REPORT

Jae Yoon Chun, Mairi McLaren,
Emer Nevin, Jun Tange

Year 4 Project
School of Mathematics
University of Edinburgh
November 12, 2016

Abstract

The Sturm-Liouville problem has been around for a mere two hundred years. Despite its young age, a surprising number of phenomena in this world can be related to the problem. The foundations section of our report will discuss grounding concepts, definitions and theorems, relevant to specific mathematical conditions surrounding the problem, such as boundary or initial conditions. The foundations section will also demonstrate some common methods of finding solutions. This will allow us to consider specific examples of the Sturm-Liouville problem in a physical context, where we explore correlations and derive key equations which are commonly found in the real world.

Each application we have presented is influential in its own regard. However, as the contents presentable in this report are limiting, we acknowledge the necessity for further investigation. Overall, we have aimed to provide some substantial insight into the Sturm-Liouville problem, by developing concrete foundations and exploring various examples and applications.

Contents

Abstract **2**

Introduction **5**

I Foundations 7

1 Lagrange's Identity **8**
- 1.1 The Sturm-Liouville Boundary Value Problem 8
- 1.2 An Adjoint Operator 9
- 1.3 A Self-Adjoint Operator 9

2 Eigenvalues and Eigenfunctions **11**
- 2.1 Eigenvalues 11
- 2.2 Eigenfunctions 12
- 2.3 Existence of Eigenvalues and Eigenfunctions 12
- 2.4 Main Properties of Eigenfunctions and Eigenvalues 14

3 Fourier Sine Series Expansion **18**
- 3.1 Functions as a Series 18
- 3.2 Series of Eigenfunctions 19

4 The Nonhomogeneous Boundary Value Problem **20**
- 4.1 Motivation 20
- 4.2 Solving Nonhomogeneous Problems 20
- 4.3 Nonhomogeneous Heat Conduction 22
- 4.4 Method of Eigenfunction Expansion Using Green's Formula 25

5 The Singular Sturm-Liouville Problem **28**
- 5.1 Motivation 28
- 5.2 Bessel's Equation 30
  - 5.2.1 Solving Bessel's Equation of Order Zero 31
- 5.3 Generalising Bessel's Equation 33
- 5.4 Bessel's Equation as an Example of a Singular Problem 33
- 5.5 General Problem 35
- 5.6 Revisiting Self-Adjointness 36
- 5.7 Continuous Spectrum 36
  - 5.7.1 Returning to our Bessel Equation Example 37

6 Series of Orthogonal Functions: Mean Convergence **38**

II Applications 41

7 Application to Partial Differential Equations **42**
7.1 Vibrations of a Circular Elastic Membrane 42
7.2 Conclusion 45

8 Kepler's Equation **46**
8.1 Motivation 46
8.2 Kepler's Equation 46
8.3 Kepler's Equation and Bessel Functions 48
8.4 Conclusion 49

9 The Korteweg-de Vries Equation **50**
9.1 Motivation 50
9.2 The Sturm-Liouville Equation as the KdV Equation 50
9.3 Fundamental Solutions 52
9.4 Determining Coefficients and the Wronskian 54
9.5 Reflection and Transmission Coefficients 55
9.6 Auxiliary Theorems 56
9.7 Prelude to Inverse Scattering 56
9.8 Conclusion 60

Conclusion **61**

Introduction

The Sturm-Liouville problem is a boundary value problem for a linear second order differential equation of a particular form. The problem first arose in the 1800s as a result of Jacques Charles-Francois Sturm and Joseph Liouville's interest in the subject of differential equations derived from the theory of heat. Based on this theory, they wrote papers involving differential equations as boundary value problems, investigating the properties of the eigenvalues and eigenfunctions of these problems [1] [2]. Their idea became known today as the Sturm-Liouville problem, and is widely applicable in the world of physics.

In this report, we will firstly introduce the "classical" Sturm-Liouville problem. The purpose of this is to cover the basic theory conceptualised by Sturm and Liouville. This involves some research into the eigenvalues and eigenfunctions of the boundary value problem. Then we define the Fourier sine series expansion, as this allows us to express the solutions of the Sturm-Liouville problem as a series of eigenfunctions, and therefore allows us to obtain the general solution.

The first issue addressed is that of the nonhomogeneous case. This is nice in the sense that the same theory described previously is applicable. The nonhomogeneous case is relevant to many real life examples such as heat conduction problems.

Then we discuss a more complicated form of the Sturm-Liouville problem, the singular problem. In order to do this, we must define Bessel functions, which we will see occur in various examples throughout the report and so are fundamental to the discussion of the Sturm-Liouville problem. The singular problem arises in the case where there is a singularity at one or both of the boundary points, or if the interval is unbounded. This can be solved using Bessel functions.

The final element of the theory which will be discussed in this report is the theory of mean convergence. This is required when we want to find a solution associated with eigenfunctions of a piecewise function that does not converge pointwise.

After discussing the theory behind the Sturm-Liouville problem, we next want to explore some of its applications, as there are many interesting examples of the problem in the real world, which come in different forms. We have investigated three examples in this report.

The first example we discuss is the applicability of the Sturm-Liouville problem to partial differential equations. This is a well known application of the problem.

Then we move on to Kepler's equation. It is an equation which describes the planetary motion related to the time and place of an object in an elliptic orbit.

Seemingly irrelevant to the Sturm-Liouville problem, here we characterise some interesting connections with planetary motions.

Finally we discuss the Korteweg-de Vries (KdV) equation. We explicitly demonstrate its relationship with the Sturm-Liouville equation, then further use the solutions to the Sturm-Liouville equation and some of the properties of eigenvalues discussed in preceding chapters to detail certain new concepts. We aim to show how our presentations thus far may be a foundation for analysing and solving the KdV equation.

Our main goal for this project was to learn about the Sturm-Liouville problem and focus on various applications to show that it plays a part in a wide range of physical problems. To do this we discuss the theory behind the problem and then explore three of its applications which we find particularly fascinating.

Part I

Foundations

Chapter 1

Lagrange's Identity

1.1 The Sturm-Liouville Boundary Value Problem

The Sturm-Liouville boundary value problem, which we will henceforth refer to simply as the Sturm-Liouville problem, is a boundary value problem for a real second order differential equation of the form,

$$[p(x)y']' - q(x)y + \lambda r(x)y = 0, \tag{1.1}$$

on the interval $0 < x < 1$, where the boundary conditions are

$$a_1 y(0) + a_2 y'(0) = 0 \ , \ b_1 y(1) + b_2 y'(1) = 0. \tag{1.2}$$

These conditions, since each involves only one of the boundary points, are called "separated." We note that the form we have given in (1.1) is homogeneous. Non-homogeneous cases shall not be our focus until later. In addition, we will refer to equations of the form (1.1) as the Sturm-Liouville equation. An identity relevant to establishing some of the properties of the Sturm-Liouville problem, known as "Lagrange's identity" will be our key focus in this section.

To begin with, Lagrange's identity, denoted L will allow us to realise a link between eigenvalues, eigenfunctions and their properties. First, we define some linear homogeneous differential operator L, as $L[y] = \lambda r(x)y$. Consequently, we mention that (1.1) can be written as follows,

$$L[y] = -[p(x)y']' + q(x)y. \tag{1.3}$$

Now for the functions we have included in both (1.1) and (1.3), here are our assumptions. The functions $p(x), p'(x), q(x)$ and $r(x)$ are continuous on $0 \leq x \leq 1$ and also $p(x) > 0$ and $r(x) > 0$ at all points in $0 \leq x \leq 1$. For the time being we will only be discussing regular Sturm-Liouville problems.

Definition 1.1.1. (1.1) is said to be a **regular Sturm-Liouville equation** in a closed finite interval, $[a, b]$, when the functions $p(x), r(x)$ are positive $\forall x \in [a, b]$, where $p(x), q(x)$ and $r(x)$ are continuous and bounded in the interval.

1.2 An Adjoint Operator

A notable feature of L is that it is self-adjoint, given that certain conditions are satisfied. To explain such a property, we introduce a tool known as the inner product and its notation as follows.

Definition 1.2.1. For two real-valued functions $u(x), v(x)$ on the interval $[0,1]$ we denote the **inner product** of $u(x)$ and $v(x)$ as (u,v) and define as

$$(u,v) = \int_0^1 u(x)v(x)dx. \tag{1.4}$$

We note that in the case of complex-valued functions as opposed to real functions, we can modify (1.4) as,

$$(u,v) = \int_0^1 u(x)\bar{v}(x)dx, \tag{1.5}$$

where $\bar{v}(x)$ represents the complex conjugate of $v(x)$. (1.5) coincides with (1.4) when $u(x), v(x)$ are real. Now we mention the notion of adjoint operators.

Definition 1.2.2. The **adjoint of an operator**, say Λ, is denoted by Λ^* such that it satisfies $(\Lambda[u], v) = (u, \Lambda^*[v])$ for all $u(x), v(x)$.

We note that this definition is not rigorous, as an idea of this concept is all which is required for our discussion.

1.3 A Self-Adjoint Operator

Using this notation, the aforementioned self-adjointness means that such an operator, be it L, satisfies $L = L^*$. Namely, if

$$(L[u], v) = (u, L[v]) \tag{1.6}$$

is true for any $u(x), v(x)$. We set our goal in this chapter clearly now; we will prove that L, as defined earlier, satisfies equation (1.6). Take functions $u(x)$ and $v(x)$, which have continuous second order derivatives along the interval $[0,1]$. Then starting from (1.3),

$$\begin{aligned}\int_0^1 L[u]v dx &= \int_0^1 [-(pu')'v + quv]dx \\ &= \Big[-p(x)u'(x)v(x)\Big]_0^1 + \Big[p(x)u(x)v'(x)\Big]_0^1 + \int_0^1 \Big[-u(pv')' + quv\Big]dx \\ &= \Big[-p(x)(u'(x)v(x) - u(x)v'(x))\Big]_0^1 + \int_0^1 uL[v]dx,\end{aligned} \tag{1.7}$$

where we have used integration by parts twice to obtain the result. Rearranging (1.7) so that the integrals are all on the left hand side,

$$\int_0^1 L[u]v - uL[v]dx = \Big[-p(x)(u'(x)v(x) - u(x)v'(x))\Big]_0^1. \tag{1.8}$$

We may stop here unless we add more assumptions; $u(x)$ and $v(x)$ satisfy the boundary conditions (1.2).

1. Further taking $a_2 \neq 0$ and $b_2 \neq 0$, we may evaluate the right-hand side of equation (1.8) as,

$$-p(1)[u'(1)v(1) - u(1)v'(1)] + p(0)[u'(0)v(0) - u(0)v'(0)], \tag{1.9}$$

then by using the boundary conditions (1.2), we rearrange them so that,

$$-\frac{b_1}{b_2}u(1) = u'(1)\ ,\ -\frac{b_1}{b_2}v(1) = v'(1)\ ,\ -\frac{a_1}{a_2}v(1) = u'(0)\ ,\ -\frac{a_1}{a_2}v(1) = v'(0). \tag{1.10}$$

Substitution of (1.10) into (1.9) yields,

$$-p(1)\Big[-\frac{b_1}{b_2}u(1)v(1) + \frac{b_1}{b_2}u(1)v(1)\Big] + p(0)\Big[-\frac{a_1}{a_2}u(0)v(0) + \frac{a_1}{a_2}u(0)v(0)\Big] = 0. \tag{1.11}$$

2. In a similar manner, should we assume $a_2 = 0$, we make the following substitution

$$-\frac{b_1}{b_2}u(1) = u'(1)\ ,\ -\frac{b_1}{b_2}v(1) = v'(1).$$

Then, safely assuming that $a_1 \neq 0$,

$$-p(1)\Big[-\frac{b_1}{b_2}u(1)v(1) + \frac{b_1}{b_2}u(1)v(1)\Big] + p(0)\Big[u'(0)v(0) + u(0)v'(0)\Big] = 0. \tag{1.12}$$

3. By identical reasoning, for $b_2 = 0$, taking $b_1 \neq 0$, we have

$$-p(1)\Big[u'(1)v(1) + u(1)v'(1)\Big] + p(0)\Big[-\frac{a_1}{a_2}u(0)v(0) + \frac{a_1}{a_2}u(0)v(0)\Big] = 0. \tag{1.13}$$

In summary, if L is defined as in (1.3) and if functions $u(x), v(x)$ satisfy the boundary conditions (1.2), equation (1.8) simplifies to

$$\int_0^1 L[u]v - uL[v]dx = 0. \tag{1.14}$$

Using definition (1.2.1), (1.14) can be written as

$$(L[u], v) - (u, L[v]) = 0. \tag{1.15}$$

Hence (1.6) is satisfied. Therefore, Lagrange's identity is self-adjoint.

Chapter 2

Eigenvalues and Eigenfunctions

2.1 Eigenvalues

Eigenvalues are most commonly introduced during the early studies of linear algebra.

Definition 2.1.1. Let A be a $n \times n$ square matrix. The scalar λ is said to be an **eigenvalue** of A if there is a nontrivial solution v of

$$Av = \lambda v. \tag{2.1}$$

where v is called the eigenvector corresponding to λ.

Without further elaboration in the context of linear algebra, in light of our topic, for a linear differential operator D on the space of infinitely differentiable real functions C^∞, the eigenvalue λ of a function $f \in C^\infty$ is a scalar that satisfies

$$Df = \lambda f. \tag{2.2}$$

We note that the notion is essentially identical to that of linear algebra from (2.1). We look at a specific example to demonstrate this.

Example 2.1.1. Let us consider the differential equation,

$$y'' + \lambda y = 0, \text{ where } y(0) = 0 \text{ and } y(1) = 0. \tag{2.3}$$

Note that this a Sturm-Liouville problem as in (1.1), (1.2). The idea which will be clear as we arrive at the solution is that there are multiple solutions of problem (2.3). In fact, infinitely many. We proceed to find $y(t)$.

Since the equation is a second order differential equation, we take $y(t) = e^{\mu t}$ for some constant $\mu \in \mathbb{C}$. Then $y''(t) = \mu^2 e^{\mu t}$ and substituting into (2.3) yields,

$$e^{\mu t}(\mu^2 + \lambda) = 0. \tag{2.4}$$

Noting that $e^{\mu t} \neq 0, \forall \mu, t \in \mathbb{C}$, we find $\mu = \pm i\sqrt{\lambda}t$. By the principle of superposition, we have $y = a_1 e^{i\sqrt{\lambda}t} + a_2 e^{-i\sqrt{\lambda}t}$ for constants $a_1, a_2 \in \mathbb{C}$. Using one of the the given boundary conditions $y(0) = 0$, we obtain $a_1 = -a_2$ and thus from the remaining condition $y(1) = 0$,

$$a_1 e^{i\sqrt{\lambda}} - a_1 e^{-i\sqrt{\lambda}} = 0. \tag{2.5}$$

Using $e^{i\theta} = \cos\theta + i\sin\theta$ and rewriting (2.5), we arrive at

$$a_1(2i\sin\sqrt{\lambda}) = 0, \tag{2.6}$$

where $a_1 \neq 0$, since otherwise we only have the trivial solution $y(t) = 0$. We then have that $\sqrt{\lambda} = n\pi$ for $n \in \mathbb{N}$. Namely, $\lambda = n^2\pi^2$. It is clear that there are n solutions, thus we express our nth solution as $\lambda_n = n^2\pi^2$. Then, since $y(t)$ is also dependent on λ_n, we see a correspondence between λ_n and solution functions, thus denote it $\phi_n(t)$; λ_n are our eigenvalues and $\phi_n(t)$ are the corresponding functions for each n. Explicitly, since $\lambda = n^2\pi^2$, we have $y_n(t) = 2ia_1 \sin n\pi t$. [3, p.631]

△

In a similar manner, eigenvalues can be obtained for various Sturm-Liouville problems. A sensible question to ask now is, does a series of eigenvalues always exist? We will answer this later.

2.2 Eigenfunctions

Immediately above, we have found that given n eigenvalues, then we must have n "corresponding" functions. These functions are called eigenfunctions denoted $\phi_n(x)$ as long as they are nontrivial solutions to the given differential equation.

Definition 2.2.1. The nonzero functions f that satisfy the differential equation $Df = \lambda f$ as in equation (2.2) are called **eigenfunctions**.

We have argued briefly that we can superposition multiple solutions in our assumption of what $y(t)$ is. Given n solutions, is there a way we can treat these solutions in a similar way, to ultimately express $y(t)$? We will answer this in the following section.

2.3 Existence of Eigenvalues and Eigenfunctions

In Chapter 2.1, we asked if there is always a series of eigenfunctions of a regular Sturm-Liouville problem, with boundary conditions (1.2). The answer is yes. We will prove this and state it as a theorem. To proceed, we newly introduce a system known as the Prüfer system. We consider the self-adjoint differential equation,

$$[P(x)y']' + Q(x)y = 0, \tag{2.7}$$

where $P(x) > 0$ is of class C^1 and $Q(x)$ is continuous along some given boundary $[a, b]$. We consider (2.7) with boundary conditions,

$$a_1 y(a) + a_2 y'(a) = 0\ ,\ b_1 y(b) + b_2 y'(b) = 0. \tag{2.8}$$

It is clear that if we let $P(x) = p(x)$, $Q(x) = \lambda r(x) - q(x)$ this is equivalent to (1.1) and taking $a = 0, b = 1$ reproduces the Sturm-Liouville problem (1.1), (1.2). Essentially, we wish to write (2.7), (2.8) in polar coordinates, since this will reduce (2.7) to a first order differential equation. New variables $r(x), \theta(x)$ are defined as

$$r^2 = y^2 + P^2 y^2 \ , \ \theta = \arctan \frac{y}{Py'}. \tag{2.9}$$

Subsequently the Prüfer substitution is given as,

$$P(x)y'(x) = r\cos\theta \ , \ y(x) = r\sin\theta. \tag{2.10}$$

Using these, we derive a system of differential equations equivalent to (1.1). Differentiating r, θ once, noting that $\cot\theta = \frac{Py'}{y}$, yields

$$-\text{cosec}^2\theta \frac{d\theta}{dx} = -Q(x) - \frac{1}{P(x)}\cot^2\theta, \tag{2.11}$$

$$\frac{dr}{dx} = \frac{1}{2}\Big[\frac{1}{P(x)} - Q(x)\Big] r\sin 2\theta. \tag{2.12}$$

We multiply equation (2.11) by $-\sin^2\theta$ and finally obtain

$$\frac{d\theta}{dx} = Q(x)\sin^2\theta + \frac{1}{P(x)}\cos^2\theta. \tag{2.13}$$

We say that system (2.12), (2.13) is equivalent to (2.7) in the sense that all nontrivial solutions of the system define a unique solution of (2.7), by the substitution (2.10), and conversely. Together, (2.12), (2.13) is called the Prüfer system.

Since $y = 0$ if and only if $\sin\theta = 0$ by (2.10), we have that the zeros of the solutions of (1.1) are $\theta = 0 \pm \pi, \pm 2\pi, ...$, where θ is the solution to the Prüfer equation (2.13),

$$\frac{d\theta}{dx} = \sin^2\theta[\lambda r(x) - q(x)] + \frac{1}{p(x)}\cos^2\theta, \tag{2.14}$$

associated to (1.1). The phase function $\theta(x, \lambda)$ will denote the solution to (2.14) that also satisfies the initial condition $\theta(a, \lambda) = \gamma$, $\forall\lambda$ where γ is found by the conditions $0 \leq \gamma \leq \pi$ and

$$\tan\gamma = \frac{y(a)}{p(a)y'(a)} = -\frac{a_2}{a_1 p(a)}. \tag{2.15}$$

Note that constants a_1, a_2 are constants as in the boundary conditions (2.8). Finally, we will state two theorems that we will use in what follows in this section.

Theorem 2.3.1. (Oscillation Theorem) The solution $\theta(x, \lambda)$ of (2.14) satisfying the initial condition $\theta(a, \lambda) = \gamma$ where $0 \leq \gamma \leq \pi$ for each λ is both a strictly increasing and continuous function of λ for some fixed x where $a < x \leq b$. Furthermore,

$$\lim_{\lambda\to\infty} \theta(x, \lambda) = \infty \ , \ \lim_{\lambda\to-\infty} \theta(x, \lambda) = 0 \tag{2.16}$$

where $a < x \leq b$.

Theorem 2.3.2. Let $n \geq 0$ be an integer. Then suppose that for some $x_n > a$ we have $\theta(x_n, \lambda) = n\pi$. Then $\theta(x, \lambda > n\pi)$, $\forall x > x_n$.

Now, the crux of this section is ready to be presented and proven.

Theorem 2.3.3. (Spectral theorem) Any regular Sturm-Liouville problem has an infinite sequence of real eigenvalues, and can be ordered according to increasing magnitude such that $\lambda_0 < \lambda_1 < ...$ with $\lim_{n\to\infty} \lambda_n = \infty$. Furthermore, the eigenfnctions corresponding to λ_n has exactly n zeros in the interval $a < x < b$.

Proof. We transform our boundary conditions (1.2) to equivalent conditions for $\theta(x, \lambda)$ introduced in our discussion of the Prüfer system. If $a_1 \neq 0$ then $\theta(x, \lambda)$ must satisfy the initial condition $\theta(x, \lambda) = \gamma$ for the smallest γ where $0 \leq \gamma \leq \pi$ such that $p(a) \tan \gamma = -\frac{a_2}{a_1}$. If $a_1 = 0$ we choose $\gamma = \frac{\pi}{2}$. For the remaining boundary b, we similarly find some $0 < \delta \leq \pi$ so that $p(b) \tan \delta = -\frac{b_2}{b_1}$.

Now, a solution $y(x)$ of (2.7) for the closed interval $a \leq x \leq b$ is an eigenfunction of the regular Sturm-Liouville problem given by imposing boundary conditions (2.8) if and only if the following holds for $\theta(x, \lambda)$,

$$\theta(a, \lambda) = \gamma \ , \ \theta(b, \lambda) = \delta + n\pi \ \text{,where } n = 0, 1, 2... \tag{2.17}$$

for γ, δ defined as above. Thus any value of λ that satisfies conditions (2.17) is an eigenvalue of the given regular Sturm-Liouville problem (2.7), (2.8) and vice versa. Now, let $\theta(x, \lambda)$ be a solution of (2.14) where the initial condition $\theta(a, \lambda) = \gamma$ is satisfied. If we consider $\theta(b, \lambda)$, this is an increasing function of λ and $\theta(b, \lambda) > 0$ by Theorem (2.3.2). Furthermore, as λ increases from $-\infty$ there is some value λ_0 which is the first to satisfy the condition $\theta(b, \lambda) = \delta + n\pi$ in (2.17). Namely, we have $n = 0$ as this is the first value to satisfy the condition so $\theta(b, \lambda_0) = \delta$. As λ increases, from λ_0, there exists an infinite sequence of λ_n where the condition $\theta(b, \lambda) = \delta + n\pi$ for some nonnegative integer n. Each of these values gives an eigenfunction

$$y_n(x) = r_n(x) \sin \theta(x, \lambda_n) \tag{2.18}$$

by (2.10), of the Sturm-Liouville problem. In addition, the eigenfunction belonging to λ_n has exactly n zeros in the interval $a < x < b$ by Theorem (2.3.1). □

2.4 Main Properties of Eigenfunctions and Eigenvalues

In relation to the Sturm-Liouville problem, a notable fact may be presented as a theorem. As aforementioned, Lagrange's identity comes to play for a proof.

Theorem 2.4.1. All eigenvalues of the Sturm-Liouville problem with boundary conditions given in equation (1.1), (1.2) are real.

Proof. Suppose that λ is a possibly complex eigenvalue of the Sturm-Liouville problem (1.1), (1.2) with corresponding eigenfunction ϕ, which is also possibly complex valued. If we let $\mu, \nu, U(x)$ and $V(x)$ be real, then let us write $\lambda = \mu + i\nu$ and $\phi(x) = U(x) + iV(x)$. Further, if $u = v = \phi$, then equation (1.6) becomes,

$$(L[\phi], \phi) = (\phi, L[\phi]). \tag{2.19}$$

By definition of L we have that $L[\phi] = \lambda r\phi$, and therefore, (2.19) can be written as,

$$(\lambda r\phi, \phi) = (\phi, \lambda r\phi). \tag{2.20}$$

Now, by definition of the inner product, (2.20) is,

$$\int_0^1 \lambda r(x)\phi(x)\bar{\phi}(x)dx = \int_0^1 \phi(x)\bar{r}(x)\bar{\lambda}\bar{\phi}(x)dx. \tag{2.21}$$

Since by initial assumption, $r(x)$ is real, rearranging (2.21) reduces to,

$$(\lambda - \bar{\lambda}) \int_0^1 r(x)\phi(x)\bar{\phi}(x)dx = 0. \tag{2.22}$$

Or equivalently,

$$(\lambda - \bar{\lambda}) \int_0^1 r(x)\big[U^2(x) + V^2(x)\big]dx = 0. \tag{2.23}$$

As $r(x) > 0$ at all points in $[0, 1]$, and $U(x), V(x)$ are real, the integrand in (2.23) is positive and not zero. It follows that $\lambda - \bar{\lambda} = 2i\nu$ is zero, therefore $\nu = 0$. Hence λ must be real. □

This theorem immediately implies that when considering the Sturm-Liouville boundary problem, one needs to look only for real eigenvalues. It is also true that the eigenfunctions of a Sturm-Liouville problem are all real.

Theorem 2.4.2. The eigenfunctions of the Sturm-Liouville problem (1.1), (1.2) are real.

Proof. Let λ be an eigenvalue and ϕ a corresponding eigenfunction of a regular Sturm-Liouville problem (1.1), (1.2). Now assume $\phi = U(x) + iV(x)$ for some real valued functions $U(x), V(x)$. By assumption we have $[p(x)\phi(x)']' - q(x)\phi(x) + \lambda r(x)\phi(x) = 0$. Expanding,

$$p'(U + iV)' + p(U + iV)'' - q(U + iV) + \lambda r(U + iV) = 0. \tag{2.24}$$

The real part of (2.24) is $p'U' + pU'' - qU + \lambda rU = 0$ and the imaginary part is $p'V' + pV'' - qV + \lambda rV = 0$. For (2.24) to hold, we need both parts to equal zero. However we notice that the parts can be written

$$[pV']' - qV + \lambda rV = 0 \ , \ [pU']'' - qU + \lambda rU = 0. \tag{2.25}$$

Namely, as in (1.1) with boundary conditions (1.2). Hence U, V are eigenfunctions corresponding to λ. Now, the Wronskian of U, V at the end point 0, is calculated as $W(U,V)(0) = U(0)V'(0) - U'(0)V(0)$. Given that U, V both satisfy (1.2),

$$W(U,V)(0) = -\frac{a_2}{a_1}U'V' + \frac{a_2}{a_1}U'V', \tag{2.26}$$

and thus zero. Therefore, U, V must be linearly dependent. We can therefore express U as a scalar multiple of V and vice versa. Let $V = c_1 U$ for some scalar $c_1 \in \mathbb{C}$. We then have $\phi = (1 + ic_1)V$, which hence shows that ϕ must be real, apart from an arbitrary multiplicative constant that may be complex. □

In the proof of (2.4.2), we have shown that given U, V are both eigenfunctions of λ, U, V are linearly dependent. The following thoeorem states another property of the eigenfunction in relation to linear independecy.

Theorem 2.4.3. For each eigenvalue of the Sturm-Liouville problem (1.1), (1.2) there corresponds a unique linearly independent eigenfunction; that is, it is simple.

Proof. See [25, p.13-14]. □

To further build on the understanding of fundamental properties of eigenfunctions and eigenvalues, we state and proof the following theorem.

Theorem 2.4.4. If ϕ_1, ϕ_2 are two eigenfunctions of the Sturm-Liouville problem (1.1), (1.2) with corresponding eigenvalues λ_1, λ_2 respectively where $\lambda_1 \neq \lambda_2$ then

$$\int_0^1 r(x)\phi_1(x)\phi_2(x)dx = 0 \tag{2.27}$$

where $r(x)$ is the weight function given in (1.1).

This theorem illustrates the notion of orthogonality of eigenfunctions with respect to $r(x)$. We say that ϕ_1, ϕ_2 are orthogonal with respect to $r(x)$ if they satisfy (2.27). The proof is as follows,

Proof. We note that ϕ_1, ϕ_2 satisfy the differential equations

$$L[\phi_1] = \lambda_1 r\phi_1 \ , \ L[\phi_2] = \lambda_2 r\phi_2, \tag{2.28}$$

respectively. Using equation (1.6),

$$(\lambda_1 r\phi_1, \phi_2) = (\phi_1, \lambda_2 r\phi_2). \tag{2.29}$$

Since $\lambda_2, r(x)$ and $\phi_2(x)$ are real, we have by definition of the inner product,

$$(\lambda_1 - \lambda_2)\int_0^1 r(x)\phi_1(x)\phi_2(x)dx = 0. \tag{2.30}$$

By assumption, $\lambda_1 \neq \lambda_2$ and therefore, (2.27) must be true. □

In accordance, we give a concept that often comes as a package with orthogonality.

Definition 2.4.1. Given an eigenfunction ϕ_n for some $n \in \mathbb{N}$, the eigenfunction is said to be **normalised** if it satisfies

$$\int_0^1 r(x)\phi_n^2(x)dx = 1, \tag{2.31}$$

This is known as the **normalisation condition**.

Definition 2.4.2. Kronecker's delta is the function δ_{mn} defined by,

$$\delta_{mn} = \begin{cases} 1 & \text{if } m = n \ , \\ 0 & \text{if } m \neq n \ . \end{cases} \tag{2.32}$$

where $m, n \in \mathbb{N}$.

Using Kronecker's delta, the normalization condition can be written as

$$\int_0^1 r(x)\phi_m(x)\phi_n(x)dx = \delta_{mn}. \tag{2.33}$$

Since by Theorem (2.4.4) all eigenfunctions of (1.1), (1.2) are orthogonal, if we can choose an arbitrary constant to multiply each eigenfunction to satisfy (2.33), we say that the eigenfunctions form an orthonormal set with respect to $r(x)$.

Example 2.4.1. We will determine the normalised eigenfunctions of example 2.1.1. We have found in example 2.1.1 that the eigenvalues are $\lambda_n = n^2\pi^2$ and corresponding eigenfunctions $\phi_n(x) = k_n \sin n\pi x$ for $n \in \mathbb{N}$. By definition, for each n we choose k_n so that

$$\int_0^1 (k_n \sin n\pi x)^2 dx = 1.$$

Computation leads to

$$\begin{aligned} k_n^2 \int_0^1 \sin^2 n\pi x dx &= k_n^2 \int_0^1 \left(\frac{1}{2} - \frac{1}{2}\cos 2n\pi x\right)dx \\ &= \frac{1}{2}k_n^2. \end{aligned}$$

Therefore, should we choose $k_n = \sqrt{2}$, $\forall n$ we see that the normalisation condition is satisfied. Hence the normalised eigenfunctions of this problem are $\phi_n(x) = \sqrt{2}\sin n\pi x$ where $n \in \mathbb{N}$. [3, p.634]

$\triangle$

Along with the numerous theorems and definitions given in this section, the Spectral theorem (2.3.3) in the previous section, will all come into play as we develop some important theories and results in Chapter 4, where we turn our focus to the nonhomogeneous Sturm-Liouville problem.

Chapter 3

Fourier Sine Series Expansion

3.1 Functions as a Series

The method of separation of variables is a commonly known technique for solving certain second order partial differential equations, including wave and heat conduction problems and Laplace's equation. By superposing the solutions obtained through this method, we are able to ultimately obtain a solution in the form of a series that satisfies particular initial or boundary conditions. Fourier sine series expansion comes into play when we wish to do this. Let us present a well-known result in the form of a theorem.

Theorem 3.1.1. Let $m, n \in \mathbb{N}$. Then, the definite integral of $\sin \frac{m\pi}{L} x \sin \frac{n\pi}{L} x$ along the interval $[0, L]$ is

$$\int_0^L \sin \frac{m\pi}{L} x \sin \frac{n\pi}{L} x dx = \begin{cases} \frac{L}{2} & \text{if } m = n \\ 0 & \text{if } m \neq n \ . \end{cases} \tag{3.1}$$

The proof of Theorem 3.1.1 is standard and thus omitted. Now we state the core theorem of this chapter:

Theorem 3.1.2. Any continuous function $f(x)$ that has piecewise continuous derivative $f'(x)$ along a finite interval $[0, L]$ can be expressed as

$$f(x) = \sum_{n=1}^{\infty} b_n \sin \frac{n\pi}{L} x, \tag{3.2}$$

where for $n \in \mathbb{N}$,

$$b_n = \frac{2}{L} \int_0^L f(x) \sin \frac{n\pi}{L} x \, dx. \tag{3.3}$$

While for functions $f(x, y)$ we have the double Fourier sine series expansion analogous to Theorem 3.1.2, we will not discuss that in this context as its relevance is scarce. Theorem 3.1.2 essentially tells us that when certain conditions are satisfied, any function $f(x)$ can be expressed as a series of sine functions. Visually, we are superposing countless sine functions with various amplitude and wave length to recreate the shape of $f(x)$.

3.2 Series of Eigenfunctions

Suppose we are given a function $f(x)$ that satisfies suitable conditions as described in the preceding section, that can be expanded in an infinite series of eigenfunctions. Then,

$$f(x) = \sum_{n=1}^{\infty} c_n \phi_n(x), \tag{3.4}$$

where $\phi_n(x)$ are eigenfunctions and thus satisfy (1.1), (1.2). We use the orthogonality property of eigenfunctions of the Sturm-Liouville problem (1.1), (1.2) to determine the coefficients in (3.4). By multiplying (3.4) by $r(x)\phi_m(x)$, for a fixed positive integer m, and then integrating along our interval $[0, 1]$, assuming term-by-term integration is valid, we have

$$\begin{aligned} \int_0^1 r(x)f(x)\phi_m(x)dx &= \sum_{n=1}^{\infty} c_n \int_0^1 r(x)\phi_m(x)\phi_n(x)dx \\ &= \sum_{n=1}^{\infty} c_n \delta_{mn}, \end{aligned} \tag{3.5}$$

which by the definition of δ_{mn} given in (2.32), implies

$$c_m = \int_0^1 r(x)f(x)\phi_m(x)dx, \tag{3.6}$$

where $m \in \mathbb{N}$. We finish this chapter with an extremely useful theorem,

Theorem 3.2.1. Let f and f' be piecewise continuous along the interval $0 \leq x \leq 1$. Then the series (3.4) with coefficients (3.6), where $\phi_1, \phi_2, ..., \phi_n$ are the normalised eigenfunctions of the Sturm-Liouville problem (1.1), (1.2), converges to $\frac{f(x+)+f(x-)}{2}$ at every point along $0 < x < 1$.

Proof. See [24, p.490-493] □

While it is apt to present Theorem (3.2.1) in this chapter where we discuss Fourier sine series expansion and how it is related to the eigenfunctions of the Sturm-Liouville problem (1.1), (1.2), we will not demonstrate nor discuss its use here. The theorem will be of great use when describing a different type of convergence of the series of eigenfunctions of the Sturm-Liouville problem, which is discussed in detail in Chapter 6. Now, we will turn to nonhomogeneous Sturm-Liouville problems.

Chapter 4

The Nonhomogeneous Boundary Value Problem

4.1 Motivation

We have seen that the solutions of the homogeneous Sturm-Liouville boundary value problem, known as the eigenfunctions, vary depending on the distinct value of the corresponding eigenvalues. We will now consider the nonhomogeneous Sturm-Liouville boundary value problem and explain the theorems with some examples. In this chapter, the nonhomogeneous Sturm-Liouville boundary value problem will be referred to as the nonhomogeneous problem for convenience.

4.2 Solving Nonhomogeneous Problems

Consider a nonhomogeneous differential equation with the boundary conditions (1.2). Note that there is an extra function $f(x)$ on the right hand side of the Sturm-Liouville problem (1.1),

$$L[y] = -[p(x)y']' + q(x)y = \mu r(x)y + f(x) \tag{4.1}$$

with a constant μ and the function f, given on $0 \leq x \leq 1$, and the boundary conditions (1.2).

We already know a homogeneous problem can be expressed as

$$L[y] = \lambda r(x)y \tag{4.2}$$

by using the linear differential operator. For the nonhomogeneous problem, (4.1), with boundary conditions (1.2), let $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n \leq ...$ and $\phi_1, \phi_2, ..., \phi_n, ...$ be the eigenvalues and corresponding normalised eigenfunctions of the problem, respectively, by Theorem (2.3.3).

Assume an unknown solution $y = \phi(x)$ of the nonhomogeneous problem (4.1), (1.2), can be expressed in a series

$$\phi(x) = \sum_{n=1}^{\infty} c_n \phi_n(x), \tag{4.3}$$

where $c_n = \int_0^1 r(x)\phi(x)\phi_n(x)dx$ for $n = 1, 2, ...$

Since we do not what ϕ is, we have to find the value of the c_n that satisfies the nonhomogeneous boundary condition. First, replacing y by ϕ in (4.1), we obtain

$$L[\phi] = \mu r(x)\phi(x) + f(x), \tag{4.4}$$

where ϕ is the solution of (4.1). Since $y = \phi(x)$, we can express the left hand side of (4.4)

$$L[\phi](x) = L\left[\sum_{n=1}^{\infty} c_n \phi_n\right](x) = \sum_{n=1}^{\infty} c_n L[\phi_n](x) = \sum_{n=1}^{\infty} c_n \lambda_n r(x)\phi_n(x). \tag{4.5}$$

Now we have the function $r(x)$ on the both sides of (4.4). It will be convenient to have the function $r(x)$ as a common factor for (4.4).Then we can take the nonhomogeneous part $f(x)$ in (4.1) as $r(x)[f(x)/r(x)]$. If the function f/r satisfies Theorem 3.2.1, then

$$\frac{f(x)}{r(x)} = \sum_{n=1}^{\infty} d_n \phi_n(x), \tag{4.6}$$

where

$$d_n = \int_0^1 r(x)\frac{f(x)}{r(x)}\phi_n(x)dx = \int_0^1 f(x)\phi_n(x)dx, \quad n = 1, 2, ... \tag{4.7}$$

Since we have all the terms of the series in (4.4), we get

$$\sum_{n=1}^{\infty} c_n \lambda_n r(x)\phi_n(x) = \mu r(x)\sum_{n=1}^{\infty} c_n \phi_n(x) + r(x)\sum_{n=1}^{\infty} d_n \phi_n(x) \tag{4.8}$$

by substitution. After collecting and cancelling the nonzero common factor $r(x)$, we have

$$\sum_{n=1}^{\infty}\Big[(\lambda_n - \mu)c_n - d_n\Big]\phi_n(x) = 0. \tag{4.9}$$

If (4.9) holds for each x in the interval $0 \le x \le 1$, then

$$(\lambda_n - \mu)c_n - d_n = 0, \quad n = 1, 2, ... \tag{4.10}$$

Now we have to consider three cases where $c_n \neq 0$:

1. For $\mu \neq \lambda_n$ where $n = 1, 2, ...,$

$$c_n = \frac{d_n}{\lambda_n - \mu} \qquad \Longrightarrow \qquad y = \phi(x) = \sum_{n=1}^{\infty} \frac{d_n}{\lambda_n - \mu}\phi_n(x), \tag{4.11}$$

where d_n is given by (4.7). This does not tell us whether or not the series converges but if any solution of (4.1), (1.2) satisfies the conditions of Theorem (3.2.1), then the series will converge at each point given that f is continuous.

2. For $\mu = \lambda_m$ and $d_m \neq 0$ where $n = m$, we cannot find c_m since (4.10) is unsolvable.

3. For $\mu = \lambda_m$ and $d_m = 0$ where $n = m$, c_m has an arbitrary multiple of the eigenfunction ϕ_m and the boundary value problem (4.1), (1.2) has a solution but it is not unique. Since $c_m = 0$,

$$\int_0^1 f(x)\phi_m(x)dx = 0. \tag{4.12}$$

Hence, if $\mu = \lambda_m$, the nonhomogeneous problem (4.1), (1.2) can be solved if f is orthogonal to the eigenfunction.

The theorem below summarises what we have looked through for the nonhomogeneous Sturm-Liouville problems. [5, p. 690].

Theorem 4.2.1. The nonhomogeneous problem (4.1), (1.2) has a unique solution (4.11) for each continuous f with distinct μ; the series converges for each x in $0 \leq x \leq 1$. If $\mu = \lambda_m$, then the nonhomogeneous problem has no solution unless f is orthogonal to ϕ_m.

Ivar Fredholm, a Swedish mathematician, stated the main part of the theorem in the following way:

Theorem 4.2.2. (Fredholm Alternative Theorem) For a given value of μ, if $\mu \neq \lambda_m$, for all m, then the nonhomogeneous problem (4.1), (1.2) has a unique solution for each continuous f. Otherwise, the homogeneous problem (4.2), (1.2) has a nontrivial solution.

4.3 Nonhomogeneous Heat Conduction

This section shows how the eigenfunction expansions can be used to solve nonhomogeneous problems for partial differential equations.

Consider the generalised heat conduction equation

$$r(x)u_t = [p(x)u_x]_x - q(x)u + F(x,t) \tag{4.13}$$

with the boundary conditions

$$u_x(0,t) - h_1u(0,t) = 0, \quad u_x(1,t) + h_2u(1,t) = 0 \tag{4.14}$$

and the initial condition

$$u(x,0) = f(x). \tag{4.15}$$

Let $u(x,t) = X(x)T(t)$ and $F(x,t) = 0$. Then (4.13) becomes

$$r(x)XT' = [p(x)X']'T - q(x)XT. \tag{4.16}$$

When we divide the both sides by $r(x)XT$, we obtain

$$\frac{T'}{T} = \frac{[p(x)X']'}{r(x)X} - \frac{q(x)}{r(x)} = -\lambda. \tag{4.17}$$

[5, p. 614]

Then $X(x)$ must be a solution of the boundary value problem

$$-[p(x)X']' + q(x)X = \lambda r(x)X \tag{4.18}$$

with the boundary conditions

$$X'(0) - h_1X(0) = 0, \quad X'(1) + h_2X(1) = 0. \tag{4.19}$$

If we assume p, q, and r are continuous and that p and r are always positive, then the problem (4.18), (4.19) is a Sturm-Liouville problem by Definition (1.1.1), and Theorem (2.3.3) applies.

We solve this problem in a similar manner as we did earlier in this chapter. Suppose

$$u(x,t) = \sum_{n=1}^{\infty} b_n(t)\phi_n(x) \tag{4.20}$$

is a series of eigenfunctions. The coefficient b_n must only depend on t, otherwise the function u would be a function of x only. Since each ϕ_n satisfies the boundary conditions (4.19), the boundary conditions (4.14) are automatically satisfied by (4.20).

Now substituting the linear combination (4.20) into (4.13) gives

$$\begin{aligned}\big[p(x)u_x\big]_x - q(x)u &= \frac{\partial}{\partial x}\Big[p(x)\sum_{n=1}^{\infty} b_n(t)\phi_n'(x)\Big] - q(x)\sum_{n=1}^{\infty} b_n(t)\phi_n(x)\\ &= \sum_{n=1}^{\infty} b_n(t)\{[p(x)\phi_n'(x)]' - q(x)\phi_n(x)\} \qquad (4.21)\\ &= -r(x)\sum_{n=1}^{\infty} b_n(t)\lambda_n\phi_n(x),\end{aligned}$$

where $[p(x)\phi_n'(x)] - q(x)\phi_n(x) = -\lambda_n r(x)\phi_n(x)$. For the left hand side of (4.13) we have

$$r(x)u_t = r(x)\frac{\partial}{\partial t}\sum_{n=1}^{\infty} b_n(t)\phi_n(x) = r(x)\sum_{n=1}^{\infty} b_n'(t)\phi_n(x). \tag{4.22}$$

Also it is convenient to look at the ratio $F(x,t)/r(x)$ for the term $F(x,t)$ since both terms $r(x)u_t$ and $[p(x)u_x]_x - q(x)u$ have a common factor, $r(x)$, in (4.13).

So we have

$$\frac{F(x,t)}{r(x)} = \sum_{n=1}^{\infty} \Gamma_n(t)\phi_n(x), \tag{4.23}$$

where

$$\Gamma_n(t) = \int_0^1 r(x)\frac{F(x,t)}{r(x)}\phi_n(x)dx = \int_0^1 F(x,t)\phi_n(x)dx, \quad n = 1, 2, ... \tag{4.24}$$

Then by substituting (4.8), (4.22) and (4.23) into (4.13), we obtain

$$r(x)\sum_{n=1}^{\infty} b_n'(t)\phi_n(x) = -r(x)\sum_{n=1}^{\infty} b_n(t)\lambda_n\phi_n(x) + r(x)\sum_{n=1}^{\infty} \Gamma_n(t)\phi_n(x). \tag{4.25}$$

We now can collect all the terms and cancel the common nonzero factor $r(x)$ in (4.25). We obtain

$$\sum_{n=1}^{\infty}[b_n'(t) + \lambda_n b_n(t) - \Gamma_n(t)]\phi_n(x) = 0. \tag{4.26}$$

If (4.26) holds for all x in the interval $0 < x < 1$, then $b_n(t)$ is a solution of the first order linear ordinary differential equation

$$b_n'(t) + \lambda_n b_n(t) = \Gamma_n(t), \quad n = 1, 2, ... \tag{4.27}$$

In order to determine $b_n(t)$, we must have an initial condition

$$b_n(0) = \alpha_n, \quad n = 1, 2, ... \tag{4.28}$$

for (4.27). We obtain this from (4.15). Setting $t = 0$ in (4.20) and using (4.15), we have

$$u(x,0) = \sum_{n=1}^{\infty} b_n(0)\phi_n(x) = \sum_{n=1}^{\infty} \alpha_n\phi_n(x) = f(x). \tag{4.29}$$

Hence the initial values α_n are the coefficients in the eigenfunction expansion for $f(x)$ where

$$\alpha_n = \int_0^1 r(x)f(x)\phi_n(x)dx, \quad n = 1, 2, ... \tag{4.30}$$

Since we know all the terms in the integrand, we consider α_n to be known.

Let the integrating factor be $\mu(t) = \exp(\lambda_n t)$. Then an initial value problem (4.27), (4.28) can be solved by method of integrating factors and becomes

$$b_n(t) = \alpha_n e^{-\lambda_n t} + \int_0^t e^{-\lambda_n(t-s)}\Gamma_n(s)ds, \quad n = 1, 2, ... \tag{4.31}$$

Hence an explicit solution of the boundary problem (4.13), (4.14), (4.15), is given by (4.20).

In summary, to solve a boundary value problem (4.13), (4.14), (4.15), we have to

1. Find the eigenvalues λ_n and the corresponding normalised eigenfunctions ϕ_n of the homogeneous problem (4.18), (4.19).
2. Calculate the coefficients α_n and $\Gamma_n(t)$ from the (4.30) and (4.24), respectively.
3. Evaluate the integral in (4.31) to find $b_n(t)$.
4. Calculate (4.20).

These steps may be difficult to follow. If the series (4.20) converges rapidly, we only need few terms to obtain an approximation to the solution. [5, p. 692-694].

4.4 Method of Eigenfunction Expansion Using Green's Formula

Consider the heat equation (4.13) when $r(x) = 1$, $p(x) = k$, and $q(x) = 0$. We have

$$\frac{\partial u}{\partial t} = k\frac{\partial^2 u}{\partial x^2} + F(x,t) \tag{4.32}$$

with the nonhomogeneous time dependent boundary conditions

$$u(0,t) = T_1(t), \quad u(1,t) = T_2(t) \tag{4.33}$$

and the initial conditions

$$u(x,0) = f(x). \tag{4.34}$$

We know how to solve this problem. First, find a function $v(0,t)$, which is the difference $w(0,t) = u(x,t) - v(x,t)$, which satisfies the homogeneous boundary conditions such that $v(0,t) = T_1(t)$, $v(1,t) = T_2(t)$. Then use the method of eigenfunction expansion to obtain $w(x,t)$ and $u(x,t) = w(x,t) + v(x,t)$.

Then we have

$$v(x,t) = T_1(t) + x[B(t) - A(t)], \tag{4.35}$$

where $v(0,t) = T_1(t)$ and $v(1,t) = T_2(t)$, and we obtain the solution $w(x,t)$ of the heat equation (4.32)

$$\frac{\partial w}{\partial t} = k\frac{\partial^2 w}{\partial x^2} + \left[F(x,t) - \frac{\partial v}{\partial t} + k\frac{\partial^2 v}{\partial x^2}\right] \equiv k\frac{\partial^2 w}{\partial x^2} + \overline{F}(x,t), \tag{4.36}$$

with homogeneous boundary conditions

$$w(0,t) = 0, \quad w(1,t) = 0, \tag{4.37}$$

and the initial condition

$$w(x,0) = u(x,0) - v(x,0) = f(x) - T_1(0) - x[T_2(0) - T_(0)] \equiv h(x). \tag{4.38}$$

Then we have to solve (4.36), (4.37) using the method of eigenfunction expansion. We will look at how this method can be applied directly to solve the problem (4.36), (4.37). Observe that term-by-term differentiation of Fourier series is not justified for the nonhomogeneous boundary conditions.

Consider a homogeneous problem

$$\frac{\partial z}{\partial t} = k\frac{\partial^2 z}{\partial x^2} \tag{4.39}$$

with the boundary condition

$$z(0,t) = 0, \quad z(1,T) = 0. \tag{4.40}$$

Using separation of variables, we obtain an eigenvalue problem

$$\frac{d^2\Phi}{dx^2} + \lambda\Phi = 0 \tag{4.41}$$

with the boundary condition

$$\Phi(0) = 0, \quad \Phi(1) = 0. \tag{4.42}$$

The eigenvalues for (4.41) are $\lambda_n = (n\pi/L)^2$ and the corresponding eigenfunctions are $\Phi_n = sin\frac{n\pi x}{L}$. We can expand any piecewise smooth function as a series of eigenfunctions

$$u(x,t) = \sum_{n=1}^{\infty} b_n(t)\Phi_n(x), \tag{4.43}$$

where the Fourier coefficients $b_n(t)$ are

$$b_n(t) = \frac{\int_0^L u\Phi_n dx}{\int_0^L \Phi_n^2 dx}. \tag{4.44}$$

Since $u(x,t)$ is not continuous at $x = 0$ and $x = 1$, we can only consider term-by-term differentiation with respect to time

$$\frac{\partial u}{\partial t} = \sum_{n=1}^{\infty} \frac{db_n}{dt}\Phi_n(x). \tag{4.45}$$

Then we obtain

$$\sum_{n=1}^{\infty} \frac{db_n}{dt}\Phi_n(x) = k\frac{\partial^2 u}{\partial x^2} + F(x,t) \tag{4.46}$$

by substituting (4.45) into (4.32). Now using the orthogonality of the eigenfunctions,

$$\frac{db_n}{dt} = \frac{\int_0^L \left[k\frac{\partial^2 u}{\partial x^2} + F(x,t)\right]\Phi_n(x)dx}{\int_0^L \Phi_n^2(x)dx}. \tag{4.47}$$

We want the coefficients $b_n(t)$ to be a first order differential equation. First, consider the generalised Fourier series,

$$F(x,t) = \sum_{n=1}^{\infty} f_n(t)\Phi_n(x), \tag{4.48}$$

where the coefficients $f_n(t)$ are

$$f_n(t) = \frac{\int_0^L F(x,t)\Phi_n(x)dx}{\int_0^L \Phi_n^2(x)dx}. \tag{4.49}$$

Then using (4.49), (4.47) can be written as

$$\frac{db_n}{dt} = f_n(t) + k\frac{\int_0^L \frac{\partial^2 u}{\partial x^2}\Phi_n(x)dx}{\int_0^L \Phi_n^2(x)dx}. \tag{4.50}$$

We can use Green's formula

$$\int_0^L \left(u\frac{\partial^2 v}{\partial x^2} - v\frac{\partial^2 u}{\partial x^2}\right)dx = \left[v\frac{\partial v}{\partial x} - v\frac{\partial u}{\partial x}\right]_0^L \tag{4.51}$$

to express the last term on the right hand side of (4.50). [7, p. 4-6].

Chapter 5

The Singular Sturm-Liouville Problem

5.1 Motivation

So far we have studied regular Sturm-Liouville problems of the form

$$-[p(x)y']' + q(x)y = \lambda r(x)y, \quad 0 < x < 1, \tag{5.1}$$

with boundary conditions

$$a_1 y(0) + a_2 y'(0) = 0, \tag{5.2}$$
$$b_1 y(1) + b_2 y'(1) = 0. \tag{5.3}$$

We will now move on to look at Sturm-Liouville problems that are classified as singular.

Definition 5.1.1. A differential equation (5.1) is considered to be a **singular Sturm-Liouville problem** if the following conditions are satisfied; that p is differentiable, q and r are continuous, and $p(x) > 0$ and $r(x) > 0$ at all points in the open interval $0 < x < 1$, but at least one of them does not satisfy these conditions at at least one of the two boundary points.

A Sturm-Liouville problem is also considered to be singular if the interval is unbounded, for example on the interval $0 \leq x < \infty$, but we have decided to focus on the classical problem in this report as numerous difficulties may arise when considering this particular set of problems.

Example 5.1.1. Legendre's equation is of the form

$$(1 - x^2)y'' - 2xy' + \lambda y = 0, \quad -1 < x < 1, \tag{5.4}$$

but can be written as a Sturm-Liouville problem in the form

$$-[(1 - x^2)y']' = \lambda y, \quad -1 < x < 1. \tag{5.5}$$

Here we have $p(x) = 1 - x^2$, $q(x) = 0$ and $r(x) = 1$. The eigenvalues of this problem are given by $\lambda = n(n + 1)$, which can easily be found by the power series method. The eigenfunctions for this particular equation are the Legendre polynomials $P_n(x)$. This problem is clearly singular as $p(1) = 0$. [3, p. 657] △

Example 5.1.2. Chebyshev's equation is of the form

$$(1 - x^2)y'' - xy' + \lambda y = 0, \quad -1 < x < 1, \tag{5.6}$$

but can be written as a Sturm-Liouville problem by the division of $\sqrt{1 - x^2}$ to give us

$$-\left[\frac{1}{\sqrt{1 - x^2}}y'\right]' = \frac{\lambda}{\sqrt{1 - x^2}}y, \quad -1 < x < 1. \tag{5.7}$$

Here we have $p(x) = (1-x^2)^{-1/2}$, $q(x) = 0$ and $r(x) = (1-x^2)^{-1/2}$. The eigenvalues of this problem are given by $\lambda = n^2$, found as above, and the eigenfunctions are Chebyshev polynomials of the first kind, $T_n(x)$. This is also a singular problem as $p(1) = 0$.[15] △

While both problems shown here are of great interest we will not be exploring either Legendre or Chebyshev polynomials, though a comprehensive exploration of these topics can be found elsewhere, see [16], [17]. We have decided to focus our report on Bessel's equation as we find this particular problem the most interesting and useful regarding practical applications of the Sturm-Liouville problem.

In order to make our main example of a singular Sturm-Liouville problem as simple to understand as possible, we must discuss Bessel's equation and subsequently, Bessel functions in some detail. Bessel's equation is a special type of Sturm-Liouville problem, as we will see momentarily, whilst Bessel functions are solutions of Bessel's equation. To give a full understanding of this section, we first give a brief introduction to the theory of singular points.

We begin with a differential equation of the form

$$S(x)y'' + T(x)y' + W(x)y = 0, \tag{5.8}$$

which can be written in standard form as

$$y'' + \frac{T(x)}{S(x)}y' + \frac{W(x)}{S(x)}y = 0. \tag{5.9}$$

We can let $\frac{T(x)}{S(x)} = s(x)$ and $\frac{W(x)}{S(x)} = t(x)$. If we are interested in a particular point, say $x = x_0$, where $S(x_0) = 0$, then it is clear that as $x \to x_0$, $s(x)$ and $t(x) \to \infty$. If this is the case, we call $x = x_0$ a singular point of the original differential equation, assuming at least one of $T(x)$ and $W(x)$ are nonzero at x_0. At a singular point, this equation is not analytic and so we cannot find a Taylor series expansion about the point x_0, in powers of $(x - x_0)$, for its solution.

There are two types of singular points: regular and irregular. When we say we have a regular singular point, we mean a point where the limits

$$\lim_{x \to x_0} (x - x_0)\frac{T(x)}{S(x)} \text{ and } \lim_{x \to x_0} (x - x_0)^2\frac{W(x)}{S(x)} \tag{5.10}$$

are finite. More generally, we can say that a point of any equation, polynomial or not, is a regular singular point if

$$(x - x_0)\frac{T(x)}{S(x)} \text{ and } (x - x_0)^2\frac{W(x)}{S(x)} \tag{5.11}$$

are analytic about $x = x_0$. A point is called an irregular singular point if the above conditions do not hold, see [6, p. 277-278]. We will be focusing on regular singular points in this report.

5.2 Bessel's Equation

Bessel's equation is a second order ordinary differential equation, and is of the form

$$x^2y'' + xy' + (x^2 - \nu^2)y = 0, \tag{5.12}$$

where ν is a constant. Here we have $S(x) = x^2, T(x) = x$ and $W(x) = x^2 - \nu^2$. In this section, we will only be discussing the case when $\nu = 0$, which corresponds to one 'case' of Bessel's equation. There are many cases of Bessel's equation, which take on different solutions for different values of ν, however for the purpose of this report, only the case when $\nu = 0$ is required.

The first thing we must do is check that $x = 0$ is a regular singular point of Bessel's equation. To do this, we consider Bessel's equation in standard form,

$$y'' + \frac{1}{x}y' + \frac{(x^2 - \nu^2)}{x^2}y = 0, \tag{5.13}$$

where we can now say $\frac{T(x)}{S(x)} = s(x) = \frac{1}{x}$, and $\frac{W(x)}{S(x)} = t(x) = \frac{x^2-\nu^2}{x^2}$.

It is clear to see that as $x \to 0, \frac{1}{x} \to \infty$. Thus, Bessel's equation is not analytic at $x = 0$ and so we cannot find a Taylor series representation for the solution about this point. Considering Bessel's equation in the form of (5.8), we can say

$$s_0 = \lim_{x\to 0} x\frac{T(x)}{S(x)} = \lim_{x\to 0} x\frac{1}{x} = 1, \tag{5.14}$$

$$t_0 = \lim_{x\to 0} x^2\frac{W(x)}{S(x)} = \lim_{x\to 0} x^2\frac{x^2 - \nu^2}{x^2} = -\nu^2. \tag{5.15}$$

Thus these limits exist and are finite, and so we can say $x = 0$ is a regular singular point of Bessel's equation. Then we can use the method of Frobenius to find the solutions. The method of Frobenius is useful in finding solutions for a second order ordinary differential equation with a regular singular point at $x = 0$. Bessel's equation satisfies the necessary criteria.

By this method, we will generate an infinite power series, and show that we have the indicial equation, $F(r)$, such that

$$F(r) = r(r - 1) + s_0r + t_0 = r(r - 1) + r - \nu^2 = r^2 - \nu^2 = 0. \tag{5.16}$$

The indicial equation is taken from the coefficient of the lowest power of x in the series. And so the roots of this equation are $r = \pm\nu$. As stated, we will consider the case when $\nu = 0$.

5.2.1 Solving Bessel's Equation of Order Zero

We say Bessel's equation is of order zero when $\nu = 0$. Then

$$L[y] = x^2y'' + xy' + x^2y = 0. \tag{5.17}$$

To find solutions for this equation, we use the method of Frobenius [14]. By this method, we assume there exists a solution of the form

$$y = \phi(r, x) = a_0x^r + \sum_{n=1}^{\infty} a_nx^{r+n}. \tag{5.18}$$

We differentiate this twice, substitute into (5.17) and then simplify to give

$$\begin{aligned} L[\phi](r, x) &= \sum_{n=0}^{\infty} a_n[(r+n)(r+n-1)+(r+n)]x^{r+n} + \sum_{n=0}^{\infty} a_nx^{r+n+2} \\ &= a_0[r(r-1)+r]x^r + a_1[(r+1)r+(r+1)]x^{r+1} + \\ &\sum_{n=2}^{\infty}\{a_n[(r+n)(r+n-1)+(r+n)] + a_{n-2}\}x^{r+n}. \end{aligned} \tag{5.19}$$

Since r is the lowest power of x in this series, we use the coefficient of x^r to give the indicial equation, $F(r)$. We know a_0 is nonzero so we can say

$$F(r) = r(r-1) + r = r^2 = 0. \tag{5.20}$$

Thus the roots of the indicial equation are $r_1 = r_2 = 0$. The next step in the method of Frobenius is to find a recurrence relation. Using the coefficient of x^{r+n}, as this contains both a_n and a_{n-2}, it follows that the recurrence relation is

$$a_n(r) = -\frac{a_{n-2}(r)}{(r+n)(r+n-1)+(r+n)} = -\frac{a_{n-2}(r)}{(r+n)^2}, \quad n \geqslant 2. \tag{5.21}$$

For both the indicial equation and the recurrence relation, it is worth mentioning that the reason we can equate the coefficients to zero is that they should be solutions for the differential equation.

Now, in order to find a solution for $y_1(x)$, we let $r = 0$ and then we see from (5.19) that for the coefficient of x^{r+1} to be zero, a_1 must be equal to zero. Then from (5.21) we can see that $a_3 = a_5 = a_7 = ... = 0$. We also have

$$a_n(0) = -\frac{a_{n-2}(0)}{n^2}, \quad n = 2, 4, 6..., \tag{5.22}$$

when $r = 0$. Then, substituting the first few even numbers in for n, and rearranging, we can see a pattern begins to form:

$$a_2(0) = -\frac{a_0}{2^2}, \quad a_4(0) = \frac{a_0}{2^4 2^2}, \quad a_6(0) = -\frac{a_0}{2^6(3 \cdot 2)^2}, \tag{5.23}$$

and generally we can let $n = 2m$ and say

$$a_{2m}(0) = \frac{(-1)^m a_0}{2^{2m}(m!)^2}, \quad m = 1, 2, 3, ... \tag{5.24}$$

And to give our first solution, we substitute these coefficients back into (5.18), which gives

$$y_1(x) = a_0 \Big[1 + \sum_{m=1}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m}(m!)^2}\Big], \quad x > 0. \tag{5.25}$$

We can abbreviate the function in brackets as $J_0(x)$, and we call this Bessel's function of the first kind of order zero.

Now we must find a second solution, $y_2(x)$. We can use the method of Frobenius to derive this, however to avoid the rigmarole, and because we have already discussed the method, we will simply state the second solution without the derivation:

$$y_2(x) = J_0(x) \ln x + \sum_{m=1}^{\infty} \frac{(-1)^{m+1} H_m}{2^{2m}(m!)^2} x^{2m}, \quad x > 0, \tag{5.26}$$

where $H_m = 1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{m}$. We want an expression for Bessel's function of the second kind of order zero, which we will denote by $Y_0(x)$. We consider $Y_0(x)$ to be a linear combination of $y_2(x)$ and $J_0(x)$.The current definition of Y_0 is known as the Weber function and is defined as

$$Y_0(x) = \frac{2}{\pi}[y_2(x) + (\gamma - \ln 2)J_0(x)], \tag{5.27}$$

where γ is the Euler-Máscheroni constant such that

$$\gamma = \lim_{n \to \infty} (H_n - \ln n) \cong 0.5772. \tag{5.28}$$

So if we substitute our expression for $y_2(x)$, we have

$$Y_0(x) = \frac{2}{\pi}\Big[\Big(\gamma + \ln \frac{x}{2}\Big) J_0(x) + \sum_{m=1}^{\infty} \frac{(-1)^{m+1} H_m}{2^{2m}(m!)^2} x^{2m}\Big], \quad x > 0. \tag{5.29}$$

For the full derivation, see [6, p. 298-299].

We can finally conclude that the general solution of the Bessel function of order zero with $x > 0$ is

$$y = c_1 J_0(x) + c_2 Y_0(x), \tag{5.30}$$

where c_1 and c_2 are constants. From Figure 5.1 we can see that $Y_0(x)$ has a singularity as $x \to 0$, and $J_0(x) \to 1$ as $x \to 0$. This means that we must have

$c_2 = 0$ if we seek solutions of Bessel's equation of order zero which are finite at the origin. We can also see that as x increases, both $J_0(x)$ and $Y_0(x)$ are oscillatory and decaying.

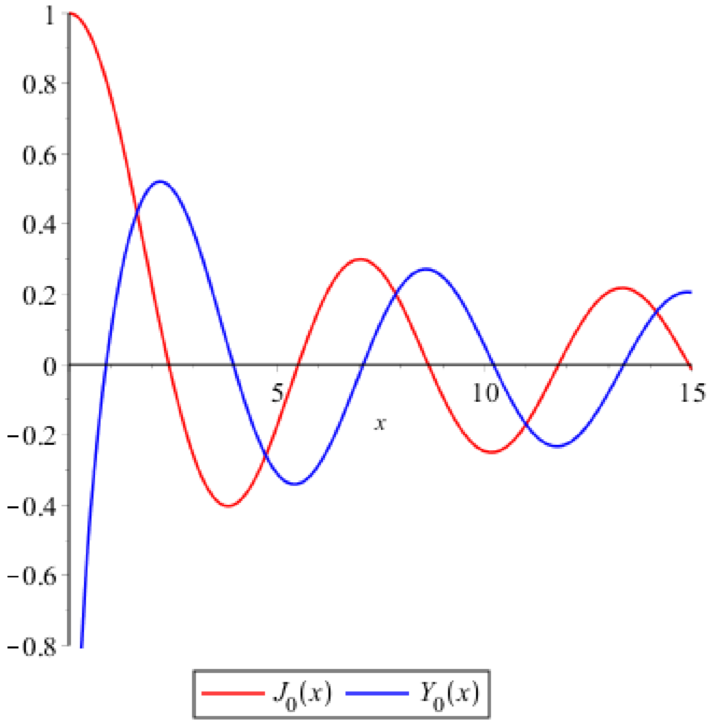

Figure 5.1: Graph showing Bessel's functions of order zero of the first and second kind (generated using Maple).

5.3 Generalising Bessel's Equation

As previously stated, we only give an in depth explanation of Bessel's equation of order zero in this report. However, it is possible to generalise Bessel's equation of the first kind to order n, see [8]. This is done using the method of Frobenius when $\nu = n$. The resulting solution is included as it is relevant to one of the applications discussed later in the report, however we omit its derivation. Bessel's function of the first kind of order n is

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+n}}{k!(k+n)!2^{2k+n}}. \tag{5.31}$$

5.4 Bessel's Equation as an Example of a Singular Problem

Now that we have established the theory of Bessel's equation of order zero, we may move on to showing its usefulness in dealing with singular Sturm-Liouville

problems. First we will show how Bessel's equation may be derived from a Sturm-Liouville problem.
The Sturm-Liouville problem

$$xy'' + y' + \lambda xy = 0, \tag{5.32}$$

also written as

$$-(xy')' = \lambda xy, \tag{5.33}$$

on the interval $0 < x < 1$, with $\lambda > 0$, can be reduced to the standard form of a Bessel equation by the substitution of the new independent variable $t = \sqrt{\lambda}x$, which leads to the equation

$$\frac{t}{\sqrt{\lambda}}\lambda\frac{d^2y}{dt^2} + \sqrt{\lambda}\frac{dy}{dt} + \lambda\frac{t}{\sqrt{\lambda}}y = 0. \tag{5.34}$$

So by multiplying through by $\frac{t}{\sqrt{\lambda}}$, we are left with the standard form of Bessel's equation of order zero,

$$t^2y'' + ty' + t^2y = 0. \tag{5.35}$$

We know the general solution to Bessel's equation of order zero is given by

$$y = c_1J_0(t) + c_2Y_0(t), \tag{5.36}$$

and so the general solution to our Sturm-Liouville problem is

$$y = c_1J_0(\sqrt{\lambda}x) + c_2Y_0(\sqrt{\lambda}x). \tag{5.37}$$

Although J_0 is bounded at zero, Y_0 is in fact singular at zero and so this causes a problem. Let us say that we seek a solution that satisfies the boundary conditions

$$y(0) = 0, \tag{5.38}$$
$$y(1) = 0. \tag{5.39}$$

As $J_0(0) = 1$ and $Y_0(x) \to -\infty$ as $x \to 0$, the first condition can only be satisfied for the trivial solution ie. $c_1 = c_2 = 0$.

Therefore we can see that the first boundary condition is in a sense too restrictive for our differential equation. Therefore we must modify the boundary condition at $x = 0$. Instead let us take the boundary condition as

$$y,\ y' \text{ bounded as } x \to 0. \tag{5.40}$$

This condition is satisfied for $c_2 = 0$ (which effectively eliminates our singular solution Y_0). Using our other boundary condition, $y(1) = 0$, we find that

$$J_0(\sqrt{\lambda}) = 0. \tag{5.41}$$

As the function J_0 is well tabulated, it is possible to find the roots of (5.41) in one such table. Therefore we can easily show that (5.41) has an infinite set of positive discrete roots which correspond to the eigenvalues $0 < \lambda_1 < \lambda_2 < ... < \lambda_n < ...$

of our problem. Thus the eigenfunctions are given by

$$\phi_n(x) = J_0(\sqrt{\lambda_n}x). \tag{5.42}$$

[3, p. 657-658]

5.5 General Problem

The methods used in this example can then be applied to the more general case of a singular Sturm-Liouville problem, so that relaxing the boundary conditions can lead to a singular problem having an infinite number of eigenvalues and eigenfunctions just as we have shown a regular Sturm-Liouville problem has.

The question remains as to how these boundary conditions can be relaxed and to what extent the eigenvalues and eigenfunctions of a singular problem compare to that of a regular problem. To answer this question, We focus on the identity

$$\int_0^1 L[u]v - uL[v]\, dx = 0. \tag{5.43}$$

If we allow (5.43) to be an improper integral, we can investigate for what conditions this identity shall still hold for singular problems. We then assume that $x = 0$ is a singular boundary point but that the point $x = 1$ is not. Therefore the boundary condition holds for our nonsingular boundary point and we purposefully do not define a boundary condition for $x = 0$ yet, as our goal here is to determine what boundary conditions we must employ in order for the identity to hold.

As our nonsingular boundary point is at $x = 0$, we define $\epsilon > 0$ and investigate the modified integral $\int_\epsilon^1 L[u]v$ by letting $\epsilon \to 0$.

We assume u and v have at least two continuous derivatives on the interval $\epsilon \leq x \leq 1$ and integrate twice by parts to find

$$\int_\epsilon^1 L[u]v - uL[v]\, dx = \Big[-p(x)[u'(x)v(x) - u(x)v'(x)]\Big]_\epsilon^1. \tag{5.44}$$

If both u and v satisfy the boundary conditions stated earlier then the boundary term at $x = 1$ is eliminated and the identity becomes

$$\int_\epsilon^1 L[u]v - uL[v]\, dx = p(\epsilon)[u'(\epsilon)v(\epsilon) - u(\epsilon)v'(\epsilon)]. \tag{5.45}$$

If we then take the limit as $\epsilon \to 0$ we get

$$\int_0^1 L[u]v - uL[v]\, dx = \lim_{\epsilon\to 0} p(\epsilon)[u'(\epsilon)v(\epsilon) - u(\epsilon)v'(\epsilon)]. \tag{5.46}$$

Therefore we can see that equation (5.43) holds, if and only if,

$$\lim_{\epsilon\to 0} p(\epsilon)[u'(\epsilon)v(\epsilon) - u(\epsilon)v'(\epsilon)] = 0, \tag{5.47}$$

for every pair of functions u and v.

This gives a clear criterion for the boundary conditions when the singular point occurs at $x = 0$. When the singular point is instead at $x = 1$, a similar process leads to the condition

$$\lim_{\epsilon \to 0} p(1-\epsilon)[u'(1-\epsilon)v(1-\epsilon) - u(1-\epsilon)v'(1-\epsilon)] = 0. \tag{5.48}$$

[3, p. 659-660]

5.6 Revisiting Self-Adjointness

Now we must revisit our definition for self-adjointness in order to include singular boundary value problems as well as the regular problems discussed earlier.

Definition 5.6.1. As stated earlier, the boundary value problem is said to be **self-adjoint** if equation (5.43) holds for each pair of functions u and v such that:

1. u and v are twice continuously differentiable on the open interval $0 < x < 1$,
2. at each regular boundary point u, v satisfy boundary condition (5.2),
3. u and v satisfy a boundary condition suitable such that they satisfy (5.47) if a singular boundary point occurs at the point $x = 0$ an (5.48) if a singular boundary point occurs at the point $x = 1$.

Therefore a singular Sturm-Liouville problem is defined if at least one of the boundary points is singular and if it satisfies two boundary conditions as detailed above. As an example, we can take a look back at Bessel's equation in its Sturm-Liouville form, as in (5.35). Here we have $p(x) = x$, so if both u and v satisfy the boundary condition specified in (5.40) at $x = 0$, we can see that (5.47) holds.

Therefore the Sturm-Liouville problem consisting of the equation (5.32), the boundary condition (5.40) at $x = 0$ and a boundary condition of the form (5.3) for the point $x = 1$ is self-adjoint.

5.7 Continuous Spectrum

What is considered the main difference between regular and singular Sturm-Liouville problems is that for a singular problem the eigenvalues may not always be discrete whereas the eigenvalues for a regular problem will be.

Definition 5.7.1. The problem is said to have a **continuous spectrum** when this occurs. This means the problem may have nontrivial solutions for every value of λ, or for every value of λ in a particular interval.

It is also possible for a problem of this form to have a mixture of both discrete eigenvalues and a continuous spectrum.

Of course it is also possible for a singular problem to have only discrete eigenvalues, as shown in the problem given by (5.32), (5.38) and (5.40).

In this case, we may use (5.43) just as we did in Chapter 2 to prove that the eigenvalues are real and that the eigenfunctions are orthogonal with respect to the weight function r. The expansion of a given function in terms of a series of eigenfunctions then follows as it did in Chapter 3. These expansions can be useful in solving nonhomogeneous boundary value problems, as in Chapter 4.

5.7.1 Returning to our Bessel Equation Example

To look into this further, let us look back at our earlier problem

$$-(xy')' = \lambda xy, \quad 0 < x < 1,$$

$$y,\, y' \text{ bounded as } x \to 0, \quad y(1) = 0.$$

We found that the eigenfunctions for this problem were given by $\phi_n(x) = J_0(\sqrt{\lambda_n}x)$. These eigenfunctions satisfy the orthogonality relation

$$\int_0^1 x\phi_m(x)\phi_n(x)\, dx = 0, \quad m \neq n \tag{5.49}$$

with respect to the weight function $r(x) = x$. Then, for a given function f, we assume that

$$f(x) = \sum_{n=1}^{\infty} c_n J_0(\sqrt{\lambda_n}x). \tag{5.50}$$

If we then multiply equation (5.50) by $xJ_0(\sqrt{\lambda_m}x)$ and integrate term-by-term from $x = 0$ to $x = 1$, we are left with

$$\int_0^1 xf(x)J_0(\sqrt{\lambda_m}x)\, dx = \sum_{n=1}^{\infty} c_n \int_0^1 xJ_0(\sqrt{\lambda_m}x)J_0(\sqrt{\lambda_n}x)\, dx. \tag{5.51}$$

Due to our orthogonality condition (5.49), the right hand side of equation (5.51) collapses into a single term. Therefore

$$c_m = \frac{\displaystyle\int_0^1 xf(x)J_0(\sqrt{\lambda_m}x)\, dx}{\displaystyle\int_0^1 xJ_0^2(\sqrt{\lambda_m}x)\, dx}, \tag{5.52}$$

which determines the coefficients in the series (5.50). [3, p. 660-661]

The convergence of the series (5.50) can be seen by extending Theorem 3.2.1 to cover this case. In a similar way, this theorem can be extended to other sets of Bessel functions, Legendre polynomials and for solutions of other singular Sturm-Liouville problems of interest. It should also be mentioned that most singular boundary value problems do not have all discrete eigenvalues, but possess some form of continuous spectra, which means series expansions of the form shown in Theorem 3.2.1 do not exist, but rather appropriate integrals must be used instead.

Chapter 6

Series of Orthogonal Functions: Mean Convergence

For a piecewise continuous f and f' on the interval $-L \leq x \leq L$ where

$$f(x) = \frac{a_0}{2} + \sum_{m=1}^{\infty} \Big(a_m \cos \frac{m\pi x}{L} + b_m \sin \frac{m\pi x}{L} \Big), \tag{6.1}$$

expanded series of eigenfunctions converges to f where f is continuous and to $\frac{[f(x+)+f(x-)]}{2}$ all points where f is discontinuous. [5, p. 595-597].

Now we shall see another practical type of convergence is more useful for series of orthogonal functions, such as eigenfunctions..

Suppose the given set of functions $\phi_1, \phi_2, ..., \phi_n$ are continuous on the interval $0 \leq x \leq 1$ and satisfy the orthonormality condition

$$\int_0^1 r(x)\phi_i(x)\phi_j(x)dx = \begin{cases} 0, & \text{if } i \neq j, \\ 1, & \text{if } i = j, \end{cases} \tag{6.2}$$

where r is a nonnegative weight function. Also, suppose a given function f defined on $0 \leq x \leq 1$ can be approximated by a linear combination of $\phi_1, \phi_2, ..., \phi_n$. Then

$$S_n(x) = \sum_{i=1}^{n} a_i \phi_i(x), \tag{6.3}$$

where $a_1, a_2, ..., a_n$, called Fourier coefficients, make the best approximation of f on the interval $0 \leq x \leq 1$. Using the mean square error, R_n, consider a function

$$R_n(a_1, a_2, ..., a_n) = \int_0^1 r(x)\big[f(x) - S_n(x)\big]^2 dx, \tag{6.4}$$

where $f(x)$ and $S_n(x)$ are shown in Figure 6.1.

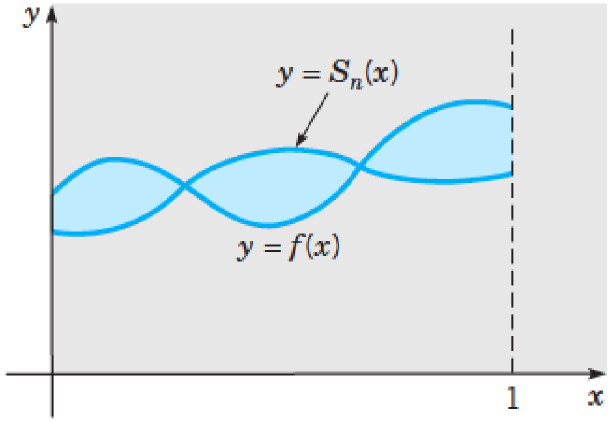

Figure 6.1: Approximation of $f(x)$ by $S_n(x)$. Source: [5, p. 718]

Note that $\partial S_n(x; a_1, a_2, ..., a_n)/\partial a_i = \phi_i(x)$. Then we must satisfy the necessary conditions

$$\frac{\partial R_n}{\partial a_i} = 0 \tag{6.5}$$

to choose the coefficients $a_1, a_2, ..., a_n$ which minimise R_n. When we write out (6.5), we obtain

$$-\frac{\partial R_n}{\partial a_i} = 2\int_0^1 r(x)[f(x) - S_n(x)]\phi_i(x)dx = 0. \tag{6.6}$$

Substituting $S_n(x)$ from (6.3) and using orthogonality relation (6.2), we obtain

$$a_i = \int_0^1 r(x)f(x)\phi_i(x)dx, \quad i = 1, 2, ..., n. \tag{6.7}$$

Since (6.5) is the only condition but not sufficient for R_n to be a minimum, we have to show R_n is minimised if $c_i = a_i$ for each i. (6.4) can be written as

$$R_n = \int_0^1 r(x)f^2(x)dx - \sum_{i=1}^{n} a_i^2 + \sum_{i=1}^{n}(c_i - a_i)^2, \tag{6.8}$$

where $S_n(x) = \sum_{n=1}^{n} c_i\phi_i(x)$ and c_i does not have to be from (6.7). [5, p.723] Also note that the Fourier coefficients (6.7) are the same as those in the eigenfunctions series that is stated in Theorem 3.2.1. Hence, S_n has the best mean sqaure approximation to $f(x)$.

Assume the Fourier coefficients in $S_n(x)$ (6.3) are given by (6.7). Then there are n formula for each a_i and those formula are independent of n. Recomputing all the coefficients for k terms, where $k > n$, is unnecessary. We only need to compute the terms $a_{n+1}, a_{n+2}, ..., a_k$ from the additional base functions $\phi_{n+1}, \phi_{n+2}, ..., \phi_k$.

Now suppose an infinite sequence of functions $\phi_1, \phi_2, ..., \phi_n, ...$ are continuous and orthonormal on $0 \leq x \leq 1$ and that R_n approaches zero as n increases. Then the infinite series

$$\sum_{i=1}^{\infty} a_i \phi_i(x) \tag{6.9}$$

is said to converge in the mean to $f(x)$. We should be aware that mean convergence is different from the pointwise convergence we have dealt with up until now.

We would like to know what class of functions on $0 \leq x \leq 1$ can be an infinite series of the orthonormal set ϕ_i for $i = 1, 2, ...$ The class of functions depends on the convergence we require.The set $\phi_1, \phi_2, ..., \phi_n, ...$ is said to be complete with respect to mean square convergence for a set if

$$f(x) = \sum_{i=1}^{\infty} a_i \phi_i(x) \tag{6.10}$$

converges in the mean.

Theorem (3.2.1) can be restated: The eigenfunctions of the Sturm-Liouville problem (5.1) with the boundary conditions (5.2), (5.3) are complete with respect to pointwise convergence for the set of continuous functions on $0 \leq x \leq 1$ which have piecewise continuous derivatives.

Then Theorem (3.2.1) can be generalised if we replace the pointwise convergence with mean convergence using the definition below.

Definition 6.0.1. A function f is square integrable on $0 \leq x \leq 1$ if f and f' are integrable on that interval. [5, p. 720].

Theorem 6.0.1. For the Sturm-Liouville problem (5.1),(1.2), the eigenfunctions ϕ_i are complete with respect to mean convergence for the square integrable set of functions on $0 \leq x \leq 1$. [5, p. 720].

Part II

Applications

Chapter 7

Application to Partial Differential Equations

In the real world, many physical problems involve more than one independent variable. These problems are modelled and solved using partial differential equations. The oldest method for solving these equations is the method of separation of variables. General second order linear partial differential equations are the partial differential equations with the most widely developed theory and the most important and diverse applications. There are three groups of these equations: the heat conduction equation, the potential equation, and the wave equation, the last of which we will be looking at closely in an example. In this example, we will be looking into the waves caused by the vibrations of a circular elastic membrane.

We want to use singular Sturm-Liouville theory to expand on the method of separation of variables, in order to apply it to a wider range of problems. For example, problems involving a variety of geometrical regions, varying boundary conditions or different kinds of differential equations. Our example will lead us to a singular Sturm-Liouville problem after applying the method of separation of variables, which we can then solve.

7.1 Vibrations of a Circular Elastic Membrane

The two dimensional wave equation,

$$a^2(u_{xx} + u_{yy}) = u_{tt}, \tag{7.1}$$

where a is a real constant, describes the behaviour of the transverse vibrations of an elastic membrane. To apply this to a circular membrane, it makes sense to write this equation in polar coordinates. Hence, we have

$$a^2(u_{rr} + \frac{1}{r}u_r + \frac{1}{r^2}u_{\theta\theta}) = u_{tt}. \tag{7.2}$$

We can make several assumptions about our membrane. We assume it is displaced by an angle θ from its neutral position, at time $t = 0$, at which point it is released. For simplicity, we can assume the membrane has unit radius and is firmly attached around its circumference. And due to the circular symmetry of its initial and boundary conditions, it can be said that u is independent of θ. Then (7.2) becomes

$$a^2(u_{rr} + \frac{1}{r}u_r) = u_{tt}, \quad 0 < r < 1, \quad t > 0. \tag{7.3}$$

At $r = 1$, we have the boundary condition

$$u(1,t) = 0, \quad t \geq 0, \tag{7.4}$$

and we have the initial conditions

$$u(r,0) = f(r), \quad 0 \leq r \leq 1, \tag{7.5}$$

$$u_t(r,0) = 0, \qquad 0 \leq r \leq 1, \tag{7.6}$$

where $f(r)$ is the function describing the initial position of the membrane. It is necessary that $f(1) = 0$, and we must say that $u(r,t)$ is bounded when r is between zero and one. We then assume that $u(r,t) = R(r)T(t)$ and substitute in to (7.3) to give

$$\frac{R'' + \frac{1}{r}R'}{R} = \frac{1}{a^2}\frac{T''}{T} = -\lambda^2. \tag{7.7}$$

We have written our separation constant as $-\lambda^2$ as we expect it to be a negative constant, written this way we know that $\lambda > 0$. (7.7) then provides us with the two ordinary differential equations

$$r^2R'' + rR' + \lambda^2r^2R = 0, \tag{7.8}$$

$$T'' + \lambda^2a^2T = 0. \tag{7.9}$$

We can easily solve (7.9) to give us

$$T(t) = k_1 \sin \lambda at + k_2 \cos \lambda at. \tag{7.10}$$

To solve (7.8) we must first introduce the new independent variable $\xi = \lambda r$. Once this is substituted in, we can see that (7.8) is reduced to

$$\xi^2\frac{d^2R}{d\xi^2} + \xi\frac{dR}{d\xi} + \xi^2R = 0, \tag{7.11}$$

which is none other than Bessel's equation of order zero. Therefore we know that

$$R = c_1J_0(\xi) + c_2Y_0(\xi), \tag{7.12}$$

which in terms of our original variable is

$$R = c_1J_0(\lambda r) + c_2Y_0(\lambda r). \tag{7.13}$$

We must have that R remains bounded as $r \to 0$ due to the boundedness condition on $u(r,t)$. As $Y_0(\lambda r) \to -\infty$ as $r \to 0$ we must have $c_2 = 0$. Our boundary condition (7.4) leaves us with the equation

$$J_0(\lambda) = 0. \tag{7.14}$$

From our previous section we know that this equation has an infinite set of discrete positive roots, given by $\lambda_1, \lambda_2, ..., \lambda_n, ...$ and that the functions $J_0(\lambda_n r)$ are the eigenfunctions of a singular Sturm-Liouville problem, and therefore can be used as a basis for a series expansion for our given function f. Therefore the fundamental solution to this problem is given by the following:

$$u_n(r,t) = J_0(\lambda_n r)\,\sin\lambda_n at, \quad n = 1, 2, ... \tag{7.15}$$

$$v_n(r,t) = J_0(\lambda_n r)\,\cos\lambda_n at, \quad n = 1, 2, ... \tag{7.16}$$

We then assume that our function $u(r,t)$ can be expressed as an infinite linear combination of the solutions (7.15), (7.16) as follows:

$$\begin{aligned} u(r,t) &= \sum_{n=1}^{\infty}[k_n u_n(r,t) + c_n v_n(r,t)] \\ &= \sum_{n=1}^{\infty}[k_n J_0(\lambda_n r)\sin\lambda_n at + c_n J_0(\lambda_n r)\cos\lambda_n at]. \end{aligned} \tag{7.17}$$

Our initial conditions (7.5) and (7.6) give us:

$$u(r,0) = \sum_{n=1}^{\infty} c_n J_0(\lambda_n r) = f(r), \tag{7.18}$$

and

$$u_t(r,0) = \sum_{n=1}^{\infty} \lambda_n a k_n J_0(\lambda_n r) = 0. \tag{7.19}$$

Then using equation (5.52) from chapter 5 we have

$$k_n = 0, \quad c_n = \frac{\displaystyle\int_0^1 rf(r)J_0(\sqrt{\lambda_n}r)\ dr}{\displaystyle\int_0^1 r[J_0(\sqrt{\lambda_n}r)]^2\ dr}; \quad n = 1, 2, ... \tag{7.20}$$

Therefore we can see that the solution to the differential equation (7.3) that satisfies the conditions (7.4),(7.5) and (7.6) is given by

$$u(r,t) = \sum_{n=1}^{\infty} c_n J_0(\lambda_n r)\cos\lambda_n at, \tag{7.21}$$

where the coefficients c_n are as defined in (7.20). [3, p. 664-666]

7.2 Conclusion

As important as the method of separation of variables is, its use is quite restricted. The problem at hand must be linear as we require linear combinations of solutions to give rise to additional solutions. Also, it can be difficult to solve the resulting ordinary differential equations, so even if it is possible to separate the variables, it may not be possible to go further than the first step when trying to obtain solutions. Another problem is that we must be using a coordinate system for which the method can be applied, and such that the partial differential equation can be replaced by a set of ordinary differential equations. Further, it is necessary to consider the boundary of the region focused on. The boundary must be made up of surfaces or curves where at least one variable is constant.

This section gives just one example as to how Bessel functions and their applicability to Sturm-Liouville theory can be applied to a partial differential equation to help solve it.

Chapter 8

Kepler's Equation

8.1 Motivation

Kepler's equation is an application using the Sturm-Liouville problem which has been around since the ancients observed the motions of the planets among the stars. Copernicus, a Renaissance mathematician and astronomer, later discovered the planets were rotating around the sun. It took more time after this to understand how the planets rotate around the sun and with what motion.

In the fifteenth century, while there were great debates about geocentric and heliocentric theory, Tycho Brahe, a Danish nobleman, had an idea that the nature of motions of the planets would be solved if they could measure the actual positions of the planets in the sky sufficiently accurately. So he studied the positions of the planets in his observatory on the island of Hven, near Copenhagen for many years.

The notes taken while Tycho Brahe was studying the planets were studied by the German mathematician Johannes Kepler, after Tycho's death. Kepler found remarkable laws regarding planetary motion which are known as Kepler's laws. One of these laws gives rise to Kepler's equation, which is related to the Sturm-Liouville problem. We will discuss how Bessel's equation, Bessel functions and Kepler's equation are connected. [13, p. 1]

8.2 Kepler's Equation

Kepler's equation uses Kepler's second law that a planet sweeps out equal area in equal times.

The planet, P, travels via an ellipse in the counter clockwise direction. Let T be the period P travels and M be the mean anomaly, where anomaly(angle) means irregularity [12]. Now denote

$$M = \frac{2\pi t}{T}, \tag{8.1}$$

where time $t \in (0, T)$.

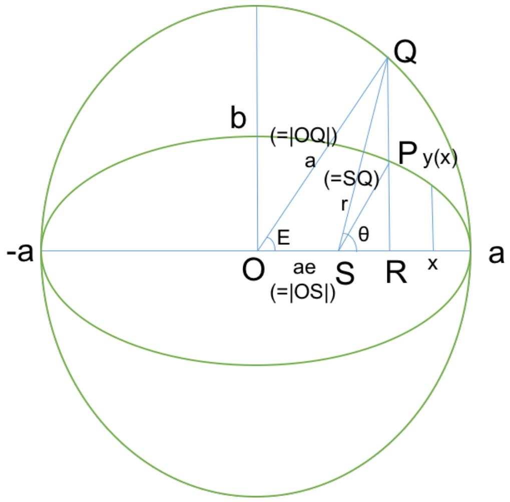

Figure 8.1: The geometric relationship between the eccentric anomaly E and the true anomaly θ. Source: [11, p. 2]

From Figure (8.1), $|OS| = ae$ and $|OR| = a\cos E$ where e is the eccentricity of the ellipse. We also know πab is the area of an ellipse and the sector SPa becomes

$$SPa = \pi ab\frac{t}{T} = \frac{ab}{2}\frac{2\pi t}{T} = \frac{ab}{2}M. \tag{8.2}$$

Observe the point P, with the coordinates $(a\cos E, y_p)$, is circling the ellipse and the point R has the coordinates $(a\cos E, 0)$. Recall the ellipse function

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{8.3}$$

and $|PR| = y_p = b\sqrt{1 - (\frac{a\cos E}{a})^2} = b\sin E$.

Then the area of the triangle SPR becomes

$$\begin{aligned}\triangle SPR &= \frac{1}{2}|PR||SR| \\ &= \frac{1}{2}(b\sin E)(|OR| - |OS|) \\ &= \frac{1}{2}(b\sin E)(a\cos E - aE) \\ &= \frac{ab}{2}(\sin E\cos E - e\sin E).\end{aligned} \tag{8.4}$$

By computing the area under the function $y(x) = b\sqrt{1 - \frac{x^2}{a^2}}$ for $x \in (a\cos E, a)$, the area of the sector RPa, where $t = \frac{x}{a}$, becomes

$$\begin{aligned} RPa &= b\int_{a\cos E}^{a}\sqrt{1-\frac{x^2}{a^2}}dx \\ &= ab\int_{\cos E}^{1}\sqrt{1-t^2}dt \\ &= \frac{ab}{2}\Big(t\sqrt{1-t^2}+\arctan\frac{t}{\sqrt{1-t^2}}\Big)\Big|_{\cos E}^{1} \\ &= \frac{ab}{2}\Big(-\cos E\sin E+\frac{\pi}{2}-\arctan\frac{\cos E}{\sin E}\Big) \\ &= \frac{ab}{2}\Big(-\sin E\cos E+\frac{\pi}{2}-\Big(-E+\frac{\pi}{2}\Big)\Big) \\ &= \frac{ab}{2}\Big(-\sin E\cos E+E\Big). \end{aligned} \tag{8.5}$$

By adding the triangle SPR and the sector RPa, the sector SPa becomes

$$\begin{aligned} SPa &= \triangle SPR + RPa \\ &= \frac{ab}{2}(\sin E\cos E - e\sin E) + \frac{ab}{2}\Big(-\sin E\cos E+E\Big) \\ &= \frac{ab}{2}(E-e\sin E). \end{aligned} \tag{8.6}$$

After substituting (8.6) into the left hand side of (8.2) and cancelling the common factor $\frac{ab}{2}$, we obtain Kepler's equation [9]

$$M = E - e\sin E. \tag{8.7}$$

8.3 Kepler's Equation and Bessel Functions

Now recall Bessel's function of the first kind of order n, (5.31). Joseph Louis Lagrange showed that an equation of the form

$$w = a + t\phi(w) \tag{8.8}$$

has a solution for w and a function f under suitable conditions. Then it is true that

$$f(w) = f(a) + \sum_{n=1}^{\infty}\frac{t^n}{n!}\frac{d^{n-1}}{da^{n-1}}f'(a)[\phi(a)]^n. \tag{8.9}$$

Kepler's equation is of this form with $f(z) = z$, $\phi(z) = \sin z$, $a = M$, $t = e$ and $w = E$, Then (8.9) becomes

$$E = M + \sum_{n=1}^{\infty}\frac{e^n}{n!}\frac{d^{n-1}}{dM^{n-1}}(\sin^n M) = M + \sum_{n=1}^{\infty}a_n(M)e^n. \tag{8.10}$$

While Lagrange used repeated differentiation, Bessel used integration for a solution of Kepler's equation

$$E = M + \sum_{n=1}^{\infty} b_n(e) \sin nM, \tag{8.11}$$

which is similar to a Fourier sine series. Now we explain how Bessel solved Kepler's equation. If $E = g(M)$ is the solution of Kepler's equation, then g has $M = 0, \pi$ as its fixed points. Also, for the interval $0 \leq M \leq \pi$, if

$$g(M) - M = \sum_{n=1}^{\infty} b_n(e) \sin nM, \tag{8.12}$$

then

$$b_n(e) = \frac{2}{\pi} \int_0^{\pi} [g(M) - M] \sin nM dM = \frac{2}{n\pi} \int_0^{\pi} \cos nM dg(M). \tag{8.13}$$

We know that $M = E - e \sin E = g(M) - e \sin(g(m))$. Then $b_n(e)$ becomes

$$b_n(e) = \frac{2}{n\pi} \int_0^{\pi} \cos(nE - ne \sin E) dE. \tag{8.14}$$

It also can be expressed in Bessel function notation, $J_n(x)$

$$J_n(x) = \frac{1}{\pi} \int_0^{\infty} \cos(nE - ne \sin E) dE \tag{8.15}$$

and Bessel's solution of Kepler's equation is

$$E = M + \sum_{n=1}^{\infty} \left(\frac{2}{n} J_n(ne) \right) \sin nM. \tag{8.16}$$

In a landmark paper of 1824, Bessel showed that (5.31) satisfies (8.15) which describes the same function. [8, p. 46-48]

8.4 Conclusion

We have seen that Kepler's equation, $M = E - e \sin E$, comes from Kepler's second law. There seems to be no relevance to the Sturm-Liouville problem, but we have seen Kepler's equation using Bessel functions so there is a relationship between Bessel functions and the Sturm-Liouville problem. Putting certain functions into the Bessel function and integrating, we found Bessel's solution of Kepler's equation. There are many ways to solve Kepler's equation such as by iteration or Newton's method. We did not discuss these in this chapter as we wish to focus on the relationship with the Sturm-Liouville problem. Kepler's equation is beautiful as it describes planetary motion in a single equation. In this section we have given another layer of significance to both Kepler's equation and the Sturm-Liouville problem by linking one to the other.

Chapter 9

The Korteweg-de Vries Equation

9.1 Motivation

Discovered by John Scott Russel as a peculiar form of wave in the suburbs of Edinburgh in 1834, solitons and relevant theories have recently started to attract attention from many areas of study [22, p. 463]. Solitons are extremely stable solitary waves, meaning they possess only a single crest; that is, the amplitude, wave length and speed do not change over time, nor when it collides with another wave. In this chapter we discuss the Korteweg-de Vries 9(KdV) equation,

$$u_t - 6uu_x + u_{xxx} = 0, \tag{9.1}$$

where $u(x,t)$ describes the perpendicular displacement of a soliton at position x and time t. Loosely, the "height" of the soliton which we shall refer to as the potential of the wave. Its applications have been noted in various areas of study, notably in relation with understanding properties of tsunamis and the designing of oceanic structures. We will explore a method known as the inverse scattering method in relation with the Sturm-Liouville equation and the properties of associated eigenfunctions and eigenvalues in this chapter.

9.2 The Sturm-Liouville Equation as the KdV Equation

Consider (1.1) in a one dimensional case where $p(x) = r(x) = 1$. For convenience in this section, replacing $q(x)$ in (1.1) with $u(x)$, (1.1) can be written

$$y'' + [\lambda - u(x)]y(x) = 0. \tag{9.2}$$

For simplicity, we will discuss the one dimensional case only. Henceforth $u(x)$ will denote the potential and is taken to be real. Now, if $y(x), u(x)$ and λ are dependent on a time parameter t, we have

$$y_{xx} + [\lambda(t) - u(x,t)]y(x,t) = 0. \tag{9.3}$$

We first show that there are infinitely many PDEs of $u(x,t)$ with respect to t and that when λ does not depend on t, one such equation is the KdV equation (9.1) [20]. Let us denote operators

$$D = \frac{\partial}{\partial x}\ ,\ L_1 = -D^2 + u(x,t). \tag{9.4}$$

We mention that L_1 is equivalent to (1.3) when $p(x) = 1$. Since $L_1 y = \lambda y$, differentiation with respect to t yields

$$(L_1 y)_t = \lambda_t y + \lambda y_t. \tag{9.5}$$

As $\lambda y = -y_{xx} + uy$ by (9.3), direct computation of $(L_1 y)_t$ yields

$$(-y_{xx} + uy)_t = L_1 y_t + u_t y, \tag{9.6}$$

thus by substituion of (9.6) to the left hand side of (9.5), we obtain

$$L_1 y_t + u_t y = \lambda_t y + \lambda y_t. \tag{9.7}$$

Recall that we wish to consider the case where λ is not time dependent, namely when $\lambda_t = 0$. To do this we let $y_t = \beta y$ for some differential operator β. Then, (9.7) can be written

$$L_1 \beta y + u_t y = \lambda_t y + \beta L_1 y. \tag{9.8}$$

Rearranging gives

$$(u_t + L_1\beta - \beta L_1)y = \lambda_t y. \tag{9.9}$$

If $\lambda_t = 0$ then,

$$(u_t + L_1\beta - \beta L_1)y = 0. \tag{9.10}$$

In other words, if we can set β to satisfy (9.10) we have $\lambda_t = 0$. Assume

$$\beta = aD^3 + fD + g \tag{9.11}$$

for some constant a, functions $f(x,t), g(x,t)$. Calculate $L_1\beta y$, $\beta L_1 y$

$$\begin{aligned} L_1\beta y &= (-D^2 + u)(aD^3 + fD + g)y \\ &= -(aD^5 + f_{xx}D + 2f_x D^2 + fD^3 + g_{xx} + 2g_x D + gD^2)y \\ &\quad + u(aD^3 + fD + g)y. \end{aligned} \tag{9.12}$$

$$\begin{aligned} \beta L_1 y &= (aD^3 + fD + g)(-D^2 + u)y \\ &= -aD^5 - fD^3 y + guy + a(u_{xxx} + 3u_{xx}D + 3u_x D^2 + uD^3)y \\ &\quad + f(u_x + uD)y - gD^2 y. \end{aligned} \tag{9.13}$$

Therefore by (9.12), (9.13)

$$\begin{aligned} L_1\beta y - \beta L_1 y = &-(2f_x + 3au_x)D^2 y \\ &- (f_x x + 2g_x + 3au_{xx})Dy - (g_x x + au_{xxx} + fu_x)y. \end{aligned} \tag{9.14}$$

If we let the coefficients of D^2y and Dy be zero we obtain the system

$$2f_x + 3au_x = 0, \tag{9.15}$$

$$f_{xx} + 2g_x + 3au_{xx} = 0. \tag{9.16}$$

We first find f by integration with respect to x and substitute into (9.16) and find

$$f = -\frac{3}{2}au + c_1(t) \tag{9.17}$$

$$g = -\frac{3}{4}au_x + c_2(t). \tag{9.18}$$

Where $c_1(t), c_2(t)$ are some functions of t. By (9.17) and (9.18), (9.14) becomes

$$L_1\beta y - \beta L_1 y = \Big\{\frac{1}{4}a(6uu_x - u_{xxx}) - c_1u_x\Big\}y. \tag{9.19}$$

For λ to be time independent, we require the coefficient of (9.10) to equal zero. Introduce new variables $\hat{x} = x + c_1t$, $\hat{t} = t$ to eliminate $c_1(t)$. Then,

$$\frac{\partial}{\partial x} = \frac{\partial \hat{x}}{\partial x}\frac{\partial}{\partial \hat{x}} + \frac{\partial \hat{t}}{\partial x}\frac{\partial}{\partial \hat{t}} = \frac{\partial}{\partial \hat{x}}, \tag{9.20}$$

$$\frac{\partial}{\partial t} = \frac{\partial \hat{x}}{\partial t}\frac{\partial}{\partial \hat{x}} + \frac{\partial \hat{t}}{\partial t}\frac{\partial}{\partial \hat{t}} = c_1\frac{\partial}{\partial \hat{x}} + \frac{\partial}{\partial \hat{t}}. \tag{9.21}$$

By (9.20) and (9.21) and rewriting the coefficient of y in (9.19) in terms of $\hat{x}, \hat{t}$ becomes

$$u_{\hat{x}} + \frac{1}{4}a(u_{\hat{x}\hat{x}\hat{x}} - 6uu_{\hat{x}}) = 0. \tag{9.22}$$

Since a is an arbitrary constant, replacing $\hat{x}, \hat{t}$ by x, t and considering the case where $a = -4$ in (9.22) gives us the KdV equation

$$u_t - 6uu_x + u_{xxx} = 0. \tag{9.23}$$

This demonstrates that when $u(x,t)$ of the Sturm-Liouville equation (9.2) satisfies the KdV eqution, its eigenvalue λ does not depend on t.

9.3 Fundamental Solutions

For later convenience let $\lambda = k^2$. Then (9.2) can be written as

$$y'' + [k^2 - u(x)]y(x) = 0. \tag{9.24}$$

Clearly the solution to (9.24) can be written as a linear combination of two fundamental solutions. Since $u(x)$ is the potential, we assume it converges to zero at $|x| \to \infty$. Considering (9.24) at $|x| \to \infty$,

$$y'' + k^2 y = 0, \tag{9.25}$$

thus solutions $y(x)$ can be written as linear combinations of $e^{\pm ikx}$. Now, consider solutions $f_1(x,k), f_2(x,k)$ such that $\lim_{x\to\infty} f_1(x,k) = e^{ikx}$ and $\lim_{x\to-\infty} f_2(x,k) = e^{-ikx}$. If we consider (9.24) as a nonhomogeneous Sturm-Liouville equation, then

$$y'' + k^2 y = u(x)y. \tag{9.26}$$

We let the solution be written in the form $y(x) = A(x)e^{ikx} + B(x)e^{-ikx}$. Then,

$$y'' = (A'e^{ikx} + B'e^{-ikx})' + A'(e^{ikx})' + A(e^{ikx})'' + B'(e^{ikx})' + B(e^{ikx})''. \tag{9.27}$$

We substitute (9.27) to (9.26) which is thus rewritten as,

$$y'' + k^2 y = (A'e^{ikx} + B'e^{-ikx})' + A'(e^{ikx})' + B'(e^{ikx})'. \tag{9.28}$$

Then our nonhomogeneous equation (9.26) can be written

$$(A'e^{ikx} + B'e^{-ikx})' + A'(e^{ikx})' + B'(e^{ikx})' = uy. \tag{9.29}$$

Here let us assume,

$$A'e^{ikx} + B'e^{-ikx} = 0, \tag{9.30}$$

which leads (9.29) to

$$A'(e^{ikx})' + B'(e^{ikx})' = uy. \tag{9.31}$$

Solving $(9.30), (9.31)$ as a system of equations of A', B' we solve this to obtain expressions

$$A' = \frac{1}{2ik}u(x)y(x)e^{-ikx}\ ,\ B' = -\frac{1}{2ik}u(x)y(x)e^{ikx} \tag{9.32}$$

which, integrating both A', B' along some interval $[0, x]$ yields

$$A(x) = \frac{1}{2ik}\int_0^x u(x)y(x')e^{-ikx'}dx' + C_1, \tag{9.33}$$

$$B(x) = -\frac{1}{2ik}\int_0^x u(x)y(x')e^{ikx'}dx' + C_2, \tag{9.34}$$

where C_1, C_2 are constants. Substituting (9.33) and (9.34) into our solution,

$$\begin{aligned} y(x) = &\Big(\frac{1}{2ik}\int_0^x u(x)y(x')e^{-ikx'}dx' + C_1\Big)e^{ikx} \\ &+ \Big(-\frac{1}{2ik}\int_0^x u(x)y(x')e^{ikx'}dx' + C_2\Big)e^{-ikx}. \end{aligned} \tag{9.35}$$

9.4 Determining Coefficients and the Wronskian

Consider the case $y = f_1(x,k)$. By definition of $f_1(x,k)$, $\lim_{x\to+\infty} y = e^{ikx}$. Therefore we can express e^{ikx} as improper integrals

$$\begin{aligned} e^{ikx} &= \Big(\frac{1}{2ik}\int_0^\infty u(x')f_1(x',k)e^{-ikx'}dx' + C_1\Big)e^{ikx} \\ &+ \Big(-\frac{1}{2ik}\int_0^\infty u(x')f_1(x',k)e^{ikx'}dx' + C_2\Big)e^{ikx}. \end{aligned} \tag{9.36}$$

Comparing coefficients of e^{ikx}, e^{-ikx} we directly obtain,

$$C_1 = 1 - \frac{1}{2ik}\int_0^\infty u(x')f_1(x',k)e^{-ikx'}dx', \tag{9.37}$$

$$C_2 = \frac{1}{2ik}\int_0^\infty u(x')f_1(x',k)e^{ikx'}dx'. \tag{9.38}$$

Substituting into (9.35) and using the identity $e^{ik(x-x')} - e^{-ik(x-x')} = 2i\sin k(x-x')$,

$$f_1(x,k) = e^{ikx} - \frac{1}{k}\int_x^\infty \sin k(x-x')u(x')f_1(x',k)dx'. \tag{9.39}$$

By an identical argument, we can derive the following for $y = f_2(x,k)$. By definition of $f_2(x,k)$, we obtain

$$f_2(x,k) = e^{-ikx} + \frac{1}{k}\int_{-\infty}^x \sin k(x-x')u(x')f_2(x',k)dx'. \tag{9.40}$$

This can be done by noting that $\lim_{x\to-\infty} y = e^{-ikx}$. Equations of the form (9.39), (9.40) are called the Volterra integral equation of the second kind. It is known that $f_1(x,k), f_2(x,k)$ are analytic in the region $\mathrm{Im}k > 0$ in the complex k-plane. See [27, p.211-212]. We chose $f_1(x,k), f_2(x,k)$ as our solutions, but since these were arbitrary, we may also arbitrarily choose functions $f_1(x,-k)$ such that $\lim_{x\to\infty} f_1(x,-k) = e^{-ikx}$, and $f_2(x,-k)$ such that $\lim_{x\to-\infty} f_2(x,-k) = e^{ikx}$. For a second order linear differential equation, if two linearly independent solutions are found, the linear combination of the two allows us to express any solution. We use the following theorem

Theorem 9.4.1. The Wronskians $W_1 = W(f_1(x,k), f_1(x,-k))$ and $W_2 = W(f_2(x,k), f_2(x,-k))$ are nonzero for $\forall x \in \mathbb{R}$,

Proof. See [20, p. 17-18]. □

This allows us to write,

$$f_2(x,k) = c_{11}(k)f_1(x,k) + c_{12}(k)f_1(x,-k), \tag{9.41}$$

$$f_1(x,k) = c_{21}(k)f_2(x,-k) + c_{22}(k)f_2(x,k), \tag{9.42}$$

where $c_{ij} \in \mathbb{C}$ for $i, j \in \{1, 2\}$, and we know that,

$$f_2(x,k) = \begin{cases} e^{-ikx} & (x \to -\infty), \\ c_{11}(k)e^{ikx} + c_{12}e^{-ikx} & (x \to \infty), \end{cases} \tag{9.43}$$

$$f_1(x,k) = \begin{cases} c_{21}(k)e^{ikx} + c_{22}e^{-ikx} & (x \to -\infty), \\ e^{ikx} & (x \to \infty). \end{cases} \tag{9.44}$$

We can take (9.43) for instance, to describe waves incident from $x = \infty$ with amplitude c_{12}, and reflect with amplitude c_{11}, transmit through and reach $x = -\infty$ with 1. (9.44) can be viewed to describe waves in a similar manner, incident from $x = -\infty$. See [18, p. 107-108].

9.5 Reflection and Transmission Coefficients

Set the positive side of the x-axis to be on the right. Then (9.43), (9.44) correspond to waves incident from the right and left sides respectively. Denote the reflection and transmission coefficients from the right as $R_R(k), T_R(k)$ respectively. Similarly, reflection and transmission coefficients from the left are denoted $R_L(k), T_L(k)$ respectively. With these notations, define

$$R_R(k) = \frac{c_{11}(k)}{c_{12}(k)},\ T_R(k) = \frac{1}{c_{12}(k)},\ R_L(k) = \frac{c_{22}(k)}{c_{21}(k)},\ T_L(k) = \frac{1}{c_{21}(k)}. \tag{9.45}$$

It can be shown that $T_L(k) = T_R(k)$ by considering the Wronskians W_1, W_2 and substituting (9.41), (9.42) respectively. see [20, p. 19-20] for proof. For simplicity, we will henceforth denote both transmission coefficients as $T(k)$. It is known that these coefficients behave as follows,

$$R_R(k) = \frac{c_{11}(k)}{c_{12}(k)} \to 0\ (|k| \to \infty)\ ,\ R_L(k) = \frac{c_{22}(k)}{c_{12}(k)} \to 0\ (|k| \to \infty), \tag{9.46}$$

$$T(k) = \frac{1}{c_{12(k)}} \to 1\ (|k| \to \infty). \tag{9.47}$$

Proof. See [20, p. 20-21] □

This implies that $R_R(k)$ and $R_L(k)$ are Fourier transformable, while only $T(k) - 1$ would converge to 0 at $|k| \to \infty$. Thus define each Fourier and inverse Fourier transform of $R_R(k), R_L(k)$ and $T(k) - 1$ respectively, as

$$r_R(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikz}R_R(k)dk,\ R_R(k) = \int_{-\infty}^{\infty} e^{-ikz}r_R(z)dz, \tag{9.48}$$

$$r_L(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ikz}R_L(k)dk,\ R_L(k) = \int_{-\infty}^{\infty} e^{ikz}r_L(z)dz, \tag{9.49}$$

$$\Omega(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikz}[T(k)-1]dk,\, T(k)-1 = \int_{-\infty}^{\infty} e^{ikz}\Omega(z)dz. \tag{9.50}$$

These will come to use in the folowing sections.

9.6 Auxiliary Theorems

We know the eigenvalue $\lambda = k^2$ of (9.2) is real. Therefore $k \in \mathbb{R}$ or purely imaginary. We state following theorems for future use. Proofs have been omitted here but references are given due to their lengths.

Theorem 9.6.1. For $c_{12}(k)$ as in (9.41), k is purely imaginary if $c_{12}(k) = 0$

Proof. See [20, p. 25-32]. □

Theorem 9.6.2. If there are N zeros of $c_{12}(k)$ along the positive imaginary line, then every pole of $T(k)$ has order 1. Equivalently, if we denote the N zeros $k_1, k_2 ... k_N$ in the order that is closer to the origin, then $c_{12}(k) = (k-k_1)...(k-k_N)$

Proof. See [20, p. 28-32]. □

Theorem 9.6.3. The residue of each $T(k_l)$ can be written as

$$\gamma_l(i) = \frac{1}{\dot{c}_{12}(k_l)} \tag{9.51}$$

for some $\gamma_l \in \mathbb{R}$ where $\dot{c}_{12}(k_l) = \frac{dc_{12}(k)}{dk}|_{k=k_l}$.

Proof. See [20, p. 33-34]. □

Now using these theorems, we compute $\Omega(z)$ given in (9.50). Since $T(k)$ is analytic at Im$k > 0$ apart from finite poles, we can take the upper half semicircle with infinite radius with centre at the origin as our integration path. Then we must have the integral to converge to zero, thus $e^{-ikz} \to 0$ as $|k| \to \infty$ is required. Here, e^{-ikz} does not possess poles in the upper half complex plane thus $\int_{-\infty}^{\infty} e^{-ikz}dk = 0$. Moreover, $e^{-ikz}T(k)$ has finite poles in the upper half complex plane. Then by Theorem (9.6.1) and (9.6.2) we express $k_l = p_l i$ where $p_l > 0, \forall l = 1, 2, ..., N$ and by using the residue theorem, we get

$$\begin{aligned} \Omega(z) &= \frac{1}{2\pi} 2\pi i \sum_{l=1}^{N} e^{p_l z} \text{Res}\Big[\frac{1}{c_1 2(x, k_l)}\Big] \\ &= -\sum_{l=1}^{N} \gamma_l e^{p_l z} \ (z < 0). \end{aligned} \tag{9.52}$$

9.7 Prelude to Inverse Scattering

Thus far we have shown that the time dependent Sturm-Liouville equation (9.3) is related to the KdV equation, where $q(x) := u(x)$ can be taken as the potential of a one dimensional wave and when the eigenvalue k^2 is not time-dependent, $u(x,t)$ satisfies the KdV equation (9.1). The time evolution of $u(x,t)$ is given by

$$y_t = -4y_{xxx} + 6uy_x + 3u_x y. \tag{9.53}$$

In other words, when the fundamental solutions of (9.3) satisfies (9.53), $u(x,t)$ satisfies (9.1). Namely, the soltuion $u(x,t)$ of (9.1) can be determined by considering the time dependency of the coefficients R_R(or R_L), $m_R l$(or m_{Ll}). We provide some definitions and a lemma we will use.

Lemma 9.7.1. For $u(x), A_R(x,x')$ and $A_L(x,x')$ we have the relationship

$$u(x) = -2\frac{d}{dx}A_R(x,x) \text{ and } u(x) = 2\frac{d}{dx}A_L(x,x). \tag{9.54}$$

Proof. See [20, p. 44-47]. □

Theorem 9.7.2. Define

$$m_{Rl} \equiv \gamma_l c_{11}(p_l i) = -\frac{ic_{11}(p_l i)}{\dot{c}_{12}(p_l i)}\ ,\ \Omega_R(z) \equiv r_R(z) + \sum_{l=1}^{N} m_{Rl}e^{-p_l z}, \tag{9.55}$$

$$m_{Ll} \equiv \gamma c_{22}(p_l i) = -\frac{ic_{22}(p_l i)}{\dot{c}_{12}(p_l i)}\ ,\ \Omega_L(z) \equiv r_L(z) + \sum_{l=1}^{N} m_{Ll}e^{p_l z}. \tag{9.56}$$

Then the Gelfand-Levitan-Marchenko (GLM) equation is

$$\Omega_R(x+y) + A_R(x,y) + \int_x^{\infty} \Omega_R(x'+y)A_R(x,x')dx' = 0 \tag{9.57}$$

for $x < y$ and,

$$\Omega_L(x+y) + A_L(x,y) + \int_x^{\infty} \Omega_L(x'+y)A_L(x,x')dx' = 0 \tag{9.58}$$

for $x > y$.

Proof. See [20, p. 36-43]. □

Definition 9.7.1. $\Psi(z), \Phi(z)$ denotes vectors
$\Psi(z) = (m_{L1}e^{p_1 z}, m_{L2}e^{p_2 z}, ..., m_{LN}e^{p_N z})$, $\Phi(z) = (e^{p_1 z}, ..., e^{p_N z})$ where m_{Ll} is as defined in (9.55).

Definition 9.7.2. $V(x)$ is a matrix defined by

$$V(x) = I + \int_{-\infty}^{x} \Phi^T(x')\Psi(x')dx', \tag{9.59}$$

where is I the identity matrix.

Now let us find $u(x,t)$ from $A_L(x,x;t)$. By Lemma 9.7.1 and using our definition of $V(x)$, (9.54) can be expressed as

$$u(x,t) = -2\frac{\partial^2}{\partial x^2}[\ln\big(\det V(x,t)\big)]. \tag{9.60}$$

We note $A_L(x,y;t)$ satisfies the GLM equation

$$\Omega_L(x+y;t) + A_L(x,y;t) + \int_x^{-\infty} \Omega_L(x'+y;t)A_L(x,x';t)dx' = 0 \tag{9.61}$$

and that

$$\Omega_L(z;t) = r_L(z;t) + \sum_{l=1}^{N} m_{Ll}(t)e^{p_l z}, \tag{9.62}$$

$$m_{Ll}(t) = -\frac{ic_{22}(p_l i;t)}{\dot{c}_{12}(p_l i;t)}, \tag{9.63}$$

$$r_L(z;t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ikz}\frac{c_{22}(k;t)}{c_{12}(k;t)}dk. \tag{9.64}$$

We must be warned that $\dot{c}_{12}$ is a partial differential with respect to t. Taking t into consideration, (9.42) is

$$f_1(x,k;t) = c_{12}(k;t)f_2(x,-k;t) + c_{22}(k;t)f_2(x,k;t). \tag{9.65}$$

To determine the time dependency of $c_{12}(k;t), c_{22}(k;t)$ we consider some solution that is proportional to $f_1(x,k;t)$,

$$y(x,k;t) = h(k;t)f_1(x,k;t). \tag{9.66}$$

Noting that $f_1(x,k;t) \to e^{ikx}$ as $x \to \infty$ and substituting 9.66 to (9.53), we obtain

$$y_t + 4y_{xxx} = [h_t(k,t) - 4ik^3 h(k,t)]e^{ikx}. \tag{9.67}$$

Also, when $|x| \to \infty$ then $u(x,t) \to 0$ and consecutively $u_x(x,t) \to 0$ as well. Then (9.53) becomes $y_t + 4y_{xxx} = 0$ at $|x| \to \infty$. Hence 9.67 implies,

$$h(k,t) = 4ik^3 h(k,t), \tag{9.68}$$

which is thus sensible to assume $h(k,t) = a(k)e^{bt}$ and substitute to (9.68). Solving, we find $a = h(k,0), b = 4ik^3$. Hence $h(k,t) = h(k,0)e^{4ik^3t}$. Next consider $x \to -\infty$. Then $f_2(x,-k;t) \to e^{ikx}$, $f_2(x,k;t) \to e^{-ikx}$ which gives

$$f_1(x,k;t) \to c_{12}(k;t)e^{ikx} + c_{22}(k;t)e^{-ikx}. \tag{9.69}$$

Substituting (9.66) to $y_t + 4y_{xxx} = 0$ and applying to (9.69),

$$y_t + 4y_{xxx} = h(k,0)e^{4ik^3t}[c_{12}e^{ikx} + (c_{22t} + 8ik^3 c_{22})e^{-ikx}]. \tag{9.70}$$

For (9.70) to equal zero for arbitrary x, k, t, we need conditions

$$c_{12t} = 0,\ c_{22t} = -8ik^3 c_{22}. \tag{9.71}$$

From (9.71) we immediately conclude that

$$c_{12}(k;t) = c_{12}(k;0),\ c_{22}(k;t) = c_{22}(k;0)e^{-8ik^3 t}. \tag{9.72}$$

Using (9.72), we can now compute $m_{Ll}(t)$ as,

$$m_{Ll}(t) = -\frac{ic_{22}(p_l i;t)}{\cdot c_{12}(p_l i;t)} = m_{Ll}(p_l i, 0)e^{-8p_l^3 t} \tag{9.73}$$

Furthermore, the reflection coefficient $R_L(k;t)$ becomes,

$$R_L(k;t) = \frac{c_{22}(k;t)}{c_{12}(k;t)} = R_L(k;0)e^{-8ik^3 t} \tag{9.74}$$

We can take $A_R(x,x;t)$ and proceed through idential calculation and reasoning as we have for $A_L(x,x;t)$ above to obtain, $c_{11}(k,t) = c_{11}(k,0)e^{8ik^3 t}$, $m_{Rl}(t) = m_{Rl}(p_l i;0)e^{8p_l^3 t}$, $R_R(k;t) = R_R(k;0)e^{8ik^3 t}$, and therefore we have determined the time dependency of each c_{ij}, m_{Ll} and m_{Rl}. We close this chapter with an example of how our discussion thus far can be applied to find a soliton solution.

Example 9.7.1. Let us look at the simplest case where $N = 1$ and $R_l(k) = R_R(k) = 0$. Thus there is no reflection, for any k. Now by definition (9.7.2), $V(x,t)$ is a 1×1 matrix. Then,

$$\det V(x,t) = 1 + \int_{-\infty}^{x} m_{L1}(t)e^{2p_1 x'}dx' = 1 + \frac{m_{L_1}(t)}{2p_1}e^{2p_1 x}. \tag{9.75}$$

Then using (9.73), we can write the right hand side of (9.75) as

$$\det V(x,t) = 1 + \frac{m_{L1}(0)}{2p_1}e^{2p_1 x - rp_1^3 t}. \tag{9.76}$$

Then,

$$\ln[\det V(x,t)] = \ln\Big[1 + \frac{m_{L1}(0)}{2p_1}e^{2p_1 x - 8p_1^3 t}\Big]. \tag{9.77}$$

Calculating the second order derivative of (9.77) we get,

$$\frac{\partial^2}{\partial x^2}\ln[\det V(x,t)] = \frac{2p_1 m_{L1}(0)}{(e^{-p_1 x + 4p_1^3 t} + \frac{m_{Ll(0)}}{2p_1}e^{p_1 x - 4p_1^3 t})^2}. \tag{9.78}$$

If we let $e^{\delta} = \sqrt{\frac{m_{L1}(0)}{2p_1}}$, then (9.78) becomes

$$\frac{\partial^2}{\partial x^2}\ln[\det V(x,t)] = p_1^2\frac{4}{(e^{-p_1 x + 4p_1^3 t - \delta} + e^{p_1 x - 4p_1^3 t + \delta})^2}. \tag{9.79}$$

Using the hyperbolic function $\operatorname{sech}\alpha$ we can finally write (9.78) as j

$$\frac{\partial^2}{\partial x^2}\ln[\det V(x,t)] = p_1^2 \operatorname{sech}^2(p_1 x - 4p_1^3 t + \delta). \tag{9.80}$$

Therefore, by (9.60), we obtain the potential $u(x,t)$ as

$$u(x,t) = -2\frac{\partial^2}{\partial x^2}\ln[\det V(x,t)] = -2p_1^2 \operatorname{sech}^2(p_1 x - 4p_1^3 t + \delta). \tag{9.81}$$

If we take t to denote time and $u(x,t)$ to denote the perpendicular displacement, let us show that (9.81) describes a solitary wave. In general, a wave propagating with speed c toward the positive x direction can be described by $f(c - xt)$. Thus (9.81) can be viewed as a wave with constant speed $c = 4p_1^2$ toward the positive x direction. Now, partial differentiation of (9.81) with respect to x gives,

$$\frac{\partial u(x,t)}{\partial x} = 16p_1^3 \frac{e^{p_1 x - 4p_1^3 t + \delta} - e^{-p_1 x + 4p_1^3 t - \delta}}{[e^{p_1 x - 4p_1^3 t + \delta} + e^{-p_1 x + 4p_1^3 t - \delta}]^3} \tag{9.82}$$

Here, $e^{p_1 x - 4p_1^3 t + \delta} + e^{-p_1 x + 4p_1^3 t - \delta} > 0$ and since $p_1 > 0$ by definition of p_l, thus the sign of $\frac{\partial u}{\partial x}$ is determined solely by the sign of the numerator of (9.82) which is $e^{p_1 x - 4p_1^3 t + \delta} - e^{-p_1 x + 4p_1^3 t - \delta}$. When $x \geq \frac{4p_1^3 t - \delta}{p_1}$, $e^{p_1 x - 4p_1^3 t + \delta} - e^{-p_1 x + 4p_1^3 t - \delta} \geq 0$ and $x < \frac{4p_1^3 t - \delta}{p_1}$ we have $e^{p_1 x - 4p_1^3 t + \delta} - e^{-p_1 x + 4p_1^3 t - \delta} < 0$. Therefore,

$$\frac{\partial u(x,t)}{\partial x} \begin{cases} > 0 & \text{if } x > \frac{4p_1^3 t - \delta}{p_1} \\ = 0 & \text{if } x = \frac{4p_1^3 t - \delta}{p_1} \\ < 0 & \text{if } x < \frac{4p_1^3 t - \delta}{p_1} \end{cases} \tag{9.83}$$

Noting that $\lim_{|x|\to\infty} \frac{\partial u}{\partial x} = 0$, we therefore see that by (9.83), $u(x,t)$ describes a solitary wave with a crest at $x = \frac{4p_1^3 t - \delta}{p_1}$. [20, p. 59-61]

$\triangle$

9.8 Conclusion

We have considered the Sturm-Liouville equation (1.1) under the specific conditions that $u(x,t)$ is real and that it converges to zero at infinity. When λ is not time dependent we concluded that the Sturm-Liouville problem is strongly related to the KdV equation. Solving this with variation of parameters leads to finding solutions $f_1(x,k), f_2(x,k)$ and relevant integral equations that further allow us to describe scattering information of (1.1) as a one dimensional wave equation. The properties of the eigenvalues of (1.1) have led to interesting conclusions of the physical coefficients defined by the coefficients of the fundamental solutions. These properties and results have enabled us to briefly discuss the inverse scattering method where we have theoretically demonstrated how to solve the KdV equation, where we have specifically presented an $N = 1$ soliton example in section 9.7.

Conclusion

Our primary goal in this report was to provide a concrete foundation of the Sturm-Liouville problem (1.1), (1.2) through specific examples where applicable. Much of our report demonstrated how certain essential theorems such as the Spectral Theorem are used to solve problems that may be difficult to tackle without them, emphasising the importance of each theorem. Furthermore, we note that we can see how each part of the foundations section leads to the next: we began with the self-adjointness property of Lagrange's identity which consequently led to discuss properties of eigenvalues and eigenfunctions, thereby ultimately providing a ground for finding solutions. Throughout this project, we replaced the trigonometric functions usually used to give a Fourier series, with other orthogonal functions. It is possible to do this for any continuous function with continuous derivative, which we are attempting to approximate on a given interval, by using Legendre's and Bessel's functions. We ended our theory section with discussion of the singular problem and mean convergence, which are applicable to a wider range of problems.

In the applications section we have seen how the Sturm-Liouville problem has relevance to ideas first conceptualised more than five hundred years ago, but also has importance for problems which are still being investigated today. In fact, the KdV equation can be expanded to two dimensions described by the Kadmotsev-Petviashvili(KP) equation that ultimately relates to the Davey-Stewartson I (DSI) equation, which is a generalisation of the two dimensional Schrödinger equation, giving solutions known as dromions. The study of the DSI equation sheds light on the relationship between the Sturm-Liouville equation and the Schrödinger equation which we were unfortunately unable to elaborate on in this report.

Hence we can say that there is much more which can be discussed on this topic. This is an indication of the applicability and wide range of related subjects, which may seem otherwise at first glance. This is what makes the Sturm-Liouville problem a truly fascinating area of study in mathematics and physics, and we believe we have succeeded in providing both insight and a foundation to act as inspiration to investigate further ideas which may lead to new discoveries and concepts.

Bibliography

[1] J. J. O'Connor and E. F. Robertson (2001) *Jacques Charles Francois Sturm.* http://www-groups.dcs.st-and.ac.uk/~history/Biographies/Sturm.html Last accessed 13th March 2016.

[2] J. J. O'Connor and E. F. Robertson (1997) *Joseph Liouville.* http://www-groups.dcs.st-and.ac.uk/~history/Biographies/Liouville.html Last accessed 13th March 2016.

[3] Boyce, W.E. and DiPrima, R.C. (2001). *Elementary Differential Equations and Boundary Value Problems. Seventh edn.* Asia: John Wiley and Sons.

[4] Boyce, W.E. and DiPrima, R.C. (2008). *Elementary Differential Equations and Boundary Value Problems. Ninth edn.* Asia: John Wiley and Sons.

[5] Boyce, W.E. and DiPrima, R.C. (2008). *Elementary Differential Equations and Boundary Value Problems. Ninth edn.*

[6] Boyce, W.E. and DiPrima, R.C. (2013). *Elementary Differential Equations and Boundary Value Problems. Tenth edn.* Asia: John Wiley and Sons.

[7] Professor Dacian N. Daescu (2012). *Nonhomogeneous Problems.* http://web.pdx.edu/~daescu/mth427_527/notes_weeks6-7.pdf Last accessed 13th March 2016.

[8] Peter Colwell (1992). *Bessel Functions and Kepler's Equation.* http://www.jstor.org/stable/2324547?seq=1#page_scan_tab_contents Last accessed 13th March 2016.

[9] rip (2011). *Elliptical Orbits – Deriving Kepler's Equation.* https://rip94550.wordpress.com/2011/05/02/ Last accessed 13th March 2016.

[10] Astronomy Answers Kepler's Equation (2016). *Astronomy Answers, Kepler's Equation.* http://aa.quae.nl/en/reken/kepler.html Last accessed 13th March 2016.

[11] Marc A. Murison (2006). *A Practical Method for Solving the Kepler Equation.* http://alpheratz.net/dynamics/twobody/KeplerIterations_summary.pdf Last accessed 13th March 2016.

[12] Jürgen Giesen (2016). *Solving Kepler's Equation of Elliptical Motion.* http://www.jgiesen.de/kepler/kepler.html Last accessed 13th March 2016.

[13] The Feynman Lectures on Physics, Volume I (1963,2006,2013). *7 The Theory of Gravitation.* http://www.feynmanlectures.caltech.edu/I_07.html Last accessed 13th March 2016.

[14] Gustafson, Grant B. (2004). *Bessel's Equation and Bessel Functions.* http://www.math.utah.edu/~gustafso/s2014/3150/asmar-def-theorem-example/ch4AsmarPDF/asmar2004Ch4.7.pdf Last accessed 13th March 2016.

[15] Charudatt Kadolkar (2003). *Sturm-Liouville System.* http://www.iitg.ernet.in/physics/fac/charu/courses/ph402/SturmLiouville.pdf Last accessed 13th March 2016.

[16] Weisstein, Eric W. *Legendre Differential Equation.* From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/LegendreDifferentialEquation.html Last accessed 13th March 2016.

[17] Weisstein, Eric W. *Chebyshev Differential Equation.* From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/ChebyshevDifferentialEquation.html Last accessed 13th March 2016.

[18] Don Hinton (1997). *Spectral Theory & Computational Methods of Sturm-Liouville Problems First Edition.* pp 105-110. CRC Press.

[19] Zabusky,N.J., Kruskal, M.D. , (1965). *Interaction of "solitons" in a collisionless plasma and recurrence of initial states.* Physical Review Letters

[20] Tatsuya Yamashita (2007). *Mathematical theory of solitons -Inverse scattering method for solving the KdV equation-.* Hokkaido University, Department of Earth Sciences

[21] Masayuki Oikawa (2012). *Nonlinear motions centering the soliton, Chapter 1 Discovery of Soliton.* Fukuoka Institution of Technology, College of Engineering

[22] S. S. Ghosh, A. Sen, and G. S. Lakhina. (2002) *Dromion solutions for nonlinear electron acoustic waves in space plasmas.* Nonlinear Processes in Geophysics, pp 463-464 © European Geosciences Union 2002

[23] Tetsu Yajima, Katsuhiro Nishinari (1999). *The solution to linear eigenalvalue problems associated to the repulsive potential and the initial value problem of the Davey-Stewartson equation*, Utsunomiya University, Mathematical Theory Research Department 1092

[24] Wilfred Kaplan (August 2002). *Advanced Calculus, Fifth Edition.* Pearson

[25] R. S. Johnson (2006). *The Notebook Series An Introduction to Sturm-Liouville theory.* School f Mathematics & Statistics, University of Newcastle upon Tyne, © R. S. Johnson 2006

[26] Yutaka Takasugi, Keishi Baba (2009). *Partial Differential Equations, The Campus Seminar Series.* Mathema

[27] Tetsu Yajima, Katsuhiro Nishinari (1997). *Numerical and exact solutions of the Davey-Stewartson equation with general initial conditions.* Utsunomiya University, Mathematical Theory Research Department 993