PART IV
SPECTRAL METHODS

- ADDITIONAL REFERENCES:
  - R. Peyret, *Spectral methods for incompressible viscous flow*, Springer (2002),
  - B. Mercier, *An introduction to the numerical analysis of spectral methods*, Springer (1989),
  - C. Canuto, M. Y. Hussaini, A. Quarteroni and T. A. Zang, *Spectral Methods — Fundamentals in Single Domains*, Springer (2006).
  - J. Shen, T. Tang and L.-L. Wang, *Spectral Methods: Algorithms, Analysis and Applications*, Springer (2011).

Agenda

- General Formulation
  - Method of Weighted Residuals
  - Approximation of Functions
  - Approximation of PDEs
- Orthonormal Systems
  - Spectral Theorem
  - Polynomial Approximation
  - Orthogonal Polynomials
- Fourier Series
  - Convergence Results
  - Spectral Differentiation
  - Numerical Quadratures

- SPECTRAL METHODS belong to the broader category of WEIGHTED RESIDUAL METHODS , for which approximations are defined in terms of series expansions, such that a measure of the error knows as the RESIDUAL is set to be zero in some approximate sense

- In general, an approximation $u_N(x)$ to $u(x)$ is constructed using a set of basis functions $\varphi_k(x)$, $k = 0, \dots, N$ (note that $\varphi_k(x)$ need not be ORTHOGONAL)

$$u_N(x) \triangleq \sum_{k \in I_N} \hat{u}_k \varphi_k(x), \quad a \le x \le b, \quad I_N = \{1, \dots, N\}$$

- Residual for two main problems:
  - APPROXIMATION of a function u:

$$R_N(x) = u - u_N$$

  - APPROXIMATE SOLUTION of a (differential) equation $\mathcal{L}u - f = 0$:

$$R_N(x) = \mathcal{L}u_N - f$$

- In general, the residual R_N in cancelled in the following sense:

$$(R_N, \psi_i)_{w_*} = \int_a^b w_* \, R_N \, \bar{\psi}_i \, dx = 0, \quad i \in I_N,$$

where $\psi_i(x)$, $i \in I_N$ are the TRIAL (TEST) FUNCTIONS and $w : [a, b] \to \mathbb{R}^+$ are the WEIGHTS

- Spectral Method is obtained by:
  - selecting the BASIS FUNCTIONS φ_k to form an ORTHOGONAL system under the weight w:

$$(\varphi_i, \varphi_k)_w = \delta_{ik}, \quad i, k \in I_N \quad \text{and}$$

  - selecting the trial functions to coincide with the basis functions:

$$\psi_k = \varphi_k, \quad k \in I_N$$

    with the weights $w_* = w$ (SPECTRAL GALERKIN APPROACH), or
  - selecting the trial functions as

$$\psi_k = \delta(x - x_k), \quad x_k \in (a, b),$$

    where x_k are chosen in a non-arbitrary manner, and the weights are $w_* = 1$ (COLLOCATION, "PSEUDO-SPECTRAL" APPROACH)
- Note that the residual R_N vanishes
  - in the mean sense specified by the weight w in the Galerkin approach
  - pointwise at the points x_k in the collocation approach

Galerkin Method

- Assume that the basis functions $\{\varphi_k\}_{k=1}^N$ form an orthogonal set
- Define the residual: $R_N(x) = u - u_N = u - \sum_{k=0}^N \hat{u}_k \varphi_k$
- Cancellation of the residual in the mean sense (with the weight w)

$$(R_N, \varphi_i)_w = \int_a^b \left(u - \sum_{k=0}^N \hat{u}_k \varphi_k \right) \bar{\varphi}_i \, w \, dx = 0, \quad i = 0, \dots, N$$

$(\bar{\cdot})$ denotes complex conjugation (cf. definition of the inner product)

- Orthogonality of the basis / trial functions thus allows us to determine the coefficients $\hat{u}_k$ by evaluating the expressions

$$\hat{u}_k = \int_a^b u \, \bar{\varphi}_k \, w \, dx, \quad k = 0, \dots, N$$

- Note that, for this problem, the Galerkin approach is equivalent to the LEAST SQUARES METHOD .

Collocation Method

- Define the residual: $R_N(x) = u - u_N = u - \sum_{k=0}^{N} \hat{u}_k \varphi_k$
- POINTWISE cancellation of the residual

$$\sum_{k=0}^{N} \hat{u}_k \varphi_k(x_i) = u(x_i), \quad i = 0, \dots, N$$

  Determination of the coefficients $\hat{u}_k$ thus requires solution of an algebraic system. Existence and uniqueness of solutions requires that $\det\{\varphi_k(x_i)\} \neq 0$ (condition on the choice of the collocation points x_j and the basis functions φ_k)

- For certain pairs of basis functions φ_k and collocation points x_j the above system can be easily inverted and therefore determination of $\hat{u}_k$ may be reduced to evaluation of simple expressions
- For this problem, the collocation method thus coincides with an INTERPOLATION TECHNIQUE based on the set of points $\{x_j\}$

Galerkin Method (I)

- Consider a generic PDE problem

$$\begin{cases} \mathcal{L}u - f = 0 & a < x < b \\ \mathcal{B}_- u = g_- & x = a \\ \mathcal{B}_+ u = g_+ & x = b, \end{cases}$$

where $\mathcal{L}$ is a linear, second-order differential operator, and $\mathcal{B}_-$ and $\mathcal{B}_+$ represent appropriate boundary conditions (Dirichlet, Neumann, or Robin)

Galerkin Method (II)

- Reduce the problem to an equivalent HOMOGENEOUS formulation via a "lifting" technique, i.e., substitute $u = \tilde{u} + v$, where $\tilde{u}$ is an arbitrary function satisfying the boundary conditions above and the new (homogeneous) problem for v is

$$\begin{cases} \mathcal{L}v - h = 0 & a < x < b \\ \mathcal{B}_{-}v = 0 & x = a \\ \mathcal{B}_{+}v = 0 & x = b, \end{cases}$$

  where $h = f - \mathcal{L}\tilde{u}$

- The reason for this transformation is that the basis functions φ_k (usually) satisfy homogeneous boundary conditions.

Galerkin Method (III)

- The residual $R_N(x) = \mathcal{L}v_N - h$, where $v_N = \sum_{k=0}^{N} \hat{v}_k \varphi_k(x)$ satisfies ("by construction") the boundary conditions
- Cancellation of the residual in the mean (cf. THE WEAK FORMULATION)

$$(R_N, \varphi_i)_w = (\mathcal{L}v_N - h, \varphi_i)_w, \quad i = 0, \dots, N$$

Thus

$$\sum_{k=0}^{N} \hat{v}_k \, (\mathcal{L}\varphi_k, \varphi_i)_w = (h, \varphi_i)_w, \quad i = 0, \dots, N,$$

where the scalar product $(\mathcal{L}\varphi_k, \varphi_i)_w$ can be accurately evaluated using properties of the basis functions φ_i and $(h, \varphi_i)_w = \hat{h}_i$

- An $(N+1) \times (N+1)$ algebraic system is obtained with the matrix determined by
  - the properties of the basis functions $\{\varphi_k\}_{k=1}^{N}$
  - the properties of the operator $\mathcal{L}$

Collocation Method (I)

- The residual (corresponding to the original inhomogeneous problem)

$$R_N(x) = \mathcal{L}u_N - f, \quad \text{where} \quad u_N = \sum_{k=0}^{N} \hat{u}_k \varphi_k(x)$$

- Pointwise cancellation of the residual, including the boundary nodes:

$$\begin{cases} \mathcal{L}u_N(x_i) = f(x_i) & \qquad i = 1, \dots, N-1 \\ \mathcal{B}_- u_N(x_0) = g_- & \\ \mathcal{B}_+ u_N(x_N) = g_+, & \end{cases}$$

This results in an $(N+1) \times (N+1)$ algebraic system. Note that depending on the properties of the basis $\{\varphi_0, \dots, \varphi_N\}$, this system may be singular.

Collocation Method (II)

- Sometimes an alternative formulation is useful, where the nodal values $u_N(x_j)$ $j = 0, \dots, N$, rather than the expansion coefficients $\hat{u}_k$, $k = 0, \dots, N$ are unknown. The advantage is a convenient form of the expression for the derivative

$$u_N^{(p)}(x_i) = \sum_{j=0}^{N} d_{ij}^{(p)} \, u_N(x_j),$$

where $d^{(p)}$ is a p-TH ORDER DIFFERENTIATION MATRIX .

Theorem

Let $\mathcal{H}$ be a separable Hilbert space and $\mathcal{T}$ a compact Hermitian operator. Then, there exists a sequence $\{\lambda_n\}_{n\in\mathbb{N}}$ and $\{W_n\}_{n\in\mathbb{N}}$ such that

1. $\lambda_n \in \mathbb{R}$,
2. *the family* $\{W_n\}_{n\in\mathbb{N}}$ *forms* A COMPLETE BASIS *in* $\mathcal{H}$
3. $\mathcal{T}W_n = \lambda_n W_n$ *for all* $n \in \mathbb{N}$

- Systems of orthogonal functions are therefore related to spectra of certain operators, hence the name SPECTRAL METHODS

Example I

- Let $\mathcal{T}: \ L_2(0,\pi) \to L_2(0,\pi)$ be defined for all $f \in L_2(0,\pi)$ by $\mathcal{T} f = u$, where u is the solution of the Dirichlet problem

  $$\begin{cases} -u'' = f \\ u(0) = u(\pi) = 0 \end{cases}$$

  Compactness of $\mathcal{T}$ follows from the Lax-Milgram lemma and compact embeddedness of $H^1(0,\pi)$ in $L_2(0,\pi)$.

- EIGENVALUES AND EIGENVECTORS

  $$\lambda_k = \frac{1}{k^2} \quad \text{and} \quad W_k = \sqrt{2}\sin(kx) \quad \text{for} \quad k \geq 1$$

Example I

- Thus, each function $u \in L_2(0,\pi)$ can be represented as

$$u(x) = \sqrt{2} \sum_{k \geq 1} \hat{u}_k W_k(x),$$

where $\hat{u}_k = (u, W_k)_{L_2} = \frac{\sqrt{2}}{\pi} \int_0^\pi u(x) \sin(kx)\, dx$.

- Uniform (pointwise) convergence is not guaranteed (only in L_2 sense)!

Example II

- Let $\mathcal{T}: \; L_2(0,\pi) \to L_2(0,\pi)$ be defined for all $f \in L_2(0,\pi)$ by $\mathcal{T}f = u$, where u is the solution of the Neumann problem

$$\begin{cases} -u'' + u = f \\ u'(0) = u'(\pi) = 0 \end{cases}$$

  Compactness of $\mathcal{T}$ follows from the Lax-Milgram lemma and compact embeddedness of $H^1(0,\pi)$ in $L_2(0,\pi)$.

- EIGENVALUES AND EIGENVECTORS

$$\lambda_k = \frac{1}{1+k^2} \quad \text{and} \quad W_0(x) = 1, \;\; W_k = \sqrt{2}\cos(kx) \quad \text{for} \quad k > 1$$

Example II

- Thus, each function $u \in L_2(0,\pi)$ can be represented as

$$u(x) = \sqrt{2}\sum_{k\geq 0} \hat{u}_k W_k(x),$$

where $\hat{u}_k = (u, W_k)_{L_2} = \frac{\sqrt{2}}{\pi}\int_0^\pi u(x)\cos(kx)\,dx$.

- Uniform (pointwise) convergence is not guaranteed (only in L_2 sense)!

Example III

- Expansion in SINE SERIES for functions vanishing on the boundaries
- Expansion in COSINE SERIES for functions with first derivatives vanishing on the boundaries
- Combining sine and cosine expansions we obtain the FOURIER SERIES EXPANSION with the basis functions (in $L_2(-\pi, \pi)$)

$$W_k(x) = e^{ikx}, \quad \text{for} \quad k \geq 0$$

W_k form a Hilbert basis more flexible then sine or cosine series alone.

Example III

- Fourier series vs. Fourier transform —
  - Fourier Transform : $\mathcal{F}_1 : L_2(\mathbb{R}) \to L_2(\mathbb{R})$,

$$\mathcal{F}_1[u](k) = \int_{-\infty}^{\infty} e^{-ikx} u(x)\, dx, \quad k \in \mathbb{R}$$

  - Fourier Series : $\mathcal{F}_2 : L_2(0, 2\pi) \to l_2$, (i.e., bounded to discrete)

$$\hat{u}_k = \mathcal{F}_2[u](k) = \int_0^{2\pi} e^{-ikx} u(x)\, dx, \quad k = 0, 1, 2, \dots$$

Theorem (Weierstrass Approximation Theorem)

To any function $f(x)$ that is continuous in $[a, b]$ and to any real number $\epsilon > 0$ there corresponds a polynomial $P(x)$ such that $\|P(x) - f(x)\|_{C(a,b)} < \epsilon$, i.e. the set of polynomials is DENSE *in the Banach space $C(a, b)$*

($C(a, b)$ is the Banach space with the norm $\|f\|_{C(a,b)} = \max_{x\in[a,b]} |f(x)|$)

- Thus the power functions x^k, $k = 0, 1, \dots$ represent a natural basis in $C(a, b)$
- QUESTION — Is this set of basis functions useful?

NO! — SEE BELOW

- Find the polynomial $\bar{P}_N$ (of order N) that best approximates a function $f \in L_2(a,b)$ [note that we will need the structure of a Hilbert space, hence we go to $L_2(a,b)$, but $C(a,b) \subset L_2(a,b)$], i.e.

$$\int_a^b [f(x) - \bar{P}_N(x)]^2 \, dx \le \int_a^b [f(x) - P_N(x)]^2 \, dx$$

  where $\bar{P}_N(x) = \bar{a}_0 + \bar{a}_1 x + \bar{a}_2 x^2 + \cdots + \bar{a}_N x^N$

- Using the formula $\sum_{j=0}^N \bar{a}_j (e_j, e_k) = (f, e_k)$, $j = 0, \dots, N$, where $e_k = x^k$

$$\sum_{k=0}^N \bar{a}_k \int_a^b x^{k+j} \, dx = \int_a^b x^j f(x) \, dx$$

$$\sum_{k=0}^N \bar{a}_k \frac{b^{k+j+1} - a^{k+j+1}}{k+j+1} = \int_a^b x^j f(x) \, dx$$

- The resulting algebraic problem is extremely ILL-CONDITIONED , e.g. for $a = 0$ and $b = 1$

$$[A]_{kj} = \frac{1}{k+j+1}$$

- Much better behaved approximation problems are obtained with the use of ORTHOGONAL BASIS FUNCTIONS
- Such systems of orthogonal basis functions are derived by applying the SCHMIDT ORTHOGONALIZATION PROCEDURE to the system $\{1, x, \dots, x^N\}$
- Various families of ORTHOGONAL POLYNOMIALS are obtained depending on the choice of:
  - the domain $[a, b]$ over which the polynomials are defined, and
  - the weight w characterizing the inner product $(\cdot, \cdot)_w$ used for orthogonalization

- Polynomials defined on the interval $[-1, 1]$
  - LEGENDRE POLYNOMIALS $(w = 1)$

$$P_k(x) = \sqrt{\frac{2k+1}{2}}\,\frac{1}{2^k\,k!}\,\frac{d^k}{dx^k}(x^2-1)^k, \quad k = 0, 1, 2, \dots$$

  - JACOBI POLYNOMIALS $(w = (1-x)^\alpha(1+x)^\beta)$

$$J_k^{(\alpha,\beta)}(x) = C_k(1-x)^{-\alpha}(1+x)^{-\beta}\frac{d^k}{dx^k}[(1-x)^{\alpha+k}(1+x)^{\beta+k}] \;\; k = 0, 1, 2, \dots,$$

    where C_k is a very complicated constant

  - CHEBYSHEV POLYNOMIALS $(w = \frac{1}{\sqrt{1-x^2}})$

$$T_n(x) = \cos(k\,\arccos(x)), \quad k = 0, 1, 2, \dots,$$

    Note that Chebyshev polynomials are obtained from Jacobi polynomials for $\alpha = \beta = -1/2$

- Polynomials defined on the PERIODIC interval $[-\pi, \pi]$
  TRIGONOMETRIC POLYNOMIALS $(w = 1)$

$$S_k(x) = e^{ikx} \quad k = 0, 1, 2, \dots$$

- Polynomials defined on the interval $[0, +\infty]$
  LAGUERRE POLYNOMIALS $(w = e^{-x})$

$$L_k(x) = \frac{1}{k!}\, e^x \, \frac{d^k}{dx^k}(e^{-x} x^k), \quad k = 0, 1, 2, \dots$$

- Polynomials defined on the interval $[-\infty, +\infty]$
  HERMITE POLYNOMIALS $(w = 1)$

$$H_k(x) = \frac{(-1)^k}{(2^k\, k!\, \sqrt{\pi})^{1/2}}\, e^{x^2}\, \frac{d^k}{dx^k} e^{-x^2}, \quad k = 0, 1, 2, \dots$$

- What is the relationship between ORTHOGONAL POLYNOMIALS and eigenfunctions of a COMPACT HERMITIAN OPERATOR (cf. Spectral Theorem)

- Each of the aforementioned families of ORTHOGONAL POLYNOMIALS forms the set of eigenvectors for the following STURM-LIOUVILLE PROBLEM

$$\begin{aligned} &\frac{d}{dx}\left[p(x)\frac{dy}{dx}\right] + [q(x) + \lambda r(x)]\, y = 0 \\ &a_1 y(a) + a_2 y'(a) = 0 \\ &b_1 y(b) + b_2 y'(b) = 0 \end{aligned}$$

for appropriately selected domain $[a, b]$ and coefficients p, q, r, a_1, a_2, b_1, b_2.

- TRUNCATED FOURIER SERIES: $u_N(x) = \sum_{k=-N}^{N} \hat{u}_k e^{ikx}$
- The series involved $2N+1$ complex coefficients (weight $w \equiv 1$):

$$\hat{u}_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} u e^{-ikx}\, dx, \quad k = -N, \dots, N$$

- The expansion is redundant for real-valued u — the property of CONJUGATE SYMMETRY $\hat{u}_{-k} = \bar{\hat{u}}_k$, which reduces the number of complex coefficients to $N+1$; furthermore, $\Im(\hat{u}_0) \equiv 0$ for real u, thus one has $2N+1$ REAL coefficients; in the real case one can work with positive frequencies only!
- Equivalent real representation:

$$u_N(x) = a_0 + \sum_{k=1}^{N} \left[a_k \cos(kx) + b_k \sin(kx)\right],$$

where $a_0 = \hat{u}_0$, $a_k = 2\Re(\hat{u}_k)$ and $b_k = 2\Im(\hat{u}_k)$.

Uniform Convergence (I)

- Consider a function u that is smooth and periodic (with the period 2π); note the following two facts:
  - The Fourier coefficients are always less than the average of u

    $$|\hat{u}_k| = \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} u(x) e^{ikx}\, dx \right| \le M(u) \triangleq \frac{1}{2\pi} \int_{-\pi}^{\pi} |u(x)|\, dx$$

  - If $v = \frac{d^\alpha u}{dx^\alpha} = u^{(\alpha)}$, then $\hat{u}_k = \frac{\hat{v}_k}{(ik)^\alpha}$

Uniform Convergence (II)

- Then, using integration by parts, we have

$$\hat{u}_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} u(x)e^{-ikx}\,dx = \frac{1}{2\pi}\left[u(x)\frac{e^{-ikx}}{-ik}\right]_{-\pi}^{\pi} - \frac{1}{2\pi}\int_{-\pi}^{\pi} u'(x)\frac{e^{-ikx}}{-ik}\,dx$$

- Repeating integration by parts p times

$$\hat{u}_k = (-1)^p\frac{1}{2\pi}\int_{-\pi}^{\pi} u^{(p)}(x)\frac{e^{-ikx}}{(-ik)^p}\,dx \quad\Longrightarrow\quad |\hat{u}_k| \le \frac{M(u^{(p)})}{|k|^p}$$

Therefore, the more regular is the function u, the more rapidly its Fourier coefficients tend to zero as $|n| \to \infty$

Uniform Convergence (III)

- We have $|\hat{u}_k| \le \frac{M(u'')}{|k|^2} \implies \sum_{k\in\mathbb{Z}} |\hat{u}_k e^{ikx}| \le \hat{u}_0 + \sum_{n\neq 0} \frac{M(u'')}{n^2}$
  The latter series converges ABSOLUTELY
- Thus, if u is TWICE CONTINUOUSLY DIFFERENTIABLE and its first derivative is CONTINUOUS AND PERIODIC with period 2π, then its Fourier series $u_N = P_N u$ CONVERGES UNIFORMLY to u for $|N| \to \infty$
- SPECTRAL CONVERGENCE – if $\phi \in C_p^\infty(-\pi,\pi)$, then for all $\alpha > 0$ there exists a positive constant C_α such that $|\hat{\phi}_k| \le \frac{C_\alpha}{|n|^\alpha}$, i.e., for a function with an infinite number of smooth derivatives, the Fourier coefficients vanish faster than algebraically
- RATE OF DECAY of Fourier transform of a function $f : \mathbb{R} \to \mathbb{R}$ is determined by its SMOOTHNESS ; functions defined on a bounded (periodic) domain are a special case

Theorem (a collection of several related results, see also Trefethen (2000))

Let $u \in L_2(\mathbb{R})$ have Fourier transform $\hat{u}$.

- *If u has $p-1$ continuous derivatives in $L_2(\mathbb{R})$ for some $p \ge 0$ and a p-th derivative of bounded variation, then $\hat{u}(k) = \mathcal{O}(|k|^{-p-1})$ as $|k| \to \infty$,*
- *If u has infinitely many continuous derivatives in $L_2(\mathbb{R})$, then $\hat{u}(k) = \mathcal{O}(|k|^{-m})$ as $|k| \to \infty$ for* EVERY *$m \ge 0$ (the converse also holds)*
- *If there exist $a, c > 0$ such that u can be extended to an* ANALYTIC *function in the complex strip $|\Im(z)| < a$ with $\|u(\cdot + iy)\| \le c$ uniformly for all $y \in (-a, a)$, where $\|u(\cdot + iy)\|$ is the L_2 norm along the horizontal line $\Im(z) = y$, then $u_a \in L_2(\mathbb{R})$, where $u_a(k) = e^{a|k|}\hat{u}(k)$ (the converse also holds)*
- *If u can be extended to an* ENTIRE *function (i.e., analytic throughout the complex plane) and there exists $a > 0$ such that $|u(z)| = o(e^{a|z|})$ as $|z| \to \infty$ for all complex values $z \in \mathbb{C}$, the $\hat{u}$ has compact support contained in $[-a, a]$; that is $\hat{u}(k) = 0$ for all $|k| > a$ (the converse also holds)*

Radii of Convergence

- DARBOUX'S PRINCIPLE [see Boyd (2001)] — for all types of spectral expansions (and for ordinary power series), both the domain of convergence in the complex plane and the rate of convergence are controlled by the location and strength of the GRAVEST SINGULARITY in the complex plane ("singularities" in this context denote poles, fractional powers, logarithms and discontinuities of $f(z)$ or its derivatives)
- Thus, given a function $f : [0, 2\pi] \to \mathbb{R}$, the rate of convergence of its Fourier series is determined by the properties of its COMPLEX EXTENSION $F : \mathbb{C} \to \mathbb{C}$!!!
- Shapes of regions of convergence:
  - Taylor series — circular disk extending up to the nearest singularity
  - Fourier (and Hermite) series — horizontal strip extending vertically up to the nearest singularity
  - Chebyshev series — ellipse with foci at $x = \pm 1$ and extending up to the nearest singularity

Periodic Sobolev Spaces

- Let $H_p^r(I)$ be a PERIODIC SOBOLEV SPACE , i.e.,

  $$H_p^r(I) = \{u \ : \ u^{(\alpha)} \in L_2(I), \alpha = 0, \dots, r\},$$

  where $I = (-\pi, \pi)$ is a periodic interval. The space $C_p^\infty(I)$ is dense in $H_p^r(I)$
- The following two norms can be shown to be EQUIVALENT in H_p^r:

  $$\|u\|_r = \left[\sum_{k\in\mathbb{Z}} (1+k^2)^r |\hat{u}_k|^2\right]^{1/2}, \qquad |||u|||_r = \left[\sum_{\alpha=0}^{r} C_r^\alpha \|u^{(\alpha)}\|^2\right]^{1/2}$$

  Note that the first definition is naturally generalized for the case when r is non-integer!
- The PROJECTION OPERATOR P_N commutes with the derivative in the distribution sense:

  $$(P_N u)^{(\alpha)} = \sum_{|k|\le N} (ik)^\alpha \hat{u}_k W_k = P_N u^{(\alpha)}$$

Estimates of Approximation Error in $H_p^s(I)$

Theorem

Let $r, s \in \mathbb{R}$ *with* $0 \le s \le r$*; then we have:*

$$\|u - P_N u\|_s \le (1+N^2)^{\frac{s-r}{2}} \|u\|_r, \quad \text{for} \ \ u \in H_p^r(I)$$

Proof.

$$\|u - P_N u\|_s^2 = \sum_{|k|>N} (1+k^2)^{s-r+r} |\hat{u}_k|^2 \le (1+N^2)^{s-r} \sum_{|k|>N} (1+k^2)^r |\hat{u}_k|^2$$

$$\le (1+N^2)^{s-r} \|u\|_r^2$$

☐

- Thus, accuracy of the approximation $P_N u$ is better when u is SMOOTHER ; more precisely, for $u \in H_p^r(I)$, the L_2 leading order error is $\mathcal{O}(N^{-r})$ which improves when r increases.

Estimates of Approximation Error in $L_\infty(I)$

Lemma (Sobolev Inequality)

let $u \in H^1_p(I)$*, then there exists a constant* C *such that*

$$\|u\|^2_{L_\infty(I)} \le C\|u\|_0 \, \|u\|_1$$

Proof.

Suppose $u \in C^\infty_p(I)$; note the following facts

- $\hat{u}_0$ is the average of u
- From the mean value theorem: $\exists x_0 \in I$ such that $\hat{u}_0 = u(x_0)$

Let $v(x) = u(x) - \hat{u}_0$, then

$$\frac{1}{2}|v(x)|^2 = \int_{x_0}^{x} v(y)v'(y)\,dy \le \left(\int_{x_0}^{x} |v(y)|^2\,dy\right)^{1/2} \left(\int_{x_0}^{x} |v'(y)|^2\,dy\right)^{1/2} \le 2\pi\|v\|\,\|v'\|$$

$$|u(x)| \le |\hat{u}_0| + |v(x)| \le |\hat{u}_0| + 2\pi^{1/2}\|v\|^{1/2}\,\|v'\|^{1/2} \le C\|u\|_0^{1/2}\,\|u\|_1^{1/2},$$

since $v' = u'$, $\|v\| \le \|u\|$ and $|\hat{u}_0| \le \|u\|$.
As $C^\infty_p(I)$ is dense in $H^1_p(I)$, the inequality also holds for any $u \in H^1_p(I)$. □

Estimates of Approximation Error in $L_\infty(I)$

- An estimate in the norm $L_\infty(I)$ follows immediately from the previous lemma and estimates in the $H_p^s(I)$ norm

  $$\|u - P_N u\|^2_{L_\infty(I)} \le C(1+N^2)^{-\frac{r}{2}}(1+N^2)^{\frac{1-r}{2}}\|u\|_r,$$

  where $u \in H_p^r(I)$
- Thus for $r \ge 1$

  $$\|u - P_N u\|^2_{L_\infty(I)} = \mathcal{O}(N^{\frac{1}{2}-r})$$

- UNIFORM CONVERGENCE for all $u \in H_p^1(I)$
  (Note that u need only to be CONTINUOUS , therefore this result is stronger than the one given earlier)

- Assume we have a truncated Fourier series of $u(x)$
$$u_N(x) = P_N u(x) = \sum_{k=-N}^{N} \hat{u}_k e^{ikx}$$
- The Fourier series of the p–th derivative of $u(x)$ is
$$u_N^{(p)}(x) = P_N u^{(p)} = \sum_{k=-N}^{N} (ik)^p \hat{u}_k e^{ikx} = \sum_{k=-N}^{N} \hat{u}_k^{(p)} e^{ikx}$$
- Thus, using the vectors $\hat{U} = [\hat{u}_{-N}, \dots, \hat{u}_N]^T$ and $\hat{U}^{(p)} = [\hat{u}_{-N}^{(p)}, \dots, \hat{u}_N^{(p)}]^T$, one can introduce the SPECTRAL DIFFERENTIATION MATRIX $\mathcal{D}^{(p)}$ defined in Fourier space as $\hat{U}^{(p)} = \hat{\mathcal{D}}^{(p)} \hat{U}$, where
$$\hat{\mathcal{D}}^{(p)} = i^p \begin{bmatrix} -N^p & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & & \ddots & \\ & & & & N^p \end{bmatrix}$$

- Properties of the spectral differentiation matrix in Fourier representation
  - $\mathcal{D}^{(p)}$ is DIAGONAL
  - $\mathcal{D}^{(p)}$ is SINGULAR (diagonal matrix with a zero eigenvalue)
  - after desingularization the 2–norm condition number of $\mathcal{D}^{(p)}$ grows in proportion to N^p (since the matrix is diagonal, this is not an issue)
- QUESTION — how to derive the corresponding spectral differentiation matrix in REAL REPRESENTATION ?

Will see shortly ...

- Let's return to the Spectral Galerkin Method
- We need to evaluate the expansion (Fourier) coefficients

$$\hat{u}_k = (u, \phi_k)_w = \int_a^b w(x) u(x) \phi_k(x)\, dx, \quad k = 0, \dots, N$$

- QUADRATURE is a method to evaluate such integrals approximately.
- GAUSSIAN QUADRATURE seeks to obtain the best numerical estimate of an integral $\int_a^b w(x)\, f(x)\, dx$ by picking OPTIMAL POINTS x_i, $i = 1, \dots, N$ at which to evaluate the function $f(x)$.

Theorem (Gauß–Jacobi Integration Theorem)

If $(N+1)$ *interpolation points* $\{x_i\}_{i=0}^N$ *are chosen to be the zeros of* $P_{N+1}(x)$*, where* $P_{N+1}(x)$ *is the polynomial of degree* $(N+1)$ *of the set of polynomials which are orthogonal on* $[a,b]$ *with respect to the weight function* $w(x)$*, then the quadrature formula*

$$\int_a^b w(x)f(x)\,dx = \sum_{i=0}^{N} w_i f(x_i)$$

is EXACT *for all* $f(x)$ *which are polynomials of at most degree* $(2N+1)$

Definition

Let K be a non-empty, Lipschitz, compact subset of $\mathbb{R}^d$. Let $l_q \geq 1$ be an integer. A quadrature on K with l_q points consists of:

- A set of l_q real numbers $\{\omega_1, \dots, \omega_{l_q}\}$ called QUADRATURE WEIGHTS
- A set of l_q points $\{\xi_1, \dots, \xi_{l_q}\}$ in K called GAUSS POINTS or QUADRATURE NODES

The largest integer k such that $\forall p \in P_k$, $\int_K p(x)\,dx = \sum_{l=1}^{l_q} \omega_l p(\xi_l)$ is called the quadrature order and is denoted by k_q

- REMARK — As regards 1D bounded intervals, the most frequently used quadratures are based on Legendre polynomials which are defined on the interval $(0,1)$ as $\mathcal{E}_k(t) = \frac{1}{k!}\frac{d^k}{dt^k}(t^2-t)^k$, $k \geq 0$. Note that they are orthogonal on $(0,1)$ with the weight $w = 1$.

Theorem

Let $l_q \geq 1$, denote by $\xi_1, \dots, \xi_{l_q}$ the l_q roots of the Legendre polynomial $\mathcal{L}_{l_q}(x)$ and set $\omega_l = \int_0^1 \prod_{\substack{j=1 \\ j\neq l}}^{l_q} \frac{t-\xi_j}{\xi_l - \xi_j}\, dt$. Then $\{\xi_1, \dots, \xi_{l_q}, \omega_1, \dots, \omega_{l_q}\}$ is a quadrature of order $k_q = 2l_q - 1$ on $[0,1]$.

Proof.

Let $\{\mathcal{L}_1, \dots, \mathcal{L}_{l_q}\}$ be the set of Lagrange polynomials associated with the Gauß points $\{\xi_1, \dots, \xi_{l_q}\}$. Then $\omega_l = \int_0^1 \mathcal{L}_l(t)\, dt$, $1 \leq l \leq l_q$

- when $p(x)$ is a polynomial of degree less than l_q, we integrate both sides of the identity $p(t) = \sum_{l=1}^{l_q} p(\xi_l)\mathcal{L}_l(t)\, dx$, $\forall t \in [0,1]$ and deduce that the quadrature is exact for $p(x)$
- when the polynomial $p(x)$ has degree less than $2l_q$ we write it in the form $p(x) = q(x)\mathcal{L}_{l_q}(x) + r(x)$, where both $q(x)$ and $r(x)$ are polynomials of degree less than l_q; owing to orthogonality of the Legendre polynomials, we conclude
$$\int_0^1 p(t)\, dt = \int_0^1 r(t)\, dt = \sum_{l=1}^{l_q} \omega_l r(\xi_l) = \sum_{l=1}^{l_q} \omega_l p(\xi_l),$$
since the points ξ_l are also roots of $\mathcal{L}_{l_q}$

- PERIODIC GAUSSIAN QUADRATURE — If the interval $[a, b] = [0, 2\pi]$ is periodic, the weight $w(x) \equiv 1$ and $P_N(x)$ is the trigonometric polynomial of degree N, the Gaussian quadrature is equivalent to the TRAPEZOIDAL RULE (i.e., the quadrature with unit weights and equispaced nodes)
- Evaluation of the spectral coefficients:
  - Assume $\{\phi\}_{k=1}^{N}$ is a set of basis functions orthogonal under the weight w

    $$\hat{u}_k = \int_a^b w(x)u(x)\phi_k(x)\,dx \cong \sum_{i=0}^{N} w_i u(x_i)\phi_k(x_i), \quad k = 0, \dots, N,$$

    where x_i are chosen so that $\phi_{N+1}(x_i) = 0$, $i = 0, \dots, N$
  - Denoting $\hat{U} = [\hat{u}_0, \dots, \hat{u}_N]^T$ and $U = [u(x_0), \dots, u(x_N)]^T$ we can write the above as

    $$\hat{U} = \mathbb{T}U,$$

    where $\mathbb{T}$ is a TRANSFORMATION MATRIX