AN INTRODUCTION TO STURM-LIOUVILLE THEORY

ERIK BÉDOS

1.

The purpose of this note is to illustrate how the spectral theorem for a compact self-adjoint operator on a Hilbert space may be used to study some classical Sturm-Liouville problems. For simplicity we will only discuss the so-called regular case.

When $[a,b] \subset \mathbb{R}$ and $n \in \mathbb{N}$, we let $C^n([a,b])$ denote the space of n times continuously differentiable complex functions on $[a,b]$. A *regular Sturm-Liouville system* on some interval $[a,b]$ is a second order linear differential equation of the form

$$-(py')' + q\,y = \lambda\,\rho\,y\,, \tag{1}$$

where

- $p \in C^1([a,b])$ is real-valued and $p(x) \neq 0$ for all $x \in [a,b]$,
- $q,\ \rho \in C([a,b])$ are real-valued and $\rho(x) \neq 0$ for all $x \in [a,b]$,
- $\lambda \in \mathbb{C}$,

and the unknown function $y = y(x)$, which necessarily has to lie in $C^2([a,b])$, is required to satisfy boundary conditions of the type

$$\alpha_1\,y(a) + \alpha_2\,y'(a) = 0\,, \quad \beta_1\,y(b) + \beta_2\,y'(b) = 0\,, \tag{2}$$

for some (α_1, α_2), $(\beta_1, \beta_2) \in \mathbb{R}^2 \setminus \{(0,0)$.

Ideally, the Sturm-Liouville problem is to determine the values of λ for which there exist non-trivial solutions of equation (1) satisfying the conditions (2), and to describe these solutions. These values of λ are called the *eigenvalues* of the system, and the corresponding solutions y are called *eigenfunctions* of the system. A concrete answer to this problem is not possible in general, but as we will see, one may still obtain some valuable theoretical information about it.

Since we only intend to give a small taste of Sturm-Liouville theory, we will assume that $p(x) = \rho(x) = 1$ for all $x \in [a,b]$, in which case equation (1) simplifies to

$$-y'' + qy = \lambda y\,. \tag{3}$$

A suitably scaled version of this equation appears for example as the one dimensional time-independent Schrödinger equation in quantum mechanics (where it is usually considered on the whole real line).

Date: May 7, 2021.

Set $Y = \{y \in C^2([a,b]) : y$ satisfies the boundary conditions (2)$\}$. We will consider Y and $C([a,b])$ as inner product spaces w.r.t. to the inner product given by

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

Letting $D : Y \to C[a,b]$ be the linear operator defined by

$$D(y) = -y'' + qy\,,$$

it is clear that our Sturm-Liouville system may be written as

$$D(y) = \lambda y \quad \text{where } y \in Y. \tag{4}$$

Although the associated Sturm-Liouville problem looks like a familiar eigenvalue/eigenvector problem, it is not obvious how to proceed. The fact that Y and $C([a,b])$ are not Hilbert spaces (they are not complete) can easily be fixed because both can be considered as dense subspaces of $L^2([a,b])$. (We leave it as an exercise to show this for Y). However, the trouble is that D is not a bounded operator (check this!), so it does not extend to a bounded operator on $L^2([a,b])$. We will have to work quite a bit to recast the problem into one involving a compact selfadjoint operator.

- We will first study the differential equation

$$-y'' + qy = \lambda y \quad \text{with } y \in C^2([a,b]) \tag{5}$$

  and show that its solution space

$$S_\lambda := \{y \in C^2([a,b]) : -y'' + qy = \lambda y\}$$

  is 2-dimensional for every $\lambda \in \mathbb{C}$.[1] Note that trying to find out when there exists some $y \in S_\lambda \setminus \{0\}$ which also belongs to Y, which would solve our problem, is not possible because a concrete description of S_λ is not available in general.
- Next, we will establish some spectral properties of the operator D.
- Thirdly, we will assume that D is 1-1. We will then show that D is onto $C([a,b])$, and that there exists a compact self-adjoint operator $T_G : L^2([a,b]) \to L^2([a,b])$ such that its restriction to $C([a,b])$ is the inverse of D. Applying the spectral theorem to T_G will lead us to a theoretical answer to our Sturm-Liouville problem in this case.
- Finally, we will explain how to handle the general case where D is not assumed to be 1-1.

[1] This fact holds for the solution space of any homogeneous second order linear ordinary differential equation, as some students may have seen in a previous course. We will give a self-contained proof in our case.

2. About the differential equation $-y'' + qy = \lambda y$, $y \in C^2([a,b])$

We recall that $\lambda \in \mathbb{C}$. In this section it is not important that the function $q \in C([a,b])$ is assumed to be real-valued.

Theorem 2.1. *Let $c \in [a,b]$ and $z_1, z_2 \in \mathbb{C}$. Then there exists a unique function $y \in C^2([a,b])$ satisfying that*

$$-y'' + qy = \lambda y \quad \text{and} \quad y(c) = z_1,\ y'(c) = z_2. \tag{6}$$

Proof. Suppose first that $y \in C^2([a,b])$ satisfies (6), that is,

$$y'' = (\lambda - q)y\,,\ y(c) = z_1 \ \text{ and } \ y'(c) = z_2\,. \tag{7}$$

For every $u \in [a,b]$, we get

$$y'(u) - z_2 = y'(u) - y'(c) = \int_c^u y''(t)\,dt = \int_c^u (\lambda - q(t))\,y(t)\,dt.$$

This gives that

$$\begin{aligned} y(x) - z_1 - z_2(x-c) &= y(x) - y(c) - z_2(x-c) \\ &= \int_c^x (y'(u) - z_2)\,du \\ &= \int_c^x \int_c^u (\lambda - q(t))\,y(t)\,dt\,du \\ &= \int_c^x \int_t^x (\lambda - q(t))\,y(t)\,du\,dt \\ &= \int_c^x (x-t)(\lambda - q(t))\,y(t)\,dt \end{aligned}$$

for all $x \in [a,b]$, hence that y satisfies that the integral equation

$$y(x) = z_1 + z_2(x-c) + \int_c^x (x-t)(\lambda - q(t))\,y(t)\,dt \quad \text{for all } x \in [a,b]. \tag{8}$$

Conversely, if $y \in C([a,b])$ satisfies (8), then it is an easy exercise to check that y belongs to $C^2([a,b])$ and satisfies (7).

Now, let $T : C([a,b]) \to C([a,b])$ be the integral operator defined for each f in $C([a,b])$ by

$$[T(f)](x) = z_1 + z_2(x-c) + \int_c^x (x-t)(\lambda - q(t))\,f(t)\,dt$$

for all $x \in [a,b]$. We consider here $C([a,b])$ as a complete metric space w.r.t. the metric $d(f,g) := \|f-g\|_\infty = \sup\{|f(x) - g(x)| : x \in [a,b]\}$.

Set $K := \sup\big\{|(x-t)(\lambda - q(t))| : x,t \in [a,b]\big\} < \infty$. Let $f, g \in C([a,b])$ and $x \in [a,b]$. By induction on $n \in \mathbb{N}$, one easily shows that

$$\Big|[T^n(f) - T^n(g)](x)\Big| \le \frac{1}{n!}\,K^n\,|x-c|^n\,\|f-g\|_\infty\,.$$

This implies that

$$\|T^n(f) - T^n(g)\|_\infty \le \frac{K^n\,(b-a)^n}{n!}\,\|f-g\|_\infty\,.$$

It clearly follows that T^n is a contraction when n is so large that $\frac{K^n\,(b-a)^n}{n!} < 1$. Hence, Banach's fixed point theorem [2] gives that T has a unique fixed point, say y, in $C([a,b])$. This means that y is the unique function in $C([a,b])$ such that $y = T(y)$, i.e., such that y satisfies (8). Taking into account what we proved in the first part of the proof, we are done. $\square$

Corollary 2.2. *Let $\lambda \in \mathbb{C}$, and recall that $S_\lambda = \{y \in C^2([a,b]) : -y'' + qy = \lambda y\}$. Let $c \in [a,b]$. Then the map $T_{\lambda,c} : S_\lambda \to \mathbb{C}^2$, defined by*

$$T_{\lambda,c}(y) = (y(c), y'(c)) \quad \textit{for every } y \in S_\lambda\,,$$

is an isomorphism. Hence, $\dim S_\lambda = 2$.

Proof. Theorem 2.1 shows that the map $T_{\lambda,c}$ is 1-1 and onto. It is obvious that it is linear. $\square$

Remark 2.3. It should be noted that Theorem 2.1 is essentially an existence result (although our method of proof gives a way to approximate the unique solution of (6) by picking some $y_0 \in C([a,b])$ and computing $T^n(y_0)$ for large enough n). Explicit formulas for a basis of S_λ are only known when q is a constant function. To illustrate Corollary 2.2, we recall these. Assume $q(x) = \omega$ for all $x \in [a,b]$ for some $\omega \in \mathbb{C}$. Then $-y''+qy = \lambda y$ can be rewritten as the homogeneous equation $y''+(\lambda-\omega)y = 0$, which we know can be solved by considering the characteristic equation $z^2 + (\lambda - \omega) = 0$:

If $\lambda \neq \omega$, then, letting $(\omega - \lambda)^{1/2}$ denote a square root of $\omega - \lambda$ in $\mathbb{C}$, we get that S_λ consists of the functions of the form

$$y(x) = C_1\, e^{(\omega-\lambda)^{1/2}x} + C_2\, e^{-(\omega-\lambda)^{1/2}x}, \quad x \in [a,b],$$

where $C_1, C_2 \in \mathbb{C}$. Thus $\{e^{(\omega-\lambda)^{1/2}x},\, e^{-(\omega-\lambda)^{1/2}x}\}$ is a basis for S_λ in this case.

If $\lambda = \omega$, then the equation is $y'' = 0$, and $\{1, x\}$ is obviously a basis for S_ω.

Remark 2.4. Suppose that q is real-valued, $\lambda \in \mathbb{R}$ and $y \in S_\lambda$. Then it is not difficult to verify that $\overline{y} \in S_\lambda$, so that $\operatorname{Re} y$ and $\operatorname{Im} y$ also lie in S_λ. Moreover, if it happens that $y(c)$ and $y'(c)$ both are real numbers for some $c \in [a,b]$, then the function y has to be real-valued: indeed, we then have $(\operatorname{Im} y)(c) = 0 = (\operatorname{Im} y)'(c)$, so Theorem 2.1 implies that $\operatorname{Im} y$ is the zero function on $[a,b]$.

Remark 2.5. Consider $y_1, y_2 \in C^2([a,b])$. Define $W_{y_1,y_2} \in C^1([a,b])$ by

$$W_{y_1,y_2}(x) = \begin{vmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{vmatrix} \quad \text{for each } x \in [a,b].$$

$W_{y_1,y_2}(x)$ is called the *Wronsky determinant* of (y_1, y_2) at x.

Assume that $y_1, y_2 \in S_\lambda$ and let $c \in [a,b]$. Corollary 2.2 implies that the set $\{y_1, y_2\}$ is a basis for S_λ if and only if the vectors $(y_1(c), y_1'(c)), (y_2(c), y_2'(c))$ are linearly independent in $\mathbb{C}^2$, i.e., $W_{y_1,y_2}(c) \neq 0$. Note that this gives that if $\{y_1, y_2\}$ is a basis for S_λ, then $W_{y_1,y_2}(x) \neq 0$ for all $x \in [a,b]$.

The Wronsky determinant appears in the following lemma (sometimes called *Lagrange's lemma*), which will be useful to us later:

[2] cf. Lindstrøm's book *Spaces*, Exercise 3.4.7.

Lemma 2.6. *Define* $\widetilde{D} : C^2([a,b]) \to C([a,b])$ *by*

$$\widetilde{D}(y) = -y'' + qy \quad \textit{for } y \in C^2([a,b]),$$

and let $f, g \in C^2([a,b])$*. Then the following identity holds:*

$$\widetilde{D}(f)g - \widetilde{D}(g)f = (fg' - gf')' = (W_{f,g})'$$

Proof. We have

$$\begin{aligned}(W_{f,g})' &= (fg' - gf')' = fg'' + f'g' - gf'' - g'f' = fg'' - gf'' \\ &= -f''g + qfg - qgf + fg'' = \widetilde{D}(f)g - \widetilde{D}(g)f.\end{aligned}$$

□

3. Some spectral properties of D

Let (α_1, α_2), $(\beta_1, \beta_2) \in \mathbb{R}^2 \setminus \{(0,0)\}$ and $q \in C([a,b])$ be real-valued. We recall that

$$Y = \{y \in C^2([a,b]) \mid y \text{ satisfies the boundary conditions (9)}\},$$

where

$$\alpha_1\, y(a) + \alpha_2\, y'(a) = 0\,, \quad \beta_1\, y(b) + \beta_2\, y'(b) = 0\,, \tag{9}$$

and $D : Y \to C[a,b]$ is the linear operator defined by $D(y) = -y'' + qy$ for $y \in Y$.

Let $\lambda \in \mathbb{C}$ and set $E_\lambda := \{y \in Y : D(y) = \lambda y\}$. We say that λ is an eigenvalue of D if the subspace E_λ is non-trivial, in which case E_λ is called the eigenspace of D associated to λ. We note that $E_\lambda \subset S_\lambda$, so Corollary 2.2 implies that $\dim E_\lambda \leq 2$.

Proposition 3.1. *Let* $f, g \in Y$*. Then we have*

i) $D(f)g - D(g)f = (fg' - gf')'$,

ii) $\langle D(f), g\rangle = \langle f, D(g)\rangle$.

Proof. i) Since $D = \widetilde{D}_{|Y}$, this identity follows from Lagrange's lemma (Lemma 2.6).
ii) It is easy to check that $\overline{g} \in Y$ and $\overline{D(g)} = D(\overline{g})$. Thus, using i), we get

$$\begin{aligned}\langle D(f), g\rangle - \langle f, D(g)\rangle &= \int_a^b [D(f)\overline{g} - fD(\overline{g})](t)\, dt \\ &= \int_a^b (f\overline{g}' - \overline{g}f')'(t)\, dt = \Big[(f\overline{g}' - \overline{g}f')(t)\Big]_a^b \\ &= f(b)\overline{g'(b)} - \overline{g(b)}f'(b) - f(a)\overline{g'(a)} + \overline{g(a)}f'(a).\end{aligned}$$

Now, since f and $\overline{g}$ both satisfy 9, we have

$$\begin{bmatrix} f(b) & f'(b) \\ \overline{g}(b) & \overline{g}'(b) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Since $(\beta_1, \beta_2) \neq (0,0)$, this implies that $f(b)\overline{g'(b)} - \overline{g(b)}f'(b) = 0$.

Arguing in a similar way, one can also show that $f(a)\overline{g'(a)} - \overline{g(a)}f'(a) = 0$. Inserting these two equalities in our computation above, we get that

$$\langle D(f), g\rangle - \langle f, D(g)\rangle = 0,$$

as desired. □

Part ii) of Proposition 3.1 shows that the operator D enjoys a property similar to self-adjointness. Proceeding exactly as we did for bounded self-adjoint operators on Hilbert spaces, one deduces that the following result holds.

Corollary 3.2. *All the possible eigenvalues of D are real, and the associated eigenspaces are orthogonal to each other.*

Note that we don't know yet whether D has eigenvalues. Anyhow, we can say more about its eigenspaces (if any).

Proposition 3.3. *All possible eigenspaces of D are one-dimensional.*

Proof. Let $\lambda \in \mathbb{C}$. Recall that $\widetilde{D} : C^2([a,b]) \to C([a,b])$ is defined by $\widetilde{D}(y) = -y'' + qy$ for $y \in C^2([a,b])$, so $D = \widetilde{D}_{|Y}$. We first consider the space

$$L_\lambda := \{y \in C^2([a,b]) : \widetilde{D}(y) = \lambda y \text{ and } y \text{ satisfies (10)}\},$$

where

$$\alpha_1\, y(a) + \alpha_2\, y'(a) = 0. \tag{10}$$

Condition (10) says that the vector $(y(a), y'(a))$ belongs to $M := \mathrm{Span}\{(-\alpha_2, \alpha_1)\}$. Now, Corollary 2.2 (with $c = a$) gives that $L_\lambda = T_{\lambda,a}^{-1}(M)$. Since $\dim M = 1$ and $T_{\lambda,a}$ is an isomorphism, we get that L_λ is a one-dimensional subspace of S_λ.

Similarly, one shows that $R_\lambda := \{y \in C^2([a,b]) : \ \widetilde{D}(y) = \lambda y$ and y satisfies (11)$\}$, where

$$\beta_1\, y(b) + \beta_2\, y'(b) = 0, \tag{11}$$

is also a one-dimensional subspace of S_λ.

Clearly, we have $E_\lambda = L_\lambda \cap R_\lambda$. So there are only two possibilities: either $E_\lambda = \{0\}$ or $\dim E_\lambda = 1$. Hence, if λ is an eigenvalue of D, we must have $\dim E_\lambda = 1$.

∎

Remark 3.4. We use the notation introduced in the proof above. Assume that $\lambda \in \mathbb{C}$ is *not* an eigenvalue of D, and pick $u \in L_\lambda$, $u \neq 0$, $v \in R_\lambda$, $v \neq 0$. It is clear from this proof that $L_\lambda = \mathrm{Span}\{u\}$ and $R_\lambda = \mathrm{Span}\{v\}$. As $L_\lambda \cap R_\lambda = E_\lambda = \{0\}$, the vectors u and v must be linearly independent. Since they both lie in S_λ, which is 2-dimensional by Corollary 2.2, we can then conclude that $\{u, v\}$ is a basis for S_λ.

4. Exploiting D^{-1} when D is 1-1

We go back to the Sturm-Liouville problem for the equation $D(y) = \lambda y$, $y \in Y$. As long as we are not able to show that D has eigenvalues, it is not possible for us to make efficient use of its spectral properties. Ideally, we would like to show that D is diagonalizable, in the sense that there exists a sequence of eigenfunctions of D in Y which forms an orthonormal basis for $L^2([a,b])$. The trick to make progress on this problem is to turn our attention to the inverse of D, whenever this makes sense.

We therefore assume throughout this section that $D : Y \to C([a,b])$ is 1-1. We will see how to get rid of this assumption in the next section.

We will first show that $D(Y) = C([a,b])$, i.e., D is onto, and that the inverse operator

$$D^{-1} : C([a,b]) \to Y$$

is an integral operator associated to a continuous kernel $G : [a,b] \times [a,b] \to \mathbb{C}$.

It should be noted here that given $f \in C([a,b])$, the standard way to show that the differential equation $-y'' + qy = f$ has a solution is to pick a basis for the associated homogeneous equation and use the method called variation of parameters. We will not discuss this method here and follow a shorter path.

Since the operator D is linear, the fact that D is 1-1 means that its kernel is trivial, that is, 0 is *not* an eigenvalue of D. As we saw in Remark 3.4, we can then pick a basis $\{u, v\}$ for $S_0 = \{y \in C^2([a,b]) : \widetilde{D}(y) = 0\} = \{y \in C^2([a,b]) : y'' = qy\}$ such that

- u satisfies the condition $\alpha_1\, u(a) + \alpha_2\, u\,'(a) = 0$,
- v satisfies the condition $\beta_1\, v(b) + \beta_2\, v\,'(b) = 0$.

Since q is real-valued, we can also assume that u and v are real-valued, cf. Remark 2.4.

We note that Remark 2.5 tells us that $W_{u,v}(x) \neq 0$ for all $x \in [a,b]$. Moreover, as $\widetilde{D}(u) = \widetilde{D}(v) = 0$, Lemma 2.6 gives that

$$(W_{u,v})' = \widetilde{D}(u)v - \widetilde{D}(v) = 0.$$

Hence, $W_{u,v}$ is a constant function on $[a,b]$. This means that

$$W_{u,v}(x) = u(x)v'(x) - v(x)u'(x) = W \qquad \text{for all } x \in [a,b]$$

for some $W \in \mathbb{R} \setminus \{0\}$.

We can now define the associated *Green's function* $G : [a,b] \times [a,b] \to \mathbb{R}$ by

$$G(x,t) = -\frac{1}{W} \cdot \begin{cases} u(x)v(t) & \text{if } a \leq x \leq t \leq b, \\ \\ u(t)v(x) & \text{if } a \leq t \leq x \leq b. \end{cases}$$

It is then straightforward to see that G is continuous. Hence we may form the associated integral operator $T_G : L^2([a,b]) \to L^2([a,b])$, which is given by

$$[T_G(f)](x) = \int_a^b G(x,t)\, f(t)\, dt$$

for all $f \in L^2([a,b])$ and $x \in [a,b]$. It is clear that T_G maps $C([a,b])$ into itself. In fact, it maps $C([a,b])$ into Y:

Proposition 4.1. *Let $f \in C([a,b])$ and set $y := T_G(f)$. Then $y \in Y$ and $D(y) = f$.*

Proof. Let $x \in [a,b]$. Using the definitions of G and T_G we get

$$y(x) = -\int_a^x W^{-1}\, v(x)u(t)f(t)\, dt - \int_x^b W^{-1}\, u(x)v(t)f(t)\, dt.$$

This implies that $-Wy(x) = v(x)A(x) + u(x)B(x)$, where

$$A(x) := \int_a^x u(t)f(t)\,dt \quad \text{and} \quad B(x) := \int_x^b v(t)f(t)\,dt.$$

Thus we get

$$\begin{aligned}-Wy'(x) &= v'(x)A(x) + v(x)A'(x) + u'(x)B(x) + u(x)B'(x)\\ &= v'(x)A(x) + v(x)u(x)f(x) + u'(x)B(x) - u(x)v(x)f(x)\\ &= v'(x)A(x) + u'(x)B(x).\end{aligned}$$

Since v', A, u' and B all lie in $C^1([a,b])$, we see that $y \in C^2([a,b])$.

Further, using that $A(a) = 0$ and $\alpha_1 u(a) + \alpha_2 u'(a) = 0$, we get

$$\begin{aligned}\alpha_1 y(a) + \alpha_2 y'(a) &= \frac{-1}{W}(\alpha_1 v(a)A(a) + \alpha_1 u(a)B(a) + \alpha_2 v'(a)A(a) + \alpha_2 u'(a)B(a))\\ &= \frac{-1}{W}(\alpha_1 u(a) + \alpha_2 u'(a))B(a) = 0.\end{aligned}$$

In a similar way we get $\beta_1\, v(b) + \beta_2\, v\,'(b) = 0$. Thus we have shown that $y \in Y$.

To verify the second assertion, we first compute $-Wy''$. Since $u'' = qu$, $v'' = qv$, and $v'u - u'v = W$ on $[a,b]$, we get

$$\begin{aligned}-Wy'' &= (v'A + u'B)' = v''A + v'A' + u''B + u'B'\\ &= q(vA + uB) + (v'u - u'v)f\\ &= -qWy + (v'u - u'v)f\\ &= W(f - qy).\end{aligned}$$

Thus, $-y'' = f - qy$, which gives

$$D(y) = -y'' + qy = f - qy + qy = f\,,$$

as desired. $\square$

The first part of Proposition 4.1 shows that D is onto $C([a,b])$. Since D is also 1-1 (by assumption), D has an inverse map $D^{-1} : C([a,b]) \to Y$, which is defined as follows:

Given some $f \in C([a,b])$, then

$$D^{-1}(f) := y\,,$$

where $y \in Y$ is the unique function in Y such that $D(y) = f$.

We now see that the second part of Proposition 4.1 tells us that

$$D^{-1}(f) = T_G(f) \quad \text{for every } f \in C([a,b]),$$

i.e., $D^{-1} = (T_G)_{|C([a,b])}$.

Since we have $G(x,t) = G(t,x)$ for all (x,t) in $[a,b] \times [a,b]$ (check!), we get that T_G is self-adjoint. As T_G is also compact (indeed, it is a Hilbert-Schmidt operator on $L^2([a,b])$, cf. ELA, Example 5.2.11), we are in the position to apply the spectral theorem to T_G. However, we will also need to know that T_G maps $L^2([a,b])$ into $C([a,b])$. This is true for any integral operator with continuous kernel:

Lemma 4.2. *Assume $K : [a,b] \times [a,b] \to \mathbb{C}$ is continuous. Then the associated integral operator $T_K : L^2([a,b]) \to L^2([a,b])$ maps $L^2([a,b])$ into $C([a,b])$.*

Proof. Let $f \in L^2([a,b])$ and let $\varepsilon > 0$. Note that the Cauchy-Schwarz inequality gives that

$$M := \int_{[a,b]} |f| \, dm \leq \Big(\int_{[a,b]} 1 \, dm \Big)^{1/2} \Big(\int_{[a,b]} |f|^2 \, dm \Big)^{1/2} = \sqrt{b-a} \, \|f\|_2 < \infty,$$

where m denotes the Lebesgue measure on $[a,b]$. As K is automatically uniformly continuous on the compact set $R := [a,b] \times [a,b]$, we can find $\delta > 0$ such that

$$\big|K(x_1,t_1) - K(x_2,t_2)\big| < \varepsilon/M$$

whenever $(x_1,t_1),(x_2,t_2) \in R$ and $|x_2 - x_1| < \delta$, $|t_2 - t_1| < \delta$.

Let now $x_0 \in [a,b]$. Then for every $t \in [a,b]$ and all $x \in [a,b]$ such that $|x - x_0| < \delta$, we have

$$\big|K(x,t) - K(x_0,t)\big| < \varepsilon/M\,.$$

Thus we get

$$\begin{aligned}
\Big|[T_K(f)](x) - [T_K(f)](x_0)\Big| &= \Big| \int_{[a,b]} \big(K(x,t) - K(x_0,t)\big) \, f(t) \, dm(t) \Big| \\
&\leq \int_{[a,b]} |K(x,t) - K(x_0,t)| \, |f(t)| \, dm(t) \\
&\leq \varepsilon/M \int_{[a,b]} |f| \, dm = \varepsilon
\end{aligned}$$

for all $x \in [a,b]$ such that $|x - x_0| < \delta$. This shows that $T_K(f)$ is continuous at x_0. Since x_0 was an arbitrary point of $[a,b]$, $T_K(f) \in C([a,b])$.

□

Theorem 4.3. *Assume that D is 1-1, and consider the Sturm-Liouville problem*

$$D(y) = \lambda y \qquad \text{with } y \in Y.$$

Then the following assertions hold:

- *The eigenvalues for this problem form a countable set $\{\lambda_k : k \in \mathbb{N}\}$ of non-zero distinct real numbers satisfying that $|\lambda_k| \to \infty$ as $k \to \infty$.*
- *For each $k \in \mathbb{N}$ the eigenspace $E_{\lambda_k} = \{y \in Y : D(y) = \lambda_k \, y\}$ is one-dimensional.*
- *If y_k is a unit vector in E_{λ_k} for each $k \in \mathbb{N}$, then $\{y_k : k \in \mathbb{N}\}$ is an orthonormal basis for $L^2([a,b])$.*

Proof. We first observe that 0 is not an eigenvalue of T_G :

Indeed, since Y is dense in $H := L^2([a,b])$, and $Y = T_G(C([a,b])$, we have

$$H = \overline{Y} \subset \overline{T_G(H)} \subset H,$$

hence $\overline{T_G(H)} = H$. Thus, we get $\ker(T_G)^\perp = \overline{T_G(H)} = H$, i.e., $\ker(T_G) = \{0\}$.

Applying the spectral theorem to T_G, we obtain that the eigenvalues of T_G form a countable set $\{\mu_k : k \in \mathbb{N}\}$ of non-zero distinct real numbers satisfying that $\mu_k \to 0$ as $k \to \infty$.

Let $k \in \mathbb{N}$, and set $\lambda_k = \mu_k^{-1} \neq 0$. Let $f_k \in H$ be an eigenfunction for T_G associated to μ_k. Since $T_G(f_k) = \mu_k\, f_k$, we get

$$f_k = \lambda_k\, T_G(f_k)\,. \tag{12}$$

As Lemma 4.2 gives that $T_G(f_k) \in C([a, b])$, this gives that $f_k \in C([a, b])$. Hence, $T_G(T_G(f_k)) \in Y$. But (12) implies that

$$f_k = \lambda_k^2\; T_G(T_G(f_k))\,,$$

so we get that $f_k \in Y$. Now, applying D to (12), we get

$$D(f_k) = \lambda_k\, f_k\,.$$

This shows that λ_k is an eigenvalue of D, and f_k is an eigenfunction for D associated to λ_k. Now, Proposition 3.3 tells us that $E_{\lambda_k} := \{y \in Y : D(y) = \lambda_k\, y\}$ is one-dimensional. Hence, we have $E_{\lambda_k} = \mathrm{Span}\,\{f_k\}$. Further, one readily checks that E_{λ_k} is also the eigenspace of T_G associated to μ_k.

We note that D can not have other eigenvalues than the λ_k's (for if D had one such eigenvalue, then T_G would have an eigenvalue different from all μ_k's, which is not the case). Further, we note that

$$\lim_{k\to\infty} |\lambda_k| = \lim_{k\to\infty} |\mu_k|^{-1} = \infty$$

since $\lim_{k\to\infty} \mu_k = 0$.

Finally, if we set $y_k := \pm(\|f_k\|_2)^{-1}\, f_k \in E_{\lambda_k}$ for each $k \in \mathbb{N}$, then we also get from the spectral theorem for T_G that $\{y_k : k \in \mathbb{N}\}$ is an orthonormal basis for H consisting of eigenfunctions for D.

□

Example 4.4. To illustrate this theorem, let us consider the Sturm-Liouville system

$$-y'' = \lambda y \text{ on } [0, \pi], \text{ with } Y = \{y \in C^2([0, \pi]) : y(0) = y(\pi) = 0\}.$$

In other words, we consider $D(y) = -y''$ with $y \in Y$.

Let $\lambda \in \mathbb{C}$. We consider first the case $\lambda = 0$. It is straightforward to check that the only function y in Y satisfying $-y'' = 0$ is $y = 0$. Thus, 0 is not an eigenvalue of D, i.e., D is 1-1, so Theorem 4.3 applies in this case. We can determine the eigenvalues and the eigenfunctions of D explicitly as follows.

Assume $\lambda \neq 0$, and write $\lambda^{1/2} = r + i\, s$ with $(r, s) \in \mathbb{R}^2 \setminus \{(0, 0)\}$.[3]
A basis for $S_\lambda := \{y \in C^2([0, \pi]) : -y'' = \lambda y\}$ is given by

$$\{e^{i\lambda^{1/2}x},\, e^{-i\lambda^{1/2}x}\} = \{e^{-sx}(\cos(rx) + i \sin(rx)),\;\; e^{sx}(\cos(rx) - i \sin(rx))\}.$$

If $y \in S_\lambda$, say $y(x) = C_1\, e^{-sx}(\cos(rx) + i\,(sx)) + C_2\, e^{sx}(\cos(rx) - i\, \sin(rx))$,
then $y \in Y$ if and only if

$$\begin{cases} C_1 + C_2 = 0, \\ C_1\, e^{-s\pi}(\cos(r\pi) + i\, \sin(r\pi)) + C_2\, e^{s\pi}(\cos(r\pi) - i\, \sin(r\pi)) = 0. \end{cases}$$

[3]We could here have used that we know that all the possible eigenvalues of D are real, so that we need only to consider $\lambda \in \mathbb{R} \setminus \{0\}$. However this would not shorten our discussion significantly.

This gives that $E_\lambda = S_\lambda \cap Y$ is non-trivial if and only if

$$e^{-s\pi}(\cos(r\pi) + i\,\sin(r\pi)) = e^{s\pi}(\cos(r\pi) - i\,\sin(r\pi)),$$

and it is elementary to deduce that this happens if and only if $s = 0$ and $r = k$ for some $k \in \mathbb{Z} \setminus \{0\}$, in which case $\lambda = k^2 \in \mathbb{N}$ and $E_\lambda = \mathrm{span}\{\sin(kx)\}$.

This means that the distinct eigenvalues of this Sturm-Liouville system are $\lambda_k = k^2$, $k \in \mathbb{N}$, with associated normalized eigenfunctions $y_k(x) = \sqrt{2/\pi}\sin(kx)$. This is in accordance with Theorem 4.3. Note that this theorem implies that the set $\{\sqrt{2/\pi}\,\sin(kx) : k \in \mathbb{N}\}$ is an orthonormal basis for $L^2([0,\pi])$.

For completeness, we also compute the Green's function G and $T_G(f)$ for $f \in C([0,\pi])$. One computes easily that $L_0 = \{y \in C^2([0,\pi]) : y'' = 0,\, y(0) = 0\} = \mathrm{span}\{u\}$, where $u(x) = x$, while $R_0 = \{y \in C^2([0,\pi]) : y'' = 0,\, y(\pi) = 0\} = \mathrm{span}\{v\}$, where $v(x) = x - \pi$. Thus we get that

$$W = W_{u,v}(x) = u(x)v'(x) - v(x)u'(x) = x - (x - \pi) = \pi$$

for all $x \in [0,\pi]$. Moreover, the Green's function $G : [0,\pi] \times [0,\pi] \to \mathbb{C}$ is given by

$$G(x,t) = \frac{1}{\pi} \cdot \begin{cases} x\,(\pi - t) & \text{if } 0 \le x \le t \le \pi, \\ t\,(\pi - x) & \text{if } 0 \le t \le x \le \pi, \end{cases}$$

and we obtain that

$$[T_G(f)](x) = \int_0^\pi G(x,t)\,f(t)\,dt = \frac{1}{\pi}\Big((\pi - x)\int_0^x t f(t)\,dt + x\int_x^\pi (\pi - t)f(t)\,dt\Big)$$

for all $f \in C([0,\pi])$ and $x \in [0,\pi]$. Note that determining the eigenvalues of T_G by direct computation is not an easy task. Anyhow, you should verify that $\sin(kx)$ is an eigenfunction for T_G associated with the eigenvalue $\mu_k = k^{-2}$ for each $k \in \mathbb{N}$.

5. The general case

In this final section, we consider the general case, i.e., we don't assume that D is 1-1. The idea now is to show that there exists some $\mu \in \mathbb{R}$ which is not an eigenvalue of D, and consider the operator $D_\mu : Y \to C([a,b])$ defined by $D_\mu(y) = -y'' + qy - \mu y$. Then 0 will not be an eigenvalue of D_μ (otherwise there would be some $y \in Y \setminus \{0\}$ such that $D_\mu(y) = 0$, i.e., $D(y) = \mu y$, and μ would be an eigenvalue of D, giving a contradiction). Hence, we will be able to apply Theorem 4.3 to the Sturm-Liouville system $D_\mu(y) = \lambda' y$ with $y \in Y$, and deduce some interesting consequences for our original Sturm-Liouville problem.

Since we know that the possible eigenvalues of D are all real numbers, one may think: why not just pick some $\mu \in \mathbb{C} \setminus \mathbb{R}$? The problem with such a choice is that the function $q_\mu(x) := q(x) - \mu$ for $x \in [a,b]$ will not be real-valued, hence that the Sturm-Liouville system associated with D_μ will not match our requirements.

Lemma 5.1. *There exists some $\mu \in \mathbb{R}$ which is not an eigenvalue of D.*

Proof. We know that $L^2([a,b])$ has a countable orthonormal basis, say $\{u_k\}_{k\in\mathbb{N}}$. (One may for example take $u_k(x) = \sqrt{2/(b-a)}\sin\left(k\pi(x-a)/(b-a)\right)$ for $x \in [a,b]$ and $k \in \mathbb{N}$, cf. Example 4.4 when $[a,b] = [0,\pi]$). We will show that this implies that D

has a countable number of distinct eigenvalues. Since $\mathbb{R}$ is uncountable, the assertion to be proven will clearly follow.

Assume (for contradiction) that D has an uncountable number of distinct eigenvalues. Then we can pick a unit vector in each of the associated eigenspaces. As all eigenspaces of D are orthogonal to each other, this means that there exists an orthonormal subset Γ of Y which is uncountable.

Let $k \in \mathbb{N}$. Then it follows from Bessel's inequality that

$$M_k := \sup_{A \subset \Gamma,\, A \,\text{finite}} \Big\{ \sum_{\gamma \in A} |\langle u_k, \gamma\rangle|^2 \Big\} \leq \|u_k\|_2^2 = 1 \,<\, \infty.$$

This implies that the set

$$U_{k,n} := \Big\{ \gamma \in \Gamma : |\langle u_k, \gamma\rangle| \,\geq\, \frac{1}{n} \Big\}$$

is finite for every $n \in \mathbb{N}$: indeed, if $U_{k,n}$ was infinite for some n, then we could find an infinite sequence $\{\gamma_m\}_{m\in\mathbb{N}}$ of distinct elements in $U_{k,n}$, and this would give that

$$M_k \,\geq\, \sup_{m\in\mathbb{N}} \Big\{ \sum_{\ell=1}^{m} |\langle u_k, \gamma_\ell\rangle|^2 \Big\} \,\geq\, \sup_{m\in\mathbb{N}} \Big\{ m \cdot \frac{1}{n^2} \Big\} = \infty\,,$$

contradicting that $M_k < \infty$. Setting now $U_k := \{\gamma \in \Gamma : \langle u_k, \gamma\rangle \neq 0\}$, we get that

$$U_k = \bigcup_{n\in\mathbb{N}} U_{k,n}$$

is countable (being a countable union of finite sets). Hence, the countable union $U := \bigcup_{k\in\mathbb{N}} U_k$ is a countable subset of Γ. As Γ is uncountable, there must exist some $\gamma \in \Gamma \setminus U$. But then $\gamma \notin U_k$ for every $k \in \mathbb{N}$, so we have

$$\langle u_k, \gamma\rangle = 0 \quad \text{for every } k \in \mathbb{N}.$$

This says that γ is orthogonal to every u_k, so we must have $\gamma = 0$ (since $\{u_k\}_{k\in\mathbb{N}}$ is an orthonormal basis for $L^2([a,b])$). But this gives a contradiction, since every element of Γ is a unit vector. We can therefore conclude that D has a countable number of distinct eigenvalues. $\square$

We can now state our main result about regular Sturm-Liouville systems:

Theorem 5.2. *Consider the Sturm-Liouville problem*

$$D(y) = \lambda y \quad \textit{with } y \in Y.$$

Then the following assertions hold:

- *The eigenvalues for this problem form a countable set* $\{\lambda_k : k \in \mathbb{N}\}$ *of distinct real numbers satisfying that* $|\lambda_k| \to \infty$ *as* $k \to \infty$.
- *For each* $k \in \mathbb{N}$ *the eigenspace* $E_{\lambda_k} = \{y \in Y : D(y) = \lambda_k\, y\}$ *is one-dimensional.*
- *If* y_k *is a unit vector in* E_{λ_k} *for each* $k \in \mathbb{N}$*, then* $\{y_k : k \in \mathbb{N}\}$ *is an orthonormal basis for* $L^2([a,b])$.

Proof. By Lemma 5.1 we can find some $\mu \in \mathbb{R}$ which is not an eigenvalue of D. We define $D_\mu : Y \to C([a,b])$ by $D_\mu(y) = -y'' + qy - \mu y$ and consider the Sturm-Liouville system $D_\mu(y) = \lambda' y$ on Y. Then, as 0 is not an eigenvalue of D_μ, we may apply Theorem 4.3 to D_μ. This gives:

- The eigenvalues of D_μ form a countable set $\{\lambda'_k : k \in \mathbb{N}\}$ of non-zero distinct real numbers satisfying that $|\lambda'_k| \to \infty$ as $k \to \infty$.
- For each $k \in \mathbb{N}$ the eigenspace $E'_{\lambda'_k} = \{y \in Y : D_\mu(y) = \lambda'_k\, y\}$ is one-dimensional.
- If v_k is a unit vector in $E'_{\lambda'_k}$ for each $k \in \mathbb{N}$, then $\{v_k : k \in \mathbb{N}\}$ is an orthonormal basis for $L^2([a,b])$.

Now, for $y \in Y$, we obviously have $D(y) = \lambda y$ if and only if $D_\mu(y) = (\lambda - \mu)y$. This implies that the set consisting of all eigenvalues of D is the countable set of distinct real numbers given by $\{\lambda_k : k \in \mathbb{N}\}$, where $\lambda_k := \lambda'_k + \mu$ for each $k \in \mathbb{N}$. Moreover, the eigenspace E_{λ_k} of D associated to each λ_k is then equal to $E'_{\lambda'_k}$, hence is one-dimensional. Finally, if y_k is a unit vector in E_{λ_k} for each $k \in \mathbb{N}$, then we have $y_k = \pm\, v_k$ for every $k \in \mathbb{N}$, so the last assertion clearly follows. □

6. Exercises

Exercise 1

Find the eigenvalues and eigenfunctions of the Sturm-Liouville system $-y'' = \lambda y$ on the given interval with the following boundary conditions:

i) $[a,b] = [0,\pi]$, $y'(0) = 0$, $y'(\pi) = 0$.

ii) $[a,b] = [0,\pi]$, $y'(0) = 0$, $y(\pi) = 0$.

iii) $[a,b] = [0,2\pi]$, $y(0) = 0$, $y(2\pi) = 0$.

iv) $[a,b] = [0,1]$, $y(0) = 0$, $y(1) + y'(1) = 0$.

Exercise 2

Consider a Sturm-Liouville system $D(y) = \lambda\, y$ on Y as in (4), with boundary conditions as in (2).

i) Assume that the following extra conditions holds:

a) $q(x) \geq 0$ for all $x \in [a,b]$,

b) $\alpha_1\alpha_2 \leq 0$ and $\beta_1\beta_2 \geq 0$.

Show that the eigenvalues of D are all non-negative.

Exercise 3

Consider a Sturm-Liouville system $D(y) = \lambda\, y$ as in (4), but where

$$Y = \{C^2([a,b]) : y(a) \text{ or } y'(a) = 0\,;\, y(b) \text{ or } y'(b) = 0\}.$$

Show that the distinct eigenvalues of D may ordered so that

$$\lambda_1 < \lambda_2 < \lambda_3 < \cdots \text{ and } \lim_{k\to\infty} \lambda_k = \infty\,.$$

Exercise 4

Assume D is 1-1, as in section 4. Show that the distinct eigenvalues of D satisfy

$$\sum_{k=1}^{\infty} \frac{1}{|\lambda_k|^2} < \infty .$$

Exercise 5

Show that the space $Y = \{y \in C^2([a,b]) : y \text{ satisfies the conditions (2)}\}$ is dense in $L^2([a,b])$.