Stürm – Liouville Theory of orthogonal functions

The Fourier representation is very useful for solving differential equations in rectangular coordinates x, y, z. We need to develop similar techniques for other coordinate systems.

The Fourier representation is appropriate in a rectangular system because $e^{i\underset{\sim}{k}\cdot\underset{\sim}{x}}$ is the eigenfunction of the ∇ operator. That is

$$\nabla\left(e^{i\underset{\sim}{k}\cdot\underset{\sim}{x}}\right) = i\underset{\sim}{k}\left(e^{i\underset{\sim}{k}\cdot\underset{\sim}{x}}\right)$$

with $\underset{\sim}{k}$ the wavevector. To extend this concept to other geometries we must develop a procedure for obtaining eigenfunctions in those geometries.

Consider the operator:

$$\mathcal{L}(y) = a_2(x)\frac{d^2}{dx^2}y + a_1(x)\frac{d}{dx}y + a_0(x)y$$

defined over (a, b) where $a_2(x) \neq 0$ over $a < x < b$. Can have $a_2(a) = 0$ or $a_2(b) = 0$. Take a_0, a_1, a_2 to be real with no singular points interior to (a, b).

We can define an adjoint operator

$$\bar{\mathcal{L}}(y) = \frac{d^2}{dx^2} a_2 y - \frac{d}{dx} a_1 y + a_0 y$$

$$= a_2 y'' + (2a_2' - a_1) y'$$

$$+ (a_2'' - a_1' + a_0) y$$

If the operator is self adjoint ($\mathcal{L} = \bar{\mathcal{L}}$) we must ~~have~~ have

$$2a_2' - a_1 = a_1 \Rightarrow a_2' = a_1$$

or

$$\mathcal{L}(y) = a_2 y'' + a_2' y' + a_0$$

$$= \frac{d}{dx} a_2 \frac{dy}{dx} + a_0 y$$

If you have a second order equation that is not self-adjoint, you can always convert it to self-adjoint.

Let

$$\mathcal{L}(y) = a_2 y'' + a_1 y' + a_0 y$$

Multiply by

$$\frac{1}{a_2}\, e^{\int^x dx' \frac{a_1(x')}{a_2(x')}}$$

$$\frac{d}{dx}\, e^{\int^x dx' \left(\frac{a_1}{a_2}\right)} \frac{d}{dx} y + \frac{a_0}{a_2} e^{\int^x dx' \frac{a_1}{a_2}} y$$

Let $\hat{a}_2 = e^{\int^x dx' \left(\frac{a_1}{a_2}\right)}$

$$\hat{a}_0 = \frac{a_0}{a_2} e^{\int^x dx' \left(\frac{a_1}{a_2}\right)}$$

$$\hat{\mathcal{L}}(y) \equiv \frac{d}{dx}\, \hat{a}_2 \frac{d}{dx} y + \hat{a}_0 y$$

Thus, we assume that $\mathcal{L}$ is self-adjoint and of the form

$$\mathcal{L}(y) \equiv \frac{d}{dx}\, p(x) \frac{d}{dx} y + q(x) y$$

with $p(x) > 0$ except possibly at a, b.

Suppose that y satisfies the equation

$$\mathcal{L}(y) + \lambda\, w(x)\, y = 0$$

with λ a constant and $w(x)$ the weight function. $w(x)$ is positive except possibly at isolated points where $w=0$.

For certain values of λ_n (the eigenvalue) the solution y_n (the eigenfunction) satisfies the boundary conditions specified at a and b.

example : Legendre's Eqn

$$(1-x^2)y'' - 2xy' + \ell(\ell+1)y = 0$$

$$\frac{d}{dx}(1-x^2)\frac{dy}{dx} + \ell(\ell+1)y = 0$$

$\Rightarrow$ Legendre's eqn is self-adjoint with

$$p = 1-x^2, \; q = 0, \; w = 1$$

$$\text{and } \lambda = \ell(\ell+1)$$

Why are self-adjoint operators important?

Consider two solutions

$$\mathcal{L}(y_1) + \lambda_1 w y_1 = 0$$

$$\mathcal{L}(y_2) + \lambda_2 w y_2 = 0$$

Consider the integral

$$\int_a^b dx\, y_2^* \mathcal{L} y_1 = \int_a^b dx\, y_2^* (p y_1')' + \int_a^b dx\, q\, y_2^* y_1$$

$$= y_2^* p y_1' \Big|_a^b - \int_a^b dx\, p y_1' y_2^{*\prime}$$

$$+ \int_a^b dx\, q\, y_2^* y_1$$

where did one integration by parts. Choose BCs such that

$$p y_2^* y_1' \Big|_a^b = 0$$

Do a second integration by parts

$$\int_a^b dx\, y_2^* \mathcal{L} y_1 = - p y_1 y_2^{*\prime} \Big|_a^b$$

$$+ \int_a^b dx\, y_1 \frac{d}{dx} p\, y_2^{*\prime} + \int_a^b dx\, q\, y_2^* y_1$$

Again, take $p y_2 y_2^{*\prime} \Big|_a^b = 0$.

$$\int_a^b dx\, y_2^* \mathcal{L}(y) = \int_a^b dx\, y_1 \mathcal{L}(y_2^*)$$

This is the important property of self-adjoint operators

$\Longrightarrow$ with BCs

that allows us to generate a series of orthogonal functions that form a complete set.

Hermitean operators

More generally the operator might be complex. For example, in quantum mechanics

$$P_x = -i\hbar \frac{d}{dx}$$

In this case, we define a Hermitean operator as

$$\int_a^b dx\, y_2^* \mathcal{L} y = \int_a^b dx\, (\mathcal{L} y_2)^* y_1$$

For $\mathcal{L}$ real this is the same as for a self-adjoint operator.

Expectation values.

The expectation value of an operator $\mathcal{L}$ is

$$\langle f \rangle = \frac{\int_a^b dx\, \psi^* f \psi}{\int_a^b dx\, \psi^* \psi}$$

Any ~~obe~~ observable must be real so

$$\langle f \rangle^* = \int_a^b dx \frac{\psi (f\psi)^*}{\int_a^b dx\, |\psi|^2}$$

If f is Hermitian then

$$\langle f \rangle^* = \int_a^b dx \frac{\psi^* (f\psi)}{\int_a^b dx\, |\psi|^2} = \langle f \rangle$$

Thus, any operator corresponding to an observable quantity must be Hermitian.

Development of eigenfunctions from self-adjoint operators

Consider the equation

$$\mathcal{L} y + \lambda\, w(x)\, y = 0$$

$$\mathcal{L} = \frac{\partial}{\partial x} P(x) \frac{\partial}{\partial x} + q(x)$$

For any value of λ have two solutions

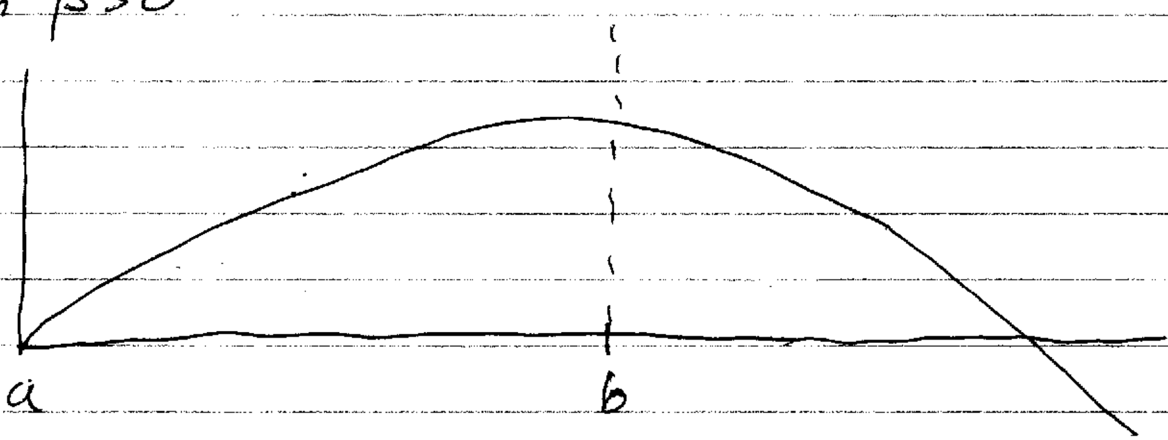
$$y = d_1 \psi_1(x) + d_2 \psi_2(x)$$

Choose a boundary at $x = a$. We can always choose $y(a) = 0$,

$$d_1 \psi_1(a) + d_2 \psi_2(a) = 0$$

$\Rightarrow$ assume not a sing. point of eqn.

Suppose $y(x)$ looks like this for some λ with $p > 0$



We want to vary λ to match the B.C at $x=b$. That is $y(b)=0$. We increase λ to cause the equation to become more oscillatory

$$\frac{d}{dx} p \frac{d}{dx} y + q y + \lambda w y = 0$$

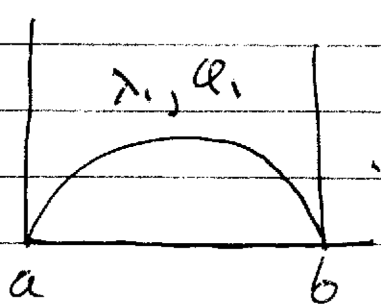
$\Rightarrow$ can think of λw like k^2

$\Rightarrow$ larger λ $\Rightarrow$ larger k^2 $\Rightarrow$ more oscillatory

$\Rightarrow$ zero point of y moves to the left

$\Rightarrow$ This can be rigorously proven for $p, w > 0$. (Morse and Feshbach Methods of Theoretical Physics)

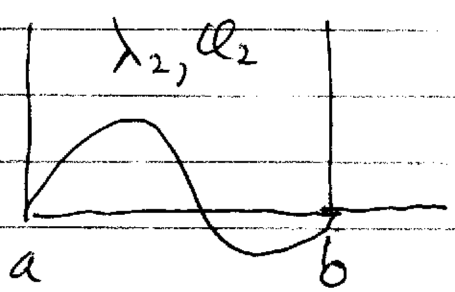
$\Rightarrow$ increase λ until $y(b)=0$



This defines the first eigenvalue λ_1 and eigenfunction $\varphi_1(x)$

$\Rightarrow$ increase λ further until the second zero intersects $x=b$.

$\Rightarrow$ yields $\lambda_2, \varphi_2(x)$



$\Rightarrow$ continue

$\Rightarrow$ generate an infinite # of eigenfunctions example $\frac{d^2}{dx^2} y_i + k_i^2 y_i = 0$

Solutions on interval (a, b)?

We require

$$y_n y_m' \Big|_a^b = 0$$

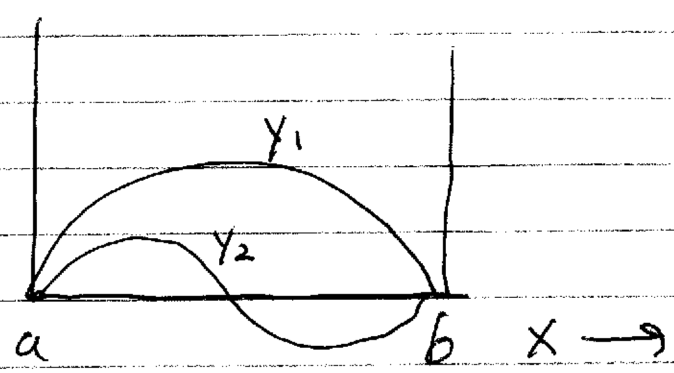
(1) Take $y_n(a), y_n(b) = 0$

$$y_n(x) = \sin k_n (x-a)$$

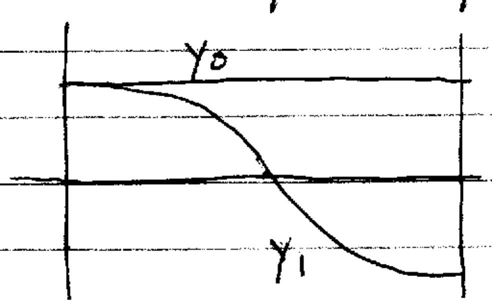
$$y_n(b) = 0 = \sin k_n(b-a)$$

$$k_n = \frac{n\pi}{b-a}$$

$$y_n = \sin \frac{(x-a)\pi n}{(b-a)}$$



(2) Take $y'(a) = y'(b) = 0$, $y_n = \cos\left[\frac{n\pi(x-a)}{b-a}\right]$



$\Rightarrow$ Different BCs yield different eigenfunctions

Properties of self-adjoint / Hermitean operators

(1) Eigenvalues of Hermitian operators are real

(2) Eigenfunctions of Hermitian operators are orthogonal

(3) Eigenfunctions of a Hermitian operator form a complete set.

Proof of (1):

$$\mathcal{L} y_i + \lambda_i \omega y_i = 0$$

Consider a second eigenfunction

$$\mathcal{L} y_j + \lambda_j \omega y_j = 0$$

$$\int_a^b dx\, y_j^* \mathcal{L} y_i + \lambda_i \int_a^b dx\, \omega y_j^* y_i = 0$$

Since $\mathcal{L}$ is Hermitian,

$$\int_a^b dx\, (\mathcal{L} y_j)^* y_i + \lambda_i \int_a^b dx\, \omega\, y_j^* y_i = 0$$

but $\mathcal{L} y_j = -\omega \lambda_j y_j$

$$-\int_a^b dx\, \lambda_j^* \omega y_j^* y_i + \lambda_i \int_a^b dx\, \omega y_j^* y_i = 0$$

$$(\lambda_i - \lambda_j^*) \int_a^b dx \; y_j^* y_i \; w = 0$$

Consider $i = j$

$$(\lambda_i - \lambda_i^*) \int_a^b dx \; w |y_i|^2 = 0$$

Since $w(x)$ is positive or at most zero at a finite # of points, must have

(1) $\lambda_i = \lambda_i^* \Rightarrow \lambda_i$ is real

If $i \neq j$ then if $\lambda_i \neq \lambda_j$, we must have

(2) $$\int_a^b dx \; w \; y_i \; y_j^* = 0$$

$\Rightarrow$ the eigenfunctions are orthogonal

$\Rightarrow$ $w(x)$ is the weight function in the orthogonality relation.

What if $i \neq j$ but $\lambda_i = \lambda_j$?

$\Rightarrow$ degenerate eigenvalues

$\Rightarrow$ eigenfunctions not necessarily orthogonal example: Consider

$$\left(\frac{d^2}{dx^2}+k^2\right)y=0$$

with interval $(0, 2\pi)$ with y periodic

$$\Rightarrow \text{BCs ok}$$

The eigenfunctions are $k_n = n$. For each n have two eigenfunctions

$$\sin nx, \cos nx$$

An arbitrary combination of $\sin nx, \cos nx$ are not orthogonal. Can alway choose eigenfunctions to be orthogonal. In this that choice is

$$\sin(nx), \cos(nx)$$

The choice

$$\sin(nx), \sin(nx)+\cos(nx)$$

would not be orthogonal.

Normalization

In general, we have

$$\int_a^b dx\, w |y_i|^2 = N_i^2$$

with N_i a real number. We can define our eigenfunctions

$$\hat{y}_i = \frac{y_i}{N_i}$$

so

$$\int_a^b dx\, w |\hat{y}_i|^2 = 1$$

Thus,

$$\int_a^b dx\, w\, \hat{y}_j^* \hat{y}_i = \delta_{ij}$$

$\Rightarrow$ this is an orthonormal set

Example Legendre Polynomials

$$\frac{d}{dx}(1-x^2)\frac{d}{dx}P_n + n(n+1)P_n = 0$$

Can illustrate these concepts using Legendre polynomials. Need Rodrigues' formula

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n$$

Can prove this from our previous series solutions.

$\Rightarrow$ note also that the solution of the Legendre eqn that behave like $\ln(1-x)$ do not satisfy the required BCs for self-adjoint operators

$$p(x)\, y_i^* y_j' \Big|_{-1}^{1} = (1-x^2)\, y_i y_j' \Big|_{-1}^{1}$$

$$\sim (1-x)\ln(1-x)\frac{1}{1-x}\Big|_{-1}^{1}$$

$$\Rightarrow \infty$$

Want to investigate the orthogonality of Legendre polynomials. BCs are

$$(1-x^2)\, P_n(x) P_m'(x) \Big|_{-1}^{1} = 0$$

Since $P_n(x)$ are bounded at $x = \pm 1$, BCs automatically satisfied for $x \in (-1, 1)$.

First take $m \neq n$. Assume $m > n$,

$$I \equiv \int_{-1}^{1} dx\, P_m(x) P_n(x) = \frac{1}{2^{m+n} n!\, m!} \int_{-1}^{1} dx \left(\frac{d^m}{dx^m}(x^2-1)^m\right)\left(\frac{d^n}{dx^n}(x^2-1)^n\right)$$

Integrate by parts m times and note that the endpoint contributions are zero

e.g. $\frac{d^{m-1}}{dx^{m-1}}(x^2-1)^m = \frac{d^{m-1}}{dx^{m-1}}\left[(x-1)^m(x+1)^m\right]$

$\Rightarrow$ even if all derivatives act on $(x-1)^m$, still have one $(x-1)$ left

$\Rightarrow 0$

$$\Rightarrow \int_{-1}^{1} dx\, (x^2-1)^m \frac{d^{m+n}}{dx^{m+n}} (x^2-1)^n$$

Since $m+n > 2n$, $\underbrace{\frac{d^{m+n}}{dx^{m+n}}(x^2-1)^n}_{\text{highest power is } x^{2n}} = 0$

Thus, P_n, P_m are orthogonal for $m \neq n$.

Normalization of Legendre polynomials:
For $m=n$

$$I_n = \int_{-1}^{1} dx\, P_n^2(x) = \frac{1}{2^{2n}(n!)^2}\int_{-1}^{1} dx \left[\frac{d^n}{dx^n}(x^2-1)^n\right]^2$$

Again integrate by parts n times

$$I_n = \frac{1}{2^{2n}(n!)^2}(-1)^n \int_{-1}^{1} dx\, (x^2-1)^n \underbrace{\frac{d^{2n}}{dx^{2n}}(x^2-1)^n}_{(2n)!}$$

$$= \frac{(2n)!}{2^{2n}(n!)^2} \int_{-1}^{1} dx\, (1-x^2)^n$$

Obtain I_n by induction

$$(1-x^2)^n = (1-x^2)(1-x^2)^{n-1} = (1-x^2)^{n-1} - x^2(1-x^2)^{n-1}$$

$$= (1-x^2)^{n-1} + \frac{x}{2n}\frac{d}{dx}(1-x^2)^n$$

$$I_n = \frac{2n(2n-1)(2n-2)!}{4\,4^{n-1}\,n^2[(n-1)!]^2} \int_{-1}^{1} dx\,(1-x^2)^{n-1} + \frac{(2n)!}{4^n(n!)^2}\underbrace{\int_{-1}^{1} dx\,\frac{x}{2n}\frac{d}{dx}(1-x^2)^n}_{\text{integrate by parts}}$$

$$I_n = \frac{2n(2n-1)}{4n^2} I_{n-1} - \frac{1}{2n} I_n$$

$$(2n+1) I_n = (2n-1) I_{n-1} = [2(n-1)+1] I_{n-1}$$

Thus, $\quad I_n = \dfrac{\text{const}}{2n+1}$

For $n=0$, $P_0 = 1 \qquad \int_{-1}^{1} dx\, P_0^2 = \int_{-1}^{1} dx = 2$

$I_n = \dfrac{2}{2n+1} \Rightarrow \sqrt{\dfrac{2n+1}{2}}\, P_n(x)$ are the normalized polynomials.

Completeness

The motivation for studying eigenfunctions solutions of Stürm-Liouville equations is that we want to use these eigenfunctions to solve diff. eqns. To do this we must be able to show that we can represent a function $f(x)$ the series

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

where $\varphi_n(x)$ are the eigenfunctions of the self adjoint operator of interest. To show that such a series converges uniformly and is unique is required. In the case where $f(x)$ is analytic this task is straight foreward since $f(x)$ can be expanded in a Taylor series

$$f(x) = \sum_{n=0}^{\infty} b_n x^n$$

and we know that such a series converges uniformly and is unique. All that needs to be shown is that x^n can be expanded as a series of eigenfunctions

$$x^n = \sum_{m=0}^{\infty} a_{nm} P_m(x)$$

<u>example</u> Legendre Polynomials

Multiply the equation by $P_\ell(x)$ and integrate over $(-1,1)$

$$\int_{-1}^{1} dx\, P_\ell(x)\, x^n = \frac{2}{2\ell+1} a_{n\ell}$$

Use Rodrigues's' formula for $P_\ell(x)$

$$a_{n\ell} = \frac{2\ell+1}{2} \int_{-1}^{1} dx\, x^n \frac{1}{2^\ell \ell!} \frac{d^\ell}{dx^\ell} (x^2-1)^\ell$$

Integrate by parts ℓ times

$$a_{n\ell} = \frac{2\ell+1}{2} (-1)^\ell \int_{-1}^{1} dx\, (x^2-1)^\ell x^{n-\ell} \frac{n!}{(n-\ell)!}$$

for $\ell \leq n$

$= 0$ for $\ell > n$

Thus,

$$x^n = \sum_{\ell=0}^{n} a_{n\ell} P_\ell(x)$$

$\Longrightarrow$ highest power of $P_\ell(x)$ is x^ℓ so don't need $\ell > n$ to represent x^n.

Thus, when $f(x)$ is a analytic function, we can expand $f(x)$ in a power series in x and each power of x can be expanded in a Legendre series so $f(x)$ can be directly expanded in a Legendre series

$$f(x) = \sum_{n=0}^{\infty} C_n P_n(x)$$

Similar results can be obtained for other polynomials (Chebyshev, Laguerre, ..)

Measures of Convergence

Consider a function $f(x)$ which we ~~can~~ represent by a series of functions $\varphi_n(x)$ over (a,b). Define

$$f_\ell(x) = \sum_{n=1}^{\ell} C_n \varphi_n(x)$$

1) Pointwise convergence: For any ϵ there exists an N such that

$$|f(x) - f_\ell(x)| < \epsilon$$

for all $\ell > N$, where N may depend on x.

2) Uniform convergence: For any ϵ there exists an N such that

$$\left| f(x) - f_\ell(x) \right| < \epsilon$$

for $\ell > N$ and for all values of x.

Often we will want to represent functions that are not ~~analyt~~ analytic. For example, $f(x) = |x|$ is not analytic anywhere yet it would seem to be reasonable to represent it by a series of eigenfunctions. On the other hand, for such cases since we can not represent $f(x)$ by a uniformly convergent power series, we should not expect

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

to converge uniformly. We introduce a weaker convergence criterion,

3) Convergence in the mean: Specifically we define a set of functions to be complete, if any function that is piecewise continuous can be represented by a series of the $\mathcal{U}_n$'s such that

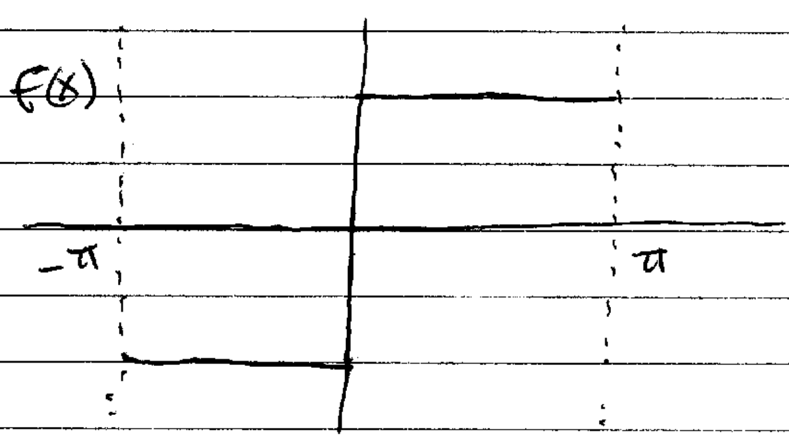
$$\lim_{m \to \infty} \int_a^b dx\, w(x) \left(f(x) - \sum_{n=0}^{m} c_n \varphi_n \right)^2 = 0$$

Convergence in the mean is weaker than pointwise or uniform convergence since $f(x)$ can deviate substantially from $\sum_{n=0}^{\infty} c_n \varphi_n(x)$ at locations, for example, where $f(x)$ is discontinuous as long as the integral above is zero.

$\Rightarrow$ All of the common functions of mathematical physics (Bessel, Legendre, ...) are complete in this sense.

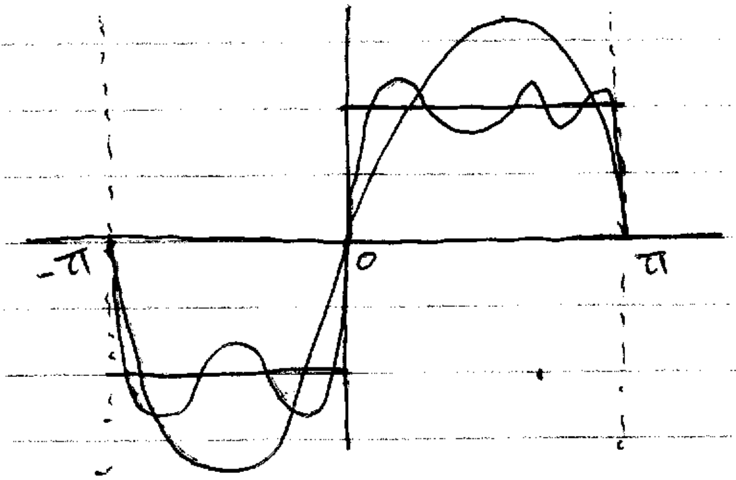
The Gibbs phenomenon (Arfken p. 959)

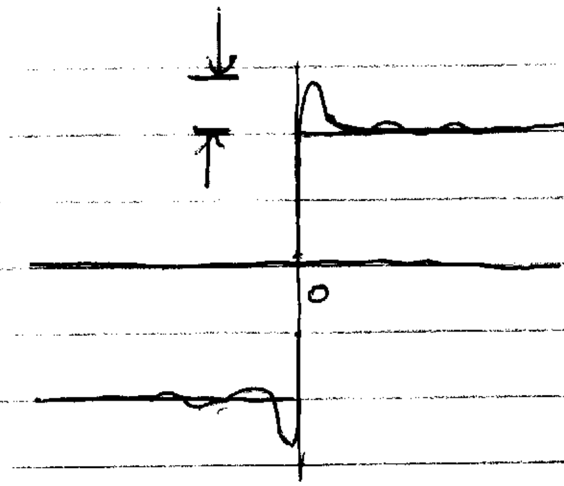
Let $f(x)$ be given by the square-wave function



$$f(x) = \sum_{n=1}^{\infty} c_n \sin(nx)$$

$$f(x) = \frac{4}{\pi} \sum_{n\,\mathrm{odd}} \frac{\sin(nx)}{n}$$





The overshoot around zero does not diminish as the number of the terms in the series increases

$\Rightarrow$ the region where the partial sum deviates from $f(x)$ becomes smaller

$\Rightarrow$ no uniform convergence

$\Rightarrow$ convergence in the mean

Calculating eigenfunctions representations of $f(x)$

Consider a function $f(x)$ and the series representation

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

The c_n's are the generalized Fourier coefficients. When the φ_n's are orthonormal, we can calculate the c_n's easily

$$\int_a^b dx\, w(x)\, \varphi_m^*(x)\, f(x) = \sum_{n=0}^{\infty} c_n \int_a^b dx\, w\, \varphi_m^*(x)\, \varphi_n(x)$$

$$= \sum_{n=0}^{\infty} c_n \delta_{mn} = c_m$$

$$c_m = \int_a^b dx\, w\, \varphi_m^*(x)\, f(x)$$

Define the inner product

$$(f_1, f_2) = \int_a^b dx\, w(x)\, f_1^*(x)\, f_2(x)$$

then

$$c_m = (\varphi_m, f)$$

$$f(x) = \sum_{n=0}^{\infty} (\varphi_n, f)\, \varphi_n$$

or

$$f(x) = \sum_{n=0}^{\infty} \int_a^b dx' \, \varphi_n^*(x') \, f(x') \, w(x') \, \varphi_n(x)$$

$$= \int_a^b dx' \, w(x') \left(\sum_{n=0}^{\infty} \varphi_n^*(x') \varphi_n(x) \right) f(x')$$

Thus, we must have

$$\delta(x-x') = \sum_{n=0}^{\infty} \varphi_n^*(x') \, \varphi_n(x) \, w(x)$$

This is called the closure relationship.

Closure

Definition: a set of orthonormal functions is closed if no non-zero function is orthogonal to every function in the set

Theorem: A set is complete if and only if it is closed.

$\Rightarrow$ If the function $f(x)$ is orthogonal to every eigenfunction φ_n, then it can not be represented a series of those eigenfunctions and the set of functions must not be complete.