\phi, f)=0$ it follows that $\left.p\left(u \phi^{\prime}-\phi u^{\prime}\right)\right|_{\mathfrak{a}} ^{b}=0$. Substituting for $u$ gives us

$$
\begin{equation*}
\int_{a}^{b} \frac{\phi(y) f(y)}{p(y) W(y)} d y=0 \tag{4.10}
\end{equation*}
$$

It is easy to check that this is exactly the condition needed for
$u$ to satisfy the boundary conditions $B_{1} u(a)=0$ and $B_{2} u(b)=0$ given that $\phi$ does.

Example 4.7 Consider the Sturm-Liouville problem
$\mathfrak{u}^{\prime \prime}+4 \mathfrak{u}=\mathrm{f}(\mathrm{x}), \mathfrak{u}(0)=\mathfrak{u}(\pi)=0$. If we try to solve the problem using the usual Green's Function method we find (check) $\mathrm{g}(\mathrm{x}, \mathrm{y})=A \sin 2 \mathrm{x}$ for $\mathrm{x}<\mathrm{y}$ and $\mathrm{g}(\mathrm{x}, \mathrm{y})=\mathrm{B} \sin 2 \mathrm{x}$ for $\mathrm{x}>\mathrm{y}$.
Continuity at $\mathrm{x}=\mathrm{y}$ requires that $\mathrm{A}=\mathrm{B}$ but when we impose the "jump" at $\mathrm{x}=\mathrm{y}$ we find that $2(\mathrm{~B}-\mathrm{A}) \cos 2 \mathrm{y}=1$ which is a contradiction. So no Green's Function exists for the problem.

You should check that the Green's Function formula (4.3) also breaks down.

What "goes wrong"? Can you find a solution for any choices of $\mathrm{f}(\mathrm{x})$ ?

Solution: Check that this Sturm-Liouville problem has the property that $\mathrm{Lu}=0$ has a non-zero solution, namely $\sin 2 x . S o \mathrm{~L}$ is non-invertible and the problem does not have a Green's Function.

Thm 4.2 does state however that an (infinite) set of solutions does exist if $(\mathrm{f}, \phi)=0$, where $\phi$ is the non-trivial solution to the homogeneous problem $\mathrm{Lu}=0$. For the current problem we have $\phi(x)=\sin 2 x$.

But check $\int_{0}^{\pi} \sin n x \sin 2 x d x=0$ for $n \neq 2$ ( $n$ an integer). Also $\int_{0}^{\pi} \cos \mathrm{nx} \sin 2 \mathrm{xdx}=0$ when $\mathfrak{n}$ is even. So we can find a solution for $\mathrm{f}(\mathrm{x})=\sin \mathfrak{n x}, \mathrm{n} \neq 2$ or $\mathrm{f}(\mathrm{x})=\cos \mathfrak{n x}, \mathrm{n}$ even.

Let's see how it all works out. We can take $\mathcal{v}(\mathrm{x})=\cos 2 \mathrm{x}$ as it satisfies $v^{\prime \prime}+4 v=0$ but not the boundary conditions. Following the recipe above we have $\mathfrak{p}(\mathrm{y})=-1$ and
$W(y)=\phi v^{\prime}-v \phi^{\prime}=\sin 2 y(-2) \sin 2 y-\cos 2 y(2) \cos 2 y=-2$.
So the pseudo-Green's Function $\mathrm{G}(\mathrm{x}, \mathrm{y})(4.9)$ is given by:

$$
G(x, y)=-\frac{1}{2} \begin{cases}\sin 2 x \cos 2 y & x<y \\ \sin 2 y \cos 2 x & x>y\end{cases}
$$

and
$u(x)=c \sin 2 x-\frac{1}{2}\left(\int_{0}^{x} \cos 2 x \sin 2 y f(y) d y+\int_{x}^{\pi} \cos 2 y \sin 2 x f(y) d y\right)$
for any real c provided that $(\mathrm{f}, \sin 2 \mathrm{x})=0$ which is equivalent to $\mathrm{f}(\mathrm{x})=\sin \mathrm{n} \mathrm{x}, \mathrm{n} \neq 2$ or $\mathrm{f}(\mathrm{x})=\cos \mathrm{nx}, \mathrm{n}$ even.

- You should check by direct differentiation that this expression for $\mathfrak{u}(\mathrm{x})$ satisfies $u^{\prime \prime}+4 \mathfrak{u}=\mathrm{f}$ and that $\mathfrak{u}(0)=u(\pi)=0$.
- You will find that the boundary condition at $\pi$ is only satisfied if $(f, \sin 2 x)=0$.
- If (for example) $\mathrm{f}(\mathrm{x})=\cos \mathfrak{n x}$ for $\mathfrak{n}$ odd then the problem has no solution.


### 4.7 Green's functions via Eigenfunctions

Suppose the Sturm-Liouville problem Lu $=\mathrm{f}$ as above has a Green's function. Consider the eigenvalue problem $L u=\lambda u$ with the same boundary conditions. It can be shown that (as L is self-adjoint see below) A has infinitely many (real) eigenvalues and that the eigenfunctions corresponding to distinct eigenvalues $\lambda_{n}$ are orthogonal. Also the eigenfunctions $\phi_{\mathrm{n}}$ form an orthogonal basis for $L^{2}[a, b]$. So the solution to $L u=f$ can be written (normalising the $\phi_{n}$ so that $\left\|\phi_{n}\right\|=1$ ):

$$
\begin{aligned}
u & =\sum u_{n} \phi_{n}(x) & \text { and } & u_{n}=\left(u, \phi_{n}\right) \\
f & =\sum f_{n} \phi_{n}(x) & & f_{n}=\left(f, \phi_{n}\right) .
\end{aligned}
$$

Then using the orthogonality of the $\phi_{n}$ we have $u_{n}=\frac{f_{n}}{\lambda n}$ for all $n$.

So

$$
\begin{aligned}
u(x) & =\sum_{1}^{\infty} \frac{1}{\lambda_{n}}\left(f, \phi_{n}\right) \phi_{n}(x) \\
& =\int_{a}^{b}\left(\sum \frac{1}{\lambda_{n}} \phi_{n}(x) \phi_{n}(y)\right) f(y) d y
\end{aligned}
$$

which allows us to express the Green's function $g(x, y)$ as

$$
g(x, y)=\sum \frac{1}{\lambda_{n}} \phi_{\mathfrak{n}}(x) \phi_{\mathfrak{n}}(y)
$$

Example 4.8 Solve the Sturm-Liouville problemu" ${ }^{\prime \prime}+4 \boldsymbol{u}=\mathrm{f}(\mathrm{x})$ on the interval $[0,1]$ with $u(0)=u(1)=0$ using the eigenfunction method.

Solution: We need to solve the eigenvalue problem
$u^{\prime \prime}+4 u=\lambda u$ with the above boundary conditions. We have
$\mathrm{u}^{\prime \prime}=(\lambda-4) \mathrm{u}$ so if $\lambda-4=\mathrm{k}^{2}>0$ we have $\mathrm{u}=A e^{\mathrm{kx}}+B \mathrm{e}^{-\mathrm{kx}}$ but
the boundary conditions give us $\mathrm{A}=\mathrm{B}=0$. So we take
$\lambda-4=-\mathrm{k}^{2}<0$.
This is straightforward: the solution is $u=A \sin k x$ with $k=n \pi$.
So $\lambda_{n}=4-n^{2} \pi^{2}$ and $\phi_{\mathrm{n}}=A_{\mathrm{n}} \sin \mathrm{n} \pi x$. We can normalise the eigenfunctions $\phi_{\mathrm{n}}$ by requiring them to have unit norm:
$\int_{0}^{\pi} A^{2} \sin ^{2} n \pi x d x=1$ which gives check $\phi_{n}(x)=\sqrt{2 / \pi} \sin n \pi x$.

Finally,

$$
\begin{aligned}
g(x, y) & =\sum_{n=1}^{\infty} \frac{1}{\lambda_{n}} \phi_{n}(x) \phi_{n}(y) \\
& =\frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1}{4-n^{2} \pi^{2}} \sin n \pi x \sin n \pi y
\end{aligned}
$$

and

$$
u(x)=\int_{0}^{1} g(x, y) f(y) d y
$$

Exercise 4.2 Can you finish the calculation when $\mathrm{f}(\mathrm{x})=\mathrm{x}$ ?
Compare your result with the result found using the standard Green's Function method.

A nice symmetry exists between the study of Green's Functions and of integral equations. Suppose the differential operator L (including boundary conditions ) has a symmetric Green's Function $g(x, t)$ Then we have just seen that the eigenvalue problem $L u=\lambda u$ is equivalent to

$$
u=\lambda L^{-1} u=\lambda \int_{a}^{b} g(x, y) u(y) d y
$$

where $g(x, y)$ is given by the eigenfunction expansion above. (This is one reason for writing an inhomogeneous Volterra integral equation of the second kind as $u=f+\lambda K u$ rather than the - at first sight - more natural $K u=\lambda u+f)$.

### 4.8 Exercises

1. Check that the following o.d.e.s can be expressed in terms of the Sturm-Liouville operator $A$ and identify $p(x)$ and $q(x)$. (Try to come up with a general method for identifying $p$ and $q$ - remember that you will in general need an integrating factor $m(x)$ so that $a u^{\prime \prime}+b u^{\prime}+c u=0$ will be replaced by $m a u^{\prime \prime}+\mathrm{mbu}^{\prime}+\mathrm{mcu}=0$ where $\mathfrak{m}(x)$ is to be determined.)

Hermite Equation $u^{\prime \prime}-2 x u^{\prime}+\lambda u=0$
Laguerre Equation $x u^{\prime \prime}+(1-x) u^{\prime}+\lambda u=0$
Chebyshev Equation $\left(1-x^{2}\right) u^{\prime \prime}-x u^{\prime}+n^{2} u=0$
2. The integrating factor method is only useful when the relevant integral can be calculated explicitly. Can you transform $x u^{\prime \prime}+\sin x u^{\prime}+x^{2} u=0$ into the form $-\left(p u^{\prime}\right)^{\prime}+q u=0$ ?
3. Check the derivation of Eq. 4.10.
4. Show that Eq. 4.10 is as claimed the condition required for $u$ to satisfy the boundary conditions $\mathrm{B}_{1} \mathfrak{u}(\mathrm{a})=0$ and $\mathrm{B}_{2} \mathfrak{u}(b)=0$ given that $\phi$ does in Thm. 4.2.
5. Re-solve Ex. 4.1 using the formula (4.3) for the Green's Function.

## 5 Distributions

We need to clarify the definition of the Green's Function $g(x, y)-$ is it a function or not? First some definitions

Definition 5.1 (Test Functions) Define the set $\mathrm{C}_{0}^{\infty}(\mathrm{a}, \mathrm{b})$ to be set of all continuous functions all of whose derivatives exist on ( $\mathbf{a}, \mathrm{b}$ ) and (crucially) which is non-zero only in a closed subset of $(a, b)$.

Informally - a test function is a very smooth function that vanishes outside a prescribed set.

We say that $f$ on $\mathbb{R}$ has "compact suport" if it is non-zero only in a closed and bounded subset of $\mathbb{R}$ - in practice one or more closed intervals $\left[a_{i}, b_{i}\right] \subseteq(a, b)$.

## Example 5.1

$$
\phi(x)= \begin{cases}e^{-x^{2}} & |x|<1 \\ 0 & |x| \geq 1\end{cases}
$$

Check that this is a $\mathbb{C}^{\infty}$ function on $\mathbb{R}$ that vanishes outside $(-1,1)-$ and so is a test function.

## Example 5.2

$$
\phi(x)= \begin{cases}e^{-\frac{a^{2}}{a^{2}-x^{2}}} & |x|<a \\ 0 & |x| \geq a\end{cases}
$$

Check that this is a $\mathrm{C}^{\infty}$ function on $(-\mathrm{a}, \mathrm{a})$ that vanishes outside $(-\mathbf{a}, \mathrm{a})-a n d$ so is a test function. (Hint; no need to differentiate - just appeal to the Chain Rule.)

A technical point - it follows from the Definition that a test function must vanish at a and b .

Definition 5.2 (Local Integrability) A function f is locally integrable on $(\mathrm{a}, \mathrm{b})$ if $\int_{\mathrm{c}}^{\mathrm{d}}|\mathrm{f}(\mathrm{x})| \mathrm{dx}$ is defined for all intervals $[\mathrm{c}, \mathrm{d}]$ in $(\mathrm{a}, \mathrm{b})$. Note that locally integrable functions need not be continuous.

We can use integration by parts to generalise the idea of the derivative to functions that are not differentiable. The is the key to the rest of this Chapter.

Definition 5.3 (Weak Derivative) Let $u$ be in $\mathrm{C}^{1}(\mathrm{a}, \mathrm{b})$ and let $\mathrm{f}=\mathrm{u}^{\prime}$. Let $\phi \in \mathrm{C}_{0}^{\infty}(\mathrm{a}, \mathrm{b})$, a test function. Then $\int_{\mathrm{a}}^{\mathrm{b}} \mathrm{u}^{\prime} \phi \mathrm{d} \mathrm{x}=-\int_{\mathrm{a}}^{\mathrm{b}} u \phi^{\prime} \mathrm{dx}$ as $\phi$ vanishes at a and b .
The important point is that the right-hand integral is defined even if $u$ is not differentiable and so can be used to define the weak derivative of an integrable function. If f and u are both locally integrable on $(\mathrm{a}, \mathrm{b})$, say that f is the weak derivative of $u$ if

$$
\int_{\mathrm{a}}^{\mathrm{b}} \mathrm{f} \phi=-\int_{\mathrm{a}}^{\mathrm{b}} \mathrm{u} \phi^{\prime} \quad \text { for all } \phi \in \mathrm{C}_{0}^{\infty}(\mathrm{a}, \mathrm{~b})
$$

Obviously if $u \in C^{1}(a, b)$ it has a weak derivative $f$ and $f \equiv u^{\prime}$.
But a function can still have a weak derivative even if it does not have an ordinary derivative.

Example 5.3 On the interval $(-1,1)$ let $\mathfrak{u}(x)=|x|$ and $\mathrm{f}(\mathrm{x})=\mathrm{H}(\mathrm{x})-\mathrm{H}(-\mathrm{x})$. The function u is not differentiable at $\mathrm{x}=0$. However $u^{\prime}=\mathrm{f}$ in the weak sense on $(-1,1)$ as for any test function $\phi \in \mathrm{C}_{0}^{\infty}(-1,1)$

$$
-\int_{-1}^{1}|x| \phi^{\prime}(x) d x=\int_{-1}^{1}(H(x)-H(-x)) \phi(x) d x
$$

so $\frac{\mathrm{d}}{\mathrm{d} x}|\mathrm{x}|=\mathrm{H}(\mathrm{x})-\mathrm{H}(-\mathrm{x})$ "in a weak sense".

Now consider the following important example.
Example 5.4 On the interval $(-1,1)$ let $\mathfrak{u}(x)=\mathrm{H}(\mathrm{x})$. Can we find a function $\mathrm{f}(\mathrm{x})$ so that $\mathrm{u}^{\prime}=\mathrm{f}$ "in the weak sense"?

Solution: $\quad$ Suppose that $u^{\prime}=\mathrm{f}$ i.w.s. then by definition

$$
-\int_{-1}^{1} \mathrm{H} \phi^{\prime}=\int_{-1}^{1} \mathrm{f} \phi, \quad \text { for any test function } \phi \in \mathrm{C}_{0}^{\infty}(-1,1)
$$

Now LHS $=-\int_{0}^{1} \phi^{\prime}(\mathrm{x}) \mathrm{d} \mathrm{x}=\phi(0)$ so for f to be the weak derivative of $\mathrm{H}(\mathrm{x})$ it must satisfy

$$
\int_{-1}^{1} f(x) \phi(x) d x=\phi(0), \quad \forall \phi \in C_{0}^{\infty}(-1,1)
$$

But there is no locally integrable function $f(x)$ satisfying this equation. We can show that there is no such function by re-considering our earlier example test function:

$$
\phi(x)= \begin{cases}e^{-\frac{a^{2}}{a^{2}-x^{2}}} & |x|<a \\ 0 & |x| \geq a\end{cases}
$$

Then if $\int_{-1}^{1} f(x) \phi(x) d x=\phi(0)$ we have

$$
e^{-1} \equiv|\phi(0)|=\left|\int_{-a}^{a} f(x) e^{-\frac{a^{2}}{a^{2}-x^{2}}} d x\right| \leq e^{-1} \int_{-a}^{a}|f(x)| d x
$$

But as if we let $a \rightarrow 0$ the RHS $\rightarrow 0$ even though the LHS has a constant value of $e^{-1}$ - which is a contradiction. So the weak derivative of a Heaviside function cannot be a locally integrable function.

Looking at the graph of $H(x)$, we expect that the derivative should be zero except at $x=0$ - where $H(x)$ has a jump - and so the derivative should be "concentrated" at $x=0$.

We will see that the weak derivative of $\mathrm{H}(\mathrm{x})$ is the "delta function" $\delta(x)$ in a general sense to be defined below.

Informally — $\mathrm{H}(\mathrm{x})$ jumps from 0 to 1 at $\mathrm{x}=0$ so its "slope" at $x=0$ is "plus infinity". It is reasonable to expect that the weak derivative of $H(x)$ should be "plus infinity" at $x=0$. As we will see the delta function can be visualised as exactly that - zero everywhere except for a "plus infinity" value at $x=0$.

This is all very vague - we will clarify these ideas in the next Section.

### 5.1 Definitions

Even the more flexible weak derivative is not flexible enough to cope with mathematical ideas as strange as the "delta function".
We need the more general concept of a distribution - a mapping from the set of test functions $C_{0}^{\infty}(a, b) \rightarrow \mathbf{R}$. So a distribution assigns a number to every test function. More precisely; a distribution is a continuous linear functional on the set of test functions. We will use D for the set of distributions and T for the set of test functions $T=C_{0}^{\infty}(a, b)$, so

$$
f \in D \Rightarrow f: T \rightarrow \mathbf{R} .
$$

Instead of the usual function notation $f: \phi \rightarrow f(\phi)$, we write $f: \phi \rightarrow(f, \phi)$. This is because a distribution is closely related to an inner product - see below.

A Point on Notation: The standard notation in most texts is to use D for the set of test functions and $\mathrm{D}^{\prime}$ for the set of distributions

- the notation adopted in these notes is more convenient.

We make these terms precise with some definitions:
Definition 5.4 (Linearity) A mapping from the set of test functions $\mathrm{T}=\mathrm{C}_{0}^{\infty}(\mathrm{a}, \mathrm{b})$ to $\mathbb{R}$ is linear if

$$
\begin{aligned}
(\mathrm{f}, \alpha \phi) & =\alpha(\mathrm{f}, \phi), \quad \forall \phi \in \mathrm{T} \\
\left(\mathrm{f}, \phi_{1}+\phi_{2}\right) & =\left(\mathrm{f}, \phi_{1}\right)+\left(\mathrm{f}, \phi_{2}\right), \quad \forall \phi \in \mathrm{T}
\end{aligned}
$$

Definition 5.5 (Convergence in T ) We say that a set $\left\{\phi_{\mathrm{n}}\right\} \subset \mathrm{T}=\mathrm{C}_{0}^{\infty}(\mathrm{a}, \mathrm{b})$ converges to zero in $\mathrm{T}\left(\phi_{\mathrm{n}} \rightarrow 0\right)$ if there exists a single closed interval $\mathrm{I} \subseteq[\mathrm{a}, \mathrm{b}]$ containing the non-zero domains (supports) of all the $\phi_{\mathrm{n}}$ and if on that interval I the sequence of test functions $\phi_{\mathrm{n}}$ (and the corresponding sequences of all higher derivatives) converge uniformly to zero as $\mathfrak{n} \rightarrow \infty$.

Definition 5.6 (Continuity) Then we say that f is a continuous mapping ("functional") on T if $\phi_{\mathrm{n}} \rightarrow 0$ implies that $\left(\mathrm{f}, \phi_{\mathrm{n}}\right) \rightarrow 0$.

Definition 5.7 (Distribution) A distribution is a continuous linear functional on the set of test functions $T=C_{0}^{\infty}(a, b)$.

Example 5.5 For every locally integrable function $u$ on $[\mathrm{a}, \mathrm{b}]$ there is a "natural distribution" $u$ defined by:

$$
(u, \phi)=\int_{a}^{b} u(x) \phi(x) d x, \quad \text { for any } \phi \in T=C_{0}^{\infty}(a, b)
$$

The linearity and continuity properties hold as a result of the properties of the integral. So every locally integrable function is a distribution.

Example 5.6 For any $y \in(a, b)$, the distribution $\delta_{y}$ defined by $\left(\delta_{y}, \phi\right)=\phi(y)$ is called the Dirac delta distribution with pole at y. (It is easy to see that $\delta_{y}$ is linear and continuous.)

The delta distribution $\delta_{y}$ maps a test function $\phi$ into its value at y , $\phi(y)$.
Note that we should not write

$$
\left(\delta_{y}, \phi\right)=\int_{a}^{b} \delta_{y}(x) \phi(x) d x=\phi(y)
$$

- because: there is no locally integrable function $\delta_{y}$ which satisfies the equation. A distribution which has this property is called a singular distribution.
Despite this, we do often write ( $\mathrm{f}, \phi)=\int_{\mathrm{a}}^{\mathrm{b}} \mathrm{f}(\mathrm{x}) \phi(\mathrm{x}) \mathrm{d} \mathrm{x}$ even when f is a singular distribution. Also the Dirac delta distribution $\delta_{y}$ is often (as in the previous section) written $\delta_{y}(x)$ or $\delta(x, y)$. This causes no problems provided we are careful!

Now return to the earlier example where $u(x)=H(x)$. We found for $f=u^{\prime}$ to exist in a weak sense, we needed

$$
-\int_{-1}^{1} H(x) \phi(x) d x=\int_{-1}^{1} f(x) \phi(x)=\phi(0) \quad \forall \phi \in T(-1,1)
$$

But we showed that there is no locally integrable function $f$ that satisfies this equation. If $f$ is taken to be the delta distribution $\delta_{0}$ (pole at $x=0$ ) then the equation is satisfied if we interpret the integral symbolically as above.

So in a distributional sense we have $\mathrm{H}^{\prime}=\delta_{0}\left(\right.$ or $\left.\mathrm{H}^{\prime}(x)=\delta_{0}(x)\right)$.

### 5.2 Formal Results

We list some properties of distributions.
Let $T=C_{0}^{\infty}(a, b)$ and $D=$ distributions on $T$. We say two distributions are equal if $\left(f_{1}, \phi\right)=\left(f_{2}, \phi\right), \quad \forall \phi \in T$ and write $f_{1}=f_{2}$. We can do algebra in the set $D$ just as in an ordinary function space. For example, for any $\alpha \in C^{\infty}$, if $f$ is a locally integrable function,

$$
(\alpha f, \phi) \equiv \int_{a}^{b} \alpha f \phi=(f, \alpha \phi)
$$

So for distributions, we define $(\alpha f, \phi)=(f, \alpha \phi)$ - we are defining the distribution $\alpha f$ by specifying how it acts on an arbitrary test function.

We already have that if $f$ and $f^{\prime}$ are locally integrable then

$$
\left(f^{\prime}, \phi\right)=\int_{a}^{b} f^{\prime} \phi=-\int_{a}^{b} f \phi^{\prime}=-\left(f, \phi^{\prime}\right)
$$

as $\phi$ vanishes outside $(a, b)$.
Definition 5.8 (Derivative of a distribution) If $\mathrm{f} \in \mathrm{D}$, define $\left(\mathrm{f}^{\prime}, \phi\right)=-\left(\mathrm{f}, \phi^{\prime}\right)$ all $\phi \in \mathrm{T}$. Then $\mathrm{f}^{\prime}$ is called the distributional derivative of f . In general, integrating by parts n times, we have

$$
\left(f^{(n)}, \phi\right)=(-)^{n}\left(f, \phi^{(n)}\right) \text { for all } \phi \in T
$$

So despite its exotic definition and behaviour, a distribution has distributional derivatives of all orders!

Example 5.7 Differentiate $\delta_{y}$. By definition
$\left(\delta_{y}^{\prime}, \phi\right)=-\left(\delta_{y}, \phi^{\prime}\right)=-\phi^{\prime}(y)$.

For any constant $c \in \mathbb{R}$, we can define for any $f(x) \in D$ (note sloppy notation) the "translated" distribution $f(x-c) \in D$ by

$$
(f(x-c), \phi)=(f(x), \phi(x+c)) \quad \forall \phi \in T,
$$

motivated by the corresponding result for locally integrable functions $f$.

So, for example, if $f(x)=\delta_{0}(x)$ (or just $\delta(x)$ ) then $\delta(x-c)$ is defined by

$$
(\delta(x-c), \phi)=(\delta, \phi(x+c))=\phi(c)
$$

But we already have $\left(\delta_{c}(x), \phi\right)=\phi(c)$ so $\delta_{\mathcal{c}}(x)=\delta(x-c)$ We often write $\int_{a}^{b} \delta(x-c) \phi(x) d x=\phi(c)$.
Think of distributions as returning a value when a test function is averaged over a region - rather than returning a value at a point as functions do.

### 5.3 Distributional Solutions to ODE's

Consider the 2nd order linear homogeneous differential operator L defined by:

$$
\mathrm{Lu}=\alpha u^{\prime \prime}+\beta u^{\prime}+\gamma u, \alpha, \beta, \gamma \in C^{\infty}(a, b) .
$$

Definition 5.9 By a classical solutionto $L \mathfrak{f}=\mathrm{f}$ we mean $a$ function $\mathfrak{u}(\mathrm{x}) \in \mathrm{C}^{2}(\mathrm{a}, \mathrm{b})$ that satisfies $\mathrm{Lu}=\mathrm{f}$ identically $\forall x \in(a, b)$.

We can also interpret $L \mathfrak{u}=f$ in a distributional sense; if $\mathfrak{u}$ and $f$ are distributions then (as we know how to differentiate a distribution) so is Lu.

Definition 5.10 If $\mathrm{Lu}=\mathrm{f}$ (equality as distributions) then we say
$\mathfrak{u}$ is a distributional solution to the equation and mean that
$(\mathrm{Lu}, \phi)=(\mathrm{f}, \phi) \quad \forall \phi \in \mathrm{T}$.
Obviously if $\mathfrak{u}$ is a classical solution then it is a distributional solution but the reverse is not true.

Definition 5.11 We define a "fundamental solution" associated with a differential operator L to be a distributional solution to

$$
\mathrm{Lu}=\delta(x-y) \quad-\text { the Dirac delta distribution. }
$$

Note that a Green's Function is a fundamental solution with a particular choice of boundary conditions.

Now we know that if $u$ is a distribution then so is $L u$ and $(\mathrm{Lu}, \phi)=\left(u,(\alpha \phi)^{\prime \prime}\right)-\left(u,(\beta \phi)^{\prime}\right)+(u, \gamma \phi)-$ using integration by parts as usual. We can write RHS as $\left(u, L^{*} \phi\right)$ where

$$
\mathrm{L}^{*} \phi=(\alpha \phi)^{\prime \prime}-(\beta \phi)^{\prime}+\gamma \phi \quad \forall \phi \in \mathrm{T}
$$

$L^{*}$ is called the formal adjoint operator.
So if $u$ is a distributional solution to $L u=f$, then $\left(u, L^{*} \phi\right)=(f, \phi) \quad \forall \phi \in T$.

If $u$ and $f$ are locally integrable then we have
$\int_{a}^{b} u(x) L^{*} \phi(x) d x=\int_{a}^{b} f(x) \phi(x) d x, \quad \forall \phi \in T$ and we say $u$ is a weak solution.

Note that $u$ can be a weak solution without having any conventional derivatives. We note in passing that an operator $L$ is "formally self-adjoint" if $L^{*}=\mathrm{L}$.

Example 5.8 The Sturm-Liouville operator A:

$$
A u=-\left(p u^{\prime}\right)^{\prime}+q u
$$

is formally self-adjoint - check.
We have extended the idea of a solution from the classical solution - valid at every point in the domain ( $a, b)$ - to weak or distributional solutions.

This allows us to accept solutions that are physically meaningful but may not have derivatives everywhere - for example we can now interpret the o.d.e. for the Green's Function $g(x, y)$; saying that $g(x, y)$ is the distributional solution to $\operatorname{Ag}(x, y)=\delta(x, y)$.

It is to be interpreted as

$$
(A g(x, y), \phi)=(\delta(x-y), \phi) \quad \forall \phi \in \mathrm{T}
$$

or equivalently

$$
\left(g(x, y), A^{*} \phi\right)=\phi(y) \quad \forall \phi \in \mathrm{T}
$$

For Sturm-Liouville problems $A^{*}$ may be replaced by $A$.

Example 5.9 Consider
$\mathrm{g}(\mathrm{x}, \mathrm{y})=\mathrm{y}(1-\mathrm{x}) \mathrm{H}(\mathrm{x}-\mathrm{y})+\mathrm{x}(1-\mathrm{y}) \mathrm{H}(\mathrm{y}-\mathrm{x})$. We will show that $\mathrm{g}(\mathrm{x}, \mathrm{y})$ is the Green's Function associated with $\mathrm{L}=-\frac{\mathrm{d}^{2}}{\mathrm{dx}^{2}}$ on $(0,1)$ subject to $u(0)=u(1)=0$.

So RTP

$$
\left\{\begin{array}{l}
-g^{\prime \prime}(x, y)=\delta(x, y) \\
g(0, y)=g(1, y)=0
\end{array}\right.
$$

in a distributional sense which means that

$$
R T P\left(\mathrm{~g}(\mathrm{x}, \mathrm{y}),-\phi^{\prime \prime}\right)=\phi(\mathrm{y}) \quad \forall \phi \in \mathrm{T}
$$

$$
\begin{aligned}
\mathrm{LHS}= & -\int_{0}^{1} g(x, y) \phi^{\prime \prime}(x) d x \\
= & -(1-y) \int_{0}^{y} x \phi^{\prime \prime}(x) d x-y \int_{y}^{1}(1-x) \phi^{\prime \prime}(x) d x \\
= & (y-1)\left[\left.x \phi^{\prime}\right|_{0} ^{y}-\int_{0}^{y} \phi^{\prime}(x) d x\right] \\
& -y\left[\left.(1-x) \phi^{\prime}\right|_{y} ^{1}+\int_{y}^{1} \phi^{\prime}\right] \\
= & (y-1)\left[y \phi^{\prime}(y)-\phi(y)\right] \\
& \quad-y\left[-(1-y) \phi^{\prime}(y)-\phi(y)\right] \\
= & \phi^{\prime}(y) \times 0+\phi(y)\{-(y-1)+y\}=\phi(y)
\end{aligned}
$$

as required.

Example 5.10 Does $(u, \phi)=\phi(0)^{2}$ define a distribution?
Example 5.11 Show $\mathfrak{u}(\mathrm{x}, \mathrm{y})=\frac{1}{2}|\mathrm{x}-\mathrm{y}|$ is a fundamental solution for $\mathrm{L}: \frac{\mathrm{d}^{2}}{\mathrm{dx}^{2}}$ on $\mathbb{R}$.
$R T P \quad(\mathrm{Lu}, \phi)=\phi(\mathrm{y})$. Now, LHS is $\left(\mathrm{u}, \phi^{\prime \prime}\right)$ so $R T P$ $\int u(x, y) \phi^{\prime \prime}(x) d x=\phi(y)$ for all $\phi \in T(\mathbb{R})$.
Now LHS is

$$
\begin{aligned}
\text { LHS }= & \frac{1}{2}\left\{\int_{-\infty}^{y}(y-x) \phi^{\prime \prime}(x) d x+\int_{y}^{\infty}(x-y) \phi^{\prime \prime}(x) d x\right\} \\
= & \frac{1}{2}\left\{\left.(y-x) \phi^{\prime}(x)\right|_{-\infty} ^{y}-\int_{-\infty}^{y} \phi^{\prime}(x)(-1)\right\} \\
& +\frac{1}{2}\left\{\left.(x-y) \phi^{\prime}(x)\right|_{y} ^{\infty}-\int_{y}^{\infty} \phi^{\prime}(x) 1 d x\right\}
\end{aligned}
$$

which reduces to $\frac{1}{2} 2 \phi(y)$ and is equal to the RHS as required.

Example 5.12 Find a fundamental solution for the operator L defined by:

$$
\mathrm{Lu} \equiv x^{2} u^{\prime \prime}+x u^{\prime}-u
$$

- We want to solve

$$
\mathrm{Lu}=\delta(x-y)
$$

- For $\mathrm{x} \neq \mathrm{y}$ solve $\mathrm{Lu}=0$. So

$$
x^{2} u^{\prime \prime}+x u^{\prime}-u=0
$$

- We could use the general substitution $\mathrm{x}=\mathrm{e}^{\mathrm{t}}$ - which works for so-called Euler-Cauchy equations of general form (a, b, c constants):

$$
x^{2} u^{\prime \prime}+b x u^{\prime}+c u=0
$$

- Check that the above substitution reduces an E-C equation to the form

$$
\ddot{u}+(b-1) \dot{u}+c u=0,
$$

where $\dot{u} \equiv \frac{\mathrm{du}}{\mathrm{dt}}$.

- It is simpler here to use the substitution $\mathfrak{u}=x^{n}$ - this gives $n^{2}-1=0$ so $u=A x+\frac{B}{x}$.
- Check that the more general substitution leads to the same result.
- Take $\mathfrak{u}=\mathrm{Ax}+\frac{\mathrm{B}}{\mathrm{x}}$ on $0<\mathrm{x}<\mathrm{y}$ and $\mathrm{u}=\mathrm{C} x+\frac{\mathrm{D}}{\mathrm{x}}$ on $\mathrm{y}<\mathrm{x}<1$.
- So continuity at $x=y$ implies that $A y+\frac{B}{y}=C y+\frac{D}{y}$.
- Also use $\int_{y_{-}}^{y_{+}} \mathrm{Lu}=\int_{y_{-}}^{y_{+}} \delta(x-y) \mathrm{d} x=1$.
- Use integration by parts: $\int_{y_{-}}^{y_{+}} x^{2} u^{\prime \prime}(x)=\left.y^{2} \Delta u^{\prime}\right|_{\mathrm{y}}$.
- The other terms go to zero when $\mathrm{y}_{-} \rightarrow \mathrm{y} \leftarrow \mathrm{y}_{+}$.
- Now $\Delta \mathrm{u}^{\prime} \mid \mathrm{y}=\mathrm{C}-\mathrm{D} / \mathrm{y}^{2}-\mathrm{A}+\mathrm{B} / \mathrm{y}^{2}$ so we have

$$
C-D / y^{2}-A+B / y^{2}=1
$$

- Solving for A and B we find that $\mathrm{A}=\mathrm{C}-\frac{1}{2 \mathrm{y}^{2}}$ and $\mathrm{B}=\mathrm{D}+\frac{1}{2}$ so finally,

$$
u(x, y)= \begin{cases}\left(C-\frac{1}{2 y^{2}}\right) x+\frac{D+\frac{1}{2}}{x}, & x<y \\ C x+\frac{D}{x} & x>y\end{cases}
$$

- Note that there are two arbitrary parameters as we would expect as boundary conditions have not been specified.

Exercise 5.1 Check that the above solution $\mathfrak{u}(\mathrm{x}, \mathrm{y})$ is indeed a fundamental solution for the operator L .

Example 5.13 Is the function $\phi=x(1-x)$ a test function on $(0,1)$ ? No - as $\phi^{\prime \prime}=$ const $(-2)$ on all of $\mathbb{R}$, so fails compact support requirement.

Example 5.14 Find the second (distributional) derivative of $u=H(x) \cos (x)$.

## Solution:

$$
\begin{array}{rlrl}
\left(u^{\prime \prime}, \phi\right) & = & & \left(u, \phi^{\prime \prime}\right) \\
& = & & \int_{-\infty}^{\infty} \mathrm{H}(x) \cos (x) \phi^{\prime \prime}(x) d x \\
& = & & \int_{0}^{\infty} \cos (x) \phi^{\prime \prime}(x) d x \\
& = & & \left.\phi^{\prime}(x) \cos x\right|_{0} ^{\infty}-\int_{0}^{\infty}(-) \sin x \phi^{\prime}(x) d x \\
& = & -\phi^{\prime}(0)+\left.\sin x \phi(x)\right|_{0} ^{\infty}-\int_{0}^{\infty} \phi(x) \cos x d x .
\end{array}
$$

So

$$
\left(u^{\prime \prime}, \phi\right)=\left(\delta_{0}^{\prime}, \phi\right)-(\mathrm{H} \cos , \phi)
$$

and so $\mathrm{u}^{\prime \prime}=\delta_{0}^{\prime}-\mathrm{H}(\mathrm{x}) \cos (\mathrm{x})$ in a distributional sense.

Example 5.15 Calculate $\left(\frac{\mathrm{d}}{\mathrm{dx}}-\lambda\right)\left(\mathrm{H}(\mathrm{x}) \mathrm{e}^{\lambda x}\right)$ in $\mathrm{D}(\mathbb{R})$. We have

$$
\begin{aligned}
\left(\left(u^{\prime}-\lambda u\right), \phi\right) & =-\left(u, \phi^{\prime}\right)-\lambda(u, \phi) ; \quad \text { where } u=H e^{\lambda x} \\
& =-\int_{0}^{\infty} e^{\lambda x} \phi^{\prime}(x) d x-\lambda \int_{0}^{\infty} e^{\lambda x} \phi(x) d x \\
& =-\left.e^{\lambda x} \phi(x)\right|_{0} ^{\infty}+\lambda \int_{0}^{\infty} e^{\lambda x} \phi(x) d x-\lambda \int_{0}^{\infty} e^{\lambda x} \phi(x) d x \\
& =\phi(0)
\end{aligned}
$$

So $(\mathrm{D}-\lambda) \mathrm{u}=\delta_{0}$. Can you" explain" this result in terms of the graph of u?

### 5.4 Partial Differential Equations

In this section we will see how Green's Function methods can be extended to P.D.E.'s. First we look at Elliptic P.D.E.'s

### 5.4.1 Elliptic Problems

We can carry over the above ideas (Green's Functions and distributions ) in a natural way to multivariate problems and so to p.d.e.'s.

Summary of definitions for $\mathbb{R}^{n}$
Definition 5.12 Given an open set $\Omega \subseteq \mathbb{R}^{n}$, define
$\mathrm{T}(\Omega)=\mathrm{C}_{0}^{\infty}(\Omega)$ the test functions on $\Omega$. If $\phi \in \mathrm{T}(\Omega)$ it must vanish outside a closed bounded subset of $\Omega$ and have continuous derivatives of all orders.

Definition 5.13 A distribution $\mathfrak{u}$ on $\mathrm{T}(\Omega)$ is a continuous linear functional on $\mathrm{T}(\Omega)$ - as before we denote the value (action) of $u$ on $\phi$ by $(u, \phi) \in \mathbb{R}$ for any $\phi \in \mathrm{T}(\Omega)$. The set of all distributions defined on $\mathrm{T}(\Omega)$ is $\mathrm{D}(\Omega)$.

Every function $u: \mathbb{R}^{n} \rightarrow \mathbb{R}$ that is locally integrable on $\Omega$ (i.e. $\int_{K}|f(x)| d x<\infty$ for all closed bounded sets $\left.K \leq \Omega\right)$ generates a dist via

$$
(u, \phi)=\int_{\Omega} u(x) \phi(x) d x, \phi \in \mathrm{~T}(\Omega)
$$

The most important example of a singular distribution (one not defined by the integral of a function on $\Omega$ ) is the delta distribution $\delta_{\xi}$ where $\left(\delta_{\xi}, \phi\right)=\phi(\xi)$ for any $\phi \in T(\Omega)$.

We often blur the important distinction between a locally integrable function $u(x)$ and its associated distribution - using the same notation for both. Even when a distribution is singular we often still write it as $u(x)$ and denote its action by $\int_{\Omega} u(x) \phi(x) d x-$ even though the integral is not defined.

So, as in $\mathbb{R}$, we write

$$
\left(\delta_{\xi}, \phi\right)=\int_{\Omega} \delta(x-\xi) \phi(x) d x=\phi(\xi)
$$

Define $(a u, \phi)=(u, a \phi)$ where $a \in C^{\infty}(\Omega)$ and the partial derivative $\partial_{k}$ of a distribution $u$ is defined by

$$
\left(\partial_{k} u, \phi\right)=-\left(u, \partial_{k} \phi\right) \text { as in } \mathbb{R} \quad \text { for any } \phi \in T(\Omega)
$$

Similarly, second partials are defined by

$$
\left(\partial_{j k} u, \phi\right)=\left(u, \partial_{j k} \phi\right) \text { as expected. }
$$

Definition 5.14 If f is a distribution, the equation $\mathrm{Lu}=\mathrm{f}$ can have a distributional solution $u$ if $(\mathrm{Lu}, \phi)=(\mathrm{f}, \phi), \quad \forall \phi \in \mathrm{T}(\Omega)$. In the special case where $\mathrm{f} \equiv \delta(\mathrm{x}-\xi)$ (delta distribution) than u is called a fundamental solution — just as in $\mathbb{R}$ - associated with the operator L .

Definition 5.15 The formal adjoint operator L* is defined by

$$
(\mathrm{Lu}, \phi)=\left(u, \mathrm{~L}^{*} \phi\right) \quad \forall \phi \in \mathrm{T}(\Omega)
$$

Definition 5.16 We say $u$ is a distributional solution of $\mathrm{Lu}=\mathrm{f}$ if $\left(\mathrm{u}, \mathrm{L}^{*} \phi\right)=(\mathrm{f}, \phi), \quad \forall \phi \in \mathrm{T}$ and $\mathfrak{u}$ is a fundamental solution (with pole at $\xi$ ) assoc. with L if

$$
\left(u, L^{*} \phi\right)=\phi(\xi), \quad \forall \phi \in T(\Omega)
$$

Definition 5.17 If $u$ and $f$ are locally integrable functions on $\Omega$ then $\mathfrak{u}$ is $a$ weak solution to $\mathrm{Lu}=\mathrm{f}$ if

$$
\int_{\Omega} u(x) L^{*} \phi(x) d x=\int_{\Omega} f(x) \phi(x) d x, \quad \forall \phi \in T(\Omega)
$$

Definition 5.18 Finally, a solution to $\mathrm{Lu}=\mathrm{f}$ that is in $\mathrm{C}^{2}(\Omega)$ is called a classical solution.

So - as on $\mathbb{R}$ - we have three levels of solution:

- classical
- weak
- distributional.

Example 5.16 Show $\mathrm{g}(\mathrm{x}, \mathrm{y})=\frac{1}{4 \pi} \ln \left(\mathrm{x}^{2}+\mathrm{y}^{2}\right)$ is a distributional solution of

$$
\Delta \mathfrak{u}=\quad \delta(x, y) \quad\left(\Delta u \equiv \nabla^{2} \mathfrak{u}\right)
$$

where the source term is the delta distribution with pole at
$(x, y)=(0,0)$.
Solution: $\quad R T P(\Delta \mathrm{~g}, \phi) \equiv(\mathrm{g}, \Delta \phi)=\phi(0,0), \quad \forall \phi \in \mathrm{T}\left(\mathbb{R}^{2}\right)$. Use polar coordinates:

$$
\begin{aligned}
(4 \pi g, \Delta \phi) & =\int_{\mathbb{R}^{2}} \ln \left(x^{2}+y^{2}\right)\left(\phi_{x x}+\phi_{y y}\right) d x d y \\
& =\int_{0}^{2 \pi} \int_{0}^{\infty} \ln \left(r^{2}\right)\left(\phi_{r r}+\frac{1}{r} \phi_{r}+\frac{1}{r^{2}} \phi_{\theta \theta}\right) r d r d \theta \\
& =\lim _{\varepsilon \rightarrow 0} \int_{\varepsilon}^{\infty} \int_{0}^{2 \pi}\left(r \ln r^{2} \phi_{r r}+\ln r^{2} \phi_{r}+\frac{1}{r} \ln r^{2} \phi_{\theta \theta}\right) d r d \theta
\end{aligned}
$$

The $\theta$-integral of the last term is zero as the test function $\phi$ must be periodic in $\theta$ (as must its derivatives).

Now integrate by parts:

$$
\begin{aligned}
\mathrm{T}_{1} & =\int_{\varepsilon}^{\infty} \mathrm{r} \ln \mathrm{r}^{2} \phi_{\mathrm{rr}} \mathrm{dr} \\
& =\left.\mathrm{r} \ln \mathrm{r}^{2} \phi_{\mathrm{r}}\right|_{\varepsilon} ^{\infty}-\int_{\varepsilon}^{\infty} \phi_{\mathrm{r}}\left(\ln r^{2}+2\right) \mathrm{dr} \\
& =\left.r \ln r^{2} \phi_{r}\right|_{\varepsilon} ^{\infty}-\left.\phi\left(\ln r^{2}+2\right)\right|_{\varepsilon} ^{\infty}+\int_{\varepsilon}^{\infty} \phi(r, \theta)\left(\frac{2}{r}\right) \mathrm{dr} \\
& =-\varepsilon \ln \varepsilon^{2} \phi_{r}(\varepsilon, \theta)+\phi(\varepsilon, \theta)\left(\ln \varepsilon^{2}+2\right)+\int_{\varepsilon}^{\infty} \phi(r, \theta)\left(\frac{2}{r}\right) d r .
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{T}_{2} & =\int_{\varepsilon}^{\infty} \ln \mathrm{r}^{2} \phi_{\mathrm{r}} \mathrm{dr} \\
& =\left.\phi \ln \mathrm{r}^{2}\right|_{\varepsilon} ^{\infty}-\int_{\varepsilon}^{\infty} \frac{2}{\mathrm{r}} \phi \mathrm{dr} \\
& =-\phi(\varepsilon, \theta) \ln \varepsilon^{2}-\int_{\varepsilon}^{\infty} \frac{2}{\mathrm{r}} \phi \mathrm{dr}
\end{aligned}
$$

So cancelling where possible, we have

$$
\mathrm{T}_{1}+\mathrm{T}_{2}=-\varepsilon \ln \varepsilon^{2} \phi_{\mathrm{r}}(\varepsilon)+2 \phi(\varepsilon)
$$

But $\varepsilon \ln \varepsilon^{2} \rightarrow 0$ as $\varepsilon \rightarrow 0$ so $\mathrm{T}_{1}+\mathrm{T}_{2}$ reduces to

$$
\int_{0}^{2 \pi} 2 \lim _{\varepsilon \rightarrow 0} \phi(\varepsilon, \theta) d \theta=4 \pi \phi(0,0)
$$

as required.

The distributional solution $\frac{1}{4 \pi} \ln \left(\mathrm{r}^{2}\right)$ is called the $\log$ potential and is a fundamental solution associated with the 2-D Laplacian.

Note that it is not a classical or weak solution to $\Delta \mathfrak{u}=0$.
As $\Delta \equiv \nabla^{2}$ is invariant under translations, a fundamental solution corresponding to the Laplacian with pole at $(\xi, \eta)$ is just $\frac{1}{4 \pi} \ln \left((x-\xi)^{2}+(y-\eta)^{2}\right)$.

Example 5.17 Show that in $\mathbb{R}^{3}$, the corresponding fundamental solution is $\mathrm{g}(\mathrm{x}, \mathrm{y}, \mathrm{z})=-\frac{1}{4 \pi \mathrm{r}}$.
Solution: Use spherical polar coordinates:

- The volume element is:

$$
d x d y d z=r^{2} d r \sin \phi d \theta d \phi
$$

- The Laplacian is:

$$
\Delta \mu=\frac{1}{r^{2}}\left(\frac{\partial\left(r^{2} \mu_{r}\right)}{\partial r}+\frac{1}{\sin \phi} \frac{\partial\left(\sin \phi \mu_{\phi}\right)}{\partial \phi}+\frac{1}{\sin ^{2} \phi} \mu_{\theta \theta}\right)
$$

$R T P \quad(\Delta \mathrm{~g}, \mu) \equiv(\mathrm{g}, \Delta \mu)=\mu(0)$, for any test function $\mu \in \mathrm{T}\left(\mathbb{R}^{3}\right)$. (We use $\mu$ to avoid confusion with the angle ф.)

So $\operatorname{RTP}\left(\frac{1}{\mathrm{r}}, \Delta \mu\right)=-4 \pi \mu(0)$.
Now,

$$
\begin{aligned}
\left(\frac{1}{r}, \Delta \mu\right)= & \int_{0}^{2 \pi} d \theta \int_{0}^{\pi} d \phi r^{2} d r\left(\frac{1}{r}\right) \sin \phi \frac{1}{r^{2}}\left(\frac{\partial\left(r^{2} \mu_{r}\right)}{\partial r}\right. \\
& \left.+\frac{1}{\sin \phi} \frac{\partial\left(\sin \phi \mu_{\phi}\right)}{\partial \phi}+\frac{1}{\sin ^{2} \phi} \mu_{\theta \theta}\right) \\
= & T_{1}+T_{2}+T_{3}
\end{aligned}
$$

As before, change the lower limit on the r -integral to $\varepsilon$ - we will take the limit as $\varepsilon \rightarrow 0$ after simplifying as much as possible.

- Now $\mathrm{T}_{1}=\int \mathrm{d} \theta \mathrm{d} \phi \sin \phi \int_{\varepsilon}^{\infty} \frac{1}{\mathrm{r}} \frac{\partial\left(\mathrm{r}^{2} \mu_{\mathrm{r}}\right)}{\partial \mathrm{r}}$. The angular integral gives a factor of $4 \pi$ once the r -integral is done. Integrating by parts we have:
$\mathrm{T}_{1}=4 \pi\left(\left.\mathrm{r} \mu_{\mathrm{r}}\right|_{\varepsilon} ^{\infty}-\int_{\varepsilon}^{\infty} \mathrm{r}^{2} \mu_{\mathrm{r}}\left(-\frac{1}{\mathrm{r}^{2}}\right) \mathrm{dr}\right)=4 \pi \int_{\varepsilon}^{\infty} \mu_{\mathrm{r}} \mathrm{dr}=-4 \pi \mu(\varepsilon)$.
- $\mathrm{T}_{2}=\int_{\varepsilon}^{\infty} \mathrm{dr} \frac{1}{\mathrm{r}} \int_{0}^{\pi} \frac{\partial\left(\sin \phi \mu_{\phi}\right)}{\partial \phi} \mathrm{d} \phi=\left.\int_{\varepsilon}^{\infty} \mathrm{dr} \sin \phi \mu_{\phi}\right|_{0} ^{\pi}=0$.
- $\mathrm{T}_{3}=\int \mathrm{dr} \int \mathrm{d} \theta \int \mathrm{d} \phi \frac{1}{\sin \phi} \mu_{\theta \theta}$. The $\theta$-integral is zero as $\mu$ and its derivatives are periodic in $\theta$ and $\phi$.

Taking the limit as $\varepsilon \rightarrow 0$ gives us the result.
Again the fundamental solution may be displaced so that $g(x, y, z ; \rho, \sigma, \tau)=-\frac{1}{4 \pi} \sqrt{(x-\rho)^{2}+(y-\sigma)^{2}+(z-\tau)^{2}}$

Note that fundamental solutions associated with $\Delta$ are radial (only dependant on distance from the pole). The Green's
Function associated with a partial differential operator like $\Delta$ is a fundamental solution that also satisfies the homogeneous boundary conditions.

So the Green's Function is the equilibrium (time invariant) response of the physical system to a unit point source. Mathematically, the Green's Function is the kernel of the integral operator that represents the inverse of the partial differential operator.

On finite domains separation of variables may be useful in calculating the Green's Function. On infinite domains transform methods may be useful or we will sometimes use geometrical or physical insights.

Example 5.18 Find the Green's Function associated with $\Delta$ on $-\infty<x<\infty ; y>0$ with Dirichlet boundary conditions on $\mathrm{y}=0$. We seek a distributional solution $\mathrm{G}(\mathrm{x}, \mathrm{y} ; \xi, \eta)$ to

$$
\begin{aligned}
\mathrm{G}_{x x}+\mathrm{G}_{y y} & =\delta_{\xi, \eta}, \quad x, \xi \in \mathbb{R} ; y, \eta>0 \\
\mathrm{G}(x, 0 ; \xi, \eta) & =0, \quad \text { all } x, \xi \in \mathbb{R} ; \eta>0
\end{aligned}
$$

Physically, G represents the potential in the half-plane $\mathrm{y}>0$ of $a$ static electric field generated by a unit positive point charge at $(\xi, \eta)$ with the condition that the potential vanishes on $\mathrm{y}=0$.

Use the method of images/reflection principle.
We already have the fundamental solution g for a 2D Laplacian but it doesn't satisfy the boundary condition on $\mathrm{y}=0$. To compensate we locate an "image charge" of opposite sign at $(\xi,-\eta)$. The potential due to this charge is $\widehat{\mathrm{g}}$ where $\widehat{\mathrm{g}}=-\mathrm{g}(\mathrm{x}, \mathrm{y} ; \xi,-\eta)$. So $\mathrm{G}=\mathrm{g}+\widehat{\mathrm{g}}=\frac{1}{4 \pi} \ln \left(\frac{(x-\xi)^{2}+(\mathrm{y}-\mathfrak{\eta})^{2}}{(\mathrm{x}-\bar{\xi})^{2}+(\mathrm{y}+\mathfrak{\eta})^{2}}\right)$.

Clearly G satisfies the homogeneous boundary conditions on $\mathrm{y}=0$. Also, $\Delta \mathrm{G}=\Delta \mathrm{g}+\Delta \widehat{\mathrm{g}}=\delta(\mathrm{x}, \mathrm{y} ; \xi, \eta)+\delta(\mathrm{x}, \mathrm{y} ; \xi,-\eta)$ but the second term is identically zero in the upper half plane.

So we can now solve the inhomogeneous problem

$$
\begin{aligned}
u_{x x}+u_{y y} & =\rho(x, y) \\
u(x, 0) & =0
\end{aligned}
$$

with

$$
u(x, y)=\int_{-\infty}^{\infty} \int_{0}^{\infty} G(x, \xi ; y, \eta) \rho(\xi, \eta) d \xi d \eta
$$

### 5.5 Transforms of distributions

Suppose we ask for the Fourier Transform of a delta distribution $\delta_{x_{0}} \equiv \delta\left(x-x_{0}\right)$. If we ignore the mathematical pitfalls we have

$$
\mathcal{F}\left(\delta\left(x-x_{0}\right)\right)(s)=\int_{-\infty}^{\infty} \delta\left(x-x_{0}\right) e^{-i x s} d x \equiv\left(\delta\left(x-x_{0}\right), e^{-i s x}\right)=e^{-i s x} .
$$

So $\delta\left(\widehat{x-x_{0}}\right)(s)=e^{-i s x_{0}}$.
This is correct but based on sloppy mathematics. Suppose that $\mathfrak{u}$ is a distribution in $D(\mathbb{R})$ generated by a locally integrable function $u$. Assume that the Fourier transform $\widehat{\mathfrak{u}}$ is also locally integrable.

Then $\forall \phi \in \mathrm{T}(\mathbb{R})$ we have

$$
\begin{aligned}
(\widehat{u}, \phi) & =\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} u(x) e^{-i s x} d x\right) \phi(s) d s \\
& =\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} \phi(s) e^{-i s x} d s\right) u(x) d x \\
& =(u, \widehat{\phi})
\end{aligned}
$$

But this requires that both $\phi$ and $\widehat{\phi}$ be test functions. It can be shown that this implies $\phi \equiv 0$ ! To resolve this difficulty we need is to use a broader class of test functions - the Schwartz class $\mathcal{S}$.

Loosely $\mathcal{S}$ is the set of functions $\phi(x)$ that (along with all their derivatives) have the property that $\phi \in \mathrm{C}^{\infty}$ and

$$
\left|\phi^{(\mathrm{k})}(\mathrm{x})\right|<\frac{\mathrm{const}}{|x|^{N}}
$$

for all k and N as $|\mathrm{x}| \longrightarrow \infty$. We say that $\phi^{(k)} \rightarrow 0$ "faster than any negative power of x " - the set $\mathcal{S}$ are called "rapidly decreasing functions". It can be shown that

$$
\phi \in \mathcal{S} \Rightarrow \widehat{\phi} \in \mathcal{S}
$$

Definition 5.19 $A$ tempered distribution is a continuous linear functional on $\mathcal{S}$.

Now we can define the Fourier Transform of a tempered distribution $u$ by $(\widehat{u}, \phi)=(u, \widehat{\phi})$ where $\phi \in \mathcal{S}$.

So define the Fourier Transform of $\delta\left(x-x_{0}\right)$ by

$$
\left(\delta\left(\widehat{x-x_{0}}\right), \phi\right)=\left(\delta\left(x-x_{0}\right), \widehat{\phi}\right)=\widehat{\phi}\left(x_{0}\right)=\int_{-\infty}^{\infty} \phi(s) e^{-i x_{0} s} d x=\left(e^{-i x_{0} s}, \phi\right)
$$

So conclude as above (but now on a firm basis) that

$$
\delta\left(\widehat{x-x_{0}}\right)=e^{-i x_{0} s}
$$

So if $x_{0}=0$, have $\widehat{\delta(x)}=1$.

### 5.6 Diffusion problems - problems involving time

We will focus on the heat equation in 1-D: $\mathrm{Lu}=\mathfrak{u}_{\mathrm{t}}-\mathrm{k} \mathfrak{u}_{\mathrm{xx}}$. As with elliptic problems, say $u$ is a fundamental solution associated with $L$ if $u$ is a distributional solution of

$$
\mathrm{Lu}=\delta(x, \mathrm{t} ; \xi, \tau)
$$

where we can interpret RHS as a unit heat source applied at $x=\xi$ at time $t=\tau$. The domain $\Omega$ is $\mathbb{R}^{2}$. To find a fundamental solution take a source at $\xi=0$ and $\tau=0$.

Consider the initial value problem:

$$
\begin{aligned}
u_{t}-k u_{x x} & =0, \quad t>0 \\
u(x, 0) & =\delta(x)
\end{aligned}
$$

Now solve using Fourier Transforms - assuming that distributions are tempered (test functions $\phi \in \mathcal{S}$ ), we have

$$
\widehat{u_{t}(s, t)}+k s^{2} \widehat{u(s, t)}=0, \quad \text { for } \quad t>0
$$

so $\widehat{u(s, t)}=c(s) e^{-k s^{2} t}$.
Taking the Fourier Transform of the initial condition $u(x, 0)=\delta(x)$ gives $u(\widehat{s, t}=0)=1$ and so $c(s)=1$.
So $\widehat{u(s, t)}=e^{-k s^{2} t}$. From tables or otherwise we have that

$$
\mathcal{F}\left(\frac{1}{\sqrt{4 \pi k t}} e^{-\frac{x^{2}}{4 k t}}\right)=e^{-k s^{2} t}
$$

as in general

$$
\mathcal{F}\left(e^{-a \times 2}\right)=\sqrt{\frac{\pi}{a}} e^{-s^{2} / 4 a}
$$

so $u(x, t)=\frac{1}{\sqrt{4 \pi k t}} e^{-x^{2} / 4 k t}$. Note, not defined at $t=0$.

To find the fundamental solution to the heat equation, just translate source to $(\xi, \tau)$ and multiply by a Heaviside function so that the solution is "off" for $t<\tau$. So

$$
K(x, t ; \xi, \tau)=\frac{H(t-\tau)}{\sqrt{4 \pi k(t-\tau)}} e^{\frac{(x-\xi)^{2}}{4 K(t-\tau)}}
$$

(Exercise: check that $\mathrm{K}(x, \mathrm{t} ; \xi, \tau)$ is a distributional solution to $\mathrm{Lu}=\delta(x, \mathrm{t} ; \xi, \tau)$.
We now try to understand the structure of this solution. Fix $\tau=0$, so we have

$$
K(x, t ; \xi, 0)=\frac{1}{\sqrt{4 \pi k t}} e^{\frac{-(x-\xi)^{2}}{4 k t}}, \quad t>0
$$

Obviously

$$
\begin{aligned}
K \rightarrow 0 & \text { as } \\
\mathrm{K} \rightarrow 0^{+}, & x \neq \xi \\
K \rightarrow 0 & \text { as } \\
\mathrm{K} \rightarrow \infty, & x \in \mathbb{R} \\
K \rightarrow \infty & \text { as } \\
\mathrm{K} \rightarrow 0^{+}, & x=\xi
\end{aligned}
$$

So the response of the system has a spike at $t=0$ at $x=\xi$ while going to zero at $t=0$ for $x \neq \xi$. Physically we interpret the fundamental solution $K$ as the temperature distribution in an infinite bar initially at $0^{\circ}$ with an instantaneous unit "pulse" of heat applied at the point $x=\xi$ at time $t=0$. We can confirm that it is a unit source of heat as

$$
\int_{-\infty}^{\infty} K(x, t ; \xi) d x=1, \quad t>0
$$

As $t$ increases, the "pulse" profile spreads, remaining centred at $x=\xi$. We can now solve initial value problems for the 1-D Heat Equation:

$$
\begin{aligned}
u_{t}-k u_{x x} & =0, \quad x \in \mathbb{R}, \quad t>0 \\
u(x, 0) & =f(x), \quad x \in \mathbb{R} \\
u(x, t) & =\int_{-\infty}^{\infty} \frac{1}{\sqrt{4 \pi k t}} e^{\frac{-(x-\xi)^{2}}{4 k t}} f(\xi) d \xi
\end{aligned}
$$

Now finally, consider the inhomogeneous problem:

$$
\begin{aligned}
u_{t}-k u_{x x} & =f(x, t) \quad x \in \mathbb{R}, t>0 \\
u(x, 0) & =0, x \in \mathbb{R} .
\end{aligned}
$$

The solution is:

$$
\begin{aligned}
u(x, t) & =\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{H(t-\tau)}{\sqrt{4 \pi k(t-\tau)}} e^{\frac{-(x-\xi)^{2}}{4 k(t-\tau)}} f(\xi, \tau) d \xi d \tau \\
& =\int_{0}^{t} d \tau \int_{-\infty}^{\infty} d \xi \frac{1}{\sqrt{4 \pi k(t-\tau)}} e^{\frac{-(x-\xi)^{2}}{4 k(t-\tau)}} f(\xi, \tau)
\end{aligned}
$$

### 5.7 Exercises

1. Let $\phi \in T=C_{0}^{\infty}(a, b)$ be a test function. For which of the following definitions of $\psi_{n}$ does $\psi_{n} \rightarrow 0$ in $T$ ?
(a) $\psi_{n}(x)=\frac{1}{n} \phi(x)$
(b) $\psi_{n}(x)=\frac{1}{n} \phi\left(\frac{x}{n}\right)$
(c) $\psi_{n}(x)=\frac{1}{n} \phi(n x)$
2. Prove the following:
(a) $x \delta^{\prime}(x)=-\delta(x)$
(b) $\alpha(x) \delta^{\prime}(x)=-\alpha^{\prime}(0) \delta(x)+\alpha(0) \delta^{\prime}(x)$ for any $\alpha \in C^{\infty}(\mathbb{R})$.
3. Show that the Sturm-Liouville operator $A u=-\left(p u^{\prime}\right)^{\prime}+q u$ is formally self-adjoint.
4. Is $f(x)=1 / x$ locally integrable on $(0,1)$ ?
5. Is the function $e^{x^{2}}$ locally integrable on $\mathbb{R}$ ? Does it generate a distribution in $D(\mathbb{R})$ ?
6. Show that for any locally integrable function $f$ on $\mathbb{R}$, the function $u(x, y)=f(x-y)$ is a weak solution to the equation $u_{x}+u_{y}=0$ on $\mathbb{R}^{2}$.
7. In the quarter plane in $\mathbb{R}^{2}$ find the Green's Function associated with the boundary value problem

$$
\begin{aligned}
\Delta u & =\delta(x, y ; \xi, \eta) \quad x, y ; \xi, \eta>0 \\
u & =(x, 0)=0, x>0 \quad u(0, y)=0, \quad y>0
\end{aligned}
$$

Hint - put image charges in the other quadrants.
8. In the upper half plane in $\mathbb{R}^{2}$ use an image charge to find the Green's Function for the Neumann problem

$$
\begin{aligned}
\Delta u & =\delta(x, y ; \xi, \eta) \quad x, \xi \in \mathbb{R} ; y, \eta>0 \\
u_{y}(x, 0) & =0, x \in \mathbb{R}
\end{aligned}
$$

## Part III

## Complex Analysis

## 6 Calculus in $\mathbb{C}$

In this final Part of the course we will study differential and integral calculus in the complex plane $\mathbb{C}$ and see how the results some of them quite unexpected based on our knowledge of calculus on $\mathbb{R}$ - can be applied. The basic ideas should be familiar and are summarised here for reference.

### 6.1 Brief Summary of Terms and Ideas

Some equations have no real solution - e.g. $x^{2}=-1$. This motivated the invention of complex numbers.

Definition 6.1 (Complex Number) $A$ complex number $z$ is an ordered pair $(\mathrm{x}, \mathrm{y})$ of real numbers; we write $\mathrm{z}=(\mathrm{x}, \mathrm{y})$. We call $x$ the real part and $y$ the imaginary part of $z$. Write $x=\mathfrak{R z}$ and $\mathrm{y}=\mathfrak{I} z$. Addition is defined as usual for ordered pairs while the product of two complex numbers $z_{1}$ and $z_{2}$ is

$$
z_{1} z_{2}=\left(x_{1} x_{2}-y_{1} y_{2}, x_{1} y_{2}+x_{2} y_{1}\right)
$$

It follows from the definition that complex numbers of the form $(x, 0)$ have exactly the same properties as the corresponding real number $x$. For that reason we write (in a slightly sloppy notation) $(x, 0)=x$.

The complex number $(0,1)$ is denoted by $i$ and it follows from the definition of the complex product that

$$
\mathfrak{i}^{2}=-1
$$

It follows that any complex number $z=(x, y)$ can be written as $z=x+i y$.

Definition 6.2 (Modulus) The modulus $|z|$ of a complex number $z=(x, y)$ is just the length of the line segment from the origin to the point in the plane whose coordinates are ( $\mathrm{x}, \mathrm{y}$ ). So we write:

$$
|z|=\sqrt{x^{2}+y^{2}} .
$$

Definition 6.3 (Complex Conjugate) The complex conjugate $\bar{z}$ of a complex number $z=(x, y)$ is just its reflection in the x -axis so

$$
\bar{z}=x-\mathfrak{i} y .
$$

It follows that $z \bar{z}=x^{2}+y^{2}=|z|^{2}$.

Definition 6.4 (Quotient) Given two two complex numbers $z_{1}=x_{1}+i y_{1}$ and $z_{2}=x_{2}+i y_{2}$ we can define the quotient $z_{1} / z_{2}$ by "multiplying above $\mathcal{B}$ below by" $\bar{z}_{2}$ so that the denominator is real:

$$
\frac{z_{1}}{z_{2}} \equiv \frac{z_{1} \bar{z}_{2}}{z_{2} \bar{z}_{2}} \quad \text { a real denominator. }
$$

It is easy to check that with this definition, $\frac{z_{1}}{z_{2}} \times z_{2}=z_{1}$ as expected.

Definition 6.5 (Polar Form) Any complex number $z=x+i y$ can be expressed in polar form by $\mathrm{x}=\mathrm{r} \cos \theta, \mathrm{y}=\mathrm{r} \sin \theta$. Then r is called the modulus of $z-$ written $|z|=\sqrt{x^{2}+y^{2}}$ and $\theta$ is called the argument of $z$, written $\theta=\arg z=\arctan \frac{y}{x}$.

The rules for multiplication and division in polar form follow directly from the definition and are:

$$
\begin{aligned}
z_{1} z_{2} & =r_{1} r_{2}\left(\cos \left(\theta_{1}+\theta_{2}\right)+i \sin \left(\theta_{1}+\theta_{2}\right)\right) \\
\frac{z_{1}}{z_{2}} & =\frac{r_{1}}{r_{2}}\left(\cos \left(\theta_{1}-\theta_{2}\right)+i \sin \left(\theta_{1}-\theta_{2}\right)\right)
\end{aligned}
$$

Definition 6.6 (Roots) The $\mathrm{n}^{\text {th }}$ roots of a complex number $\boldsymbol{z}$ may be easily checked to be the set of n complex numbers

$$
\sqrt[n]{z}=\sqrt[n]{r}\left(\cos \frac{\theta+2 k \pi}{n}+i \sin \frac{\theta+2 k \pi}{n}\right), \quad k=0,1,2, \ldots, n-1
$$

Example 6.1 A important special case are the $\mathfrak{n} n^{\text {th }}$ roots of unity. For example, using the above definition we find that $\sqrt[3]{1}=1,-\frac{1}{2} \pm \frac{1}{2} \sqrt{3} i$.

Finally some properties of sets in $\mathbb{C}$ :
Definition 6.7 (Open Set) $A$ set $\mathrm{S} \subset \mathbb{C}$ is open if every point in S has a neighbourhood consisting entirely of points that are in S .

Informally an open set has a "fuzzy boundary". Given any point in the set, no matter how close to the boundary, we can draw a small circle round it that is still entirely contained in the set.

The most important example of an open set is an "open disk" $\{z \in \mathbb{C}|\quad| z-a \mid<r\}$.

Definition 6.8 (Disk) $A$ disk in the complex plane is the set of complex numbers satisfying $|z-a| \leq \rho$ where the complex number $\mathbf{a}$ is the centre and $\rho$ is the radius. The set $|z-a|<\rho$ is called an open disk as the boundary points are excluded. An open disk centred at a point a is often called a neighbourhood of a.

Definition 6.9 (Connected Set) An open set S is connected if any two of its points can be joined by a "zig-zag" line of finitely many straight line segments all of whose points are in S .

An open connected set is called a domain. Informally, a domain can have "holes" but must not consist of separate pieces.


Figure 1: A domain $\mathrm{D} \subseteq \mathbb{C}$

### 6.2 Limits and Derivatives

A complex function $f$ is just a function $f: \mathbb{C} \rightarrow \mathbb{C}$. We usually refer to the real $u(x, y)$ and imaginary $v(x, y)$ parts of a complex function $f(z)$ where $w=f(z)=u(x, y)+\mathfrak{i} v(x, y)$. Limits are defined exactly as on $\mathbb{R}$, except that now the term $|z-a|<\rho$ means the open disk centred at a rather than the open interval $(a-\rho, a+\rho)$. Just as on $\mathbb{R}$, a function is continuous at $z=z_{0}$ if $\lim _{z \rightarrow z_{0}} f(z)=f\left(z_{0}\right)$.

Definition 6.10 (Derivative) Just as on $\mathbb{R}$, we define

$$
f^{\prime}\left(z_{0}\right)=\lim _{\Delta z \rightarrow 0} \frac{f\left(z_{0}+\Delta z\right)-f\left(z_{0}\right)}{\Delta z}
$$

provided the limit exists. Alternatively we can write

$$
f^{\prime}\left(z_{0}\right)=\lim _{z \rightarrow z_{0}} \frac{f(z)-f\left(z_{0}\right)}{z-z_{0}}
$$

It is important to remember that the rules for differentiation are exactly the same as for real calculus as the proofs are exactly the same, line for line.

The first major difference from real analysis is the crucial idea of analyticity - a bit more than differentiability, as we will see.

Definition 6.11 (Analytic Functions) A function f is analytic in a domain D if $\mathrm{f}(\mathrm{z})$ is defined and differentiable at all points of D. We say that a function is analytic at a point $z_{0}$ if f is analytic in some (perhaps very small) neighbourhood of $z_{0}$.

We are asking a lot for $f$ to be analytic at a point $z_{0}$ - $f$ must be differentiable at every point in some neighbourhood of $z_{0}$ and we can take the limit $\Delta x \rightarrow 0$ along any path towards zero. We will see that if a function is analytic in a domain we will be able to deduce some interesting properties as a consequence.

### 6.3 Cauchy-Riemann Equations

We now derive a very important result which will give us a simple test for analyticity of a complex function

$$
w=f(z)=u(x, y)+\mathfrak{i} v(x, y)
$$

We will show that a complex function $f$ is analytic in a domain $D$ if and only if the first partial derivatives of $u$ and $v$ satisfy the two Cauchy-Riemann equations:

$$
\begin{equation*}
u_{x}=v_{y} ; \quad u_{y}=-v_{x} \tag{6.1}
\end{equation*}
$$

at every point in $D$.
Example 6.2 We know that the function $\mathrm{f}(\mathrm{z})=z^{2}=\mathrm{x}^{2}-\mathrm{y}^{2}+2 \mathrm{ixy}$ is analytic for all $z$ as its derivative $2 z$ is defined everywhere on $\mathbb{C}$. We find $u_{x}=2 x=v_{y}$ and $u_{y}=-2 \mathrm{y}=-v_{\mathrm{x}}$ so the $C-R$ Eqs are satisfied.

Example 6.3 For $f(z)=\bar{z}=x-\mathfrak{i} y$, we have $u=x$ and $v=-y$ so $u_{x}=1 \neq v_{y}=-1$. So the complex conjugate function is not analytic - we could also show this from first principles with a lot more work.

We break the "if and only if" into two separate Theorems.
Theorem 6.1 (Cauchy-Riemann Equations) Let
$\mathrm{f}(\mathrm{z})=\mathfrak{u}(\mathrm{x}, \mathrm{y})+\mathfrak{i v}(\mathrm{x}, \mathrm{y})$ be defined and continuous in some neighbourhood of a point $z=x+i y$ and differentiable at $z$. Then at that point, the first partial derivatives of $u$ and $v$ exist and satisfy the Cauchy-Riemann equations (6.1). As a consequence, if f is analytic in a domain D , the first partial derivatives of $u$ and $v$ exist and satisfy the Cauchy-Riemann equations at all points of D .

Proof: We are given that $f^{\prime}(z)$ exists for all $z \in D$. It is given by

$$
\begin{equation*}
f^{\prime}(z)=\lim _{\Delta z \rightarrow 0} \frac{f(z+\Delta z)-f(z)}{\Delta z} . \tag{6.2}
\end{equation*}
$$

Writing $f(z)=u(x, y)+\mathfrak{i v}(x, y)$ and $\Delta z=\Delta x+\mathfrak{i} \Delta y$, we have

$$
\begin{align*}
& f^{\prime}(z)= \\
& \lim _{\Delta z \rightarrow 0} \frac{[\mathfrak{u}(x+\Delta x, y+\Delta y)-\mathfrak{u}(x, y)]+\mathfrak{i}[v(x+\Delta x, y+\Delta y)-v(x, y)]}{\Delta x+\mathfrak{i} \Delta y} \tag{6.3}
\end{align*}
$$

As the limit may be taken along any path to $\Delta z=0$, we just take two particular choices (I \& II) and equate the results - letting $\Delta z=\Delta x+i \Delta y$ we have
(I) $\Delta y \rightarrow 0$ first then $\Delta x \rightarrow 0$
(II) $\Delta x \rightarrow 0$ first then $\Delta y \rightarrow 0$

Now take the limit over each path separately:
(I) So after $\Delta y \rightarrow 0, \Delta z=\Delta x$. Then (6.3) becomes

$$
\begin{aligned}
f^{\prime}(z)= & \lim _{\Delta x \rightarrow 0} \frac{[u(x+\Delta x, y)-u(x, y)]}{\Delta x} \\
& +i \lim _{\Delta x \rightarrow 0} \frac{[v(x+\Delta x, y)-v(x, y)]}{\Delta x}
\end{aligned}
$$

Since $f^{\prime}(z)$ exists, the two (one real, one imaginary) limits on the RHS must exists. They are just the partial derivatives of $u$ and $v$ w.r.t. $x$. So

$$
\begin{equation*}
f^{\prime}(z)=u_{x}+i v_{x} \tag{6.4}
\end{equation*}
$$

(II) Now after $\Delta x \rightarrow 0, \Delta z=i \Delta y$. Using the same steps as for path (I), we find

$$
\begin{equation*}
f^{\prime}(z)=-i u_{y}+i v_{y} \tag{6.5}
\end{equation*}
$$

Equating real parts and imaginary parts respectively in (6.4) and (6.5) gives us the result.

We now state and prove the converse result.
Theorem 6.2 (Cauchy-Riemann Equations) If two real-valued functions $\mathfrak{u}(\mathrm{x}, \mathrm{y})$ and $v(\mathrm{x}, \mathrm{y})$ have continuous first partials that satisfy the Cauchy-Riemann equations in some domain D , then the complex function $\mathrm{f}(\mathrm{z})=\mathfrak{u}(\mathrm{x}, \mathrm{y})+\mathfrak{i} v(\mathrm{x}, \mathrm{y})$ is analytic in D .

Before the proof: what does this mean? We are assuming that the two real-valued $u(x, y)$ and $v(x, y)$ are differentiable on $D$. What we need to prove is that the complex-valued function $f(z) \equiv \mathfrak{u}(x, y)+\mathfrak{i v}(x, y)$ is analytic in $D$. This seems "obvious" but takes some work to prove. The interesting point is that it is not enough for the partial derivatives of $u$ and $v$ to exist - we need the C-R Eqs to hold as well. This is not "obvious"!

Proof: We take the proof in three steps.
(A) Let $\mathrm{P}(\mathrm{x}, \mathrm{y})$ be any point in D . As D is a domain by taking the horizontal and vertical steps $|\Delta x|$ and $|\Delta y|$ sufficiently small we can choose a point $Q(x+\Delta x, y+\Delta y)$ such that the line segment PQ also lies in D. Now apply the Mean Value Theorem:
$\mathfrak{R} \Delta \mathrm{f} \equiv \mathrm{u}(\mathrm{x}+\Delta \mathrm{x}, \mathrm{y}+\Delta \mathrm{y})-\mathrm{u}(\mathrm{x}, \mathrm{y})=\Delta \mathrm{x} \cdot \mathrm{u}_{\mathrm{x}}\left(\mathrm{M}_{1}\right)+\Delta \mathrm{y} \cdot \mathrm{u}_{\mathrm{y}}\left(M_{1}\right.$ $\mathfrak{I} \Delta f \equiv v(x+\Delta x, y+\Delta y)-v(x, y)=\Delta x \cdot v_{x}\left(M_{2}\right)+\Delta y \cdot v_{y}\left(M_{2}\right)$
where $M_{1}$ and $M_{2}$ are some unknown points on the line PQ. (See Fig. 2 on the next Slide.)


Figure 2: The line seqment $\mathrm{P}-\mathrm{Q}$
(B) Use the C-R equations $\left(u_{x}=v_{y}\right.$ and $\left.u_{y}=-v_{x}\right)$ to eliminate $u_{y}$ and $v_{y}$ from from the above equations for $\mathfrak{R} \Delta f$ and $\mathfrak{I} \Delta f$. Now, using $\Delta f \equiv \mathfrak{R} \Delta f+i \Im \Delta f$, we have
$\Delta f=\left[\Delta x \cdot u_{x}\left(M_{1}\right)-\Delta y \cdot v_{x}\left(M_{1}\right)\right]+i\left[\Delta x \cdot v_{x}\left(M_{2}\right)+\Delta y \cdot u_{x}\left(M_{2}\right)\right]$.
If we replace $\Delta x$ by $\Delta z-i \Delta y$ in the first term and $\Delta y$ by $(\Delta z-\Delta x) / i$ in the second term we find that:

$$
\begin{aligned}
\Delta f=(\Delta z-i \Delta y) u_{x}\left(M_{1}\right) & +i(\Delta z-\Delta x) v_{x}\left(M_{1}\right) \\
& +i\left[\Delta x \cdot v_{x}\left(M_{2}\right)+\Delta y \cdot u_{x}\left(M_{2}\right)\right]
\end{aligned}
$$

Expanding \& re-ordering:

$$
\begin{aligned}
\Delta f=\Delta z \cdot \mathbf{u}_{x}( & \left.\mathbf{M}_{1}\right)-i \Delta y\left\{\mathbf{u}_{\mathbf{x}}\left(\mathbf{M}_{1}\right)-\mathbf{u}_{\mathbf{x}}\left(\mathbf{M}_{\mathbf{2}}\right)\right\} \\
& +\mathfrak{i}\left[\Delta z \cdot v_{x}\left(\mathbf{M}_{1}\right)-\Delta x\left\{\mathbf{v}_{\mathbf{x}}\left(\mathbf{M}_{\mathbf{1}}\right)-\mathbf{v}_{\mathbf{x}}\left(\mathbf{M}_{\mathbf{2}}\right)\right\}\right]
\end{aligned}
$$

(C) The terms coloured blue (in braces $\{\ldots\}$ ) go to zero as $|\Delta z| \rightarrow 0$ because as $\Delta z \rightarrow 0, \mathrm{Q}$ approaches P and so also must $M_{1}$ and $M_{2}$. Therefore as the partials are assumed to be continuous, the differences of $u_{x}$ and $v_{x}$ at $M_{1}$ and $M_{2}$ must go to zero.
Also $\left|\frac{\Delta x}{\Delta z}\right|<1$ and $\left|\frac{\Delta y}{\Delta z}\right|<1$ so dividing $\Delta f$ by $\Delta z$ and taking the limit we find that

$$
\begin{aligned}
f^{\prime}(z) \equiv \lim _{\Delta z \rightarrow 0} \frac{\Delta f}{\Delta z} & =\lim _{\Delta z \rightarrow 0}\left(u_{x}\left(M_{1}\right)+i v_{x}\left(M_{1}\right)\right)+0+0 \\
& =u_{x}(x, y)+i v_{x}(x, y)
\end{aligned}
$$

So $f^{\prime}(z)$ is defined everywhere in $D$ and therefore $f$ is analytic in $D$.

One reason for the great importance of complex analysis in Applied Mathematics is that the real and imaginary parts $u$ and $v$ of any analytic function $f(z)=u(x, y)+\mathfrak{i v}(x, y)$ both satisfy Laplace's Equation:

$$
\begin{align*}
\nabla^{2} u \equiv u_{x x}+u_{y y} & =0  \tag{6.6}\\
\nabla^{2} v \equiv v_{x x}+v_{y y} & =0 \tag{6.7}
\end{align*}
$$

This is easy to check - we state the result as a Theorem.
Theorem 6.3 (Laplace's Equation) If $f(z)=u(x, y)+\mathfrak{i v}(x, y)$ is analytic in a domain D , then $u$ and $v$ satisfy Laplace's Equation in D and have continuous second derivatives in D .

Proof: If we differentiate the equation $u_{x}=v_{y}$ w.r.t. $x$ and the equation $u_{y}=-v_{x}$ w.r.t. $y$, we have

$$
\begin{equation*}
u_{x x}=v_{y x}, \quad u_{y y}=-v_{x y} \tag{6.8}
\end{equation*}
$$

We will see later that derivatives of an analytic function are themselves analytic so $u$ and $v$ have continuous partials of all orders. In particular the mixed second derivatives are equal: $v_{x y}=v_{y x}$. Adding the two equations in (6.8) gives us (6.6). The proof of (6.7) is left as an exercise.

## 7 Complex Integration

We will study complex integration for two reasons:

- Some real integrals can be more easily evaluated in terms of complex integrals.
- Some basic properties of analytic functions need complex integration for their proof.

In this Chapter we will define line integrals in $\mathbb{C} \&$ study their properties. The main result will be Cauchy's Theorem from which most of the properties of analytic functions follow.

### 7.1 Line integrals in $\mathbb{C}$

In real analysis, a definite integral is taken over an interval of the real line. For definite integrals in $\mathbb{C}$, we integrate along a curve $\mathbb{C}$ in $\mathbb{C}$ - called the path of integration. A curve $\mathbb{C}$ in $\mathbb{C}$ can be represented as

$$
\begin{equation*}
z(t)=x(t)+i y(t), \quad a \leq t \leq b \tag{7.1}
\end{equation*}
$$

where $t$ is a real parameter.
Example 7.1

$$
\begin{array}{lrl}
z(t)=2 t+3 i t, & 0 \leq t \leq 1 & \text { segment of line } y=3 / 2 x \\
z(t)=2 \cos t+i 2 \sin t & 0 \leq t \leq 2 \pi & \text { circle }|z|=2
\end{array}
$$

Definition 7.1 A curve C is smooth if it has a derivative w.r.t.t;

$$
\dot{z}(\mathrm{t})=\frac{\mathrm{d} z}{\mathrm{dt}}=\dot{x}(\mathrm{t})+\mathrm{i} \dot{y}(\mathrm{t})
$$

that is continuous and nowhere zero. Geometrically this means that at every point on the curve there is a tangent vector that varies continuously as we move along the curve. If the derivative $\dot{z}\left(\mathrm{t}_{0}\right)=0$ for some $\mathrm{t}=\mathrm{t}_{0}$ then the curve does not have a tangent at the point $z\left(\mathrm{t}_{0}\right)$ and so is not smooth.

Draw a sketch graph of the curve with the tangent vector $\dot{z}$ at $z(t)$, together with the arrows representing the vectors $z(t), z(t+\Delta t)$ and $z(t+\Delta t)-z(t)$. The latter becomes parallel with $\dot{z}$ in the limit as $\Delta t \rightarrow 0$. (Why?)

Definition 7.2 (Complex Line Integral) This definition is included for reference purposes and is very similar to the corresponding definition on $\mathbb{R}$.

- Let $\mathbb{C}$ be a smooth curve in $\mathbb{C}$ with endpoints $z_{\mathrm{a}}$ and $z_{\mathrm{b}}$. Let $f(z)$ be a continuous function defined at each point of the curve.
- We partition the interval $\mathrm{a} \leq \mathrm{t} \leq \mathrm{b}$ into sub-intervals by the points $a=t_{0}, t_{1}, \ldots, t_{n-1}, t_{n}=b$.
- To these t -values there correspond the points $z_{\mathrm{a}}=z_{0}, z_{1}, \ldots, z_{\mathrm{n}}=z_{\mathrm{b}}$ on the curve C .
- On each sub-curve with end points $\boldsymbol{z}_{\mathfrak{i}}$ and $\boldsymbol{z}_{\mathfrak{i}+1}$ we choose an arbitrary point $\xi_{i+1}-$ so that $\xi_{i+1}=z(t)$ for some $t$ such that $\mathrm{t}_{\mathrm{i}} \leq \mathrm{t} \leq \mathrm{t}_{\mathrm{i}+1}$ - see Fig. 3 on the next Slide.


Figure 3: Partitioning a curve

- We can partition the curve arbitrarily - but we impose the condition that the largest of the $\left|\Delta z_{i}\right| \rightarrow 0$ as $n$ (the number of pieces) goes to infinity.
- In other words the length of all the pieces goes to zero in the limit as the number of pieces goes to infinity.
- Now form the sum

$$
\begin{equation*}
S_{n}=\sum_{i=1}^{n} f\left(\xi_{i}\right) \Delta z_{i} \tag{7.2}
\end{equation*}
$$

where $\Delta z_{\mathfrak{i}}=z_{\mathfrak{i}}-z_{\mathfrak{i}-1}$.

- The limit of the sequence $\mathrm{S}_{\mathrm{n}}$ is called the line integral and we write (with the above restriction on the maximum chord length)

$$
\begin{equation*}
\lim _{n \rightarrow \infty} S_{n}=\int_{C} f(z) d z \tag{7.3}
\end{equation*}
$$

Definition 7.3 The curve C is called the path of integration. If the start point and the end point are the same (e.g. a circle or ellipse) we say that the path is closed and write:

$$
\oint_{C} f(z) d z
$$

It can be shown that if $C$ is piecewise smooth and $f$ is continuous then the complex line integral exists - based on theorems for the existence of the real integral.

Three important properties follow directly from Def. 7.2 we list them without proof as they are easily checked:

1. Integration is linear:

$$
\int_{C}\left(k_{1} f_{1}(z)+k_{2} f_{2}(z)\right) d z=k_{1} \int_{C} f_{1}(z) d z+k_{2} \int_{C} f_{2}(z) d z
$$

2. The integral over two successive sub-paths is the sum of the integrals: when a curve $C$ is decomposed into two sub-paths $C_{1}$ and $\mathrm{C}_{2}$, then

$$
\int_{C} f(z) d z=\int_{C_{1}} f(z) d z+\int_{C_{2}} f(z) d z
$$

3. Reversing the direction (sense) of integration changes the sign of the integral. Where $C_{+}$and $C_{-}$are the same curve $C$ traversed in opposite directions:

$$
\int_{C_{+}} f(z) d z=-\int_{C_{-}} f(z) d z
$$

### 7.2 Integration Methods

There are two standard methods:

1. Use a parameterisation of the path. We summarise the method with a Theorem.

Theorem 7.1 Let C is piecewise smooth $(z=z(t)$, for $\mathrm{a} \leq \mathrm{t} \leq \mathrm{b}$ ). If f is continuous on C then

$$
\begin{equation*}
\int_{C} f(z) d z=\int_{a}^{b} f(z(t)) \dot{z}(t) d t \tag{7.4}
\end{equation*}
$$

Proof: Simply substitute $f(z)=u(z)+\mathfrak{i v}(z)$ and $z=x+\mathfrak{i y}$ and use the corresponding result for real integrals.

## Example 7.2 (Integral of $\frac{1}{z}$ round the unit circle)

Parameterise the unit circle $C$ by $z(t)=\cos t+i \sin t$, where $0 \leq t \leq 2 \pi$ (anti-clockwise). Then $\dot{z}(\mathrm{t})=-\sin \mathrm{t}+\mathrm{i} \cos \mathrm{t}$, and of course $\mathrm{f}(\mathrm{z}(\mathrm{t}))=\frac{1}{\mathrm{z}(\mathrm{t})}=\frac{1}{\cos \mathrm{t}+\mathrm{i} \sin \mathrm{t}}=\cos \mathrm{t}-\mathrm{i} \sin \mathrm{t}$. So

$$
\begin{aligned}
\oint_{|z|=1} \frac{1}{z} d z & =\int_{0}^{2 \pi}(\cos t-i \sin t)(-\sin t+i \cos t) d t \\
& =i \int_{0}^{2 \pi} d t=2 \pi i
\end{aligned}
$$

Integrals like this are easier if we use the representation $z(t) \equiv \cos t+i \sin t=e^{i t}$ as then immediately we have $\dot{z}(\mathrm{t})=i e^{i t}$ and $\frac{1}{z(\mathrm{t})}=e^{-i t}$.

## Example 7.3 (Integral of integer powers) Let

$\mathrm{f}(z)=\left(z-z_{0}\right)^{n}, \mathfrak{n}$ an integer and $z_{0}$ a constant. Let C be a circle radius r centred at $z_{0}$. Parameterise C by $z(\mathrm{t})=z_{0}+\mathrm{re}^{\mathrm{it}}$. Then $\left(z-z_{0}\right)^{\mathrm{n}}=\mathrm{r}^{\mathrm{n}} \mathrm{e}^{\mathrm{int}}$ and $\dot{z}(\mathrm{t})=\mathrm{ir}^{\mathrm{n}} \mathrm{e}^{\mathfrak{i t}}$. Assembling the pieces we have

$$
\oint f(z) d z=\int_{0}^{2 \pi} r^{n} e^{i n t} i r e^{i t} d t=i r^{n+1} \int_{0}^{2 \pi} e^{i(n+1) t} d t
$$

Using the Euler formula $e^{i t}=\cos t+i \sin t$ we find that the real and imaginary parts of the complex integral above vanish unless $\mathrm{n}=-1$ when we recover the result from the previous Example.

Example 7.4 (Integral of a non-analytic function) Let $\mathrm{f}(\mathrm{z})=\mathfrak{R z}=\mathrm{x}$. Take the integral along two different paths:
(a) Let $\mathrm{C}_{1}$ be the line segment from $z=0$ to $z=1+i$. Then $z(\mathrm{t})=(1+\mathfrak{i}) \mathrm{t}, 0 \leq \mathrm{t} \leq 1$ and $\dot{z}=1+\mathrm{i}$. Obviously $\mathrm{f}(\mathrm{z})=\mathrm{x}(\mathrm{t})=\mathrm{t}$. So

$$
\int_{C} f(z) d z=\int_{0}^{1} t(1+\mathfrak{i}) d t=(1+\mathfrak{i}) / 2 .
$$

(b) Let $\mathrm{C}_{2}$ be the path formed by a unit step $\mathrm{C}_{2 \mathrm{a}}$ along the x -axis from $\mathrm{z}=0$ followed by a unit step $\mathrm{C}_{2 \mathrm{~b}}$ along the y -direction. Note that the start and finish points are the same as for $\mathrm{C}_{1}$. Then on $\mathrm{C}_{2 \mathrm{a}}$ we have $\mathrm{z}(\mathrm{t})=\mathrm{t}, 0 \leq \mathrm{t} \leq 1$ and $\mathrm{f}(\mathrm{z})=\mathrm{x}(\mathrm{t})=\mathrm{t}$. So $\dot{\mathrm{z}}=1$ and so

$$
\int_{C_{2 a}} f(z) d z=\int_{0}^{1} t d t=1 / 2 .
$$

On $\mathrm{C}_{2 \mathrm{~b}}$ we have $\mathrm{z}(\mathrm{t})=1+\mathrm{it}, 0 \leq \mathrm{t} \leq 1$ and $f(z)=x(t)=1$. So $\dot{z}=\mathfrak{i}$ and so

$$
\int_{C_{2 b}} f(z) d z=\int_{0}^{1} 1 i d t=i
$$

So $\int_{C} f(z) d z \equiv \int_{C_{2 a}} f(z) d z+\int_{C_{2 b}} f(z) d z=1 / 2+i$. The answer is different to that found for path $\mathrm{C}_{1}$ - this is not a surprise.

We will see later that analytic functions have the nice property that the value of a line integral depends only on the start and end points, not on the choice of path between.
2. Our second integration method: just as in real calculus, we can use indefinite integrals - in other words, if we know that $f(z)=F^{\prime}(z)$, then

$$
\int_{C} f(z) d z=F(b)-F(a), \quad \text { where } C \text { is any path from } a \text { to } b
$$

We state a Theorem to make this precise:
Theorem 7.2 Let $f(z)$ be analytic in a simply connected domain D . Then there is an indefinite integral of $\mathrm{f}(\mathrm{z})$ in D , that is an analytic function $F(z)$ such that $F^{\prime}(z)=f(z)$ in $D$. Also, for any path in D connecting two points a and b in D we have

$$
\begin{equation*}
\int_{a}^{b} f(z) d z=F(b)-F(a) \tag{7.5}
\end{equation*}
$$

We will prove the Theorem after first proving Cauchy's Theorem - needed for the proof.

Example $7.5 \int_{0}^{1+i} z^{2} d z=\left.\frac{z^{3}}{3}\right|_{0} ^{1+i}=(1+i)^{3} / 3=-\frac{2}{3}+\frac{2}{3} i$.
Example $7.6 \int_{-\pi i}^{\pi i} \cos z d z=\left.\sin z\right|_{-\pi i} ^{\pi i}=2 \sin \pi i=2 i \sinh \pi$.
Example 7.7
$\int_{8+\mathfrak{i} \pi}^{8-3 i \pi} e^{z / 2} \mathrm{~d} z=\left.2 e^{z / 2}\right|_{8+\mathfrak{i} \pi} ^{8-3 i \pi}=2\left(e^{4-3 / 2 i \pi}-e^{4+\mathfrak{i} \pi / 2}\right)=0$ as $e^{z}$ is periodic with period $2 \pi$.

A result we need on line integrals before Cauchy's Theorem:

## Theorem 7.3 (ML-inequality)

$$
\begin{equation*}
\left|\int_{C} f(z) d z\right| \leq M L \tag{7.6}
\end{equation*}
$$

where L is the length of C and M is a constant such that $|f(z)| \leq M$ everywhere on the path C .

Proof: Check that it follows directly from the definition of $\int_{C} f(z) d z$.
Example 7.8 Find an upper bound for the absolute value of the integral $\int_{\mathrm{C}} z^{2} \mathrm{~d} z$ where C is the line segment from 0 to $1+\mathrm{i}$. Obviously $\mathrm{L}=\sqrt{2}$. The maximum value (check) that $\left|z^{2}\right|$ takes is at $z=1+\mathfrak{i}$, namely $M=2$ so our upper bound on the modulus of the integral is $2 \sqrt{2} \approx 2.8$. The modulus of the integral (Example 7.5) is $|-2 / 3+2 / 3 i|=2 / 3 \sqrt{2} \approx 0.94$ so the $M L$-inequality considerably over-estimates the result.

### 7.3 Cauchy's Theorem

First, two more definitions.
Definition 7.4 A simple closed path C is one that neither touches nor intersects itself. A circle is a simple closed path, a figure-eight is not.

Definition 7.5 A domain D in $\mathbb{C}$ is a simply connected domain if every simple closed path in D encloses only points in D .

Drawing a few sketches should convince you that the interior of a simple closed path is a simply connected set.

Definition 7.6 A simple closed path is called a contour and an integral over a contour is called a contour integral.

Now to the Theorem. First we'll state it, then look at some examples and finally prove it.

Theorem 7.4 (Cauchy's) If $f(z)$ is analytic in a simply connected domain D , then for every simple closed path C

$$
\begin{equation*}
\oint_{C} f(z) d z=0 \tag{7.7}
\end{equation*}
$$

This is a very general result and has some surprising consequences. First, some examples.

Example 7.9 For any contour $C, \oint_{C} e^{z} d z=0, \oint_{C} \cos z d z=0$, $\oint_{C} z^{n} \mathrm{~d} z=0, \quad z=0,1,2, \ldots$ as all three functions are entire, i.e. analytic on $\mathbb{C}$.

Example 7.10 Take $C$ to be the unit circle, then $\oint_{C} \sec z \mathrm{~d} z=0$ even though sec is not analytic at $z= \pm \pi / 2, \pm 3 \pi / 2, \ldots$ as all these points lie outside C. Similarly $\oint \frac{\mathrm{dz}}{z^{2}+4}=0$ as the non-analytic points $\pm 2 i$ lie outside C .

Example 7.11 Take $C$ to be the unit circle anti-clockwise. Then $\oint_{C} \bar{z} \mathrm{~d} z=2 \pi i$ despite Cauchy's Thm. as $\mathrm{f}(\mathrm{z})=\bar{z}$ is not analytic so the Thm. does not apply.

Example 7.12 Take C to be the unit circle anti-clockwise. Then (check) $\oint_{C} \frac{\mathrm{~d} z}{z^{3}}=0$. This does not follow from Cauchy's Thm. as $f(z)=\frac{1}{z^{3}}$ is not analytic at $z=0$.

Example 7.13 Take C to be the unit circle anti-clockwise and D to be the ring or annulus $\frac{1}{2}<|z|<\frac{3}{2}$. Despite $\mathrm{f}(\mathrm{z})$ being analytic in D , we have already seen that $\oint_{\mathrm{C}} \frac{\mathrm{d} z}{z}=2 \pi \mathrm{i}$. This does not contradict Cauchy's Thm. as D is not simply connected.

Example 7.14 Take $C$ to be the unit circle anti-clockwise. Then $\oint_{C} \frac{7 z-6}{z^{2}-2 z} \mathrm{~d} z=\oint_{C} \frac{3}{z} \mathrm{~d} z+\oint_{C} \frac{4}{z-2} \mathrm{~d} z=3 \times 2 \pi i+0=6 \pi i$. Here the value of the first integral comes from Example 7.2 and that of the second from Cauchy's Thm.

## Proof: (Cauchy's Theorem 7.4)

- First suppose (for simplicity) that $C$ is the boundary of a triangle, oriented anti-clockwise. By joining the mid-points of the sides we divide $C$ into four sub-triangles $C_{1}$ to $C_{4}$.
It is easy to see that

$$
\begin{equation*}
\oint_{C} f(z) d z=\oint_{C_{1}} f(z) d z+\oint_{C_{2}} f(z) d z+\oint_{C_{3}} f(z) d z+\oint_{C_{4}} f(z) d z \tag{7.8}
\end{equation*}
$$

as the legs along the sides of the sub-triangles that are in the interior of C cancel - see Fig. 4 on the next Slide.


Figure 4: Construction for proof of Cauchy's Thm.

Now, from the four integrals on the RHS pick the one that is largest in modulus. Re-labelling if necessary, call its path $\mathrm{C}_{1}$. Then by the triangle inequality

$$
\left|\oint_{C} f(z) d z\right| \leq 4\left|\oint_{C_{1}} f(z) d z\right|
$$

Now subdivide $C_{1}$ as we did $C$ and again select the sub-triangle (call it $C_{2}$ ) such that

$$
\left|\oint_{C_{1}} f(z) d z\right| \leq 4\left|\oint_{C_{2}} f(z) d z\right|
$$

so

$$
\left|\oint_{C} f(z) d z\right| \leq 4^{2}\left|\oint_{C_{2}} f(z) d z\right|
$$

This process can be repeated indefinitely. We obtain a sequence of similar triangles $T_{1}, T_{2}, \ldots$, with boundaries $C_{1}, C_{2}, \ldots$, such that $T_{n} \subset T_{m}$ whenever $n>m$ and

$$
\left|\oint_{C} f(z) d z\right| \leq 4^{n}\left|\oint_{C_{n}} f(z) d z\right|
$$

Let $z_{0}$ be a point common to all these nested triangles again, see Fig. 4 above.
As $f$ is differentiable at $z=z_{0}, f^{\prime}\left(z_{0}\right)$ is defined.

Define

$$
h(z)=\frac{f(z)-f\left(z_{0}\right)}{\left(z-z_{0}\right)}-f^{\prime}\left(z_{0}\right) .
$$

Obviously $h(z)$ can be made as small as we want by taking $z$ close enough to $z_{0}$. So let's make this precise: for any $\varepsilon>0$ we can find a number $\delta>0$ so that

$$
\begin{equation*}
|h(z)|<\varepsilon \quad \text { when }\left|z-z_{0}\right|<\delta \tag{7.9}
\end{equation*}
$$

Now, solve for $f(z)$ from (7.3):

$$
f(z)=f\left(z_{0}\right)+\left(z-z_{0}\right) f^{\prime}\left(z_{0}\right)+h(z)\left(z-z_{0}\right)
$$

Now integrate this over the boundary $C_{n}$ of the triangle $T_{n}$ :

$$
\oint_{C_{n}} f(z) d z=\oint_{C_{n}} f\left(z_{0}\right) d z+\oint_{C_{n}}\left(z-z_{0}\right) f^{\prime}\left(z_{0}\right) d z+\oint_{C_{n}} h(z)\left(z-z_{0}\right) d z .
$$

As $f\left(z_{0}\right)$ and $f^{\prime}\left(z_{0}\right)$ are constants and $C_{n}$ is a closed path, the first two integrals on the RHS are zero. (See Q 1a.) So

$$
\oint_{C_{n}} f(z) d z=\oint_{C_{n}} h(z)\left(z-z_{0}\right) d z .
$$

Now take $n$ large enough that the triangle $T_{n}$ lies inside the disk $\left|z-z_{0}\right|<\delta$.
Let $L_{n}$ be the length of $C_{n}$. Then $\left|z-z_{0}\right|<L_{n} / 2$ for all $z$ on $C_{n}$ and $z_{0}$ inside $T_{n}$. (See Q 1b.) Using this and (7.9) we have $\left|h(z)\left(z-z_{0}\right)\right|<\varepsilon L_{n}$. The ML-inequality gives us:

$$
\begin{equation*}
\left|\oint_{C_{n}} f(z) d z\right|=\left|\oint_{C_{n}} h(z)\left(z-z_{0}\right) d z\right|<\varepsilon L_{n} \times L_{n}=\varepsilon L_{n}^{2} . \tag{7.10}
\end{equation*}
$$

Now the sides of the sub-triangles are halved at each iteration so we also have

$$
L_{n}=\frac{L}{2^{n}} \quad \text { So } L_{n}^{2}=\frac{L^{2}}{4^{n}} .
$$

Assembling the pieces we have

$$
\begin{equation*}
\left|\oint_{C} f(z) d z\right| \leq 4^{n}\left|\oint_{C_{n}} f(z) d z\right|<4^{n} \varepsilon L_{n}^{2}=4^{n} \varepsilon \frac{L^{2}}{4^{n}}=\varepsilon L^{2} . \tag{7.11}
\end{equation*}
$$

The RHS can be made as small as we wish by taking $\varepsilon$ sufficiently small so we must have $\oint_{\mathrm{C}} \mathrm{f}(z) \mathrm{d} z=0$ as required.

- The proof for the case where C is the boundary of a polygon follows by subdividing the polygon into triangles.
- Finally it can be shown that any simple closed path C can be approximated as accurately as we wish by a polygon with sufficiently many sides.


### 7.4 Consequences of Cauchy's Theorem

1. One immediate consequence of Cauchy's Theorem is the principle of deformation of path - namely that the integral of an analytic function from (say) $z_{0}$ to $z_{1}$ depends only on the values of $z_{0}$ and $z_{1}$ and not on the path between. The proof is simple: divide the simple closed path $C$ in Cauchy's Theorem into two sub-curves $\mathrm{C}_{1}$ and $\mathrm{C}_{2}$, both in the counter-clockwise sense, say. Then

$$
\oint_{C} f(z) d z=\int_{C_{1}} f(z) d z+\int_{C_{2}} f(z) d z=\int_{C_{1}} f(z) d z-\int_{C_{2}^{-}} f(z) d z=0
$$

where $C_{2}^{-}$is just $C_{2}$ traversed in the opposite direction - so that both $C_{1}$ and $C_{2}^{-}$have the same start and end points; $z_{0}$ and $z_{1}$ say.

It follows that

$$
\begin{equation*}
\int_{C_{1}} f(z) d z=\int_{C_{2}^{-}} f(z) d z \tag{7.12}
\end{equation*}
$$

But $C_{1}$ and $C_{2}^{-}$are any two paths with the same start and end points (on which $f(z)$ is analytic).
Another way of viewing this result is that for a given integral we may deform the path from $z_{0}$ to $z_{1}$ without changing the value of the integral - provided we do not cross a point where $f(z)$ fails to be analytic. Hence the term principle of deformation of path.
2. Our next application of Cauchy's Theorem is the existence of an indefinite integral for any analytic function. We re-state \& prove Theorem 7.2:
Theorem 7.5 Let $\mathrm{f}(\boldsymbol{z})$ be analytic in a simply connected domain D . Then there is an indefinite integral of $\mathrm{f}(\mathrm{z})$ in D , that is an analytic function $\mathrm{F}(z)$ such that $\mathrm{F}^{\prime}(z)=\mathrm{f}(z)$ in D . Also, for any path in D connecting two points a and b in D we have

$$
\begin{equation*}
\int_{a}^{b} f(z) d z=F(b)-F(a) \tag{7.13}
\end{equation*}
$$

Proof: The requirements for Cauchy's Theorem are satisfied . So the line integral of $f(z)$ from any $z_{0}$ in $D$ to any $z$ in $D$ is independent of the path chosen in D. Now fix $z_{0}$. Then

$$
F(z)=\int_{z_{0}}^{z} f\left(z^{\prime}\right) d z^{\prime}
$$

is well-defined as a function on D .

RTP that $F(z)$ is analytic on $D$ and that $F^{\prime}(z)=f(z)$. We will differentiate $F(z)$ from first principles. Choose any $z$ in $D$ and $\Delta z$ small enough so that $z+\Delta z \in \mathrm{D}$.
So

$$
\begin{aligned}
\frac{F(z+\Delta z)-F(z)}{\Delta z} & =\frac{1}{\Delta z}\left[\int_{z_{0}}^{z+\Delta z} f\left(z^{\prime}\right) d z^{\prime}-\int_{z_{0}}^{z} f\left(z^{\prime}\right) d z^{\prime}\right] \\
& =\frac{1}{\Delta z} \int_{z}^{z+\Delta z} f\left(z^{\prime}\right) d z^{\prime}
\end{aligned}
$$

Now RTP that this ratio less $f(z)$ goes to zero as $\Delta z \rightarrow 0$. Use the fact that

$$
\int_{z}^{z+\Delta z} f(z) d z^{\prime}=f(z) \int_{z}^{z+\Delta z} d z^{\prime}=f(z) \Delta z
$$

So

$$
\frac{F(z+\Delta z)-F(z)}{\Delta z}-f(z)=\frac{1}{\Delta z} \int_{z}^{z+\Delta z}\left(f\left(z^{\prime}\right)-f(z)\right) d z^{\prime}
$$

Now $f(z)$ is analytic and therefore continuous - so for any $\varepsilon>0$ we can find $\delta>0$ such that

$$
\left|f\left(z^{\prime}\right)-f(z)\right|<\varepsilon \quad \text { when } \quad\left|z-z^{\prime}\right|<\delta
$$

So, letting $\Delta z<\delta$ and using the ML-inequality,

$$
\begin{array}{r}
\left|\frac{F(z+\Delta z)-F(z)}{\Delta z}-f(z)\right|=\frac{1}{|\Delta z|}\left|\int_{z}^{z+\Delta z}\left(f\left(z^{\prime}\right)-f(z)\right) d z^{\prime}\right| \\
\leq \frac{1}{|\Delta z|} \varepsilon|\Delta z|=\varepsilon
\end{array}
$$

So $F^{\prime}(z)=f(z)$ as required and (7.13) follows immediately..
3. The most important consequence of Cauchy's Theorem is Cauchy's integral formula - it will allow us to show that analytic functions have integrals of all orders and is useful in evaluating integrals.

## Theorem 7.6 (Cauchy's Integral Formula) Let $f(z)$ be

 analytic in a simply connected domain D . Then for any point $z_{0}$ in D and any contour (simple closed path) C in D we have (taking C counter-clockwise)$$
\begin{equation*}
\oint_{C} \frac{f(z)}{z-z_{0}} d z=2 \pi i f\left(z_{0}\right) \tag{7.14}
\end{equation*}
$$

Proof: The proof is surprisingly simple. Writing $f(z)=f\left(z_{0}\right)+\left[f(z)-f\left(z_{0}\right)\right]$, we have

$$
\oint_{C} \frac{f(z)}{z-z_{0}} d z=f\left(z_{0}\right) \oint_{C} \frac{d z}{z-z_{0}}+\oint_{C} \frac{f(z)-f\left(z_{0}\right)}{z-z_{0}} d z
$$

The first term on RHS is $2 \pi i f\left(z_{0}\right)$ so RTP second term is zero.. The integrand in $T_{2}$ is analytic except at $z=z_{0}$. So by the principle of deformation of path we can replace $C$ by a small circle $K$ of radius $r$ centred at $z_{0}$ without changing the value of the integral. As $f(z)$ is analytic it follows that (as in the proof of Thm 7.5) for any $\varepsilon>0$ we can find $\delta>0$ such that

$$
\left|f(z)-f\left(z_{0}\right)\right|<\varepsilon \quad \text { when } \quad\left|z-z_{0}\right|<\delta
$$

Now if we choose the radius $r$ of $K$ smaller than $\delta$ we have

$$
\left|\frac{f(z)-f\left(z_{0}\right)}{z-z_{0}}\right|<\frac{\varepsilon}{r}
$$

at each point $z$ on the circle $K$. The length of $K$ is $2 \pi r$ so by the ML inequality we have

$$
\left|\oint_{K} \frac{f(z)-f\left(z_{0}\right)}{z-z_{0}} d z\right|<\frac{\varepsilon}{r} 2 \pi r=2 \pi \varepsilon
$$

As $\varepsilon$ can be as small as needed we have that $T_{2}=0$.

Example $7.15 \oint \frac{e^{z}}{z-2} \mathrm{~d} z=2 \pi i e^{2}$ for any contour enclosing $z_{0}=2$.

## Example $\mathbf{7 . 1 6}$

$\oint_{C} \frac{z^{3}-6}{2 z-i} \mathrm{~d} z=\frac{1}{2} \oint_{C} \frac{z^{3}-6}{z-i / 2} \mathrm{~d} z=\left.\frac{1}{2} 2 \pi i\left(z^{3}-6\right)\right|_{z=i / 2}=\pi / 8-6 \pi i$
for any contour C enclosing $z_{0}=\mathfrak{i} / 2$.
Example 7.17 Integrate $\mathrm{g}(\mathrm{z})=\frac{z^{2}+1}{z^{2}-1}$ around a circle with radius 1 centred at each of the points:

$$
(\mathrm{a}) z=1 ; \quad(\mathrm{b}) z=1 / 2, \quad(\mathrm{c}) z=-1+\mathfrak{i} / 2, \quad(\mathrm{~d}) z=\mathrm{i} .
$$

An immediate consequence of the Cauchy Integral Formula (7.14) is that analytic functions have derivatives of all orders. There is no corresponding result for real differentiable functions.
Theorem 7.7 (Derivatives of an analytic function) If $f(z)$ is analytic in a domain D then it has derivatives of all orders in D which are also analytic in D . If C is any contour enclosing $z_{0}$ that is fully contained in D , the values of these derivatives are given by:

$$
\begin{align*}
f^{\prime}\left(z_{0}\right) & =\frac{1}{2 \pi i} \oint_{C} \frac{f(z)}{\left(z-z_{0}\right)^{2}} d z  \tag{7.15a}\\
f^{\prime \prime}\left(z_{0}\right) & =\frac{2!}{2 \pi i} \oint_{C} \frac{f(z)}{\left(z-z_{0}\right)^{3}} d z  \tag{7.15b}\\
\vdots & =\vdots \\
f^{(n)}\left(z_{0}\right) & =\frac{n!}{2 \pi i} \oint_{C} \frac{f(z)}{\left(z-z_{0}\right)^{n+1}} d z . \tag{7.15c}
\end{align*}
$$

Before proving Thm 7.7 it is useful to note that these formulas can be obtained (though this is not a proof) by differentiating the Cauchy Integral Formula (7.14) w.r.t. $z_{0}$.

Proof: We will prove (7.15a) for $f^{\prime}\left(z_{0}\right)$. We differentiate from first principles as before.

$$
f^{\prime}\left(z_{0}\right)=\lim _{\Delta z \rightarrow 0} \frac{f\left(z_{0}+\Delta z\right)-f\left(z_{0}\right)}{\Delta z}
$$

In the RHS we represent $f\left(z_{0}+\Delta z\right)$ and $f\left(z_{0}\right)$ by the Cauchy Integral Formula (7.14). We can combine the two integrals into a single integral.

$$
\begin{aligned}
\frac{f\left(z_{0}+\Delta z\right)-f\left(z_{0}\right)}{\Delta z} & =\frac{1}{2 \pi i \Delta z}\left[\oint \frac{f(z)}{z-\left(z_{0}+\Delta z\right)} d z-\oint \frac{f(z)}{z-z_{0}} d z\right] \\
& =\frac{1}{2 \pi i} \oint \frac{f(z)}{\left(z-z_{0}-\Delta z\right)\left(z-z_{0}\right)} d z
\end{aligned}
$$

Finally, we need to show that as $\Delta z \rightarrow 0$, the above integral converges to (7.15a). Take the difference between the two integrals:

$$
\begin{array}{r}
\frac{1}{2 \pi i \Delta z}\left[\oint \frac{f(z)}{z-\left(z_{0}+\Delta z\right)} d z-\oint \frac{f(z)}{z-z_{0}} d z\right]-\frac{1}{2 \pi i} \oint_{C} \frac{f(z)}{\left(z-z_{0}\right)^{2}} d z \\
=\oint \frac{f(z) \Delta z}{\left(z-z_{0}-\Delta z\right)\left(z-z_{0}\right)^{2}} d z
\end{array}
$$

We can use the ML-inequality to show that this goes to zero as $\Delta z \rightarrow 0$. The only problem is the term in blue in the denominator on the RHS. As usual we have $f(z)$ continuous so that it is bounded on $D$, say $|f(z)| \leq K$. Let $d$ be the smallest distance from $z_{0}$ to the contour $C$. Then, for any point $z$ on $C,\left|z-z_{0}\right|^{2} \geq d^{2}$ and so $\frac{1}{\left|z-z_{0}\right|^{2}} \leq \frac{1}{\mathrm{~d}^{2}}$.

For the term in blue we need to be slightly cleverer. Take $|\Delta z| \leq \mathrm{d} / 2$. Then for all $z$ on C we have (using the alternate form for the triangle inequality)

$$
\left|z-z_{0}-\Delta z\right| \geq\left|z-z_{0}\right|-|\Delta z| \geq \mathrm{d}-\mathrm{d} / 2=\mathrm{d} / 2 \text { so }
$$

$$
\frac{1}{\left|z-z_{0}-\Delta z\right|} \leq \frac{2}{d}
$$

Assembling the pieces gives us:

$$
\left|\oint \frac{f(z) \Delta z}{\left(z-z_{0}-\Delta z\right)\left(z-z_{0}\right)^{2}} d z\right| \leq K|\Delta z| \frac{2}{d} \frac{1}{d^{2}}
$$

The RHS goes to zero as $\Delta z \rightarrow 0$.
The general equation (7.15c) for $f^{(n)}(z)$ follows as we can now use our formula (7.15a) for $f^{\prime}(z)$ instead of the Cauchy Integral Formula to assemble a similar proof for (7.15b) - and so on by induction.

Example 7.18 From (7.15a), for any contour enclosing $z_{0}=\pi \mathfrak{i}$, (taken counter-clockwise)
$\oint_{C} \frac{\cos z}{(z-\pi i)^{2}} \mathrm{~d} z=\left.2 \pi i \cos ^{\prime} z\right|_{z=\pi i}=-2 \pi i \sin \pi i=2 \pi \sinh \pi .$.
Example 7.19 For any contour that contains 1 but excludes $\pm 2 \mathrm{i}$ (taken counter-clockwise)

$$
\oint \frac{e^{z}}{(z-1)^{2}\left(z^{2}+4\right)} \mathrm{d} z=\left.2 \pi i\left(\frac{e^{z}}{z^{2}+4}\right)^{\prime}\right|_{z=1}=\frac{6 e \pi}{25} i .
$$

An immediate consequence of Thm. 7.7 is Cauchy's inequality

$$
\begin{equation*}
\left|f^{(n)}\left(z_{0}\right)\right| \leq \frac{n!M}{r^{n}} \tag{7.16}
\end{equation*}
$$

which follows from (7.15c) taking $C$ to be a circle centre $z_{0}$ and radius $r$ and using the ML inequality. (Check.)

From (7.16) we can immediately conclude
Theorem 7.8 (Liouville's Theorem) If an entire function is bounded in absolute value for all Z , then it is a constant.

Proof: Use (7.16) with $\mathfrak{n}=1$.
So the surprising result is that for a function to be non-trivial (non-constant) and analytic for all finite $z$ it must not be bounded as $|z| \rightarrow \infty$ - i.e. not too well-behaved...

### 7.5 Exercises

1. Some exercises on Cauchy's Theorem :
(a) Show from the definition of the complex integral (7.2) and (7.3) that the integral around a simple closed path (contour) of a $f(z)=$ constant and $f(z)=z$ is zero. (Hint: in the latter case take first $\xi_{i}=z_{i-1}$ then $\xi_{i}=z_{i}$ and calculate half the total of the two sums from $i=1 \ldots n$.)
(b) Show that (as used in the proof of Cauchy's Theorem on Slide 232) the maximum distance between a point on the boundary of a triangle and a point in the interior is half the perimeter. (Hint: use a simple geometrical argument.)
(c) Verify Cauchy's Theorem for $\oint_{C} z^{2} d z$ where $C$ is the boundary of the triangle with vertices $0,2,2 i$.
(d) For what simple closed paths $C$ is $\oint_{C} 1 / z \mathrm{~d} z=0$ ?
(e) For each of the following find the integral round the unit
circle anti-clockwise and say whether Cauchy's Theorem may be used.

$$
\begin{array}{cccc}
1 / z^{4} & e^{-z} & \Im z & \Re z \\
1 /\left(z^{2}+2\right) & 1 / \bar{z} & z^{2} \sec z & 1 /\left(z^{3}+4 z\right)
\end{array}
$$

(f) Evaluate $\oint_{C} \frac{2 z-1}{z^{2}-z} d z$ where $C$ is the ellipse $\frac{\left(x-\frac{1}{2}\right)^{2}}{4}+y^{2}=1$.
(g) Integrate $f(z)=\bar{z} / z$ round the circles $|z|=2$ and $|z|=4$, anti-clockwise. Can the second result be derived from the first using the principle of deformation of path?
(h) Evaluate the following integrals (anticlockwise $=\mathrm{AC}$, clockwise $=\mathrm{C}$ )

$$
\begin{array}{lll}
\oint_{C} \frac{d z}{z} & C:|z-2|=1 & A C \\
\oint_{C} \frac{z^{2}-z+1}{z^{3}-z^{2}} d z & C:|z|=2, \quad|z|=\frac{1}{2} & C \\
\oint_{C} \frac{d z}{z^{2}-1} & C:|z|=2, \quad|z-1|=1 & A C \\
\oint_{C} \frac{d z}{z^{2}+1} & C:|z|=2, \quad|z+i|=1 & A C \\
\oint_{C} \frac{e^{z}}{z} d z & C:|z|=2 & A C \\
\oint_{C} \frac{d z}{z^{4}+4 z^{2}} d z & C:|z|=\frac{3}{2} & A C
\end{array}
$$

2. Some exercises using indefinite integration - evaluate the following integrals:

$$
\begin{array}{llll}
\int_{i}^{2+i} z \mathrm{~d} z & \int_{i}^{1}(z+1)^{2} \mathrm{~d} z & \int_{\pi i}^{2 \pi i} e^{2 z} \mathrm{~d} z & \int_{0}^{i} z^{z^{2}} \mathrm{~d} z \\
\int_{0}^{2 \pi i} \sin 2 z \mathrm{~d} z & \int_{0}^{\pi i} z \cos z^{2} \mathrm{~d} z & \int_{0}^{\pi i} z \cos z \mathrm{~d} z & \int_{-\pi i}^{\pi i} z \cosh z \mathrm{~d} z
\end{array}
$$

3. Some exercises using the Cauchy Integral Formula:
(a) Integrate $z^{2} /\left(z^{2}+1\right)$ anti-clockwise round each of:

$$
|z+\mathfrak{i}|=1 \quad|z-\mathfrak{i}|=\frac{1}{2} \quad|z|=2 \quad|z|=\frac{1}{2}
$$

(b) Integrate $z^{2} /\left(z^{4}-1\right)$ anti-clockwise round each of:

$$
|z-1|=1 \quad|z+i|=1 \quad|z-i|=\frac{1}{2} \quad|z|=2
$$

(c) Evaluate the following integrals anticlockwise round a unit circle.

$$
\begin{array}{llllll}
\frac{1}{z} & \frac{1}{z^{2}+4} & \frac{1}{4 z-i} & \frac{e^{z}}{z} & \frac{e^{2 z}}{z+2 i} & \frac{e^{z^{2}}}{2 z-i} \\
\frac{\cos z}{z} & \frac{\sin z}{z} & \frac{e^{z}-1}{z} & \frac{\sinh z}{z} & \frac{\cosh 3 z}{z} & \frac{\sin z}{z-2}
\end{array}
$$

## 8 The Method of Residues

Before explaining what the term residue means and its significance, we list some results about Taylor Series in $\mathbb{C}$ and related results.

Definition 8.1 A power series is a sum of the form

$$
\begin{equation*}
\sum_{n=0}^{\infty} a_{n}\left(z-z_{0}\right)^{n} \tag{8.1}
\end{equation*}
$$

Here $z_{0}$ is called the centre. A power series converges in general for $\left|z-z_{0}\right|<R$, the radius of convergence. It can be shown that $R=\lim _{n \rightarrow \infty}\left|a_{n} / a_{n+1}\right|$ if this limit exists. If $\mathrm{R}>0$, the series is an analytic function for $\left|z-z_{0}\right|<R$. The derivatives $f^{\prime}(z), f^{\prime \prime}(z), \ldots$ can be obtained by differentiating (8.1) term-by-term and have the same radius of convergence.

Every analytic function $f(z)$ may be expressed as a power series called a Taylor series - within its radius of convergence.

Definition 8.2 A Taylor series for an analytic function $\mathrm{f}(\mathrm{z})$ is a power series of the form

$$
\begin{equation*}
f(z)=\sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}\left(z_{0}\right)\left(z-z_{0}\right)^{n}, \quad\left|z-z_{0}\right|<R \tag{8.2}
\end{equation*}
$$

Taylor series converge for all $\boldsymbol{z}$ if $\mathrm{f}(\boldsymbol{z})$ is entire or in the open disk with centre $z_{0}$ and radius equal to the distance from $z_{0}$ to the nearest singularity (point where $\mathrm{f}(\mathrm{z})$ ceases to be analytic).

The familiar functions $e^{z}, \cos z, \sinh z$ etc. all have Taylor Series identical to those found for their real-valued versions.

Definition 8.3 $A$ Laurent series is of the form

$$
\begin{equation*}
f(z)=\sum_{n=-\infty}^{\infty} a_{n}\left(z-z_{0}\right)^{n}, \quad a_{n}=\frac{1}{2 \pi i} \oint_{C} \frac{f\left(z^{\prime}\right)}{\left(z^{\prime}-z_{0}\right)^{n+1}} d z^{\prime} \tag{8.3}
\end{equation*}
$$

This series converges in a ring or annulus $\mathcal{A}$ with centre $z_{0}$. In the ring the function $\mathrm{f}(\mathrm{z})$ is analytic. The sum from $\mathrm{n}=0$ to $\infty$ is a power series in $z-z_{0}$. The second series (in negative powers of $\left.z-z_{0}\right)$ is called the principal part of the Laurent series. In a given annulus, a Laurent series is unique but $f(z)$ may have different Laurent series in different annuli with the same centre.
Example 8.1 Find all Laurent series of $\frac{1}{z^{3}-z^{4}}$ with centre 0.
(i) $\frac{1}{z^{3}-z^{4}}=\frac{1}{z^{3}} \frac{1}{1-z}=\sum_{n=0}^{\infty} z^{n-3}=\frac{1}{z^{3}}+\frac{1}{z^{2}}+\frac{1}{z}+1+z+\ldots$ for $0<|z|<1$
(ii) $\frac{1}{z^{3}-z^{4}}=\frac{1}{z^{4}} \frac{-1}{1-1 / z}=-\sum_{n=0}^{\infty} \frac{1}{z^{n+4}}=-\frac{1}{z^{4}}-\frac{1}{z^{5}}-\frac{1}{z^{6}} \ldots$ for $|z|>1$.

Example 8.2 Find the Laurent series for $f(z)=1 /\left(1-z^{2}\right)$ that converges in the annulus $1 / 4<|z-1|<1 / 2$ and determine the precise region of convergence. The annulus has centre 1 so we must expand

$$
f(z)=\frac{-1}{(z-1)(z+1)}
$$

as a sum of (both negative and positive) powers of $z-1$. We have

$$
\begin{aligned}
\frac{1}{z+1} & =\frac{1}{2+(z-1)}=\frac{1}{2} \frac{1}{\left(1-\left(-\frac{z-1}{2}\right)\right)} \\
& =\frac{1}{2} \sum_{n=0}^{\infty}\left(-\frac{z-1}{2}\right)^{n}=\sum_{n=0}^{\infty} \frac{(-1)^{n}}{2^{n+1}}(z-1)^{n}
\end{aligned}
$$

which converges in the disk $|(z-1) / 2|<1$ or just $|z-1|<2$.
Multiplying by $-1 /(z-1)$ gives us the series (singular at $z=-1$ ) $f(z)=\sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{2^{n+1}}(z-1)^{n-1}$ with $0<|z-1|<2$.

Definition 8.4 If $f(z)$ has an isolated singularity at $z=z_{0}$ (for small enough disks centred at $z_{0}, z_{0}$ is the only singular point) the Laurent series of $f(z)$ that converges on $0<\left|z-z_{0}\right|<R$ can be used to classify this singularity as either

- a pole; if the principal part of the Laurent series has only a finite number of terms
- otherwise an essential singularity.

Example 8.3 The function $e^{\frac{1}{z}}$ has Laurent series

$$
\sum_{0}^{\infty} \frac{\left(\frac{1}{z}\right)^{n}}{n!}=\sum_{0}^{\infty} \frac{1}{z^{n} n!}-\text { an essential singularity at } z=0
$$

and converges in the region $|z|>0$.
Definition 8.5 A pole is of order $\mathfrak{n}$ when the largest negative power of $z-z_{0}$ in the principal part is $\frac{1}{\left(z-z_{0}\right)^{n}}$. A first-order pole is called a simple pole.

### 8.1 Residues

We know from Cauchy's Theorem that if a function is analytic inside and on a contour $C$ then $\oint_{C} f(z) d z=0$. What if $f$ has a singular point $z_{0}$ inside (but not on) $C$ ?

Then $f(z)$ has a Laurent series

$$
f(z)=\sum_{n=0}^{\infty} a_{n}\left(z-z_{0}\right)^{n}+\frac{b_{1}}{z-z_{0}}+\frac{b_{2}}{\left(z-z_{0}\right)^{2}}+\ldots
$$

that converges for all $z$ near $z_{0}$ (except at $z=z_{0}$ itself) in a domain of form $0<\left|z-z_{0}\right|<R$. The coefficient $b_{1}$ is given by (8.3) as

$$
b_{1}=\frac{1}{2 \pi i} \oint_{C} f(z) d z
$$

The key idea is to evaluate $b_{1}$ by some other method and so evaluate the contour integral indirectly.

Definition 8.6 The coefficient $\mathrm{b}_{1}$ is called the residue of $\mathrm{f}(z)$ at $z=z_{0}$ and we will write it as

$$
\begin{equation*}
\mathrm{b}_{1}=\operatorname{Res}_{z=z_{0}} f(z) \tag{8.4}
\end{equation*}
$$

## Example 8.4 (Evaluating an integral using a residue)

Integrate the function $f(z)=z^{-4} \sin z$ around the unit circle $C$ (anti-clockwise). As we know the Taylor series for $\sin$ it is easy to write the Laurent series for $\mathrm{f}(\mathrm{z})$ as

$$
f(z)=\frac{\sin z}{z^{4}}=\frac{1}{z^{3}}-\frac{1}{3!z}+\frac{z}{5!}-\frac{z^{3}}{7!} \ldots
$$

that converges for all $|z|>0$. The residue $\mathrm{b}_{1}$ can be read off as $\mathrm{b}_{1}=-\frac{1}{3!}$. So

$$
\oint_{C} \frac{\sin z}{z^{4}}=2 \pi i b_{1}=-\frac{\pi i}{3}
$$

## Example 8.5 (Use the right Laurent Series !) Integrate

 $f(z)=1 /\left(z^{3}-z^{4}\right)$ around the circle $C:|z|=\frac{1}{2}$ clockwise. Now $f(z)$ is singular at $z=0$ and $z=1$. The latter lies outside C so is irrelevant. We need the residue of $\mathrm{f}(\mathrm{z})$ at $z=0$. We find it from the Laurent Series :$$
\frac{1}{z^{3}-z^{4}}=\frac{1}{z^{3}} \frac{1}{1-z}=\frac{1}{z^{3}}+\frac{1}{z^{2}}+\frac{1}{z}+1+z+\ldots
$$

Obviously $\mathrm{b}_{1}=1$ so
$\oint_{C} \frac{1}{z^{3}-z^{4}}=-2 \pi i \quad$ (the minus sign as we are integrating clockwise)
Note that if we used the "wrong" Laurent Series - the second case in Example 8.1 -
$\frac{1}{z^{3}-z^{4}}=\frac{1}{z^{4}} \frac{-1}{1-1 / z}=-\sum_{n=0}^{\infty} \frac{1}{z^{n+4}}=-\frac{1}{z^{4}}-\frac{1}{z^{5}}-\frac{1}{z^{6}} \ldots$ for $|z|>1$ gives $\mathrm{b}_{1}=0$. Is there a contradiction here?

Next, an obvious question: to calculate the residue, do we really need to find the whole series or can we short-circuit the process?
Provided the singularity is a pole, we can.
Let $f(z)$ have a simple pole at $z=z_{0}$. Then the corresponding
Laurent Series is

$$
f(z)=\frac{b_{1}}{z-z_{0}}+a_{0}+a_{1}\left(z-z_{0}\right)+\ldots, \quad 0<\left|z-z_{0}\right|<R
$$

Multiplying both sides by $z-z_{0}$ and taking the limit as $z \rightarrow z_{0}$ we have

$$
\begin{equation*}
\operatorname{Res}_{z=z_{0}} f(z)=b_{1}=\lim _{z \rightarrow z_{0}}\left(z-z_{0}\right) f(z) \tag{8.5}
\end{equation*}
$$

## Example 8.6 (Residue at a simple pole)

$$
\operatorname{Res}_{z=i} \frac{9 z+i}{z\left(z^{2}+1\right)}=\lim _{z \rightarrow i}(z-i) \frac{9 z+i}{z\left(z^{2}+1\right)}=\left(\frac{9 z+i}{z(z+i)}\right)_{z=i}=\frac{10 i}{-2}=-5 i
$$

An alternative method for finding the residue for a simple pole is based on the fact that as $f(z)=\frac{p(z)}{q(z)}$ and $q(z)$ must have a factor of $z-z_{0}$ we must have

$$
\mathrm{q}(z)=\left(z-z_{0}\right) \mathrm{q}^{\prime}\left(z_{0}\right)+\frac{1}{2}\left(z-z_{0}\right)^{2} \mathrm{q}^{\prime \prime}\left(z_{0}\right)+\ldots
$$

so using (8.5) we have, cancelling the factor $\left(z-z_{0}\right)$ and taking the limit as $z \rightarrow z_{0}$ (check)

$$
\begin{equation*}
\operatorname{Res}_{z=z_{0}} f(z)=\operatorname{Res}_{z=z_{0}} \frac{p(z)}{q(z)}=\frac{p\left(z_{0}\right)}{q^{\prime}\left(z_{0}\right)} . \tag{8.6}
\end{equation*}
$$

Example 8.7 (Residue at a simple pole using (8.6))
$\operatorname{Res}_{z=i} \frac{9 z+i}{z\left(z^{2}+1\right)}=\left(\frac{9 z+i}{3 z^{2}+1}\right)_{z=i}=\frac{10 i}{-2}=-5 i$.

Example 8.8 Find all poles and the corresponding residues of

$$
f(z)=\frac{\cosh \pi z}{z^{4}-1} .
$$

The numerator is entire and $z^{4}-1$ has zeroes at $1, \mathfrak{i},-1,-\mathfrak{i}$. So these are the (simple) poles of $f(z)$. Now $\mathrm{q}^{\prime}(z)=4 z^{3}$ so the residues are just the values of $\cosh z / 4 z^{3}$ at these points:

$$
\frac{\cosh \pi}{4} ; \quad \frac{\cosh \pi i}{4 \mathfrak{i}^{3}}=\frac{\cos \pi}{-4 i}=-\frac{i}{4} ; \quad \frac{-\cosh \pi}{4} ; \quad \frac{\cosh (-\pi)}{4(-i)^{3}}=\frac{i}{4}
$$

Poles of Higher Order If $f(z)$ has a pole of order $m>1$ at $z=a$ then its Laurent Series is of the form:

$$
\begin{aligned}
& f(z)= \frac{c_{m}}{(z-a)^{m}}+ \\
& \frac{c_{m-1}}{(z-a)^{m-1}}+\cdots+ \\
& \frac{c_{2}}{(z-a)^{2}}+\frac{c_{1}}{(z-a)}+b_{0}+b_{1}(z-a)+\ldots
\end{aligned}
$$

Multiplying both sides by $(z-a)^{m}$ gives

$$
\begin{array}{r}
(z-a)^{m} f(z)=c_{m}+c_{m-1}(z-a)+\cdots+c_{2}(z-a)^{m-2}+c_{1}(z-a)^{m-1}+ \\
b_{0}(z-a)^{m}+b_{1}(z-a)^{m+1}+\ldots
\end{array}
$$

Examining this Taylor Series it is clear that if we differentiate $(z-a)^{m} f(z) m$ times and evaluate the result at $z=a$ the only term that survives is $c_{1}$, multiplied by a factor $(m-1)$ !.

So we have a general result

$$
\begin{equation*}
\operatorname{Res}_{z=a} f(z)=\frac{1}{(m-1)!} \lim _{z \rightarrow a}\left(\frac{d^{m-1}}{d z^{m-1}}\left[(z-a)^{m} f(z)\right]\right) \tag{8.7}
\end{equation*}
$$

Example 8.9 (Residue at a pole of higher order) The function $\mathrm{f}(\mathrm{z})=\frac{2 z}{(z+4)(z-1)^{2}}$ has a pole of second order at $z=1$. Applying (8.7) we find:

$$
\operatorname{Res}_{z=1} f(z)=\frac{1}{1!} \lim _{z \rightarrow 1} \frac{d}{d z} \frac{2 z}{z+4}=\left.\frac{8}{(z+4)^{2}}\right|_{z=1}=\frac{8}{25}
$$

Example 8.10 Evaluate the following integrals where $C$ is the unit circle anticlockwise:

$$
\begin{array}{ccc}
\oint_{C} e^{1 / z} d z & \oint_{C} z e^{1 / z} d z & \oint_{C} \cot z d z \\
\oint_{C} \frac{1}{\cosh z} d z & \oint_{C} \frac{\sin \pi z}{z^{4}} d z & \oint_{C} \frac{\left(z^{2}+1\right)}{e^{z} \sinh z} d z
\end{array}
$$

### 8.2 The Residue Theorem

With the techniques above we can evaluate contour integrals whose integrands have a single pole inside the contour. This can be easily extended to the general case where the integrand has several poles inside the contour.

Theorem 8.1 (Residue Theorem) Let $\mathrm{f}(\boldsymbol{z})$ be a function that is analytic inside a contour C and on C except at a finite number of singular points $\mathrm{a}_{1}, \ldots, \mathrm{a}_{\mathrm{m}}$ inside C . Then (taking the integral round C anti-clockwise)

$$
\begin{equation*}
\oint_{C} f(z) d z=2 \pi i \sum_{j=1}^{m} \operatorname{Res}_{z=a_{j}} f(z) \tag{8.8}
\end{equation*}
$$

## Proof:

Enclose each of the poles $a_{j}$ in a circle $C_{j}$ small enough so that none of these circles nor $C$ intersect. (Draw a sketch.)

Then $f(z)$ is analytic in the (multiply connected) domain $D$ bounded by $C$ and the $m$ small circles and on the boundary of $D$. From Cauchy's Theorem (Thm 7.4)

$$
\oint_{C} f(z) d z+\oint_{C_{1}} f(z) d z+\oint_{C_{2}} f(z) d z+\cdots+\oint_{C_{m}} f(z) d z=0
$$

where the integral round $C$ is taken anticlockwise and the integrals round the small circles clockwise. Now reverse the direction of integration round the small circles (which has the effect of flipping their signs) giving:

$$
\oint_{C} f(z) d z=\oint_{C_{1}} f(z) d z+\oint_{C_{2}} f(z) d z+\cdots+\oint_{C_{m}} f(z) d z
$$

But as the integral round each of the small circles is $2 \pi i \underset{z=\mathfrak{a}_{j}}{\operatorname{Res}} f(z)$, the result follows.

Example 8.11 The function $\frac{4-3 z}{z^{2}-z}$ is analytic except at the points 0 and 1 where it has simple poles. The residues are -4 and 1 respectively. So

$$
\oint_{C} \frac{4-3 z}{z^{2}-z} d z=2 \pi i(-4+1)=-6 \pi i
$$

for every contour $C$ that encloses the points 0 and 1.
Example 8.12 Integrate $\frac{1}{\left(z^{3}-1\right)^{2}}$ anti-clockwise round the circle $|z-1|=1$. The function has poles of second order at 1 , $e^{2 \pi i / 3}$ and $e^{-2 \pi i / 3}$. Only the pole at $z=1$ lies inside C. Using (8.7) we have

$$
\oint_{C} \frac{d z}{\left(z^{3}-1\right)^{2}}=2 \pi i \operatorname{Res}_{z=1} \frac{1}{\left(z^{3}-1\right)^{2}}=2 \pi i\left(-\frac{2}{9}\right)=-\frac{4 \pi i}{9}
$$

### 8.3 Evaluating Real Integrals

Many difficult real integrals can be quite easily evaluated using the Residue Theorem. In this short section we will just consider two types of real integrals:

## 1. Rational functions of $\cos \theta, \sin \theta$

$$
I=\int_{0}^{2 \pi} F(\cos \theta, \sin \theta) d \theta
$$

where $F(\cos \theta, \sin \theta)$ is a real rational function of $\cos \theta, \sin \theta$. The first step is to note that with $z=e^{i \theta}$, we have $\cos \theta=\frac{1}{2}\left(z+\frac{1}{z}\right)$ and $\sin \theta=\frac{1}{2 i}\left(z-\frac{1}{z}\right)$. So $F(\cos \theta, \sin \theta)$ is a rational function of $z$, say $f(z)$. As $\theta$ ranges from 0 to $2 \pi$, the variable $z$ ranges once round the unit circle $|z|=1$ anti-clockwise.

Since $\frac{d z}{d \theta}=i e^{i \theta}$ we have $d \theta=\frac{1}{i z} d z$ and

$$
I=\oint_{C} \frac{f(z)}{i z} d z
$$

Example 8.13 Let $\mathrm{I}=\int_{0}^{2 \pi} \frac{\mathrm{~d} \theta}{\sqrt{2}-\cos \theta}$. Use $\cos \theta=\frac{1}{2}\left(z+\frac{1}{z}\right)$ and $\mathrm{d} \theta=\frac{\mathrm{dz}}{\mathrm{iz}}$. Then we have

$$
\begin{aligned}
\mathrm{I} & =\oint_{C} \frac{\mathrm{~d} z / \mathrm{i} z}{\sqrt{2}-\frac{1}{2}\left(z+\frac{1}{z}\right)} \\
& =\oint_{C} \frac{d z}{-\frac{i}{2}\left(z^{2}-2 \sqrt{2} z+1\right)} \\
& =-\frac{2}{i} \oint_{C} \frac{d z}{(z-\sqrt{2}-1)(z-\sqrt{2}+1)}
\end{aligned}
$$

The integrand has two simple poles at $z_{1}=\sqrt{2}+1$ (outside C) and $z_{2}=\sqrt{2}-1$ (inside C).

At $z_{2}$, the residue is just $-\left.\frac{2}{i}\left(\frac{1}{z-\sqrt{2}-1}\right)\right|_{z=\sqrt{2}-1}=-i$ so the solution is $\mathrm{I}=2 \pi \mathrm{i}(-\mathfrak{i})=2 \pi$.

Example 8.14 Evaluate the following integrals:

$$
\int_{0}^{2 \pi} \frac{d \theta}{2+\cos \theta} \quad \int_{0}^{2 \pi} \frac{2 \sin ^{2} \theta}{5-4 \cos \theta} d \theta \quad \int_{0}^{2 \pi} \frac{1+\sin \theta}{3+\cos \theta} d \theta .
$$

2. Infinite (improper) integrals of rational functions

$$
I=\int_{-\infty}^{\infty} f(x) d x
$$

where $f(x)$ is rational and the degree of the denominator is at least two greater than that of the numerator. Consider the corresponding contour integral

$$
\oint_{C} f(z) d z
$$

where $C$ is the line segment $-R \leq x \leq R$ followed by the semi-circle $S: z=R e^{i \theta}$ for $0 \leq \theta \leq \pi$.

As $f(z)$ is rational. it has a finite number of poles in the upper half plane so by taking $R$ large enough we can be sure that all these poles are inside C. By the Residue Theorem we have

$$
\oint_{C} f(z) d z=\int_{S} f(z) d z+\int_{-R}^{R} f(x) d x=2 \pi i \sum \operatorname{Res} f(x)
$$

where the sum is over all the poles of $f(z)$ in the upper half plane. Re-writing, we have

$$
\int_{R}^{R} f(x) d x=2 \pi i \sum \operatorname{Res} f(x)-\int_{S} f(z) d z
$$

We need to show that as $R \rightarrow \infty$, the integral over the semi-circle $S$ goes to zero.

As the degree of the denominator is at least two greater than that of the numerator we have on the semi-circle $S$ that

$$
|f(z)| \leq \frac{k}{|z|^{2}}
$$

for R sufficiently large. Therefore by the ML-inequality,

$$
\left|\int_{S} f(z) d z\right|<\frac{k}{|z|^{2}} \pi R=\frac{k \pi}{R}
$$

So in the limit as $R \rightarrow \infty$ we have

$$
\int_{-\infty}^{\infty} f(x) d x=2 \pi i \sum \operatorname{Res} f(z)
$$

where the sum is over all the poles of $f(z)$ in the upper half plane.

Example 8.15 Let $\mathrm{I}=\int_{0}^{\infty} \frac{\mathrm{dx}}{1+\mathrm{x}^{4}}$. First note that we can write $\mathrm{I}=\frac{1}{2} \int_{-\infty}^{\infty} \frac{\mathrm{dx}}{1+x^{4}}$ as $\mathrm{f}(\mathrm{x})$ is even. The complex function $f(z)=\frac{1}{1+z^{4}}$ has four simple poles at $z_{1}=e^{i \pi / 4}, z_{2}=e^{i 3 \pi / 4}$, $z_{3}=e^{-i \pi / 4}$ and $z_{4}=e^{-i 3 \pi / 4}$. The first two lie in the upper half plane and have residues as follows:

$$
\begin{aligned}
\operatorname{Res}_{z=z_{1}} f(z) d z=\left.\left(\frac{1}{\left(1+z^{4}\right)^{\prime}}\right)\right|_{z=z_{1}} & =\left.\left(\frac{1}{4 z^{3}}\right)\right|_{z=z_{1}} \\
& =\frac{1}{4} e^{-3 \pi i / 4}=-\frac{1}{4} e^{i \pi / 4}
\end{aligned}
$$

and

$$
\begin{aligned}
\operatorname{Res}_{z=z_{2}} f(z) d z=\left.\left(\frac{1}{\left(1+z^{4}\right)^{\prime}}\right)\right|_{z=z_{2}} & =\left.\left(\frac{1}{4 z^{3}}\right)\right|_{z=z_{2}} \\
& =\frac{1}{4} e^{-i 9 \pi / 4}=-\frac{1}{4} e^{-i \pi / 4}
\end{aligned}
$$

So we have
$\mathrm{I}=\frac{1}{2} \int_{-\infty}^{\infty} \frac{\mathrm{dx}}{1+x^{4}}=\frac{1}{2} \frac{2 \pi i}{4}\left(-e^{i \pi / 4}+e^{-i \pi / 4}\right)=\frac{1}{2} \pi \sin \pi / 4=\frac{\pi}{2 \sqrt{2}}$.
Note that a complex answer indicates an error in the algebra as the integral of a real-valued function on the real line must give a real result.

Example 8.16 Evaluate the following integrals:

$$
\begin{array}{ccc}
\int_{-\infty}^{\infty} \frac{d x}{1+x^{2}} & \int_{-\infty}^{\infty} \frac{d x}{\left(1+x^{2}\right)^{2}} & \int_{-\infty}^{\infty} \frac{x^{2}}{\left(1+x^{2}\right)\left(x^{2}+4\right)} d x \\
\int_{0}^{\infty} \frac{1+x^{2}}{1+x^{4}} d x & \int_{0}^{\infty} \frac{d x}{\left(x^{2}+1\right)\left(x^{2}+9\right)} & \int_{-\infty}^{\infty} \frac{x^{3}}{1+x^{8}} d x
\end{array}
$$

### 8.4 Inverting the Laplace Transform

In this section, we will see that the Residue Theorem can be used to evaluate Inverse Laplace Transforms. First, a reminder:

Definition 8.7 (Laplace Transform) Let $\mathrm{f}(\mathrm{t})$ be a real or complex valued function defined for all $\mathrm{t}>0$ and let $\mathrm{s}=\sigma+\mathrm{i} \omega$ be a complex variable. Then the Laplace transform of $\mathrm{f}(\mathrm{t})$, written $\mathrm{F}(\mathrm{s})$ is

$$
\begin{equation*}
F(s)=\int_{0}^{\infty} f(t) e^{-s t} d t . \tag{8.9}
\end{equation*}
$$

This operation is often written $F(s)=\mathcal{L} f(t)$. The function $f(t)$ whose Laplace transform is $F(s)$ is written $f(t)=\mathcal{L}^{-1} F(s)$. We say $f(t)$ is the inverse Laplace transform of $F(s)$.

### 8.4.1 Deriving the Inverse Laplace transform

First we remind ourselves that

$$
\begin{equation*}
\mathcal{L} e^{-a t}=\frac{1}{s+a}, \quad \text { if } \mathfrak{R s}>-\mathfrak{R a} . \tag{8.10}
\end{equation*}
$$

To derive a formula for $\mathcal{L}^{-1} \mathrm{~F}(\mathrm{~s})$, we use two steps:

1. we use the Cauchy Integral Formula (7.14) to express $F(z)$ in terms of an integral of $F(s)$ times a factor $\frac{1}{z-s}$,
2. we will then apply the inverse Laplace transform to this integral - the $z$-dependence is all in the factor $\frac{1}{z-s}$ whose inverse Laplace transform we know from (8.10).

We will develop a formula for $\mathcal{L}^{-1} \mathrm{~F}(\mathrm{~s})$ by proving a Theorem.

Theorem 8.2 (Bromwich Integral Formula) Suppose that
$\mathrm{F}(z) \equiv \int_{0}^{\infty} \mathrm{f}(\mathrm{t}) \mathrm{e}^{-z \mathrm{t}} \mathrm{dt}$ is analytic in $\mathbb{C}$ everywhere along the line $x=\mathrm{a}$ and to the right of this line (so all the singularities are to the left of $\mathrm{x}=\mathrm{a}$ ). We also assume that $\mathrm{F}(z) \rightarrow 0$ as $|z| \rightarrow \infty$ along any path in the half-plane $\mathfrak{R z} \geq \mathrm{a}$; more precisely we assume that for some $\mathrm{m}, \mathrm{k}$ and $\mathrm{R}_{0}$ (all positive) we have when $|z|>\mathrm{R}_{0}$ and $\Re z \geq \mathrm{a}$

$$
\begin{equation*}
|F(z)| \leq \frac{m}{|z|^{k}} \tag{8.11}
\end{equation*}
$$

Then for any $\mathrm{t}>0$

$$
\begin{equation*}
f(t) \equiv \mathcal{L}^{-1} F(s)=\frac{1}{2 \pi i} \int_{a-i \infty}^{a+i \infty} F(z) e^{z t} d z \tag{8.12}
\end{equation*}
$$

Proof: Begin by applying the Cauchy Integral Formula (7.14) to $F(z)$. We use the closed contour $C$ consisting of a semi-circle $C_{1}$ of radius $b$ centred at $z=a$ followed by the line segment $a+i b \rightarrow a-i b$. Our choice of $a$ ensures that there are no singularities in or on C .


Figure 5: Bromwich contour

Take $s$ to be any point inside the contour $C$. Then from the Cauchy Integral Formula (7.14):

$$
\begin{equation*}
F(s)=\frac{1}{2 \pi i} \oint \frac{F(z)}{z-s} d z=\frac{1}{2 \pi i}\left[\int_{a+i b}^{a-i b} \frac{F(z)}{z-s} d z+\int_{C_{1}} \frac{F(z)}{z-s} d z\right] \tag{8.13}
\end{equation*}
$$

RTP that the integral over the semi-circle $C_{1}$ goes to zero as $\mathbf{b} \rightarrow \infty$. We have from (8.11) that

$$
\begin{equation*}
|F(z)| \leq \frac{m}{b^{k}}, \quad \text { for } z \text { on } C_{1} \tag{8.14}
\end{equation*}
$$

as on $\mathrm{C}_{1},|z| \geq \mathrm{b}$ so $\frac{1}{|z|} \leq \frac{1}{\mathrm{~b}}$.
Now examine $|z-s|$ on $C_{1}$. It is easy to check that the minimum value for $|z-s|\left(z\right.$ on $C_{1}$ and $s$ in $\left.C\right)$ occurs when $z$ lies on the radius of the semi-circle that contains $s$. The minimum value is just $b-|s-a|$. So, for all $z$ on $C_{1}$, we have

$$
\begin{equation*}
|z-s| \geq b-|s-a| \tag{8.15}
\end{equation*}
$$

Now, using the Triangle Inequality $|s-a| \leq|s|+a$ together with (8.15) we find that

$$
\begin{equation*}
|z-s| \geq b-(|s|+a) \tag{8.16}
\end{equation*}
$$

We can ensure that the RHS in (8.16) is positive by taking $b$ sufficiently large. Now taking the reciprocal of each side in (8.16) gives

$$
\begin{equation*}
\frac{1}{|z-s|} \leq \frac{1}{b-(|s|+a)} \tag{8.17}
\end{equation*}
$$

and finally, multiplying by $F(s)$ and using (8.14) we have

$$
\begin{equation*}
\frac{F(s)}{|z-s|} \leq \frac{m}{b^{k}(b-|s|-a)} \tag{8.18}
\end{equation*}
$$

Now apply the ML-inequality to the integral over $C_{1}$ (the path length $L$ is $\pi b)$ :

$$
\begin{equation*}
\left|\oint_{C_{1}} \frac{F(z)}{z-s} d z\right| \leq \frac{m}{b^{k}(b-|s|-a)} \pi b \tag{8.19}
\end{equation*}
$$

Clearly, for any $k>0$, the RHS in (8.20) goes to zero as $b \rightarrow \infty$. So we have

$$
\begin{equation*}
F(s)=\frac{1}{2 \pi i} \int_{a+i \infty}^{a-i \infty} \frac{F(z)}{z-s} d z=\frac{1}{2 \pi i} \int_{a-i \infty}^{a+i \infty} \frac{F(z)}{s-z} d z \tag{8.20}
\end{equation*}
$$

Now, finally, consider the Inverse Laplace transform of (8.20). We know from (8.10) that

$$
\mathcal{L} e^{z t}=\frac{1}{s-z}, \quad \text { if } \mathfrak{R s}>\mathfrak{R z}
$$

The inverse of this result is that

$$
\begin{equation*}
\mathcal{L}^{-1} \frac{1}{s-z}=e^{z \mathrm{t}}, \quad \Re s>\Re z \tag{8.21}
\end{equation*}
$$

Now, ignoring any possible mathematical difficulties, apply the inverse operator $\mathcal{L}^{-1}$ to both sides of (8.20).

We finally have that

$$
\begin{align*}
\mathcal{L}^{-1} F(s) & =\frac{1}{2 \pi i} \int_{a-i \infty}^{a+i \infty} \mathcal{L}^{-1}\left[\frac{F(z)}{s-z}\right] d z  \tag{8.22}\\
& =\frac{1}{2 \pi i} \int_{a-i \infty}^{a+i \infty} F(z) \mathcal{L}^{-1}\left[\frac{1}{s-z}\right] d z  \tag{8.23}\\
& =\frac{1}{2 \pi i} \int_{a-i \infty}^{a+i \infty} F(z) e^{z t} d z . \tag{8.24}
\end{align*}
$$

So

$$
f(t) \equiv \mathcal{L}^{-1} F(s)=\frac{1}{2 \pi i} \int_{a-i \infty}^{a+i \infty} F(z) e^{z t} d z .
$$

The constraint $\mathfrak{R s}>\mathfrak{R z}$ is satisfied as $s$ lies to the right of the vertical path from $\mathfrak{a}-i \infty$ to $a+i \infty$ along which $z$ varies.

### 8.4.2 Examples of Inverse Laplace transform

Example 8.17 Take the simple case $\mathrm{F}(\mathrm{s})=\frac{1}{(\mathrm{~s}+1)^{2}}$. As $\mathrm{F}(\mathrm{s})$ has a double pole at $\mathrm{s}=-1$, we need $\mathrm{a}>-1-$ we can take $\mathrm{a}=0$. So

$$
f(t)=\frac{1}{2 \pi i} \int_{-i \infty}^{+i \infty} \frac{1}{(s+1)^{2}} e^{s t} d s
$$

Now, to evaluate this integral - using contour integral methods we "close the contour" with a semi-circular arc to the left of the vertical path from $-\mathfrak{i} \infty$ to $+\mathfrak{i} \infty$. See Fig. 6 on the next slide. We will use the Residue Theorem to evaluate the integral around the closed path (contour) formed.
We expect - and will check - that the integral around the semicircle will go to zero as its radius $\mathrm{R} \rightarrow \infty$.


Figure 6: Contour for Example 8.17
So consider

$$
\frac{1}{2 \pi i} \oint \frac{e^{s t}}{(s+1)^{2}} \mathrm{~d} s
$$

We have

$$
\begin{aligned}
\frac{1}{2 \pi i} \oint \frac{e^{s t}}{(s+1)^{2}} d s & =f(t)+\frac{1}{2 \pi i} \int_{C_{1}} \frac{e^{s t}}{(s+1)^{2}} d s \\
& =\frac{1}{2 \pi i} 2 \pi i \operatorname{Res}_{s=-1} \frac{e^{s t}}{(s+1)^{2}} \\
& =t e^{-t}
\end{aligned}
$$

Let's evaluate the integral round the semi-circle $\mathrm{C}_{1}$.

$$
\int_{C_{1}} \frac{e^{s t}}{(s+1)^{2}} d s=\int_{\theta=\pi / 2}^{\theta=3 \pi / 2} \frac{e^{R e^{i \theta} t}}{\left(\operatorname{Re}^{i \theta}+1\right)^{2}} i R e^{i \theta} d \theta
$$

Now $\left|e^{R e^{i \theta}} \mathrm{t}\right|=\left|e^{R \mathrm{t} \cos \theta} e^{i R \mathrm{t} \sin \theta}\right|=e^{R \mathrm{t} \cos \theta}$ and as $\pi / 2 \leq \theta \leq 3 \pi / 2$ we have $\cos \theta \leq 0$ so as $R, t>0$ it follows that $\left|e^{R e^{i \theta}}\right| \leq 1$.

The magnitude of the $B L$ is given by $|\mathrm{BL}|=\left|\left(\operatorname{Re}^{i \theta}+1\right)^{2}\right|$. Now using $|z|^{2}=\left|z^{2}\right|$ we have

$$
\begin{aligned}
|B L| & =|(R \cos \theta+1)+i R \sin \theta|^{2} \\
& =(R \cos \theta+1)^{2}+R^{2} \sin ^{2} \theta \\
& =R^{2}+2 R \cos \theta \\
& \geq R^{2}-2 R .
\end{aligned}
$$

Using the M-L inequality;

$$
\left|\int_{C_{1}} \frac{e^{s t}}{(s+1)^{2}} d s\right| \leq \frac{\pi R}{R^{2}-2 R}=\frac{\pi}{R-2}
$$

which goes to 0 as $\mathrm{R} \rightarrow \infty$.
So $\mathrm{f}(\mathrm{t})=\mathrm{t} \mathrm{e}^{-\mathrm{t}}$.

The calculations needed to show that $\int_{C_{1}} F(s) e^{s t} d s \rightarrow 0$ as $R \rightarrow \infty$ can be generalised. We will just state the result as a Theorem without proof.

Theorem 8.3 Let $\mathrm{F}(\mathrm{s})$ be analytic in the s-plane except for a finite number of poles to the left of some vertical line $\mathfrak{R}=\mathrm{a}$. Suppose that there are positive constants $\mathrm{m}, \mathrm{R}_{0}$ and k such that for all s in the half-plane $\mathfrak{R} \leq \mathbf{a}$ which satisfy $|\mathrm{s}|>\mathrm{R}_{0}$ we have $|\mathrm{F}(\mathrm{s})| \leq \mathrm{m} /|\mathrm{s}|^{\mathrm{k}}$. Then for $\mathrm{t}>0$.

$$
\mathcal{L}^{-1} \mathrm{~F}(s)=\sum \operatorname{Res}\left(\mathrm{F}(\mathrm{~s}) e^{s t}\right)
$$

where the sum is over the poles of $\mathrm{F}(\mathrm{s})$.

Example 8.18 Find $\mathcal{L}^{-1} \frac{1}{(s-2)(s+1)^{2}}=f(t)$. The function $F(s)$ has poles at $\mathrm{s}=2$ and $\mathrm{s}=-1$. So

$$
\begin{aligned}
f(t) & =\operatorname{Res}_{z=2} \frac{e^{s t}}{(s-2)(s+1)^{2}}+\operatorname{Res}_{z=-1} \frac{e^{s t}}{(s-2)(s+1)^{2}} \\
& =\frac{e^{2 t}}{9}+\frac{-3 t e^{-t}-e^{-t}}{9}
\end{aligned}
$$

The above examples could all have been solved using partial fraction expansions. To finish with; an example for which the Bromwich contour integral is necessary.
Example 8.19 Find $\mathcal{L}^{-1}\left(\frac{1}{s^{\frac{1}{2}}}\right)$. The conditions of Theorem 8.3 do not hold as the singularity at $\mathrm{s}=0$ is not a pole. We need to amend the contour to avoid the discontinuity across the negative x -axis in the s-plane. We make a "cut" along the negative x -axis - this amounts to making a particular choice of definition for the inherently multivalued"function" $\frac{1}{s^{\frac{1}{2}}}$.
When s is real and positive, we take $\frac{1}{\mathrm{~s}^{\frac{1}{2}}}=\frac{1}{\sqrt{\mathrm{~s}}}>0$. Now define the "keyhole" contour as in Figure 7. As it contains no singularities, the total integral round the contour is zero.


Figure 7: Contour for Example 8.19

Take $\delta$ to be small and positive. The constant a is any positive real number greater than $\delta$. The parameter b is given $b y \mathrm{~b}=\sqrt{\mathrm{R}^{2}-\mathrm{a}^{2}}$.

- Along the upper branch of the cut, we have $s=\sigma e^{i \pi}$, $\mathrm{R} \rightarrow \sigma \rightarrow \delta$.
- Along the lower branch of the cut, we have $s=\sigma e^{-i \pi}$, $\delta \rightarrow \sigma \rightarrow R$.
- Along the arc $\mathrm{C}_{3}$ of the small circle radius $\delta$, we have $\mathrm{s}=\delta e^{\mathfrak{i} \theta}$, $\pi \rightarrow \theta \rightarrow-\pi$.
- Along the arc $\mathrm{C}_{1}$ of the circle radius R , we have $\mathrm{s}=\mathrm{Re}^{i \theta}$, $\theta_{1} \rightarrow \theta \rightarrow \pi$. (The angle $\theta_{1}$ is just $\tan ^{-1} \mathrm{~b} / \mathrm{a}$ where $\mathrm{b}=\sqrt{\mathrm{R}^{2}-\mathrm{a}^{2}}$ )
- Along the arc $\mathrm{C}_{2}$ of the circle radius R , we have $\mathrm{s}=\mathrm{Re}^{\mathrm{i} \mathrm{\theta} \theta}$, $-\pi \rightarrow \theta \rightarrow \theta_{2}$. (The angle $\theta_{2}$ is just $-\tan ^{-1} \mathrm{~b} / \mathrm{a}$.)
- The conditions for Theorem 8.3 in respect of the behaviour of $\mathrm{F}(\mathrm{s})$ for large $|\mathrm{s}|$ hold so we can conclude that the contributions from the arcs $\mathrm{C}_{1}$ and $\mathrm{C}_{2}$ of radius R go to zero as $\mathrm{R} \rightarrow \infty$.
- It is easy to check that the contribution from the arc $C_{3}$ of the small circle radius $\delta$ goes to zero as $\delta \rightarrow 0$.
- Along the top of the branch cut the integral takes the form (as $s=\sigma e^{i \pi}$ )

$$
\int_{R}^{\delta} \frac{e^{-t \sigma}}{i \sqrt{\sigma}}(-1) d \sigma .
$$

- Along the bottom of the branch cut the integral takes the form (as $s=\sigma e^{-i \pi}$ )

$$
\int_{\delta}^{R} \frac{e^{-t \sigma}}{-i \sqrt{\sigma}}(-1) d \sigma
$$

- Finally, the contribution from the integral along the vertical line from $\mathrm{a}-\mathrm{ib}$ to $\mathrm{a}+\mathrm{ib}$ is

$$
\int_{a-i b}^{a+i b} \frac{e^{s t}}{s^{\frac{1}{2}}} d s
$$

Now, as the two branch cut integrals are equal; we have, taking the limit as $\mathrm{R} \rightarrow \infty($ so $\mathrm{b} \rightarrow \infty)$ and dividing across by $2 \pi \mathrm{i}$

$$
\frac{1}{2 \pi i} \int_{a-i \infty}^{a+i \infty} \frac{e^{s t}}{s^{\frac{1}{2}}} d s=\frac{1}{\pi} \int_{0}^{\infty} \frac{e^{-t \sigma}}{\sqrt{\sigma}} d \sigma
$$

The RHS can be re-written as $\frac{2}{\pi} \int_{0}^{\infty} e^{-x^{2} t} \mathrm{~d} x-a$ standard integral which equals $\frac{1}{2} \sqrt{\pi / \mathrm{t}}$ so we have

$$
\mathcal{L}^{-1} \frac{1}{s^{\frac{1}{2}}}=\frac{1}{\sqrt{\pi / t}}
$$

### 8.5 Exercises

1. Questions on Laurent Series .
(a) Expand the following functions in Laurent Series that converge for $0<|z|<R$ and find the exact region of convergence.

$$
\frac{e^{-z}}{z^{3}} \quad \frac{e^{1 / z^{2}}}{z^{6}} \quad \frac{\cos 2 z}{z^{2}} \quad \frac{1}{z^{4}(1+z)} \quad \frac{1}{z^{2}\left(1-z^{2}\right)} \quad \frac{1}{z^{2}(z-3)}
$$

(b) Does $\tan (1 / z)$ have a Laurent Series that converges in a region $0<|z|<R$ for any $R$ ?
(c) Find all the Laurent Series with centre at $z=2$ for the function $\frac{4 z^{2}+2 z-4}{z^{3}-4 z}$ and find their regions of convergence.
2. Questions on residues
(a) Find the residues of the following functions at their singular points

$$
\frac{1}{1-z} \quad \frac{z+3}{z+1} \quad \frac{1}{z^{2}} \quad \frac{1}{(z-1)^{2}} \quad \frac{z}{z^{4}-1} \quad \cot z
$$

(b) For each of the following functions, find the residue at the singular points that lie inside the unit circle

$$
\frac{3 z^{2}}{1-z^{4}} \quad \frac{z-\frac{1}{4}}{z^{2}+3 z+2} \quad \frac{6-4 z}{z^{3}+3 z^{2}} \quad \frac{1}{\left(z^{4}-1\right)^{2}} \quad \frac{z+2}{(z+1)\left(z^{2}+16\right)} \quad \frac{4-3 z}{z^{3}-3 z^{2}+2 z}
$$

(c) Evaluate the integrals of the following functions round the unit circle anti-clockwise

$$
\begin{array}{llllll}
e^{1 / z} & z e^{1 / z} & \cot z & \tan z & \frac{1}{\sin z} & \frac{z}{2 z+i} \\
\frac{1}{\cosh z} & \frac{z^{2}-4}{(z-2)^{4}} & \frac{z^{2}+1}{z^{2}-2 z} & \frac{\sin \pi z}{z^{4}} & \frac{1}{1-e^{z}} & \frac{z^{2}+1}{e^{z} \sin z}
\end{array}
$$

3. Questions on the Residue Theorem
(a) Integrate $\frac{3 z^{2}+2 z-4}{z^{3}-4 z}$ round each of the following paths anti-clockwise: $|z|=1,|z|=3$ and $|z-4|=1$.
(b) Integrate $\frac{z+1}{z(z-1)(z-2)}$ round each of the following paths
clockwise: $|z-2|=\frac{1}{2},|z|=\frac{3}{2}$ and $\left|z-\frac{1}{2}\right|=\frac{1}{4}$.
(c) Integrate each of the following functions round the unit circle anti-clockwise:

$$
\begin{array}{llllll}
\frac{3 z}{3 z-1} & \frac{z+1}{4 z^{3}-z} & \frac{z^{5}-3 z^{3}+1}{(2 z+1)\left(z^{2}+4\right)} & \frac{z}{1+9 z^{2}} & \frac{(z+4)^{3}}{z^{4}+5 z^{3}+6 z^{2}} & \frac{6 z^{2}-4+1}{(z-2)\left(1-4 z^{2}\right)} \\
\tan 2 \pi z & \frac{\tan \pi z}{z^{3}} & \frac{e^{z}}{z^{2}-5 z} & \frac{e^{z}}{\sin z} & \frac{\cot z}{z} & \frac{e^{z^{2}}}{\cos \pi z}
\end{array}
$$

4. Questions on evaluating real integrals
(a) Evaluate the following real integrals (involving cos and $\sin$ ):

$$
\begin{array}{lll}
\int_{0}^{2 \pi} \frac{d \theta}{2+\cos \theta} & \int_{0}^{\pi} \frac{d \theta}{1+\frac{1}{3} \cos \theta} & \int_{0}^{\pi} \frac{d \theta}{k+\cos \theta}(k>1) \\
\int_{0}^{2 \pi} \frac{\cos \theta}{3+\sin \theta} d \theta & \int_{0}^{2 \pi} \frac{\sin ^{2} \theta}{5-4 \cos \theta} d \theta & \int_{0}^{2 \pi} \frac{\cos ^{2} \theta}{26-10 \cos 2 \theta} d \theta
\end{array}
$$

(b) Evaluate the following infinite (improper) integrals:

$$
\begin{array}{lll}
\int_{-\infty}^{\infty} \frac{d x}{1+x^{2}} & \int_{-\infty}^{\infty} \frac{d x}{\left(1+x^{2}\right)^{2}} & \int_{-\infty}^{\infty} \frac{d x}{1+x^{6}} \\
\int_{-\infty}^{\infty} \frac{x^{3}}{1+x^{8}} d x & \int_{0}^{\infty} \frac{1+x^{2}}{1+x^{4}} d x & \int_{-\infty}^{\infty} \frac{d x}{\left(x^{2}+1\right)\left(x^{2}+9\right)}
\end{array}
$$

## 9 Conformal Mappings and their Applications

It is easy to forget that a complex function $w=f(z)$ is a mapping from $\mathbb{C}$ to $\mathbb{C}$ and therefore from $\mathbb{R}^{2}$ to $\mathbb{R}^{2}$. So a domain $D$ in the $z$-plane maps into a region $E$ in the $\mathcal{w}$-plane (not necessarily a domain).

Before looking at the details of the geometry, remember that we saw in Theorem 6.3 that both the real and imaginary parts of any analytic function satisfy Laplace's Equation - used in Applied Mathematics to model (for example) fluid flow and electrical currents in two dimensions. We will see that this harmonic property is preserved when we apply certain mappings from the $z$-plane into the $\mathcal{w}$-plane. Suitable choices of mapping transform complicated boundary conditions into simple ones.

### 9.1 The Conformal Property

To see how curves in the $z$-plane are affected by an analytic function, let's begin with an example. Consider the case $w=\log z$ (taking the principal value of the multi-valued $\log$ function obtained by using the principal argument of $z$, namely $-\pi<\theta \leq \pi)$. We apply this mapping to the arc $A$ defined by $|z|=1, \pi / 6 \leq \arg z \leq \pi / 4$ and also to the line segment L defined by $\arg z=\pi / 6,1 \leq|z| \leq 2$. See Figures 8 and 9 below.


Figure 8: Before Log mapping


Figure 9: After Log mapping

- Under the Log transform, each point on the arc $A:|z|=1$, $\pi / 6 \leq \arg z \leq \pi / 4$ has an image on $A^{\prime}$ :

$$
w=i \arg z
$$

As $\arg z$ advances from $\pi / 6$ to $\pi / 4$, the image point $w$ moves from $\boldsymbol{\nu}^{\prime}$ along the vertical line $A^{\prime}$ in Fig 9.

- Under the Log transform, each point on the line $\mathrm{L}: \arg z=\pi / 6$, $1 \leq|z| \leq 2$ has an image on $L^{\prime}:$

$$
w=\log |z|+i \arg z=\log |z|+i \pi / 6
$$

As $z$ advances from 1 to 2 , the image point $w$ moves from $v^{\prime}$ along the horizontal line $\mathrm{L}^{\prime}$ in Fig 9.

Note that the line $L$ is perpendicular to the arc $\mathcal{A}$ at $v$ and the same is true of $L^{\prime}$ and $A^{\prime}$ in the $\mathcal{w}$-plane. We will see that this is always true for analytic mappings $w=f(z)$.

Definition 9.1 (Conformal mapping) A mapping $w=f(z)$ that preserves the size and sense (clockwise or anticlockwise) of the angle of intersection between any two curves intersecting at $z_{0}$ is conformal at $z_{0}$. If the mapping is conformal everywhere in a domain D we say that it is conformal in D .

We can now state and prove an important Theorem.
Theorem 9.1 (Conformal mapping) Let $\mathbf{f}(\boldsymbol{z})$ be analytic in a domain D . Then $\mathrm{f}(z)$ is conformal at every point in D where $f^{\prime}(z) \neq 0$.

Proof: Let a smooth curve $C$ be parameterised as
$z(t)=x(t)+i y(t)$. We take $x(t)$ and $y(t)$ to be differentiable (real) functions of $t$. Then $C$ is transformed into an "image curve" $C^{\prime}$ in the $w$-plane by $f(z)$ :

$$
w=f(z(t))=u(x(t), y(t))+i v(x(t), y(t))
$$

At any point $z_{0} \equiv z\left(\mathrm{t}_{0}\right)$ on the curve C , the real and imaginary parts $\dot{x}$ and $\dot{y}$ of the complex number $\dot{z} \equiv \frac{\mathrm{~d} z}{\mathrm{dt}}$ are the components of the tangent to the curve $C$ at $z_{0}$ (to see this consider $\Delta z \equiv z(\mathrm{t}+\Delta \mathrm{t})-z(\mathrm{t})$ for small $\Delta \mathrm{t}$ - then the limit as $\Delta \mathrm{t} \rightarrow 0$ (of $\frac{\Delta z}{\Delta \mathrm{t}}$ ) is just $\dot{z}$. In particular $\left.\left.\frac{d y}{d x}\right|_{z_{0}} \equiv \frac{\dot{y}}{\dot{x}}\right|_{\mathcal{O}_{0}}$ is the slope of the curve at $z_{0}$.

Similarly at the corresponding point $\mathcal{w}_{0}=f\left(z_{0}\right)$ in the $\mathcal{w}$-plane, $\dot{w} \equiv \frac{\mathrm{~d} w}{\mathrm{dt}}$ is the tangent to the image curve $\mathrm{C}^{\prime}$ at $w_{0}$. Se Figs 10 and 11 below.


Figure 10: Before Conformal Mapping


Figure 11: After Conformal Mapping

Now, using the Chain Rule,

$$
\frac{d w}{d t}=\frac{d w}{d z} \frac{d z}{d t}=f^{\prime}(z) \frac{d z}{d t}
$$

At $t=t_{0}$,

$$
\left.\frac{d w}{d t}\right|_{w_{0}}=\left.f^{\prime}\left(z_{0}\right) \frac{d z}{d t}\right|_{t_{0}}
$$

Equating the arguments of each side we have

$$
\left.\arg \frac{d w}{d t}\right|_{w_{0}}=\arg f^{\prime}\left(z_{0}\right)+\left.\arg \frac{d z}{d t}\right|_{t_{0}}
$$

or just $\phi=\alpha+\theta$, where $\phi=\left.\arg \frac{d w}{d t}\right|_{w_{0}}, \alpha=\arg f^{\prime}\left(z_{0}\right)$ and $\theta=\left.\arg \frac{d z}{d t}\right|_{t_{0}}$.

Now, $\theta$ and $\phi$ are the angles made (with the positive $\chi$-axis) by the tangents to the curves $C$ and $C^{\prime}$ at $z_{0}$ and $w_{0}$ respectively. So the latter equation tells us that under the mapping $w=f(z)$ the tangent to the curve $C$ at $z_{0}$ is rotated through an angle $\alpha \equiv \arg f^{\prime}\left(z_{0}\right)$.

Now consider another curve $C_{1}$ passing through the same point $z_{0}$; where the tangent at $z_{0}$ to $C_{1}$ makes an angle $\psi$ with the tangent to $C$. Then (and this is the key idea) this tangent is rotated through exactly the same angle $\alpha$ as this angle depends only on the mapping $f$ and not on the choice of curve $C$ or $C_{1}$.

So $C^{\prime}$ and $C_{1}^{\prime}$ have the same angle of intersection as do $C$ and $C_{1}$.

Note that the Theorem breaks down if $f^{\prime}(z)=0$.

Example 9.1 Consider the curve $C$ defined by $x=y$, for $x \geq 0$ and the curve $\mathrm{C}_{1}$ defined by $\mathrm{x}=1, \mathrm{y} \geq 1$. They intersect at the point $(1,1)$ and the angle between the tangents is $\pi / 4$ measured anticlockwise from C to $\mathrm{C}_{1}$. We will map these two curves using the mapping $w=1 / z$ and check that the angle of intersection is preserved.

We have

$$
w=\frac{1}{z}=u+i v=\frac{1}{x+i y}=\frac{x}{x^{2}+y^{2}}-\frac{i y}{x^{2}+y^{2}}
$$

On $\mathrm{C}, \mathrm{y}=\mathrm{x}$ so $\mathrm{u}=\frac{1}{2 \mathrm{x}}$ and $\mathrm{v}=-\frac{1}{2 \mathrm{x}}$ and therefore C is mapped into the line $\mathrm{C}^{\prime}: \mathbf{u}=-v$ in the $w$-plane. As $\mathrm{x} \geq 0$ we have $\mathrm{u} \geq 0$ and $v \leq 0$.

On $\mathrm{C}_{1}, \mathrm{x}=1$ so $\mathrm{u}=\frac{1}{1+\mathrm{y}^{2}}$ and $v=-\frac{\mathrm{y}}{1+\mathrm{y}^{2}}$ so $\mathrm{C}_{1}$ is mapped into $\mathrm{C}_{1}^{\prime}: v=-\mathrm{yu}$ in the $\boldsymbol{w}$-plane. Using $\mathfrak{u}=\frac{1}{1+\mathrm{y}^{2}}$ to eliminate y gives us $v=-\sqrt{u-u^{2}}$. Finally, squaring and re-arranging gives us $\mathrm{C}_{1}^{\prime}$ : $\left(u-\frac{1}{2}\right)^{2}+v^{2}=\left(\frac{1}{2}\right)^{2}-a$ circle centerd at $\left(\frac{1}{2}, 0\right)$ in the $w$-plane. The two image curves intersect at $\left(\frac{1}{2},-\frac{1}{2}\right)$. It is easy to check that the angle between the tangents to $\mathrm{C}^{\prime}$ and $\mathrm{C}_{1}^{\prime}$ is still $\pi / 4$ as predicted by the Theorem.

You should sketch the various curves.

### 9.2 The Bilinear Transformation

The Theorem just proved applies to any analytic mapping from the $z$-plane to the $w$-plane. A particular analytic mapping, the Bilinear mapping, is widely used as we will see to transform complicated boundary conditions into simpler ones. Its special virtue is that as we will prove - it transforms lines and circles into either lines or circles.

Definition 9.2 ( Bilinear Transformation) The bilinear transformation is defined by

$$
\begin{equation*}
w=f(z)=\frac{\mathrm{az}+\mathrm{b}}{\mathrm{cz}+\mathrm{d}}, \quad \text { where } \mathrm{a}, \mathrm{~b}, \mathrm{c} \text { and } \mathrm{d} \text { are complex constants. } \tag{9.1}
\end{equation*}
$$

Obviously the transformation is well defined provided $z \neq-\frac{d}{c}$.
It is useful to extend our terminology to allow us to refer to the "complex number $\infty$ ". We call the complex numbers $\mathbb{C}$ together with $\infty$ "the extended complex plane $\mathbb{C}^{*}$ ". (Think of $\infty$ as the point reached if we move along any path from the origin along which $|z|$ grows without bound.)

Now we can say that $-\frac{d}{c}$ is mapped into $\infty$.
It is easy to check that

$$
\begin{equation*}
\frac{d w}{d z} \equiv f^{\prime}(z)=\frac{a d-b c}{(c z+d)^{2}} \tag{9.2}
\end{equation*}
$$

so, if $a d-b c=0$ the mapping is a trivial constant map, i.e. every point in the $z$-plane is mapped into the same point in the $\mathcal{w}$-plane.

Provided that this degenerate case is excluded, the bilinear map is one to one and onto, to see this we can just invert the mapping by solving for $z$, giving

$$
\begin{equation*}
z=\frac{-\mathrm{d} w+\mathrm{b}}{\mathrm{c} w-\mathrm{a}} \tag{9.3}
\end{equation*}
$$

which is also a bilinear mapping and gives a finite value of $z$ for all $w \neq a / c$.

When c is zero, the mapping is linear and maps finite values of $z$ into finite values of $w$ and maps $\infty$ into $\infty$.

For $c \neq 0$, we have

- $-\mathrm{d} / \mathrm{c}$ maps into $\infty$
- $\infty$ maps into $a / c$
- all other values of $z$ map into a finite value of $w$.

So the bilinear mapping is one-to one and onto from $\mathbb{C}^{*} \rightarrow \mathbb{C}^{*}$.

It is useful to note that unbounded straight lines in $\mathbb{C}$ can be regarded as circles of infinite radius (and therefore unspecified centre). So in the following we will use the term "circle" in quotes to mean circles or unbounded straight lines. Omitting the quotes will mean a circle in the conventional sense.

Theorem 9.2 The bilinear mapping maps "circles" into "circles".
Proof: Begin by writing our bilinear map as (we assume that $c \neq 0$ )

$$
\begin{equation*}
w=\frac{a}{c}+\frac{(b c-a d)}{c} \frac{1}{c z+d} \tag{9.4}
\end{equation*}
$$

So the bilinear map can be expressed as a sequence or composition of simpler maps:

$$
\begin{aligned}
w_{1} & =c z \\
w_{2} & =w_{1}+d \\
w_{3} & =\frac{1}{w_{2}} \\
w_{4} & =\frac{(b c-a d)}{c} w_{3} \\
w & =\frac{a}{c}+w_{4}
\end{aligned}
$$

All of the above are either

- translations $w=z+k$,
- rotation/magnifications $w=k z$
- or inversions $w=\frac{1}{z}$.

If we can show that "circles" are mapped into "circles" by each of the above three operations we will have proved the Theorem.

- The case of translations is trivial.
- For rotation/magnifications $w=k z$ write $w=|k| e^{i \theta_{k}} z$. It is now easy to check the result algebraically; just write the equation of a straight line (use $a x+b y=c$ with $x=\frac{z+\bar{z}}{2}$ and $\left.y=\frac{z-\bar{z}}{2 i}\right)$ or a circle $\left|z-z_{0}\right|=r$ and apply the transformation. (Exercise.)
- Now we examine the case $w=\frac{1}{z}$. Consider the following quadratic equation in $x$ and $y$ :

$$
\begin{equation*}
A\left(x^{2}+y^{2}\right)+B x+C y+D=0 \tag{9.5}
\end{equation*}
$$

where $A, B, C$ and $D$ are all real.

Now if $A$ is zero, we just have the equation of a line. In this case it is easy to check (using $x=\frac{z+\bar{z}}{2}$ and $y==\frac{z-\bar{z}}{2 i}$ as before) that straight lines map into either straight lines or circles under the mapping $w=1 / z$. (Exercise.)
So assume that $A \neq 0$. Now, completing squares we have:

$$
\begin{equation*}
\left(x+\frac{B}{2 A}\right)^{2}+\left(y+\frac{C}{2 A}\right)^{2}=-\frac{D}{A}+\left(\frac{B}{2 A}\right)^{2}+\left(\frac{C}{2 A}\right)^{2} \tag{9.6}
\end{equation*}
$$

This is the equation of a circle provided that

$$
-\frac{D}{A}+\left(\frac{B}{2 A}\right)^{2}+\left(\frac{C}{2 A}\right)^{2} \geq 0
$$

as the radius squared cannot be negative. Rearranging gives the condition

$$
B^{2}+C^{2} \geq 4 A D
$$

Again, make the substitutions $x=\frac{z+\bar{z}}{2}$ and $y=\frac{z-\bar{z}}{2 i}$ :

$$
\begin{equation*}
A z \bar{z}+\frac{B}{2}(z+\bar{z})+\frac{C}{2 i}(z-\bar{z})+D=0 . \tag{9.7}
\end{equation*}
$$

As we saw above, this is the equation of a circle if $A \neq 0$ and $B^{2}+C^{2} \geq 4 A D$. It is the equation of a straight line if $A=0$.
Now set $z=1 / w$ to apply the mapping. The "circle" is transformed into:

$$
\mathrm{A} \frac{1}{w \bar{w}}+\frac{\mathrm{B}}{2}\left(\frac{1}{w}+\frac{1}{\bar{w}}\right)+\frac{\mathrm{C}}{2 i}\left(\frac{1}{w}-\frac{1}{\bar{w}}\right)+\mathrm{D}=0 .
$$

Rearranging gives:

$$
\begin{equation*}
D w \bar{w}+\frac{\mathrm{B}}{2}(w+\bar{w})-\frac{\mathrm{C}}{2 i}(w-\bar{w})+\mathrm{A}=0 . \tag{9.8}
\end{equation*}
$$

Note that this is the same as (9.7) except that $A \leftrightarrow D$ and $C \leftrightarrow-C$ (which leaves the condition $B^{2}+C^{2} \geq 4 A D$ unchanged) so this is a circle if $D \neq 0$ and $B^{2}+C^{2} \geq 4 A D$.

If $\mathrm{D}=0$ then (9.7) represents a straight line.
Finally, if $\mathrm{D}=0$, (9.6) is satisfied by $z=0$, so the "circle" in the $z$-plane passes through the origin and is transformed into a straight line in the $\mathcal{w}$-plane.

If $\mathrm{c}=0$ then $w=\mathrm{f}(z)=\mathrm{az}+\mathrm{b}$ is just a rotation followed by a translation - we have already checked the result for these mappings.

### 9.3 Constructing the "Right" Bilinear Transformation

We often need to find a specific bilinear transformation that will map certain points in the $z$-plane into a particular curve in the $w$-plane. A related problem is mapping a given line or circle into some other specific line or circle. So we need to choose the "right" values for $\mathfrak{a}, \mathrm{b}, \mathrm{c}$ and d .

If $a \neq 0$ we can divide through by $a$ and write $w=\frac{z+c_{1}}{c_{2} z+c_{3}}$ so specifying three points in the $z$-plane and three corresponding points in the $w$-plane constitutes three equations in the three unknowns $\mathfrak{c}_{1}, c_{2}$ and $c_{3}$. If no solution exists it must be because $a=0$.

A more direct way of solving for a bilinear transformation uses the "cross ratio".

Definition 9.3 (Cross Ratio) The cross ratio of four distinct complex numbers $z_{1}, z_{2}, z_{3}$ and $z_{4}$ is defined by

$$
\begin{equation*}
\left(z_{1}, z_{2}, z_{3}, z_{4}\right)=\frac{\left(z_{1}-z_{2}\right)\left(z_{3}-z_{4}\right)}{\left(z_{1}-z_{4}\right)\left(z_{3}-z_{2}\right)} \tag{9.9}
\end{equation*}
$$

If any of these numbers $\left(s a y z_{\mathfrak{i}}\right)$ is $\infty$ then the cross ratio is redefined so that the factors in the numerator and denominator containing $z_{i}$ are cancelled.. For example if $z_{4}=\infty$,

$$
\left(z_{1}, z_{2}, z_{3}, z_{4}\right)=\frac{\left(z_{1}-z_{2}\right)}{\left(z_{3}-z_{2}\right)}
$$

Example 9.2 The order of the numbers is significant - check that $(1,2,3,4)=-1 / 3$ while $(3,1,2,4)=4$.

The following Theorem will give us a neat method for constructing bilinear transformations.

Theorem 9.3 [Invarience of Cross Ratio] Under the bilinear transformation (9.4) the cross ratio of four points is unchanged; so that

$$
\begin{equation*}
\left(w_{1}, w_{2}, w_{3}, w_{4}\right)=\left(z_{1}, z_{2}, z_{3}, z_{4}\right) \tag{9.10}
\end{equation*}
$$

Proof: We have $w_{i}=\frac{\mathrm{a} z_{i}+\mathrm{b}}{\mathrm{c} z_{i}+\mathrm{d}}$ so

$$
w_{i}-w_{j}=\frac{a z_{i}+b}{c z_{i}+d}-\frac{a z_{j}+b}{c z_{j}+d}=\frac{(a d-b c)\left(z_{i}-z_{j}\right)}{\left(c z_{i}+d\right)\left(c z_{j}+d\right)} .
$$

Now just form the cross ratio $\left(w_{1}, w_{2}, w_{3}, w_{4}\right)$ and note that all the terms involving the parameters $a, b, c$ and $d$ cancel. (The case where one of the $z_{\mathfrak{i}}$ is $\infty$ is left as an exercise.)

We use this result to find an expression for $w$ in terms of $z$ as follows: rewrite (9.10) with $z_{4}=z$ and $w_{4}=w$ so that

$$
\begin{equation*}
\frac{\left(w_{1}-w_{2}\right)\left(w_{3}-w\right)}{\left(w_{1}-w\right)\left(w_{3}-w_{2}\right)}=\frac{\left(z_{1}-z_{2}\right)\left(z_{3}-z\right)}{\left(z_{1}-z\right)\left(z_{3}-z_{2}\right)} \tag{9.11}
\end{equation*}
$$

Once the values of $z_{1}, z_{2}, z_{3}, w_{1}, w_{2}$ and $w_{3}$, are supplied Eq. 9.11 can be solved for $w$ in terms of $z$.

Example 9.3 Find the bilinear transformation for:

$$
\begin{array}{ccc} 
& z & w \\
1 & 1 & 0 \\
2 & i & -1 \\
3 & 0 & -i
\end{array}
$$

Substituting into (9.11) gives

$$
\frac{(0-(-1))(-i-w)}{(0-w)(-i-(-1))}=\frac{(1-i)(0-z)}{(1-z)(0-i)}
$$

which simplifies to

$$
w=\frac{\mathfrak{i}(z-1)}{z+1}
$$

Example 9.4 For this mapping, what is the image of the circle passing through $z_{1}=1, z_{2}=\mathfrak{i}$ and $z_{3}=0$ ? The circle in the $z$-plane has centre at $(1+i) / 2$ as the three given points define a right angle (a sketch helps) and the radius is $1 / \sqrt{2}$. So we can write in the z-plane:

$$
\left|z-\frac{1+i}{2}\right|=\frac{1}{\sqrt{2}} .
$$

The image in the $\mathcal{w}$-plane must be either a circle or a straight line. The three given points $\mathcal{w}_{1}, w_{2}$ and $w_{3}$ are not collinear so the image must be a circle:

$$
\left|w+\frac{1+i}{2}\right|=\frac{1}{\sqrt{2}} .
$$

Finally, does the disk $\left|z-\frac{1+\mathfrak{i}}{2}\right|<\frac{1}{\sqrt{2}}$ map into the disk $\left|w+\frac{1+i}{2}\right|<\frac{1}{\sqrt{2}}$ (interior of the circle) or the annulus (exterior to circle) $\left|w+\frac{1+\mathfrak{i}}{2}\right|>\frac{1}{\sqrt{2}}$ ? The easiest way to answer the question is to take a convenient interior point in the z-plane and check whether it maps into the interior of the circle in the $\mathcal{w}$-plane. Take $z_{0}=\frac{1}{2}$ - it is easy to check that $\mathfrak{w}_{0}=f\left(z_{0}\right)=-\mathfrak{i} / 3$ and so the disk in the $z$-plane is mapped into the disk in the $w$-plane.

Example 9.5 Find the bilinear transformation that maps $z_{1}=1$, $z_{2}=\mathfrak{i}$ and $z_{3}=0$ (as in previous Example) into $w_{1}=0, w_{2}=\infty$ and $\mathfrak{w}_{3}=-\mathrm{i}$. As $\mathfrak{w}_{2}=\infty$, we have

$$
\frac{-\mathfrak{i}-w}{-w}=\frac{(1-\mathfrak{i})(0-z)}{(1-z)(0-\mathfrak{i})}
$$

which simplifies to $w=\frac{1-z}{i-z}$.
In this case the circle $\left|z-\frac{1+i}{2}\right|=\frac{1}{\sqrt{2}}$ maps into a "circle" passing through $0, \infty$ and $-i-i . e . a$ straight line along the imaginary axis in the w-plane. Check that the half-plane to the left of this line is the image of the interior of the circle in the z-plane.

For more complicated geometries we need to be more creative in choosing our mapping.

Example 9.6 Find the transformation that maps the domain $0<\arg z<\pi / 2$ (First quadrant) from the $z$-plane onto $|z|<1$ (unit disk) in the $w$-plane.

We need to transform the boundary of our z-plane domain (the positive x and y axes) into the unit circle $|\boldsymbol{w}|=1$. A bilinear transformation cannot transform a line with a right angle turn into a circle. (Why?) So the answer cannot be a bilinear transformation.

But the mapping $w=z^{2}$ maps our "first quadrant" onto the upper half of the z-plane. If we can then find a second transform that maps the upper half of the z-plane onto the interior of the unit circle in the $\mathcal{w}$-plane, we can compose them to get the required mapping.

Take the points:

|  | $s$ | $w$ |
| :---: | :---: | :---: |
| 1 | -1 | -1 |
| 2 | 1 | $i$ |
| 3 | $\infty$ | 1 |

You should check that the corresponding mapping is

$$
w=\frac{s+1+2 i}{s+1-2 i}
$$

and that it maps the upper half of the s-plane into the unit circle in the $w$-plane. So composing this mapping with $s=z^{2}$ we find that

$$
w=\frac{z^{2}+1+2 i}{z^{2}+1-2 i}
$$

is the required result.

Example 9.7 The choice made was one of (infinitely) many possibilities. A general result can be found for the mapping from the s-plane to the $w$-plane. Write

$$
w=\left(\frac{a}{c}\right) \frac{z+b / a}{z+d / c} \equiv\left(\frac{a}{c}\right) \frac{z+u}{z+V}
$$

where $\mathrm{U}=\mathrm{b} / \mathrm{a}$ and $\mathrm{V}=\mathrm{d} / \mathrm{c}$. Now, we need $|\mathcal{w}|=1$ for all real $z$ so taking $|z| \rightarrow \infty$ we must have

$$
\left|\frac{\mathrm{a}}{\mathrm{c}}\right|=1, \quad \text { and so we can write } \frac{\mathrm{a}}{\mathrm{c}}=e^{\mathfrak{i} \psi}
$$

Now

$$
w=e^{i \psi} \frac{z+U}{z+V}
$$

Takingg $z$ real $(z=x)$ we must have

$$
1=\left|\frac{x+\mathrm{U}}{\mathrm{x}+\mathrm{V}}\right|, \quad \text { for all } x \in \mathbb{R}
$$

Writing $\mathrm{U}=\mathrm{u}_{1}+\mathfrak{i} \mathfrak{u}_{2}$ and $\mathrm{V}=v_{1}+\mathfrak{i} v_{2}$ it is easy to see that we must have

$$
2 x u_{1}+u_{1}^{2}+u_{2}^{2}=2 x v_{1}+v_{1}^{2}+v_{2}^{2}, \quad \text { for all } x \in \mathbb{R}
$$

So we must have $u_{1}=v_{1}$ and therefore $u_{2}^{2}=v_{2}^{2}$ or $u_{2}= \pm v_{2}$. This gives us $\mathrm{U}=\mathrm{V}$ or $\mathrm{U}=\overline{\mathrm{V}}$. The first choice must be discarded as it corresponds to a constant mapping.

So the most general form for a mapping from the x -axis in the $z$-plane to the unit circle in the $w$-plane is

$$
w=e^{i \psi} \frac{(z-\overline{\mathrm{U}})}{(z-\overline{\mathrm{U}})}, \quad \text { where } \psi \in \mathbb{R} \text { and } \mathrm{U} \in \mathbb{C}
$$

Note that the mapping derived previously

$$
w=\frac{s+1+2 i}{s+1-2 i}
$$

is a special case of this with $\psi=0$ and $\mathrm{U}=-1-2 \mathrm{i}$.

### 9.4 Conformal Mappings and Boundary Value Problems

We have already seen in Thm. 6.3 that the real and imaginary parts of an anlaytic function are harmonic - satisfy Laplace's Equation $u_{x x}+u_{y y}=0$. We can now prove a Theorem that shows that this property is preserved under the action of a conformal mapping. This will allow us to transform a problem with a complicated boundary into one with a simpler boundary where a solution to Laplace's Equation can be more easily found.

Theorem 9.4 Let the analytic function $\mathcal{w}=\mathrm{f}(z)$ map the domain D from the z-plane to a domain E in the $w$-plane. Suppose that a function $\phi_{\mathrm{E}}(u, v)$ is harmonic in E so that at any point $w=u+\mathfrak{i} v \in E$,

$$
\begin{equation*}
\frac{\partial^{2} \phi_{\mathrm{E}}}{\partial u^{2}}+\frac{\partial^{2} \phi_{\mathrm{E}}}{\partial v^{2}}=0 \tag{9.12}
\end{equation*}
$$

Then $\phi(\mathrm{x}, \mathrm{y}) \equiv \phi_{\mathrm{E}}(\mathrm{u}(\mathrm{x}, \mathrm{y}), v(\mathrm{x}, \mathrm{y}))$ is harmonic in D , in other words for any point $z=x+\mathfrak{i y} \in \mathrm{D}$,

$$
\begin{equation*}
\frac{\partial^{2} \phi}{\partial x^{2}}+\frac{\partial^{2} \phi}{\partial y^{2}}=0 \tag{9.13}
\end{equation*}
$$

Proof: Use the Chain Rule:

$$
\phi_{\chi}=\phi_{\mathrm{E}_{\mathfrak{u}}} \cdot \mathrm{u}_{\chi}+\phi_{\mathrm{E}_{v}} \cdot v_{\chi}
$$

so that (differentiating the $u_{x}$ and $v_{x}$ factors wrt $x$ )

$$
\begin{aligned}
& \phi_{x x}=\phi_{\mathbf{E}_{\mathbf{u}}} \cdot \mathbf{u}_{\mathrm{xx}}+\phi_{\mathrm{E}_{\mathbf{v}}} \cdot \mathbf{v}_{\mathrm{xx}} \\
& \quad+u_{x} \cdot\left\{u_{x} \cdot \phi_{\mathrm{E}_{u \mathfrak{u}}}+v_{x} \phi_{\mathrm{E}_{\mathfrak{u} v}}\right\}+v_{x} \cdot\left\{u_{x} \cdot \phi_{\mathrm{E}_{v u}}+v_{x} \phi_{\mathrm{E}_{v v}}\right\} .
\end{aligned}
$$

Similarly so that (differentiating the $u_{x}$ and $v_{x}$ factors wrt $x$ )

$$
\begin{aligned}
& \phi_{y y}=\phi_{\mathrm{E}_{u}} \cdot u_{\mathrm{yy}}+\phi_{\mathrm{E}_{v}} \cdot v_{\mathrm{yy}} \\
& \quad+u_{y} \cdot\left\{u_{y} \cdot \phi_{\mathrm{E}_{u u}}+v_{y} \phi_{\mathrm{E}_{u v}}\right\}+v_{y} \cdot\left\{u_{\mathrm{y}} \cdot \phi_{\mathrm{E}_{v u}}+v_{y} \phi_{\mathrm{E}_{v v}}\right\} .
\end{aligned}
$$

The blue and red terms both sum to zero as both $u$ and $v$ are harmonic given that $w=f(z)=\mathfrak{u}+\mathfrak{i} v$ is analytic.

Also both $\phi_{\mathrm{E}_{\mathfrak{u} v}}$ and $\phi_{\mathrm{E}_{v u}}$ have coefficient $u_{x} v_{x}+u_{y} v_{y}=0$ by the Cauchy-Riemann equations.

We are left with

$$
\phi_{x x}+\phi_{y y}=\phi_{\mathrm{E}_{u u}}\left(u_{x}^{2}+u_{y}^{2}\right)+\phi_{\mathrm{E}_{v v}}\left(v_{x}^{2}+v_{y}^{2}\right)
$$

Using the Cauchy-Riemann equations again allows us to write

$$
\phi_{x x}+\phi_{y y}=\left(\phi_{\mathrm{E}_{u u}}+\phi_{\mathrm{E}_{v v}}\right)\left(u_{x}^{2}+v_{x}^{2}\right)
$$

So, as $\phi_{\mathrm{E}_{\mathfrak{u}}}+\phi_{\mathrm{E}_{v \nu}}=0$ we have $\phi_{x x}+\phi_{y y}=0$ as required.

Example 9.8 Take $\phi_{\mathrm{E}}=\mathrm{e}^{\mathfrak{u}} \cos \boldsymbol{v}$ which is $\mathfrak{R} e^{\mathfrak{w}}$. It is easy to check that (9.12) is satisfied - as it must be given that $\mathrm{e}^{w}$ is analytic in $w$. Let $w=z^{2}=\left(x^{2}-y^{2}\right)+\mathfrak{i} 2 x y=u+\mathfrak{i v}$. Now $\phi(x, y)=e^{x^{2}-y^{2}} \cos 2 x y$ satisfies (9.13) as expected.

To see the usefulness of Thm. 9.4 imagine that we are given a domain D in the $z$-plane. We seek a function $\phi(x, y)$ that is harmonic in D (satisfies Laplace's equation) and that takes certain given boundary conditions on the boundary of D. Suppose that we can find a clever conformal mapping $w=u+i v$ that maps $D$ onto a domain $E$ in the $\mathcal{w}$-plane - chosen so that $E$ has a simpler shape than D.

Now if we can find a harmonic function $\phi_{E}(u, v)$ that satisfies the boundary conditions on the boundary of $E$ corresponding to the given boundary conditions on the corresponding parts of the boundary of D we know by Thm. 9.4 that the function $\phi(x, y)$ as defined above is harmonic and satisfies the boundary conditions on the boundary of D .


Figure 12: Transforming a boundary value problem

The Riemann mapping Theorem guarantees that any simply connected domain D can be mapped onto the unit disc - but provides no way of constructing such a mapping. We will consider only simple transformations such as the bilinear map, powers of $z$ and log.

Example 9.9 Two cylinders are maintained at temperatures of $0^{\circ}$ and $100^{\circ}$, as shown in Fig. 13 on the next Slide. (An infinitesimal gap separates the cylinders at the origin.) Find $\phi(x, y)$, the temperature in the domain between the cylinders. Obviously the steady-state heat equation is just Laplace's Equation.


Figure 13: Simplifying a boundary value problem

Solution: The shape of the domain between the cylinders is complicated. But as the bilinear map can transform circles into straight lines we can transform the domain into an infinite strip as in the Figure.

We will use the bilinear mapping. Take the points (as in Fig. 13):

|  |  | $z$ | $w$ |  |
| :---: | :---: | :---: | :---: | :---: |
| 1 | $A$ | $\frac{1}{2}$ | 1 | $A^{\prime}$ |
| 2 | B | 1 | 0 | $\mathrm{~B}^{\prime}$ |
| 3 | O | 0 | $\infty$ | $\mathrm{O}^{\prime}$ |

Then the cross-ratio formula (9.9) becomes (note that the terms involving $w_{3}=\infty$ have to be cancelled above and below):

$$
\frac{1}{1-w}=\frac{\left(\frac{1}{2}-1\right)(0-z)}{\left(\frac{1}{2}-z\right)(0-1)}
$$

Simplifying and solving for $w$ gives

$$
w=\frac{1-z}{z}
$$

The cylindrical boundary kept at $100^{\circ}$ is transformed into the vertical line $\mathrm{u}=1$ and the cylinder at $0^{\circ}$ becomes the vertical line $u=0$.

Our problem is now the (much) simpler one of finding a function $\phi_{\mathrm{E}}(\mathrm{u}, \mathrm{v})$ that is harmonic in the strip - and satisfies the boundary conditions. The "obvious" choice is $\phi_{\mathrm{E}}(u, v)=100 u$ as the symmetry of the geometry in the z-plane implies that $\phi_{\mathrm{E}}$ is independent of $v$.

Now Thm. 9.4 tells us that $\phi(x, y)=100 u=100 \Re w=100 \Re \frac{1-z}{z}$.
The latter simplifies to

$$
\phi(x, y)=100\left(\frac{x}{x^{2}+y^{2}}-1\right)
$$

In the last Example, the isotherms (curves of constant temperature) are given by

$$
T_{0}=100\left(\frac{x}{x^{2}+y^{2}}-1\right)
$$

Completing the square gives:

$$
\left(x-\frac{1 / 2}{1+\frac{T_{0}}{100}}\right)^{2}+y^{2}=\left(\frac{1 / 2}{1+\frac{T_{0}}{100}}\right)^{2}
$$

So the isotherms are cylinders (circles in the $x-y$ plane) of increasing radius offset along the $x$-axis.

The "stream function" is just the imaginary part of the complex potential of which $\phi$ is the real part. For the Example above,

$$
\Phi(z)=\phi(x, y)+\mathfrak{i} \psi(x, y)
$$

In general we can find $\psi$ using the Cauchy-Riemann equations but in the present case we obviously have $\Phi(z)=\frac{1-z}{z}$ and so

$$
\psi(x, y)=\mathfrak{I} \frac{1-z}{z}=-\frac{y}{x^{2}+y^{2}}
$$

The "streamlines" for a complex potential $\Phi$ are the curves $\psi(x, y)=$ const and give the direction of flow of the quantity represented by the potential, in this case heat. It is easy to check that the streamlines $\psi(x, y)=C$ for the above problem are the cylinders (circles in the $x-y$ plane)

$$
x^{2}+\left(y+\frac{1}{2 C}\right)^{2}=\frac{1}{4 C^{2}}
$$

Example 9.10 Suppose a heat-conducting material occupies a wedge $0 \leq \arg z \leq \alpha$. Let the horizontal boundary be kept at a temperature of $\mathrm{T}_{1}^{\circ}$ and the oblique boundary be kept at a temperature of $\mathrm{T}_{2}^{\circ}$.
(i) Show that the conformal mapping $w=u+\mathfrak{i v}=\log z$ transforms the wedge into a strip in the uv plane parallel to the u-axis.
(ii) Show that the solution in the $u v$ plane must be of the form $\mathrm{A} v+\mathrm{B}$ and find A and B .
(iii) Show that the temperature in the xy plane is given by

$$
\phi(x, y)=\frac{T_{2}-T_{1}}{\alpha} \tan ^{-1}\left(\frac{y}{x}\right)+T_{1}
$$

(iv) Describe the streamlines and isotherms in the wedge.

Example 9.11 (The Last One) Suppose that a cylinder has its axis a distance H from a plane (see the Figure).


Figure 14: Final example
Suppose that the skin of the cylinder is maintained at $100^{\circ}$ and the temperature of the plane is maintained at $0^{\circ}$. Find the temperature between the cylinder and the plane.

As in the previous example we seek an appropriate conformal transformation to map the region in the z-plane into a simpler region in the w-plane.

1. Construct a bilinear map that transforms the points $z_{1}, z_{2}$ and $z_{3}$ in the z-plane into the corresponding points $\mathcal{w}_{1}, \mathcal{w}_{2}$ and $\mathfrak{w}_{3}$ (see the Figure). You should find:

$$
w=-\frac{(\mathrm{R}-\rho \mathrm{H}) z+\rho\left(\mathrm{H}^{2}-\mathrm{R}^{2}\right)}{(\rho \mathrm{R}-\mathrm{H}) z+\left(\mathrm{H}^{2}-\mathrm{R}^{2}\right)} .
$$

2. The parameter $\rho$ is still arbitrary - to fix it we can require that $z_{4}$ (the point at infinity) maps into $w_{4}=\rho$. This gives a quadratic equation for $\rho-$ choose the root that gives $\rho>1$.
3. Check that our transform maps the y -axis $\mathrm{z}=\mathrm{it},-\infty<\mathrm{t}<\infty$ into the larger circle and the circle $z=\mathrm{H}+\mathrm{Re}^{i t}, 0 \leq \mathrm{t} \leq 2 \pi$ into the smaller one (difficult).
4. Now use the polar form for Laplace's equation in the r, $\theta$-plane:

$$
\phi_{\mathrm{E}_{r r}}+\frac{1}{\mathrm{r}^{2}} \phi_{\mathrm{E}_{\theta \theta}}+\frac{1}{\mathrm{r}} \phi_{\mathrm{E}_{r}}=0
$$

5. The symmetry in the $w$-plane suggests that we take $\phi_{\mathrm{E}} a$ function of r only - so

$$
\phi_{\mathrm{E}_{r r}}+\frac{1}{\mathrm{r}} \phi_{\mathrm{E}_{r}}=0
$$

Solve this equation - use the boundary conditions on the two circles $|w|=1$ and $|w|=\rho$.
6. Finally, write $\phi(x, y)=\phi_{\mathrm{E}}(\mathrm{u}(\mathrm{x}, \mathrm{y}), v(\mathrm{x}, \mathrm{y}))$. What are the isotherms of $\phi$ ?

