7 The Sturm-Liouville problem.

In this section we shall discuss the differential operator

$$Ly = \frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) + q(x)y$$

acting on functions y defined on a closed bounded interval $[a, b]$. We shall assume that $p(x) > 0$ and $q(x)$ real for $a \le x \le b$.

We make further assumptions that may be summarized, broadly speaking, by saying that "everything makes sense". Specifically we need L to act on functions that are twice differentiable and whose second derivatives are in $L^2[a, b]$. We also need to have that p is differentiable with p' continuous on $[a, b]$.

We shall be concerned with solving the problem

$$Ly = \frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) + q(x)y = f(x) \qquad (*)$$

subject to boundary conditions

$$\left.\begin{aligned} \alpha_1 y(a) + \alpha_2 y'(a) &= 0 \\ \beta_1 y(b) + \beta_2 y'(b) &= 0 \end{aligned}\right\} \qquad (\dagger)$$

Where $\alpha_1, \alpha_2, \beta_1$ and β_2 are real and $\alpha_1\alpha_2 \neq 0,\ \beta_1\beta_2 \neq 0$.

Note. The following calculation is of interest because it shows that L satisfies a symmetry condition that would, for a bounded operator, make it self-adjoint. However, it will not be used in the sequel. If L is restricted to act on the set of functions that satisfy the boundary conditions, then $\langle Ly, z\rangle = \langle y, Lz\rangle$. Indeed,

$$\begin{aligned} \langle -Ly, z\rangle + \langle y, Lz\rangle &= \int_a^b \left[\frac{d}{dt}\left(p(t)\frac{d\bar{z}(t)}{dt}\right)y - \frac{d}{dt}\left(p(t)\frac{dy}{dt}\right)\bar{z}\right] dt + \int_a^b (-qy\bar{z} + yq\bar{z})\, dt \\ &= \left[p(t)\frac{d\bar{z}}{dt}y(t) - p(t)\frac{dy}{dt}\bar{z}(t)\right]_a^b + \int_a^b p(t)\left(\frac{dy}{dt}\frac{d\bar{z}}{dt} - \frac{d\bar{z}}{dt}.\frac{dy}{dt}\right) dt \end{aligned}$$

and, as the integrals on the right of each line are 0, this will vanish if

$$p(b)\left[\bar{z}'(b)y(b) - \bar{z}(b)y'(b)\right] = p(a)\left[\bar{z}'(a)y(a) - \bar{z}(a)y'(a)\right].$$

But $\bar{z}'(a)y(a) - \bar{z}(a)y'(a)$ is the determinant of the 2×2 system

$$\begin{aligned} \xi\bar{z}(a) + \eta\bar{z}'(a) &= 0, \\ \xi y(a) + \eta y'(a) &= 0, \end{aligned}$$

which has the non-trivial solution $(\xi, \eta) = (\alpha_1, \alpha_2)$ when y and z satisfy the boundary conditions $(\dagger)$. So $\bar{z}'(a)y(a) - \bar{z}(a)y'(a) = 0$ and similarly $\bar{z}'(b)y(b) - \bar{z}(b)y'(b) = 0$. Therefore $\langle Ly, z\rangle = \langle y, Lz\rangle$.

We shall be looking for eigenvalues and eigenfunctions of L that satisfy the conditions $(\dagger)$; that is, for scalars λ and corresponding functions f that satisfy $(\dagger)$ and the equation $Lf = \lambda f$. We make the additional assumption that $\lambda = 0$ is not an eigenvalue of the system. This is quite a reasonable assumption, since if it fails then the problem (*), subject to $(\dagger)$ does not have a unique solution [an arbitrary multiple of the eigenfunction corresponding to $\lambda = 0$ could be added to any solution to obtain another solution].

Theorem 7.1 (Existence of the Green's function.) *Under the assumptions stated above, the problem (*), subject to (†) has the solution*

$$y(x) = \int_a^b k(x,t)\, f(t)\, dt$$

where $k(x,t)$ is real-valued and continuous on the square $[a,b] \times [a,b]$.

Proof. From the elementary theory of the initial value problem for linear differential equations, (also from Questions 1,2 and 3 of Exercises 6) we have that there is a unique function u such that $Lu = 0,\ u(a) = -\alpha_2, u'(a) = \alpha_1$. It follows easily that every solution of $Ly = 0,\ \alpha_1 y(a) + \alpha_2 y'(a) = 0$ is a scalar multiple of u. Similarly we have a unique v such that $Lv = 0, v(b) = -\beta_2, v'(b) = \beta_1$. The assumption that 0 is not an eigenvalue implies that u and v are linearly independent [if u were a multiple of v then it would be an eigenfunction].

Let

$$k(x,t) = \begin{cases} l\, u(x) & a \le x \le t \\ m\, v(x) & t \le x \le b \end{cases}$$

where (for fixed t) l, m are constants to be chosen. [Our motivational work suggests that we require $k(x,t)$ to be continuous and $p(x).\frac{\partial}{\partial x}k(x,t)$ to have a unit discontinuity at $x = t$.] Choose l, m such that

$$\begin{aligned} m.v(t) - l.u(t) &= 0, \\ p(t)[m.v'(t) - l.u'(t)] &= 1. \end{aligned}$$

Solving for l, m gives

$$l = \frac{v(t)}{\Delta} \qquad m = \frac{u(t)}{\Delta}$$

where $\Delta = p(t)(v'u - u'v) = pJ(u,v)$, where J is the Jacobean and hence non-zero [since u and v are independent]. Also,

$$\frac{d\Delta}{dt} = u(pv')' + u'(pv') - v(pu')' - v'(pu') = -quv + vqu = 0\,,$$

so Δ is a constant (i.e. also independent of t). [One can see, independently of the theory of Jacobeans, that $\Delta \neq 0$ since otherwise, at some point t_0

$$\begin{aligned} \xi.u(t_0) + \eta.v(t_0) &= 0, \\ \xi.u'(t_0) + \eta.v'(t_0) &= 0 \end{aligned}$$

has a non-trivial solution (ξ, η). Then $y = \xi.u + \eta.v$ is a solution of $Ly = 0, y(t_0) = y'(t_0) = 0$ and so $\xi.u + \eta.v$ is identically 0, contradicting the linear independence of u and v.] Hence we have that

$$k(x,t) = \begin{cases} \frac{u(x).v(t)}{\Delta} & a \le x \le t\,, \\ \frac{u(t).v(x)}{\Delta} & t \le x \le b\,. \end{cases}$$

To complete the proof, we just verify directly that

$$y(x) = \int_a^b k(x,t)\, f(t)\, dt$$

is the required solution. First note that when $x = a$ we have $x \le t$ throughout the range of integration and so

$$y(a) = \frac{1}{\Delta} \int_a^b u(a).v(t)\, f(t)\, dt$$

and, since $y'(x) = \displaystyle\int_a^b \frac{\partial k(x,t)}{\partial x}\, f(t)\, dt$,

$$y'(a) = \frac{1}{\Delta} \int_a^b u'(a).v(t)\, f(t)\, dt\,.$$

Therefore $\alpha_1 y(a) + \alpha_2 y'(a) = 0$ since u satisfies the boundary condition at $x = a$. Similarly $\beta_1 y(b) + \beta_2 y'(b) = 0$.

We now substitute into the equation. For notational convenience we substitute $\Delta\, y(x)$ (remembering that Δ is constant). Since $u(x), v(x)$ can be taken outside the integration, we obtain

$$\begin{aligned}
\Delta\, y(x) &= \Delta \int_a^b k(x,t)\, f(t)\, dt = v(x) \int_a^x u(t)\, f(t)\, dt + u(x) \int_x^b v(t)\, f(t)\, dt \\
(\Delta\, y(x))' &= v(x)u(x)f(x) + v'(x) \int_a^x u(t)\, f(t)\, dt \\
&\qquad - u(x)v(x)f(x) + u'(x) \int_x^b v(t)\, f(t)\, dt \\
(p(x)(\Delta\, y(x))')' &= (pv')' \int_a^x u(t)\, f(t)\, dt + pv'uf + (pu')' \int_x^b v(t)\, f(t)\, dt - pu'vf\,.
\end{aligned}$$

Therefore

$$\begin{aligned}
(p(x)(\Delta\, y(x))')' + q\Delta y &= [(pv')' + qv] \int_a^x u(t)\, f(t)\, dt \\
&\qquad + [(pu')' + qu] \int_x^b v(t)\, f(t)\, dt + pf(v'u - u'v) \\
&= f.\Delta
\end{aligned}$$

since u and v are solutions of $Ly = 0$. ■

We can now apply the results of Section 6 to draw conclusions about the eigenfunctions and eigenvalues of the Sturm-Liouville system (*), (†).

Define the operator K by

$$(Kf)(x) = \int_a^b k(x,t)\, f(t)\, dt\,.$$

Since k is continuous on $[a,b] \times [a,b]$ it is clear that $\int\int |k|^2 < \infty$ and so, as shown in Section 4, K is compact.

Let $\mathcal{D}$ be the set of functions y such that Ly exists as a function in $L^2[a,b]$ and satisfies the boundary conditions (†). (There is a little technical hand waving here. A more precise statement is: $y \in \mathcal{D}$ (which is an equivalence class of functions) if there is a representative y such that y is differentiable and $p.(y)'$ has a derivative almost everywhere such that $(p.(y)')' \in L^2[a,b]$. Informally $\mathcal{D}$ is the all $y \in L^2[a,b]$ that qualify as solutions of (*),(†) for some right hand side.) If $y \in \mathcal{D}$ and $f = Ly$ then Theorem 1.1 shows that $Kf = K(Ly) = y$, that is KL acts like the identity on $\mathcal{D}$.

In the other order, it follows from the proof of Theorem 1.1, that for every $f \in L^2[a,b]$ $Kf \in \mathcal{D}$. (The verification that Kf is a solution of $Ly = f$ explicitly shows this. Naturally, for the most general f, the differentiation of expressions like $\int_a^x u(t)\, f(t)\, dt$ one requires the relevant background from Lebesgue integration.) Also, from Theorem 1.1, $LKf = f$. Thus $LK = I$.

Note that L fails to be an inverse of K since it is not defined on the whole of $L^2[a,b]$, the Hilbert space in question. Indeed, since K is compact, it cannot be invertible. However, L *is* defined on a dense subset.

We use these notations and observations in the statements a proofs below.

Theorem 7.2 *(i)* *The operator K does not have $\lambda = 0$ as an eigenvalue.*
(ii) *λ is an eigenvalue of K if an only if $\mu = \frac{1}{\lambda}$ is an eigenvalue of the system* (*),(†).

Consequently, the system (*),(†)

1. *has a countable sequence (μ_i) of real eigenvalues such that $(|\mu_i|) \to \infty$;*
2. *has eigenfunctions which form an orthonormal basis of $L^2[a,b]$;*
3. *has finite-dimensional eigenspaces.*

Proof. (i) If $f \neq 0$ the solution of $Ly = f$ is Kf and cannot be $y = 0$. Therefore 0 is not an eigenvalue of K.

(ii) If λ is an eigenvalue of K the $K\phi = \lambda\phi$ and since ϕ is in the range of K, from the discussion above, $\phi \in \mathcal{D}$. Then $\lambda L\phi = LK\phi = \phi$ so that

$$L\phi = \frac{1}{\lambda}\phi = \mu\phi,$$

and μ is an eigenvalue of (*),(†).

Conversely, if μ is an eigenvalue of (*),(†), by assumption $\mu \neq 0$. We then have $L\phi = \mu\phi$ and $\phi \in \mathcal{D}$. Then

$$KL\phi = \phi = \mu K\phi$$

and so $K\phi = \lambda\phi$ where $\lambda = \frac{1}{\mu}$ is an eigenvalue of K.

The consequences are immediate deductions from the results of Section 6 (principally 6.3, 6.4 and 6.7). Note that in this case the set of eigenvalues of K cannot be finite because this would imply (by Corollary 6.8) that K vanishes on a non-zero (in fact, infinite-dimensional) subspace. ■

The most important result arising from this is consequence 2, since, for example, this is what justifies the expansions that are required in solving partial differential equations by the method of separation of variables.

Note. The assumption that 0 is not an eigenvalue of the system is not an essential restriction. For any constant c, the eigenfunctions of L and $L+c$ are the same and the eigenvalues of $L+c$ are $\lambda + c$ whenever λ is an eigenvalue L. It is a fact that, by adding a suitable constant to q we can always ensure that $\lambda = 0$ is not an eigenvalue of the system. For example, if the boundary conditions are $y(a) = y(b) = 0$, choose c so that $c + q$ does not change sign in $[a, b]$; for definiteness, assume $c + q(t) < 0$ for $a \leq t \leq b$. Let u be the (unique) solution of

$$L_c y = \frac{d}{dt}\left(p(t)\frac{dy}{dt}\right) + [c + q(t)]y = 0, \qquad y(a) = 0, y'(a) = 1\,.$$

Any solution of $L_c y = 0,\ y(a) = 0$ is a multiple of u so to show that 0 is not an eigenvalue we must show that $u(b) \neq 0$.

Since $u'(a) > 0$ and $u(a) = 0$, it follows that u is strictly positive in some interval $(a, a+\delta)$. Suppose z is the smallest zero of u that is $> a$ (if any). Then u' must vanish between a and z. If $z \leq b$ then $(pu')' = -[c+q]u$ is positive in (a, z) and so pu' is increasing in (a, z). But pu' is strictly positive at a so it is strictly positive in (a, z) contradicting that u' vanishes between a and z. Hence 0 is not an eigenvalue of $L_c y = 0, y \in \mathcal{D}$.

Similar, but more complicated arguments can be used for other boundary conditions (see Dieudonne, "Foundations of modern analysis", Chapter XI, Section 7).

The trigonometric functions form an orthonormal basis of $L^2[-\pi, \pi]$. This fact can be deduced from the work of the present section. [The trigonometric functions are actually the eigenfunctions of $y'' = 0$ subject to periodic boundary conditions $y(-\pi) = y(\pi), y'(-\pi) = y'(\pi)$, and such systems are not covered by our work; however the device below gives us the result.]

The eigenfunctions of the system $y'' = 0, y(0) = y(\pi) = 0$ are the functions $\{\sin nt : n = 1, 2, 3, \cdots\}$. Therefore, from Theorem 1.2 these form an orthonormal basis of $L^2[0, \pi]$. Similarly the functions $\{\cos nt : n = 0, 1, 2, 3, \cdots\}$ are the eigenfunctions of the system $y'{,}' = 0, y'(0) = y'(\pi) = 0$ and so also form an orthonormal basis of $L^2[0, \pi]$. (It is true that

0 is an eigenvalue of the latter system, but this is covered by the note above. Alternatively, on can consider the system $y'' + ky = 0, y'(0) = y'(\pi) = 0$ for a suitable constant k – any non-integral k will do).

Now suppose that $f \in L^2[-\pi, \pi]$ is orthogonal to all the trigonometric functions. Then a simple change of variable shows that for each integer n,

$$0 = \int_{-\pi}^{\pi} f(t) \sin nt \, dt = \int_{-\pi}^{0} f(t) \sin nt \, dt + \int_{0}^{\pi} f(t) \sin nt \, dt = \int_{0}^{\pi} [f(t) - f(-t)] \sin nt \, dt \, .$$

Therefore $f(t) = f(-t)$ (almost everywhere) for $0 \leq t \leq \pi$, showing that f is an even function on $[-\pi, \pi]$. A similar calculation with cosines shows that f is also an odd function on $[-\pi, \pi]$. Thus $f = 0$ (almost everywhere) and the fact is proved.

1. Let X be a Banach space (that is, a normed linear space that is complete). A series $\sum_n x_n$ in X is said to be **absolutely convergent** if the series $\sum_n \| x_n \|$ of real numbers is converegent. Prove that an absolutely convergent series in a Banach space is convergent. [Hint: prove that the sequence of partial sums is Cauchy.]

Existence theory for linear initial value problems using operator theory.

2. Let $k(x,t)$ be bounded and square integrable over the square $[a,b] \times [a,b]$, (in particular this will hold if k is continuous on $[a,b] \times [a,b]$). Define $K : L^2[a,b] \to L^2[a,b]$ by

$$(Kf)(x) = \int_a^x k(x,t) f(t)\, dt.$$

Prove that $\| K \| \leq (b-a)M$ where M is a bound for k in $[a,b] \times [a,b]$.

Let k_n be defined inductively by $k_1 = k$ and $k_n(x,t) = \int_t^x k(x,s) k_{n-1}(s,t)\, ds$. Prove that

$$(K^n f)(x) = \int_a^x k_n(x,t) f(t)\, dt.$$

Show by induction that

$$|k_n(x,t)| \leq M^n \frac{|x-t|^{n-1}}{(n-1)!}.$$

Using this result and the formal binomial expansion of $(I-K)^{-1}$, deduce from Question 1 that $(I-K)$ has an inverse in $\mathcal{B}(\mathcal{H})$. [Hint : after proving absolute convergence, verify by multiplication that the sum of the formal expansion is the inverse of $(I-K)$.] For each $\lambda \neq 0$, observe that $\frac{K}{\lambda}$ is an operator of the same type as K and deduce that $(\lambda - K)$ has an inverse in $\mathcal{B}(\mathcal{H})$.

3. Consider the initial value problem

$$(*) \qquad \begin{cases} y^{(n)}(x) + p_1(x) y^{(n-1)}(x) + \ldots + p_n(x) y(x) = f(x) \\ y(0) = y'(0) = \ldots = y^{(n-1)}(0) = 0 \end{cases}$$

where p_i and f are continuous. By putting $u(x) = y^{(n)}(x)$, show that, for any $b > 0$, this problem reduces to

$$(\dagger) \qquad (I-K)u = f$$

where K is an operator on $L^2[0,b]$ of the type considered in Question 6.

[Hint : show that $y^{(n-r)}(x) = \int_0^x \frac{(x-t)^{r-1}}{(r-1)!} u(t)\, dt$.]

Prove that $(\dagger)$ has a unique solution in $L^2[0,b]$ for each $b > 0$. By quoting appropriate theorems show that this solution is continuous (strictly, that the equivalence class of this solution contains a continuous function). Deduce that there is a unique continuous function with n continuous derivatives that satisfies $(*)$ in $[0,\infty)$.

Note that the general initial value problem

$$(**) \qquad \begin{cases} y^{(n)}(x) + p_1(x) y^{(n-1)}(x) + \ldots + p_n(x) y(x) = f(x) \\ y(0) = a_0, y'(0) = a_1, \ldots = y^{(n-1)}(0) = a_{n-1} \end{cases}$$

can be transformed into the form (*) by changing the dependent variable from y to z where

$$z(x) = y(x) - \sum_{k=0}^{n-1} a_k x^k$$

and thus (**) also has a unique solution.

4. Find a Green's function for the system

$$y'' = f, \qquad y(0) = y(1) = 0\,.$$

Check your answer by verifying that it gives $x(x-1)$ as the solution when $f = 2$. Evaluate the eigenvalues and eigenfunctions of

$$y'' = \lambda y, \qquad y(0) = y(1) = 0\,,$$

and consequently find an orthonormal basis of $L^2[0,1]$.

5. Repeat the above question with the system

$$y'' = f, \qquad y(0) + y'(0) = y(1) - y'(1) = 0\,.$$