Chapter 12

Eigenfunction Expansions

The Sturm-Liouville theory provides an infinite sequence of eigenvalue-eigenfunction pairs. Among the triumphs of nineteenth-century mathematics was the realization that these sequences of eigenfunctions can be used to represent arbitrary functions[1] via infinite series. Thus if $\{u_n\}_{n=0}^{\infty}$ is the sequence of eigenfunctions of a Sturm-Liouville problem on an interval $[a, b]$ of the real axis, and f is any function defined on this interval, one seeks a representation of f in the form

$$f(x) = \sum_{0}^{\infty} c_n u_n(x), \tag{12.1}$$

where the constants $\{c_n\}$ must characterize the function f. This is the subject of the present chapter.

12.1 Kinds of Convergence

For a sequence of functions $\{f_k\}$, where each function $f_k(x)$ is defined on an interval I of the real axis, there is a variety of possibilities where convergence is concerned. Some examples are

1. The sequence converges to a function f at some points of the interval but not at others;

2. The sequence converges to a function f at each point of the interval;

3. The sequence converges to a function f *uniformly* on the entire interval.

[1] Well, almost arbitrary.

These examples make different mathematical demands on the convergence. The last kind of convergence, uniform convergence, demands that the preceding kind (pointwise convergence) hold, and more. We have encountered it previously, in Chapter 6, and shall encounter it again below.

What, if anything, can we make of convergence number 1 above? If the subset of points of the interval where convergence fails is not too large, there is a natural way to make sense of this situation. It involves a introducing a certain *norm* of a function f on I in the form of a definite integral[2]:

Definition 12.1.1 *For any function f that is integrable on I, define*

$$\|f\| = \left\{ \int_I f(x)^2 \, dx \right\}^{1/2}. \tag{12.2}$$

This definition assigns to a function f a single number, $\|f\|$, which is a measure of the size of the function; in particular, for continuous functions u the vanishing of $\|u\|$ implies that $u \equiv 0$ on the interval I, and this endows the function $\|\cdot\|$ with the properties usually associated with a norm: it vanishes when and only when $u \equiv 0$, it satisfies the triangle inequality (Corollary 12.2.1 below) and it satisfies the homogeneity condition $\|\lambda u\| = |\lambda| \, \|u\|$ for any constant λ. The norm (12.2) is often referred to as the L^2 norm of the function f and written $\|f\|_2$. We'll use this terminology but dispense with the identifying subscript since we shall not need it. In terms of this norm, we can define a more general kind of convergence: the sequence f_k of functions on I will be said to *converge in* L^2 to a function f if

$$\lim_{k\to\infty} \|f_k - f\| = 0. \tag{12.3}$$

This L^2 convergence is also sometimes referred to as *mean-square* convergence. It is easy to see that it can hold in the absence of pointwise convergence on the entire interval, (number 2 above), provided the set where the pointwise convergence fails is small.

Example 12.1.1 Consider the functions on $[0,1]$ defined as follows: $f_1(x) = 1$, and, for $k = 2, 3, \ldots$,

$$f_k(x) = \begin{cases} 1 \text{ if } x \neq j/k \\ 0 \text{ if } x = j/k \end{cases}, \; j = 1, 2, \ldots k-1.$$

[2] We implicitly assume here and below that all integrals that appear are Riemann integrals.

Then if x is any irrational number in $(0,1)$ $f_k(x) = 1$ for each k and therefore, if we define $f(x) = 1$, then $f_k(x) \to f(x)$ for each irrational x. However, if x is rational, say $x = p/q$ for integers p and $q > p$ without common factors, then $f_k(x) = 0$ if $k = q, 2q, 3q, \ldots$. Thus $\{f_k(p/q)\}$ cannot converge to $f(p/q) \equiv 1$, since a subsequence converges to zero (indeed $\{f_k(p/q)\}$ cannot converge at all, since other subsequences converge to one). This sequence therefore cannot converge pointwise to the function f, but

$$\|f_k - f\|^2 = \int_0^1 (f_k(x) - f(x))^2 \, dx = 0,$$

because the integrand differs from zero only at finitely many points. Hence $f_k \to f$ in L^2. □

L^2 convergence is weaker than uniform convergence in the sense that it allows convergence of some sequences of functions that are not uniformly convergent, as in the preceding example. On the other hand, as the following theorem shows, any uniformly convergent sequence is also convergent in L^2.

Theorem 12.1.1 *If the sequence of integrable functions $\{f_k\}$ on I converges uniformly there to an integrable function f, then $f_k \to f$ in L^2.*

Proof: Given $\epsilon > 0$ choose K such that $k \geq K$ implies that $|f_k(x) - f(x)| < \epsilon/\sqrt{l}$, where l is the length of I. Then

$$\|f_k - f\|^2 = \int_I |f_k(x) - f(x)|^2 \, dx \leq \epsilon^2$$

if $k \geq K$. This proves the L^2 convergence. □

One can define a norm for bounded, continuous functions on an interval I that is parallel to that given for integrable functions in Definition 12.1.1 above. If f is any bounded, continuous function on I, we define its *uniform norm* to be the non-negative real number

$$\|f\|_U = \sup_{x \in I} |f(x)|.$$

Then if $\{f_k\}$ is a sequence of bounded, continuous functions on I, it is said to converge to the function f on I in the uniform norm if

$$\lim_{k \to \infty} \|f_k - f\|_U = 0.$$

This means that if $\epsilon > 0$ is given, there is $K > 0$ such that $k > K$ implies that $\|f_k - f\| < \epsilon$. In other words, for any $x \in I$,

$$|f_k(x) - f(x)| \leq \sup_{x\in I} |f_k(x) - f(x)| = \|f_k - f\| < \epsilon$$

if $k > K$. Since K, by the manner of its definition, cannot depend on x, this shows that convergence in the uniform norm implies uniform convergence. The reverse, that uniform convergence implies convergence in the uniform norm is likewise true, so the two are equivalent.

12.2 Inner Product and Orthogonality

Recall that the sequence of Sturm-Liouville eigenfunctions is orthogonal with respect to the weight function ρ:

$$\int_a^b u_m(x)\, u_n(x)\, \rho(x)\ dx = 0 \text{ if } m \neq n. \tag{12.4}$$

Consider equation (12.1) above. The coefficients in this formula can be determined with the aid of this orthogonality relation as follows. Multiply each side of equation (12.1) by $\rho(x)\, u_m(x)$ and integrate from a to b; on the right-hand side assume term-by-term integration is allowed. Because of the orthogonality relation all the terms on right vanish except that one for which $n = m$, and one obtains

$$c_m = \frac{\int_a^b f(x)\, u_m(x)\, \rho(x)\ dx}{\int_a^b u_m(x)^2\, \rho(x)\ dx}. \tag{12.5}$$

Irrespective of the legitimacy of the procedure for obtaining this formula, these coefficients, called the Fourier coefficients of the function f with respect to the given system of orthogonal functions $\{u_n\}$, provide us with a precisely defined convergence problem: that of finding under what conditions, and in what sense, the series $\sum c_n u_n$, with the coefficients $\{c_n\}$ given by equation (12.5), converges to the function f.

Before proceeding with this problem, we introduce some conventions. First, we simplify the expression (12.5) above and further expressions that we obtain below by assuming that

$$\rho(x) = 1. \tag{12.6}$$

This is done principally for convenience and we discuss the relaxation of this assumption in §12.6 below. Second, for the kind of integral appearing on the right-hand side of equation (12.5), we introduce the notation

$$(u, v) = \int_a^b u\,(x)\,v\,(x)\;dx. \tag{12.7}$$

This number, determined by the pair of functions u and v, will be referred to as the *inner product* of these two functions. In this notation, the square of the norm of u is given by $\|u\|^2 = (u, u)$.

The inner product satisfies a frequently used inequality:

Theorem 12.2.1 *(The Schwarz Inequality)*

$$|(u, v)| \leq \|u\|\|v\|. \tag{12.8}$$

Proof: The identity

$$\|u - tv\|^2 = \|u\|^2 - 2t\,(u, v) + t^2\|v\|^2$$

shows that the right-hand side is non-negative for any real value of the parameter t. The result (12.8) can easily be shown to hold trivially if $\|v\| = 0$; assuming then that $\|v\| \neq 0$, we may set $t = (u, v)\,/\|v\|^2$. This gives the stated inequality. □

Corollary 12.2.1 *(The Triangle Inequality)*

$$\|u + v\| \leq \|u\| + \|v\|. \tag{12.9}$$

Proof:

$$\|u+v\|^2 = \|u\|^2+\|v\|^2+2\,(u, v) \leq \|u\|^2+\|v\|^2+2\|u\|\|v\| = (\|u\| + \|v\|)^2\,. \quad \square$$

Corollary 12.2.2 *Suppose the sequence $\{f_k\}$ converges to f in L^2. Then*

1. $(f_k, f) \to \|f\|^2$ *and*
2. $\|f_k\| \to \|f\|$.

Proof: Our assumption is that $\|f_k - f\| \to 0$ as $k \to \infty$. Write $(f_k, f) = (f_k - f, f) + \|f\|^2$. Then, by the Scwharz inequality,

$$\left|(f_k, f) - \|f\|^2\right| = |(f_k - f, f)| \leq \|f_k - f\|\|f\| \to 0$$

as $k \to \infty$, proving the first statement. The identities

$$\|f_k - f\|^2 = \|f_k\|^2 + \|f\|^2 - 2\,(f_k, f) = \|f_k\|^2 - \|f\|^2 + 2\left(\|f\|^2 - (f_k, f)\right)$$

together with the first statement prove the second. □

In the notation of equation (12.7) the formula for the Fourier coefficient c_m becomes

$$c_m = \frac{(u_m, f)}{(u_m, u_m)} = \frac{(u_m, f)}{\|u_m\|^2}, \tag{12.10}$$

Next, we recall (cf. §10.2) that, if $u(x)$ is an eigenfunction of a linear, homogeneous boundary-value problem, then so is $ku(x)$ for any constant $k \neq 0$. Choosing $k = k_n \equiv \|u_n\|^{-1}$ we find that $v_n(x) = k_n u_n(x)$ is an eigenfunction belonging to the same eigenvalue, but with norm one: $\|v_n\| = 1$. We say that such an eigenfunction is *normalized*, and the sequence of eigenfunctions, which now satisfies the conditions

$$(v_n, v_m) = \begin{cases} 0 \text{ if } m \neq n \\ 1 \text{ if } m = n \end{cases},$$

is said to be *orthonormal*. We shall usually assume below that orthogonal systems of functions that we encounter are in fact orthonormal; this then further simplifies the expression for the Fourier coefficients to $c_m = (u_m, f)$. The following theorem helps to characterize the Fourier coefficients of a function f relative to an orthonormal set of functions $\{u_n\}$ on an interval $[a, b]$:

Theorem 12.2.2 *(Bessel's inequality) Let $\{u_k\}$ be an orthonormal sequence of functions on $[a, b]$, suppose f is any Riemann-integrable function on that interval, and denote the Fourier coefficients of f with respect to this sequence by $\{c_k\} : c_k = (u_k, f)$. Then the series $\sum_0^\infty c_k^2$ converges and satisfies the inequality*

$$\sum_{k=0}^{\infty} c_k^2 \leq \|f\|^2. \tag{12.11}$$

Proof: The identities

$$\int_a^b \left(f(x) - \sum_{k=0}^{n} c_k u_k(x)\right)^2 dx = \|f\|^2 - 2\sum_{k=0}^{n} c_k (u_k, f) + \sum_{k=0}^{n} c_k^2 = \|f\|^2 - \sum_{k=0}^{n} c_k^2$$

follow from the definition of the Fourier coefficients. Since the leftmost expression is non-negative, it follows that

$$\sum_{k=0}^{n} c_k^2 \leq \|f\|^2.$$

This shows that the partial sums of the series $\sum c_k^2$ are bounded above. Since they are nondecreasing, they converge to their supremum. It also shows that the inequality (12.11) holds. $\square$

Remarks:

- The considerations of this section are independent of the theory of Sturm-Liouville systems, or indeed of any boundary-value problems. Theorem 12.2.2 relates to orthonormal sequences of functions, however they may be obtained.

- A consequence of the convergence of the series $\sum c_k^2$ is that $c_k \to 0$ as $k \to \infty$, i.e.
$$\lim_{k\to\infty} \int_a^b u_k\left(x\right) f\left(x\right)\, dx = 0$$
for any integrable function f, and any orthonormal sequence $\{u_k\}$.

- We'll use this theorem principally for continuous functions on $[a, b]$ but we have stated it more generally, for Riemann-integrable functions; it is true with greater generality yet, but we shall not pursue this.

12.3 Extremal Properties of Eigenvalues

It has been known since the nineteenth century that the eigenvalues of a self-adjoint problem have an extremal character. For definiteness, consider the Sturm-Liouville problem, equation (10.11) above with $\rho \equiv 1$,

$$Lu = \lambda u, \text{ for } a < x < b \tag{12.12}$$

together with the separated boundary conditions (10.20). Functions u and v that are twice differentiable and satisfy these boundary conditions will be called *admissible*. For these admissible functions, $(v, Lu) = (Lv, u)$. If λ, u are an eigenvalue-eigenfunction pair, then, multiplying each side of equation (12.12) with u and integrating, one obtains for λ the expression

$$\lambda = \frac{(u, Lu)}{(u, u)}. \tag{12.13}$$

Now consider this formula, not as an expression for the eigenvalue in terms of the eigenfunction, but merely as the definition of a number λ associated to any admissible function u, not necessarily an eigenfunction. We'll denote this number by $\lambda[u]$ to emphasize its dependence on the choice of the function

u. Our claim is that this *functional* of u has an extremal character when u is an eigenfunction of L. We need to make precise what we mean by this.

A function $F(x) = F(x_1, x_2, \ldots, x_n)$ of n variables has an extremal character (maximum, minimum or saddle) at a point $a = (a_1, a_2, \ldots, a_n)$ if its partial derivatives all vanish there or, alternatively, if the directional derivative of F at a vanishes in any direction v:

$$\frac{d}{dt} F(a + tv)\bigg|_{t=0} = 0.$$

It is this idea that we use to characterize the extremal character of the functional λ. Let u be given and let v be any other admissible function. Then $u + tv$ is also admissible for any choice of the parameter t, and we consider $l(t) \equiv \lambda[u + tv]$. Then

$$l(t) = \frac{(u, Lu) + 2t(Lu, v) + t^2(v, Lv)}{(u, u) + 2t(u, v) + t^2(v, v)}$$

and dl/dt evaluated at $t = 0$ represents the directional derivative, in the "direction" v, at the "point" u. A straightforward computation gives the formula

$$\frac{dl}{dt}\bigg|_{t=0} = \frac{2}{(u, u)}(Lu - \lambda[u]u, v).$$

It is clear that this vanishes for arbitrary v if u is an eigenfunction; it is plausible that if it vanishes for all admissible choices of v, then u must be an eigenfunction. This plausibilty can be turned into certainty, but we forego the proof since the development below does not depend on it. This is the extremal property of eigenvalues. It can be developed into a powerful tool for the approximation of eigenvalues, and we shall exploit it below to prove the validity of eigenfunction expansions.

Remarks:
This extremal property is not restricted to the Sturm-Liouville problem, as can be seen by abstracting from the calculation above the assumptions that were used in it:

1. The operator L is self-adjoint.

2. The class of admissible functions forms a linear family (if u and v are members, so is $u + tv$ for any constant t).

PROBLEM SET 12.3.1

1. Show that if u is continuous on I the vanishing of $\|u\|$ as defined in Definition 12.1.1 implies that $u(x) \equiv 0$ on I.

2. Prove that a uniformly convergent sequence of functions converges in the uniform norm.

3. Prove the assertion made in the proof of Theorem 12.2.1: if $\|v\| = 0$ then $(u, v) = 0$.

4. On the interval $[0, 1]$, let

$$f_k(x) = \begin{cases} x^k \text{ if } 0 \leq x < 1, \\ (-1)^k \text{ if } x = 1. \end{cases}$$

   Show that $\{f_k(x)\}$ fails to converge if $x = 1$ but that with $f \equiv 0$, f_k converges to f in L^2.

5. Let $\{u_n\}$ be an orthonormal sequence of functions on an interval I, and let f be a given integrable function on I. It is required to find the constants $c_1, c_2, \ldots c_N$ providing the best approximation to f by a sum $\sum_{i=1}^{N} c_i u_n$. Here 'best' means that the chosen constants minimize the expression

$$\phi(c) = \phi(c_1, c_2, \ldots, c_N) = \int_I \left\{ f(x) - \sum_{i=1}^{N} c_i u_i(x) \right\}^2 dx.$$

6. Let $\{u_n\}$ be an orthonormal sequence of functions on an interval I. Show that they are linearly independent on I (recall Definition 2.1.1).

> The Gram-Schmidt procedure produces from a linearly independent set of functions $u_0, u_1, \ldots, u_n$ on the interval I an orthonormal set $v_0, v_1, \ldots, v_n$ as follows: put $v_0 = \|u_0\|^{-1} u_0$. Then, successively for $k = 1, 2, \ldots, n$,
>
> (a) First put $\tilde{v}_k = u_k - \sum_{j=0}^{k-1} c_j v_j$ and choose $\{c_j\}$ so that $(\tilde{v}_k, v_j) = 0$ for $j = 0, 1, \ldots, k-1$.
>
> (b) Then put $v_k = \|\tilde{v}_k\|^{-1} v_k$.

7. Explain why linear independence of the original set $\{u_n\}$ is needed to make the Gram-Schmidt procedure 'work.'

8. Let $u_k(x) = x^k$ and take $I = [-1, 1]$. Find v_0, v_1 and v_2.

9. Let $u_k(x) = \cos(kx)$ and take $I = [0, \pi/2]$. Find v_0, v_1 and v_2.

10. Consider the Sturm-Liouville problem $-u'' = \lambda u$ on $(0,1)$ with boundary data $u(0) = u(1) = 0$. The least eigenvalue is π^2. Approximate π^2 as follows: the functions

$$u(x) = x(1-x)\left(1 + \alpha_1 x + \alpha_2 x^2 + \cdots\right)$$

are admissible. Evaluate the functional 12.13 using this expression with $\alpha_k = 0$ for all k, i.e., without any parameters, and compare the result to $\pi^2 = 9.8696\cdots$. Try a second approximation, exploiting the knowledge that the eigenfunction should be symmetric in the line $x = 1/2$, i.e. that $u(1-x) = u(x)$. To do this, choose the trial function u above to have the form

$$u(x) = x(1-x) + \alpha x^2(1-x)^2,$$

and choose the best value of α.

12.4 The Norm of an Operator

The Sturm-Liouville problem involves a differential equation $Lu = \lambda u$. The differential operator L acts on functions in $C^2[a,b]$. The latter is a linear family of functions in that whenever u and v are C^2 functions, so also is $\alpha u + \beta v$ for arbitrary constants α and β. This is one of many examples of linear operators acting on linear families of functions.

Example 12.4.1 *Consider the family of bounded, continuous functions on an interval I and the operator Γ defined by the formula $(\Gamma u)(x) = \gamma u(x)$ where γ is a specified real constant.*

Example 12.4.2 *Let the family be C^1 functions on an interval I and define Γ by $(\Gamma u)(x) = u'(x)$.*

We wish to define the norm of an operator – a number that characterizes how big the operator can get. This must be closely connected to the norm (12.1.1) of functions. A realistic and widely used definition of the norm of a linear operator is

$$\|\Gamma\| = \sup_{\|u\| \neq 0} \frac{\|\Gamma u\|}{\|u\|} \equiv \sup_{\|u\|=1} \|\Gamma u\|. \tag{12.14}$$

This definition embodies the implicit assumption that the linear family to which u belongs is such that both $\|u\|$ and $\|\Gamma u\|$ are well defined; this will be so in our applications below.

The supremum in the definition (12.14) might be infinite; such an operator is *unbounded*. If the supremum is finite, the operator is *bounded* and its norm is the non-negative real number $\|\Gamma\|$.

Example 12.4.3 *For the operator of Example 12.4.1 above it is easy to see that* $\|\Gamma\| = |\gamma|$.

Example 12.4.4 *The operator of Example 12.4.2 is unbounded. To see this, let the interval be* $[0,1]$ *and consider the sequence of normalized functions*

$$u_k(x) = \sqrt{2k+1}x^k, \; k = 1, 2, \ldots$$

A simple calculation shows that

$$\|u_k'\| = k\sqrt{\frac{2k+1}{2k-1}},$$

which is unbounded as $k \to \infty$.

The reasoning of the preceding example can be extended to other differential opeators, like the second-order, self-adjoint operator of the Sturm-Liouville theory, and we conclude that the latter operator is unbounded. This presents an obstacle to a rigorous demonstration of the convergence of the expansion (12.1). However, we can overcome this obstacle. The key to doing this is the Green's function.

Denote by $G(x,\xi)$ the Green's function for the boundary-value problem

$$Lu \equiv -\left(p(x)\,u'\right)' + q(x)\,u = r, \; a < x < b \tag{12.15}$$

together with the separated end-point conditions (10.20). It is described in detail in §10.4.1 where it is derived as a function not only of the variables x and ξ but also of the parameter λ. In the present case we shall set the parameter λ equal to zero and provisionally assume:

> Zero is not an eigenvalue of the Sturm-Liouville problem (10.11), (10.20).

This is again an assumption made for convenience and we show how to relax it in §12.6 below. Now, in view of Theorem 10.4.1 we may write the solution of the problem (12.15) in the form

$$u(x) = \int_a^b G(x,\xi)\,r(\xi)\;d\xi \equiv (\Gamma r)(x)\,. \tag{12.16}$$

It follows that if (u, λ) are an eigenfunction-eigenvalue pair for the eigenvalue problem $Lu = \lambda u$ together with the separated end-point conditions (10.20), then

$$u(x) = \lambda \int_a^b G(x, \xi)\, u(\xi)\; d\xi. \tag{12.17}$$

Conversely, if (u, λ) represent a solution of the linear integral equation (12.17), then application of the operator L shows that they represent a solution of the original, differential version of the problem as well.

The operator Γ is a bounded operator. To see this, note that since G is a continuous function on a closed, bounded set ($[a, b] \times [a, b]$), it is bounded there; say $|G(x, \xi)| < M$. Therefore

$$|\Gamma u(x)| \leq M \int_a^b |u(\xi)|\; d\xi \leq M\sqrt{b-a}\|u\|, \tag{12.18}$$

where we have used the Schwarz inequality to get the last term. Then

$$\|\Gamma u\|^2 = \int_a^b (\Gamma u)^2\; dx \leq M^2 (b-a)^2 \|u\|^2,$$

so, for $\|u\| = 1$, $\|\Gamma u\| \leq M(b-a)$. We now rewrite equation (12.17) in the form

$$(\Gamma u)(x) = \mu u(x), \tag{12.19}$$

where $\mu = \lambda^{-1}$.

This reformulates the original differential equation plus boundary data as a single integral equation, in which the boundary data have been incorporated. The eigenfunctions of problem (12.19) are the same as those of the Sturm-Liouville problem, the eigenvalues are the reciprocals, and the operator is now bounded. By virtue of the symmetry of the Green's function in x and ξ (cf. Problem Set 12.3.1, Problem 6), the operator Γ is self-adjoint on the family $C[a, b]$ consisting of all continuous functions on the interval $[a, b]$:

$$(u, \Gamma v) = (\Gamma u, v) \text{ for } u, v \in C[a, b].$$

From this it follows that the eigenvalues $\{\mu\}_k$ enjoy the extremal property described in the preceding section. Another consequence of the self-adjointness is an alternative characterization of the norm. It is true in general that if u is normalized,

$$|(u, \Gamma u)| \leq \|\Gamma\|; \tag{12.20}$$

when Γ is self-adjoint we have the following

Theorem 12.4.1 *For Γ self-adjoint and bounded,*

$$\sup_{\|u\|=1} |(u, \Gamma u)| = \|\Gamma\|.$$

Proof: In view of the inequality (12.20) above, it suffices to show that

$$\gamma \equiv \sup_{\|u\|=1} |(u, \Gamma u)| \geq \|\Gamma\|.$$

To this end we consider the inequalities

$$\begin{aligned} (\Gamma (u+v), (u+v)) &\leq \gamma \|u+v\|^2 \text{ and} \\ (\Gamma (u-v), (u-v)) &\geq -\gamma \|u-v\|^2. \end{aligned}$$

Expanding the left-hand sides, using the self-adjointness and subtracting, we find

$$(\Gamma u, v) \leq \frac{\gamma}{2} \left(\|u\|^2 + \|v\|^2 \right).$$

If $\|\Gamma u\| = 0$ then certainly $\|\Gamma u\| \leq \gamma$. If $\|\Gamma u\| \neq 0$, set $v = k (\Gamma u)$, where k is as yet unchosen. Then

$$2k \|\Gamma u\|^2 \leq \gamma \left(\|u\|^2 + k^2 \|\Gamma u\|^2 \right).$$

This implies that $\gamma > 0$. Now choose $k = 1/\gamma$; this gives

$$\gamma \geq \frac{\|\Gamma u\|}{\|u\|},$$

where we may assume the $\|u\| \neq 0$ since $\Gamma u \neq 0$. This shows that $\gamma \geq \|\Gamma\|$. □

Remarks:

1. Theorem 12.4.1 does not refer explicitly to the Sturm-Liouville problem, but is more generally a property of bounded, self-adjoint operators.

2. It follows from Theorem 12.4.1 that either $\sup_{\|u\|=1} (u, \Gamma u) = \|\Gamma\|$ or $\inf_{\|u\|=1} (u, \Gamma u) = -\|\Gamma\|$.

12.5 The Sequence of Eigenfunctions

In this section we obtain further properties of the eigenvalues and eigenfunctions and demonstrate the convergence of the series (12.1) under certain conditions on the function f (Theorem 12.5.3 below). We focus on the self-adjoint eigenvalue problem (12.19) in integral form, since its eigenfunctions are the same as those of the Sturm-Liouville problem if G is the Green's function for that problem. The development provides convergence of eigenfunction expansions for bounded, self-adjoint integral operators and therefore is more general than that which is strictly needed for the Sturm-Liouville theory. We shall need some further ideas from classical, real analysis; we now proceed to these.

A family of functions is *equicontinuous* on I if, given $\epsilon > 0$ there is $\delta > 0$ such that $|f(x_1) - f(x_2)| < \epsilon$ whenever $|x_1 - x_2| < \delta$, where f is any member of the family. Since the number δ is independent of x_1 and x_2, each member of an equicontinuous family is uniformly continuous. More is true of such a family: the number δ can also be chosen independently of the member of the family. A family of functions is said to be *uniformly bounded* if there is a fixed number M such that $|f(x)| \leq M$ for every member of the family and every $x \in I$.

Lemma 12.5.1 *Let Γ be defined as the integral operator in equation (12.16), where the function G is continuous on $[a,b] \times [a,b]$. Denote by U the family of functions*

$$U = \{u : u \textit{ is continuous on } [a,b] \textit{ and } \|u\| = 1\}.$$

Then the family of functions

$$\Gamma U = \{v : v = \Gamma u \textit{ for } u \in U\}$$

is equicontinuous and uniformly bounded.

Proof: Consider first the equicontinuity. For v in ΓU we have

$$v(x_1) - v(x_2) = \int_a^b \{G(x_1, \xi) - G(x_2, \xi)\} u(\xi)\, d\xi.$$

Since G is continuous on $[a,b] \times [a,b]$ it is uniformly continuous there and therefore, given $\epsilon > 0$, there is $\delta > 0$ such that

$$|G(x_1, \xi) - G(x_2, \xi)| < \epsilon/\sqrt{b-a} \text{ if } |x_1 - x_2| < \delta.$$

Therefore

$$|v(x_1) - v(x_2)| < \frac{\epsilon}{\sqrt{b-a}} \int_a^b |u(\xi)|\, d\xi \leq \epsilon \|u\|,$$

where we have used Schwarz's inequality in obtaining the last term. Since $\|u\| = 1$, this demonstrates the equicontinuity. As to the uniform boundedness, since $|G(x, \xi)| \leq M$, we find

$$|v(x)| \leq M \int_a^b |u(\xi)|\, d\xi \leq M\sqrt{b-a}\|u\| = M\sqrt{b-a}. \quad \square$$

The properties of the family ΓU expressed in this Lemma are important because of the following result of real analysis[3]

Theorem 12.5.1 *(Ascoli-Arzela Theorem) Let F be an infinite, equicontinuous, uniformly bounded family of functions on a bounded set Ω in R^n. Then there exists a sequence $\{f_n\}$ of distinct functions in F which converges uniformly on Ω to a function $f \in F$.*

Of course, the limit function f of this theorem is necessarily continuous on Ω since it is the uniform limit of continuous functions. The form in which we use this theorem is

Corollary 12.5.1 *Any equicontinuous, uniformly bounded sequence of functions on a bounded set has a uniformly convergent subsequence.*

We are now ready to give a more general proof of the extremal property of eigenvalues discussed in §12.3 above.

Theorem 12.5.2 *Let Γ, U and ΓU be as in Lemma 12.5.1 above, and assume further that the function G is symmetric in x and ξ so that Γ is self-adjoint. Then either $\|\Gamma\|$ or $-\|\Gamma\|$ is an eigenvalue.*

Proof: Since Γ is self-adjoint we may use the alternative characterization of the norm $\|\Gamma\|$ of Theorem 12.4.1. Assume for definiteness that $\sup_{\|u\|=1}(u, \Gamma u) = \|\Gamma\|$; the alternative possibility $\inf_{\|u\|=1}(u, \Gamma u) = -\|\Gamma\|$ is treated similarly. Then there exists a sequence $\{u_m\}$ in U such that $(u_m, \Gamma u_m) \to \gamma$, where we have written $\gamma = \|\Gamma\|$. By the Ascoli-Arzela theorem there exists a subsequence of the sequence $\{\Gamma u_m\}$ that converges uniformly. Denote the limit function by v and re-index so that we may continue to designate by $\{u_m\}$

[3]For a proof, see for example *Principles of Mathematical Analysis* by Walter Rudin, ??.

the sequence such that $\Gamma u_m \to v$. We'll show that (γ, v) is an eigenvalue-eigenfunction pair.

We may as well assume that $\gamma > 0$ since otherwise we can easily infer that $\Gamma u = 0$ for $u \in U$ and the conclusion of the theorem holds trivially. From this we infer that $\|v\| > 0$, as follows:

$$\|\Gamma u_m - \gamma u_m\|^2 = \|\Gamma u_m\|^2 + \gamma^2 \|u_m\|^2 - 2\gamma\left(\Gamma u_m, u_m\right) \to \|v\|^2 - \gamma^2,$$

where we appeal to Theorem 12.1.1 and Corollary 12.2.2. Since this is evidently non-negative, $\|v\| \geq \gamma$ is indeed positive. Since $\|\Gamma u_m\|^2 \leq \gamma^2$, the preceding inequality also shows that

$$0 \leq \|\Gamma u_m - \gamma u_m\|^2 \leq 2\gamma\left(\gamma - \left(\Gamma u_m, u_m\right)\right).$$

Since the term on the right tends to zero, it follows that

$$\lim_{m \to \infty} \|\Gamma u_m - \gamma u_m\| = 0. \tag{12.21}$$

To see that v is an eigenfunction with eigenvalue γ we first decompose the difference $\Gamma v - \gamma v$ as follows:

$$\Gamma v - \gamma v = \left(\Gamma v - \Gamma^2 u_m\right) + \left(\Gamma^2 u_m - \gamma \Gamma u_m\right) + \gamma\left(\Gamma u_m - v\right), \tag{12.22}$$

where Γ^2 represents the operator applied twice:

$$\left(\Gamma^2 u\right)(x) = \int_a^b G(x, \xi)\left(\Gamma u\right)(\xi)\, d\xi.$$

We may estimate the first term on the right-hand side of equation (12.22) as follows:

$$\|\Gamma^2 u_m - \Gamma v\| = \|\Gamma\left(\Gamma u_m - v\right)\| \leq \gamma \|\Gamma u_m - v\|.$$

We find exactly the same estimate for the last term and, estimating the middle term similarly, obtain from equation 12.22) the inequalities

$$0 \leq \|\Gamma v - \gamma v\| \leq 2\gamma \|\Gamma u_m - v\| + \gamma \|\Gamma u_m - \gamma u_m\|.$$

It now follows from the L^2 convergence of $\{\Gamma u_m\}$ to v and from equation (12.21) above that $\|\Gamma v - \gamma v\| = 0$ and therefore (since v and Γv are continuous,

$$\left(\Gamma v\right)(x) = \gamma v(x).$$

In accordance with the observation above that $\|v\| \neq 0$, the function v is not identically zero and is therefore indeed an eigenfunction of Γ with eigenvalue γ. $\square$

This theorem can be iterated to generate a succession of eigenvalues and eigenfunctions in the following way. Denote by μ_0 the eigenvalue γ guaranteed by the theorem, and by u_0 the normalized eigenfunction: $u_0(x) = \|v\|^{-1} v(x)$. Define

$$G_1(x,\xi) = G(x,\xi) - \mu_0 u_0(x) u_0(\xi),$$

and by Γ_1 the corresponding operator on the family of continuous functions on $[a,b]$:

$$(\Gamma_1 u)(x) = \int_a^b G_1(x,\xi) u(\xi)\, d\xi.$$

Since G_1 is bounded and symmetric, the operator Γ_1 is bounded and self-adjoint and the theorem can be applied with Γ_1 in place of Γ, guaranteeing an eigenvalue μ_1 and a normalized eigenfunction u_1, provided $\|\Gamma_1\| = |\mu_1| \neq 0$; provisionally assume this. The operator equation takes the form

$$\mu_1 u_1 = \Gamma_1 u_1 = \Gamma u_1 - \mu_0 (u_0, u_1) u_0. \tag{12.23}$$

Note that

$$\|\Gamma_1 u\|^2 = \|\Gamma u\|^2 - \mu_0^2 (u, u_0)^2 \leq \|\Gamma u\|^2,$$

showing that $|\mu_1| \leq |\mu_0|$. Taking the inner product of each side with u_0 gives

$$\mu_1 (u_0, u_1) = (u_0, \Gamma u_1) - \mu_0 (u_0, u_1).$$

But the right-hand side vanishes by virtue of the self-adjointness of Γ and, since $\mu_1 \neq 0$, we find that $(u_0, u_1) = 0$. Returning to equation (12.23), we find that u_1 is also an eigenfunction of the original operator Γ, belonging to the eigenvalue μ_1, and that it is orthogonal to u_0. Next consider the symmetric function

$$G_2(x,\xi) = G_1(x,\xi) - \mu_1 u_1(x) u_1(\xi).$$

Define the bounded, self-adjoint operator Γ_2 in the now-obvious fashion, and assume that $\|\Gamma_2\| > 0$. We then infer the existence of an eigenvalue μ_2 and a corresponding eigenfunction u_2, satisfying

$$\mu_2 u_2 = \Gamma_2 u_2 = \Gamma u_2 - \mu_0 (u_0, u_2) u_0 - \mu_1 (u_1, u_2) u_1.$$

Taking the inner product first with u_0 and then with u_1 we find that u_2 is orthogonal to each of these and is an eigenfunction of Γ_1 and also of Γ. Since μ_1 is extremal with respect to Γ_1 and μ_2 is also an eigenvalue of Γ_1, we have $|\mu_2| \leq |\mu_1| \leq |\mu_0|$. We summarize this as follows:

Proposition 12.5.1 *Suppose in the successive definitions*

$$G_m(x,\xi) = G_{m-1}(x,\xi) - \mu_{m-1} u_{m-1}(x)\, u_{m-1}(\xi)$$

above, one has $\|\gamma_m\| > 0$ for each $m = 0, 1, 2, \ldots$. Then there is an infinite sequence of eigenvalues $\{\mu_m\}$ and eigenfunctions with $|\mu_m| \leq |\mu_{m-1}|$, and a corresponding sequence of mutually orthogonal eigenfunctions.

This presents in a weaker form what we already know about the Sturm-Liouville problem. There the eigenvalues $\lambda_m \to +\infty$ so, for the corresponding integral operator Γ, the eigenvalues $\mu_m \to 0+$ and the reduced operators Γ_m all must have positive norm. Another item of information about the Sturm-Liouville problem is that there can be no repeated eigenvalues: they are all simple in the sense that to an eigenvlaue μ there corresponds a unique (up to multiplication by a scalar) eigenfunction. But in the description above, there is nothing to prevent multiple eigenvalues, and indeed this may occur for some bounded, self-adjoint operators. However, there is always the limit imposed by the following result:

Proposition 12.5.2 *Under the conditions of Proposition 12.5.1 series $\sum_{m=0}^{\infty} \mu_m^2$ converges; in particular, any nonzero eigenvalue μ has finite multiplicity.*

Proof: In the relation

$$\int_a^b G(x,\xi)\, u_m(\xi)\; d\xi = \mu_m u_m(x),$$

view the left hand side as the Fourier coefficient $c_m(x)$ of G, where the latter is regarded as a function of ξ for fixed x. Then Bessel's inequality implies that

$$\sum_{k=0}^{m} \mu_k^2 u_m(x)^2 \leq \int_a^b G(x,\xi)^2\; d\xi$$

for an integer $m \geq 0$. Integrating from a to b, we find, with M as a bound for $|G|$,

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

This shows that the series converges and also shows that any eigenvalue $\mu \neq 0$ must have finite multiplicity. $\square$

We are now ready to prove a basic convergence result for an expansion of the form (12.1) under restrictions on the function f allowing it to be expressed in the form $f = \Gamma u$ for some continuous function u. For the Sturm-Liouville problem (and for other self-adjoint boundary value problems) it is sufficient that

> f is in $C^2[a,b]$ and satisfies the same homogeneous boundary conditions as the eigenfunctions.

For then the equation $Lf = u$ defines u as a continuous function on $[a,b]$ and, since f satisfies the boundary conditions, it has the (unique) solution $f = \Gamma u$. We discuss below in §12.6 the 'reasonableness' of this condition.

Theorem 12.5.3 *Let f be expressible in the form Γu for a continuous function u on $[a,b]$. Then the series (12.1) converges to f uniformly on $[a,b]$.*

Proof: For the reduced Green's function

$$G_m(x,\xi) = G(x,\xi) - \sum_{k=0}^{m-1} \mu_k u_k(x)\, u_k(\xi),$$

the maximal eigenvalue μ_m satisfies the equation $|\mu_m| = \|\Gamma_m\|$. Therefore, for any continuous function u on $[a,b]$,

$$\|\Gamma_m u\| = \|\Gamma u - \sum_{k=0}^{m-1} \mu_k (u, u_k)\, u_k\| \le |\mu_m|\, \|u\|$$

and we infer from Lemma 12.5.2 that

$$\lim_{m\to\infty} \|\Gamma u - \sum_{k=0}^{m-1} \mu_k (u, u_k)\, u_k\| = 0. \tag{12.24}$$

Consider now the sum

$$\sum_{k=n}^{n+p} \mu_k (u, u_k)\, u_k(x) = \Gamma\left(\sum_{k=n}^{n+p} (u, u_k)\, u_k(x)\right),$$

where n and p are arbitrary positive integers. Referring to inequality 12.18 above and mindful of the orthonormality of the family of eigenfunctions, we obtain from this the important inequality

$$\left|\sum_{k=n}^{n+p} \mu_k (u, u_k) u_k (x)\right| \leq M\sqrt{b-a}\left\{\sum_{k=n}^{n+p} (u, u_k)^2\right\}^{1/2}.$$

The infinite series $\sum (u, u_k)^2$ converges by virtue of Bessel's inequality, and therefore the right-hand side can be made arbitrarily small by choosing n sufficiently large; since this n is independent of x we infer that the series $\sum \mu_k (u, u_k) u_k (x)$ converges uniformly. Denote the limit function by v; it is necessarily continuous since it is the uniform limit of continuous functions. It must in fact be equal to Γu:

$$\|\Gamma u - v\| \leq \|\Gamma u - \sum_{k=0}^{m} \mu_k (u, u_k) u_k\| + \|v - \sum_{k=0}^{m} \mu_k (u, u_k) u_k\|,$$

and each term on the right tends to zero as $m \to \infty$. This shows that $\|\Gamma u - v\|$ vanishes and, as both Γu and v are continuous, $\Gamma u - v = 0$.

Finally, suppose $f = \Gamma u$. Then

$$(f, u_k) = (\Gamma u, u_k) = (u, \Gamma u_k) = \mu_k (u, u_k).$$

This shows that series (12.1) with coefficients $c_k = (f, u_k)$ converges uniformly to f on $[a, b]$. □

Corollary 12.5.2 *(Parseval's Formula) Under the conditions of the Theorem,*

$$\sum_{k=0}^{\infty} c_k^2 = \|f\|^2.$$

Proof: Multiply each side of equation (12.1) by $f(x)$ and integrate. The uniformity of the convergence permits term-by-term integration, and the conclusion follows from this. □

A system of functions $\{u_k\}$ for which Parseval's formula holds is said to be *complete*.

12.6 Relaxation of Assumptions

In this section we reconsider some of the assumptions employed above.

12.6.1 The Weight Function

The Sturm-Liouville equation (10.11) includes a function ρ, assumed positive but not necessarily constant, whereas in §12.2 we introduced the simplifying assumption that $\rho = 1$. We now wish to relax that assumption. The eigenvalue problem consisting of equation (10.11) together with separated endpoint conditions (10.20) is converted via the Green's function $G(x,\xi)$ to the integral equation

$$u(x) = \lambda(\Gamma\rho u)(x) = \lambda\int_a^b G(x,\xi)\,\rho(\xi)\,u(\xi)\,d\xi, \tag{12.25}$$

where G is the Green's function for the problem of solving the equation $Lu = r$ with the same boundary conditions. The eigenfunctions are orthogonal with respect to the weight function, as in equation (12.4) above; we'll write the integral appearing there as $(u_m, \rho u_n)$. They may be assumed normalized: $(u_n, \rho u_n) = 1$. The eigenfunction expansion problem may be formulated as follows: what conditions need to be imposed on the function $g(x)$ on $[a,b]$ in order that the eigenfunction expansion

$$\sum_{k=0}^{\infty} a_k u_k(x) \text{ with } a_k = (u_k, \rho g)$$

converge to g?

This question can be addressed in the following way. Rewrite the eigenvalue problem in the form

$$\mu_k u_k(x) = \int_a^b G(x,\xi)\,\rho(\xi)\,u(\xi)\,d\xi,$$

with $\mu_k = \lambda_k^{-1}$ (we continue to assume that $\lambda = 0$ is not an eigenvalue). Now set $v_k(x) = \rho(x)^{1/2}\,u_k(x)$. Then the eigenvalue problem takes the form

$$\mu_k v_k(x) = \int_a^b H(x,\xi)\,v_k(\xi)\,d\xi,$$

where

$$H(x,\xi) = (\rho(x)\,\rho(\xi))^{1/2}\,G(x,\xi).$$

This function H is symmetric in x and ξ and the eigenvalue problem for $\{v_k\}$ is therefore one for a bounded, self-adjoint operator. Assuming that ρ

is continuous on $[a, b]$, we find from Theorem 12.5.3 that any function f on $[a, b]$ that may be written in the form

$$f(x) = \int_a^b H(x, \xi)\, v(\xi)\, d\xi, \quad v \in C^2[a, b],$$

has the uniformly convergent expansion

$$f(x) = \sum_{k=0}^{\infty} c_k v_k(x), \ c_k = (v_k, f);$$

we assume that $(v_k, v_k) = 1$.

This expansion may be written in terms of the sequence $\{u_k\}$:

$$f(x) = \rho(x)^{1/2} \sum_{k=0}^{\infty} c_k u_k(x).$$

The condition on the function f implies that $f(x) = \rho(x)^{1/2} g(x)$ where

$$g(x) = \int_a^b G(x, \xi)\, \rho(\xi)^{1/2}\, v(\xi)\, d\xi.$$

The expansion for f now becomes the expansion for g given by

$$g(x) \ \sum_{k=0}^{\infty} c_k u_k(x), \ c_k = (u_k, \rho g).$$

Theorem 12.6.1 *Suppose ρ is continuous on $[a, b]$ and positive there. Then the eigenfunction expansion converges uniformly for any function g that is C^2 on (a, b) and satisfies the boundary conditions.*

This generalizes Theorem 12.5.3.

12.6.2 Zero Eigenvalue

In §12.4 above we introduced the assumption that zero is not an eigenvalue of the boundary-value problem consisting of the differential equation together with the assigned, homogeneous boundary conditions. This assumption can be replaced by the assumption:

There exists a real number λ_0 which is not an eigenvalue.

It is obvious that this is true for the Sturm-Liouville problem; it is true for a more general class of boundary-value problems as well.

The equation $Lu = \lambda u$ is equivalent to the equation $\hat{L}u = \hat{\lambda}u$ if $\hat{\lambda} = \lambda - \lambda_0$ and the operator $\hat{L}$ is defined by the formula $\hat{L}u = Lu - \lambda_0 u$. The operator $\hat{L}$ is easily seen to be self-adjoint, and $\hat{\lambda} = 0$ cannot be an eigenvalue because, if it were, λ_0 would be an eigenvalue of the original problem. This problem therefore possesses a symmetric Green's function $\hat{G}(x, \xi)$ converting the boundary-value problem into the integral equation $\hat{\Gamma}u = \hat{\mu}u$ for the bounded, self-adjoint operator $\hat{\Gamma}$. The theory is now applied to this operator. The sequence of eigenfunctions is the same as for the original problem, and Theorem 12.5.3 therefore holds.

12.7 Singular boundary-value problems

There are various respects in which the boundary-value problems that we have considered might be generalized so as to be viewed as singular, but we shall consider only those two that are the most frequently encountered in applications: either the positive function p vanishes at one or both endpoints, or the interval is unbounded. In fact, although we will not pursue a general treatment of either of these problems, it is essentially the same for these two notions of singular behavior. As usual, we consider the operator

$$Lu \equiv -\left(p(x)\, u'\right)' + q(x)\, u = \lambda \rho(x)\, u, \tag{12.26}$$

but now allowing either for p to vanish at a finite endpoint or for one or the other of the endpoints to be at infinity; we also reserve judgment on the nature of the boundary conditions. We begin with some examples.

Example 12.7.1 *Legendre's equation* The equation is

$$Lu = -\left(\left(1 - x^2\right) u'\right)' = \lambda u,$$

and the interval is $(-1, 1)$. The coefficient $p = 1 - x^2$ vanishes at both endpoints. On the usual (L^2) inner product, integrating by parts, one finds

$$(v, Lu) = -\left(1 - x^2\right) vu'|_{-1}^{1} + \int_{-1}^{1} \left(1 - x^2\right) u'v'\, dx.$$

If the integrated part vanishes, the inner product is symmetric in u and v and the problem is of the self-adjoint kind. This will be the case if the functions u and v are required to be bounded, with bounded derivatives on $[-1,1]$. Thus for this problem to be self-adjoint, the algebraic boundary conditions of the form (??) are replaced by these conditions of boundedness.

We can find power-series solutions $u = \sum_0^\infty a_k x^k$ of the differential equation that are convergent for $|x| < 1$. The coefficients $\{a_k\}$ satisfy the two-term recursion relation

$$a_{k+2} = \frac{k\,(k+1) - \lambda}{(k+1)\,(k+2)} a_k. \tag{12.27}$$

This shows that there are even solutions and odd solutions and also that, if $\lambda = n\,(n+1)$, the series terminates and there is a solution of the differential equation in the form of a polynomial $p_n\,(x)$ of degree n (even or odd according as n is even or odd). These *Legendre polynomials* are eigenfunctions, with eigenvalues $\lambda_n = n\,(n+1)$. The eigenfunctions necessarily possess the "Sturm-Liouville" property that p_n has precisely n zeros in the interval $(-1,1)$. In the present context this can be derived from the corresponding theorem regarding orthogonal polynomials, since it is clear that $(p_n, p_m) = 0$ if $n \neq m$. $\square$