26.10.210

(12)

VIII Compact operators

VIII.1 Definitions & Basic Properties

The class of compact operators

i) Appear quite often in applications
ii) resemble ordinary matrices.

In particular, they do $\overbrace{\text{have}}^{\text{usually}}$ eigenvalues and eigenvectors and a discrete spectrum (only countably many points in the spectrum).

8.1 Defn. An operator $A \in \mathcal{L}(E;F)$ (E and F are normed spaces, and F typically Banach) is compact if for every bounded sequence $\{x_n\}_{n=1}^{\infty}$ in E, the sequence $\{Ax_n\}_{n=1}^{\infty}$ contains a convergent subsequence in F.

(Note: The limit point of this convergent sequence need not belong to the range of A).

8.2a Ex Every bounded operator with finite rank is compact. Finite rank means that the range is finite-dimensional.

Proof: Let $A \in \mathcal{L}(E;F)$ have finite rank. Then range(A) is finite dimensional. Call range$(A) = V$. Then we actually have $A \in \mathcal{L}(E;V)$.
Take any bounded $\{x_n\}_{n=1}^{\infty}$. Then $\{Ax_n\}_{n=1}^{\infty}$ is a bounded sequence in V. Since V is finite-dimensional it is isomorphic to $\mathbb{C}^m$ (or $\mathbb{R}^m$), and every bounded sequence contains a convergent subsequence. Thus, $\{Ax_n\}_{n=1}^{\infty}$ contains a bounded subsequence, so A is compact. □

8.2b Ex. The identity operator in E (= Hilbert space) is compact $\Longleftrightarrow$ $\dim(E) < \infty$.

Proof: $\dim(E) < \infty \Rightarrow I$ is compact by Ex. 8.2.a.
If $\dim(E) = \infty$, then we can find an orthonormal sequence $\{e_n\}_{n=1}^{\infty}$ in E with infinitely many e_n. Then $\{e_n\}_{n=1}^{\infty}$ is a bounded sequence in E, but for $n \neq m$, $\|e_n - e_m\|^2 = \|e_n\|^2 + \|e_m\|^2 = 2$ (by Pythagoras). Thus, $\{e_n\}_{n=1}^{\infty}$ contains no convergent subsequence.

8.3c Ex. If $\overbrace{\text{the operator } D \text{ in homework 25 is compact, then}}^{\lambda_n \text{ does not} \to 0 \text{ as } n \to \infty \text{, then}}$ $\lambda_n \to 0$ as $n \to \infty$.

Proof: If " ", then we can find some subsequence $\{\lambda_{n_k}\}_{k=1}^{\infty}$ with $|\lambda_{n_k}| \geq \varepsilon > 0$ for all k. The sequence $\{e_{n_k}\}$ is bounded, and $De_{n_k} = \lambda_{n_k} e_{n_k}$, so

$$\|De_{n_k} - De_{n_l}\|^2 = \|\lambda_{n_k} e_{n_k} - \lambda_{n_l} e_{n_l}\|^2$$

$$\text{(by Pythagoras)} = |\lambda_{n_k}|^2 + |\lambda_{n_l}|^2 \geq 2\varepsilon^2 > 0.$$

As above, this sequence contains no converging subsequence.

8.3d Ex. If $\lambda_n \to 0$ as $n \to \infty$ then the operator D in homework 14 is compact.

Thus: $\boxed{D \text{ compact} \Longleftrightarrow \lambda_n \to 0 \text{ as } n \to \infty.}$

Proof later. It is based on the foll. thm:

8.3 Thm. Let E, F be Banach spaces. The set of compact operators in $\mathcal{L}(E;F)$ is a closed subset of $\mathcal{L}(E;F)$.

Thus, if A_n are compact, $\|A - A_n\| \to 0$ as $n\to\infty$, then A is compact.

Proof. By a diagonalization argument. See the book.

8.2d Example continues. Let D be the operator in homework 25. Put

$$D_N x = \sum_{n=1}^{N} \lambda_n (x, e_n) e_n.$$

Then the range of D_N lies in the linear span of $\{e_1, e_2, \dots, e_N\}$, so it is finite-dimensional, and D_N is compact.

$$\text{Now } (D - D_N)(x) = \sum_{n=N+1}^{\infty} \lambda_n (x, e_n) e_n.$$

By Homework 25, $\|D - D_N\| = \sup_{n \ge N+1} |\lambda_n|$

$\to 0$ as $N \to \infty$. By Thm 8.3,

D is compact. $\square$

8.4 Lemma If $A, B \in \mathcal{L}(E;F)$, and if both A and B are compact, then $A+B$ is compact.

Proof: Let $\{x_n\}_{n=1}^{\infty}$ be bounded. Take some subsequence so that $\{A x_{n_k}\}$ converges. Take a subsequence of this subseq. so that also $\{B x_{n_{k_\ell}}\}$ converges. Then $(A+B) x_{n_{k_\ell}}$ converges. This means that $A+B$ is compact. $\square$

VIII.2 Hilbert-Schmidt operators

Problem: It is often difficult to determine if an operator is compact.

One solution: Try to show that the operator is Hilbert-Schmidt, and use Thm 8.7.

8.5 Defn. $A \in \mathcal{L}(E;F)$ is a Hilbert-Schmidt operator (here E & F are Hilbert spaces) if there exists some orthonormal basis $\{e_n\}_{n=1}^{\infty}$ for E such that

$$\sum_{n=1}^{\infty} \|Ae_n\|^2 < \infty.$$

Note: E is necessarily separable. We do allow E to be finite dimensional (then the sum is finite).

Note: E finite-dimensional $\Rightarrow$ every $A \in \mathcal{L}(E;F)$ is Hilbert-Schmidt. Proof obvious (finite sum).

8.5a Lemma (HS = Hilbert-Schmidt). Let $A \in \mathcal{L}(E;F)$ be a HS-oper. If $\{e_n\}_{n=1}^{\infty}$ is a complete orthonormal sequence in E and $\{f_m\}_{m=1}^{\infty}$ is a complete orthonormal sequence in F, then

$$\sum_{n=1}^{\infty} \|Ae_n\|^2 = \sum_{m=1}^{\infty} \|A^* f_m\|^2.$$

In particular, if $\{e_n'\}_{n=1}^{\infty}$ is another orthonormal basis in E, then

$$\sum_{n=1}^{\infty} \|Ae_n\|^2 = \sum_{n=1}^{\infty} \|Ae_n'\|^2$$

Proof: By Thm 4.14, for every $\phi \in F$ we have $\|\phi\|^2 = \sum_{m=1}^{\infty} |(\phi, f_m)|^2$. Take $\phi = Ae_n$. Then

$$\|Ae_n\|^2 = \sum_{m=1}^{\infty} |(Ae_n, f_m)|^2 = \sum_{m=1}^{\infty} |(e_n, A^* f_m)|^2.$$

Thus $\sum_{n=1}^{\infty} \|Ae_n\|^2 = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |(e_n, A^* f_m)|^2$

$$= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |(A^* f_m, e_n)|^2 = \sum_{m=1}^{\infty} \|A^* f_m\|^2,$$

since it is true for every $x \in E$ that

$$\|x\|^2 = \sum_{n=1}^{\infty} |(x, e_n)|^2 \quad (\text{take } x = A^* f_m). \quad \square$$

8.5b Defn. We call $\left(\sum_{n=1}^{\infty} \|Ae_n\|^2\right)^{1/2}$ the Hilbert-Schmidt-norm of A. (By Lemma 8.5a, this number does not depend on how we choose the orthonormal basis $\{e_n\}_{n=1}^{\infty}$ of E).

8.5c Cor. A is a Hilbert-Schmidt operator $\Longleftrightarrow$ A^* is a Hilbert Schmidt operator.

Proof: By Lemma 8.5a. $\square$

8.5d lemma Let $A \in L(E; F)$ be a HS-operator, let $\{e_n\}$ be a complete orthonormal basis in E, and let $\{\phi_n\}$ be a complete orthonormal basis in F. Let $A_{nm} = (Ae_n, \phi_m)$ be the matrix representation of A (see p. 78), and let $A^*_{mn} = (A^* \phi_m, e_n)$ be the matrix representation of A^*. Then $A_{nm} = \overline{A^*_{mn}}$ (= complex conjugation, and change of indeces), and the Hilbert-Smidt norm of A is

$$\|A\|_{HS}^2 = \sum_{m,n} |A_{n,m}|^2 = \sum_{m,n} |A^*_{mn}|^2 = \|A^*\|_{HS}^2.$$

Proof. We continue the proof of Lemma 8.5a: On the bottom of p. 84 we had

$$\|Ae_n\|^2 = \sum_{m=1}^{\infty} |(Ae_n, \phi_m)|^2 = \sum_{m=1}^{\infty} |A_{nm}|^2$$

Add over n to get

$$\|A\|_{HS}^2 = \sum_{n=1}^{\infty} \|Ae_n\|^2 = \sum_{m,n} |(Ae_n, \phi_m)|^2 = \sum_{m,n} |A_{n,m}|^2.$$

From the top of p. 85 we see that this is equal to

$$\sum_{m=1}^{\infty} \|A^* f_m\|^2 = \sum_{m,n} |(A^* f_m, e_n)|^2 = \sum_{m,n} |A^*_{m,n}|^2. \quad \square$$

Lemma For every Hilbert-Schmidt operator A, its HS-norm $\|A\|_{HS}$ and operator norm $\|A\|_{op}$ satisfy

$$\|A\|_{op} \le \|A\|_{HS}$$

Proof: Expand $x \in E$ into $x = \sum_{n=1}^{\infty} (x, e_n) e_n$, and also expand $Ax \in F$ into

$$Ax = \sum_{m=1}^{\infty} (Ax, f_m) \phi_m.$$

(Here $\{e_n\}$ is a basis for E, $\{\phi_m\}$ a basis for F)

Then $\|Ax\|^2 = \sum_{m=1}^{\infty} |(Ax, f_m)|^2$

$$= \sum_{m=1}^{\infty} |(x, A^* \phi_m)|^2 \quad (\text{by Cauchy-Schw})$$

$$\le \sum_{m=1}^{\infty} \|x\|^2 \|A^* f_m\|^2$$

$$= \|x\|^2 \sum_{m=1}^{\infty} \|A^* f_m\|^2$$

$$= \|x\|^2 \text{ times } \|A\|_{HS}$$

(See Lemma 8.5a)

Alternative proof: $\|Ax\| = \|A\sum_{n=1}^{\infty}(x,e_n)e_n\|$

$= \|\sum_{n=1}^{\infty}(x,e_n)Ae_n\|$ (triangle-ineq)

$\leq \sum_{n=1}^{\infty}|(x,e_n)|\,\|Ae_n\|$ (Cauchy-Schwartz in ℓ^2)

$\leq \left(\sum_{n=1}^{\infty}|(x,e_n)|^2\right)^{1/2}\left(\sum_{n=1}^{\infty}\|Ae_n\|^2\right)^{1/2}$

$= \|x\|$ times HS-norm of A. $\square$

Note: The Hilbert-Smidt norm is actually induced by a Hilbert space inner product: IO

$A \to A_{ij}$ is the matrix repr. of A

$B \to B_{ij}$ is the matrix repr. of B,

then we can define the inner product

$$(A,B)_{HS} = \sum_{ij} A_{ij}\overline{B_{ij}}.$$

This is related to Homework 1.

Note: The standard operator norm defined on p. 53 is never induced by an inner product (except when both E and F have dimension $=1$).

27.10.2010 (13)

8.7. Thm Every Hilbert-Schmidt operator $A \in \mathcal{L}(E;F)$ is compact.

Proof: Choose some orthonormal basis $\{e_i\}_{i=1}^{\infty}$ in E. Then for each $x \in E$ we have $x = \sum_{i=1}^{\infty}(x,e_i)e_i$, and since $A \in \mathcal{L}(E;F)$, we get

$$Ax = A\sum_{i=1}^{\infty}(x,e_i)e_i = \sum_{i=1}^{\infty}(x,e_i)Ae_i$$
$$= \lim_{k\to\infty}\sum_{i=1}^{k}(x,e_i)Ae_i.$$

Define $A_k x = \sum_{i=1}^{k}(x,e_i)Ae_i$. Each A_k is a linear operator with finite rank, and for each $x \in E$ we have $A_k x \to Ax$ as $k\to\infty$.

Claim: $\|A_k - A\| \to 0$ as $k\to\infty$.

Proof:

$$(A_k - A)(x) = \sum_{i=k+1}^{\infty}(x,e_i)Ae_i, \text{ so}$$

$$\|(A_k-A)x\|_F^2 = \|\sum_{i=k+1}^{\infty}(x,e_i)Ae_i\|_F^2 \quad \text{(use } \Delta\text{-ineq)}$$

$$\leq \left(\sum_{i=k+1}^{\infty}|(x,e_i)|\,\|Ae_i\|_F\right)^2 \quad \text{(use Cauchy-Sch)}$$

$$\leq \sum_{i=k+1}^{\infty}|(x,e_i)|^2\sum_{i=k+1}^{\infty}\|Ae_i\|_F^2 \quad \text{(use Bessel's ineq)}$$

$$\leq \|x\|^2\sum_{i=k+1}^{\infty}\|Ae_i\|_F^2, \text{ so}$$

$$\|A_k - A\|^2 = \sup_{\|x\|\leq 1}\|(A_k-A)x\|^2 \leq \sum_{i=k+1}^{\infty}\|Ae_i\|_F^2$$
$$\to 0 \text{ as } k\to\infty.$$

By Theorem 8.3, A is compact (since each A_k is compact, and $\|A_k - A\| \to 0$ as $k \to \infty$). $\square$

8.8 Thm Let $k : (c,d) \times (a,b) \to \mathbb{C}$ be a Lebesgue measurable function (for example, continuous), and suppose that

$$M^2 := \int_{t=c}^{d} \int_{s=a}^{b} |k(t,s)|^2 \, ds\,dt < \infty.$$

Then the integral operator

$$(Kx)(t) = \int_a^b k(t,s)\, x(s)\, ds$$

is a Hilbert–Schmidt operator from $L^2([a,b])$ to $L^2([c,d])$. In particular, it is compact. The HS-norm of this operator is $\|K\|_{HS} = M$.

Proof. See the book, p. 94.

Note: Not every compact operator is HS. Counterexample: The operator D in homework 25 with $\lambda_n = 1/\sqrt{n}$. Here

$$\sum_{n=1}^{\infty} \|De_n\|^2 = \sum_{n=1}^{\infty} |\lambda_n|^2 = \sum_{n=1}^{\infty} \frac{1}{n} = \infty,$$

but by Ex 8.3d, D is compact.

An $n \times n$ matrix has at most n and at least one eigenvector. For example, the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

has only one eigenvector, namely $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, and one eigenvalue namely $\lambda = 0$. Thus, in this case we cannot get any information about $\|A\|$ by studying eigenvectors alone.

However, the situation is different if A is selfadjoint, both in the matrix case, and in the case of a compact operator.

8.10 Thm. Let K be a compact selfadjoint operator on the Hilbert space H. Then either $\|K\|$ or $-\|K\|$ is an eigenvalue of K.

(In particular, this eigenvalue is real.) This theorem is not true for non-selfadj. operators. See the matrix example above.

Proof: See the book, p. 96. (by self-adjointness)

Idea: We know from Thm 7.18 that $\|K\|$ is the maximum of the two numbers $\sup_{\|x\| \le 1} (Kx, x)$ and $-\min_{\|x\| \le 1} (Kx, x)$.

(see page 70). For example, suppose that $\|A\| = \sup_{\|x\| \le 1} (Kx, x)$. Call this number λ.

Take some sequence $x_n \in H$, $\|x_n\| = 1$, so that $(Kx_n, x_n) \to \lambda$ as $n \to \infty$.

By the compactness there is some subsequence x_{n_k} so that $Kx_{n_k} \to y \in H$. Use a "simple estimate" to show that also x_{n_k} converges. The limit will be an eigenvector x with eigenvalue λ. $\square$

The analogy between self-adjoint matrices and self-adjoint compact operators goes even further:

8.11 Thm. The eigenvalues of a self-adjoint operator are real, and eigenvectors corresponding to different eigenvalues are orthogonal.

(Familiar from matrix theory??)

Proof: Let λ be an eigenvalue, and $\varphi \neq 0$ an eigenvector. Then

$$\begin{aligned} 0 &= (A\varphi, \varphi) - (\varphi, A\varphi) \quad \text{(since self-adj.)} \\ &= (\lambda\varphi, \varphi) - (\varphi, \lambda\varphi) \\ &= (\lambda - \bar\lambda)(\varphi, \varphi). \end{aligned}$$

Thus, $\lambda - \bar\lambda = 0$, and λ is real.

Now take two different eigenvalues λ and μ with eigenvectors φ and ψ. Then

$$\begin{aligned} 0 &= (A\varphi, \psi) - (\varphi, A\psi) \\ &= (\lambda\varphi, \psi) - (\varphi, \mu\psi). \\ &= (\lambda - \bar\mu)(\varphi, \psi). \end{aligned}$$

Here $\mu = \bar\mu \neq \lambda$, so $(\varphi, \psi) = 0$. $\square$

Remark: Even more is true: The whole spectrum of a self-adjoint operator is real, that is, if $\operatorname{Im} \lambda \neq 0$ and A is self-adjoint, then $(\lambda I - A)$ is invertible.

Our next thm generalizes Ex. 8.3c:

8.12 Thm. Let K be compact and self adjoint. Then
- either K has only finitely many eigenvalues, or
- K has infinitely many eigenvalues λ_n, where $\boxed{\lambda_n \to 0 \text{ as } n \to \infty}$.

Proof: The same as the proof of Ex. 8.3c (since the eigenvectors corresponding to different eigenvalues are orthogonal). See the book for details.

Our next thm needs a preliminary lemma:

8.14 lemma. Let M be a subspace of a Hilbert space H, let $T \in \mathcal{L}(H)$, and suppose that M is invariant under T, i.e., T maps M into M. Then $M^\perp$ is invariant under T^*, i.e., T^* maps $M^\perp$ into itself.

Proof. Take $x \in M^\perp$ and $m \in M$. Then $Tm \in M$, so $Tm \perp x$, i.e.,

$$0 = (Tm, x) = (m, T^*x).$$

This is true for all $m \in M$, so $T^*x \in M^\perp$ $\square$

The following theorem says: The spectral theory for compact self adjoint operators is "the same" as for self-adjoint matrices.

8.15 Spectral thm for compact self-adj. oper. Let K be a compact self-adjoint operator on a Hilbert space H. Then K has a (finite or infinite) sequence of orthonormal eigenvectors φ_n with corresponding real eigenvalues λ_n such that K is of the type described in homework 25, i.e.

$$(*) \quad Kx = \sum_n \lambda_n (x, \varphi_n)\varphi_n$$

The sequence is finite iff K is degenerate, and otherwise $\lambda_n \to 0$ as $n \to \infty$.

Note: It is possible that $\lambda = 0$ is an eigenvalue but there is no need to include this term in ⊛ since it gives a zero. We do not claim that the sequence φ_n is complete. For example the operator $K = 0$ is compact and self-adjoint, and in this case we need no vectors in the sum. Another example is $Kx = (\varphi, x)\varphi$ where $\|\varphi\| = 1$ is an arbitrary vector. Then $n = 1$.

Note: If we take all $\lambda_n \neq 0$ in ⊛ (by dropping eigenvectors with eigenvalue zero), then the orthogonal complement to the span of $\{\varphi_n\}$ is the null space of K.

Proof. See the book for the details.
Idea: We know that K has an eigenvector φ_0 with eigenvalue $\lambda_0 = \pm\|K\|$. (see Thm. 8.10). Scale φ_0 so that $\|\varphi_0\| = 1$ (i.e., divide φ_0 by $\|\varphi_0\|$ to make the norm one).

Subtract off the term $\lambda_0 (x, \varphi_0)\varphi_0$ from K, i.e., study the operator

$$K_1 x = Kx - \lambda_0 (x, \varphi_0)\varphi_0 .$$

Clearly $K_1 \varphi_0 = 0$. If we project an arbitrary $x \in H$ onto φ_0, i.e., we write

$$x = x_0 + x_1 , \quad \text{where} \quad x_0 = (x, \varphi_0)\varphi_0, \quad x_1 \perp \varphi_0,$$

then

$$Kx = K(x_0 + x_1) = \lambda_0 x_0 + K_1 x_1 .$$

Now we repeat the argument from the beginning, replacing the original space H by the orthogonal complement to φ_0, and replacing K by K_1.
(For every x orthogonal to φ_0 we have $K_1 x = Kx$). □

8.16 Coroll. Let K be a compact self-adjoint operator on a separable Hilbert space H. Then it is possible to find a complete set of orthonormal eigenvectors $\{\varphi_n\}_{n=1}^{\infty}$ of K so that

$$⊛ \quad Kx = \sum_{n=1}^{\infty} \lambda_n (x, \varphi_n)\varphi_n ,$$

where λ_n is the eigenvalue corresponding to φ_n.

Proof (outline): The representation that the construction in the proof of Thm 8.15 produces has $\lambda_n \neq 0$ for all n (because we can stop when the remaining operator K_{n+1} is zero). The orthogonal complement is the null space of K, so every nonzero vector in this space is an eigenvector with eigenvalue zero ($Kx = 0 = 0 \cdot x$). Choose an arbitrary orthonormal basis for this subset, and add this to the collection of all eigenvectors with nonzero eigenvalues. This makes the collection complete at the expense of adding a lot of zeros to the sum in ⊛. □

28.10.2010

(14)

IX. Sturm-Liouville Problems

IX.1 Definitions

A Sturm-Liouville system is a second order differential equation with two boundary conditions (one at each end of the interval in which we solve the equation). These systems appear in many physical problems, where one solves partial differential equations by separating the variables.

9.1 Defn. A regular Sturm-Liouville system is a differential equation on an interval $[a,b]$ together with two boundary conditions:

$$\text{RSL}: \quad \frac{d}{dx}\left(p\frac{d\phi}{dx}\right) + (\lambda\rho + q)\phi = 0, \qquad (1)$$

$$\left.\begin{aligned} \alpha_0\phi(a) + \alpha_1\phi'(a) &= 0, \\ \beta_0\phi(b) + \beta_1\phi'(b) &= 0. \end{aligned}\right\} \qquad (2)$$

Here all functions and numbers are real (except that ϕ is allowed to be complex), and

- ρ, q are continuous on $[a,b]$
- p is continuously differentiable on $[a,b]$
- ϕ (= the solution) is two times cont. diff. on $[a,b]$.
- $p(x) > 0$ and $\rho(x) > 0$ for $x \in [a,b]$.
- we do not allow both $\alpha_0 = 0$ and $\alpha_1 = 0$
- we do not allow both $\beta_0 = 0$ and $\beta_1 = 0$.

If we differentiate (1) and divide by p we get

$$p(x)\phi''(x) + p'(x)\phi'(x) + [\lambda\rho(x) + q(x)]\phi(x) = 0, \quad a \le x \le b,$$

$$\Rightarrow \quad \phi''(x) + \frac{p'(x)}{p(x)}\phi'(x) + \frac{\lambda\rho(x) + q(x)}{p(x)}\phi(x) = 0, \quad a \le x \le b.$$

We cannot divide by $p(x)$ unless $p(x) \neq 0$ for all $x \in [a,b]$, and this is why we require $p(x) > 0$ for $x \in [a,b]$. In a nonregular problem we relax this requirement:

9.2 Defn. A singular Sturm-Liouville problem is of the same type as the regular problem in Defn 9.1, with the following exceptions:

- we require $p(x) > 0$ for $x \in (a,b)$, but we do allow $p(a) = 0$ or $p(b) = 0$
- the differential equation need not be satisfied at an end point where $p(x) = 0$.
- we drop the boundary condition (2) at an end point where $p(x) = 0$, and replace it by the condition that ϕ is bounded (close to this point).

9.2b Ex.

$$\frac{d}{dx}\left(x\frac{d\phi}{dx}\right) + \lambda\phi(x) = 0, \quad 0 \le x \le L$$
$$\phi(L) = 0$$
$$\phi \text{ is bounded.}$$

Here $[a,b] = [0,L]$, $p(x) = x$, $p(0) = 0$, $q(x) = 0$, $\rho(x) = 1$, α_0 and α_1 are not needed, $\beta_0 = 1$, $\beta_1 = 0$. This is a singular problem which pops up when one tries to solve the equations of motions of a hanging chain. (pp. 105–110 in the book).

Why such a very special form of (1)? Why not the more general

$$\phi''(x) + P(x)\phi'(x) + (\lambda R(x) + Q(x))\phi(x) = g(x)$$

instead?

Answer 1: Later we shall allow $g(x) \neq 0$.

Answer 2: It is possible to rewrite (*) so that it resembles (1): Multiply the equation by ("variation of constants"):

$$p(x) = e^{\int_a^x P(s)\,ds} \quad \Rightarrow$$

$$p(x)\phi''(x) + P(x)p(x)\phi'(x) + p(x)(\lambda R(x) + Q(x))\phi(x) = p(x)g(x)$$

$$\Longleftrightarrow \quad \frac{d}{dx}\left[p(x)\phi'(x)\right] + (\lambda\rho(x) + q(x))\phi(x) = p(x)g(x),$$

where $\rho(x) = p(x)R(x)$, $q(x) = p(x)Q(x)$. This is of the form (1) with a nonzero right hand side.

IX.2 Eigenvalues and eigenfunctions

We shall introduce eigenvalues and eigenfunctions of the Sturm-Liouville problem. These are connected to a particular "differential operator" L, defined as follows.

We let $\mathcal{D}(L)$ = "domain of L" consist of all "two times differentiable functions" whose second derivative is in $L^2(0,1)$, and which satisfy the boundary conditions of either defn. 9.1 or 9.2, depending on whether the problem is regular or singular. This can be shown to be a Hilbert space with inner product

$$(\phi, g) = \int_0^1 \phi(x)\bar{g}(x)\,dx + \int_0^1 \phi''(x)\overline{g''(x)}\,dx.$$

On this space we define two operators:

$$(*)\begin{cases} L\phi = \frac{d}{dx}(p\phi') + q\phi & \text{(combined multiplication and differentiation)} \\ J\phi = \rho(x)\phi(x) & \text{(multiplication).} \end{cases}$$

Both of these map $\mathcal{D}(L)$ continuously into $L^2(a,b)$. We can write the Sturm-Liouville problem (1) as

$$\boxed{\underset{\text{S-L-equation}}{(\lambda J + L)\phi = 0}, \quad \underbrace{\phi \in \mathcal{D}(L)}_{\text{smoothness and boundary conditions}}}$$

This motivates us to define eigenvalues and eigenfunctions of the S-L-problem as follows:

9.3 Defn. An eigenfunction ϕ in $\mathcal{D}(L)$ is a nonzero solution of the (regular or singular) Sturm-Liouville problem. (These only exist for some values of the parameter λ). The corresponding value of λ is called an eigenvalue.

Note: This function can be interpreted as a vector in the space $L^2(0,1)$, so we could also call it an eigenvector.

9.4 Ex. $[a,b] = [0,\pi]$,

$$\begin{cases} \phi''(x) + \lambda\phi(x) = 0, & 0 \le x \le \pi \\ \phi(0) = \phi(\pi) = 0. \end{cases}$$

We shall see later (Thm 9.8) that all eigenvalues must be real, so we study only real λ.

Case A: $\lambda = 0$: $\Rightarrow \phi'' = 0 \Rightarrow \phi(x) = Ax + B$, where A and B are constants.

$\phi(0) = 0 \Rightarrow B = 0$

$\phi(\pi) = 0 \Rightarrow A\pi + B = 0 \Rightarrow A = 0$

$\Rightarrow$ $\lambda = 0$ is not an eigenvalue

Case B: $\lambda < 0$, so that $\lambda = -|\lambda|$. By the course in differential eqs: the general solution is $\phi(x) = Ae^{\mu x} + Be^{-\mu x}$, where $\mu = +\sqrt{|\lambda|}$.

$\phi(0) = 0 \Rightarrow A + B = 0 \Rightarrow B = -A$

$\phi(\pi) = 0 \quad Ae^{\mu\pi} - Ae^{-\mu\pi} = 0 \Rightarrow A(e^{\mu\pi} - e^{-\mu\pi}) = 0$

$\Rightarrow A = B = 0 \Rightarrow$

$\lambda < 0$ is not an eigenvalue.

Case C: $\lambda > 0$. This time the general solution is

$\phi(x) = A\cos(\mu x) + B\sin(\mu x)$, $\mu = \sqrt{\lambda}$

$\phi(0) = 0 \Rightarrow A = 0$

$\phi(\pi) = 0 = B\sin(\mu\pi) = 0$. We need $B \ne 0$ because otherwise $\phi(x) \equiv 0$. Therefore $\sin(\mu\pi) = 0 \Rightarrow$

$\mu =$ integer. $\mu = 0$ gives $y(x) \equiv 0$, so
$\mu \neq 0$. $\mu = \{\pm 1, \pm 2, \pm 3, \dots\}$

$$\Rightarrow \lambda = \mu^2 = \{1, 4, 9, 16, \dots\}$$

Thus, the set of all eigenvalues is $\boxed{\{n^2 \mid n \in \mathbb{N}\}}$.

Note: This is an unbounded sequence (tending to infinity), so they cannot be eigenvalues of a bounded linear operator.

The corresponding eigenfunctions are $\varphi_n(x) = \sin(nx)$. If we normalize the norm to be 1 we get:

$$\int^{\pi} \sin^2(nx)\,dx = \frac{\pi}{2},$$ i.e., the normalized functions are
$$\varphi_n(x) = \sqrt{\frac{2}{\pi}}\sin(nx)$$

We next prove some results on these eigenvalues and -functions. First one lemma. (Recall the definition of L given on p. 87).

9.6 Lagrange's identity For all $u, v \in \mathcal{D}(L)$,

$$uLv - vLu = \left(p(uv' - u'v)\right)'.$$

Proof: Use ⊛ on page 87, and the rule for the derivative of a product.

The operator maps a subset $\mathcal{D}(L)$ of $L^2(a,b)$ into $L^2(a,b)$, and we can use this fact to extend the definition of self-adjointness to L:

9.7 Thm. The operator L is self-adjoint in the sense that

$$⊛ \quad (Lu, v) = (u, Lv) \quad \text{for all } u, v \in \mathcal{D}(L).$$

Note: Here we use the standard inner product in L^2. The standard definition of self-adjointness would require $L \in \mathcal{B}(L^2(a,b))$, and would require ⊛ to hold for all $u, v \in L^2(a,b)$. Here we restrict u, v to the domain of L.

Proof: If $v \in \mathcal{D}(L)$, then $\bar{v} \in \mathcal{D}(L)$ (complex conjugate) because $\alpha_1, \alpha_2, \beta_1, \beta_2$ are real, hence also $\bar{v}$ satisfies the boundary conditions. Since p, q, r are real, $\overline{Lv} = L\bar{v}$, and

$$(Lu, v) - (u, Lv) = \int_a^b \bar{v}Lu - uL\bar{v}\,dx \quad (\text{by } 9.6)$$

$$= \Big[p(u\bar{v}' - u'\bar{v})\Big]_a^b. \quad ⊛$$

Singular case: This vanishes at the right end point if $p(b) = 0$ and at the lower bound if $p(a) = 0$.

Regular case: At the left end point we have $\alpha_1 u(a) + \alpha_2 u'(a) = 0$ and $\alpha_1 \bar{v}(a) + \alpha_2 \bar{v}'(a) = 0$. If $\alpha_1 = 0$, then $\alpha_2 \neq 0$, and $u'(a) = \bar{v}'(a) = 0$ so $u(a)\bar{v}'(a) - u'(a)\bar{v}(a) = 0$. If $\alpha_1 \neq 0$, then $u(a) = -\frac{\alpha_2}{\alpha_1}u'(a)$ and $\bar{v}(a) = -\frac{\alpha_2}{\alpha_1}\bar{v}'(a)$, and
$u(a)\bar{v}'(a) - u'(a)\bar{v}(a) = -\frac{\alpha_2}{\alpha_1}u'(a)\bar{v}'(a) + \frac{\alpha_2}{\alpha_1}u'(a)\bar{v}'(a) = 0.$

Thus, in all cases the substitution in ⊛ is zero, and $(Lu, v) = (u, Lv)$. $\square$

9.8 Thm. The eigenvalues of a Sturm–Liouville system are real.

Proof: Let ϕ be an eigenfunction with eigenvalue λ, i.e., $L\phi = -\lambda\rho\phi$. Then

$$0 = (L\phi,\phi) - (\phi,L\phi) = (-\lambda\rho\phi,\phi) - (\phi,-\lambda\rho\phi)$$
$$= (\bar\lambda - \lambda)\int_a^b \rho(x)|\phi(x)|^2\,dx. \quad \text{Thus, } \lambda = \bar\lambda. \quad \square$$

IX.3 Orthogonality of eigenfunctions

9.9 Thm Let u and v be two eigenfunctions of a S-L-system corresponding to *different* eigenfunctions. Then $\sqrt{\rho}\,u$ is orthogonal to $\sqrt{\rho}\,v$.

Proof: We have $Lu = -\lambda\rho u$, $Lv = -\mu\rho v$, $\lambda \neq \mu$, and

$$0 = (Lu,v) - (u,Lv) = (-\lambda\rho u, v) - (u, -\mu\rho v)$$
$$= (\mu-\lambda)\int_a^b \rho(x)\,u(x)\,\bar v(x)\,dx$$
$$= (\mu-\lambda)\int_a^b \sqrt{\rho(x)}\,u(x)\,\sqrt{\rho(x)}\,\bar v(x)\,dx, \quad \text{so}$$

$\sqrt{\rho}\,u \perp \sqrt{\rho}\,v$. $\square$

9.10 Thm. A S-L-system has only countably many eigenvalues.

Proof (outline): If there would be uncountably many different eigenvalues, then there would also be uncountably many orthogonal eigenfunctions. But this contradicts the fact that $L^2(a,b)$ is separable: there do not exist orthonormal systems with uncountably many members in $L^2(a,b)$.

Note: The same argument applies to any self-adjoint operator in a separable space. The left-shift operator S^* which has uncountably many eigenvalues is *not* self-adjoint.

X. Green's functions

One popular way to solve S-L-problems is to "diagonalize" the problem by using the eigenfunctions of the operator L. Unfortunately, the theory in Chapter VIII does *not apply* since L is not a *bounded* linear operator from a *space* back into *itself* (It only maps the subspace $D(L)$ of $L^2(a,b)$ into $L^2(a,b)$).

The solution to this problem is to study the *inverse* of L instead of L itself. This will be a bounded linear operator.

X.1 An example

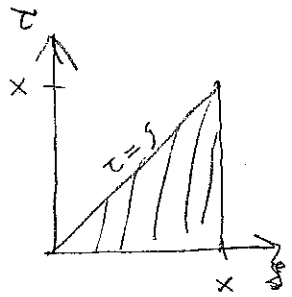
Before inverting a "general" L we look at a specific example (cf. Ex 9.4 with $\lambda = 0$):

$$(*) \qquad \frac{d^2\phi}{dx^2} = g\ ;$$
$$\phi(0) = \phi(1) = 0.$$

Here $L\phi = \frac{d^2\phi}{dx^2}$ on $D(L)$, so $L\phi = g$, and $\phi = L^{-1}g$.

To solve $(*)$ we integrate two times:

$$\phi(x) = \int_0^x \int_0^\xi g(\tau)\,d\tau\,d\xi + Ax + B$$
$$= \int_0^x \int_\tau^x d\xi\; g(\tau)\,d\tau + Ax + B$$
$$= \int_0^x (x-\tau)\,g(\tau)\,d\tau + Ax + B.$$



$\phi(0) = 0 \Rightarrow B = 0$

$\phi(1) = 0 \Rightarrow A = -\int_0^1 (1-\tau)\,g(\tau)\,d\tau.$

$$\Rightarrow \phi(x) = \int_0^x (x-\tau)\,g(\tau)\,d\tau - \int_0^x x(1-\tau)\,g(\tau)\,d\tau$$
$$- \int_x^1 x(1-\tau)\,g(\tau)\,dt$$

(over)