CHAPTER 5. Generalization of separation of variables: Sturm-Liouville theory & eigenfunction expansions

In Chapter 4, we studied some very simple linear PDEs (with constant coefficients) with simple boundary conditions (rectangular domains) which lend themselves particularly well to the separation of variables

In this chapter, we generalize the method to any homogeneous linear PDE of the form

$$\begin{cases} m(t)\, u_t = \mathcal{L}_x^{(2)}(u) \\ m(t)\, u_{tt} = \mathcal{L}_x^{(2)}(u) \end{cases} \quad \text{where } \mathcal{L}_x^{(2)}(u) = a(x)u_{xx} + b(x)u_x + c(x)u$$

and formalize the notion of boundary conditions.

5.1 Separation of variables in this case

Let, as usual, $u(x,t) = A(x)B(t)$ then we have (in the parabolic case, for example)

$$\frac{m(t)}{B}\frac{dB}{dt} = \frac{1}{A}\left[a(x)\frac{d^2A}{dx^2} + b(x)\frac{dA}{dx} + c(x)A\right] = \text{constant}.$$

$$\rightarrow \begin{cases} \dfrac{dB}{dt} = \dfrac{KB}{m(t)} \\ a(x)\dfrac{d^2A}{dx^2} + b(x)\dfrac{dA}{dx} + c(x)A = KA \end{cases}$$

For a given set of boundary conditions, The x-equation is an eigenvalue problem, which typically has an infinite number of solutions $A_n(x)$ each associated with a particular value K_n

$A_n(x)$ is called an eigen-mode

K_n is the associated eigen-value

The eigenmodes characterize the intrinsic spatial properties of the PDE. The eigenvalues characterize its intrinsic temporal properties.
(see previous chapter for examples).

5.2 Classification of the boundary conditions

Since we may be interested in a variety of domain shapes and associated BCs, we need a new classification system.

For a given spatial domain Ω, we can apply the following BCs to the contour $\partial\Omega$ of the domain:

(a) Dirichlet conditions

$$u(\underline{r},t) = f(\underline{r},t) \quad \forall\, r \in \partial\Omega$$

i.e the value of the function is fixed on the contour

examples
- $u(\underline{r},t) = 0$ (null condition)

  cf. Guitar string pinned at $x=0$ and $x=L$ (edge of domain)
- $u(\underline{r},t) = K$ (constant condition)

  cf. Ends of a rod held at some temperature K

(b) Von Neumann conditions

for $\underline{r} \in \partial\Omega$, $\underline{n}\cdot\nabla u = f(\underline{r},t)$ where $\underline{n}$ is the vector normal to the contour/edge of the domain.

i.e. the flux of u through the boundary is fixed.

example: $\dfrac{\partial u}{\partial z} = 0$ at $z = 0, L$.

(c) Robin conditions = mixed conditions

$$\alpha(\underline{r},t)\,\underline{n}\cdot\nabla u + \beta(\underline{r},t)\,u(r,t) = f(r,t) \quad \forall\, \underline{r} \in \partial\Omega.$$

Note: • This nomenclature applies to domains in any number of (spatial) dimensions.

→ For a 1D interval, then

Dirichlet conditions on $[a,b]$: $\begin{cases} u(a,t) = u_1(t) \\ u(b,t) = u_2(t) \end{cases}$

Neumann conditions on $[a,b]$: $\begin{cases} \frac{\partial u}{\partial x}(a,t) = u_1(t) \\ \frac{\partial u}{\partial x}(b,t) = u_2(t) \end{cases}$

Robin conditions: $\begin{cases} \alpha u(a,t) + \beta u_x(a,t) = u_1(t) \\ \gamma u(b,t) + \delta u_x(b,t) = u_2(t) \end{cases}$

$|\alpha| + |\beta| > 0 \quad |\gamma| + |\delta| > 0.$

6.3 Reformulation of the PDE

We now reformulate the problem by

Saying $\mathcal{L}_x^{(2)} = a(x) u_{xx} + b(x) u_x + c(x) u$ with $a \neq 0$

Multiply by $\frac{p(x)}{a(x)}$ with $p(x) = e^{\int \frac{b(x)}{a(x)} dx}$

then

$$\frac{p(x)}{a(x)} \mathcal{L}_x^{(2)} = p(x) u_{xx} + \frac{p(x)}{a(x)} b(x) u_x + \frac{c(x)}{a(x)} p(x) u$$

$$= p(x) u_{xx} + \frac{dp}{dx} u_x + \frac{c(x)}{a(x)} p(x) u$$

$$= (p u_x)_x + \frac{c(x)}{a(x)} p(x) u.$$

since $\frac{dp}{dx} = \frac{b(x)}{a(x)} e^{\int \frac{b(x)}{a(x)} dx} = \frac{b(x)}{a(x)} p(x)$

So the original PDES can be rewritten as

$$u_t = \frac{1}{m(t)r(x)}\left[(p(x)u_x)_x + q(x)u \right]$$

where
$$r(x) = \frac{p(x)}{a(x)}$$
$$q(x) = \frac{c(x)}{a(x)}p(x)$$
$$\left(\text{and } p(x) = e^{\int \frac{b(x)}{a(x)}dx}\right)$$

and similarly for the u_{tt} case.

As a result, separation of variables leads to (for the parabolic case, for example)

$$\begin{cases} \dfrac{dB}{dt} = \dfrac{KB(t)}{m(t)} \\ \dfrac{1}{r(x)}\left[\dfrac{d}{dx}\left(p(x)\dfrac{dA}{dx}\right) + q(x)A\right] = KA \end{cases}$$

let $K = -\lambda$ (a simple re-definition) then

$$\begin{cases} \dfrac{dB}{dt} = -\lambda\dfrac{B(t)}{m(t)} \\ \dfrac{d}{dx}\left[p(x)\dfrac{dA}{dx}\right] + q(x)A = -\lambda r(x)A \end{cases}$$

(similarly for the hyperbolic case)

The x-equation is a special type of eigenvalue ODE called a Sturm-Liouville equation, which has been extensively studied mathematically and for which there exist many important results.

5.4 Introduction to Sturm-Liouville Pbs.

- The eigenvalue problem

$$(p(x)u')' + q(x)u + \lambda r(x)u = 0$$

on the open interval $x \in (a,b)$ with

$$\begin{cases} \alpha u(a) + \beta u'(a) = 0 \\ \gamma u(b) + \delta u'(b) = 0 \end{cases}$$

is called a Sturm-Liouville problem provided
- $p(x)$, $p'(x)$, $q(x)$ and $r(x)$ are continuous
- $p(x)$, $r(x) > 0$ in (a,b)

and
- $|\alpha| + |\beta| > 0$, $|\gamma| + |\delta| > 0$

- if $p(x)$ or $r(x)$ vanish at $x=a$ or $x=b$, or if the interval (a,b) is unbounded (i.e either a or $b \to \pm\infty$) then the problem is called a singular Sturm-Liouville problem; otherwise the problem is regular

- The function $r(x)$ is called the weight function

Examples

① $\begin{cases} \dfrac{d^2u}{dx^2} + \lambda u = 0 \\ u(0) = u(L) = 0 \end{cases}$ is a regular S-L problem with $p(x) = 1$ $q(x) = 0$ $r(x) = 1$

Bessel eq. ② $\begin{cases} x^2 \dfrac{d^2u}{dx^2} + x\dfrac{du}{dx} + (x^2 - \nu^2)u = 0 \quad r \in (0,+\infty) \\ |u(0)| < +\infty \quad u(L) = 0 \end{cases}$ is a singular S.L problem with $r(x) = +\frac{1}{x}$ $p(x) = x$ $q(x) = x$ $\lambda = -\nu^2$

- Note that we may also consider periodic S.L problems: where $u(a) = u(b)$ and $u'(a) = u'(b)$ are the BCs.

5.5 Properties of Sturm–Liouville problems (ODEs)

① Symmetry of the operator

Given two functions u and v satisfying

$$\begin{cases} \alpha v(a)+\beta v'(a)=0 \\ \gamma v(b)+\beta v'(b)=0 \end{cases} \qquad \begin{cases} \alpha u(a)+\beta u'(a)=0 \\ \gamma u(b)+\beta u'(b)=0 \end{cases}$$

then $$\int_a^b [u\mathcal{L}(v)-v\mathcal{L}(u)]\,dx = 0$$

<u>Proof</u> $$\int_a^b [u\mathcal{L}(v)-v\mathcal{L}(u)]\,dx$$

$$= \int_a^b \{u[(pv')'+qv]-v[(pu')'+qu]\}\,dx$$

$$= \int_a^b \{u(pv')'-v(pu')'\}\,dx$$

integrate by parts.

$$\Big[upv'\Big]_a^b - \int_a^b pu'v'\,dx - \Big[pvu'\Big]_a^b + \int_a^b pu'v'\,dx$$

$$= \Big[p(uv'-vu')\Big]_a^b$$

$$= p(b)\{u(b)v'(b)-v(b)u'(b)\} - p(a)\{u(a)v'(a)-v(a)u'(a)\}$$

$$= 0 \qquad \text{using the bcs.}$$

② Orthogonality of the eigenfunctions

Eigenfunctions corresponding to $\neq$ eigenvalues λ are orthogonal wrt the inner product

$$\langle u,v\rangle = \int_a^b u(x)v(x)r(x)\,dx.$$

Proof: Let u_n be an eigenfunction with λ_n e.value
u_m ______ with λ_m e.value

$$\Rightarrow \begin{cases} \mathcal{L}(u_n) = -\lambda_n r u_n \\ \mathcal{L}(u_m) = -\lambda_m r u_m \end{cases}$$

then $\displaystyle\int_a^b \left[u_m \mathcal{L}(u_n) - u_n \mathcal{L}(u_m)\right] dx = 0$ by symmetry

$$= \int_a^b (\lambda_m - \lambda_n) r \, u_n u_m \, dx$$

$$= (\lambda_m - \lambda_n) \langle u_n, u_m \rangle$$

so unless $\lambda_m = \lambda_n$, $\langle u_n, u_m \rangle = 0$ □

③ The eigenvalues of the Sturm-Liouville problem are real

Proof Suppose λ is a complex eigenvalue, corresponding to a complex solution u.

then $\mathcal{L}(u) = -\lambda r u = (pu')' + qu$

then taking the C.C on both sides $\Rightarrow$

$\mathcal{L}(u^*) = -\lambda^* r u^*$ $\Rightarrow \lambda^*$ is the eigenvalue corresponding to the eigenfunction u^*.

$\Rightarrow$ if $\lambda \notin \mathbb{R}$ then $\lambda \neq \lambda^*$ and so

$\langle u, u^* \rangle = 0$

But $\displaystyle\int_a^b u u^* r \, dx = \int_a^b |u|^2 r \, dx > 0$ unless u is identically 0.

$\rightarrow$ So we reach a contradiction, implying that $\lambda \in \mathbb{R}$.

(4) The eigenvalues of a regular Sturm Liouville problem are simple
i.e: if two functions have the same eigenvalue then these functions are linearly dependent.

Proof: let v_1 and v_2 be two eigenfunctions belonging to the same eigenvalue

$$\mathcal{L}(v_1) = \lambda r v_1$$
$$\mathcal{L}(v_2) = \lambda r v_2$$

$$\Rightarrow \quad v_2 \mathcal{L}(v_1) = v_1 \mathcal{L}(v_2) = \lambda r v_1 v_2$$

so $\quad v_2 \mathcal{L}(v_1) - v_1 \mathcal{L}(v_2) = 0 \quad$ for all x.

Recall that $\quad v_2 \mathcal{L}(v_1) - v_1 \mathcal{L}(v_2) = v_2[(pv_1')' + q v_1]$
$$- v_1 [(pv_2')' + q v_2]$$
$$= v_2 (pv_1')' - v_1 (pv_2')'$$
$$= (p(v_2 v_1' - v_1 v_2'))'$$

So $\quad v_2 v_1' - v_1 v_2' = \text{constant}$.

However, on the boundaries this quantity is 0
$\Rightarrow$
$$v_2 v_1' = v_1 v_2'$$
$$\Rightarrow \left(\frac{v_1}{v_2}\right)' = 0 \Rightarrow \boxed{v_1 = \alpha v_2}$$

(5) The set of all eigenvalues for a regular Sturm-Liouville problem forms an unbounded, strictly monotone sequence:
$$\lambda_0 < \lambda_1 < \lambda_2 \dots < \lambda_n < \lambda_{n+1} < \dots < +\infty.$$

and $\lim_{n \to \infty} \lambda_n = +\infty$; λ_0 is called the principal eigenvalue