university of groningen

faculty of science and engineering

The Sturm-Liouville Problem

Master Project Mathematics (Science, Business & Policy track)

July 2017

Student: D.E. Koning

First supervisor: Dr. A.E. Sterk

Second supervisor: Prof. dr. H. Waalkens

Abstract

This thesis treats the Sturm-Liouville problem, a typical case of some differential equation subject to certain boundary conditions. After a short introduction this problem is explored using the separation of variables method to solve a differential equation associated with the Sturm-Liouville theory. Then some basic properties of vector spaces and inner product spaces will be given and the $\mathcal{L}^2$ space will be discussed, which is one of the most common examples of a Hilbert space. This enables us to describe the convergence, completeness and orthogonality of functions in the $\mathcal{L}^2$ space. Using this, differential operators which are self-adjoint will be explored since this concept applies to the operator associated with the Sturm-Liouville problem. Thereafter the Sturm-Liouville problem itself shall be discussed and some spectral properties concerning the eigenvalues and eigenfunctions will be proven. Finally the theory discussed thus far will be illustrated by working out an example and some applications of this subject will be given.

Contents

1 Introduction

Probably everyone has heard or played an instrument, for example a guitar. The sound of a guitar comes from plucking a string. The vibration of the plucked string changes the initial position and initial velocity. The motion of the vibrating string can be described by a wave equation, which is a typical problem that can be solved by the Sturm-Liouville theory. Exploring some of its the properties gives us insight in the harmonics of the instrument, which clarifies the pleasing sound of a guitar.
Not only the sound of a guitar, but many other important physical processes and mechanical systems from classical physics and quantum physics can be described by means of a Sturm-Liouville equation. This Sturm-Liouville theory deals with linear second-order differential equations subject to particular boundary conditions. Often these differential equations describe the oscillation of e.g. pendula, vibrations, resonances or waves.
The Sturm-Liouville problem is named after Jacques Charles François Sturm (1803–1855) and Joseph Liouville (1809–1882). Sturm described how a partial differential equation corresponding to a Sturm-Liouville problem can be solved by the separation of variables method. He discussed the heat conduction in an inhomogeneous thin bar as an example to illustrate this. However, before him people were mostly interested in the specific solution itself. In contrast, Sturm and Liouville came up with a theory which focused on investigating the properties of the solution of the differential equation, such as the eigenfunctions. In view of the guitar problem the eigenfunctions correspond to the resonant frequencies of vibration. For a more detailed and very interesting depiction of the origin and the history of the Sturm-Liouville problem the reader is referred to pages 423-475 of [7].
The simplest example of a Sturm-Liouville operator is the constant-coefficient second-derivative operator. This example will be elaborated at the end of this thesis and we shall see the eigenfunctions are the trigonometric functions. Other functions associated to Sturm-Liouville operators are the well-known Airy function and Bessel function.

First in chapter 2 a brief overview of ordinary differential equations and the general Sturm-Liouville theory will be given. The separation of variables method will be utilized to give a solution to this problem and to solve a particular case, namely the heat equation. Then in chapter 3 some basic properties of vector spaces and inner product spaces will be given. This enables us to discuss the $\mathcal{L}^2$ space, which is a particular case of a Hilbert space. Finally the convergence, completeness and orthogonality of functions in the $\mathcal{L}^2$ space will be described in this section. Thereafter this will be used to explore the self-adjointness property of differential operators in chapter 4. We will show the Sturm-Liouville operator treated in this thesis is also self-adjoint. In chapter 5 the Sturm-Liouville problem itself will be discussed together with its spectral properties. Several statements concerning the eigenvalues and eigenfunctions will be proven. Finally in section 6 some applications of the Sturm-Liouville theory will be given, whereafter a Sturm-Liouville problem will be illustrated. The second-derivative operator will be studied in detail as a concrete example, followed by some concluding remarks.

2 The Sturm-Liouville Problem

In this section we start with some basic concepts of ordinary differential equations. Then we will give the general theory behind the Sturm-Liouville problem and afterwards we shall shortly give a review of the *separation of variables* method for solving partial differential equations. However, we assume the reader is somewhat familiar with this method and partial differential equations in general. At the end of this section this method will be utilized to solve a differential equation associated with the Sturm-Liouville problem.

2.1 Ordinary differential equations

Let's first look at ordinary differential equations in general. Consider an ordinary differential equation of second order on the real interval I given by the following equation

$$\alpha_0(x)u'' + \alpha_1(x)u' + \alpha_2(x)u = f(x), \tag{1}$$

where $\alpha_0, \alpha_1, \alpha_2$ and f are complex functions on I. If the function f equals zero, equation (1) is *homogeneous*. Otherwise it is considered as *inhomogeneous*, i.e. $f \neq 0$. If we define the following second-order *differential operator*

$$L = \alpha_0(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \alpha_1(x)\frac{\mathrm{d}}{\mathrm{d}x} + \alpha_2(x),$$

we are able to rewrite equation (1) in the following form

$$Lu = f.$$

Consider any two functions f, g for which the first and second derivatives are continuous on the interval I, i.e. $f, g \in C^2(I)$. If they are solutions of differential equation (1), then for constants $c_1, c_2 \in \mathbb{C}$ we have

$$L(c_1 f + c_2 g) = c_1 Lf + c_2 Lg,$$

which implies the operator L is linear. Therefore we call (1) a *linear* differential equation. If f and g are solutions to the linear homogeneous differential equation, i.e.

$$Lf = 0 \qquad \text{and} \qquad Lg = 0,$$

then any linear combination of solutions is also a solution. This implies

$$L(c_1 f + c_2 g) = c_1 Lf + c_2 Lg = 0.$$

If the function α_0 is nowhere zero on the interval I we can divide the ordinary differential equation (1) by α_0 which gives

$$u'' + q(x)u' + r(x)u = g(x), \tag{2}$$

where $q = \frac{\alpha_1}{\alpha_0}$, $r = \frac{\alpha_2}{\alpha_0}$ and $g = \frac{f}{\alpha_0}$. Obviously, the ordinary differential equations (1) and (2) have the same set of solutions and in this case equation (1) is called *regular*. The equation is said to be *singular* if there exists some point

$z \in I$ such that $\alpha_0(z) = 0$ and in this case z is called the *singular point* of the ordinary differential equation. Since we assume the reader is familiar with linear second-order differential equations we neither discuss *initial conditions* and *boundary conditions*, nor the properties of the solutions of equation (2). Later on, especially from section 4 about self-adjoint operators onwards, we will make use of the statements made in this section.

2.2 The Sturm-Liouville problem

The Sturm-Liouville problem is a differential equation over the finite interval $[a, b]$ which is of the following form

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\frac{\mathrm{d}u(x)}{\mathrm{d}x}\right] + r(x)u(x) = \lambda w(x)u(x). \tag{3}$$

The Sturm-Liouville problem is called *regular* if the functions $p(x)$ and $w(x)$ are both strict positive, $p(x)$, $p'(x)$, $r(x)$ and $w(x)$ are continuous over the interval $[a, b]$ and moreover it has the following non-trivial boundary conditions

$$\begin{aligned} \alpha_1 u(a) + \alpha_2 u'(a) &= 0 \qquad\qquad (\alpha_1^2 + \alpha_2^2 > 0), \\ \beta_1 u(b) + \beta_2 u'(b) &= 0 \qquad\qquad (\beta_1^2 + \beta_2^2 > 0). \end{aligned} \tag{4}$$

Of course the trivial solution $u(x) = 0$ is always a solution. However, the non-trivial solutions of the differential equation given in (3) which satisfy the boundary conditions in (4) only exist for specific values of λ. These values are called the *eigenvalues* of the boundary value problem and we denote them by λ_n. The corresponding *eigenfunctions* are denoted by u_n. Before solving the problem using the separation of variables method we will first give some main results of the Sturm-Liouville theory which among others can be found on pages 147-148 of [8] or pages 72-73 of [13].

1. The eigenvalues can be ordered in the following way

   $$\lambda_1 < \lambda_2 < \lambda_3 < \cdots < \lambda_n < \cdots,$$

   where $\lim_{n\to\infty} \lambda_n \to \infty$. These eigenvalues show asymptotic behaviour as can be seen in the example given in section 6. However, this will not be proven in this thesis but can be found on page 275 of [12].

2. The eigenfunction u_n corresponding to the eigenvalue λ_n is unique up to a constant normalization factor and it has exactly $n - 1$ zeros in the open interval (a, b).

3. After normalizing $u_n(x)$, the eigenfunctions form an *orthonormal basis*, i.e.

   $$\langle u_n, u_m \rangle = \int_a^b u_n(x)u_m(x)w(x)\,\mathrm{d}x = \delta_{mn}.$$

   This is an orthonormal basis for the weighting function $w(x)$ over the interval $[a, b]$ in the Hilbert space $\mathcal{L}^2([a, b], w(x)dx)$. Later on in this thesis we will clarify the notion of *Hilbert spaces* and the $\mathcal{L}^2$ *space*.

2.3 The separation of variables method

We will now give a short review of the separation of variables method. This method is used for solving linear partial differential equations with boundary and initial conditions. Examples are the heat equation, wave equation, Laplace equation and Helmholtz equation, which we assume the reader are familiar with. This method will be illustrated in section 2.3.1 by solving a Sturm-Liouville partial differential equation. Then in section 2.3.2 we will consider a special case where the separation of variable method will be utilized to give a solution to the homogeneous heat equation.

2.3.1 Solution of a Sturm-Liouville equation

Let's consider a partial differential equation of the following form with corresponding boundary and initial conditions

$$\begin{aligned} f(x)\frac{\partial^2 u}{\partial x^2} + g(x)\frac{\partial u}{\partial x} + h(x)u = \frac{\partial u}{\partial t} + k(t)u, \\ u(a,t) = u(b,t) = 0, \\ u(x,0) = s(x). \end{aligned} \tag{5}$$

The separation of variables method is based on u being writable as a product of two functions X and T where the variables are separated, as the name already suggested. So a solution would be of the following form

$$u(x,t) = X(x)T(t).$$

Using this expression for $u(x,t)$ and taking the appropriate derivatives enables us to write the equation given in (5) in the following separated form

$$\frac{\hat{L}X(x)}{X(x)} = \frac{\hat{M}T(t)}{T(t)}, \tag{6}$$

where we used the following substitutions

$$\begin{aligned} \hat{L} &= f(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + g(x)\frac{\mathrm{d}}{\mathrm{d}x} + h(x), \\ \hat{M} &= \frac{\mathrm{d}}{\mathrm{d}t} + k(t). \end{aligned} \tag{7}$$

We see $\hat{L}$ and $X(x)$ only depend on the variable x and $\hat{M}$ and $T(t)$ only on t. Since the right hand side only depends on x and the left hand side only on t, we know both sides of (6) have to be equal to some constant which we will denote by λ. Therefore we can write our differential equation in the following form

$$\begin{aligned} \hat{L}X(x) &= \lambda X(x), \\ \hat{M}T(t) &= \lambda T(t). \end{aligned} \tag{8}$$

Using the boundary conditions given in (5) we obtain

$$X(a) = X(b) = 0.$$

The Sturm-Liouville theory enables us to solve the first formula of (8) and for now we assume this solution is known, i.e. we already have the eigenfunctions X_n and eigenvalues λ_n. However, we are already able to solve the second formula of (8). Again assuming the eigenvalues are known and substituting the formula for $\hat{M}$ given in (7) we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}T_n(t) = (\lambda_n - k(t))T_n(t).$$

This leads to the following solution for T_n

$$T_n(t) = a_n e^{\left(\lambda_n t - \int_0^t k(\tau)\mathrm{d}\tau\right)}.$$

Using this formula and substituting it in the equation for $u(x,t)$ results in

$$u(x,t) = \sum_n a_n X_n(x) e^{\left(\lambda_n t - \int_0^t k(\tau)\mathrm{d}\tau\right)}.$$

The coefficient a_n can be determined using the Hilbert space $\mathcal{L}^2$ with its corresponding *inner product*, which will be discussed later on in this thesis.

2.3.2 The heat equation

Let's examine the heat equation, which is a special case of a Sturm-Liouville problem. In the one-dimensional form it is given by the following formula

$$\frac{\partial u}{\partial t} - \alpha\frac{\partial^2 u}{\partial x^2} = 0, \tag{9}$$

and we consider homogeneous boundary conditions, i.e.

$$u(0,t) = 0 = u(L,t). \tag{10}$$

Now we will separate the variables, which implies we consider a solution of x of the following form

$$u(x,t) = X(x)T(t), \tag{11}$$

as we have seen before. Taking the first and second partial derivative of the function $u(x,t)$ given in (11) with respect to t and x, making use of the product rule, and substituting this in the heat equation (9) we obtain the following result

$$\frac{T'(t)}{\alpha T(t)} = \frac{X''(x)}{X(x)}.$$

Again this implies both sides of the equation above are equal to some constant $-\lambda$, leading to the following expressions

$$\begin{aligned} T'(t) &= -\lambda\alpha T(t), \\ X''(x) &= -\lambda X(x). \end{aligned}$$

For non-positive λ there are no solutions for $X(x)$, as we shall see when considering the following two cases. Let us first assume $\lambda < 0$. Then a solution for X is given by

$$X(x) = Be^{\sqrt{-\lambda}\,x} + Ce^{-\sqrt{-\lambda}\,x},$$

where $B, C \in \mathbb{R}$. Then, using the homogeneous boundary conditions given in (10) we obtain

$$X(0) = 0 = X(L),$$

which implies $B = 0 = C$. Therefore $u = 0$.
Assuming λ equals 0 we have the following solution for X

$$X(x) = Bx + C,$$

where $B, C \in \mathbb{R}$. Using the boundary conditions again we obtain $u = 0$.
This implies λ has to be positive. This gives the following solutions for T and X

$$\begin{aligned} T(t) &= Ae^{-\lambda \alpha t}, \\ X(x) &= B\sin(\sqrt{\lambda}\,x) + C\cos(\sqrt{\lambda}\,x), \end{aligned}$$

where $A, B, C \in \mathbb{R}$. Then from the boundary conditions, i.e. $X(0) = 0 = X(L)$ we obtain

$$\begin{aligned} C &= 0 \qquad \text{and} \\ \sqrt{\lambda} &= n\frac{\pi}{L}, \end{aligned}$$

for some $n \in \mathbb{N}$. So this is the solution to the heat equation if we can write $u(x,t)$ as a product where the variables x and t are separated as in (11).
Summing over the solutions satisfying the heat equation with the boundary conditions given in (10) yields

$$u(x,t) = \sum_{n=1}^{\infty} D_n \sin\frac{n\pi x}{L} \exp\left(-\frac{n^2\pi^2\alpha t}{L^2}\right). \tag{12}$$

This is also a solution to the heat equation satisfying the boundary conditions and therefore we can regard it as the final solution. Using an initial condition enables us to determine D_n. If we are given the following initial condition

$$u(x,0) = f(x),$$

we obtain the following solution for $f(x)$ by substituting $t = 0$ in equation (12)

$$f(x) = \sum_{n=1}^{\infty} D_n \sin\frac{n\pi x}{L}.$$

If we multiply both sides of the last equation by $\sin\frac{n\pi x}{L}$ and integrate them from 0 to L we obtain the following result for D_n

$$\begin{aligned} f(x)\sin\frac{n\pi x}{L} &= \sum_{n=1}^{\infty} D_n \left(\sin\frac{n\pi x}{L}\right)^2 \\ \int_0^L f(x)\sin\frac{n\pi x}{L}\,\mathrm{d}x &= \int_0^L \sum_{n=1}^{\infty} D_n \left(\sin\frac{n\pi x}{L}\right)^2 \mathrm{d}x \\ \int_0^L f(x)\sin\frac{n\pi x}{L}\,\mathrm{d}x &= \frac{L}{2}D_n \\ D_n &= \frac{2}{L}\int_0^L f(x)\sin\frac{n\pi x}{L}\,\mathrm{d}x. \end{aligned}$$

Now the Sturm-Liouville theory plays an important role since it assures us the eigenfunctions, which in this case are given by $\left\{\sin\frac{n\pi x}{L}\right\}_{n=1}^{\infty}$, are *orthogonal* and *complete*. The orthogonality and completeness of the eigenfunctions are necessary conditions for this method and will be discussed later on in this thesis.

3 Vector Spaces

In this section some basic concepts shall be given which are necessary to better understand the rest of the thesis. Definitions and theorems from *metric spaces* will be omitted, since they should be known by the reader. For a detailed treatment of this subject, the reader is referred to [11].
We shall consider a linear vector space X, or just called a *vector space*. This is a set over $\mathbb{K}$ whose elements are called vectors and we have two operations, namely addition and scalar multiplication. $\mathbb{K}$ here stands for the field of real numbers $\mathbb{R}$ or the field of complex numbers $\mathbb{C}$. We assume the reader is familiar with the basic properties of linear algebra, but for an accessible course on this topic we recommend [6].
First in section 3.1 some basic properties of an inner product space will be given, which is a special type of a vector space. This enables us to discuss the $\mathcal{L}^2$ space in chapter 3.2, which is a particular case of a Hilbert space. At the end of this section we will describe the convergence, completeness and orthogonality of functions in the $\mathcal{L}^2$ space.

3.1 The inner product space

The vector space X is called an *inner product space* if for every pair of elements $x, y \in X$ we have a corresponding *inner product*, denoted as $\langle x, y\rangle \in \mathbb{K}$. This can be seen as a function from $X \times X \to \mathbb{K}$ such that for $x, y \in X$ we have $(x, y) \mapsto \langle x, y\rangle \in \mathbb{K}$. The inner product has to satisfy the following properties

1. The inner product of two elements equals the complex conjugate of the reversed inner product, i.e
$$\langle x, y\rangle = \overline{\langle y, x\rangle}.$$
2. The inner product is linear, so for $x, y, z \in X$ and $\alpha \in \mathbb{K}$ this implies
$$\begin{aligned}\langle x + y, z\rangle &= \langle x, z\rangle + \langle y, z\rangle \quad \text{and}\\ \langle \alpha x, y\rangle &= \alpha\langle x, y\rangle.\end{aligned}$$
3. The inner product of an element with itself is positive definite, so for all $x \in X$ we have
$$\begin{aligned}&\langle x, x\rangle \geq 0 \quad \text{and}\\ &\langle x, x\rangle = 0 \quad \text{if and only if} \quad x = 0.\end{aligned}$$

An inner product space satisfies the following condition, which is also known as the *Cauchy-Schwarz Inequality.*

Theorem 3.1. *If X is an inner product space, then for all x and y in X we have*
$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle.$$

Proof. Let α be any complex number. Then, using the properties an inner product has to satisfy on an inner product space, we get the following inequality
$$\begin{aligned}0 &\leq \langle x - \alpha y, x - \alpha y\rangle\\ &= \langle x, x\rangle + \langle x, -\alpha y\rangle + \langle -\alpha y, x\rangle + \langle -\alpha y, -\alpha y\rangle \\ &= \langle x, x\rangle - \overline{\alpha}\langle x, y\rangle - \alpha\overline{\langle x, y\rangle} + \alpha\overline{\alpha}\langle y, y\rangle,\end{aligned} \tag{13}$$

where $\overline{\alpha}$ denotes the complex conjugate of α.
Let us first consider the case where $\langle y, y\rangle \neq 0$. Taking $\alpha = \frac{\langle x,y\rangle}{\langle y,y\rangle}$ and substituting into the inequality given in (13) gives

$$\begin{aligned}0 &\leq \langle x, x\rangle - \frac{\overline{\langle x, y\rangle}}{\overline{\langle y, y\rangle}}\langle x, y\rangle - \frac{\langle x, y\rangle}{\langle y, y\rangle}\overline{\langle x, y\rangle} + \frac{\langle x, y\rangle}{\langle y, y\rangle}\frac{\overline{\langle x, y\rangle}}{\overline{\langle y, y\rangle}}\langle y, y\rangle \\ &= \langle x, x\rangle - \frac{|\langle x, y\rangle|^2}{\overline{\langle y, y\rangle}} - \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle} + \frac{|\langle x, y\rangle|^2}{\overline{\langle y, y\rangle}} \\ &= \langle x, x\rangle - \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle},\end{aligned}$$

which is exactly the inequality we wanted to prove.
Now we consider the case where $\langle y, y\rangle = 0$. This is actually trivial since it implies $y = 0$ and therefore the inequality always holds. This makes the proof complete. □

The vector space X is a *normed space* if for every element of X there exists a *norm*, which is a non-negative real number. If we have a vector $x \in X$ we will denote the norm by $\|x\|$ and it satisfies the following properties

1. $\|x + y\| \leq \|x\| + \|y\|$
2. $\|\alpha x\| = |\alpha|\|x\|$
3. $\|x\| > 0$ whenever $x \neq 0$,

for all $x, y \in X$ and scalars α.
Now let's define the following norm for $x \in X$

$$\|x\| = \langle x, x\rangle^{\frac{1}{2}}, \tag{14}$$

where x is called *normalized* if we have $\|x\| = 1$. Using this norm the Cauchy-Schwarz inequality given in theorem 3.1 takes on the following form

$$|\langle x, y\rangle| \leq \|x\|\|y\|. \tag{15}$$

This defines a norm on an inner product space X if we prove the following theorem.

Theorem 3.2. *If X is an inner product space, then for all x and y in X we have*

$$\|x + y\| \leq \|x\| + \|y\|.$$

Proving this theorem implies every pair of elements of X satisfies the three properties above. Therefore this defines a norm on an inner product space X.

Proof. Using the norm given in (14) we obtain

$$\begin{aligned}\|x + y\|^2 &= \langle x + y, x + y\rangle \\ &= \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2 \\ &= \|x\|^2 + 2\ \mathrm{Re}\langle x, y\rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 && \text{using (15)} \\ &= (\|x\| + \|y\|)^2\end{aligned}$$

Therefore, taking square roots of both sides, we obtain the following inequality

$$\|x+y\| \leq \|x\| + \|y\|,$$

which proves the theorem. □

Theorem 3.2 also leads to the well-known *triangle inequality.* defining the distance between two elements x and y as $\|x-y\|$, we obtain for three elements $x, y, z \in X$ the following inequality

$$\|x-y\| = \|x-z+z-y\| \leq \|x-z\| + \|z-y\|. \tag{16}$$

Now we are able to explore the idea of *orthogonal sets.*

Definition 3.3. *If we have $\langle x, y\rangle = 0$ for $x, y \in X$, we call x orthogonal to y, denoted as $x \perp y$. A subset $S \subseteq X$ is called an orthogonal set if every pair of elements in this subset is orthogonal.*

We can see definition 3.3 for the n-dimensional real space also intuitively. If we have two vectors $x, y \in \mathbb{R}^n$ we can define the angle θ between these two vectors using the Cauchy-Schwarz inequality given in (15) in the following way

$$\langle x, y\rangle = \|x\|\|y\| \cos\theta.$$

Notice we used the term $\cos\theta$ to make an equality of the inequality. This angle θ is defined uniquely and if the vectors x and y are both nonzero we must have

$$\langle x, y\rangle = 0 \quad \text{if and only if} \quad \cos\theta = 0,$$

which is exactly the condition for the two vectors x and y to be orthogonal in $\mathbb{R}^n$. Using the definition of an orthogonal set we are able to give the concept of an *orthonormal set.*

Definition 3.4. *An orthogonal subset S of an inner product space X is called orthonormal if every element of this subset is normalized, i.e. the norm of every element $x \in S$ equals 1 (or symbolically $\|x\| = 1$).*

3.2 The $\mathcal{L}^2$ Space

Now we are able to move on to the $\mathcal{L}^2$ *space*, which is named after the French mathematician Henri Lebesgue (1875 – 1941). First let us define the vector space denoted by $C([a, b])$. This vector space consists of complex continuous functions on the real interval $[a, b]$. For two functions $f, g \in C([a, b])$ we define the following inner product

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\, \mathrm{d}x. \tag{17}$$

Using the norm given in (14) we obtain

$$\|f\| = \langle f, f\rangle^{\frac{1}{2}} = \sqrt{\int_a^b f(x)\overline{f(x)}\, \mathrm{d}x} = \sqrt{\int_a^b |f(x)|^2\, \mathrm{d}x}. \tag{18}$$

We can now show that the Cauchy-Schwarz inequality given in theorem 3.1 and the triangle inequality given in equation (16) both still hold in the vector space $C([a, b])$. This boils down to proving (17) is an inner product, so we have to check the three properties an inner product has to satisfy.

Proof.

1. The inner product of two elements equals the complex conjugate of the reversed inner product, i.e.

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\,\mathrm{d}x = \int_a^b \overline{g(x)\overline{f(x)}}\,\mathrm{d}x = \overline{\langle g, f\rangle}.$$

2. The inner product is linear, so for $f, g, h \in C\left([a,b]\right)$ and $\alpha \in \mathbb{K}$ we have

$$\begin{aligned}\langle f+g, h\rangle &= \int_a^b \left(f(x)+g(x)\right)\overline{h(x)}\,\mathrm{d}x \\ &= \int_a^b f(x)\overline{h(x)}\,\mathrm{d}x + \int_a^b g(x)\overline{h(x)}\,\mathrm{d}x = \langle f, h\rangle + \langle g, h\rangle,\end{aligned}$$

and

$$\langle \alpha f, g\rangle = \int_a^b \alpha f(x)\overline{g(x)}\,\mathrm{d}x = \alpha\int_a^b f(x)\overline{g(x)}\,\mathrm{d}x = \alpha\langle f, g\rangle.$$

3. The inner product is positive definite, so for $f, g \in C\left([a,b]\right)$ we have

$$\langle f, f\rangle = \int_a^b f(x)\overline{f(x)}\,\mathrm{d}x = \int_a^b |f(x)|^2\,\mathrm{d}x \geq 0,$$

Furthermore since g is continuous we must have $\langle g, g\rangle \geq 0$, so in case this is an equality this must imply $g \equiv 0$.

Therefore, the inner product defined by (17) satisfies both the Cauchy-Schwarz inequality and the triangle inequality. Thus we have

$$|\langle f, g\rangle| \leq \|f\|\|g\| \qquad \text{and} \qquad \|f+g\| \leq \|f\| + \|g\|. \tag{19}$$

□

We will now use $\mathcal{L}^2(a,b)$ to denote the set of functions $f : [a,b] \to \mathbb{C}$ for which the following hold

$$\int_a^b |f(x)|^2\,\mathrm{d}x < \infty. \tag{20}$$

However we have to note that these functions have to be *Lebesgue measurable.* This Lebesgue measure is a natural extension of the classical notions of lengths and areas to more general sets. It is a standard way of measuring a set by covering it with intervals and then taking the sum of the lengths of these intervals as an approximation of the measure of the set. This subject will not be further discussed, but a brief review can be found on pages 110-112 of [4] or for a more extensive description the reader is referred to sections 16-20 of [3].

Now we will define the same inner product as in (17) and norm as in (18) on $\mathcal{L}^2(a,b)$. Applying the triangle inequality on the following norm yields

$$\|\alpha f + \beta g\| \leq \|\alpha f\| + \|\beta g\| = |\alpha|\|f\| + |\beta|\|g\|,$$

for all $f, g \in \mathcal{L}^2(a, b)$ and $\alpha, \beta \in \mathbb{C}$. This implies $\alpha f + \beta g \in \mathcal{L}^2(a, b)$, following from the fact that the norm of f and g is less than infinity since $f, g \in \mathcal{L}^2(a, b)$. Therefore it is a vector space and with this inner product it becomes an inner product space. Actually $C([a, b])$ is a subspace of $\mathcal{L}^2(a, b)$ and from now on we will use the latter space since the integrals are interpreted as *Lebesgue integrals*. This is necessary since from the Cauchy-Schwarz inequality it follows that the inner products of f and g are well defined if $\|f\|$ and $\|g\|$ exist. This implies $|f|^2$ and $|g|^2$ have to be integrable and this is exactly the case for $\mathcal{L}^2(a, b)$ since it consists of those functions f such that $|f|^2$ is integrable on $[a, b]$. Again more information about Lebesgue integrals can be found on pages 108-110 of [4] or pages 322-323 of [3].

Two functions $f, g \in \mathcal{L}^2(a, b)$ are called *equal almost everywhere* if $\|f - g\| = 0$. This term comes from the fact f and g do not have to be equal *pointwise* on this interval, i.e. $f(x) - g(x) = 0$ at every $x \in [a, b]$. This is due to the fact that in this space $\|f\| = 0$ does not always imply $f(x) = 0$ at every point $x \in [a, b]$. Therefore we can regard the space $\mathcal{L}^2$ as been made up out of *equivalence classes* of functions which are equal almost everywhere.

3.2.1 Convergence in $\mathcal{L}^2$

Before describing convergence in the space $\mathcal{L}^2$ we will first give the definitions of *pointwise* and *uniform* convergence of a sequence of functions.

Definition 3.5. i) *A sequence of functions $f_n : I \to \mathbb{F}$ converges pointwise to the function $f : I \to \mathbb{F}$ if for every $x \in I$ we have $\lim_{n\to\infty} f_n(x) = f(x)$.*

i)) *This sequence of functions is said to be uniformly convergent if for every $\epsilon > 0$ there exists a natural number N such that*

$$n \geq N \quad \textit{implies} \quad |f_n(x) - f(x)| < \epsilon \quad \textit{for all} \quad x \in I.$$

We will denote pointwise convergence by $f_n \to f$ and uniform convergence by $f_n \overset{u}{\to} f$. One nice consequence is that uniform convergence implies pointwise convergence (but not necessarily the other way around).

Now we have given the definitions of pointwise and uniform convergence of a sequence of functions, we will mostly be looking at convergence in the space $\mathcal{L}^2$ in this section.

Definition 3.6. *If we have a sequence of functions (f_n) in $\mathcal{L}^2(a, b)$, it converges in $\mathcal{L}^2$ if there exists a function $f \in \mathcal{L}^2(a, b)$ such that*

$$\lim_{n\to\infty} \|f_n - f\| = 0.$$

In other words, for every $\epsilon > 0$ there exists a natural number N such that

$$n \geq N \quad \textit{implies} \quad \|f_n - f\| < \epsilon.$$

We will denote convergence in $\mathcal{L}^2$ by $f_n \overset{\mathcal{L}^2}{\to} f$, where f is said to be the *limit in* $\mathcal{L}^2$ of the sequence (f_n).

Pointwise convergence does not imply convergence in $\mathcal{L}^2$ and convergence in $\mathcal{L}^2$ does also not imply pointwise convergence. This is due to the limit in $\mathcal{L}^2$ is an equivalence class of functions, i.e. they are equal in $\mathcal{L}^2$ but not pointwise. However, if a sequence of functions converges pointwise as well as in $\mathcal{L}^2$, then these limits are equal.

Uniform convergence on the other hand does imply convergence in $\mathcal{L}^2$ under the following conditions. The sequence (f_n) and its limit f must lie in $\mathcal{L}^2(I)$ and this interval I should be bounded. The proof will not be given since it uses some basic properties of uniform convergence which have not been discussed in this thesis.

In the preliminaries we assumed the reader is familiar with metric spaces and therefore knows concepts such as completeness and the Cauchy sequence. Now we shall give these definitions in terms of the $\mathcal{L}^2$ space.

Definition 3.7. *A sequence in $\mathcal{L}^2$ is called a Cauchy sequence if for each $\epsilon > 0$ there exists a natural number N such that $\|f_n - f_m\| < \epsilon$ whenever $n, m \geq N$.*

A sequence (f_n) which converges in $\mathcal{L}^2$ is a Cauchy sequence. Namely, using the triangle inequality we have

$$\|f_n - f_m\| \leq \|f_n - f\| + \|f - f_m\|.$$

If we take m and n large enough, the right-hand side of the last inequality can be made as small as desired since we have $f_n \overset{\mathcal{L}^2}{\to} f$. Therefore the left hand side can be made arbitrarily small, which implies this convergent sequence is a Cauchy sequence.

Furthermore the space $\mathcal{L}^2$ is complete. This is because it is one of the most common examples of a *Hilbert space*, which is a complete inner product space under its norm induced by the inner product. Completeness of the $\mathcal{L}^2$ space actually means the following.

Theorem 3.8. *For every Cauchy sequence (f_n) in $\mathcal{L}^2$ there exists a function $f \in \mathcal{L}^2$ such that $f_n \overset{\mathcal{L}^2}{\to} f$.*

The proof can be found on pages 329-330 of [10]. Moreover, the set of continuous functions is dense in this space, i.e. $C([a, b])$ is dense in $\mathcal{L}^2(a, b)$. This density property is given by the following theorem and can also be found in [10] on page 326.

Theorem 3.9. *For any $f \in \mathcal{L}^2(a, b)$ and any $\epsilon > 0$ there exist a continuous function g on $[a, b]$ such that $\|f - g\| < \epsilon$.*

In section 5.2 this theorem will be used to prove the eigenfunctions of a Sturm-Liouville problem are complete, but for now we will first have a look at the orthogonality property in the $\mathcal{L}^2$ space.

3.2.2 Orthogonal functions

Let's say we want to write a function $f \in \mathcal{L}^2$ in terms of other functions in the complex space $\mathcal{L}^2$. So suppose we have an orthogonal set of functions in

$\mathcal{L}^2$, i.e. $\{\varphi_1, \varphi_2, \varphi_3, \dots\}$. Note that this set could be finite or infinite. If the function f is a finite linear combination of elements in this set $\{\varphi_k\}$, then f can be represented as follows

$$f = \sum_{i=1}^{n} \alpha_i \varphi_i \qquad \text{with } \alpha_i \in \mathbb{C}. \tag{21}$$

Since we are dealing with an orthogonal set the inner product of every element with another element equals zero as we have seen in definition 3.3, i.e.

$$\langle \varphi_i, \varphi_j \rangle = 0, \qquad \text{whenever } i \neq j.$$

Therefore, taking the inner product of φ_k with f defined as in (21) yields

$$\langle f, \varphi_k \rangle = \alpha_k \, \|\varphi_k\|^2. \tag{22}$$

Dividing both sides by $\|\varphi_k\|^2$ we obtain an expression for the coefficient α_k

$$\alpha_k = \frac{\langle f, \varphi_k \rangle}{\|\varphi_k\|^2}. \tag{23}$$

Substituting this into equation (21) we see the function f can be represented in $\mathcal{L}^2$ as follows

$$f = \sum_{k=1}^{n} \frac{\langle f, \varphi_k \rangle}{\|\varphi_k\|^2} \, \varphi_k. \tag{24}$$

So we see the coefficients α_k are determined by what we call the *projections* of f on φ_k. We can also obtain an orthonormal set by dividing each φ_i by its norm, i.e. $\left\{\psi_k = \frac{\varphi_k}{\|\varphi_k\|}\right\}$. This defines an orthonormal set since every element ψ_k is normalized. In other words $\|\psi_k\| = 1$ as we saw in definition 3.4. Therefore equation (22) becomes $\langle f, \psi_k \rangle = \alpha_k \|\psi_k\|^2 = \alpha_k$ and it enables us to rewrite our function f as in (24) as follows

$$f = \sum_{k=1}^{n} \langle f, \psi_k \rangle \, \psi_k,$$

where the coefficients are the projections of f on ψ_k.

On the other hand, if f is not a linear combination of elements of $\{\varphi_k\}$ we would like to determine its best approximation in $\mathcal{L}^2$, again using a finite linear combination of the elements of $\{\varphi_k\}$. This comes down to determining the coefficients α_k for which the following norm is minimized

$$\left\| f - \sum_{k=1}^{n} \alpha_k \varphi_k \right\|. \tag{25}$$

If we apply the definition of the norm to the square of the norm given above we get

$$\begin{aligned}
\left\| f - \sum_{k=1}^{n} \alpha_k \varphi_k \right\|^2 &= \left\langle f - \sum_{k=1}^{n} \alpha_k \varphi_k, f - \sum_{k=1}^{n} \alpha_k \varphi_k \right\rangle \\
&= \|f\|^2 - 2\sum_{k=1}^{n} \operatorname{Re} \overline{\alpha}_k \langle f, \varphi_k \rangle + \sum_{k=1}^{n} |\alpha_k|^2 \|\varphi_k\|^2 \\
&= \|f\|^2 - \sum_{k=1}^{n} \frac{|\langle f, \varphi_k\rangle|^2}{\|\varphi_k\|^2} \\
&\quad + \sum_{k=1}^{n} \|\varphi_k\|^2 \left[|\alpha_k|^2 - 2\operatorname{Re} \overline{\alpha}_k \frac{\langle f, \varphi_k\rangle}{\|\varphi_k\|^2} + \frac{|\langle f, \varphi_k\rangle|^2}{\|\varphi_k\|^4} \right] \\
&= \|f\|^2 - \sum_{k=1}^{n} \frac{|\langle f, \varphi_k\rangle|^2}{\|\varphi_k\|^2} + \sum_{k=1}^{n} \|\varphi_k\|^2 \left| \alpha_k - \frac{\langle f, \varphi_k\rangle}{\|\varphi_k\|^2} \right|^2 .
\end{aligned}$$

We only have to deal with the last term of this equation since we want to determine the coefficients α_k which minimize the norm given in (25). This last term on the right hand side above is always greater or equal to zero so to minimize $\|f - \sum_{k=1} \alpha_k \varphi_k\|^2$ (and therefore also its square root) we want to choose α_k such that the last term equals zero, i.e.

$$\alpha_k = \frac{\langle f, \varphi_k\rangle}{\|\varphi_k\|^2}.$$

This makes the last term $\sum_{k=1}^{n} \|\varphi_k\|^2 \left| \alpha_k - \frac{\langle f, \varphi_k\rangle}{\|\varphi_k\|^2} \right|^2$ equal to zero and therefore we obtained the minimum of the norm given in (25). Substituting this α_k in our formula for the squared norm yields

$$\begin{aligned}
\left\| f - \sum_{k=1}^{n} \frac{\langle f, \varphi_k\rangle}{\|\varphi_k\|^2} \varphi_k \right\|^2 &= \|f\|^2 - 2\sum_{k=1}^{n} \frac{\operatorname{Re}\overline{\langle f, \varphi_k\rangle}\langle f, \varphi_k\rangle}{\|\varphi_k\|^2} + \sum_{k=1}^{n} \frac{|\langle f, \varphi_k\rangle|^2}{\|\varphi_k\|^4} \|\varphi_k\|^2 \\
&= \|f\|^2 - 2\sum_{k=1}^{n} \frac{|\langle f, \varphi_k\rangle|^2}{\|\varphi_k\|^2} + \sum_{k=1}^{n} \frac{|\langle f, \varphi_k\rangle|^2}{\|\varphi_k\|^2} \\
&= \|f\|^2 - \sum_{k=1}^{n} \frac{|\langle f, \varphi_k\rangle|^2}{\|\varphi_k\|^2}.
\end{aligned}$$

Since the left hand side has to be greater or equal to zero we can rewrite the right hand side to obtain the following result

$$\sum_{k=1}^{n} \frac{|\langle f, \varphi_k\rangle|^2}{\|\varphi_k\|^2} \leq \|f\|^2.$$

Since this holds for any n we can take the limit as $n \to \infty$ which gives

$$\sum_{k=1}^{\infty} \frac{|\langle f, \varphi_k\rangle|^2}{\|\varphi_k\|^2} \leq \|f\|^2. \tag{26}$$

This last relation is called *Bessel's inequality* and it holds for any orthogonal set $\{\varphi_k : k \in \mathbb{N}\}$ and any $f \in \mathcal{L}^2$.
This inequality becomes an equality if and only if

$$\left\| f - \sum_{k=1}^{\infty} \frac{\langle f, \varphi_k \rangle}{\|\varphi_k\|^2} \varphi_k \right\| = 0.$$

This actually means

$$f = \sum_{k=1}^{\infty} \frac{\langle f, \varphi_k \rangle}{\|\varphi_k\|^2} \varphi_k \quad \text{in } \mathcal{L}^2.$$

Therefore we see the sum $\sum_{k=1}^{\infty} \alpha_k \varphi_k$ with $\alpha_k = \frac{\langle f, \varphi_k \rangle}{\|\varphi_k\|^2}$ is a representation of the function f in $\mathcal{L}^2$.
We will now use these representations of f to describe the completeness of an orthogonal set making use of theorem 3.8. Note that this should not to be confused with the completeness of a space.

Definition 3.10. *An orthogonal set $\{\varphi_k : k \in \mathbb{N}\}$ in $\mathcal{L}^2$ is called complete if for any function $f \in \mathcal{L}^2$ we have*

$$\sum_{k=1}^{n} \frac{\langle f, \varphi_k \rangle}{\|\varphi_k\|^2} \varphi_k \overset{\mathcal{L}^2}{\to} f.$$

This actually implies that a complete orthogonal set in $\mathcal{L}^2$ is a basis for the space and this complete orthogonal set is an infinite set since $\mathcal{L}^2$ is infinite-dimensional.
So if Bessel's inequality given in (26) is an equality we have

$$\|f\|^2 = \sum_{k=1}^{\infty} \frac{|\langle f, \varphi_k \rangle|^2}{\|\varphi_k\|^2}, \tag{27}$$

which is called *Parseval's identity* and we will relate this identity to definition 3.10 about completeness in the following theorem.

Theorem 3.11. *An orthogonal set $\{\varphi_k : k \in \mathbb{N}\}$ is complete if and only if it satisfies Parseval's identity given in (27) for any $f \in \mathcal{L}^2$.*

As we have seen before the orthogonal set $\{\varphi_k\}$ can be normalized to the orthonormal set $\left\{\psi_k = \frac{\varphi_k}{\|\varphi_k\|}\right\}$. Then Bessel's inequality can be written as

$$\sum_{k=1}^{\infty} |\langle f, \psi_k \rangle|^2 \leq \|f\|^2, \tag{28}$$

and Parseval's identity turns into

$$\|f\|^2 = \sum_{k=1}^{\infty} |\langle f, \psi_k \rangle|^2. \tag{29}$$

Additionally, since for all $f \in \mathcal{L}^2$ we have $\|f\| < \infty$, recall formula (20), Bessel's inequality implies $\langle f, \psi_n \rangle \to 0$. So it converges in $\mathcal{L}^2$, independent of this orthonormal set $\{\psi_k\}$ being complete.

Furthermore, Parseval's identity can be seen as the *Pythagorean theorem* for inner product spaces since it is an abstraction from $\mathbb{R}^n$ to the more general setting of the $\mathcal{L}^2$ space. In this case, $\sum_{k=1}^{\infty} |\langle f, \psi_k \rangle|^2$ is a representation of the sum of the squares of the projections of a vector on an orthonormal basis which, by Pythagoras, has to equal the squared length of this vector, now represented by $\|f\|^2$.

4 Self-Adjoint Differential Operators

We are now able to define self-adjoint differential operators in the $\mathcal{L}^2$-space. Therefore we want to find the adjoint of the operator $L : \mathcal{L}^2(I) \cap C^2(I) \to \mathcal{L}^2(I)$ defined as follows

$$L = p(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + q(x)\frac{\mathrm{d}}{\mathrm{d}x} + r(x), \tag{30}$$

which is exactly the same operator as we have seen in section 2.1, though using a more convenient notation. Again we assume p, q and r are real C^2 functions on I. In order to examine the adjoint of the operator L, which will be denoted by L', we have to look at the definition of the adjoint. It is given by the following equation

$$\langle Lf, g\rangle = \langle f, L'g\rangle \quad \text{for all } f, g \in \mathcal{L}^2(I) \cap C^2(I). \tag{31}$$

We will now further investigate $\langle Lf, g\rangle$ to shift the differential operator from f to g. Then, using equation (31), we obtain an expression for the adjoint operator L'. Setting $I = (a, b)$ and using operator (30), inner product (17) and integration by parts, we have

$$\begin{aligned}
\langle Lf, g\rangle &= \int_a^b (pf'' + qf' + rf)\overline{g}\,\mathrm{d}x \\
&= pf'\overline{g}|_a^b - \int_a^b f'(p\overline{g})'\,\mathrm{d}x + qf\overline{g}|_a^b - \int_a^b f(q\overline{g})'\,\mathrm{d}x + \int_a^b fr\overline{g}\,\mathrm{d}x \\
&= (pf'\overline{g} - f(p\overline{g})')\,|_a^b + \int_a^b f(p\overline{g})''\,\mathrm{d}x + qf\overline{g}|_a^b - \int_a^b f(q\overline{g})'\,\mathrm{d}x + \int_a^b fr\overline{g}\,\mathrm{d}x \\
&= \langle f, (\overline{p}g)'' - (\overline{q}g)' + \overline{r}g\rangle + (p(f'\overline{g} - f\overline{g}') + (q - p')f\overline{g})\,|_a^b,
\end{aligned}$$

which is well defined if $p \in C^2(a, b)$, $q \in C^1(a, b)$ and $r \in C(a, b)$. If we rewrite the inner product on the right-hand side using a new differential operator denoted by L^*, we obtain

$$\langle Lf, g\rangle = \langle f, L^*g\rangle + (p(f'\overline{g} - f\overline{g}') + (q - p')f\overline{g})\,|_a^b. \tag{32}$$

Rewriting the inner product this way implies the operator L^* has to satisfy the following equality

$$\begin{aligned}
L^*g &= (\overline{p}g)'' - (\overline{q}g)' + \overline{r}g \\
&= \overline{p}g'' + (2\overline{p}' - \overline{q})g' + (\overline{p}'' - \overline{q}' + \overline{r})g.
\end{aligned}$$

Therefore this operator is defined as follows

$$L^* = \overline{p}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + (2\overline{p}' - \overline{q})\frac{\mathrm{d}}{\mathrm{d}x} + (\overline{p}'' - \overline{q}' + \overline{r}), \tag{33}$$

and we will refer to it as the *formal adjoint* of L. L is *formally self-adjoint* if $L^* = L$. So comparing equations (33) and (30) yields

$$\overline{p} = p, \qquad 2\overline{p}' - \overline{q} = q \qquad \text{and} \qquad \overline{p}'' - \overline{q}' + \overline{r} = r.$$

This implies $q = p'$ since we assumed p and q are real. Substituting this in the definition of our operator L gives

$$Lf = pf'' + p'f' + rf = (pf')' + rf,$$

which enables us to write the operator L in the following form if it is formally self-adjoint

$$L = \frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}}{\mathrm{d}x}\right) + r.$$

Since $L = L^*$ and $q = p'$ when L is formally self-adjoint, we can rewrite equation (32) to

$$\langle Lf, g\rangle = \langle f, Lg\rangle + p(f'\overline{g} - f\overline{g}')|_a^b. \tag{34}$$

Now we are able to determine the form of the self-adjoint operator. We have already defined the adjoint of L in equation (31) and the operator L is called self-adjoint if in addition $L' = L$. Looking at equation (34) we see this requires the following condition

$$p(f'\overline{g} - f\overline{g}')|_a^b = 0 \qquad \text{for all} \quad f, g \in \mathcal{L}^2(I) \cap C^2(I). \tag{35}$$

So we have seen the conditions for an operator to be self-adjoint and we will now apply this to the *Sturm-Liouville eigenvalue problem.* This is a system consisting of the following equation

$$Lu + \lambda u = 0, \tag{36}$$

combined with certain boundary conditions. This equation will be investigated to find the the eigenvalues λ and their corresponding eigenfunctions for the operator $-L$. Since L turns out to have negative eigenvalues for positive p, we look for the eigenvalues of $-L$. This is possible since $-L$ is (formally) self-adjoint if and only if the operator L itself is (formally) self-adjoint.

Thus far we have obtained some conditions for the self-adjoint operator in this section and they will be summarized in the following theorem.

Theorem 4.1. *Let $L : \mathcal{L}^2(a,b) \cap C^2(a,b) \to \mathcal{L}^2(a,b)$ be a second-order linear differential operator defined as follows*

$$Lu = p(x)u'' + q(x)u' + r(x)u, \quad x \in (a,b),$$

where $p \in C^2(a,b)$, $q \in C^1(a,b)$ and $r \in C(a,b)$. Then the following statements hold

i) L is formally self-adjoint, i.e. $L^* = L$, if the coefficients p, q and r are real and $q = p'$.

ii) L is self-adjoint, i.e. $L' = L$, if L is formally self-adjoint and equation (35) holds.

iii) If L is self-adjoint this implies the eigenvalues λ of equation (36) are all real and when two eigenvalues are distinct, their corresponding eigenfunctions are orthogonal in $\mathcal{L}^2(a,b)$.

Proof. The first two statements have already been discussed and demonstrated in this section, so we're left with proving the last statement. Let's assume $\lambda \in \mathbb{C}$ is an eigenvalue of $-L$. This implies there exists an associated eigenfunction $f \in \mathcal{L}^2(a,b) \cap C^2(a,b)$ which is nonzero and for which the following equation holds

$$Lf + \lambda f = 0.$$

Therefore we must have

$$\lambda \|f\|^2 = \langle \lambda f, f\rangle = -\langle Lf, f\rangle, \tag{37}$$

where we substituted $\lambda f = -Lf$. Since we assumed L is self-adjoint we also have

$$-\langle Lf, f\rangle = -\langle f, Lf\rangle = \langle f, \lambda f\rangle = \overline{\lambda}\|f\|^2. \tag{38}$$

Combining equations (37) and (38) gives

$$\overline{\lambda}\|f\|^2 = \lambda\|f\|^2,$$

which implies $\overline{\lambda} = \lambda$ since $\|f\| \neq 0$. Since λ was chosen arbitrarily, all eigenvalues have to be real.
Suppose we have another eigenvalue μ of $-L$ with corresponding eigenfunction $g \in \mathcal{L}^2(a,b) \cap C^2(a,b)$. Then we have

$$\lambda\langle f, g\rangle = -\langle Lf, g\rangle = -\langle f, Lg\rangle = \mu\langle f, g\rangle,$$

which implies $(\lambda - \mu)\langle f, g\rangle = 0$. Since we assumed $\lambda \neq \mu$ we must have $\langle f, g\rangle = 0$, so the eigenfunctions f and g are orthogonal. This completes the proof. $\square$

The third statement of theorem 4.1 can be generalized to differential operators which are not even formally self-adjoint. This can be done by defining a positive function w such that the operator wL is formally self-adjoint. This w can be regarded as a sort of weight function. Multiplying the eigenvalue equation (36) by w yields

$$wLu + \lambda wu = 0.$$

Now we want $\tilde{L} := wL$ to be formally self-adjoint. Therefore we multiply the operator L given in (30) by w which gives

$$\tilde{L} = wp\frac{\mathrm{d}^2}{\mathrm{d}x^2} + wq\frac{\mathrm{d}}{\mathrm{d}x} + wr.$$

We have seen that an operator is formally self-adjoint if the second coefficient equals the derivative of the first coefficient. For this operator $\tilde{L}$ that is

$$wq = (wp)' = w'p + wp',$$

and its solution is given by

$$w(x) = \frac{c}{p(x)} exp\left(\int_a^x \frac{q(t)}{p(t)}\,\mathrm{d}t\right),$$

where c is a constant. We see when L itself is formally self-adjoint, i.e. $q = p'$, $w(x)$ becomes a constant as we expected. Furthermore if the equivalent of (35), given by

$$wp(f'\overline{g} - f\overline{g}')|_a^b = 0$$

still holds, the operator wL is even self-adjoint. Now it's not hard to show that for positive function w which is continuous on $[a, b]$ the following statements again hold

i) every eigenvalue of $Lu + \lambda wu = 0$ *is real and*

ii) if two eigenfunctions are distinct, they are orthogonal.

This is a generalisation of the third statement in theorem 4.1 since it is not required for the operator L to be (formally) self-adjoint anymore.

5 The Spectral Theory

In this section we will discuss the Sturm-Liouville problem and the spectral properties related to the Sturm-Liouville differential operator. Again we consider the following formally self-adjoint operator

$$L = \frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}}{\mathrm{d}x}\right) + r(x), \tag{39}$$

with the following eigenvalue equation

$$Lu + \lambda w(x)u = 0.$$

If it satisfies the following separated homogeneous boundary conditions

$$\begin{aligned} \alpha_1 u(a) + \alpha_2 u'(a) &= 0, & |\alpha_1| + |\alpha_2| &> 0, \\ \beta_1 u(b) + \beta_2 u'(b) &= 0, & |\beta_1| + |\beta_2| &> 0, \end{aligned} \tag{40}$$

with $\alpha_i, \beta_i \in \mathbb{R}$, we call this the *Sturm-Liouville eigenvalue problem.* From the last section we know the function w makes L self-adjoint and therefore the eigenvalues are real and their corresponding eigenfunctions are orthogonal. Furthermore we note that 0 is not an eigenvalue of this operator. The proof is elementary and will therefore be omitted. As stated in section 2.1 the problem is called *regular* if the interval (a, b) is bounded and $p \neq 0$ on $[a, b]$, otherwise it's called *singular*. We will actually only consider the regular problem and in this case we are able to assume without loss of generality that $p(x) > 0$. Then the eigenfunctions of the operator $-\frac{L}{w}$ solve the problem.

The goal of this section is to prove that the solutions of the Sturm-Liouville problem span the whole space $\mathcal{L}^2(a, b)$. Throughout this section we shall take $w(x) = 1$ since it simplifies proofs and calculations, but still covers the general idea. Eventually we will analyse the spectral properties of L^{-1} which is related to our original operator in the following way. The eigenfunctions of $-L$ coincide with the eigenfunctions of $-L^{-1}$ and its eigenvalue equation is given by

$$L^{-1}u + \mu u = 0.$$

In this case the eigenvalues are related by the identity $\mu = \frac{1}{\lambda}$. This analysis of the spectral properties will be done by exploring an integral expression for L^{-1} denoted by the operator T, making use of *Green's function.* This is a continuous C^2 function $G : [a, b] \times [a, b] \to \mathbb{R}$ which is symmetric for the Sturm-Liouville eigenvalue problem and satisfies the following equation

$$L_x G(x, \xi) = p(x)G_{xx}(x, \xi) + p'(x)G_x(x, \xi) + r(x)G(x, \xi) = 0, \tag{41}$$

whenever $x \neq \xi$. Furthermore its derivative with respect to x has a jump discontinuity at ξ which is given by

$$\frac{\partial G}{\partial x}\left(\xi^+, \xi\right) - \frac{\partial G}{\partial x}\left(\xi^-, \xi\right) = \frac{1}{p\left(\xi\right)}. \tag{42}$$

5.1 Existence of the eigenvalues and eigenfunctions

First the existence of the eigenvalues and the eigenfunctions of the Sturm-Liovuille problem will be proven. Therefore we will build up this Green's function for the operator L satisfying the boundary conditions given in (40). Using the existence and uniqueness theorem for second-order differential equations, we know there must exist two solutions v_1 and v_2 to the eigenvalue equation $Lu = 0$ which are unique and nontrivial. Furthermore they satisfy

$$\begin{aligned} v_1(a) &= \alpha_2, \qquad & v_1'(a) &= -\alpha_1, \\ v_2(b) &= \beta_2, \qquad & v_2'(b) &= -\beta_1. \end{aligned}$$

Using these values for v_1 and v_2 we see they satisfy the boundary conditions given in (40) since we have

$$\begin{aligned} \alpha_1 v_1(a) + \alpha_2 v_1'(a) &= \alpha_1\alpha_2 - \alpha_2\alpha_1 = 0, \\ \beta_1 v_2(b) + \beta_2 v_2'(b) &= \beta_1\beta_2 - \beta_2\beta_1 = 0. \end{aligned}$$

Since 0 is not an eigenvalue of the operator L, the solutions v_1 and v_2 must be linearly independent. Now the Green's function is defined in the following way

$$G(x,\xi) = \begin{cases} c^{-1} v_1(\xi) v_2(x), & a \le \xi \le x \le b, \\ c^{-1} v_1(x) v_2(\xi), & a \le x \le \xi \le b, \end{cases} \tag{43}$$

where $c = p(x)\left(v_1(x)v_2'(x) - v_1'(x)v_2(x)\right)$. Since $p \neq 0$, proving this c is a nonzero constant boils down to showing the derivative of $p(x)(v_1(x)v_2'(x) - v_1'(x)v_2(x))$ equals zero as follows.

Proof.

$$\begin{aligned} 0 = v_1 L v_2 - v_2 L v_1 &= v_1\left((pv_2')' + rv_2\right) - v_2\left((pv_1')' + rv_1\right) \\ &= v_1 (pv_2')' + v_1 r v_2 - v_2 (pv_1')' - v_2 r v_1 = v_1 (pv_2')' - v_2 (pv_1')' \\ &= v_1 p' v_2' + v_1 p v_2'' - v_2 p' v_1' - v_2 p v_1'' \\ &= p' v_1 v_2' - p' v_1' v_2 + p v_1 v_2'' + p v_1' v_2' - p v_1' v_2' - p v_1'' v_2 \\ &= \left(p\left(v_1 v_2' - v_1' v_2\right)\right)', \end{aligned}$$

which is known as the Langrange identity. □

As noted before, the Green's function defined in (43) is indeed symmetric and actually satisfies equations (41) and (42). Finally, using G we are able to define this operator T which we wanted to be an integral expression for L^{-1}. We will prove this by showing that the function Tf solves the differential equation $Lu = f$ and also the reverse, i.e. Lu solves $Tf = u$. Therefore we must show Tf is a C^2 function as well, but let us first define this operator T as follows

$$(Tf)(x) = \int_a^b G(x,\xi) f(\xi)\,\mathrm{d}\xi. \tag{44}$$

Using the continuity of G and f at $\xi = x$, the continuity of v_1 and v_2 and the property of G given in (42) we obtain the following three expressions by differentiating equation (44)

$$
\begin{aligned}
(Tf)(x) &= \int_a^x G(x,\xi)f(\xi)\,\mathrm{d}\xi + \int_x^b G(x,\xi)f(\xi)\,\mathrm{d}\xi, \\
(Tf)'(x) &= \int_a^x G_x(x,\xi)f(\xi)\,\mathrm{d}\xi + \int_x^b G_x(x,\xi)f(\xi)\,\mathrm{d}\xi, \\
(Tf)''(x) &= \int_a^x G_{xx}(x,\xi)f(\xi)\,\mathrm{d}\xi + \int_x^b G_{xx}(x,\xi)f(\xi)\,\mathrm{d}\xi + \frac{f(x)}{p(x)}.
\end{aligned}
$$

The elaborations of the first and second derivative can be found on page 13 of [1]. Because of the jump discontinuity of the Green's function at $x = \xi$ given in (42) these proofs involve some subtleties wherefore they will be left out of this thesis. Moreover, in the second expression for the first derivative $(Tf)'(x)$ we made use of the continuity of G and f at $\xi = x$ and we note that the derivatives of limits cancel out.
From the last expression it follows that $Tf \in C^2([a,b])$. Using the formulas above and the equation for $L_xG(x,\xi)$ given in (41), applying the operator L on the function Tf yields

$$
\begin{aligned}
L(Tf)(x) &= p(x)(Tf)''(x) + p'(x)(Tf)'(x) + r(x)(Tf)(x) \\
&= \int_a^x L_xG(x,\xi)f(\xi)\,\mathrm{d}\xi + \int_x^b L_xG(x,\xi)f(\xi)\,\mathrm{d}\xi + f(x) \\
&= f(x),
\end{aligned}
$$

as desired. We also made use of the property of the Green's function given by equation (41) from which we know $L_xG(x,\xi) = 0$ for every $\xi \neq x$. Furthermore, because G is symmetric it follows that Tf satisfies the boundary conditions of the Sturm-Liouville eigenvalue problem given in (40).
On the other hand, if we have another function $u \in C^2([a,b])$ which satisfies the same boundary conditions, we can prove the reverse statement. Making use of the fact that p, u and u' are continuous and also using the properties which the Green's function satisfies, we obtain

$$T(Lu)(x) = u(x).$$

This is just a matter of substituting $Lu(x)$ in equation (44) and applying integration by parts to the result. Therefore the proof will be omitted.
So we have actually seen $L(Tf) = f$ and $T(Lu) = u$, thus the operator T can be considered as an inverse of L. This is exactly the operator we searched for since it is an integral expression of L^{-1}. So we see the Sturm-Liouville eigenvalue equation given by

$$Lu + \lambda u = 0,$$

subject to the separated homogeneous boundary conditions given in (40) corresponds to the eigenvalue equation

$$Tu = \mu u,$$

with $\mu = -\frac{1}{\lambda}$. Therefore the eigenfunction u is a solution of the Sturm-Liouville problem corresponding to the the eigenvalue λ if and only if u is an eigenfunction of T corresponding to the eigenvalue $-\frac{1}{\lambda}$. Since it takes significantly less effort, we shall explore the spectral properties of the integral operator T, as they present information about the spectral properties of the Sturm-Liouville problem itself.

Let us first look at the eigenvalues of the Sturm-Liouville problem. We have already noted 0 is not an eigenvalue of $-L$ and the next lemma proves there exist more real numbers which are not an eigenvalue of this operator.

Lemma 5.1. *The eigenvalues of $-L$ are bounded below by a real constant.*

Proof. Consider a function $u \in C^2([a,b])$ satisfying the boundary conditions given in (40). To say something about the value of the eigenvalues we will first evaluate the inner product of $-Lu$ with u using (17). Applying (39) and using the boundary conditions to replace u' we get

$$\begin{aligned}\langle -Lu, u\rangle &= \langle -(pu')' - ru, u\rangle = \int_a^b \left(-(pu')'\overline{u} - r|u|^2\right) \mathrm{d}x \\ &= \left[-p(x)u'(x)\overline{u}(x)\right]_a^b - \int_a^b -p|u'|^2 \mathrm{d}x + \int_a^b -r|u|^2 \mathrm{d}x \\ &= \int_a^b \left(p|u'|^2 - r|u|^2\right) \mathrm{d}x + p(a)u'(a)u(a) - p(b)u'(b)u(b) \\ &= \int_a^b \left(p|u'|^2 - r|u|^2\right) \mathrm{d}x + p(b)\frac{\beta_1}{\beta_2}u^2(b) - p(a)\frac{\alpha_1}{\alpha_2}u^2(a),\end{aligned}$$

where we made use of integration by parts in the second step. If α_2 or β_2 equals 0 the boundary conditions imply $u(a) = 0$ and $u(b) = 0$ respectively. Therefore the second or third term above drops out. In case they both equal zero we can already state the following concerning the eigenvalue λ corresponding to the eigenfunction u

$$\begin{aligned}\lambda\|u\|^2 = \langle -Lu, u\rangle &= \int_a^b p(x)|u'(x)|^2 \,\mathrm{d}x - \int_a^b r(x)|u(x)|^2 \,\mathrm{d}x \\ &\geq -\|u\|^2 \max\{|r(x)| : a \leq x \leq b\}.\end{aligned}$$

So if we define $\ell := -\max\{|r(x)| : a \leq x \leq b\}$, we see this ℓ is a lower bound for λ since $u \in C^2([a,b])$ was chosen arbitrary.
On the other hand if α_2 and β_2 do not equal zero, we can show by contradiction $-L$ does not have more than two linearly independent eigenfunctions less than this lower bound ℓ. So let's assume our operator has three linearly independent eigenfunctions u_1, u_2 and u_3. Corresponding to them we have the eigenvalues λ_1, λ_2 and λ_3 respectively, which all three are less than ℓ. Furthermore we suppose the eigenfunctions are orthonormal. As a reminder, this implies the inner product of u_i with u_j for $i \neq j$ equals zero and the norm of every u_i equals 1. Now we define the following eigenfunction

$$v(x) = c_1u_1(x) + c_2u_2(x) + c_3u_3(x),$$

where c_1, c_2 and c_3 are constants and for which we have $v(a) = v(b) = 0$. This follows directly from the fact that u_1, u_2 and u_3 must satisfy the boundary conditions given in (40), which implies the following two identities hold

$$c_1 u_1(a) + c_2 u_2(a) + c_3 u_3(a) = 0,$$
$$c_1 u_1(b) + c_2 u_2(b) + c_3 u_3(b) = 0.$$

Since $v(a) = v(b) = 0$ we must have from the argument above that the eigenvalue corresponding to v is bounded below by ℓ. However, here we stumble upon a contradiction since the following inequality must also hold

$$\langle -Lv, v\rangle = \lambda_1 |c_1|^2 + \lambda_2 |c_2|^2 + \lambda_3 |c_3|^2 < \ell \left(|c_1|^2 + |c_2|^2 + |c_3|^2\right) = \ell \|v\|^2,$$

where we used the orthonormality property of the eigenfunctions u_i. Therefore there exist at most two linearly independent eigenvalues less than ℓ in the case of α_2 and β_2 being unequal to zero and none less than ℓ if α_2 and β_2 both equal zero. So the eigenvalues of $-L$ are bounded below by a real constant. □

By determining an eigenvalue of T we are able to show its existence. Therefore we first want to prove the set of functions Tu contains a sequence which is uniformly convergent on $[a, b]$ to a continuous function. We shall see this function is an eigenfunction of the operator T corresponding to the eigenvalue we are searching for. To prove that it contains a uniformly convergent sequence we will use the following theorem, also known as the Ascoli-Arzela theorem which can be found on pages 28-31 of [9].

Theorem 5.2. *Let F be an infinite, uniformly bounded, and equicontinuous set of functions on the bounded interval $[a, b]$. Then F contains a sequence $(f_k : k \in \mathbb{N})$ which is uniformly convergent on $[a, b]$ to a function which is continuous on $[a, b]$.*

Before proving $\{Tu\}$ is *uniformly bounded* and *equicontinuous* we first want to clarify these concepts. Again consider the infinite set F consisting of functions which are continuous on $[a, b]$ satisfying the following inequality

$$|f(x)| \leq M \quad \text{for all}\, f \in F \text{ and all}\, x \in [a, b],$$

where M is a positive number. Then we call the set F uniformly bounded. On the other hand, if for every $\epsilon > 0$ there exists a $\delta > 0$ such that $|f(x) - f(\xi)| < \epsilon$ for all $f \in F$ and all $x, \xi \in [a, b]$ whenever $|x - \xi| < \delta$, then we call the infinite set F equicontinuous on $[a, b]$. Note that δ may depend on ϵ, but not on x, ξ or f. Making use of these concepts, we are now able to prove the next theorem by showing Tu is uniformly bounded and equicontinuous, and subsequently using the Ascoli-Arzela theorem.

Theorem 5.3. *The set of functions Tu, with $u \in C([a, b])$ and $\|u\| \leq 1$, contains a sequence which is uniformly convergent to a continuous function on $[a, b]$.*

Proof. First of all we know $|G(x, \xi)|$ is uniformly continuous and bounded by some positive constant M since the Green's function G is continuous on $[a, b] \times [a, b]$. Therefore, consecutively using the expression for the operator T given in (44), the inner product given in (17) and the Cauchy-Schwarz inequality (19), we obtain

$$\begin{aligned} |Tu(x)| &= \left| \int_a^b G(x,\xi)u(\xi)\,\mathrm{d}\xi \right| = \left| \int_a^b G(x,\xi)\overline{\overline{u(\xi)}}\,\mathrm{d}\xi \right| \\ &= |\langle G(x,\xi), u(\xi)\rangle| \leq \|G\|\|u\| \leq M\sqrt{b-a}\|u\|. \end{aligned} \tag{45}$$

Since we assumed $\|u\| \leq 1$ we see the set Tu is uniformly bounded since $|Tu(x)| \leq M\sqrt{b-a}\|u\| \leq M\sqrt{b-a}$. On the other hand, again using the uniform continuity of G we must have that for any $\epsilon > 0$ there exists a $\delta > 0$ such that for $x_1, x_2 \in [a,b]$, $|x_2 - x_1| < \delta$ implies $|G(x_2,\xi) - G(x_1,\xi)| < \epsilon$ for all $\xi \in [a,b]$. Since $u \in C([a,b])$ we therefore have that whenever $|x_2 - x_1| < \delta$ this implies $|Tu(x_2) - Tu(x_1)| \leq \epsilon\sqrt{b-a}\|u\| \leq \epsilon\sqrt{b-a}$. So besides being uniformly bounded we notice the set of functions Tu for which we have $u \in C([a,b])$ and $\|u\| \leq 1$, is also equicontinuous since our choice for δ is independent of x_1, x_2 and u. ☐

Applying the Ascoli-Arzela theorem given in theorem 5.2 we see we have proven Tu contains a uniformly convergent subsequence and this enables us to make a claim concerning the existence of an eigenvalue of T. Therefore we want to define the norm of this operator as follows

$$\|T\| = \sup\{\|Tu\| : u \in C([a,b]), \|u\| = 1\}.$$

Using this norm it is not too hard to show the first of the following two identities hold for all continuous functions u on $[a,b]$

$$\begin{aligned} &1. \quad \|Tu\| \leq \|T\|\|u\|, \\ &2. \quad \|T\| = \sup_{\|u\|=1} |\langle Tu, u\rangle|. \end{aligned} \tag{46}$$

The second statement is not completely trivial but can be found on pages 234-235 in [4]. These identities will be useful in proving the following theorem about the existence of an eigenvalue of T.

Theorem 5.4. *Either $\|T\|$ or $-\|T\|$ is an eigenvalue of the operator T.*

Proof. First of all, since $\langle Tu, u\rangle$ is a real number we have using the second identity in (46) that either $\|T\| = \sup\langle Tu, u\rangle$ or $\|T\| = -\inf\langle Tu, u\rangle$ with the norm of u being equal to 1. We will assume $\|T\| = \sup\langle Tu, u\rangle$ since the proof for $-\|T\| = \inf\langle Tu, u\rangle$ is similar. This assumption implies there must exist a sequence of continuous functions $\{u_k\}$ on $[a,b]$ for which $\langle Tu_k, u_k\rangle \to \|T\|$ as $k \to \infty$. According to theorem 5.2 the sequence $\{Tu_k\}$ contains a subsequence $\{Tu_{k_j}\}$ which converges uniformly to a continuous function which we will denote by φ_0. We shall see this φ_0 actually is an eigenfunction of T and its corresponding eigenvalue, which we will denote by μ_0, indeed equals $\|T\|$. So since $\{Tu_{k_j}\}$ converges uniformly to φ_0 we know

$$\sup_{x\in[a,b]} |Tu_{k_j}(x) - \varphi_0(x)| \to 0.$$

Therefore, using (46) we have $\|Tu_{k_j} - \varphi_0\| \to 0$ which in turn implies $\|Tu_{k_j}\| \to \|\varphi_0\|$. Using this and the fact that $\langle Tu_{k_j}, u_{k_j}\rangle \to \|T\| = \mu_0$ we obtain the following identity

$$\begin{aligned}\|Tu_{k_j} - \mu_0 u_{k_j}\|^2 &= \|Tu_{k_j}\|^2 + \|\mu_0 u_{k_j}\|^2 - 2\langle Tu_{k_j}, \mu_0 u_{k_j}\rangle \\ &= \|Tu_{k_j}\|^2 + |\mu_0|^2\|u_{k_j}\|^2 - 2\mu_0\langle Tu_{k_j}, u_{k_j}\rangle \\ &= \|Tu_{k_j}\|^2 + \mu_0^2 - 2\mu_0\langle Tu_{k_j}, u_{k_j}\rangle \\ &\to \|\varphi_0\|^2 + \mu_0^2 - 2\mu_0^2 = \|\varphi_0\|^2 - \mu_0^2.\end{aligned} \tag{47}$$

Since $\|Tu_{k_j} - \mu_0 u_{k_j}\|^2$ is always greater than or equal to zero on $[a, b]$ we must have $\|\varphi_0\|^2 \geq \mu_0^2 \geq 0$. Furthermore, using (46) we have $\|Tu_{k_j}\|^2 \leq \|T\|^2\|u_{k_j}\|^2 = \|T\|^2 = \mu_0^2$. This implies (47) boils down to

$$0 \leq \|Tu_{k_j} - \mu_0 u_{k_j}\|^2 \leq \mu_0^2 + \mu_0^2 - 2\mu_0\langle Tu_{k_j}, u_{k_j}\rangle = 2\mu_0^2 - 2\mu_0\langle Tu_{k_j}, u_{k_j}\rangle.$$

Since $\langle Tu_{k_j}, u_{k_j}\rangle \to \mu_0$ the right hand side of the expression above goes to zero. Therefore we obtain $\|Tu_{k_j} - \mu_0 u_{k_j}\| \to 0$. Now we will rewrite the norm of $T\varphi_0 - \mu_0\varphi_0$ in the following way using the triangle inequality given in (19)

$$\begin{aligned}0 \leq \|T\varphi_0 - \mu_0\varphi_0\| &= \|T\varphi_0 - T(Tu_{k_j}) + T(Tu_{k_j}) - \mu_0 Tu_{k_j} + \mu_0 Tu_{k_j} - \mu_0\varphi_0\| \\ &\leq \|T\varphi_0 - T(Tu_{k_j})\| + \|T(Tu_{k_j}) - \mu_0 Tu_{k_j}\| + \|\mu_0 Tu_{k_j} - \mu_0\varphi_0\|.\end{aligned}$$

If we use the just obtained results we see the left hand side tends to zero as j goes to infinity. Therefore it follows that $\|T\varphi_0 - \mu_0\varphi_0\|$ must equal zero, which implies $T\varphi_0(x) = \mu_0\varphi_0(x)$ for all $x \in [a, b]$ because $T\varphi_0 - \mu_0\varphi_0$ is continuous. This completes the proof since we have shown that $\mu_0 = \|T\|$ is an eigenvalue of the operator T corresponding to the eigenfunction φ_0. □

We shall now use this result in the proof of the following theorem regarding the existence of eigenfunctions of T.

Theorem 5.5. *The operator T has an infinite sequence of orthonormal eigenfunctions in the $\mathcal{L}^2(a, b)$ space.*

Proof. We will prove this theorem inductively, creating a sequence of eigenfunctions in the following way. Using φ_0, μ_0, the Green's function G and the self-adjoint operator T we shall define the normalized eigenfunction ψ_0, a new function G_1 and a new self-adjoint operator T_1 as follows

$$\begin{aligned}\psi_0 &:= \frac{\varphi_0}{\|\varphi_0\|}, \\ G_1(x, \xi) &:= G(x, \xi) - \mu_0\psi_0(x)\overline{\psi_0}(\xi), \\ (T_1 u)(x) &:= \int_a^b G_1(x, \xi)u(\xi)\,\mathrm{d}\xi = \int_a^b \big(G(x, \xi) - \mu_0\psi_0(x)\overline{\psi_0}(\xi)\big)\,u(\xi)\,\mathrm{d}\xi \\ &= \int_a^b G(x, \xi)u(\xi)\,\mathrm{d}\xi - \mu_0\psi_0(x)\int_a^b u(\xi)\overline{\psi_0}(\xi)\,\mathrm{d}\xi \\ &= Tu(x) - \mu_0\langle u, \psi_0\rangle\psi_0(x),\end{aligned} \tag{48}$$

which holds for all continuous functions u on $[a, b]$. Defining it this way we see G and G_1 satisfy the same regularity properties and moreover G_1 is symmetric, i.e. $G_1(x,\xi) = G_1(\xi,x)$ for all $x,\xi \in [a,b]$ following from G being symmetric itself. This implies the set of functions $T_1 u$ is also uniformly bounded and equicontinuous for $\|u\| \leq 1$. Furthermore, if we define $|\mu_1| := \sup |\langle T_1 u, u\rangle|$ for $\|u\| = 1$ and assume $\|T_1\| \neq 0$, we see $\|T_1\| = \mu_1$ is an eigenvalue of the self-adjoint operator T_1. It corresponds to the continuous eigenfunction φ_1 and therefore it satisfies the following equation

$$T_1\varphi_1 = \mu_1\varphi_1.$$

Because φ_0 is an eigenfunction of T corresponding to the eigenvalue μ_0 we must also have $T\psi_0 = \mu_0\psi_0$ for the normalized eigenfunction ψ_0. Using this we obtain

$$\begin{aligned}\langle T_1 u, \psi_0\rangle &= \langle Tu - \mu_0\langle u,\psi_0\rangle\psi_0, \psi_0\rangle = \langle Tu, \psi_0\rangle - \mu_0\langle\langle u,\psi_0\rangle\psi_0,\psi_0\rangle \\ &= \langle u, T\psi_0\rangle - \mu_0\langle u,\psi_0\rangle = \langle u, \mu_0\psi_0\rangle - \mu_0\langle u,\psi_0\rangle = 0,\end{aligned} \tag{49}$$

where we made use of the operator T being self-adjoint and the eigenfunction ψ_0 being normalized, thus its norm being equal to 1. Since this holds for all continuous functions u on $[a,b]$ we can now normalize the eigenfunction φ_1 of T_1, i.e. $\psi_1 := \frac{\varphi_1}{\|\varphi_1\|}$ and see ψ_1 is orthogonal to ψ_0. This results from substituting ψ_1 for u in expression (49) as follows

$$0 = \langle T_1\psi_1, \psi_0\rangle = \langle \mu_1\psi_1, \psi_0\rangle = \mu_1\langle\psi_1,\psi_0\rangle.$$

Since we assumed $\mu_1 \neq 0$ we see the inner product of ψ_1 and ψ_0 equals zero which implies they are orthogonal. Using the definition of T_1 given in (48), applying this operator to ψ_1 yields

$$T_1\psi_1 = T\psi_1 - \mu_0\langle\psi_1,\psi_0\rangle\psi_0 = T\psi_1, \tag{50}$$

where we used the orthogonality of ψ_0 with ψ_1. Since $T_1\psi_1 = \mu_1\psi_1$ we see $T\psi_1 = T_1\psi_1 = \mu_1\psi_1$ which implies ψ_1 is also an eigenfunction of the operator T with corresponding eigenvalue μ_1. Using expression (50), the definitions of μ_0 and μ_1, and the inequality given in (46) we obtain

$$|\mu_1| = |\mu_1|\|\psi_1\| = \|\mu_1\psi_1\| = \|T\psi_1\| \leq \|T\|\|\psi_1\| = \|T\| = |\mu_0|.$$

In the same way we shall construct G_2 and T_2 for which we obtain a new eigenfunction. Let's define them as follows

$$\begin{aligned}G_2(x,\xi) &:= G_1(x,\xi) - \mu_1\psi_1(x)\overline{\psi_1}(\xi) = G(x,\xi) - \mu_0\psi_0(x)\overline{\psi_0}(\xi) - \mu_1\psi_1(x)\overline{\psi_1}(\xi) \\ &= G(x,\xi) - \sum_{k=0}^{1}\mu_k\psi_k(x)\overline{\psi_k}(\xi), \\ (T_2u)(x) &:= \int_a^b G_2(x,\xi)u(\xi)\,\mathrm{d}\xi = \int_a^b \left(G(x,\xi) - \sum_{k=0}^{1}\mu_k\psi_k(x)\overline{\psi_k}(\xi)\right)u(\xi)\,\mathrm{d}\xi \\ &= \int_a^b G(x,\xi)u(\xi)\,\mathrm{d}\xi - \sum_{k=0}^{1}\left(\mu_k\psi_k(x)\int_a^b u(\xi)\overline{\psi_k}(\xi)\,\mathrm{d}\xi\right) \\ &= Tu(x) - \sum_{k=0}^{1}\mu_k\langle u,\psi_k\rangle\psi_k(x).\end{aligned}$$

From this we obtain another normalized eigenfunction of T which we will denote by ψ_2 corresponding to the eigenvalue μ_2. Again we are able to show this eigenfunction is orthonormal to the other two eigenfunctions ψ_0 and ψ_1 and moreover the eigenvalue satisfies $|\mu_2| \leq |\mu_1|$. Proceeding this way we can construct a sequence of eigenfunctions $\psi_0, \psi_1, \psi_2, \dots$ of the operator T which are orthonormal, and they are associated with the sequence of eigenvalues $|\mu_0| \geq |\mu_1| \geq |\mu_2| \geq \dots$. Then we obtain the following expressions for the iterated Green's function and integral operator

$$\begin{aligned} G_n(x,\xi) &= G(x,\xi) - \sum_{k=0}^{n-1} \mu_k \psi_k(x)\overline{\psi_k}(\xi), \\ (T_n u)(x) &= \int_a^b G_n(x,\xi)u(\xi)\,\mathrm{d}\xi = Tu(x) - \sum_{k=0}^{n-1} \mu_k \langle u, \psi_k\rangle \psi_k(x), \end{aligned} \tag{51}$$

where the norm of the operator T_n satisfies

$$\|T_n\| = |\mu_n|. \tag{52}$$

Obviously, if we have $|\mu_n| = \|T_n\| = 0$ for some n, this sequence of eigenvalues ends. However, it is not too hard to prove that $\|T_n\|$ is greater than zero for all $n \in \mathbb{N}$. This can be shown by a contradiction, so assuming $|\mu_n| = 0$ we have

$$\begin{aligned} 0 = L(0) = L(\mu_n u) = L(T_n u) &= L(Tu) - L\left(\sum_{k=0}^{n-1} \mu_k \langle u, \psi_k\rangle \psi_k\right) \\ &= u - \sum_{k=0}^{n-1} \mu_k \langle u, \psi_k\rangle L\psi_k \quad \Rightarrow \quad u = \sum_{k=0}^{n-1} \mu_k \langle u, \psi_k\rangle L\psi_k \\ &= \sum_{k=0}^{n-1} \langle u, \psi_k\rangle L\mu_k\psi_k = \sum_{k=0}^{n-1} \langle u, \psi_k\rangle LT\psi_k = \sum_{k=0}^{n-1} \langle u, \psi_k\rangle \psi_k. \end{aligned}$$

But this must hold for all continuous functions u on $[a,b]$ which contradicts the fact that there cannot exist a finite set of functions which spans whole $C([a,b])$. Therefore there must exist an infinite sequence of eigenfunctions, which completes the proof. □

5.2 Completeness of the eigenfunctions

In this section we shall prove that the eigenfunctions of the operator T are complete. Therefore we make use of theorem 3.11 in which Parseval's identity is related to the completeness property of an orthogonal set. Thus we have to prove that Bessel's inequality given in (28) is an equality, i.e.

$$f = \sum_{k=0}^{\infty} \langle f, \psi_k\rangle \psi_k.$$

The following theorem will be used to show this holds for any $f \in \mathcal{L}^2(a,b)$.

Theorem 5.6. *Given any $f \in C^2([a,b])$ subject to the separated homogeneous boundary conditions given in (40) the infinite series $\sum \langle f, \psi_k\rangle \psi_k$ is uniformly convergent to f on $[a,b]$.*

Proof. First we will prove $\sum_{k=0}^{\infty} \mu_k \langle u, \psi_k\rangle \psi_k$ is uniformly convergent to a continuous function on $[a, b]$. As we constructed the eigenfunctions for the operator T we saw for any function $\langle u, \psi_k\rangle\psi_k$ we have $T(\langle u, \psi_k\rangle\psi_k) = \mu_k(\langle u, \psi_k\rangle\psi_k)$. This implies

$$T\left(\sum_{k=m}^{n} \langle u, \psi_k\rangle\psi_k\right) = \sum_{k=m}^{n} \mu_k \langle u, \psi_k\rangle\psi_k,$$

for $n > m$. From expression (45) we know the set Tu is uniformly bounded since $|(Tu)(x)| \leq M\sqrt{b-a}\|u\|$ for all continuous functions u on $[a, b]$. Therefore, using the identity above and the definition of the norm given in (18) we obtain

$$\begin{aligned}
\left|\sum_{k=m}^{n} \mu_k \langle u, \psi_k\rangle\psi_k\right| = \left|T\left(\sum_{k=m}^{n} \langle u, \psi_k\rangle\psi_k\right)\right| &\leq M\sqrt{b-a}\left\|\sum_{k=m}^{n} \langle u, \psi_k\rangle\psi_k\right\| \\
&= M\sqrt{b-a}\sqrt{\sum_{k=m}^{n} \int_a^b |\langle u, \psi_k\rangle\psi_k|^2 \, \mathrm{d}x} \\
&= M\sqrt{b-a}\sqrt{\sum_{k=m}^{n} |\langle u, \psi_k\rangle|^2},
\end{aligned}$$

since ψ_k is normalized. From Bessel's inequality given in (28) we know the last expression must be less than or equal to $M\sqrt{b-a}\|u\|$. So if $m, n \to \infty$ this implies $|\sum_{k=m}^{n} \mu_k \langle u, \psi_k\rangle\psi_k| \to 0$ and therefore $\sum_{k=m}^{n} \mu_k \langle u, \psi_k\rangle\psi_k$ is uniformly convergent to a continuous function on $[a, b]$. Now we will show this continuous function turns out to be Tu. Therefore we will apply Bessel's inequality to the Green's function. First we notice by definition of the operator T given in (44) that

$$\langle G(x, \xi), \psi_k\rangle = \int_a^b G(x, \xi)\overline{\psi_k} \, \mathrm{d}\xi = T\overline{\psi_k}(x) = \mu_k \overline{\psi_k}(x),$$

for every $x \in [a, b]$. Using this expression and applying Bessel's inequality we obtain

$$\sum_{k=0}^{n} \mu_k^2 |\psi_k(x)|^2 = \sum_{k=0}^{n} |\mu_k \psi_k(x)|^2 = \sum_{k=0}^{n} |\langle G, \psi_k\rangle|^2 \leq \|G\|^2 = \int_a^b |G(x, \xi)|^2 \, \mathrm{d}\xi,$$

where the last equality follows from (18). As noticed before we know the Green's function is bounded by some positive constant M. Therefore, using the fact that ψ_k is normalized and integrating the expression above with respect to x yields

$$\sum_{k=0}^{\infty} \mu_k^2 \leq M^2(b-a)^2,$$

as $n \to \infty$. This implies we must have that $\lim_{n\to\infty} |\mu_n| = 0$. Now using the expression for T_n given in (51), the value of its norm given in (52) and the inequality given in (46) we obtain for any continuous function u on $[a, b]$

$$\|Tu - \sum_{k=0}^{n-1} \mu_k \langle u, \psi_k\rangle\psi_k\| = \|T_n u\| \leq \|T_n\|\|u\| = |\mu_n|\|u\|.$$

Since we just determined $\lim_{n\to\infty} |\mu_n| = 0$, the right hand side of the inequality above also tends to zero as $n \to \infty$. Because Tu is continuous we therefore have

$$Tu(x) = \sum_{k=0}^{\infty} \mu_k \langle u, \psi_k \rangle \psi_k(x), \tag{53}$$

for all $x \in [a, b]$. As we saw before, if f is continuous function on $[a, b]$ satisfying the boundary conditions of the Sturm-Liouville problem given in (40) we can define another continuous function u as follows

$$u = Lf \qquad \text{and} \qquad f = Tu,$$

since T acts as an inverse operator of L. Using this and the fact that T is formally self-adjoint we obtain

$$\mu_k \langle u, \psi_k \rangle = \langle u, \mu_k \psi_k \rangle = \langle u, T\psi_k \rangle = \langle Tu, \psi_k \rangle = \langle f, \psi_k \rangle,$$

as μ_k is the eigenvalue of the operator T corresponding to the normalized eigenfunction ψ_k. Substituting this into equation (53) we finally obtain

$$f(x) = Tu(x) = \sum_{k=0}^{\infty} \mu_k \langle u, \psi_k \rangle \psi_k(x) = \sum_{k=0}^{\infty} \langle f, \psi_k \rangle \psi_k(x),$$

which holds for all $x \in [a, b]$. Therefore $\sum \langle f, \psi_k \rangle \psi_k$ is uniformly convergent to $f \in C^2([a, b])$ and this completes the proof. □

Now using the density of $C^2([a, b])$ in $\mathcal{L}^2(a, b)$ already given in theorem 3.9 we are able to prove

$$\left\| f - \sum_{k=0}^{n} \langle f, \psi_k \rangle \psi_k \right\| \to 0$$

as $n \to \infty$. The proof goes too much into detail for this thesis and will therefore be omitted, but it can be found on pages 81-83 in [2]. However, we have shown that

$$\lim_{n\to\infty} \left\| f - \sum_{k=0}^{n} \langle f, \psi_k \rangle \psi_k \right\| = 0 \qquad \Longrightarrow \qquad f = \sum_{k=0}^{\infty} \langle f, \psi_k \rangle \psi_k,$$

which is equivalent to Parseval's identity given in (29). As proved before we see this shows that the set of eigenfunctions of the operator T is orthonormal and furthermore complete in $\mathcal{L}^2(a, b)$. We are now able to summarize this whole section. Using theorem 5.5 and the relation between the eigenvalues of the operators T and L given by $\mu = -\frac{1}{\lambda}$, we see $\lim_{n\to\infty} |\mu_n| = 0$ implies $\frac{1}{|\lambda_n|} = |\mu_n| \to 0$. Considering the eigenvalues are bounded below (lemma 5.1) we therefore have $\lambda_n \to \infty$ as $n \to \infty$. Also taking the weight function w into account we obtain the following fundamental theorem regarding the Sturm-Liouville eigenvalue problem, which was already stated in section 2.2.

Theorem 5.7. *Consider the following formally self-adjoint operator*

$$L = \frac{\mathrm{d}}{\mathrm{d}x} \left(p(x) \frac{\mathrm{d}}{\mathrm{d}x} \right) + r(x),$$

with the following eigenvalue equation

$$Lu + \lambda w(x)u = 0,$$

satisfying the following separated homogeneous boundary conditions

$$\begin{aligned} \alpha_1 u(a) + \alpha_2 u'(a) &= 0, & |\alpha_1| + |\alpha_2| > 0, \\ \beta_1 u(b) + \beta_2 u'(b) &= 0, & |\beta_1| + |\beta_2| > 0, \end{aligned}$$

with $\alpha_i, \beta_i \in \mathbb{R}$. *Suppose* $p', r, w \in C([a,b])$ *with* $p, w > 0$ *and* $x \in [a,b]$. *Then the Sturm-Liouville eigenvalue problem has an infinite sequence of real eigenvalues which can be ordered in the following way*

$$\lambda_0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n < \cdots ,$$

where $\lambda_n \to \infty$ *as* $n \to \infty$. *The eigenfunctions* φ_n *corresponding to each eigenvalue* λ_n *are unique and after normalizing the eigenfunctions to* ψ_n *they form an orthonormal basis of the* $\mathcal{L}^2(a,b)$ *space.*

6 Application of the Sturm-Liouville Theory

In this section we shall have a look at some applications of the Sturm-Liouville theory and an example will be elaborated. This illustrates the main results obtained in this thesis. The section ends with some concluding remarks concerning the Sturm-Liouville problem and its applications.

6.1 Applications

Besides the example of the sound of a guitar mentioned in the introduction, the Sturm-Liouville theory has many more applications. These equations occur frequently in applied mathematics as well as in physics. They describe the vibrations of a particular system, e.g. the vibrations of the plucked string of the guitar. Another example is the one-dimensional time-independent Schrödinger wave equation, where the eigenvalues represent the energy levels of the atomic system. These Sturm-Liouville problems do not only occur in one space dimension but also in higher dimensions and the results obtained in this thesis then still apply. Considering the wave equation, an example for dimension 2 is given by the resonant frequencies of a drum. If the dimension equals 3 we can think of the resonant frequencies of sound waves in a room.
One illustration of the one-dimensional Schrödinger equation is given by a crystal structure. This wave equation then represents the motion of a conduction electron in the crystal structure. The spectrum of the Sturm-Liouville eigenvalue equation can be seen as existing of intervals. The location of an eigenvalue in one of these intervals determines whether a solution is bounded or unbounded, where in the latter case it grows exponentially. Therefore some electrons can move freely through the crystal and the motion of others is bounded, since their energy lies in either of both intervals. This in turn clarifies whether the crystal acts like an insulator or conducts electricity. An example of a semiconductor is given by silicon, which is among other things used to create computer chips.
Another example of Sturm-Liouville equations are the Airy functions, which describe the change of a solution from oscillatory to exponential behaviour. This can be clarified by the oscillatory integrals for the Airy functions which have two stationary phase points. Illustrations of this transition from one state to the other are the caustics in light reflections or the passage of a particle from one region to another in semi-classical quantum mechanics. A very familiar example of a caustic is a rainbow.
The Airy functions can also be found in the study on linear dispersive water waves. For example, Diederik Korteweg and Gustav de Vries came up with a model consisting of a nonlinear, dispersive partial differential equation which describes waves on shallow water surfaces. Also Kelvin, who actually invented the method of stationary phase, studied the pattern of dispersive water waves made by a ship with constant speed.
A few other Sturm-Liouville equations are Bessel's equation, Legendre equations and Laguerre equations. However, they will not be discussed in this thesis but a very interesting lecture on Sturm-Liouvile eigenvalue problems can be found in chapter 4 of [5].

6.2 An example: the second derivative

At last we shall discuss an example which illustrates the concepts treated so far. We will have a look at one of the simplest forms of a Sturm-Liouville problem, i.e. the second derivative $-\left(\mathrm{d}^2/\mathrm{d}x^2\right)$. In view of equation (30) for the operator L we see this is the case for $p = -1$, $q = 0$ and $r = 0$. Theorem 4.1 now implies this operator is formally self-adjoint since all the coefficients are real and furthermore $p' = \frac{\mathrm{d}(-1)}{\mathrm{d}x} = 0 = q$. To determine its eigenvalues and eigenfunctions we need to solve the equation $Lu = \lambda u$, which implies

$$u'' + \lambda u = 0. \tag{54}$$

Let's assume it's subject to the following boundary conditions

$$u(0) = u(l) = 0,$$

for $0 \leq x \leq l$. We will explore this for different values of λ, so let's start with the case where the eigenvalue is strict positive. From ordinary differential equations we know the general solution is then given by

$$u(x) = c_1 \cos \sqrt{\lambda} x + c_2 \sin \sqrt{\lambda} x,$$

where c_1 and c_2 are constants. Using the boundary conditions we obtain

$$\begin{aligned} u(0) &= c_1 = 0, \\ u(l) &= c_2 \sin \sqrt{\lambda} l = 0. \end{aligned}$$

The last equality implies $\lambda_n = \frac{n^2\pi^2}{l^2}$ where $n \in \mathbb{N}$. Since both the eigenvalue equation and the boundary conditions are homogeneous we can take $c_2 = 1$ and obtain the following expression for the eigenfunctions corresponding to the eigenvalues λ_n

$$u_n(x) = \sin \frac{n\pi}{l} x.$$

Now considering the case where $\lambda \leq 0$, we obtain the following solutions for the differential equation

$$\begin{aligned} u(x) &= c_1 x + c_2 && \text{if } \lambda = 0, \\ u(x) &= c_1 \cosh \sqrt{-\lambda} x + c_2 \sinh \sqrt{-\lambda} x && \text{if } \lambda < 0. \end{aligned}$$

Using the boundary conditions both cases lead to the trivial solution $u = 0$ and therefore we conclude equation (54) only has positive eigenvalues. Thus the eigenvalues and eigenfunctions of this particular Sturm-Liouville problem are given by

$$\lambda_n = \frac{n^2\pi^2}{l^2} \qquad \text{and} \qquad u_n(x) = \sin \frac{n\pi}{l} x \qquad \text{for} \qquad n \in \mathbb{N}.$$

So we see the eigenvalues are real, form an infinite sequence and tend to infinity as $n \to \infty$, which was also concluded in theorem 5.7. Furthermore we can prove the eigenfunctions are orthogonal in the space $\mathcal{L}^2(0,l)$ by showing each inner product equals zero. For if we have $m \neq n$ and use definition (17) we get

$$\begin{aligned}\langle u_m, u_n \rangle &= \int_0^l \sin\frac{m\pi}{l}x \sin\frac{n\pi}{l}x \,\mathrm{d}x \\ &= \frac{1}{2}\int_0^l \left(\cos(m-n)\frac{\pi}{l}x - \cos(m+n)\frac{\pi}{l}x\right)\,\mathrm{d}x = 0,\end{aligned}$$

since m and n are integers. Finally, using theorem 5.7 we know the sequence $\{u_n\}$ of eigenfunctions is complete in $\mathcal{L}^2(0,l)$ and spans this space.
The completeness of the eigenfunctions implies we are able to represent any function $f \in \mathcal{L}^2(0,l)$ as follows

$$f(x) = \sum_{n=1}^{\infty} \alpha_n \sin\frac{n\pi}{l}x. \tag{55}$$

Note the equality holds in the $\mathcal{L}^2(0,l)$ space, which means it should be interpreted as follows

$$\left\| f(x) - \sum_{n=1}^{\infty} \alpha_n \sin\frac{n\pi}{l}x \right\| = 0.$$

In section 3.2.2 we already found an expression for the coefficient α_n given in (23), which for this example turns into

$$\alpha_n = \frac{\langle f, \sin(n\pi x/l)\rangle}{\|\sin(n\pi x/l)\|^2}, \tag{56}$$

where we substituted the eigenfunction $u_n(x)$ we just found. Evaluating the denominator using (18) yields

$$\left\|\sin\frac{n\pi}{l}x\right\|^2 = \int_0^l \sin^2\frac{n\pi}{l}x\,\mathrm{d}x = \frac{l}{2}.$$

Substituting this in our formula for the coefficient α_n given in (56) and evaluating the inner product we then obtain

$$\alpha_n = \frac{2}{l}\int_0^l f(x)\sin\frac{n\pi}{l}x\,\mathrm{d}x.$$

This can be illustrated by showing how to represent the constant function $f(x) = 1$ on $[0,l]$. The inner product of $f(x)$ with our eigenfunction u_n is given by

$$\langle 1, \sin(n\pi x/l)\rangle = \int_0^l \sin\frac{n\pi}{l}x\,\mathrm{d}x = \frac{l}{n\pi}(1-\cos n\pi) = \frac{l}{n\pi}(1-(-1)^n).$$

Therefore our coefficient α_n turns into

$$\alpha_n = \frac{2}{l}\langle 1, \sin(n\pi x/l)\rangle = \frac{2}{n\pi}(1-(-1)^n)$$

Substituting this in our formula for the representation of f given in (55) yields

$$1 = \frac{2}{\pi} \sum_{n=1}^{\infty} \left(\frac{1-(-1)^n}{n} \right) \sin \frac{n\pi}{l} x,$$

where again the equality is in the space $\mathcal{L}^2(0, l)$. Therefore we should interpret it as

$$\frac{2}{\pi} \sum_{k=1}^{n} \left(\frac{1-(-1)^k}{k} \right) \sin \frac{k\pi}{l} x \overset{\mathcal{L}^2}{\to} 1,$$

when $n \to \infty$.

6.3 Concluding remarks

In this section we have discussed some applications of the Sturm-Liouville theory and we looked at an example. This illustrated some of the spectral properties such as the infinite sequence of real eigenvalues and the representation of a function in the $\mathcal{L}^2$ space. We saw the set of functions $\{\sin(n\pi x/l) : n \in \mathbb{N}\}$ spans $\mathcal{L}^2(0, l)$. It can be shown that $\{\cos(n\pi x/l) : n \in \mathbb{N}\}$ also spans this space and their combination spans $\mathcal{L}^2(-l, l)$, leading to the Fourier series. However we shall not further digress on this, but it shows how comprehensive the Sturm-Liouville problem is. We could also have considered the singular case, i.e. when some of the initial conditions are not satisfied. For example the function $p(x)$ being equal to zero at any of the endpoints a or b, or the interval (a, b) being infinite. Then the spectral properties given in this paper are still satisfied and it even leads to a generalisation of the theory and its conditions. Well-known examples follow from this singular problem, e.g. Legendre's and Hermite's equation, or the one-dimensional time-independent Schrödinger equation. Thus we see the Sturm-Liouville theory is deeply investigated and utilized and still is to this day. Hence this thesis did not cover the entire research on this topic but it did work towards some of its main results to make the reader familiar with the Sturm-Liouville theory.

7 Acknowledgements

I would like to express my gratitude to my first supervisor dr. A.E. Sterk. Knowing my research would take longer, I chose to ask him to supervise my master thesis since he also guided me through my bachelor's project. I want to thank him for his adequate feedback and for his patience to let me work at my own time. He motivated me with his own passion for this topic and I appreciate his advise and suggestions when I encountered difficulties. It was an honour and pleasure to have him as my supervisor again.

I am also thankful to my family and friends for their support during my education. I could always rely on them and I appreciate their encouragements.

Groningen, July 2017
David Koning

8 References

[1] *Stochastic Systems*. Mathematics in Science and Engineering. Elsevier Science, 1983.

[2] M. Al-Gwaiz. *Sturm-Liouville Theory and its Applications*. Springer Undergraduate Mathematics Series. Springer London, 2008.

[3] N.L. Carothers. *Real Analysis*. Cambridge University Press, 2000.

[4] M. Haase. *Functional Analysis: An Elementary Introduction*. Graduate Studies in Mathematics. American Mathematical Society, 2014.

[5] John K. Hunter. *Lecture Notes On Applied Mathematics*. University of California, 2009.

[6] S.J. Leon. *Linear Algebra with Applications: Pearson New International Edition*. Pearson Education, Limited, 2013.

[7] J. Lützen. *Joseph Liouville, 1809-1882, master of pure and applied mathematics*. Studies in the history of mathematics and physical sciences. Springer-Verlag, 1990.

[8] Y. Pinchover and J. Rubinstein. *An Introduction to Partial Differential Equations*. Number v. 10 in An introduction to partial differential equations. Cambridge University Press, 2005.

[9] M. Reed and B. Simon. *I: Functional Analysis*. Methods of Modern Mathematical Physics. Elsevier Science, 1981.

[10] W. Rudin. *Principles of Mathematical Analysis*. International series in pure and applied mathematics. McGraw-Hill, 1976.

[11] W.A. Sutherland. *Introduction to Metric and Topological Spaces*. OUP Oxford, 2009.

[12] R. Thompson and W. Walter. *Ordinary Differential Equations*. Graduate Texts in Mathematics. Springer New York, 2013.

[13] A. Zettl. *Sturm-Liouville Theory*. Mathematical Surveys and Monographs. American Mathematical Society, 2005.