Notes on the Eigenfunction Method for solving differential equations

Reminder: We are considering the infinite-dimensional Hilbert space $L^2([a,b])$ of all square-integrable functions over the interval $[a,b]$ (i.e., $\int_a^b |f(x)|^2 \mathrm{d}x < \infty$) with the *inner product* define as

$$\boxed{\langle f|g\rangle := \int_a^b f^*(x)g(x)\mathrm{d}x}$$

and the *norm* induced by this inner product

$$\|f\| := \sqrt{\langle f|f\rangle} = \left[\int_a^b |f(x)|^2 \mathrm{d}x\right]^{1/2}.$$

One can introduce a set of linearly independent basis functions $\{y_n\}_{n=0}^{\infty}$, where $y_n \in L^2([a,b])$ for $n = 0, 1, \dots,$ such that any function $f \in L^2([a,b])$ can be represented as a linear combination of these functions

$$f(x) = \sum_{n=0}^{\infty} c_n y_n(x).$$

Such a set is not unique and, in general, the functions y_n are not orthonormal, but they do form a complete basis over the interval $[a,b]$. Furthermore, one can "produce" an orthonormal basis for the Hilbert space $L^2([a,b])$ from $\{y_n\}_{n=0}^{\infty}$ using the *Gram-Schmidt procedure* (actually, this procedure can be used to generate an orthonormal basis for any inner product space)

$$\boxed{\begin{cases} \phi_0 = y_0 & , \quad \hat{\phi}_0 = \frac{\phi_0}{\|\phi_0\|} \\ \phi_1 = y_1 - \hat{\phi}_0\langle\hat{\phi}_0|y_1\rangle & , \quad \hat{\phi}_1 = \frac{\phi_1}{\|\phi_1\|} \\ \vdots & \\ \phi_n = y_n - \sum_{k=0}^{n-1}\hat{\phi}_k\langle\hat{\phi}_k|y_n\rangle & , \quad \hat{\phi}_k = \frac{\phi_k}{\|\phi_k\|} \\ \vdots & \end{cases}}$$

Using the set of orthonormal functions $\{\hat{\phi}_n\}_{n=0}^{\infty}$ one can represent any function $f \in L^2([a,b])$ as

$$f(x) = \sum_{n=0}^{\infty} c_n \hat{\phi}_n(x)$$

with the coefficients c_n given by

$$c_n = \langle\hat{\phi}_n|f\rangle = \int_a^b \hat{\phi}_n^*(x) f(x)\mathrm{d}x.$$

This is true only if the basis is orthonormal.

1 Hermitian Conjugate of an Operator

Definition. The *adjoint* of an operator $\mathscr{L}$, denoted by $\mathscr{L}^\dagger$, is defined by

$$\boxed{\int_a^b f^*(x)[\mathscr{L}g(x)]\mathrm{d}x = \int_a^b [\mathscr{L}^\dagger f(x)]^* g(x)\mathrm{d}x + \text{boundary terms},} \tag{1}$$

where the boundary terms (b.t.) are evaluated at the end-points of the interval $[a, b]$. Using the definition of an inner product one may re-write it as

$$\langle f|\mathscr{L}g\rangle = \langle \mathscr{L}^\dagger f|g\rangle + b.t.$$

For a given linear differential operator, the method for formally finding the adjoint is to integrate by parts enough times to move the derivatives from g to f^*.

Example. Compute the Hermitian conjugate of the following operators:

1. $\mathscr{L}_1 = \frac{\mathrm{d}}{\mathrm{d}x}$.

$$\begin{aligned}
\langle f|\frac{\mathrm{d}}{\mathrm{d}x}g\rangle &= \int_a^b \underbrace{f^*(x)}_{u}\underbrace{\left[\frac{\mathrm{d}}{\mathrm{d}x}g(x)\right]}_{v'}\mathrm{d}x = \left\{\begin{array}{cc} u = f^*(x) & v' = \frac{\mathrm{d}}{\mathrm{d}x}g(x) \\ u' = \frac{\mathrm{d}}{\mathrm{d}x}f^*(x) & v = g(x) \end{array}\right\} \\
&= (f^*(x)g(x))\Big|_a^b - \int_a^b \frac{\mathrm{d}}{\mathrm{d}x}f^*(x)g(x)\mathrm{d}x \\
&= \int_a^b \left[-\frac{\mathrm{d}}{\mathrm{d}x}f(x)\right]^* g(x)\mathrm{d}x + \underbrace{(f^*(x)g(x))\Big|_a^b}_{\text{b.t.}}
\end{aligned}$$

$$\Rightarrow \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^\dagger = -\frac{\mathrm{d}}{\mathrm{d}x}$$

2. $\mathscr{L}_2 = \imath\frac{\mathrm{d}}{\mathrm{d}x}$.

$$\begin{aligned}
\langle f|\imath\frac{\mathrm{d}}{\mathrm{d}x}g\rangle &= \int_a^b \underbrace{f^*(x)}_{u}\underbrace{\left[\imath\frac{\mathrm{d}}{\mathrm{d}x}g(x)\right]}_{v'}\mathrm{d}x = \left\{\begin{array}{cc} u = f^*(x) & v' = \imath\frac{\mathrm{d}}{\mathrm{d}x}g(x) \\ u' = \frac{\mathrm{d}}{\mathrm{d}x}f^*(x) & v = \imath g(x) \end{array}\right\} \\
&= \imath(f^*(x)g(x))\Big|_a^b - \int_a^b \imath\frac{\mathrm{d}}{\mathrm{d}x}f^*(x)g(x)\mathrm{d}x \\
&= \int_a^b \left[\imath\frac{\mathrm{d}}{\mathrm{d}x}f(x)\right]^* g(x)\mathrm{d}x + \underbrace{\imath(f^*(x)g(x))\Big|_a^b}_{\text{b.t.}}
\end{aligned}$$

$$\Rightarrow \left(\imath\frac{\mathrm{d}}{\mathrm{d}x}\right)^\dagger = \imath\frac{\mathrm{d}}{\mathrm{d}x}$$

Remark. If $\mathscr{L}^\dagger = \mathscr{L}$ then it is said that $\mathscr{L}$ is *self-adjoint.*

Definition 1. If in Eq. (1) the b.t. vanishes, then an operator $\mathscr{L}$ is said to be *Hermitian* over the interval

$[a, b]$ and

$$\boxed{\int_a^b f^*(x)[\mathscr{L} g(x)]\mathrm{d}x = \int_a^b [\mathscr{L} f(x)]^* g(x)\mathrm{d}x} \tag{2}$$

(equivalently, $\langle f|\mathscr{L} g\rangle = \langle \mathscr{L} f|g\rangle$).

2 Properties of Hermitian operators

From now on we are going to assume that the operator $\mathscr{L}$ is Hermitian.

- **Reality of the eigenvalues**

  Let y_p and y_q be two eigenfunctions of an operator $\mathscr{L}$ corresponding to eigenvalues λ_p and λ_q, respectively.

$$\boxed{\mathscr{L}\, y_p(x) = \lambda_p\, y_p(x)}$$
$$\Downarrow \quad \langle \cdot | y_q\rangle$$
$$\langle \mathscr{L} y_p|y_q\rangle = \langle \lambda_p y_p|y_q\rangle$$
$$\boxed{\langle \mathscr{L} y_p|y_q\rangle = \lambda_p^*\langle y_p|y_q\rangle}$$

$$\boxed{\mathscr{L}\, y_q(x) = \lambda_q\, y_q(x)}$$
$$\Downarrow \quad \langle y_p|\cdot\rangle$$
$$\langle y_p|\mathscr{L} y_q\rangle = \langle y_p|\lambda_q y_q\rangle$$
$$\langle \mathscr{L}^\dagger y_p|y_q\rangle = \lambda_q\langle y_p|y_q\rangle$$
$$\Downarrow \quad \mathscr{L}^\dagger = \mathscr{L}$$
$$\boxed{\langle \mathscr{L} y_p|y_q\rangle = \lambda_q\langle y_p|y_q\rangle}$$

  Thus one has

$$\lambda_p^*\langle y_p|y_q\rangle - \lambda_q\langle y_p|y_q\rangle = 0 \quad \Rightarrow \quad \boxed{\left(\lambda_p^* - \lambda_q\right)\langle y_p|y_q\rangle = 0}.$$

  We get the same result using the integral form of the inner product. Multiplying the equation $\mathscr{L}\, y_p(x) = \lambda_p\, y_p(x)$ by y_q^* and equation $\mathscr{L}\, y_q(x) = \lambda_q\, y_q(x)$ by y_p^* , and then integrating over the interval $[a, b]$ leads to

$$\int_a^b y_q^*(x)[\mathscr{L}\, y_p(x)]\mathrm{d}x \quad = \quad \lambda_p \int_a^b y_q^*(x)\, y_p(x)\mathrm{d}x, \tag{3}$$
$$\int_a^b y_p^*(x)[\mathscr{L}\, y_q(x)]\mathrm{d}x \quad = \quad \lambda_q \int_a^b y_p^*(x)\, y_q(x)\mathrm{d}x. \tag{4}$$

  The complex conjugate of Eq. (3) becomes

$$\int_a^b y_q(x)[\mathscr{L}\, y_p(x)]^*\mathrm{d}x = \lambda_p^* \int_a^b y_q(x)\, y_p^*(x)\mathrm{d}x.$$

  Using now the definition of Hermitian operator (Eq. 2) one gets

$$\int_a^b [\mathscr{L} y_q(x)]\, y_p^*(x)\mathrm{d}x = \lambda_p^* \int_a^b y_q(x)\, y_p^*(x)\mathrm{d}x. \tag{5}$$

Comparing Eq. (4) and Eq. (5) leads to

$$\lambda_p^* \int_a^b y_q(x)\, y_p^*(x) \mathrm{d}x - \lambda_q \int_a^b y_p^*(x)\, y_q(x) \mathrm{d}x = \boxed{\left(\lambda_p^* - \lambda_q\right) \int_a^b y_p^*(x)\, y_q(x) \mathrm{d}x = 0}.$$

Now, if $p = q$, then it follows that $\lambda_p^* = \lambda_p$, i.e., λ_p is real, since $\int_a^b y_p^*(x)\, y_p(x)\mathrm{d}x = \langle y_p | y_p \rangle \geq 0$ (for all $n = 0, 1, \ldots,$ functions $y_n \neq 0$).

And what if $p \neq q$?

- **Orthogonality and normalization of the eigenfunctions**

We are going to consider two cases: 1. when y_p and y_q correspond to different eigenvalues, λ_p and λ_q respectively, and 2. when y_p and y_q correspond to the same eigenvalue λ_0.

1. If y_p and y_q correspond to different eigenvalues, then

$$\left(\lambda_p^* - \lambda_q\right) \langle y_p | y_q \rangle = 0 \quad \Rightarrow \quad \boxed{\langle y_p | y_q \rangle = 0}$$

or, in terms of integrals,

$$\left(\lambda_p^* - \lambda_q\right) \int_a^b y_p^*(x)\, y_q(x) \mathrm{d}x = 0 \quad \Rightarrow \quad \boxed{\int_a^b y_p^*(x)\, y_q(x) \mathrm{d}x = 0}$$

and we get two equivalent statements of the orthogonality of y_p and y_q.

2. If λ_0 is degenerate eigenvalue, i.e.,

$$\mathscr{L} y_\ell(x) = \lambda_0 y_\ell \qquad \text{for } \ell = 0, 1, \ldots, k-1$$

then it may happen that y_ℓ are not orthogonal. However, by performing the Gram-Schmidt orthonormalization procedure one can obtain k orthogonal vectors.

To normalize the set of basis functions $\{y_n\}_{n=0}^{\infty}$ it is enough to divide each function by its norm

$$\hat{y}_n(x) = \frac{y_n(x)}{\|y_n(x)\|},$$

obtaining in this way orthonormal set of basis functions $\{\hat{y}_n\}_{n=0}^{\infty}$ that are satisfying the condition

$$\boxed{\langle \hat{y}_p | \hat{y}_q \rangle = \int_a^b \hat{y}_p^*(x)\, \hat{y}_q(x) \mathrm{d}x = \begin{cases} 0, & \text{if } p \neq q \\ 1, & \text{if } p = q \end{cases}.}$$

- **Completeness of the eigenfunctions**

The infinite set of normalized eigenfunctions of a Hermitian operator form a *complete basis* set over $[a, b]$, i.e., any function $y \in L^2([a, b])$ obeying the boundary conditions may be expressed as a linear combination of these functions as

$$y(x) = \sum_{n=0}^{\infty} c_n \hat{y}_n(x),$$

where, as we know,

$$c_n = \langle \hat{y}_n | y \rangle = \int_a^b \hat{y}_n^*(x)\, y(x) \mathrm{d}x.$$

Thus

$$y(x) = \sum_{n=0}^{\infty} \hat{y}_n(x) \int_a^b \hat{y}_n^*(z)\, y(z) \mathrm{d}z = \int_a^b y(z) \sum_{n=0}^{\infty} \hat{y}_n(x)\, \hat{y}_n^*(z) \mathrm{d}x.$$

On the other hand

$$y(x) = \int_a^b y(z)\, \delta(x - z) \mathrm{d}z,$$

where $\delta(x - z)$ stands for the Dirac delta function. Comparing these two equalities leads to the *completeness* (or *closure*) property of the eigenfunctions

$$\boxed{\sum_{n=0}^{\infty} \hat{y}_n(x)\, \hat{y}_n^*(z) = \delta(x - z).}$$

For continuous spectrum one has to replace summation by integration

$$\int_a^b \hat{y}_n(x)\, \hat{y}_n^*(z) = \delta(x - z).$$

3 Why is it important?

We are concerned with the solution to the inhomogeneous equation of the form

$$\mathscr{L} y(x) = f(x), \tag{6}$$

where f is some known function and the boundary conditions are given. The idea for solving this problem is to use the linearity and Hemiticity of an operator $\mathscr{L}$. How do we do it?

Step-1: Find all eigenfunctions of an operator $\mathscr{L}$, i.e., solve the equation

$$\mathscr{L} y_n(x) = \lambda_n y_n(x).$$

Transform eigenfunctions $\{y_n\}_{n=0}^{\infty}$ so that they form an orthonormal basis.

Step-2: Re-write both y and f as a superpositions of eigenfunctions of $\mathscr{L}$

$$\begin{aligned} y(x) &= \sum_{n=0}^{\infty} c_n y_n(x), \\ f(x) &= \sum_{n=0}^{\infty} f_n y_n(x) \end{aligned}$$

where c_n and f_n stand for some constant.

Step-3: Make use of the linearity of the operator $\mathscr{L}$

$$\mathscr{L}y(x) = \mathscr{L}\left(\sum_{n=0}^{\infty} c_n y_n(x)\right) = \sum_{n=0}^{\infty} c_n \mathscr{L} y_n(x) = \boxed{\sum_{n=0}^{\infty} c_n \lambda_n y_n(x) = \sum_{n=0}^{\infty} f_n y_n(x)} = f(x).$$

Step-4: Take advantage of the orthonormality of functions $\{y_n\}_{n=0}^{\infty}$ - take an inner product $\langle y_m|\cdot\rangle$ of both sides in the above equality

$$\sum_{n=0}^{\infty} c_n \lambda_n \langle y_m|y_n\rangle = \sum_{n=0}^{\infty} f_n \langle y_m|y_n\rangle.$$

Since $\langle y_m|y_n\rangle = \delta_{mn}$, thus

$$c_m \lambda_m = f_m \quad \Rightarrow \quad \boxed{c_m = \frac{f_m}{\lambda_m}}.$$

Step-5: Put your coefficients in to get general solution of the problem (6)

$$\boxed{y(x) = \sum_{n=0}^{\infty} \frac{f_n}{\lambda_n} y_n(x)}.$$

4 The Sturm-Liouville equations

Let us consider the following equation

$$p(x)\frac{\mathrm{d}^2 y(x)}{\mathrm{d}x^2} + r(x)\frac{\mathrm{d}y(x)}{\mathrm{d}x} + q(x)y(x) + \lambda\rho(x)y(x) = 0,$$

where $r(x) = \frac{\mathrm{d}p(x)}{\mathrm{d}x}$, p, q and r are real functions and $\rho(x) > 0$ is called a *"weight"* or *"density"* function. It is clear that this equation can be written as

$$\mathscr{L}y(x) = \lambda\rho(x)y(x),$$

where

$$\mathscr{L} = -\left[p(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + r(x)\frac{\mathrm{d}}{\mathrm{d}x} + q(x)\right] = -\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}}{\mathrm{d}x}\right) + q(x)\right]$$

($\mathscr{L}$ is called the Sturm-Liouville operator).

It turns out the with an appropriate boundary conditions the Sturm-Liouville operator is Hermitian over the interval $[a, b]$. These boundary conditions are

$$\left[y_i^*(x)p(x)y_j'(x)\right]\Big|_{x=a} = \left[y_i^*(x)p(x)y_j'(x)\right]\Big|_{x=b} \quad \text{for all } i, j,$$

where $y_i$and y_j stand for two eigenfunctions of the operator $\mathscr{L}$. To check it let us put the Sturm-Liouville form $\mathscr{L} = -(py')' - qy$ into the Eq. (2)

$$\begin{aligned}
\int_a^b y_i^*(x)[\mathscr{L}y_j(x)]\mathrm{d}x &= -\int_a^b y_i^*\left[(py_j')' + qy_j\right]\mathrm{d}x = -\int_a^b y_i^*(py_j')'\mathrm{d}x - \int_a^b y_i^* q y_j \mathrm{d}x \\
&= \left\{\begin{array}{cc} u = y_i^* & v' = (py_j')' \\ u' = (y_i^*)' & v = py_j' \end{array}\right\} \\
&= -\underbrace{\left[y_i^* p y_j'\right]\Big|_a^b}_{\text{vanishes}} + \int_a^b (y_i^*)'\, p y_j' \mathrm{d}x - \int_a^b y_i^* q y_j \mathrm{d}x \\
&= \left\{\begin{array}{cc} u = (y_i^*)'p & v' = y_j' \\ u' = ((y_i^*)'p)' & v = y_j \end{array}\right\} \\
&= \underbrace{\left[(y_i^*)' p y_j\right]\Big|_a^b}_{\text{vanishes}} - \int_a^b ((y_i^*)'p)' y_j \mathrm{d}x - \int_a^b y_i^* q y_j \mathrm{d}x = -\int_a^b \left[((y_i^*)'p)' y_j + y_i^* q y_j\right]\mathrm{d}x \\
&= -\int_a^b \left[(p(y_i^*)')' + q y_i^*\right] y_j \mathrm{d}x = \int_a^b [\mathscr{L}y_i(x)]^* y_j(x)\mathrm{d}x
\end{aligned}$$

Why is the Sturm-Liouville equation important? It is because all second-order linear ordinary differential equations can be recast in the form of the LHS of the Sturm-Liouville equation. How? Consider an ordinary differential equation of the form

$$P(x)\frac{\mathrm{d}^2 y(x)}{\mathrm{d}x^2} + Q(x)\frac{\mathrm{d}y(x)}{\mathrm{d}x} + R(x)y(x) = 0.$$

We want to rewrite this equation as

$$-\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}}{\mathrm{d}x}\right) + q(x)\right] y(x) = \lambda\rho(x)y(x).$$

To do so, one must find p, q and ρ in the following way:

$$\begin{aligned}
p(x) &= \mathrm{e}^{\int^x \frac{Q(t)}{P(t)}\mathrm{d}t}, \\
\lambda\rho(x) - q(x) &= p(x)\frac{R(x)}{P(x)}.
\end{aligned}$$

Example 2. Let us consider Bessel's equation

$$x^2 y'' + xy' + (\lambda^2 x^2 - v^2)y = 0.$$

Here

$$\begin{aligned}
p(x) &= \mathrm{e}^{\int^x \frac{t}{t^2}\mathrm{d}t} = \mathrm{e}^{\ln x} = x \\
\lambda\rho(x) - q(x) &= x \cdot \frac{(\lambda^2 x^2 - v^2)}{x^2} = \frac{(\lambda^2 x^2 - v^2)}{x}
\end{aligned}$$

Thus we get

$$(xy')' + (\lambda x - \frac{v^2}{x})y = 0.$$