Appendix A
Properties of the Gamma Functions

We list here some basic properties of the Gamma function (see, e.g., Abramowitz and Stegun (1964)), defined by

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt, \quad \forall z \in \mathbb{C} \text{ with } \mathrm{Re}(z) > 0. \tag{A.1}$$

In particular, we have $\Gamma(1) = 1$ and $\Gamma(1/2) = \sqrt{\pi}$.

- Recursion formula:

$$\Gamma(z+1) = z\Gamma(z), \quad \Gamma(n+1) = n!, \tag{A.2}$$

and

$$\Gamma(2z) = (2\pi)^{-1/2} 2^{2z-1/2} \Gamma(z)\Gamma(z+1/2). \tag{A.3}$$

- Connection with binomial coefficient:

$$\binom{z}{w} = \frac{\Gamma(z+1)}{\Gamma(w+1)\Gamma(z-w+1)}. \tag{A.4}$$

- Relation with Beta function:

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1} dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}, \ \ x,y > 0. \tag{A.5}$$

In particular, for $\alpha, \beta > -1$,

$$\begin{aligned} \int_{-1}^1 (1-x)^\alpha (1+x)^\beta dx &= 2^{\alpha+\beta+1} \int_0^1 t^\beta (1-t)^\alpha dt \\ &= 2^{\alpha+\beta+1} \frac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta+2)}. \end{aligned} \tag{A.6}$$

J. Shen et al., *Spectral Methods: Algorithms, Analysis and Applications*,
Springer Series in Computational Mathematics 41, DOI 10.1007/978-3-540-71041-7,

- Stirling's formula:

$$\Gamma(x) = \sqrt{2\pi} x^{x-1/2} e^{-x} \Big\{ 1 + \frac{1}{12x} + \frac{1}{288x^2} + O(x^{-3}) \Big\}, \quad x \gg 1. \tag{A.7}$$

Moreover, we have

$$\sqrt{2\pi} n^{n+1/2} < n! e^n < \sqrt{2\pi} n^{n+1/2} \Big(1 + \frac{1}{4n} \Big), \quad n \geq 1. \tag{A.8}$$

Appendix B
Essential Mathematical Concepts

We provide here some essential mathematical concepts which have been used in the mathematical analysis throughout the book. For a more comprehensive presentation, we refer to Yosida (1980) and Adams (1975).

Let $(X;d)$ be a metric space. A sequence $\{x_k\}$ in X is called a *Cauchy sequence*, if

$$d(x_k,x_l)\to 0 \quad \text{as} \quad k,l\to\infty.$$

The space $(X;d)$ is said to be a *complete space* if every Cauchy sequence in X converges to an element in X.

B.1 Banach Space

Definition B.1. *Given a (real) vector space X, a* norm *on X is a function $\|\cdot\|: X\to\mathbb{R}$ such that*

- $\|u+v\|\le\|u\|+\|v\|, \quad \forall u,v\in X;$
- $\|\alpha u\|=|\alpha|\|u\|, \quad \forall u\in X \text{ and } \forall\alpha\in\mathbb{R};$
- $\|u\|\ge 0, \quad \forall u\in X;$
- $\|u\|=0$ *if and only if* $u=0.$

In particular, a semi-norm *on X is a function $|\cdot|: X\to\mathbb{R}$ satisfying the first three conditions.*

The space $(X,\|\cdot\|)$ is called a normed vector space. A Banach space is a normed vector space which is complete with respect to the metric:

$$d(u,v)=\|u-v\|, \quad \forall u,v\in X.$$

J. Shen et al., *Spectral Methods: Algorithms, Analysis and Applications*,
Springer Series in Computational Mathematics 41, DOI 10.1007/978-3-540-71041-7,

B.2 Hilbert Space

Definition B.2. *Let X be a real vector space. An inner product on X is a function $(u,v) : X \times X \to \mathbb{R}$ such that*

- $(u,v) = (v,u), \quad \forall u,v \in X;$
- $(\alpha u + \beta v, w) = \alpha(u,w) + \beta(v,w), \quad \forall u,v,w \in X$ *and* $\forall \alpha,\beta \in \mathbb{R};$
- $(u,u) \geq 0, \quad \forall u \in X;$
- $(u,u) = 0$ *if and only if* $u = 0.$

Two elements $u,v \in X$ are said to be *orthogonal* in X, if $(u,v) = 0$. The inner product $(\cdot,\cdot)$ induces a *norm* on X, given by

$$\|u\| = \sqrt{(u,u)}, \quad \forall u \in X.$$

Correspondingly, the metric on X can be defined by $d(u,v) = \|u - v\|$.

A *Hilbert space* is a Banach space endowed with an inner product (i.e., every Cauchy sequence in X is convergent with respect to the induced norm).

In a Hilbert space, the Cauchy–Schwarz inequality holds:

$$|(u,v)| \leq \|u\|\|v\|, \quad \forall u,v \in X. \tag{B.1}$$

Remark B.1. *If X is a complex vector space, the inner product (u,v) is a complex valued function. In the Definition B.2, the first condition should be replaced by*

$$(u,v) = \overline{(v,u)}, \quad \forall u,v \in X.$$

Next, we introduce the dual space of a Banach/Hilbert space X.

Definition B.3. *A functional $F : X \to \mathbb{R}$ is said to be* linear *or* continuous*, if there exists a constant $c > 0$ such that*

$$|F(u)| \leq c\|u\|, \quad \forall u \in X. \tag{B.2}$$

Let X' be the set of all linear functionals on X, and define the norm

$$\|F\|_{X'} = \sup_{u \in X; u \neq 0} \frac{|F(u)|}{\|u\|}.$$

Then the space X' is a Banach space, which is called the *dual space* of X.

The bilinear form $F(u) = \langle F,u \rangle : X' \times X \to \mathbb{R}$, is called the *duality pairing* on $X' \times X$. If X is a Hilbert space, then its dual space X' is a Hilbert space as well. Moreover, according to the *Riesz Representation Theorem*, X and X' are isometric, and X' can be canonically identified to X. More precisely, for any linear functional $F \in X'$, there exists a unique $u \in X$ such that

$$F(v) = \langle F,v \rangle = (u,v), \quad \forall v \in X \quad \text{and} \quad \|F\|_{X'} = \|u\|.$$

In a normed space X, a sequence $\{v_n\} \subset X$ is (strongly) convergent to $v \in X$, if $\|v_n - v\| \to 0$ as $n \to \infty$. It is possible to introduce another type of convergence in a weaker sense.

Definition B.4. *A sequence $\{v_n\}$ in X is called weakly convergent to $v \in X$, if $F(v_n) \to F(v)$ in $\mathbb{R}$ for all $F \in X'$.*

If a sequence $\{v_n\}$ converges to v in X, it is also weakly convergent. The converse is not true unless X is a finite dimensional space.

In a dual space X', a sequence of functional $\{F_n\}$ in X' is called *weakly* convergent* to $F \in X'$, if $\{F_n(v)\}$ converges to $F(v)$ for all $v \in X$. The weak convergence implies the weak* convergence.

B.3 Lax-Milgram Lemma

Definition B.5. *Let X be a Hilbert space with norm $\|\cdot\|$. A functional $a(\cdot,\cdot) : X \times X \to \mathbb{R}$ defines a* bilinear form, *if for any $u, v, w \in X$ and $\alpha, \beta \in \mathbb{R}$,*

$$a(\alpha u + \beta v, w) = \alpha a(u, w) + \beta a(v, w),$$
$$a(u, \alpha v + \beta w) = \alpha a(u, w) + \beta a(v, w).$$

That is, for any fixed u, both the functionals $a(u,\cdot) : X \to \mathbb{R}$ and $a(\cdot,u) : X \to \mathbb{R}$ are linear. The bilinear form is symmetric, *if $a(u,v) = a(v,u)$ for any $u, v \in X$.*

Definition B.6. *A bilinear form $a(\cdot,\cdot)$ on a Hilbert space X is said to be* continuous, *if there exists a constant $C > 0$ such that*

$$|a(u,v)| \le C\|u\|\|v\|, \quad \forall u, v \in X, \tag{B.3}$$

and coercive *on X, if there exists a constant $\alpha > 0$ such that*

$$a(u,u) \ge \alpha\|u\|^2, \quad \forall u \in X. \tag{B.4}$$

It is clear that if $a(\cdot,\cdot)$ is symmetric, continuous and coercive on the Hilbert space X, then $a(\cdot,\cdot)$ defines an inner product on X.

Theorem B.1. (Lax-Milgram lemma). *Let X be a Hilbert space, let $a(\cdot,\cdot) : X \times X \to \mathbb{R}$ be a continuous and coercive bilinear form, and let $F : X \to \mathbb{R}$ be a linear functional in X'. Then the variational problem:*

$$\begin{cases} \text{Find } u \in X \text{ such that} \\ a(u,v) = F(v), \ \ \forall v \in X, \end{cases} \tag{B.5}$$

has a unique solution. Moreover, we have

$$\|u\| \le \frac{1}{\alpha}\|F\|_{X'}. \tag{B.6}$$

B.4 L^p-Space

Let Ω be a Lebesgue-measurable subset of $\mathbb{R}^d$ ($d = 1,2,3$) with non-empty interior, and let u be a Lebesgue measurable function on Ω. In what follows, the integration is in the Lebesgue sense.

Definition B.7. *For* $1 \leq p \leq \infty$, *let*

$$L^p(\Omega) := \big\{u \ : \ u \text{ is measurable on } \Omega \text{ and } \|u\|_{L^p(\Omega)} < \infty\big\}, \tag{B.7}$$

where for $1 \leq p < \infty$,

$$\|u\|_{L^p(\Omega)} := \Big(\int_\Omega |u(x)|^p dx\Big)^{1/p}, \tag{B.8}$$

and

$$\|u\|_{L^\infty(\Omega)} := \operatorname{ess\ sup}_{x\in\Omega} |u(x)|. \tag{B.9}$$

Remark B.2. *Some remarks are in order.*

(i) *The space $L^\infty(\Omega)$ consists all functions that are essentially bounded on Ω. A function u is said to be* essentially bounded *on Ω, if there exists a constant K such that $|u(x)| \leq K$ a.e. on Ω. The greatest lower bound of such constants K is called the* essential supremum *of $|u(x)|$ on Ω, denoted by $\|u\|_{L^\infty(\Omega)}$.*

(ii) *We identify functions in $L^p(\Omega)$ that are equal almost everywhere on Ω. The elements of $L^p(\Omega)$ are equivalence classes of measurable functions that satisfy* (B.7) *with the equivalence relation: $u \equiv v$, if they only differ on a measurable subset of measure zero.*

Equipped with the norm $\|\cdot\|_{L^p(\Omega)}$, the space $L^p(\Omega)$ with $1 \leq p \leq \infty$ is a Banach space. In particular, the space $L^2(\Omega)$ is a Hilbert space equipped with the inner product

$$(u,v)_{L^2(\Omega)} = \int_\Omega u(x)v(x)dx, \quad \forall u,v \in L^2(\Omega). \tag{B.10}$$

Definition B.8. *If p and q are positive real numbers such that*

$$p+q = pq \quad \text{or} \quad \frac{1}{p}+\frac{1}{q} = 1, \tag{B.11}$$

then we call (p,q) a pair of conjugate exponents. *As $p \to 1$, (B.11) forces $q \to \infty$. Consequently, $(1,\infty)$ is also regarded as a pair of conjugate exponents.*

Theorem B.2.

- **Minkowski's inequality**. *If $u,v \in L^p(\Omega)$ with $1 \leq p \leq \infty$, then $u+v \in L^p(\Omega)$, and*

$$\|u+v\|_{L^p(\Omega)} \leq \|u\|_{L^p(\Omega)} + \|v\|_{L^p(\Omega)}. \tag{B.12}$$

- **Hölder's inequality**. *Let p and q be conjugate exponents with $1 \le p \le \infty$. If $u \in L^p(\Omega)$ and $v \in L^q(\Omega)$, then $uv \in L^1(\Omega)$, and*

$$\int_\Omega |u(x)v(x)|dx \le \|u\|_{L^p(\Omega)} \|v\|_{L^q(\Omega)}. \tag{B.13}$$

In particular, if $p = 2$, the Hölder's inequality reduces to the Cauchy–Schwarz inequality (B.1).

It follows from (B.13) that $L^q(\Omega) \subset L^p(\Omega)$, if $p \le q$ and Ω has a finite measure.

As a final remark, given a weight function $\omega(x)$, which is almost everywhere positive and Lebesgue integrable on Ω, $\omega(x)dx$ also defines a Lebesgue measure on Ω. Replacing dx in (B.8) by $\omega(x)dx$, we define the norm $\|\cdot\|_{L^p_\omega(\Omega)}$ and the space $L^p_\omega(\Omega)$ with $1 \le p < \infty$, which is a Banach space. In particular, the space $L^2_\omega(\Omega)$ is a Hilbert space with the inner product and norm given by

$$(u,v)_\omega = \int_\Omega u(x)v(x)\omega(x)dx, \quad \|u\|_\omega = \sqrt{(u,u)_\omega}.$$

One verifies that the inequalities (B.12) and (B.13) hold in the weighted norms.

B.5 Distributions and Weak Derivatives

A multi-index $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$ is a d-tuple of non-negative integers $\{\alpha_i\}$. Denote $|\boldsymbol{\alpha}| = \sum_{i=1}^d \alpha_i$, and define the partial derivative operator

$$D^{\boldsymbol{\alpha}} = \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}}.$$

For any $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, we define $x^{\boldsymbol{\alpha}} = x_1^{\alpha_1} \ldots x_d^{\alpha_d}$.

For any $\Omega \subset \mathbb{R}^d$, let $\mathscr{D}(\Omega)$ (or $C_0^\infty(\Omega)$) be the set of all infinitely differentiable functions with compact support in Ω. With the aid of $\mathscr{D}(\Omega)$, we can extend the conventional derivatives to the notion of generalized (weak) derivatives.

We first recall the topology on the vector space $C_0^\infty(\Omega)$.

Definition B.9. *A sequence of functions $\{v_n\}$ in $C_0^\infty(\Omega)$ is said to be* convergent in the sense of $\mathscr{D}(\Omega)$ *to the function $v \in C_0^\infty(\Omega)$ provided that*

(i) *there exists $K \subset\subset \Omega$, which means the closure $\bar{K} \subset \Omega$ and $\bar{K}$ is compact (i.e., closed and bounded), such that the support of $v_n - v \subset K$ for every n;*
(ii) $\lim_{n\to\infty} D^{\boldsymbol{\alpha}} v_n(x) = D^{\boldsymbol{\alpha}} v(x)$ *uniformly on K for each multi-index $\boldsymbol{\alpha}$.*

Definition B.10. *The dual space $\mathscr{D}'(\Omega)$ of $\mathscr{D}(\Omega)$ is called the space of (Schwarz)* distributions. *A sequence of distributions $\{T_n\}$ in $\mathscr{D}'(\Omega)$ is called weakly* convergent to a distribution $T \in \mathscr{D}'(\Omega)$, if $T_n(v) \to T(v)$ in $\mathbb{R}$ for every $v \in \mathscr{D}(\Omega)$.*

For example, we consider the distributions induced by *locally integrable* functions.

Definition B.11. *Given a domain $\Omega \subset \mathbb{R}^d$, the set of all* locally integrable *functions is denoted by*

$$L^1_{\mathrm{loc}}(\Omega) = \big\{ u \,:\, u \in L^1(K),\ \forall \text{ compact } K \subset \text{interior } \Omega \big\}. \tag{B.14}$$

Corresponding to every $u \in L^1_{\mathrm{loc}}(\Omega)$, there is a distribution $T_u \in \mathscr{D}'(\Omega)$ defined by

$$T_u(v) = \int_\Omega u(x)v(x)dx, \quad \forall v \in \mathscr{D}(\Omega). \tag{B.15}$$

However, not every distribution $T \in \mathscr{D}'(\Omega)$ is of the form (B.15), and for instance, the Delta function δ is such a distribution.

Definition B.12. *Let T be a distribution in $\mathscr{D}'(\Omega)$, and let $\boldsymbol{\alpha}$ be a multi-index. Then $D^{\boldsymbol{\alpha}}T$ is also a distribution in $\mathscr{D}'(\Omega)$, defined as follows*

$$\langle D^{\boldsymbol{\alpha}}T, v\rangle = (-1)^{|\boldsymbol{\alpha}|}\langle T, D^{\boldsymbol{\alpha}}v\rangle, \quad \forall v \in \mathscr{D}(\Omega), \tag{B.16}$$

where $\langle \cdot,\cdot \rangle$ is the duality paring of $\mathscr{D}'(\Omega)$ and $\mathscr{D}(\Omega)$.

Notice that by definition, a distribution is infinitely differentiable. Moreover, if T is a smooth function, its generalized derivative coincides with the usual derivative.

Definition B.13. *A given function $u \in L^1_{\mathrm{loc}}(\Omega)$ has a weak derivative $D^{\boldsymbol{\alpha}}u$, if there exists a function $w \in L^1_{\mathrm{loc}}(\Omega)$ such that*

$$\int_\Omega w(x)v(x)dx = (-1)^{|\boldsymbol{\alpha}|}\int_\Omega u(x)D^{\boldsymbol{\alpha}}v(x)dx, \quad \forall v \in \mathscr{D}(\Omega). \tag{B.17}$$

If such a w exists, we define $D^{\boldsymbol{\alpha}}u = w$.

We can extend the above discussion to periodic distributions. Let $\Omega = (0, 2\pi)^d$, and define the space $C^\infty_p(\bar{\Omega})$ as the vector space of functions that are infinitely differentiable with all derivatives being 2π-periodic in each space direction. A sequence $\{\phi_n\}$ in $C^\infty_p(\bar{\Omega})$ converges to a function ϕ in $C^\infty_p(\bar{\Omega})$, if $D^{\boldsymbol{\alpha}}\phi_n \to D^{\boldsymbol{\alpha}}\phi$ uniformly on $\bar{\Omega}$ for every multi-index $\boldsymbol{\alpha}$. Similarly, a *periodic distribution* is a continuous linear form $T : C^\infty_p(\bar{\Omega}) \to \mathbb{C}$, that is, $\langle T, \phi_n\rangle \to \langle T, \phi\rangle$ in $\mathbb{C}$ whenever $\phi_n \to \phi$ in $C^\infty_p(\bar{\Omega})$. The derivative of a periodic distribution T can be defined by (B.16) with $C^\infty_p(\bar{\Omega})$ in place of $\mathscr{D}(\Omega)$.

B.6 Sobolev Spaces

Using the notion of weak derivatives, we define the Sobolev spaces on the L^p-spaces. Such spaces are most often used for the variational theory of partial differential

equations. In what follows, we restrict the discussions to the Hilbert spaces (i.e., with $p = 2$), and refer to Adams (1975) for a comprehensive presentation of general Sobolev spaces.

Definition B.14. *The Sobolev space $H^m(\Omega)$ with $m \in \mathbb{N}$ is the space of functions $u \in L^2(\Omega)$ such that all the distributional derivatives of order up to m can be represented by functions in $L^2(\Omega)$. That is,*

$$H^m(\Omega) = \left\{ u \in L^2(\Omega) \ : \ D^{\boldsymbol{\alpha}} u \in L^2(\Omega) \text{ for } 0 \le |\boldsymbol{\alpha}| \le m \right\}, \tag{B.18}$$

equipped with the norm and semi-norm

$$\|u\|_{m,\Omega} = \Big(\sum_{|\boldsymbol{\alpha}|=0}^{m} \|D^{\boldsymbol{\alpha}} u\|^2_{L^2(\Omega)} \Big)^{1/2}, \ \ |u|_{m,\Omega} = \Big(\sum_{|\boldsymbol{\alpha}|=m} \|D^{\boldsymbol{\alpha}} u\|^2_{L^2(\Omega)} \Big)^{1/2}. \tag{B.19}$$

The space $H^m(\Omega)$ is a Hilbert space endowed with the inner product

$$(u, v)_{m,\Omega} = \sum_{|\boldsymbol{\alpha}|=0}^{m} \int_{\Omega} D^{\boldsymbol{\alpha}} u(x) D^{\boldsymbol{\alpha}} v(x) dx.$$

The following density property holds (see, e.g., Brenner and Scott (2008)).

Theorem B.3. *For any $\Omega \in \mathbb{R}^d$, $C^\infty(\bar{\Omega})$ is dense in $H^m(\Omega)$ for any integer $m \ge 0$.*

Definition B.15. *For any positive integer m, the space $H_0^m(\Omega)$ is the closure of $C_0^\infty(\Omega)$ with respect to the norm $\|\cdot\|_{m,\Omega}$. The dual space of $H_0^m(\Omega)$ is denoted by $H^{-m}(\Omega)$ with the norm*

$$\|u\|_{-m,\Omega} = \sup_{0 \neq v \in H_0^m(\Omega)} \frac{\langle u, v \rangle}{\|v\|_{m,\Omega}}. \tag{B.20}$$

We have the following *Poincaré-Friedrichs inequality* (see, e.g., Ciarlet (1978)).

Theorem B.4. *Let Ω be a bounded open subset of $\mathbb{R}^d$. Then there exists a positive constant $c(\Omega)$ such that*

$$\|u\|_{0,\Omega} \le c(\Omega) |u|_{1,\Omega}, \quad \forall u \in H_0^1(\Omega), \tag{B.21}$$

which implies that the semi-norm $|\cdot|_{m,\Omega}$ is a norm of $H_0^m(\Omega)$, equivalent to the norm $\|\cdot\|_{m,\Omega}$.

For any real $r > 0$, the Sobolev space $H^r(\Omega)$ can be defined by space interpolation (see, e.g., Bergh and Löfström (1976), Adams (1975) and Lions and Magenes (1968)).

An important property of the Sobolev spaces is the *embedding result*, which indicates the close connection with continuous functions.

Theorem B.5. *Let Ω be a domain in $\mathbb{R}^d$ with Lipschitz boundary $\partial\Omega$, and let $r > n/2$. Then there exists a positive constant C such that*

$$\|u\|_{L^\infty(\Omega)} \leq C\|u\|_{r,\Omega}. \tag{B.22}$$

Moreover, there is a continuous function in the $L^\infty(\Omega)$ equivalence class of u.

Another important result of Sobolev spaces is the so-called *trace theorem*. The trace of a function $u \in H^r(\Omega)$ on the boundary $\partial\Omega$ is meaningful by defining it as the restriction of $\tilde{u}$ on the boundary, where $\tilde{u} \in C^0(\bar{\Omega})$ is among the equivalence class of u.

Theorem B.6. *Let Ω be a domain in $\mathbb{R}^d$ with Lipschitz boundary $\partial\Omega$, and $r > 1/2$.*

(i) *There exits a unique linear continuous map $\gamma_0 : H^r(\Omega) \to H^{r-1/2}(\partial\Omega)$ such that $\gamma_0 v = v|_{\partial\Omega}$ for each $v \in H^r(\Omega) \cap C^0(\bar{\Omega})$.*

(ii) *There exists a linear continuous map $\tau_0 : H^{r-1/2}(\partial\Omega) \to H^r(\Omega)$ such that $\gamma_0\tau_0\phi = \phi$ for each $\phi \in H^{r-1/2}(\partial\Omega)$.*

Analogous results also hold if we consider the trace γ_Γ over a Lipschitz continuous subset Γ of the boundary $\partial\Omega$.

We see that any function in $H^{r-1/2}(\partial\Omega), r > 1/2$, is the trace on $\partial\Omega$ of a function in $H^r(\Omega)$. This provides a characterization of the space $H^{r-1/2}(\partial\Omega)$. In particular, the proceeding theorem indicates that there exists a positive constant c such that

$$\|v\|_{L^2(\partial\Omega)} \leq c\|v\|_{1,\Omega}, \quad \forall v \in H^1(\Omega). \tag{B.23}$$

In addition, if the boundary $\partial\Omega$ is Lipschitz continuous, we can characterize the space $H_0^1(\Omega)$ by

$$H_0^1(\Omega) = \{v \in H^1(\Omega) \ : \ \gamma_0 v = 0\}. \tag{B.24}$$

If Γ is part of $\partial\Omega$, we define

$$H_\Gamma^1(\Omega) = \{v \in H^1(\Omega) \ : \ \gamma_\Gamma v = 0\}. \tag{B.25}$$

It is worthwhile to point out that the Poincaré-Friedrichs inequality (B.21) is also valid for functions in $H_\Gamma^1(\Omega)$, provided that Γ is non-empty.

The *interpolation theorem* is also found useful in the analysis.

Theorem B.7. *Assume that Ω is an open subset of $\mathbb{R}^d$ with a Lipschitz boundary $\partial\Omega$. Let $r_1 < r_2$ be two real numbers and set $r = (1-\theta)r_1 + \theta r_2$ with $0 \leq \theta \leq 1$. Then there exists a constant $c > 0$ such that*

$$\|u\|_{r,\Omega} \leq c\|u\|_{r_1,\Omega}^{1-\theta}\|u\|_{r_2,\Omega}^{\theta}, \quad \forall u \in H^{r_2}(\Omega). \tag{B.26}$$

In the Definition B.14 of the Sobolev space, one can require the functions as well as its distributional derivatives to be square integrable with respect to the measure $\omega(x)dx$ on Ω. This provides a natural framework to deal with the Chebyshev and

Jacobi spectral methods. In a very similar fashion, we define the space $H^m_\omega(\Omega)$, the norm $\|\cdot\|_{m,\omega,\Omega}$ and the semi-norm $|\cdot|_{m,\omega,\Omega}$ by replacing the L^2-space and the norm $\|\cdot\|_{L^2(\Omega)}$ in (B.18) and (B.19) by the weighted L^2-space and the weighted norm $\|\cdot\|_{L^2_\omega(\Omega)}$, respectively. For real $r>0$, the space $H^r_\omega(\Omega)$ is defined by space interpolation as usual. Moreover, its subspace $H^r_{0,\omega}(\Omega)$ can be defined as the closure of $C^\infty_0(\Omega)$ in $H^r_\omega(\Omega)$ as before.

In the analysis of Fourier methods, it is necessary to define Sobolev space of periodic functions. In this framework, the functions are complex-valued and their weak derivatives are in the sense of periodic distribution. In particular, for $\Omega = (0,2\pi)$ and for any integer $m>0$,

$$H^m_p(0,2\pi) = \big\{u \in H^m(0,2\pi) \,:\, u^{(k)}(0) = u^{(k)}(2\pi),\ 0 \le k \le m-1\big\}. \tag{B.27}$$

In this context, the norm can be characterized by the Fourier coefficients of the underlying function (cf. (2.46) in Chap. 2).

B.7 Integral Identities: Divergence Theorem and Green's Formula

We collect some integral identities of advanced calculus in the setting of Sobolev spaces, which are useful in the formulation of multi-dimensional variational problems.

In what follows, let $\boldsymbol{u} = (u_1,\ldots,u_d)$ be a vector, and let $\nabla = (\partial_{x_1},\ldots,\partial_{x_d})$ be the d-dimensional gradient operator. Assume that Ω is a domain with Lipschitz boundary, and $\boldsymbol{\nu}$ denotes its unit outward normal to $\partial\Omega$.

Lemma B.1. (Divergence Theorem). *Let $\boldsymbol{u}$ be a Lebesgue integrable function on Ω. Then*

$$\int_\Omega \nabla\cdot\boldsymbol{u}\,dx = \int_{\partial\Omega} \boldsymbol{u}\cdot\boldsymbol{\nu}\,d\gamma. \tag{B.28}$$

Applying (B.28) with $\boldsymbol{u} = uw\boldsymbol{e}_i$ with $\boldsymbol{e}_i$ is the ith unit coordinate vector, leads to the following identity.

Lemma B.2. (Green's formula). *Let $u,w \in H^1(\Omega)$. Then for $i=1,\ldots,n$,*

$$\int_\Omega \partial_{x_i} uw\,dx = -\int_\Omega u\partial_{x_i} w\,dx + \int_{\partial\Omega} uw\nu_i\,d\gamma. \tag{B.29}$$

In a vector form, if $\boldsymbol{u} \in (H^1(\Omega))^d$ and $w \in H^1(\Omega)$, we have

$$\int_\Omega \nabla\cdot\boldsymbol{u}w\,dx = -\int_\Omega \boldsymbol{u}\cdot\nabla w\,dx + \int_{\partial\Omega} \boldsymbol{u}\cdot\boldsymbol{\nu}\,w\,d\gamma. \tag{B.30}$$

Applying (B.30) with $\boldsymbol{u} = \nabla\phi$ yields the following formulas.

Lemma B.3. *Let $\phi \in H^2(\Omega)$ and $w \in H^1(\Omega)$. Then*

$$\int_\Omega (-\Delta\phi) w \, dx = \int_\Omega \nabla\phi \cdot \nabla w \, dx - \int_{\partial\Omega} \partial_{\boldsymbol{\nu}}\phi \, w \, d\gamma. \tag{B.31}$$

If, in addition, $w \in H^2(\Omega)$, we have

$$\int_\Omega \big(w\Delta\phi - \phi\Delta w\big) dx = \int_{\partial\Omega} \big(w\partial_{\boldsymbol{\nu}}\phi - \phi\partial_{\boldsymbol{\nu}} w\big) d\gamma. \tag{B.32}$$

B.8 Some Useful Inequalities

We present below some useful embedding inequalities on finite/infinite intervals.

B.8.1 Sobolev-Type Inequalities

Let (a,b) be a finite interval. There holds the *Sobolev inequality*:

$$\max_{x\in[a,b]} |u(x)| \le \Big(\frac{1}{b-a} + 2\Big)^{1/2} \|u\|^{1/2}_{L^2(a,b)} \|u\|^{1/2}_{H^1(a,b)}, \quad \forall u \in H^1(a,b), \tag{B.33}$$

which is also known as the *Gagliardo-Nirenberg interpolation inequality*.

This inequality may take the form.

Lemma B.4. *For any $u \in H^1(a,b)$,*

$$\max_{x\in[a,b]} |u(x)| \le \frac{1}{\sqrt{b-a}} \|u\|_{L^2(a,b)} + \sqrt{b-a}\|u'\|_{L^2(a,b)}. \tag{B.34}$$

Proof. For any $x_1, x_2 \in [a,b]$,

$$|u(x_1) - u(x_2)| \le \int_{x_1}^{x_2} |u'(x)| dx \le \sqrt{b-a}\|u'\|_{L^2(a,b)},$$

which implies $u \in C[a,b]$. Denote $|u(x_*)| = \min_{x\in[a,b]} |u(x)|$, and we have

$$|u(x)| - |u(x_*)| \le \sqrt{b-a}\|u'\|_{L^2(a,b)}.$$

Moreover,

$$|u(x_*)| \le \frac{1}{b-a}\int_a^b |u(x)| dx \le \frac{1}{\sqrt{b-a}} \|u\|_{L^2(a,b)}.$$

A combination of the above two inequalities leads to (B.34). □

Lemma B.5. *Let $\omega = e^{-x}$. For any $u \in H^1_\omega(0,\infty)$ with $u(0) = 0$, we have*

$$\|e^{-x/2}u\|_{L^\infty(0,\infty)} \le \sqrt{2}\|u\|_\omega^{1/2}|u|_{1,\omega}^{1/2}, \tag{B.35a}$$

$$\|u\|_\omega \le 2|u|_{1,\omega}. \tag{B.35b}$$

Proof. Since $u(0) = 0$, we have

$$\begin{aligned} e^{-x}u^2(x) &= \int_0^x \partial_y(e^{-y}u^2(y))dy \\ &= 2\int_0^x e^{-y}u(y)u'(y)dy - \int_0^x e^{-y}u^2(y)dy, \quad \forall x \in (0,\infty), \end{aligned}$$

from which we derive

$$\begin{aligned} e^{-x}u^2(x) + \int_0^x e^{-y}u^2(y)dy &\le 2\int_0^\infty e^{-y}|u(y)u'(y)|dy \\ &\le 2\|u\|_\omega|u|_{1,\omega}. \end{aligned}$$

This implies the first inequality, and letting $x \to \infty$ leads to the second one. □

Lemma B.6. *Let $\omega(x) = e^{-x^2}$. Then for any $u \in H^1_\omega(-\infty,\infty)$, we have*

$$\|e^{-x^2/2}u\|_{L^\infty(-\infty,\infty)} \le 2\|u\|_\omega^{1/2}|u|_{1,\omega}^{1/2}, \tag{B.36a}$$

$$\|xu\|_\omega \le \|u\|_{1,\omega}. \tag{B.36b}$$

Proof. Applying integration by parts and the Schwarz inequality yields

$$\int_{-\infty}^\infty xu^2(x)\omega(x)dx = \int_{-\infty}^\infty u(x)u'(x)\omega(x)dx \le \|u\|_\omega|u|_{1,\omega}, \tag{B.37}$$

which implies $xu^2(x)\omega(x) \to 0$ as $|x| \to \infty$. Therefore, we have

$$\begin{aligned} \int_{-\infty}^\infty (xu(x))^2\omega(x)dx &= -\frac{1}{2}\int_{-\infty}^\infty xu^2(x)d\omega(x) \\ &= \frac{1}{2}\int_{-\infty}^\infty u^2(x)\omega(x)dx + \int_{-\infty}^\infty xu(x)u'(x)\omega(x)dx \\ &\le \frac{1}{2}\|u\|_\omega^2 + \frac{1}{2}\|xu\|_\omega^2 + \frac{1}{2}|u|_{1,\omega}^2 \\ &= \frac{1}{2}\|u\|_{1,\omega}^2 + \frac{1}{2}\|xu\|_\omega^2, \end{aligned}$$

which gives (B.36b).

Next, since

$$\begin{aligned} e^{-x^2}u^2(x) &= \int_{-\infty}^x \partial_y(e^{-y^2}u^2(y))dy \\ &= 2\int_{-\infty}^x u(y)u'(y)\omega(y)dy - 2\int_{-\infty}^x yu^2(y)\omega(y)dy, \end{aligned}$$

we deduce from the Schwarz inequality that

$$e^{-x^2}u^2(x)+2\int_{-\infty}^{x} yu^2(y)\omega(y)dy \le 2\|u\|_\omega |u|_{1,\omega}.$$

Thus, by (B.37),

$$e^{-x^2}u^2(x) \le 4\|u\|_\omega |u|_{1,\omega},$$

which yields (B.36a). □

B.8.2 Hardy-Type Inequalities

Let $a<b$ be two real numbers, and let $\alpha<1$. Then for any $\phi \in L^2_\omega(a,b)$ with $\omega=(x-a)^\alpha$, we have the following Hardy inequality (see Hardy et al. (1952)):

$$\int_a^b \Big(\frac{1}{x-a}\int_a^x \phi(y)dy\Big)^2 (x-a)^\alpha dt \le \frac{4}{1-\alpha}\int_a^b \phi^2(x)(x-a)^\alpha dx. \tag{B.38}$$

Similarly, for any $\phi \in L^2_\omega(a,b)$ with $\omega=(b-x)^\alpha$, we have

$$\int_a^b \Big(\frac{1}{b-x}\int_x^b \phi(y)dy\Big)^2 (b-x)^\alpha dt \le \frac{4}{1-\alpha}\int_a^b \phi^2(x)(b-x)^\alpha dx. \tag{B.39}$$

Next, we apply the above Hardy inequality to derive some useful inequalities associated with the Jacobi weight function $\omega^{\alpha,\beta}(x)=(1-x)^\alpha(1+x)^\beta$ with $x\in I:=(-1,1)$.

Lemma B.7. *If* $-1<\alpha,\beta<1$, *then*

$$\|u\|_{\omega^{\alpha-2,\beta-2}} \le c\|u'\|_{\omega^{\alpha,\beta}}, \quad \forall u\in H^1_{0,\omega^{\alpha,\beta}}(I), \tag{B.40}$$

which implies the Poincaré-type inequality:

$$\|u\|_{\omega^{\alpha,\beta}} \le c\|u'\|_{\omega^{\alpha,\beta}}, \quad \forall u\in H^1_{0,\omega^{\alpha,\beta}}(I). \tag{B.41}$$

Proof. Taking $a=-1, b=1$ and $\phi=\partial_x u$ in (B.39) yields that for $\alpha<1$,

$$\int_0^1 u^2(x)(1-x)^{\alpha-2}dx \le c\int_0^1 (u'(x))^2(1-x)^\alpha dx.$$

Hence,

$$\begin{aligned}\int_0^1 u^2(x)(1-x)^{\alpha-2}(1+x)^{\beta-2}dx &\le c\int_0^1 u^2(x)(1-x)^{\alpha-2}dx\\ &\le c\int_0^1 (u'(x))^2(1-x)^\alpha dx \le c\int_0^1 (u'(x))^2(1-x)^\alpha(1+x)^\beta dx.\end{aligned}$$

Similarly, for $\beta < 1$, we use (B.39) to derive

$$\int_{-1}^{0} u^2(x)(1-x)^{\alpha-2}(1+x)^{\beta-2}dx \le c\int_{-1}^{0}(u'(x))^2(1-x)^{\alpha}(1+x)^{\beta}dx.$$

A combination of the above two inequalities leads to (B.40).

In view of $\omega^{\alpha,\beta}(x) < \omega^{\alpha-2,\beta-2}(x)$, (B.41) follows from (B.40). □

A consequence of Lemma B.7 is the following result.

Corollary B.1. *If $-1 < \alpha, \beta < 1$, then for any $u \in H^1_{0,\omega^{\alpha,\beta}}(I)$, we have $u\omega^{\alpha,\beta} \in H^1_{0,\omega^{-\alpha,-\beta}}(I)$.*

Proof. A direct calculation shows that

$$|u\omega^{\alpha,\beta}|^2_{1,\omega^{-\alpha,-\beta}} \le c\big(|u|^2_{1,\omega^{\alpha,\beta}} + \|u\|^2_{\omega^{\alpha-2,\beta-2}}\big) \overset{(B.40)}{\le} c|u|^2_{1,\omega^{\alpha,\beta}}.$$

On the other hand, since $\|u\omega^{\alpha,\beta}\|_{\omega^{-\alpha,-\beta}} = \|u\|_{\omega^{\alpha,\beta}}$, we have

$$\|u\omega^{\alpha,\beta}\|_{1,\omega^{-\alpha,-\beta}} \le c\|u\|_{1,\omega^{\alpha,\beta}}.$$

This ends the proof. □

The following inequalities can be found in Guo (2000).

Lemma B.8. *Let $\alpha, \beta > -1$. Then for any function $u \in H^1_{\omega^{\alpha+2,\beta+2}}(I)$ with $u(x_0) = 0$ for some $x_0 \in (-1,1)$, we have*

$$\|u\|_{\omega^{\alpha,\beta}} \le c\|u'\|_{\omega^{\alpha+2,\beta+2}}, \tag{B.42}$$

which implies

$$\|u\|_{\omega^{\alpha,\beta}} \le c\|u'\|_{\omega^{\alpha,\beta}}. \tag{B.43}$$

Proof. The inequality (B.43) follows directly from (B.42), so it suffices to prove the first one. For any $x \in [x_0, 1]$,

$$u^2(x)(1-x)^{\alpha+1} = \int_{x_0}^{x} \partial_y\big(u^2(y)(1-y)^{\alpha+1}\big)dy.$$

Hence, by the Schwarz inequality,

$$\begin{aligned} &u^2(x)(1-x)^{\alpha+1} + (\alpha+1)\int_{x_0}^{x} u^2(y)(1-y)^{\alpha}dy \\ &\quad = 2\int_{x_0}^{x} u(y)u'(y)(1-y)^{\alpha+1}dy \\ &\quad \le 2\Big(\int_{x_0}^{x} u^2(y)(1-y)^{\alpha}dy\Big)^{1/2}\Big(\int_{x_0}^{x}\big(u'(y)\big)^2(1-y)^{\alpha+2}dy\Big)^{1/2}, \end{aligned}$$

which implies

$$\int_{x_0}^{x} u^2(y)(1-y)^{\alpha}dy \le \frac{4}{(\alpha+1)^2}\int_{x_0}^{x}\big(u'(y)\big)^2(1-y)^{\alpha+2}dy.$$

Letting $x \to 1$ in the above inequality leads to

$$\begin{aligned}\int_{x_0}^{1} u^2(x)(1-x)^{\alpha}dx &\le \frac{4}{(\alpha+1)^2}\int_{x_0}^{1}\big(u'(x)\big)^2(1-x)^{\alpha+2}dx\\ &\le c\int_{x_0}^{1}\big(u'(x)\big)^2(1-x)^{\alpha+2}(1+x)^{\beta+2}dx.\end{aligned}$$

Similarly, we can derive

$$\int_{-1}^{x_0} u^2(x)(1+x)^{\beta}dx \le c\int_{-1}^{x_0}\big(u'(x)\big)^2(1-x)^{\alpha+2}(1+x)^{\beta+2}dx.$$

Finally, (B.42) follows from the above two inequalities. □

A direct consequence of (B.43) is as follows.

Corollary B.2. *For any $u \in H^1(a,b)$ with $u(x_0)=0$ for some $x_0 \in (a,b)$, the following Poincaré inequality holds:*

$$\|u\|_{L^2(a,b)} \le c\|u'\|_{L^2(a,b)}, \quad \forall u \in H^1(a,b). \tag{B.44}$$

Remark B.3. *In fact, the Poincaré inequality* (B.44) *holds, when the condition in Corollary B.2 is replaced by*

$$\int_a^b u(x)dx = 0. \tag{B.45}$$

B.8.3 Gronwall Inequalities

The Gronwall type inequalities are very useful in the stability and convergence analysis of initial-boundary value problems. The following is a typical Gronwall inequality:

Lemma B.9. *Let $f(t)$ be a non-negative integrable function over $(t_0,T]$, and let $g(t)$ and $E(t)$ be continuous functions on $[t_0,T]$. If $E(t)$ satisfies*

$$E(t) \le g(t) + \int_{t_0}^{t} f(\tau)E(\tau)d\tau, \quad \forall t \in [t_0,T], \tag{B.46}$$

then we have

$$E(t) \le g(t) + \int_{t_0}^{t} f(s)g(s)\exp\Big(\int_s^t f(\tau)d\tau\Big)ds, \quad \forall t \in [t_0,T]. \tag{B.47}$$

If, in addition, g is non-decreasing, then

$$E(t) \le g(t)\exp\Big(\int_{t_0}^{t} f(\tau)d\tau\Big), \quad \forall t \in [t_0, T]. \tag{B.48}$$

On the other hand, discrete Gronwall inequalities are often used in the stability and convergence analysis of time discretization schemes. In particular, a useful discrete analogue of Lemma B.9 is:

Lemma B.10. *Let y^n, h^n, g^n, f^n be four nonnegative sequences satisfying*

$$y^m + k\sum_{n=0}^{m} h^n \le B + k\sum_{n=0}^{m}(g^n y^n + f^n), \text{ with } k\sum_{n=0}^{T/k} g^n \le M, \ \forall\, 0 \le m \le T/k.$$

We assume $kg^n < 1$ and let $\sigma = max_{0\le n\le T/k}(1-kg^n)^{-1}$. Then

$$y^m + k\sum_{n=1}^{m} h^n \le \exp(\sigma M)(B + k\sum_{n=0}^{m} f^n), \ \ \forall m \le T/k.$$

We refer to, for instance, Quarteroni and Valli (2008) for a proof of the above two lemmas.

Appendix C
Basic Iterative Methods and Preconditioning

We review below some basic iterative methods for solving the linear system

$$Ax = b, \tag{C.1}$$

where $A \in \mathbb{R}^{n\times n}$ is an invertible matrix and $b \in \mathbb{R}^n$ is a given vector. We refer to Barrett et al. (1994), Golub and Van Loan (1996) and Saad (2003) for more detailed presentation in this matter.

C.1 Krylov Subspace Methods

The basic idea of Krylov subspace methods is to form an orthogonal basis of the sequence of successive matrix powers times the initial residual (the Krylov sequence), and then look for the approximation to the solution by minimizing the residual over the subspace formed by these orthogonal basis.

In what follows, we shall mainly discuss two proto-type of Krylov subspace methods: the *Conjugate Gradient (CG)* method and the *Generalized Minimal Residual (GMRES)* method. The CG method of Hestenes and Stiefel (1952) is the *method of choice* for solving large *symmetric positive definite* linear systems, while the GMRES method proposed by Saad and Schultz (1986) is popular for solving non-symmetric linear systems.

C.1.1 Conjugate Gradient (CG) Method

Throughout this section, let $A \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix. It can be verified that $\hat{x}$ is the solution of $Ax = b$ if and only if $\hat{x}$ minimizes the quadratic functional

$$J(x) = \frac{1}{2}x^T Ax - x^T b. \tag{C.2}$$

J. Shen et al., *Spectral Methods: Algorithms, Analysis and Applications*,
Springer Series in Computational Mathematics 41, DOI 10.1007/978-3-540-71041-7,

Suppose that $x^{(k)}$ has been obtained. Then $x^{(k+1)}$ can be found by

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}, \tag{C.3}$$

where the scalar α_k is called the step size factor and the vector $p^{(k)}$ is called the search direction. The coefficient α_k in (C.3) is selected such that

$$\alpha_k = \arg\min_{\alpha} J\left(x^{(k)} + \alpha p^{(k)}\right). \tag{C.4}$$

A simple calculation shows that

$$\alpha_k = \frac{(r^{(k)}, p^{(k)})}{(Ap^{(k)}, p^{(k)})}, \tag{C.5}$$

where the inner product of two column vectors is defined by $(u,v) = u^T v$, and the residual is given by

$$r^{(k)} := b - Ax^{(k)}. \tag{C.6}$$

Notice that the residual at the $(k+1)$th step is updated by

$$\begin{aligned} r^{(k+1)} &:= b - Ax^{(k+1)} = b - A\left(x^{(k)} + \alpha_k p^{(k)}\right) \\ &= b - Ax^{(k)} - \alpha_k Ap^{(k)} = r^{(k)} - \alpha_k Ap^{(k)}. \end{aligned} \tag{C.7}$$

In the conjugate gradient method, we select the *next search direction* $p^{(k+1)}$ satisfying the orthogonality

$$(p^{(k+1)}, Ap^{(k)}) = 0, \tag{C.8}$$

i.e.,

$$p^{(k+1)} = r^{(k+1)} + \beta_k p^{(k)}, \tag{C.9}$$

One verifies the orthogonality

$$\beta_k = -\frac{(Ap^{(k)}, r^{(k+1)})}{(Ap^{(k)}, p^{(k)})}. \tag{C.10}$$

It is important to notice the orthogonality

$$\left(p^{(i)}, Ap^{(j)}\right) = 0, \quad \left(r^{(i)}, r^{(j)}\right) = 0, \quad i \neq j. \tag{C.11}$$

Moreover, it can be shown that if A is a real $n \times n$ positive definite matrix, then, assuming exact arithmetic, the iteration converges in at most n steps, i.e., $x^{(m)} = \hat{x}$ for some $m \leq n$. Furthermore, the residual vectors satisfy

$$\left(r^{(k)}, p^{(j)}\right) = 0 \quad \text{for each} \quad j = 1, \ldots, k-1. \tag{C.12}$$

Therefore, we can reformulate the scalers α_k and β_k as

$$\alpha_k = \frac{(r^{(k)}, r^{(k)})}{(Ap^{(k)}, p^{(k)})}, \quad \beta_k = \frac{(r^{(k+1)}, r^{(k+1)})}{(r^{(k)}, r^{(k)})}. \tag{C.13}$$

We summarize the **CG Algorithm** below.

CG Algorithm

1. Initialization: choose $x^{(0)}$, compute $r^{(0)} = b - Ax^{(0)}$ and set $p^{(0)} = r^{(0)}$.
2. For $k = 0, 1, \ldots,$

   (i) Compute
   $$\alpha_k = (r^{(k)}, r^{(k)})/(Ap^{(k)}, p^{(k)}).$$
   (ii) Set
   $$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}.$$
   (iii) Compute
   $$r^{(k+1)} = r^{(k)} - \alpha_k A p^{(k)}.$$
   (iv) If the stopping rule does not apply, continue.
   (v) Compute
   $$\beta_k = (r^{(k+1)}, r^{(k+1)})/(r^{(k)}, r^{(k)}).$$
   (vi) Set
   $$p^{(k+1)} = r^{(k+1)} + \beta_k p^{(k)}.$$
3. endFor

The following theorem on the rate of convergence of the CG method can be found in e.g., Golub and Van Loan (1996).

Theorem C.1. *Suppose that $A \in \mathbb{R}^{n\times n}$ is symmetric positive definite matrix and $b \in \mathbb{R}^n$. Then the CG Algorithm produces iterates $\{x^{(k)}\}$ satisfying*

$$\|\hat{x} - x^{(k)}\|_A \le 2\Big(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\Big)^k \|\hat{x} - x^{(0)}\|_A, \tag{C.14}$$

where $\hat{x}$ is the exact solution of $Ax = b$, $\|x\|_A = \sqrt{x^T Ax}$, and $\kappa = \|A\|_2\|A^{-1}\|_2$ (the condition number of A).

Some remarks are in order.

(i) For a symmetric positive definite matrix, we have $\|A\|_2 = \lambda_n$ and $\|A^{-1}\|_2 = \lambda_1^{-1}$, where λ_n and λ_1 are the largest and smallest eigenvalues of A, respectively. One derives from Theorem C.1 the estimate in the 2-norm:

$$\|\hat{x} - x^{(k)}\|_2 \le 2\sqrt{\kappa}\Big(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\Big)^k \|\hat{x} - x^{(0)}\|_2. \tag{C.15}$$

(ii) The CG method involves one matrix–vector multiplication, three vector updates, and two inner products per iteration. If the matrix is sparse or has a special structure, these operators can be performed efficiently.
(iii) Unlike the traditional SOR type method, there is no free parameter to choose in the CG algorithm.

C.1.2 BiConjugate Gradient (BiCG) Method

The Conjugate Gradient method is not suitable for non-symmetric systems because the residual vectors can not be made orthogonal with short recurrences (see Faber and Manteuffel (1984) for the proof). The BiConjugate Gradient (BiCG) method takes another approach, replacing the orthogonal sequence of residuals by two mutually orthogonal sequences, at the price of no longer providing a minimization.

The BiCG method augments the update relations for residuals in the CG method by relations based on both A and A^T. More precisely, given two pairs: $(p^{(j)}, \tilde{p}^{(j)})$ and $(r^{(j)}, \tilde{r}^{(j)})$, we update

$$x^{(j+1)} = x^{(j)} + \alpha_j p^{(j)}, \tag{C.16}$$

and the two sequences of residuals

$$r^{(j+1)} = r^{(j)} - \alpha_j A p^{(j)}, \quad \tilde{r}^{(j+1)} = \tilde{r}^{(j)} - \alpha_j A^T \tilde{p}^{(j)}. \tag{C.17}$$

Require that $(r^{(j+1)}, \tilde{r}^{(j)}) = 0$ and $(r^{(j)}, \tilde{r}^{(j+1)}) = 0$ for all j. This leads to

$$\alpha_j = (r^{(j)}, \tilde{r}^{(j)})/(Ap^{(j)}, \tilde{p}^{(j)}). \tag{C.18}$$

The two sequences of search directions are updated by

$$p^{(j+1)} = r^{(j+1)} + \beta_j p^{(j)}, \quad \tilde{p}^{(j+1)} = \tilde{r}^{(j+1)} + \beta_j \tilde{p}^{(j)}. \tag{C.19}$$

By requiring that $(Ap^{(j+1)}, \tilde{p}^{(j)}) = 0$ and $(Ap^{(j)}, \tilde{p}^{(j+1)}) = 0$, we obtain

$$\beta_j = (r^{(j+1)}, \tilde{r}^{(j+1)})/(r^{(j)}, \tilde{r}^{(j)}). \tag{C.20}$$

The above derivations lead to the **BiCG Algorithm** outlined below.
Some remarks are in order.

(i) The BiCG algorithm is particularly suitable for matrices which are positive definite, i.e., $(Ax, x) > 0$ for all $x \neq 0$, but not necessary to be symmetric.
(ii) If A is symmetric positive definite and $\tilde{r}^{(0)} = r^{(0)}$, then the BiCG algorithm delivers the same results as the CG method, but at twice of the cost per iteration.
(iii) The algorithm breaks down if $(Ap_j, \tilde{p}_j) = 0$. Otherwise, the amount of work and storage is of the same order as the CG algorithm.

BiCG Algorithm

1. Initialization: choose $x^{(0)}$, compute $r^{(0)} = b - Ax^{(0)}$ and set $p^{(0)} = r^{(0)}$; choose $\tilde{r}^{(0)}$ (such that $(r^{(0)}, \tilde{r}^{(0)}) \neq 0$, e.g., $\tilde{r}^{(0)} = r^{(0)}$).
2. For $j = 0, 1, \ldots,$

   (i) Compute
   $$\alpha_j = (r^{(j)}, \tilde{r}^{(j)})/(Ap^{(j)}, \tilde{p}^{(j)}).$$
   (ii) Set
   $$x^{(j+1)} = x^{(j)} + \alpha_j p^{(j)}.$$
   (iii) Compute
   $$r^{(j+1)} = r^{(j)} - \alpha_j A p^{(j)}, \quad \tilde{r}^{(j+1)} = \tilde{r}^{(j)} - \alpha_j A^T \tilde{p}^{(j)}.$$
   (iv) If the stopping rule does not apply, continue.
   (v) Compute
   $$\beta_j = (r^{(j+1)}, \tilde{r}^{(j+1)})/(r^{(j)}, \tilde{r}^{(j)}).$$
   (vi) Set
   $$p^{(j+1)} = r^{(j+1)} + \beta_j p^{(j)}, \quad \tilde{p}^{(j+1)} = \tilde{r}^{(j+1)} + \beta_j \tilde{p}^{(j)}.$$

3. endFor

C.1.3 Conjugate Gradient Squared (CGS) Method

The BiCG algorithm requires multiplication by both A and A^T at each iteration. Obviously, this demands extra work, and in addition, it is sometimes cumbersome to multiply by A^T than by A. For example, there may be a special formula for the product of A with a given vector when A represents, say, a Jacobian, but a corresponding formula for the product of A^T with a given vector may not be available. In other cases, data may be stored on a parallel machine in such a way that multiplication by A is efficient but multiplication by A^T involves extra communication between processors. For these reasons it is desirable to have an iterative method that requires multiplication only by A and that generates good approximate solutions. A method for such purposes is the Conjugate Gradient Squared (CGS) method.

From the recurrence relations of BiCG algorithms, we see that

$$r^{(j)} = \Phi_j^a(A) r^{(0)} + \Phi_j^b(A) p^{(0)},$$

where $\Phi_j^a(A)$ and $\Phi_j^b(A)$ are polynomials of degree j of A. Choosing $p^{(0)} = r^{(0)}$ gives

$$r^{(j)} = \Phi_j(A) r^{(0)} \quad \text{where} \quad \Phi_j = \Phi_j^a + \Phi_j^b,$$

with $\Phi_0 \equiv 1$. Similarly,

$$p^{(j)} = \pi_j(A) r^{(0)},$$

where $\pi_j(A)$ is a polynomial of degree j of A. As $\tilde{r}^{(j)}$ and $\tilde{p}^{(j)}$ are updated, using the same recurrence relation as for $r^{(j)}$ and $p^{(j)}$, we have

$$\tilde{r}^{(j)} = \Phi_j(A^T) \tilde{r}^{(0)}, \quad \tilde{p}^{(j)} = \pi_j(A^T) \tilde{r}^{(0)}. \tag{C.21}$$

Hence,

$$\alpha_j = \frac{(\Phi_j(A)r^{(0)}, \Phi_j(A^T)\tilde{r}^{(0)})}{(A\pi_j(A)r^{(0)}, \pi_j(A^T)\tilde{r}^{(0)})} = \frac{(\Phi_j^2(A)r^{(0)}, \tilde{r}^{(0)})}{(A\pi_j^2(A)r^{(0)}, \tilde{r}^{(0)})}. \tag{C.22}$$

From the BiCG algorithm,

$$\Phi_{j+1}(t) = \Phi_j(t) - \alpha_j t\pi_j(t), \quad \pi_{j+1}(t) = \Phi_{j+1}(t) + \beta_j\pi_j(t). \tag{C.23}$$

Observe that

$$\Phi_j\pi_j = \Phi_j(\Phi_j + \beta_{j-1}\pi_{j-1}) = \Phi_j^2 + \beta_{j-1}\Phi_j\pi_{j-1}. \tag{C.24}$$

It follows from the above results that

$$\begin{aligned}
\Phi_{j+1}^2 &= \Phi_j^2 - 2\alpha_j t(\Phi_j^2 + \beta_{j-1}\Phi_j\pi_{j-1}) + \alpha_j^2 t^2\pi_j^2,\\
\Phi_{j+1}\pi_j &= \Phi_j\pi_j - \alpha_j t\pi_j^2 = \Phi_j^2 + \beta_{j-1}\Phi_j\pi_{j-1} - \alpha_j t\pi_j^2,\\
\pi_{j+1}^2 &= \Phi_{j+1}^2 + 2\beta_j\Phi_{j+1}\pi_j + \beta_j^2\pi_j^2.
\end{aligned} \tag{C.25}$$

Define

$$\begin{aligned}
&r^{(j)} = \Phi_j^2(A)r^{(0)}, \quad p^{(j)} = \pi_j^2(A)r^{(0)},\\
&q^{(j)} = \Phi_{j+1}(A)\pi_j(A)r^{(0)}, \quad d^{(j)} = 2r^{(j)} + 2\beta_{j-1}q^{(j-1)} - \alpha_j Ap^{(j)}.
\end{aligned}$$

It can be verified that

$$\begin{aligned}
r^{(j+1)} &= r^{(j)} - \alpha_j Ad^{(j)},\\
q^{(j)} &= r^{(j)} + \beta_{j-1}q^{(j-1)} - \alpha_j Ap^{(j)},\\
p^{(j+1)} &= r^{(j+1)} + 2\beta_j q^{(j)} + \beta_j^2 p^{(j)},\\
d^{(j)} &= 2r^{(j)} + 2\beta_{j-1}q^{(j-1)} - \alpha_j Ap^{(j)}.
\end{aligned}$$

Correspondingly,

$$x^{(j+1)} = x^{(j)} + \alpha_j d^{(j)}. \tag{C.26}$$

This gives the **CGS Algorithm** as summarized below.

The CGS method requires two matrix–vector multiplications at each step but no multiplications by the transpose. For problems where the BiCG method converges well, the CGS method typically requires only about half as many steps and, therefore, half the work of the BiCG method (assuming that multiplication by A or A^T requires the same amount of work). When the norm of the BiCG residual increases at a step, however, that of the CGS residual usually increases by approximately the square of the increase of the BiCG residual norm. The CGS algorithm convergence curve may therefore show wild oscillations that can sometimes lead to numerical instability and break down.

CGS Algorithm

1. Initialization: choose $x^{(0)}$, compute $r^{(0)} = b - Ax^{(0)}$ and set $p^{(0)} = r^{(0)} = u^{(0)}, q^{(0)} = 0$; choose $\tilde{r}^{(0)}$ such that $(r^{(0)}, \tilde{r}^{(0)}) \neq 0$.
2. For $j = 0, 1, \ldots,$

   (i) Compute
   $$\alpha_j = (r^{(j)}, \tilde{r}^{(0)})/(Ap^{(j)}, \tilde{r}^{(0)}),$$
   and
   $$q^{(j+1)} = u^{(j)} - \alpha_j A p^{(j)}.$$
   (ii) Set
   $$x^{(j+1)} = x^{(j)} + \alpha_j (u^{(j)} + q^{(j+1)}).$$
   (iii) Compute
   $$r^{(j+1)} = r^{(j)} - \alpha_j A (u^{(j)} + q^{(j+1)}).$$
   (iv) If the stopping rule does not apply, continue.
   (v) Compute
   $$\beta_j = (r^{(j+1)}, \tilde{r}^{(0)})/(r^{(j)}, \tilde{r}^{(0)}),$$
   and
   $$u^{(j+1)} = r^{(j+1)} + \beta_j q^{(j+1)}.$$
   (vi) Set
   $$p^{(j+1)} = u^{(j+1)} + \beta_j (q^{(j+1)} + \beta_j p^{(j)}).$$
3. endFor

C.1.4 BiConjugate Gradient Stabilized (BiCGStab) Method

The BiConjugate Gradient Stabilized (BiCGStab) method was developed by Van der Vorst (1992) to solve non-symmetric linear systems while avoiding the irregular convergence patterns of the CGS method. The main idea is to produce a residual of the form

$$r^{(j)} = \Psi_j(A)\Phi_j(A)r^{(0)}, \tag{C.27}$$

where Φ_j is again the BiCG polynomial but Ψ_j is chosen to keep the residual norm small at each step while retaining the rapid overall convergence of the CGS method. For example, $\Psi_j(t)$ could be of the form

$$\Psi_{j+1}(t) = (1 - w_j t)\Psi_j(t). \tag{C.28}$$

In the BiCGStab algorithm, the solution is updated in such a way that $r^{(j)}$ is of the form (C.27), where $\Psi_j(A)$ is a polynomial of degree j satisfying (C.28). Then

$$\begin{aligned} \Psi_{j+1}\Phi_{j+1} &= (1 - w_j t)\Psi_j(\Phi_j - \alpha_j t \pi_j) \\ &= (1 - w_j t)(\Psi_j\Phi_j - \alpha_j t \Psi_j \pi_j), \end{aligned} \tag{C.29}$$

and

$$\begin{aligned}\Psi_j\pi_j &= \Psi_j(\Phi_j + \beta_{j-1}\pi_{j-1}) \\ &= \Psi_j\Phi_j + \beta_{j-1}(1 - w_{j-1}t)\Psi_{j-1}\pi_{j-1}.\end{aligned} \tag{C.30}$$

Let $r^{(j)} = \Phi_j(A)\Psi_j(A)r^{(0)}$ and $p^{(j)} = \Psi_j(A)\pi_j(A)r^{(0)}$. It can be verified that

$$\begin{aligned}r^{(j+1)} &= (I - w_jA)(r^{(j)} - \alpha_jAp^{(j)}), \\ p^{(j+1)} &= r^{(j+1)} + \beta_j(I - w_jA)p^{(j)}.\end{aligned} \tag{C.31}$$

Letting $s^{(j)} = r^{(j)} - \alpha_jAp^{(j)}$, we obtain

$$r^{(j+1)} = (I - w_jA)s^{(j)}. \tag{C.32}$$

The parameter w_j is chosen to minimize the 2-norm of $r^{(j+1)}$, i.e.,

$$w_j = \frac{(As^{(j)}, s^{(j)})}{(As^{(j)}, As^{(j)})}. \tag{C.33}$$

We also need to find an updating formula for α_j and β_j, which ideally only involves $r^{(k)}, p^{(k)}$ and $s^{(k)}$. This seems to be rather complicated, so we omit the derivation here.

The **BiCGStab Algorithm** is summarized below.

BiCGStab Algorithm

1. Initialization: choose $x^{(0)}$, compute $r^{(0)} = b - Ax^{(0)}$ and set $p^{(0)} = r^{(0)}$; choose $\tilde{r}^{(0)}$ such that $(r^{(0)}, \tilde{r}^{(0)}) \neq 0$.
2. For $j = 0, 1, \ldots,$

   (i) Compute
   $$\alpha_j = \frac{(r^{(j)}, \tilde{r}^{(0)})}{(Ap^{(j)}, \tilde{r}^{(0)})}.$$

   (ii) Set
   $$s^{(j)} = r^{(j)} - \alpha_jAp^{(j)},$$
   and compute
   $$w_j = \frac{(As^{(j)}, s^{(j)})}{(As^{(j)}, As^{(j)})}.$$

   (iii) Set
   $$x^{(j+1)} = x^{(j)} + \alpha_jp^{(j)} + w_js^{(j)}; \quad r^{(j+1)} = s^{(j)} - w_jAs^{(j)}.$$

   (iv) If the stopping rule does not apply, continue.

   (v) Compute
   $$\beta_j = \frac{\alpha_j}{w_j}\frac{(r^{(j+1)}, \tilde{r}^{(0)})}{(r^{(j)}, \tilde{r}^{(0)})}.$$

   (vi) Set
   $$p^{(j+1)} = r^{(j+1)} + \beta_j(p^{(j)} - w_jAp^{(j)}).$$

3. endFor

In general, the BiCGStab method often converges about as fast as the CGS algorithm. We also notice that the BiCGStab method requires two matrix–vector products and four inner products, i.e., two inner products more than the BiCG and CGS methods.

C.1.5 Generalized Minimal Residual (GMRES) Method

The Generalized Minimal Residual method proposed by Saad and Schultz (1986) is one of the most important tools for solving general *non-symmetric* system: $Ax = b$. In the k-th iteration of the GMRES method, we need to find $x^{(k)}$ that minimizes $\|b - Ax\|_2$ over the set

$$S_k := x^{(0)} + \text{span}\{r^{(0)}, Ar^{(0)}, \ldots, A^{k-1}r^{(0)}\}, \tag{C.34}$$

where $r^{(0)} = b - Ax^{(0)}$. In other words, for any $x \in S_k$, we have

$$x = x^{(0)} + \sum_{j=0}^{k-1} \gamma_j A^j r^{(0)}. \tag{C.35}$$

Moreover, it can be shown that

$$r = b - Ax = r^{(0)} - \sum_{j=1}^{k} \gamma_{j-1} A^j r^{(0)}. \tag{C.36}$$

Like the CG algorithm, the GMRES method will obtain the *exact* solution of $Ax = b$ within n iterations. Moreover, if b is a linear combination of k eigenvectors of A, say $b = \sum_{p=1}^{k} \gamma_p u_{i_p}$, then the GMRES method will terminate within at most k iterations.

The first important issue is to find a basis for S_k. Suppose that we have a matrix $V_k = [v_1^k, v_2^k, \ldots, v_k^k]$, whose columns form an orthogonal basis of S_k. Then any $z \in S_k$ can be expressed as

$$z = \sum_{p=1}^{k} u_p v_p^k = V_k u, \tag{C.37}$$

where $u = (u_1, \ldots, u_k) \in \mathbb{R}^k$. Once we have found V_k, we can convert the original least-squares problem: $\min_{x \in S_k} \|b - Ax\|_2$ into a least-squares problem in $\mathbb{R}^k$. More precisely, let $x^{(k)}$ be the solution after the k-th iteration. Then we have $x^{(k)} = x^{(0)} + V_k y^{(k)}$, where the vector $y^{(k)}$ minimizes

$$\min_{y \in \mathbb{R}^k} \|b - A(x^{(0)} + V_k y)\|_2 = \min_{y \in \mathbb{R}^k} \|r^{(0)} - AV_k y\|_2. \tag{C.38}$$

This is a standard linear least-squares problem that can be solved by a QR decomposition.

To find an orthonormal basis of S_k, one can use the **modified Gram-Schmidt orthogonalization** as highlighted below.
This algorithm produces the columns of the matrix V_k, which also form an orthonormal basis for S_k. Note that the algorithm breaks down when a division by zero occurs.

If the modified Gram-Schmidt process does not break down, we can use it to carry out the GMRES method (i.e., to solve the minimization problem (C.38)) in an efficient way. More precisely, define

Modified Gram-Schmidt Orthogonalization

1. Initialization: choose $x^{(0)}$, and set $r^{(0)} = b - Ax^{(0)}$ and $v^{(1)} = r^{(0)}/\|r^{(0)}\|_2$.
2. For $j = 1, 2, \ldots, k$,

   Compute
   $$v^{(j+1)} = \frac{Av^{(j)} - \sum_{l=1}^{j} \left(Av^{(j)}, v^{(l)}\right) v^{(l)}}{\|Av^{(j)} - \sum_{l=1}^{j} \left(Av^{(j)}, v^{(l)}\right) v^{(l)}\|_2}.$$
3. endFor

$$h_{ij} = \left(Av^{(j)}, v^{(i)}\right), \quad 1 \le i \le j \le k.$$

From the modified Gram-Schmidt algorithm, we obtain a $k \times k$ matrix $H_k = (h_{ij})$, which is upper Hessenberg, i.e., its entries satisfy $h_{ij} = 0$ if $i > j+1$. Moreover, this process produces a matrix $V_k = [v_1^k, v_2^k, \ldots, v_k^k]$, whose columns form an orthonormal basis for S_k, and we have

$$AV_k = V_{k+1}\tilde{H}_k, \tag{C.39}$$

where $\tilde{H}_k$ is generated by H_k (see P. 548 of Golub and Van Loan (1996)). This allows us to convert the problem: updating $x^{(k)} = x^{(0)} + V_k y^{(k)}$ by solving (C.38), into an alternative formulation. Using the fact $r^{(0)} = b - Ax^{(0)}$, and (C.39), one verifies that

$$\begin{aligned} r^{(k)} &= b - Ax^{(k)} = r^{(0)} - A\left(x^{(k)} - x^{(0)}\right) \\ &= \beta V_{k+1} e_1 - AV_k y^{(k)} = V_{k+1}\left(\beta e_1 - \tilde{H}_k y^{(k)}\right), \end{aligned} \tag{C.40}$$

where e_1 is the first unit k-vector $(1, 0, \ldots, 0)^T$, and $\beta = \|r^{(0)}\|_2$. Therefore, the problem (C.38) becomes

$$\min_{y \in \mathbb{R}^k} \left\|\beta e_1 - \tilde{H}_k y\right\|_2. \tag{C.41}$$

To find the minimizer $y^{(k)}$ of (C.41), it is necessary to we look at the linear system $\tilde{H}_k y = \beta e_1$, which can be solved by using rotation matrices to perform Gauss-elimination for $\tilde{H}_k$ (see, e.g., Saad (2003)). Here, we skip the details.

The pseudocode of the **GMRES Algorithm** for solving $Ax = b$ with A being a non-symmetric matrix is given below.

GMRES Algorithm

1. Initialization: choose $x^{(0)}$, and set $r^{(0)} = b - Ax^{(0)}, \beta = \|r^{(0)}\|_2$ and $v^{(1)} = r^{(0)}/\|r^{(0)}\|_2$.
2. For a given k, find the basis for S_k by e.g., the modified Gram-Schmidt orthogonalization process.
3. Form $\tilde{H}_k$, and solve (C.41) to find $y^{(k)}$.
4. Set $x^{(k)} = x^{(0)} + V_k y^{(k)}$.
5. Check convergence; if necessary, set $x^{(0)} = x^{(k)}, v^{(1)} = r^{(k)}/\|r^{(k)}\|_2$, and go to Step 2.

C.2 Preconditioning

The convergence rate of iterative methods depends on spectral properties of the coefficient matrix. Hence, one may attempt to transform the linear system into an equivalent system that has more favorable spectral properties. A *preconditioner* is a matrix for such a transformation. A good preconditioner is a matrix M that is easy to invert and the condition number of $M^{-1}A$ is small. In other words, the preconditioned system $M^{-1}Ax = M^{-1}b$ can be solved efficiently by an appropriate iterative method.

C.2.1 Preconditioned Conjugate Gradient (PCG) Method

Based on this idea, the Preconditioned Conjugate Gradient (PCG) method can be derived with a slight modification of the CG method as described below.

In this algorithm, we need to solve the system $M\bar{r} = r$, which might be as complicated as the original system. The idea for reducing the condition number of $M^{-1}A$ is to choose M such that M^{-1} is close to A^{-1}, while the system $M\bar{r} = r$ is easy to solve. The following theorem provides a choice of M.

Theorem C.2. *Let A be an $n \times n$ nonsingular matrix, and let $A = P - Q$ be a splitting of A such that P is nonsingular. If $H = P^{-1}Q$ and $\rho(H) < 1$, then*

$$A^{-1} = \Big(\sum_{k=0}^{\infty} H^k\Big)P^{-1}. \tag{C.42}$$

Based on this theorem, we can regard the matrices

$$\begin{aligned} M &= P\big(I + H + \ldots + H^{m-1}\big)^{-1}, \\ M^{-1} &= \big(I + H + \ldots + H^{m-1}\big)P^{-1}, \end{aligned} \tag{C.43}$$

PCG Algorithm

1. Initialization: choose $x^{(0)}$, compute $r^{(0)} = b - Ax^{(0)}$ and solve $M\bar{r}^{(0)} = r^{(0)}$. Set $p^{(0)} = \bar{r}^{(0)}$.
2. For $k = 0, 1, \ldots,$

   (i) Compute
   $$\alpha_k = (\bar{r}^{(k)}, r^{(k)})/(Ap^{(k)}, p^{(k)}).$$
   (ii) Set
   $$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}.$$
   (iii) Compute
   $$r^{(k+1)} = r^{(k)} - \alpha_k A p^{(k)}.$$
   (iv) If the stopping rule does not apply, continue.
   (v) Solve
   $$M\bar{r}^{(k+1)} = r^{(k+1)}.$$
   (vi) Compute
   $$\beta_k = (\bar{r}^{(k+1)}, r^{(k+1)})/(\bar{r}^{(k)}, r^{(k)}).$$
   (vii) Set
   $$p^{(k+1)} = \bar{r}^{(k+1)} + \beta_k p^{(k)}.$$

3. endFor

as the approximations of A and A^{-1}, respectively. Thus the solution of the system $M\bar{r} = r$ becomes

$$\bar{r} = M^{-1}r = \left(I + H + \ldots + H^{m-1}\right)P^{-1}r.$$

Equivalently, the solution $\bar{r} = r_m$ is the result of applying m steps of the iterative scheme

$$Pr^{(i+1)} = Qr^{(i)} + r, \quad i = 0, 1, \ldots, m-1; \quad r_0 = 0.$$

If $P = D$ and $Q = L + U$, the above iteration is the standard Jacobi method. Then in the PCG method, we replace the system $M\bar{r}^{(k+1)} = r^{(k+1)}$ with *do m Jacobi iterations on* $Ar = r^{(k+1)}$ *to obtain* $\bar{r}^{(k+1)}$. The resulting method is called the ***m*-step Jacobi PCG** method.

In practice, we may just use the one-step Jacobi PCG method, i.e., $M = D$. Similarly, the symmetric Gauss-Seidel and symmetric Successive Over-Relaxation (SSOR) methods can also be used as preconditioners:

- Symmetric Gauss-Seidel preconditioner:

$$M = (D - L)D^{-1}(D - U), \quad M^{-1} = (D - U)^{-1}D(D - L)^{-1}.$$

- SSOR preconditioner:

$$M = \frac{\omega}{2 - \omega}(\omega^{-1}D - L)D^{-1}(\omega^{-1}D - U),$$
$$M^{-1} = \omega(2 - \omega)(D - \omega U)^{-1}D(D - \omega L)^{-1}.$$

C.2.2 Preconditioned GMRES Method

If we use M as a left preconditioner for the GMRES method, then we are trying to minimize the residual in the space:

$$S_m(A, r^{(0)}) = \text{span}\{r^{(0)}, M^{-1}Ar^{(0)}, \ldots, (M^{-1}A)^{m-1}r^{(0)}\}. \tag{C.44}$$

The resulting algorithm is the very same as the original GMRES method.

If M is used as a right preconditioner, we just need to replace A in the original GMRES algorithm by AM^{-1}. Also we need to update $x^{(k)}$ by

$$x^{(k)} = x^{(0)} + M^{-1}V_k y^{(k)}. \tag{C.45}$$

In practice, for the GMRES method, the Gauss-Seidel and SOR methods can also be used as preconditioners:

- Gauss-Seidel preconditioner:

$$M = D - L, \quad M^{-1} = (D - L)^{-1}.$$

- SOR preconditioner:

$$M = \omega^{-1}D - L,\ M^{-1} = \omega(D - \omega L)^{-1}.$$

The preconditioned CGS or BiCGStab algorithms can be constructed similarly. In general, to use preconditioners for the CGS method or the BiCGStab method, it is only necessary to replace the matrix A in the original algorithms by $M^{-1}A$ or AM^{-1}.

Appendix D
Basic Time Discretization Schemes

We describe below several standard methods for ordinary differential equations (ODEs), and present some popular time discretization schemes which are widely used in conjunction with spectral methods for partial differential equations (PDEs). We refer to by Gear (1971), Lambert (1991), Hairer et al. (1993), Hairer and Wanner (1996), LeVeque (2007), and Butcher (2008) for thorough discussions on numerical ODEs.

D.1 Standard Methods for Initial-Valued ODEs

Consider the initial value problem (IVP):

$$\frac{dU}{dt} = F(U,t), \quad t > 0; \quad U(0) = U_0. \tag{D.1}$$

In general, it may represent a system of ODEs, i.e., $U, F, U_0 \in \mathbb{R}^d$. To this end, let τ be the time step size, and let U^n be the approximation of U at $t_n = n\tau, n \geq 1$.

The simplest method is to approximate dU/dt by the finite difference quotient $U'(t) \approx (U(t+\tau) - U(t))/\tau$. This gives the *forward Euler's method*:

$$U^{n+1} = U^n + \tau F(U^n, t_n), \quad n \geq 0; \quad U^0 = U_0. \tag{D.2}$$

From the initial data U^0, we can compute U^1, then U^2, and so on. Accordingly, it is called a *time marching scheme.*

The *backward Euler's method* is similar, but it is based on approximating $U'(t_{n+1})$ by the backward difference:

$$U^{n+1} = U^n + \tau F(U^{n+1}, t_{n+1}), \quad n \geq 0; \quad U^0 = U_0. \tag{D.3}$$

J. Shen et al., *Spectral Methods: Algorithms, Analysis and Applications*,
Springer Series in Computational Mathematics 41, DOI 10.1007/978-3-540-71041-7,

In contrast with (D.2), to march from U^n to U^{n+1}, (D.3) requires to solve for U^{n+1}. It can be viewed as looking for a zero of the function:

$$g(u) = u - \tau F(u, t_{n+1}) - U^n,$$

which can be located by using an iterative method such as the *Newton's method.*

In view of this, the backward scheme (D.3) is an *implicit* method as U^{n+1} must be solved at each iteration, whereas the forward Euler method (D.2) is an *explicit method*.

Another implicit method is the *trapezoidal method*, obtained by averaging two Euler's methods as follows:

$$\frac{U^{n+1} - U^n}{\tau} = \frac{1}{2}\big\{F(U^n, t_n) + F(U^{n+1}, t_{n+1})\big\}, \quad n \geq 0; \quad U^0 = U_0. \tag{D.4}$$

As one might expect, this symmetric approximation is second-order accurate, whereas the Euler's methods are only first-order accurate.

The conceptually simplest approach to construct higher-order methods is to use more terms in the Taylor expansion. For example, we consider

$$U(t_{n+1}) \approx U(t_n) + \tau U'(t_n) + \frac{\tau^2}{2} U''(t_n), \tag{D.5}$$

where the remainder of $O(\tau^3)$ has been dropped. In view of (D.1), $U'(t_n)$ can be replaced by $F(U^n, t_n)$, and notice that

$$U''(t) = \frac{d}{dt} F(U(t), t) = F_U(U, t) U'(t) + F_t(U, t),$$

which motivates the approximation:

$$U''(t_n) \approx F_U(U^n, t_n) F(U^n, t_n) + F_t(U^n, t_n).$$

Consequently, we obtain the scheme:

$$U^{n+1} = U^n + \tau F(U^n, t_n) + \frac{\tau^2}{2}\big\{F_t(U^n, t_n) + F_U(U^n, t_n) F(U^n, t_n)\big\}. \tag{D.6}$$

It can be shown the above scheme has a second-order accuracy provided that F and the underlying solution U are smooth. However, this can result in very messy and problematic expressions that must be worked out for each equation.

D.1.1 Runge–Kutta Methods

The Taylor method outlined in the previous part has the desirable property of high-order accuracy, but the disadvantage of requiring the computation and evaluation

of the derivatives of F, makes it less attractive in practice. One important class of higher-order methods without derivative computations is known as the Runge–Kutta methods.

The second-order Runge–Kutta method is of the form:

$$\begin{aligned} &K_1 = F(U^n, t_n), \quad K_2 = F(U^n + a\tau K_1, t_n + b\tau), \\ &U^{n+1} = U^n + \tau\big(\alpha K_1 + \beta K_2\big), \end{aligned} \tag{D.7}$$

where the parameters satisfy

$$\alpha + \beta = 1, \quad a\beta = b\beta = \frac{1}{2}. \tag{D.8}$$

Notice that there are three equations for four unknowns, so we are free to choose one parameter. This results in different schemes.

- If $\alpha = 0, \beta = 1$ and $a = b = 1/2$, we have the *midpoint method*:

$$U^{n+1} = U^n + \tau F\Big(U^n + \frac{\tau}{2} F(U^n, t_n), t_n + \frac{\tau}{2}\Big). \tag{D.9}$$

- If $\alpha = \beta = 1/2$ and $a = b = 1$, the scheme (D.7) is known as the *modified Euler's method*:

$$U^{n+1} = U^n + \frac{\tau}{2}\Big\{F(U^n, t_n) + F\big(U^n + \tau F(U^n, t_n), t_{n+1}\big)\Big\}. \tag{D.10}$$

- If $\alpha = 1/4, \beta = 3/4$ and $a = b = 2/3$, the scheme (D.7) is known as the *Heun method*:

$$U^{n+1} = U^n + \frac{\tau}{4}\Big\{F(U^n, t_n) + 3F\Big(U^n + \frac{2}{3}\tau F(U^n, t_n), t_n + \frac{2}{3}\tau\Big)\Big\}. \tag{D.11}$$

The third-order Runge–Kutta method is given by:

$$\begin{cases} K_1 = F(U^n, t_n), \\ K_2 = F\Big(U^n + \dfrac{\tau}{2} K_1, t_n + \dfrac{\tau}{2}\Big), \\ K_3 = F\big(U^n - \tau K_1 + 2\tau K_2, t_n + \tau\big), \\ U^{n+1} = U^n + \dfrac{\tau}{6}\big(K_1 + 4K_2 + K_3\big). \end{cases} \tag{D.12}$$

The classical fourth-order Runge–Kutta (RK4) method is

$$\begin{cases} K_1 = F(U^n, t_n), \\ K_2 = F\left(U^n + \dfrac{\tau}{2} K_1, t_n + \dfrac{\tau}{2}\right), \\ K_3 = F\left(U^n + \dfrac{\tau}{2} K_2, t_n + \dfrac{\tau}{2}\right), \\ K_4 = F(U^n + \tau K_3, t_{n+1}), \\ U^{n+1} = U^n + \dfrac{\tau}{6}\left(K_1 + 2K_2 + 2K_3 + K_4\right). \end{cases} \tag{D.13}$$

The above formula requires four levels of storage, i.e., K_1, K_2, K_3 and K_4. An equivalent formulation is:

$$\begin{cases} U = U^n, \quad G = U, \quad P = F(U, t_n), \\ U = U + \tau P/2, \quad G = P, \quad P = F\big(U, t_n + \tau/2\big), \\ U = U + \tau(P - G)/2, \quad G = G/6, \quad P = F\big(U, t_n + \tau/2\big) - P/2, \\ U = U + \tau P, \quad G = G - P, \quad P = F(U, t_{n+1}) + 2P, \\ U^{n+1} = U + \tau\,(G + P/6)\,. \end{cases} \tag{D.14}$$

This version of the RK4 method requires only three levels (U, G and P) of storage.

As we saw in the derivation of the Runge–Kutta method of order 2, a number of parameters must be selected. A similar situation occurs in finding higher-order Runge–Kutta methods. Consequently, there is not just one Runge–Kutta method for each order, but a family of methods. It is worthwhile to point out that the number of required *function evaluations* increases more rapidly than the order of the Runge–Kutta methods. This makes the higher-order Runge–Kutta methods less attractive than some other classical fourth-order methods.

D.1.2 Multi-Step Methods

The methods discussed to this point are called *one-step methods* because the approximation at time t_{n+1} involves the information from only one of the previous time t_n. Although these methods might use functional evaluation information at points between t_n and t_{n+1}, they do not retain that information for direct use in future approximations.

We next review some methods using approximations at more than one previous approximations, which are called *multi-step methods.* In general, the r-step linear multi-step methods (LLMs) takes the form

$$\sum_{j=0}^{r} \alpha_j U^{n+j} = \tau \sum_{j=0}^{r} \beta_j F(U^{n+j}, t_{n+j}), \tag{D.15}$$

where U^{n+r} is computed from the equation in terms of the previous approximations $U^{n+r-1}, U^{n+r-2}, \ldots, U^n$ and F at these points (which can be stored and reused if F is expensive to evaluate).

If $\beta_r = 0$, the method (D.15) is explicit; otherwise it is implicit. Note that we can multiply both sides by any non-zero constant and have essentially the same method, so the normalization $\alpha_r = 1$ is often assumed.

The *leap frog method* is a second-order, two-step scheme given by

$$U^{n+1} = U^{n-1} + 2\tau F(U^n, t_n). \tag{D.16}$$

Some special classes of methods are particularly useful with distinctive names. We list a few of them.

The *Adams methods* take the form

$$U^{n+r} = U^{n+r-1} + \tau \sum_{j=0}^{r} \beta_j F(U^{n+j}, t_{n+j}). \tag{D.17}$$

These methods all have

$$\alpha_r = 1, \quad \alpha_{r-1} = -1, \quad \alpha_j = 0 \ \text{ for all } \ j < r-1. \tag{D.18}$$

The coefficients β_j are chosen to maximize the order of accuracy. If we require $\beta_r = 0$, the method is explicit and the r coefficients $\{\beta_j\}_{j=0}^{r-1}$ can be chosen so that the method has order r. This can be done by using Taylor expansion of the local truncation error and then choosing the coefficients to eliminate as many terms as possible. This process leads to explicit *Adams-Bashforth methods*.

Another way to derive the Adams-Bashforth methods is by writing

$$U(t_{n+1}) - U(t_n) = \int_{t_n}^{t_{n+1}} U'(t)dt = \int_{t_n}^{t_{n+1}} F(U(t),t)dt, \tag{D.19}$$

and by applying a quadrature formula to approximate

$$\int_{t_n}^{t_{n+1}} F(U(t),t)dt \approx \int_{t_n}^{t_{n+1}} L_{n,r-1}(t)dt, \tag{D.20}$$

where $L_{n,r-1}(t)$ is the Lagrange interpolating polynomial of degree $r-1$ at the points $t_n, t_{n-1}, \ldots, t_{n+r-1}$. The first few schemes obtained from this procedure are listed below, where we denote $F_k := F(U^k, t_k)$.

Explicit Adams-Bashforth methods

1-step:

$$U^{n+1} = U^n + \tau F_n \quad \text{(forward Euler)}$$

2-step:

$$U^{n+2} = U^{n+1} + \frac{\tau}{2}\{3F_{n+1} - F_n\} \quad \text{(AB2)}$$

3-step:

$$U^{n+3} = U^{n+2} + \frac{\tau}{12}\{23F_{n+2} - 16F_{n+1} + 5F_n\} \quad \text{(AB3)}$$

4-step:

$$U^{n+4} = U^{n+3} + \frac{\tau}{24}\{55F_{n+3} - 59F_{n+2} + 37F_{n+1} - 9F_n\} \quad \text{(AB4)}$$

If $\beta_r \neq 0$, then we have one more free parameter. This allows us to derive the r-step *Adams-Moulton methods* of order $r+1$. They are implicit, and we list several such schemes in the box.

One difficulty with LLMs if $r > 1$ is that we have to provide the r initial values $U^0, U^1, \ldots, U^{r-1}$ before we apply the multi-step methods. The initial value U^0 is given, but the other values are not and typically must be generated by some other methods, such as the Runge–Kutta methods.

Implicit Adams-Moulton methods

1-step:

$$U^{n+1} = U^n + \frac{\tau}{2}\{F_{n+1} + F_n\} \quad \text{(Crank-Nicolson method)}$$

2-step:

$$U^{n+2} = U^{n+1} + \frac{\tau}{12}\{5F_{n+2} + 8F_{n+1} - F_n\} \quad \text{(AM3)}$$

3-step:

$$U^{n+3} = U^{n+2} + \frac{\tau}{24}\{9F_{n+3} + 19F_{n+2} - 5F_{n+1} + F_n\} \quad \text{(AM4)}$$

4-step:

$$U^{n+4} = U^{n+3} + \frac{\tau}{720}\{251F_{n+4} + 646F_{n+3} - 264F_{n+2} + 106F_{n+1} - 19F_n\}$$

We have not touched on the theoretical issues of convergence and stability of these methods, and refer the readers to the books of Gear (1971), Lambert (1991), LeVeque (2007), and Butcher (2008) for more detail.

D.1.3 Backward Difference Methods (BDF)

The Adams-Bashforth methods might be unstable due to the fact they are obtained by integrating the interpolating polynomial outside the interval of the data that define the polynomial. This can be remedied by using multilevel implicit methods:

- Second-order backward difference method (BDF2):

$$\frac{1}{2\tau}\left(3U^{n+1} - 4U^n + U^{n-1}\right) = F(U^{n+1}, t_{n+1}). \tag{D.21}$$

- Third-order backward difference method (BDF3):

$$\frac{1}{6\tau}\left(11U^{n+1} - 18U^n + 9U^{n-1} - 2U^{n-2}\right) = F(U^{n+1}, t_{n+1}). \tag{D.22}$$

- Fourth-order backward difference method (BDF4):

$$\frac{1}{12\tau}\left(25U^{n+1} - 48U^n + 36U^{n-1} - 16U^{n-2} + 3U^{n-3}\right) = F(U^{n+1}, t_{n+1}). \tag{D.23}$$

In some applications, $F(u,t)$ is often the sum of linear and nonlinear terms. In this case, some combination of the backward difference method and extrapolation method can be used. To fix the idea, let us consider

$$U_t = L(U) + N(U), \tag{D.24}$$

where L is a linear operator and N is a nonlinear operator. By combining a second-order backward differentiation (BDF2) for the time derivative term and a second-order extrapolation (EP2) for the explicit treatment of the nonlinear term, we obtain a second-order scheme (BDF2/EP2) for (D.24):

$$\frac{1}{2\tau}\left(3U^{n+1} - 4U^n + U^{n-1}\right) = L(U^{n+1}) + N(2U^n - U^{n-1}). \tag{D.25}$$

A third-order scheme for solving (D.24) can be constructed in a similar manner, which leads to the so-called BDF3/EP3 scheme:

$$\begin{aligned}\frac{1}{6\tau}\left(11U^{n+1} - 18U^n + 9U^{n-1} - 2U^{n-2}\right)\\ = L(U^{n+1}) + N(3U^n - 3U^{n-1} + U^{n-2}).\end{aligned} \tag{D.26}$$

D.2 Operator Splitting Methods

In some situations, $F(u,t)$ is the sum of several terms with different nature. It is oftentimes advisable to use an operator splitting method (also called fractional step method) (cf. Yanenko (1971), Marchuk (1974), Godunov (1959), Strang (1968), Goldman and Kaper (1996)). To demonstrate the main idea, we consider

$$\frac{\partial u}{\partial t} = F(u) = Au + Bu; \quad u(t_0) = u_0, \tag{D.27}$$

where $F(u) = Au + Bu$ is frequently split according to physical components, such as density, velocity, energy or dimension.

We first consider the Strang's operator splitting method. For a given time step size $\tau > 0$, let $t_n = n\tau$, and let u^n be the approximation of $u(t_n)$. Let us formally write the solution $u(x,t)$ of (D.27) as

$$u(t) = e^{t(A+B)}u_0 =: S(t)u_0. \tag{D.28}$$

Similarly, denote by $S_1(t) := e^{tA}$ the solution operator for $u_t = Au$, and by $S_2(t) := e^{tB}$ the solution operator for $u_t = Bu$. Then the first-order operator splitting is based on the approximation

$$u^{n+1} \approx S_2(\tau)S_1(\tau)u^n, \tag{D.29}$$

or on the one with the roles of S_2 and S_1 reversed. To construct a second-order scheme, the Strang splitting (cf. Strang (1968)) can be used, in which the solution $S(t_n)u_0$ is approximated by

$$u^{n+1} \approx S_2(\tau/2)S_1(\tau)S_2(\tau/2)u^n, \tag{D.30}$$

or by the one with the roles of S_2 and S_1 reversed.

A fourth-order symplectic time integrator (cf. Yoshida (1990), Lee and Fornberg (2003)) for (D.27) is as follows:

$$\begin{aligned} &u^{(1)} = e^{2w_1 A\tau}\, u^n, \quad u^{(2)} = e^{2w_2 B\tau}\, u^{(1)}, \quad u^{(3)} = e^{2w_3 A\tau}\, u^{(2)}, \\ &u^{(4)} = e^{2w_4 B\tau}\, u^{(3)}, \quad u^{(5)} = e^{2w_3 A\tau}\, u^{(4)}, \quad u^{(6)} = e^{2w_2 B\tau}\, u^{(5)}, \\ &u^{n+1} = e^{2w_1 A\tau}\, u^{(6)}; \end{aligned} \tag{D.31}$$

or equivalently,

$$\begin{aligned} u^{n+1} \approx & S_1(2w_1\tau)S_2(2w_2\tau)S_1(2w_3\tau)S_2(2w_4\tau) \\ & S_1(2w_3\tau)S_2(2w_2\tau)S_1(2w_1\tau)u^n, \end{aligned} \tag{D.32}$$

where

$$\begin{aligned} w_1 &= 0.33780\,17979\,89914\,40851, \\ w_2 &= 0.67560\,35959\,79828\,81702, \\ w_3 &= -0.08780\,17979\,89914\,40851, \\ w_4 &= -0.85120\,71979\,59657\,63405. \end{aligned} \tag{D.33}$$

References

Abramowitz M, Stegun I (1964) Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover publications

Adams R (1975) Sobolov Spaces. Acadmic Press, New York

Aguirre J, Rivas J (2005) Hermite pseudospectral approximations. An error estimate. Journal of Mathematical Analysis and Applications 304(1):189–197

Ali I, Brunner H, Tang T (2009) Spectral methods for pantograph-type differential and integral equations with multiple delays. Front Math China 4:49–61

Allen S, Cahn J (1979) A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening. Acta Metall Mater 27:1085–1095

Alpert B, Rokhlin V (1991) A fast algorithm for the evaluation of Legendre expansions. SIAM J Sci Stat Comput 12:158–179

Anderson D, McFadden G, Wheeler A (1998) Diffuse-interface methods in fluid mechanics. Annual review of fluid mechanics, Vol 30 30:139–165

Arrow K, Hurwicz L, Uzawa H (1958) Studies in Nonlinear Programming. Stanford University Press

Ascher U, Mattheij R, Russell R (1995) Numerical solution of boundary value problems for ordinary differential equations, Classics in Applied Mathematics, vol 13. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, corrected reprint of the 1988 original

Auteri F, Quartapelle L (2000) Galerkin-Legendre spectral method for the 3D Helmholtz equation. J Comput Phys 161(2):454–483

Azaiez M, Shen J, Xu C, Zhuang Q (2008) A Laguerre-Legendre spectral method for the Stokes problem in a semi-infinite channel. SIAM J Numer Anal 47:271–292

Babuška I (1973) The finite element method with Lagrangian multipliers. Numer Math 20:179–192

Babuška I, Aziz A (1972) Survey lectures on the mathematical foundations of the finite element method. In: The mathematical foundations of the finite element method with applications to partial differential equations (Proc. Sympos., Univ. Maryland, Baltimore, Md., 1972), Academic Press, New York, pp 1–359, with the collaboration of G. Fix and R. B. Kellogg

Bao W, Shen J (2005) A Fourth-order time-splitting Laguerre-Hermite pseudo-spectral method for Bose-Einstein condensates. SIAM J Sci Comput 26:2110–2028

Bao W, Jaksch D, Markowich P (2003a) Numerical solution of the Gross-Pitaevskii equation for Bose-Einstein condensation. J Comput Phys 187(1):318–342

Bao W, Jin S, Markowich P (2003b) Numerical study of time-splitting spectral discretizations of nonlinear Schrödinger equations in the semiclassical regimes. SIAM J Sci Comput 25(1):27–64 (electronic)

Barrett R, Berry M, Chan T, et al (1994) Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA

J. Shen et al., *Spectral Methods: Algorithms, Analysis and Applications*,
Springer Series in Computational Mathematics 41, DOI 10.1007/978-3-540-71041-7,

Barthelmann V, Novak E, Ritter K (2000) High dimensional polynomial interpolation on sparse grids. Adv Comput Math 12:273–288
Bellen A, Zennaro M (2003) Numerical Methods for Delay Differential Equations. Oxford University Press, Oxford
Bellman R (1961) Adaptive Control Processes: A Guided Tour. Princeton University Press, Princeton, N.J.
Berenger J (1994) A perfectly matched layer for the absorption of electromagnetic waves. J Comput Phys 114(2):185–200
Bergh J, Löfström J (1976) Interpolation Spaces. An Introduction. Springer-Verlag, Berlin, grundlehren der Mathematischen Wissenschaften, No. 223
Bernardi C, Maday Y (1992a) Approximations Spectrales de Problèmes aux Limites Elliptiques. Springer-Verlag, Paris
Bernardi C, Maday Y (1992b) Polynomial interpolation results in Sobolev spaces. Journal of Computational and Applied Mathematics 43(1-2):53–80
Bernardi C, Maday Y (1997) Spectral Method. In: Ciarlet P, Lions L (eds) Handbook of Numerical Analysis, V. 5 (Part 2), North-Holland
Bernardi C, Maday Y (1999) Uniform inf-sup conditions for the spectral discretization of the Stokes problem. Math Models Methods Appl Sci 9(3):395–414
Bernardi C, Coppoletta G, Maday Y (1992) Some spectral approximations of two-dimensional fourth-order problems. Math Comp 59:63–76
Bernardi C, Dauge M, Maday Y (1999) Spectral Methods for Axisymmetric Domains. Gauthier-Villars, Éditions Scientifiques et Médicales Elsevier, Paris
Bjørstad P, Tjostheim B (1997) Efficient algorithms for solving a fourth order equation with the spectral-galerkin method. SIAM J Sci Comput 18:621–632
Boyd J (1978) The choice of spectral functions on a sphere for boundary and eigenvalue problems: a comparison of Chebyshev, Fourier and associated Legendre expansions. Monthly Weather Rev 106:1184–1191
Boyd J (1980) The rate of convergence of Hermite function series. Math Comp 35(152):1309–1316
Boyd J (1982) The optimization of convergence for Chebyshev polynomial methods in an unbounded domain. Journal of Computational Physics 45(1):43–79
Boyd J (1987a) Orthogonal rational functions on a semi-infinite interval. J Comput Phys 70:63–88
Boyd J (1987b) Spectral methods using rational basis functions on an infinite interval. J Comput Phys 69:112–142
Boyd J (2001) Chebyshev and Fourier Spectral Methods, 2nd edn. Dover Publications Inc., Mineola, NY
Boyd J (2004) Prolate spheroidal wavefunctions as an alternative to Chebyshev and Legendre polynomials for spectral element and pseudospectral algorithms. J Comput Phys 199(2):688–716
Brenner S, Scott L (2008) The Mathematical Theory of Finite Element Methods, 3rd edn. Springer Verlag
Brezzi F (1974) On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. Rev Française Automat Informat Recherche Opérationnelle Sér Rouge 8(R-2):129–151
Brezzi F, Fortin M (1991) Mixed and Hybrid Finite Element Methods. Springer-Verlag, New York
Brown DL, Cortez R, Minion ML (2001) Accurate projection methods for the incompressible Navier-Stokes equations. J Comput Phys 168(2):464–499
Brunner H (2004) Collocation Methods for Volterra Integral and Related Functional Differential Equations, Cambridge Monographs on Applied and Computational Mathematics, vol 15. Cambridge University Press, Cambridge
Brunner H, Hu Q (2007) Optimal superconvergence results for delay integro-differential equations of pantograph type. SIAM J Numer Anal 45:986–1004
Brunner H, Tang T (1989) Polynomial spline collocation methods for the nonlinear Basset equation. Comput Math Appl 18(5):449–457
Bungartz H, Griebel M (2004) Sparse grids. Acta Numer 13:147–269

Butcher J (2008) Numerical Methods for Ordinary Differential Equations, 2nd edn. John Wiley & Sons Ltd., Chichester
Butzer P, Nessel R (1971) Fourier Analysis and Approximation. Academic Press, New York, volume 1: One-dimensional theory, Pure and Applied Mathematics, Vol. 40
Caffarelli L, Muler N (1995) An L^∞ bound for solutions of the Cahn-Hilliard equation. Arch Rational Mech Anal 133(2):129–144
Cahn J, Hilliard J (1958) Free energy of a nonuniform system, I: Interfacial free energy. J Chem Phys 28:258
Canuto C, Quarteroni A (1982) Approximation results for orthogonal polynomials in Sobolev spaces. Math Comp 38:67–86
Canuto C, Quarteroni A (1985) Preconditioner minimal residual methods for Chebyshev spectral calculations. J Comput Phys 60:315–337
Canuto C, Quarteroni A (1994) Variational methods in the theoretical analysis of spectral approximations. In: Voigt R, Gottlieb D, Hussaini M (eds) Spectral Methods for Partial Differential Equations, SIAM, pp 55–78
Canuto C, Hussaini M, Quarteroni A, Zang T (1987) Spectral Methods in Fluid Dynamics. Springer-Verlag
Canuto C, Hussaini M, Quarteroni A, Zang T (2006) Spectral Methods. Scientific Computation, Springer-Verlag, Berlin, fundamentals in single domains
Canuto C, Hussaini M, Quarteroni A, Zang T (2007) Spectral Methods. Scientific Computation, Springer, Berlin, evolution to complex geometries and applications to fluid dynamics
Céa J (1964) Approximation variationnelle des problèmes aux limites. Ann Inst Fourier (Grenoble) 14(fasc. 2):345–444
Chen F, Shen J (2011) Efficient spectral Galerkin methods for coupled systems of elliptic equations with separable boundary conditions. Preprint
Chen L (2002) Phase-field models for microstructure evolution. Annual Review of Material Research 32:113
Chen L, Shen J (1998) Applications of semi-implicit Fourier-spectral method to phase-field equations. Comput Phys Comm 108:147–158
Chen Q, Gottlieb D, Hesthaven J (2005) Spectral methods based on prolate spheroidal wave functions for hyperbolic PDEs. SIAM J Numer Anal 43(5):1912–1933 (electronic)
Chen Y, Tang T (2010) Convergence analysis of the Jacobi spectral-collocation methods for Volterra integral equations with a weakly singular kernel. Math Comp 79:147–167
Cheney E (1998) Introduction to Approximation Theory. AMS Chelsea Publishing, Providence, RI, reprint of the second (1982) edition
Chorin A (1968) Numerical solution of the Navier-Stokes equations. Math Comp 22:745–762
Christov C (1982) A complete orthogonal system of functions in $l^2(-\infty,\infty)$ space. SIAM J Appl Math 42:1337–1344
Ciarlet P (1978) The Finite Element Method for Elliptic Problems. North-Holland
Ciskowski R, Brebbia C (1991) Boundary Element Methods in Acoutics. Kluwer Academic Publishers
Cloot A, Weideman J (1992) An adaptive algorithm for spectral computations on unbounded domains. J Comput Phys 102(2):398–406
Colin T, Ghidaglia J (2001) An initial-boundary value problem for the Korteweg-de Vries equation posed on a finite interval. Adv Differential Equations 6(12):1463–1492
Condette N, Melcher C, Süli E (2011) Spectral approximation of pattern-forming nonlinear evolution equations with double-well potentials of quadratic growth. Math Comp 80(273):205–223
Cooley J, Tukey J (1965) An algorithm for the machine calculation of complex Fourier series. Math Comp 19:297–301
Cummings P, Feng X (2006) Sharp regularity coefficient estimates for complex-valued acoustic and elastic Helmholtz equations. Math Models Methods Appl Sci 16(1):139–160
Davis P (1975) Interpolation and Approximation. Dover Publications, Inc, New Year
Davis P, Rabinowitz P (1984) Methods of Numerical Integration, 2nd edn. Computer Science and Applied Mathematics, Academic Press Inc., Orlando, FL

Delves L, Mohanmed J (1985) Computational Methods for Integral Equations. Cambridge University Press

Demaret P, Deville M (1991) Chebyshev collocation solution of the Navier-Stokes equations using multi-domain decomposition and finite element preconditioning. J Comput Phys 95:359–386

Demkowicz L, Ihlenburg F (2001) Analysis of a coupled finite-infinite element method for exterior Helmholtz problems. Numer Math 88(1):43–73

Deville M, Mund E (1985) Chebyshev pseudospectral solution of second-order elliptic equations with finite element preconditioning. J Comput Phys 60:517–533

Deville M, Fischer P, Mund E (2002) High-Order Methods for Incompressible Fluid Flow, Cambridge Monographs on Applied and Computational Mathematics, vol 9. Cambridge University Press, Cambridge

Djidjeli K, Price W, Twizell E, Wang Y (1995) Numerical methods for the solution of the third- and fifth-order dispersive Korteweg-de-Vries equations. J Comput Appl Math 58:307–336

Dobrovol'skiĭ N, Roshchenya A (1998) On the number of points in a lattice in a hyperbolic cross. Mat Zametki 63(3):363–369

Don W, Gottlieb D (1994) The Chebyshev-Legendre method: implementing Legendre methods on Chebyshev points. SIAM J Numer Anal 31:1519–1534

Douglas J, Santos J, Sheen D, Bennethum L (1993) Frequency domain treatment of one-dimensional scalar waves. Math Models Methods Appl Sci 3(2):171–194

E W, Liu JG (1996) Projection method. II. Godunov-Ryabenki analysis. SIAM J Numer Anal 33(4):1597–1621

Eisen H, Heinrichs W, Witsch K (1991) Spectral collocation methods and polar coordinate singularities. J Comput Phys 96:241–257

El-Daou M, Ortiz E (1998) The tau method as an analytic tool in the discussion of equivalence results across numerical methods. Computing 60(4):365–376

Elman H, Silvester D, Wathen A (2005) Finite elements and fast iterative solvers: with applications in incompressible fluid dynamics. Numerical Mathematics and Scientific Computation, Oxford University Press, New York

Elnagar G, Kazemi M (1996) Chebyshev spectral solution of nonlinear Volterra-Hammerstein integral equations. J Comput Appl Math 76(1-2):147–158

Engquist B, Majda A (1977) Absorbing boundary conditions for the numerical simulation of waves. Math Comp 31(139):629–651

Eyre D (1998) Unconditionally gradient stable time marching the Cahn-Hilliard equation. In: Computational and mathematical models of microstructural evolution (San Francisco, CA, 1998), Mater. Res. Soc. Sympos. Proc., vol 529, MRS, Warrendale, PA, pp 39–46

Faber V, Manteuffel T (1984) Necessary and sufficient conditions for the existence of a conjugate gradient method. SIAM J Numer Anal 21(2):352–362

Fang Q, Nicholls D, Shen J (2007) A stable, high-order method for three-dimensional, bounded-obstacle, acoustic scattering. J Comput Phys 224(2):1145–1169

Fang Q, Shen J, Wang L (2009) An efficient and accurate spectral method for acoustic scattering in elliptic domains. Numer Math: Theory, Methods Appl 2:258–274

Finlayson B (1972) The Method of Weighted Residuals and Variational Principles. Academic Press, New York

Fok J, Guo B, Tang T (2002) Combined Hermite spectral-finite difference method for the Fokker-Planck equation. Math Comp 71(240):1497–1528 (electronic)

Folland G (1992) Fourier Analysis and Its Applications. The Wadsworth & Brooks/Cole Mathematics Series, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA

Fornberg B (1995) A pseudospectral approach for polar and spherical geometries. SIAM J Sci Comput 16:1071–1081

Fox L, Mayers D, Ockendon J, Taylor A (1971) On a functional differential equation. J Inst Math Appl 8:271–307

Franken P, Deville M, Mund E (1990) On the spectrum of the iteration operator associated to the finite element preconditioning of chebyshev collocation calculations. Comput Methods in Appl Mech Eng 80:295–304

Fujiwara H (2006) High-accurate numerical method for integral equations of the first kind under multiple-precision arithmetic. Preprint, RIMS, Kyoto University
Funaro D (1992) Polynomial Approxiamtions of Differential Equations. Springer-Verlag
Funaro D, Kavian O (1990) Approximations of some diffusion evolution equations in unbounded domains by Hermite functions. Math Comp 57:597–619
Gautschi W (2004) Orthogonal polynomials: computation and approximation. Numerical Mathematics and Scientific Computation, Oxford University Press, New York, oxford Science Publications
Gear C (1971) Numerical Initial Value Problems in Ordinary Differential Equations. Prentice Hall, NJ, USA
Gerdes K, Demkowicz L (1996) Solution of 3D-Laplace and Helmholtz equations in exterior domains using $hp-$infinite elements. Comput Methods Appl Mech Engrg 137:239–273
Girault V, Raviart P (1986) Finite Element Methods for Navier-Stokes Equations. Springer-Verlag
Glowinski R (2003) Finite Element Methods for Incompressible Viscous Flow. In: Handbook of Numerical Analysis, Vol. IX, Handb. Numer. Anal., IX, North-Holland, Amsterdam, pp 3–1176
Godunov S (1959) Finite difference methods for numerical computations of discontinuous solutions of the equations of fluid dynamics. Mat Sb 47(3):271–295, in Russian
Goldman D, Kaper T (1996) Nth-order operator splitting schemes and nonreversible systems. SIAM J Numer Anal 33(1):349–367
Golub G, Van Loan C (1996) Matrix Computations, 3rd edn. Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD
Gottlieb D, Lustman L (1983) The spectrum of the Chebyshev collocation operator for the heat equation. SIAM J Numer Anal 20:909–921
Gottlieb D, Orszag S (1977) Numerical Analysis of Spectral Methods: Theory and Applications. SIAM-CBMS, Philadelphia
Gottlieb D, Shu C (1997) On the Gibbs phenomenon and its resolution. SIAM Rev 39(4):644–668
Gottlieb D, Hussaini MY, Orszag SA (1984) Theory and applications of spectral methods. In: Spectral methods for partial differential equations (Hampton, Va., 1982), SIAM, Philadelphia, PA, pp 1–54
Gottlieb D, Shu C, Solomonoff A, Vandeven H (1992) On the Gibbs phenomenon I: recovering exponential accuracy from the Fourier partial sum of a nonperiodic analytic function. Journal of Computational and Applied Mathematics 43(1-2):81–98
Goubet O, Shen J (2007) On the dual Petrov-Galerkin formulation of the KdV equation on a finite interval. Adv Differential Equations 12(2):221–239
Greengard L, Rokhlin V (1987) A fast algorithm for particle simulations. J Comput Phys 73:325–348
Griebel M, Hamaekers J (2007) Sparse grids for the Schrödinger equation. M2AN Math Model Numer Anal 41(2):215–247
Grosch C, Orszag S (1977) Numerical solution of problems in unbounded regions: coordinates transforms. J Comput Phys 25:273–296
Gross E (1961) Structure of a quantized vortex in boson systems. Nuovo Cimento (10) 20:454–477
Grote M, Keller J (1995) On nonreflecting boundary conditions. J Comput Phys 122(2):231–243
Guermond J, Shen J (2003) A new class of truly consistent splitting schemes for incompressible flows. J Comput Phys 192(1):262–276
Guermond J, Shen J (2004) On the error estimates of rotational pressure-correction projection methods. Math Comp 73:1719–1737
Guermond J, Minev P, Shen J (2006) An overview of projection methods for incompressible flows. Comput Methods Appl Mech Engrg 195:6011–6045
Guo B (1998a) Gegenbauer approximation and its applications to differential equations on the whole line. J Math Anal Appl 226:180–206
Guo B (1998b) Spectral Methods and Their Applications. World Scientific Publishing Co. Inc., River Edge, NJ
Guo B (1999) Error estimation of Hermite spectral method for nonlinear partial differential equations. Math Comp 68(227):1067–1078

Guo B (2000) Jacobi approximations in certain Hilbert spaces and their applications to singular differential equations. J Math Anal Appl 243:373–408
Guo B, Shen J (2001) On spectral approximations using modified Legendre rational functions: application to the Korteweg-de Vries equation on the half line. Indiana Univ Math J 50:181–204
Guo B, Shen J (2008) Irrational approximations and their applications to partial differential equations in exterior domains. Adv Comput Math 28(3):237–267
Guo B, Wang L (2001) Jacobi interpolation approximations and their applications to singular differential equations. Advances in Computational Mathematics 14:227–276
Guo B, Wang L (2004) Jacobi approximations in non-uniformly Jacobi-weighted Sobolev spaces. J Approx Theory 128(1):1–41
Guo B, Xu C (2000) On two-dimensional unsteady incompressible fluid flow in an infinite strip. Math Methods Appl Sci 23(18):1617–1636
Guo B, Shen J, Wang Z (2000) A rational approximation and its applications to differential equations on the half line. J Sci Comp 15:117–147
Guo B, Shen J, Wang Z (2002) Chebyshev rational spectral and pseudospectral methods on a semi-infinite interval. Internat J Numer Methods Engrg 53(1):65–84, *p* and *hp* finite element methods: mathematics and engineering practice (St. Louis, MO, 2000)
Guo B, Shen J, Xu C (2003) Spectral and pseudospectral approximations using Hermite functions: Application to the Dirac equation. Adv Comput Math 19:35–55
Guo B, Shen J, Xu C (2005) Generalized Laguerre approximation and its applications to exterior problems. J Comput Math 23(2):113–130
Guo B, Shen J, Wang L (2006a) Optimal spectral-Galerkin methods using generalized Jacobi polynomials. J Sci Comput 27(1-3):305–322
Guo B, Wang L, Wang Z (2006b) Generalized Laguerre interpolation and pseudospectral method for unbounded domains. SIAM J Numer Anal 43(6):2567–2589 (electronic)
Guo B, Shen J, Wang L (2009) Generalized Jacobi polynomials/functions and their applications. Appl Numer Math 59(5):1011–1028
Haidvogel D, Zang T (1979) The accurate solution of Poisson's equation by expansion in Chebyshev polynomials. J Comput Phys 30:167–180
Hairer E, Wanner G (1996) Solving Ordinary Differential Equations. II, Springer Series in Computational Mathematics, vol 14, 2nd edn. Springer-Verlag, Berlin, stiff and differential-algebraic problems
Hairer E, Nørsett S, Wanner G (1993) Solving Ordinary Differential Equations. I. Nonstiff Problems, 2nd edition, Springer Series in Computational Mathematics, vol 8. Springer-Verlag, Berlin
Haldenwang P, Labrosse G, Abboudi S, Deville M (1984) Chebyshev 3-D spectral and 2-D pseudospectral solvers for the Helmholtz equation. J Comput Phys 55:115–128
Harari I, Hughes T (1992) Analysis of continuous formulations underlying the computation of time-harmonic acoustics in exterior domains. Comput Methods Appl Mech Engrg 97(1): 103–124
Hardy G, Littlewood J, Pólya G (1952) Inequalities. Cambridge University Press, UK
Heinrichs W (1989) Spectral methods with sparse matrices. Numer Math 56:25–41
Hestenes M, Stiefel E (1952) Methods of conjugate gradients for solving linear systems. J Res Nat Bur Stand 49(6):409–436
Hesthaven J, Warburton T (2008) Nodal Discontinuous Galerkin Methods, Texts in Applied Mathematics, vol 54. Springer, New York, algorithms, analysis, and applications
Hesthaven J, Gottlieb S, Gottlieb D (2007) Spectral Methods for Time-Dependent Problems. Cambridge Monographs on Applied and Computational Mathematics, Cambridge
Hu Z, Wise S, Wang C, Lowengrub J (2009) Stable and efficient finite-difference nonlinear-multigrid schemes for the phase field crystal equation. J Comput Phys 228(15):5323–5339
Huang W, Sloan D (1992) The pseudospectral method for third-order differential equations. SIAM J Numer Anal 29(6):1626–1647

Huang W, Sloan D (1993) Pole condition for singular problems: the pseudospectral approximation. J Comput Phys 107:254–261
Hyman J, Nicolaenko B (1986) The Kuramoto-Sivashinsky equation: a bridge between PDEs and dynamical systems. Phys D 18(1-3):113–126
Ihlenburg F, Babuška I (1995) Finite element solution of the Helmholtz equation with high wave number, part I: the h-version of FEM. Computers Math Applic 30:9–37
Ihlenburg F, Babuška I (1997) Finite element solution of the Helmholtz equation with high wave number. II. The *h-p* version of the FEM. SIAM J Numer Anal 34(1):315–358
Iserles A (1993) On the generalized pantograph functional differential equation. Europ J Appl Math 4:1–38
Johnston H, Liu JG (2004) Accurate, stable and efficient Navier-Stokes solvers based on explicit treatment of the pressure term. J Comput Phys 199(1):221–259
Karniadakis G, Sherwin S (1999) Spectral/hp Element Methods for CFD. Oxford UniversityPress
Karniadakis G, Sherwin S (2005) Spectral/*hp* Element Methods for Computational Fluid Dynamics, 2nd edn. Numerical Mathematics and Scientific Computation, Oxford University Press, New York
Kassam A, Trefethen L (2005) Fourth-order time stepping for stiff PDEs. SIAM J Sci Comput 26(4):1214–1233
Kawahara R (1972) Oscillatory solitary waves in dispersive media. J Phys Soc Japan 33:260–264
Keller J, Givoli D (1989) Exact non-reflecting boundary conditions. J Comput Phys 82:172–192
Kessler D, Nochetto R, Schmidt A (2004) A posteriori error control for the Allen-Cahn problem: circumventing Gronwall's inequality. M2AN Math Model Numer Anal 38(1):129–142
Kichenassamy S, Olver P (1992) Existence and nonexistence of solitary wave solutions to higher-order model evolution equations. SIAM J Math Anal 23(5):1141–1166
Kim S, Parter S (1997) Preconditioning Chebyshev spectral collocation by finite difference operators. SIAM J Numer Anal 34, No. 3:939–958
Körner T (1988) Fourier Analysis. Cambridge University Press, Cambridge
Korobov N (1992) Exponential Sums and Their Applications, Mathematics and its Applications (Soviet Series), vol 80. Kluwer Academic Publishers Group, Dordrecht, translated from the 1989 Russian original by Y.N. Shakhov
Korteweg D, de Vries G (1895) On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. Philosophical Magazine 39:422–443
Kreiss HO, Oliger J (1979) Stability of the Fourier method. SIAM J Numer Anal 16(3):421–433
Krylov V (1962) Approximate Calculation of Integrals. Macmillan New York
Kuramoto Y, Tsuzuki T (1976) Persistent propagation of concentration waves in dissipative media far from thermal equilibrium. Prog Theoret Phys 55(2):356–369
Lambert J (1991) Numerical Methods for Ordinary Differential Systems: The initial value problem. John Wiley & Sons Ltd., Chichester
Le Maître O, Knio O (2010) Spectral Methods for Uncertainty Quantification. Scientific Computation, Springer, New York, with applications to computational fluid dynamics
Leboeuf P, Pavloff N (2001) Bose-Einstein beams: coherent propagation through a guide. Phys Rev A 64:aritcle 033,602
Lee J, Fornberg B (2003) A split step approach for the 3-d Maxwell's equations. J Comput and Appl Math 158(2):485–505
Lether F (1978) On the construction of Gauss-Legndre quadrature rules. J Comp Appl Math 4(1):47–52
LeVeque R (2007) Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA
Levin A, Lubinsky D (1992) Christoffel functions, orthogonal polynomials, and Nevai's conjecture for Freud weights. Constr Approx 8(4):461–533
Lions J, Magenes E (1968) Problemes aux limites non homogenes et applications. Vol. 1-2. Dunod, Paris

Liu C, Shen J (2003) A phase field model for the mixture of two incompressible fluids and its approximation by a Fourier-spectral method. Physica D 179(3-4):211–228
Liu JG, Liu J, Pego RL (2007) Stability and convergence of efficient Navier-Stokes solvers via a commutator estimate. Comm Pure Appl Math 60(10):1443–1487
Liu Y, Liu L, Tang T (1994) The numerical computation of connecting orbits in dynamical systems: a rational spectral approach. J Comput Phys 111(2):373–380
Lopez J, Perry A (1992) Axisymmetric vortex breakdown. III. Onset of periodic flow and chaotic advection. J Fluid Mech 234:449–471
Lopez J, Shen J (1998) An efficient spectral-projection method for the Navier-Stokes equations in cylindrical geometries I. axisymmetric cases. J Comput Phys 139:308–326
Lopez J, Marques F, Shen J (2002) An efficient spectral-projection method for the Navier-Stokes equations in cylindrical geometries. II. Three-dimensional cases. J Comput Phys 176(2):384–401
Lowengrub J, Truskinovsky L (1998) Quasi-incompressible Cahn-Hilliard fluids and topological transitions. R Soc Lond Proc Ser A Math Phys Eng Sci 454(1978):2617–2654
Lynch R, Rice J, Thomas D (1964) Direct solution of partial differential equations by tensor product methods. Numer Math 6:185–199
Ma H, Sun W (2000) A Legendre-Petrov-Galerkin method and Chebyshev-collocation method for the third-order differential equations. SIAM J Numer Anal 38(5):1425–1438
Ma H, Sun W (2001) Optimal error estimates of the Legendre-Petrov-Galerkin method for the Korteweg-de Vries equation. SIAM J Numer Anal 39(4):1380–1394 (electronic)
Ma H, Sun W, Tang T (2005) Hermite spectral methods with a time-dependent scaling for parabolic equations in unbounded domains. SIAM J Numer Anal 43(1):58–75
Maday Y, Pernaud-Thomas B, Vandeven H (1985) Reappraisal of Laguerre type spectral methods. La Recherche Aerospatiale 6:13–35
Maday Y, Meiron D, Patera A, Rønquist E (1993) Analysis of iterative methods for the steady and unsteady Stokes problem: application to spectral element discretizations. SIAM J Sci Comput 14(2):310–337
Marchuk G (1974) Numerical Methods in Weather Prediction. Academic Press, New York
Marion M, Temam R (1998) Navier-Stokes Equations: Theory and Approximation. In: Handbook of Numerical Analysis, Vol. VI, Handb. Numer. Anal., VI, North-Holland, Amsterdam, pp 503–688
Mastroianni G, Occorsio D (2001a) Lagrange interpolation at Laguerre zeros in some weighted uniform spaces. Acta Math Hungar 91(1-2):27–52
Mastroianni G, Occorsio D (2001b) Optimal systems of nodes for lagrange interpolation on bounded intervals. a survey. Journal of Computational and Applied Mathematics 134:325–341
McLachlan N (1951) Theory and Applications of Mathieu Functions. Oxford Press, London
Melenk J (1995) On generalized finite element methods. PhD thesis, University of Maryland, College Park
Merryfield W, Shizgal B (1993) Properties of collocation third-derivative operators. J Comput Phys 105(1):182–185
Morse P, Feshback H (1953) Methods of Theoretical Physics. McGraw-Hill, New York
Nagashima H, Kawahara M (1981) Computer simulation of solitary waves of the nonlinear wave equations. J Phys Soc Japan 50:3792–3800
Nicholls D, Reitich F (2003) Analytic continuation of Dirichlet-Neumann operators. Numer Math 94(1):107–146
Nicholls D, Shen J (2006) A stable, high–order method for two–dimensional bounded–obstacle scattering. SIAM J Sci Comput 28:1398–1419
Nicholls D, Shen J (2009) A rigorous numerical analysis of the transformed field expansion method. SIAM J Numer Anal 47(4):2708–2734
Nicolaenko B, Scheurer B, Temam R (1985) Some global dynamical properties of the Kuramoto-Sivashinsky equations: nonlinear stability and attractors. Phys D 16(2):155–183

Orszag S (1971) Accurate solution of the Orr-Sommerfeld equation. Journal of Fluid Mechanics 50:689–703, [Combines Chebyshev method with QR or QZ matrix eigensolver to solve a linear stability problem.]
Orszag S (1974) Fourier series on spheres. Monthly Weather Rev 102:56–75
Orszag S (1980) Spectral methods for complex geometries. J Comput Phys 37:70–92
Orszag S, Patera A (1983) Secondary instability of wall-bounded shear flows. J Fluid Mech 128:347–385
Pan V (1997) Solving a polynomial equation: some history and recent progress. SIAM Rev 39(2):187–220
Parkes E, Zhu Z, Duffy B, Huang H (1998) Sech-polynomial travelling solitary-wave solutions of odd-order generalized KdV equations. Physics Letters A 248(2-4):219–224
Parter S, Rothman E (1995) Preconditioning lengendre spectral collocation approximations to elliptic problems. SIAM J Numer Anal 32, No. 2:333–385
Pitaevskii L (1961) Vortex lines in an imperfect bose gase. Sov Phys JETP 13:451–454
Potts D, Steidl G, Tasche M (1998) Fast algorithms for discrete polynomial transforms. Math Comp 67(224):1577–1590
Qu C, Wong R (1988) Szego's conjecture on Lebesgue constants for Legendre series. Pacific J Math 135(1):157–188
Quarteroni A, Valli A (2008) Numerical Approximation of Partial Differential Equations. Springer Verlag
Ragozin D (1970) Polynomial approximation on compact manifolds and homogeneous spaces. Trans Amer Math Soc 150:41–53
Ragozin D (1971) Constructive polynomial approximation on spheres and projective spaces. Trans Amer Math Soc 162:157–170
Reddy S, Weideman J (2005) The accuracy of the Chebyshev differencing method for analytic functions. SIAM J Numer Anal 42(5):2176–2187
Rivlin T (1974) The Chebyshev Polynomials. Jong Wiley and Sons
Rokhlin V, Tygert M (2006) Fast algorithms for spherical harmonic expansions. SIAM J Sci Comput 27(6):1903–1928 (electronic)
Saad Y (2003) Iterative Methods for Sparse Linear Systems, 2nd edn. Society for Industrial and Applied Mathematics, Philadelphia, PA
Saad Y, Schultz M (1986) GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM J Sci Statist Comput 7(3):856–869
Shen J (1991) Hopf bifurcation of the unsteady regularized driven cavity flows. J Comput Phys 95:228–245
Shen J (1994) Efficient spectral-Galerkin method I. direct solvers for second- and fourth-order equations by using Legendre polynomials. SIAM J Sci Comput 15:1489–1505
Shen J (1995) Efficient spectral-Galerkin method II. direct solvers for second- and fourth-order equations by using Chebyshev polynomials. SIAM J Sci Comput 16:74–87
Shen J (1996) Efficient Chebyshev-Legendre Galerkin methods for elliptic problems. In: Ilin AV, Scott R (eds) Proceedings of ICOSAHOM'95, Houston J. Math., pp 233–240
Shen J (1997) Efficient spectral-Galerkin methods III. polar and cylindrical geometries. SIAM J Sci Comput 18:1583–1604
Shen J (1999) Efficient spectral-Galerkin methods IV. spherical geometries. SIAM J Sci Comput 20:1438–1455
Shen J (2000) A new fast Chebyshev-Fourier algorithm for the Poisson-type equations in polar geometries. Appl Numer Math 33:183–190
Shen J (2003) A new dual-Petrov-Galerkin method for third and higher odd-order differential equations: application to the KDV equation. SIAM J Numer Anal 41:1595–1619
Shen J, Tang T (2006) Spectral and High-Order Methods with Applications. Science Press, Beijing
Shen J, Wang L (2004) Error analysis for mapped Legendre spectral and pseudospectral methods. SIAM J Numer Anal 42:326–349
Shen J, Wang L (2005) Spectral approximation of the Helmholtz equation with high wave numbers. SIAM J Numer Anal 43(2):623–644

Shen J, Wang L (2006) Laguerre and composite Legendre-Laguerre dual-Petrov-Galerkin methods for third-order equations. Discrete Contin Dyn Syst Ser B 6(6):1381–1402 (electronic)
Shen J, Wang L (2007a) Analysis of a spectral-Galerkin approximation to the Helmholtz equation in exterior domains. SIAM J Numer Anal 45:1954–1978
Shen J, Wang L (2007b) Fourierization of the Legendre-Galerkin method and a new space-time spectral method. Appl Numer Math 57(5-7):710–720
Shen J, Wang L (2007c) Legendre and Chebyshev dual-Petrov-Galerkin methods for hyperbolic equations. Computer Methods in Applied Mechanics and Engineering 196(37-40):3785–3797
Shen J, Wang L (2008) On spectral approximations in elliptical geometries using Mathieu functions. Math Comp 78(266):815–844
Shen J, Wang L (2009) Some recent advances on spectral methods for unbounded domains. Communcations in Computational Physics 5(2-4):195–241
Shen J, Wang L (2010) Sparse spectral approximations of high-dimensional problems based on hyperbolic cross. SIAM J Numer Anal 48:1087–1109
Shen J, Yang X (2010) Numerical Approximations of Allen-Cahn and Cahn-Hilliard Equations. Discrete and Continuous Dynamical Systems Series A 28:1669–1691
Shen J, Yu H (2010) Efficient spectral sparse grid methods and applications to high dimensional elliptic problems. SIAM J Sci Comput 32:3228–3250
Slepian D, Pollak H (1961) Prolate spheroidal wave functions, Fourier analysis and uncertainty. I. Bell System Tech J 40:43–63
Smolyak S (1960) Quadrature and interpolation formulas the classes W_s^a and E_s^a. Dokl Akad Nauk SSSR 131:1028–1031, russian, Engl. Transl.: Soviet Math. Dokl. 1:384–387, 1963
Stein E, Shakarchi R (2003) Fourier Analysis: An Introduction, Princeton Lectures in Analysis, vol 1. Princeton University Press, Princeton, NJ
Strang G (1968) On the construction and comparison of difference schemes. SIAM J Numer Anal 5(3):506–517
Strang G, Fix G (1973) An analysis of the finite element method. Prentice-Hall Inc., Englewood Cliffs, N. J., prentice-Hall Series in Automatic Computation
Swarztrauber P, Spotz W (2000) Generalized discrete spherical harmonic transforms. J Comput Phys 159(2):213–230
Szegö G (1975) Orthogonal Polynomials (fourth edition), vol 23. AMS Coll. Publ.
Tadmor E (1986) The exponential accuracy of Fourier and Chebyshev differencing methods. SIAM J Numer Anal 23(1):1–10
Tang T (1993) The Hermite spectral method for Gaussian-type functions. SIAM J Sci Comput 14:594–606
Tang T, Xu X (2009) Accuracy enhancement using spectral postprocessing for differential equations and integral equations. Commun Comput Phys 5(2-4):779–792
Temam R (1969) Sur l'approximation de la solution des équations de Navier-Stokes par la méthode des pas fractionnaires ii. Arch Rat Mech Anal 33:377–385
Temam R (1984) Navier-Stokes Equations: Theory and Numerical Analysis. North-Holland, Amsterdam
Timan A (1994) Theory of Approximation of Functions of a Real Variable. Dover Publications Inc., New York, translated from the Russian by J. Berry, Translation edited and with a preface by J. Cossar, Reprint of the 1963 English translation
Timmermans L, Minev P, Van De Vosse F (1996) An approximate projection scheme for incompressible flow using spectral elements. Int J Numer Methods Fluids 22:673–688
Trefethen L (2000) Spectral Methods in MATLAB, Software, Environments, and Tools, vol 10. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA
Trott M (1999) Graphica I: The World of Mathematica Graphics. The Imaginary Made Real: The Images of Michael Trott. Wolfram Media, Champaign, IL
Tygert M (2010) Recurrence relations and fast algorithms. Appl Comput Harmon Anal 28(1): 121–128
Van der Vorst H (1992) Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems. SIAM J Sci Statist Comput 13(2):631–644

Wang C, Wang X, Wise S (2010) Unconditionally stable schemes for equations of thin film epitaxy. Discrete Contin Dyn Syst 28(1):405–423
Wang L (2010) Analysis of spectral approximations using prolate spheroidal wave functions. Math Comp 79(270):807–827
Wang L, Guo B (2009) Interpolation approximations based on Gauss-Lobatto-Legendre-Birkhoff quadrature. J Approx Theory 161(1):142–173
Wang L, Shen J (2005) Error analysis for mapped Jacobi spectral methods. J Sci Comput 24(2):183–218
Wang L, Guo B, Wang Z (2002) A generalized Gauss-type quadrature formula and its applications to pseudospectral method. Numer Math J Chinese Univ (English Ser) 11(2):179–196
Wang Z, Guo B (2002) A rational approximation and its applications to nonlinear partial differential equations on the whole line. J Math Anal Appl 274(1):374–403
Wang Z, Guo B (2004) Modified Legendre rational spectral method for the whole line. J Comput Math 22(3):457–474
Watson G (1966) A Treatise on the Theory of Bessel Functions. Cambridge University Press
Wu H, Ma H, Li H (2003) Optimal error estimates of the Chebyshev-Legendre spectral method for solving the generalized Burgers equation. SIAM J Numer Anal 41(2):659–672 (electronic)
Xiu D (2010) Numerical Methods for Stochastic Computations. Princeton University Press, Princeton, NJ, a spectral method approach
Xu C, Guo B (2002) Mixed Laguerre-Legendre spectral method for incompressible flow in an infinite strip. Adv Comput Math 16(1):77–96
Yanenko N (1971) The Method of Fractional Steps: The Solution of Problems of Mathematical Physics In Several Variables. Springer-Verlag, New York
Yoshida H (1990) Construction of higher order symplectic integrators. Phys Lett A 150(5-7):262–268
Yosida K (1980) Functional Analysis. Springer-Verlag, Berlin
Yuan J, Shen J, Wu J (2008) A dual-Petrov-Galerkin method for the Kawahara-type equations. J Sci Comput 34(1):48–63
Zabusky N, Galvin C (1971) Shallow water waves, the Korteveg-de Vries equation and solitons. J Fluid Mech 47:811–824
Zhang X, Guo B (2006) Spherical harmonic-generalized Laguerre spectral method for exterior problems. J Sci Comput 27(1-3):523–537
Zygmund A (2002) Trigonometric Series, 3rd edn. Cambridge Mathematical Library, Cambridge University Press

Index

A
a priori estimates, 175
acoustic scattering, 367
algebraic decay, 276
Allen–Cahn equation, 38, 367, 381
Aubin-Nitsche technique, 121

B
Backward Difference Method, 452
Bernstein inequality, 33
BiConjugate Gradient Stabilized (BiCGStab) Method, 439
BiConjugate Gradient(BiCG) Method, 436
Burgers' equation, 179

C
Cahn–Hilliard equation, 206, 227, 367, 381
Chebyshev-Galerkin Method, 148, 312
Chebyshev-Gauss points, 107
Chebyshev-Gauss-Lobatto points, 107
Chebyshev-Gauss-Radau points, 107
Chebyshev-Legendre dual-Petrov-Galerkin method, 217
Chebyshev-Legendre Galerkin Method, 150
Chebyshev-Legendre transform, 151, 217
Christoff-Darboux formula, 51
Circular Domain, 307
collocation method in the strong form, 154, 163
collocation method in the weak form, 153
compact combination of Legendre polynomials, 145, 203
compact combination of orthogonal polynomials, 7, 9
Conjugate Gradient (CG) Method, 433
Conjugate Gradient (CG) method, 379
Consistent Splitting Scheme, 394
Continuous Fourier Series, 24
Convex Splitting, 384
Crank-Nicolson leap-frog scheme, 2, 228, 229
curse of dimensionality, 346
Cylindrical Domain, 307

D
Delay Differential Equation, 197
Differentiation in the Frequency Space, 31
  (Generalized) Laguerre Polynomial/Function, 252
  Chebyshev Polynomial, 111
  Fourier Series, 31
  Hermite Polynomial/Function, 262
  Jacobi Polynomial, 92
  Legendre Polynomial, 105
  Orthogonal Polynomial, 66
Differentiation in the Physical Space, 29
  (Generalized) Laguerre Polynomial/Function, 251
  Chebyshev Polynomial, 110
  Fourier Series, 29
  Hermite Polynomial/Function, 261
  Jacobi Polynomial, 88
  Legendre Polynomial, 103
  Orthogonal Polynomial, 64
differentiation matrix, 5, 30, 88, 110, 220, 251, 261, 290
  Chebyshev-Gauss, 111
  Chebyshev-Gauss-Lobatto, 110
  Chebyshev-Gauss-Radau, 110
  Fourier method, 30
  Hermite-Gauss, 261
  Jacobi-Gauss, 91
  Jacobi-Gauss-Lobatto, 88
  Jacobi-Gauss-Radau, 90
  Laguerre-Gauss, 251

J. Shen et al., *Spectral Methods: Algorithms, Analysis and Applications*,
Springer Series in Computational Mathematics 41, DOI 10.1007/978-3-540-71041-7,

Laguerre-Gauss-Radau, 251
Legedre-Gauss, 105
Legedre-Gauss-Lobatto, 103
Legedre-Gauss-Radau, 104
modified Hermite-Gauss, 261
modified Laguerre-Gauss, 252
modified Laguerre-Gauss-Radau, 252
Dimension Reduction, 307
Dirichlet kernel, 25
Dirichlet-to-Neumann (DtN) Map, 369
Discrete Fourier Series, 25
Discrete Transform
(Generalized) Laguerre Polynomial/Function, 249
Chebyshev Polynomial, 108
Hermite Polynomial/Function, 260
Jacobi Polynomial, 86
Legendre Polynomial, 100
discrete transform, 1, 7, 68, 147, 150, 151, 254
Chebyshev polynomial, 109
Fourier method, 27, 28
Distribution and Weak Derivative, 421
distributions, 122
Dual-Petrov-Galerkin Method, 210
duality argument, 121, 167, 335

E
embedding inequalities, 426
essential pole conditions, 308
Even-Order Equation, 208
exponential convergence, 19
exponential decay, 276

F
Fast Fourier transform (FFT), 23
fifth-order equation, 219
Fifth-Order KdV Type Equations, 232
Finite Difference Preconditioning, 163
Finite Element Preconditioning, 164
Fisher equation, 180
Fourier Approximation, 33
Interpolation Errors, 35
Inverse Inequalities, 33
Orthogonal Projection Errors, 34
Fourier collocation points, 26
Fourier-Chebyshev Galerkin Method, 315
fourth-0rder equation, 305
Fourth-Order Equation, 206

G
Gagliardo-Nirenberg inequality, 167
Galerkin Method, 206
Galerkin Method with Numerical Integration, 9, 152
Galerkin method with numerical integration, 216
Galerkin Reformulation, 156
Gamma function, 71, 415
Gauss-Type Quadrature, 57
Gauss-type quadrature
Gauss quadrature, 57
Gauss-Lobatto quadrature, 62
Gauss-Radau quadrature, 60
Gegenbauer polynomials, 73
general boundary conditions, 141
generalized Gauss-Lobatto quadrature, 218
Generalized Jacobi Polynomials, 201, 356
generalized Laguerre functions, 241
generalized Laguerre polynomials, 238
Generalized Minimal Residual (GMRES) method, 441
Ginzburg-Landau double-well potential, 383
Gronwall Inequality, 430
Gross-Pitaevskii Equation, 403
Gross-Pitaevskii equation, 367

H
Hankel function, 370
Hardy Inequality, 428
Helmholtz Equation, 174
Helmholtz equation, 367
Hermite functions, 256
Hermite polynomials, 254
Hermite-Galerkin Method, 275
Hermite-Gauss points, 257
Hermite-Gauss Quadrature, 257
Computation of Nodes and Weights, 258
Hierarchical Basis, 358
High-Dimensional Problem, 346
high-order boundary value problems, 201
Hilbert space, 418
Hyperbolic Cross, 346

I
inf-sup condition, 377
inf-sup constant, 377
Interpolation and Discrete Transforms, 63
irrational approximation, 296

J
Jacobi approximation, 113
interpolation errors, 129
inverse inequalities, 113
orthogonal projection errors, 116
Jacobi Polynomials, 70
Jacobi polynomials
maximum value, 78
recurrence formulas, 74

Rodrigues' formula, 72
Sturm-Liouville equation, 70
Jacobi weight function, 70, 117, 156, 281
Jacobi-collocation method, 289
Jacobi-Galerkin method, 289
Jacobi-Gauss points, 80
Jacobi-Gauss-Lobatto points, 83, 156
Jacobi-Gauss-Radau points, 81
Jacobi-Gauss-Type Quadrature, 80
Computation of Nodes and Weights, 83
Jacobi-Gauss-type quadrature
Jacobi-Gauss quadrature, 80
Jacobi-Gauss-Lobatto quadrature, 83
Jacobi-Gauss-Radau quadrature, 81
Jacobi-weighted Korobov-type space, 347
Jacobi-weighted Sobolev space, 117, 120, 283, 330

K
Korteweg–de Vries (KdV) Equation, 229
Korteweg–de Vries (KdV) equation, 37, 227
Krylov Subspace Method, 433
Kuramoto–Sivashinsky (KS) equation, 38, 206

L
Lagrange basis polynomial, 3, 4, 57, 58, 61, 66
Laguerre and Hermite Approximation
Interpolation Errors, 271
Inverse Inequalities, 263
Orthogonal Projection Errors, 265
Laguerre and Hermite approximation, 263
Laguerre Polynomials/Functions, 238
Laguerre-Galerkin Method, 273
Laguerre-Gauss points, 243
Laguerre-Gauss-Radau points, 243
Laguerre-Gauss-Type Quadratures, 243
Computation of Nodes and Weights, 247
Laguerre-Gauss-type quadratures
Laguerre-Gauss, 243
Laguerre-Gauss-Radau, 243
modified Laguerre-Gauss, 246
modified Laguerre-Gauss-Radau, 246
Laplace-Beltrami operator, 324
Lax-Milgram Lemma, 419
Lax-Milgram lemma, 10, 128, 142, 207, 215
Lebesgue constant, 184
Legendre dual-Petrov-Galerkin method, 211, 213, 216, 229
Legendre Polynomials, 93
Legendre-Galerkin Method, 145, 311
Legendre-Galerkin method, 207, 228
Legendre-Gauss points, 95
Legendre-Gauss-Lobatto points, 95
Legendre-Gauss-Radau points, 95
Legendre-Gauss-Type Quadratures, 95
Computation of Nodes and Weights, 98
Lyapunov energy functional, 381

M
mapped Jacobi approximation
interpolation errors, 286
orthogonal projection errors, 282
Mapped Spectral Method, 279
General Mapping, 279
Mapped Jacobi Polynomials, 281
mapped spectral method
algebraic mapping, 281
exponential mapping, 281
logarithmic mapping, 281
Mathieu function, 299
Matrix Diagonalization, 301
matrix diagonalization
full diagonalization, 302
partial diagonalization, 301
Modal Basis, 158
modal basis, 7, 145, 161
modified Legendre-rational function, 295
Multi-Step Method, 450
Multivariate Jacobi Approximation, 328

N
natural pole conditions, 308
Navier–Stokes Equations, 392
Navier–Stokes equations, 367
nested grids, 358
new weighted weak formulation, 144
Nikolski's inequality, 33
Nodal Basis, 162
nodal basis, 4, 152

O
Odd-Order Equation, 210
Operator Splitting Method, 453
Optimized Hyperbolic Cross, 352
Orthogonal Polynomials, 47
Computation of Zeros, 55
Properties of Zeros, 53
orthogonal polynomials
computation of zeros
eigenvalue method, 55
iterative method, 56
orthogonal projection, 13

P
Petrov-Galerkin Method, 8
Petrov-Galerkin Reformulation, 157
Poincaré inequality, 423, 424
Poincaré inequality, 160, 167, 206

Pole Conditions, 307
Preconditioned Iterative Method, 157
Preconditioning, 158, 162, 443
prolate spheroidal wave function, 299
pseudo-spectral method, 3
pseudo-spectral technique, 68

R
Rational Approximation, 279
Rectangular Domain, 300
regularity index, 119
Riesz representation theorem, 121
Rodrigues' formula, 239, 255
Rotational Pressure-Correction Scheme, 392
Runge–Kutta Method, 448
Runge–Kutta method, 39

S
scaling, 278, 280
self-adjoint, 71, 158
semi-implicit scheme, 39
separable multi-dimensional domain, 299
Separation of Variables, 303
separation of variables, 369
separation property, 53
Smolyak's construction, 360
Smolyak's sparse grid, 357
Sobolev inequality, 161, 169, 178, 426
Sobolev Space, 422
soliton, 38, 229, 233
Sommerfeld radiation condition, 369
Sparse Spectral-Galerkin Method, 357, 361
sparsity, 346
spectral accuracy, 18
spectral differentiation, 1, 68, 254
Spectral-Collocation Method, 4
Spectral-Galerkin Method, 6
spectral-Galerkin method, 305
Spherical Domain, 323
spherical Hankel function, 369
spherical harmonic function, 324, 369
Spherical Shell, 325
stability of interpolation, 129, 131, 134, 136
stabilized semi-implicit scheme, 44
Stirling's formula, 119, 239, 416
Stokes equations, 367
Stokes pair, 378
Sturm-Liouville equation, 239, 242, 255

T
tensor product, 300
Third-Order Equation, 210
three-term recurrence formula, 49
three-term recurrence relation, 242, 254, 256
Time-Harmonic Wave Equation, 368
transformed field expansion, 368
two-point boundary value problem, 141

U
unbounded domain, 237

V
Variable Coefficient Problem, 216
variable coefficient problem, 157
Volterra integral equation, 181

W
weakly singular kernel, 181
Weierstrass theorem, 68
Weighted Galerkin Formulation, 143
weighted Galerkin formulation, 211, 300
weighted residual methods (WRMs), 1
weighted Sobolev space, 265