AN ALTERNATING-DIRECTION SINC-GALERKIN METHOD FOR ELLIPTIC PROBLEMS ON FINITE AND INFINITE DOMAINS

by

Nicomedes Alonso III

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Mathematics

MONTANA STATE UNIVERSITY
Bozeman, Montana

April 2009

APPROVAL

of a dissertation submitted by

Nicomedes Alonso III

This dissertation has been read by each member of the dissertation committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the MSU Division of Graduate Education.

Dr. Kenneth L. Bowers

Approved for the Department of Mathematical Sciences

Dr. Kenneth L. Bowers

Approved for the Division of Graduate Education

Dr. Carl Fox

STATEMENT OF PERMISSION TO USE

Nicomedes Alonso III

April 2009

ACKNOWLEDGEMENTS

With profound gratitude, I humbly dedicate this dissertation in remembrance of my mother, Candelaria de la Caridad Benitez Abascal de Alonso, who was “Mima” to all of us; to the memory of my father, Nicomedes Alonso Acosta, who sacrificed so much so his children could be free; to my soul mate, Giuseppina Yvonna Audisio, and our children, Orion, Altair, Indi, and Lyra, whose love gives my life meaning; to the memory of our son, Rigel, and to my advisor, Professor Kenneth L. Bowers, for his time, his generous support and his kind guidance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES – CONTINUED

LIST OF FIGURES

ABSTRACT

Alternating-Direction Implicit (ADI) schemes are a class of very efficient algorithms for the numerical solution of differential equations. Sinc-Galerkin schemes employ a sinc basis to produce exponentially accurate approximate solutions to differential equations even in the presence of singularities. In this dissertation we begin with a broad overview of sinc methods for problems posed on both finite and infinite, one- and two-dimensional domains. We then present a variety of finite difference methods that lead to the introduction of a new Alternating-Direction Sinc-Galerkin scheme based on the classic ADI scheme for a linear matrix system. We note that when a Sinc-Galerkin method is used to solve a Poisson equation, the resulting matrix system is a Sylvester equation. We discuss ADI *model problems* in general and then prove that when a *symmetric* Sinc-Galerkin method is employed, the resulting Sylvester equation can be classified as an ADI model problem. Finally, we derive our Alternating-Direction Sinc-Galerkin (ADSG) method to solve this resulting Sylvester equation, specifying the use of a constant iteration parameter to avoid costly eigenvalue computations. We end by applying ADSG to a variety of problems, comparing its performance to the standard technique that uses the Kronecker product, the Kronecker sum, and the concatenation operator.

CHAPTER 1

INTRODUCTION

The subject of this thesis is the introduction of an Alternating-Direction Sinc-Galerkin (ADSG) method and its application to the numerical solution of elliptic partial differential equation boundary-value problems posed on two-dimensional finite domains and on two-dimensional infinite domains.

Sinc methods were introduced by Frank Stenger in [23] and expanded upon by him in [24]. Sinc functions were first analyzed in [29] and [30]. An extensive treatment of sinc methods for two-point boundary-value problems is given in [9] and [11], the latter of which also discusses elliptic and parabolic partial differential equations. Parabolic and hyperbolic problems are discussed in detail in [7] and [14], respectively. A variety of singular elliptic problems are treated in [2], and the symmetric Sinc-Galerkin method in this context is introduced in [10]. Sinc domain decomposition is presented in depth in [12], [13], [16], and [17]. Iterative methods for symmetric Sinc-Galerkin systems are discussed in [1], [18] and [19]. In [19] we also find an alternating-direction Sinc-Galerkin method that is quite different from the one presented here. Sinc methods are discussed thoroughly in [26]. Applications of sinc methods may also be found in [6] and [15]. A thorough review up to 1981 is found in [25]. The works cited above generally involve applications on finite domains. Although sinc methods for problems posed on the nonnegative real line are presented in [11], very little work has been done to extend these methods to more general infinite domains as is done in Chapter 2 and Chapter 4 below.

According to Stenger (see [26, p. vi]), sinc methods are based on the *sine cardinal* (or sinc) function which dates back to the works of Bohr, de la Vall*é*e Poissin and E. T. Whittaker (see [29], for example). Symmetrization of the Sinc-Galerkin method for two-point boundary-value problems was introduced by John Lund in [9]. The numerical solution of elliptic problems via the Sinc-Galerkin method gives rise to an algebraic system of linear equations whose matrix representation may be given by the Sylvester equation

$$AX + XB = C \tag{1.1}$$

or its equivalent Kronecker sum formulation,

$$(I \otimes A + B^T \otimes I)co(X) = co(C).$$

The ADSG method presented in this thesis is a direct application of the Alternating-Direction Implicit (ADI) iteration equations (introduced by D. Peaceman and H. Rachford in [20]) to the Sylvester equation that results when a symmetric Sinc-Galerkin method is used to solve an elliptic boundary-value problem. This ADSG method is an iterative scheme that may be used to solve (1.1) and which avoids both the large matrix systems produced by using the Kronecker product and Kronecker sum, as well as the computation of eigenvalues and eigenvectors characteristic of the standard matrix diagonalization procedure for solving (1.1).

Chapter 2 presents the basic definitions and theorems on interpolation and quadrature that form the foundation for the discussion of sinc methods in general and the symmetric Sinc-Galerkin method for elliptic partial differential equations in particular. To enhance the clarity of the presentation, sinc methods are first used to illustrate the solution of a variety of one-dimensional problems on the unit interval. The Sinc-Galerkin solution of boundary-value problems on the nonnegative real line is then introduced. Little work has been done on the Sinc-Galerkin method for this type of problem, so care is taken to develop this method thoroughly. These problems illustrate the exponential convergence rate such methods exhibit even when applied to problems that have singular solutions. We then proceed with a thorough discussion of elliptic problems and how to solve the resulting matrix systems using a diagonalization procedure and a procedure involving the Kronecker product and Kronecker sum. After presenting several examples of the Sinc-Galerkin method applied to problems posed on the unit square, we conclude Chapter 2 with the introduction of the Sinc-Galerkin solution of elliptic partial differential equations posed on the first quadrant, an approach which is again relatively new.

Chapter 3 outlines the historical development of the ADI method by considering finite difference schemes of ever-increasing complexity, applied to a parabolic problem on the unit square. The resulting ADI scheme is then adapted to solve elliptic problems, introducing iteration parameters and culminating with the numerical solution of the specific elliptic problem that motivated the use of a constant iteration parameter for the ADSG method.

Chapter 4 begins with a discussion of Galerkin methods in general and continues with an exposition from [8] on the application of ADI iteration to the solution of the matrix equation $M\mathbf{y} = \mathbf{b}$. The subsequent direct application of ADI iteration to the solution of the Sylvester equation introduces the ADSG method. Once the ADSG method is derived in detail, we choose a constant iteration parameter and select appropriate stopping criteria for the examples to be presented. We end the chapter by employing the ADSG method to solve several elliptic problems on both the unit square and on the first quadrant. We note that the ADSG method exhibits rapid convergence with exponential accuracy and as such is a practical alternative to the comparable methods used in the previous chapters.

CHAPTER 2

THE SINC-GALERKIN METHOD

Introduction

Since their introduction in [24], sinc methods have been used to solve a variety of differential equations ([2], [6], [9], [11], [14], [15], [25], [26]). Their exponential convergence rate have made such methods excellent tools for accurately approximating the solution to partial differential equation boundary-value problems that are posed on infinite domains or whose solutions are singular on all or part of the boundary of a finite or infinite domain.

Sinc Basis Functions

Let $\mathbb{C}$ denote the set of all complex numbers and for all $z \in \mathbb{C}$ define the *sine cardinal* or *sinc* function by

$$\operatorname{sinc}(z) \equiv \begin{cases} \dfrac{\sin(\pi z)}{\pi z}, & z \neq 0 \\ 1, & z = 0. \end{cases} \tag{2.1}$$

For $h > 0$ and any integer k, the translated sinc function with evenly spaced nodes at $\{kh\}_{k=-\infty}^{\infty}$ is denoted $S(k,h)(z)$ and defined by

$$S(k,h)(z) = \operatorname{sinc}\left(\frac{z-kh}{h}\right). \tag{2.2}$$

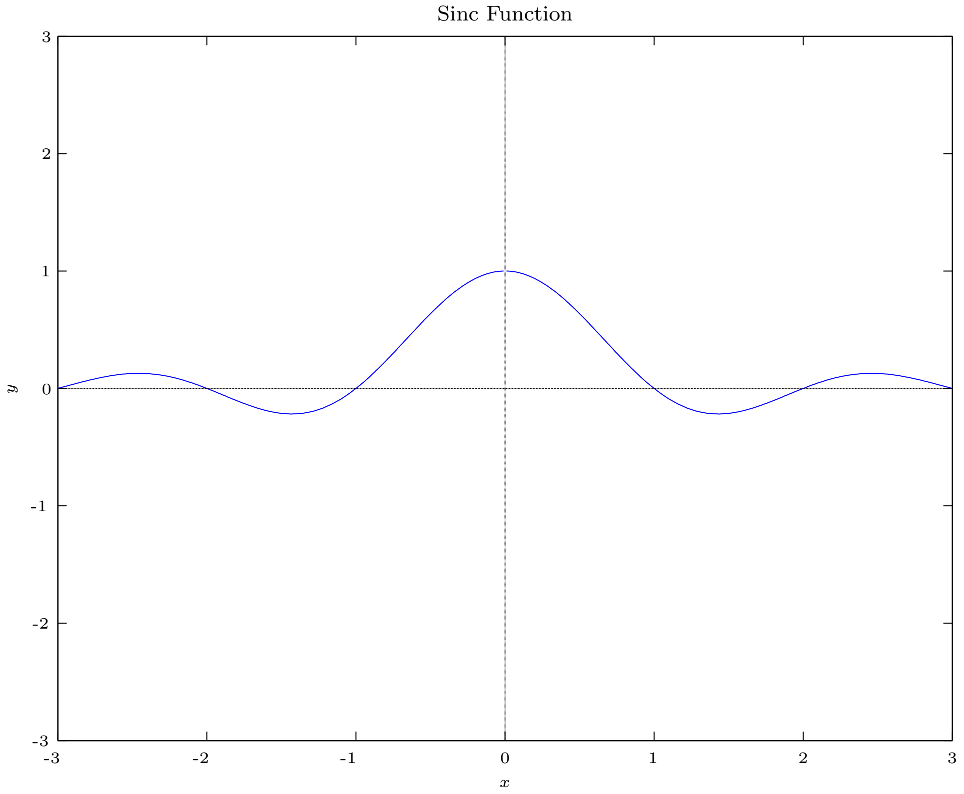


Figure 1. The sinc function for real values of x on the interval $[-3, 3]$.

A sketch of the sinc function for real values of x on the interval $[-3, 3]$ is shown in Figure 1. A sketch of sinc translates, $S(k, \pi/4)(x)$ for $k = -1$, 0, 1 and for real values of x on the interval $[-3, 3]$ is shown in Figure 2. Note these translates are centered at $-h$, 0, and h, respectively.

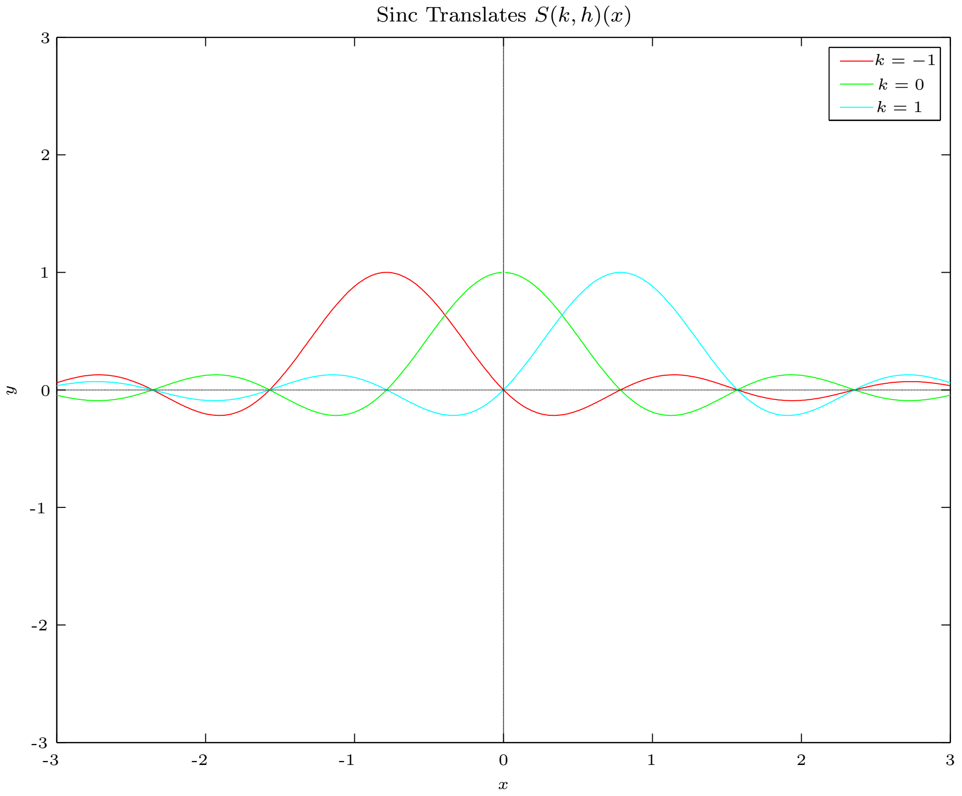


Figure 2. The translated sinc functions ($k = -1$, 0, 1) for real values of x on the interval $[-3, 3]$.

Exact Interpolation and Quadrature

If $\mathbb{R}$ represents the set of all real numbers and f is an arbitrary function defined on $\mathbb{R}$ then the *Whittaker cardinal expansion* of f is denoted $C(f,h)(x)$ and is defined by

$$C(f,h)(x) = \sum_{k=-\infty}^{\infty} f(kh)S(k,h)(x) \tag{2.3}$$

whenever the series converges. The Whittaker cardinal function interpolates f at the points $\{kh\}_{k=-\infty}^{\infty}$. The series was originally discussed in [29] and later thoroughly analyzed in [30]. A class of functions in which the cardinal function of f converges to f is the *Paley-Wiener* class $B(h)$ defined below. (Recall that a complex-valued function is called *entire* if it is analytic at all finite points in the complex plane.)

Definition 1. *For any* $h > 0$, *the Paley-Wiener class of functions* $B(h)$ *is the family of entire functions* f *which when restricted to the real line satisfy* $f \in L^2(\mathbb{R})$ *and on the complex plane are of exponential type* π/h, *i.e.,* f *satisfies*

$$|f(z)| \leq K e^{\pi|z|/h} \tag{2.4}$$

for some $K > 0$ *and for all* $z \in \mathbb{C}$.

The remarkable fact that a function that interpolates f at a countable number of points on the real line can be exact for all $z \in \mathbb{C}$ is expressed in the theorem below, which is proven in [11, p. 24].

Theorem 1. *If* $f \in B(h)$ *then for all* $z \in \mathbb{C}$ *we have*

$$f(z) = C(f,h)(z) = \sum_{k=-\infty}^{\infty} f(kh) S(k,h)(z). \tag{2.5}$$

Furthermore, if $f \in L^1(\mathbb{R})$, *then we have the trapezoidal rule*

$$\int_{-\infty}^{\infty} f(t)\, dt = h \sum_{k=-\infty}^{\infty} f(kh). \tag{2.6}$$

Approximate Interpolation and Quadrature

In an attempt to find a less restrictive class of functions than $B(h)$ for which the formulas of Theorem 1, while no longer exact, remain exponentially accurate, we begin with

Definition 2. *In the complex plane* $\mathbb{C}$, *let* D_S *denote the infinite strip domain of width* $2d, \quad d > 0, \quad$ *given by*

$$D_S \equiv \{w \in \mathbb{C} : w = u + iv,\ |v| < d\}. \tag{2.7}$$

Let $B^p(D_S)$ *be the set of functions analytic in* D_S *that satisfy*

$$\int_{-d}^{d} |f(t+iv)|dv = \mathcal{O}(|t|^a), \quad t \to \pm\infty, \quad 0 \le a < 1$$

and

$$N^p(f, D_S) \equiv \lim_{v \to d^-} \left\{ \left(\int_{-\infty}^{\infty} |f(t+iv)|^p dt \right)^{1/p} + \left(\int_{-\infty}^{\infty} |f(t-iv)|^p dt \right)^{1/p} \right\} < \infty.$$

For $p = 1$ *let* $N(f, D_S) \equiv N^1(f, D_S)$ *and* $B(D_S) \equiv B^1(D_S)$.

The following theorem which is proven in [11, pp. 35, 48], shows that sinc interpolation and quadrature methods are exponentially accurate for functions in $B^p(D_S)$.

Theorem 2. *If* $f \in B^p(D_S)$, $p = 1$ *or* 2, *then there exists a positive constant* K_1 *such that on the real line* $\mathbb{R}$

$$\left\| f - \sum_{k=-\infty}^{\infty} f(kh) S(k,h) \right\|_{\infty} \le K_1 e^{-\pi d/h}. \tag{2.8}$$

Additionally, if $p = 1$ *then there exists a positive constant* K_2 *such that*

$$\left| \int_{-\infty}^{\infty} f(t)\, dt - h \sum_{k=-\infty}^{\infty} f(kh) \right| \le K_2 e^{-2\pi d/h}. \tag{2.9}$$

In practice, however, only finite sums can be calculated. Below, we define the *truncated cardinal series* and the *truncated trapezoidal rule* and then minimize the effect of truncation on convergence by assuming appropriate growth conditions on f.

The truncated cardinal series is denoted $C_{M,N}(f,h)(z)$ and defined by

$$C_{M,N}(f,h)(z) \equiv \sum_{k=-M}^{N} f(kh)S(k,h)(z). \tag{2.10}$$

The truncated trapezoidal rule is denoted $T_{M,N}(f,h)$ and defined by

$$T_{M,N}(f,h) \equiv h \sum_{k=-M}^{N} f(kh). \tag{2.11}$$

The quadrature error for functions in $B(D_S)$ is given in the following theorem whose proof may be found in [11, p. 49]

Theorem 3. *Suppose $f \in B(D_S)$ and there are positive constants α, β, and C so that*

$$|f(x)| \le C \left\{ \begin{array}{ll} e^{-\alpha|x|}, & x \in (-\infty, 0) \\ e^{-\beta|x|}, & x \in [0, \infty). \end{array} \right.$$

Choose the positive integer M arbitrarily, let

$$N = \left[\left| \frac{\alpha}{\beta}M + 1 \right|\right] \tag{2.12}$$

(where $[|\cdot|]$ denotes the greatest integer function) and let

$$h = \sqrt{\frac{2\pi d}{\alpha M}} \le \frac{2\pi d}{\ln(2)} \tag{2.13}$$

for the truncated trapezoidal rule $T_{M,N}(f,h)$ given in (2.11)*. Then*

$$\left| \int_{-\infty}^{\infty} f(t)\,dt - h \sum_{k=-M}^{N} f(kh) \right| \le K_3 e^{-\sqrt{2\pi d\alpha M}} \tag{2.14}$$

where K_3 is a constant depending on f.

For problems on a subinterval, Γ, of the real line we employ a conformal map ϕ for which $\phi(\Gamma) = \mathbb{R}$. Suppose $d > 0$ and let ϕ be a one-to-one conformal map of D onto D_S where

$$D = \left\{ z = x + iy : \Big| \arg\left(\phi(z)\right) \Big| < d \leq \frac{\pi}{2} \right\}. \tag{2.15}$$

Then over a contour $\Gamma = \phi^{-1}(\mathbb{R})$ we have the following methods of interpolation

$$f(z) \approx C(f,h) \circ \phi(z) \equiv \sum_{k=-\infty}^{\infty} f(z_k)(S(k,h) \circ \phi)(z) \tag{2.16}$$

and quadrature

$$\int_{\Gamma} f(x)\, dx \approx h \sum_{k=-\infty}^{\infty} f(z_k)/\phi'(z_k) \tag{2.17}$$

where $z_k = \phi^{-1}(kh)$.

<u>Examples of Domains and Maps</u>

Example 1. *If* $\phi(z) = \ln\left(\dfrac{z}{1-z}\right)$ *then* D *is the eye-shaped domain denoted* D_E*, defined by*

$$D_E = \left\{ z \in \mathbb{C} : \left| \arg\left(\frac{z}{1-z}\right) \right| < d \leq \pi/2 \right\}$$

and shown in Figure 3. *The contour* Γ *is the interval* $(0,1)$ *and* $z_k = \phi^{-1}(kh) = \dfrac{e^{kh}}{1+e^{kh}}$. (See [11, pp. 63-64].) *Note that when* $d = \pi/2$, D_E *becomes the circle* $|z - \frac{1}{2}| < \frac{1}{2}$.

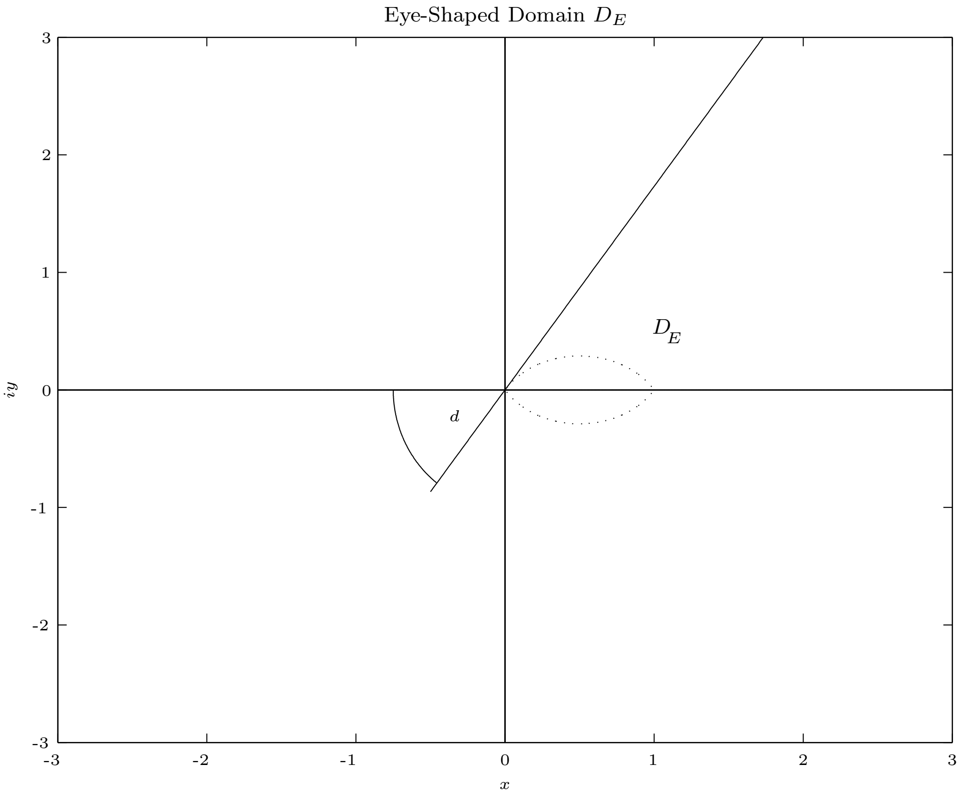


Figure 3. The domain D_E for $d = \pi/3$ and $\Gamma = (0, 1)$.

Example 2. *If* $\phi(z) = \ln(z)$ *then* D *is the wedge-shaped domain denoted* D_W*, defined by*

$$D_W = \left\{ z \in \mathbb{C} : z = re^{i\theta}, |\theta| < \ d \leq \pi/2 \right\},$$

and shown in Figure 4*. The contour* Γ *is the interval* $(0, \infty)$*, and* $z_k = \phi^{-1}(kh) = e^{kh}$*.* (See [11, pp. 71-72].) *Note that when* $d = \pi/2$*,* D_W *becomes the right half-plane.*

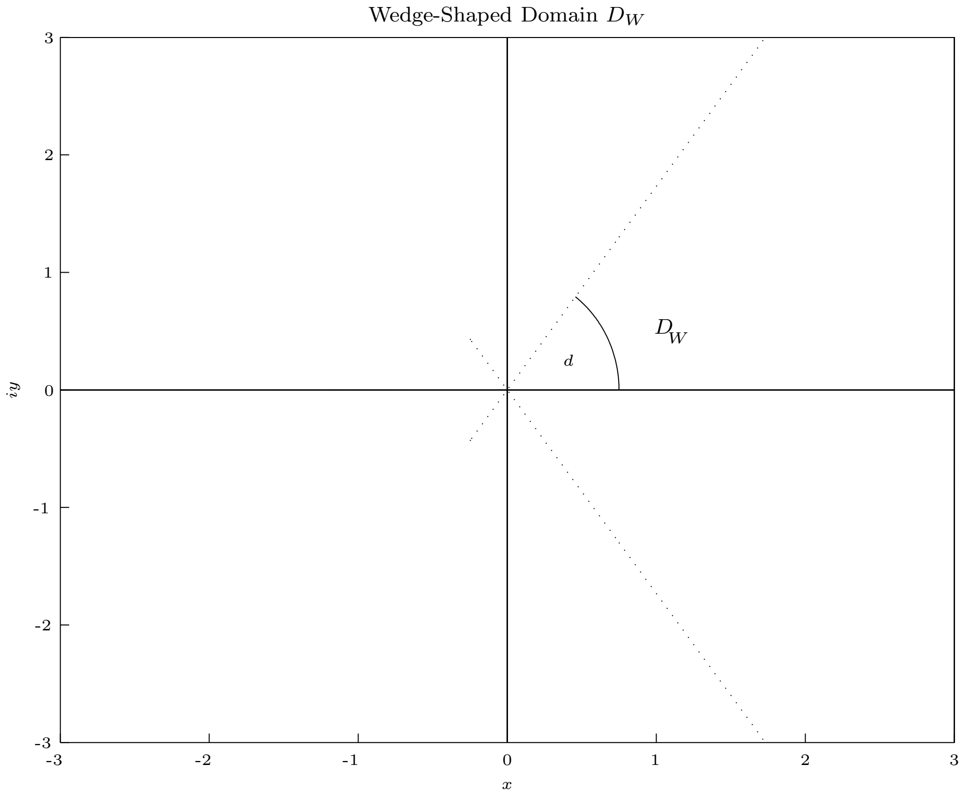


Figure 4. The domain D_W for $d = \pi/3$ and $\Gamma = (0, \infty)$.

Example 3. *If* $\phi(z) = \ln(\sinh(z))$ *then* D *is the domain denoted* D_B*, defined by*

$$D_B = \{z \in \mathbb{C} : |\arg(\sinh(z))| < \ d \le \pi/2\} \, ,$$

and shown in Figure 5*. The contour* Γ *is the interval* $(0, \infty)$*, and* $z_k = \phi^{-1}(kh) = \ln(e^{kh} + \sqrt{e^{2kh} + 1})$. (See [11, pp. 80-81].) *Note that when* $d = \pi/2$*,* D_B *becomes the right half-plane.*

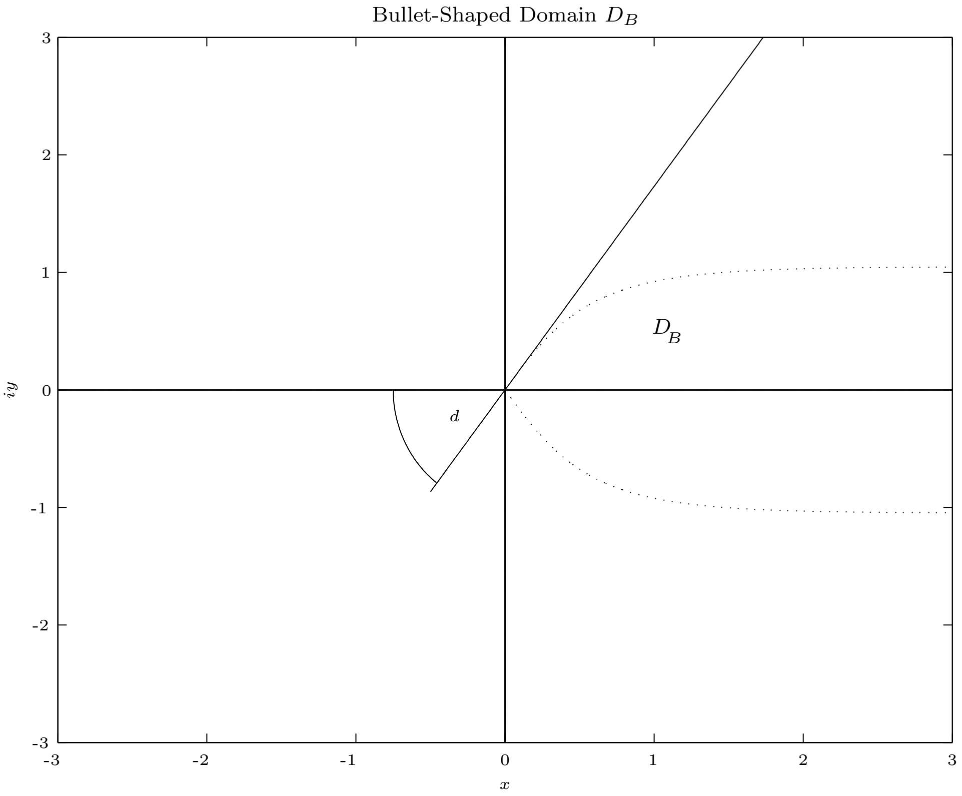


Figure 5. The domain D_B for $d = \pi/3$ and $\Gamma = (0, \infty)$.

A short list of conformal mappings may be found in Table 1 below, [11].

To develop sinc methods on arcs Γ in the complex plane or on the real line we begin with the following.

Definition 3. *Let D be a domain in the complex plane with boundary points $a \neq b$. Suppose ϕ is a one-to-one conformal map of D onto the infinite strip D_S satisfying $\phi(a) = -\infty$ and $\phi(b) = \infty$. Let ψ denote the inverse of the mapping ϕ and let $\Gamma = \psi(\mathbb{R})$. We define the function space $B(D)$ to be the class of functions F analytic*

Table 1. Conformal mappings and nodes for several subintervals of $\mathbb{R}$.

a	b	$\phi(z)$	z_k
a	b	$\ln\left(\frac{z-a}{b-z}\right)$	$\frac{a+be^{kh}}{1+e^{kh}}$
0	1	$\ln\left(\frac{z}{1-z}\right)$	$\frac{e^{kh}}{1+e^{kh}}$
0	∞	$\ln(z)$	e^{kh}
0	∞	$\ln(\sinh(z))$	$\ln(e^{kh}+\sqrt{e^{2kh}+1})$
$-\infty$	∞	z	kh
$-\infty$	∞	$\sinh^{-1}(z)$	$\sinh(kh)$

in D *which satisfy for some constant* a *with* $0 \leq a < 1$,

$$\int_{\psi(x+L)} |F(z)dz| = \mathcal{O}(|x|^a), \quad x \to \pm\infty \tag{2.18}$$

where $L = \{iy : |y| < d\}$ *and for* γ *a simple closed contour in* D

$$N(F, D) \equiv \lim_{\gamma \to \partial D} \int_\gamma |F(z)dz| < \infty. \tag{2.19}$$

Further, for $h > 0$, *define the nodes*

$$z_k = \psi(kh) = \phi^{-1}(kh), \quad k = 0, \pm 1, \pm 2, \ldots. \tag{2.20}$$

The next theorem which is proven in [11, p. 70] shows that the exponential convergence rate of the truncated trapezoidal rule is preserved under conformal mappings.

Theorem 4. *Let $F \in B(D)$ and let $h > 0$. Suppose ϕ is a one-to-one conformal mapping of the domain D onto D_S. Let $\psi = \phi^{-1}$, $z_k = \psi(kh)$ and $\Gamma = \psi(\mathrm{I\!R})$. Assume further that there are positive constants α, β, and C such that*

$$\left|\frac{F(z)}{\phi'(z)}\right| \leq C \left\{ \begin{array}{ll} e^{-\alpha|\phi(z)|}, & z \in \Gamma_a \\ e^{-\beta|\phi(z)|}, & z \in \Gamma_b \end{array} \right. \tag{2.21}$$

where

$$\Gamma_a \equiv \{z \in \Gamma : \phi(z) \in (-\infty, 0)\} \tag{2.22}$$

and

$$\Gamma_b \equiv \{z \in \Gamma : \phi(z) \in [0, \infty)\}. \tag{2.23}$$

Choose the positive integer M arbitrarily, let

$$N = \left[\left|\frac{\alpha}{\beta}M + 1\right|\right] \tag{2.24}$$

and let

$$h = \sqrt{\frac{2\pi d}{\alpha M}} \leq \frac{2\pi d}{\ln(2)} \tag{2.25}$$

for the truncated (mapped) trapezoidal rule

$$T_{M,N}(F, h, \phi) \equiv h \sum_{k=-M}^{N} \frac{F(z_k)}{\phi'(z_k)}. \tag{2.26}$$

Then

$$\left| \int_\Gamma F(z)dz - T_{M,N}(F,h,\phi) \right| \leq K_4 e^{-\sqrt{2\pi d\alpha M}} \tag{2.27}$$

where K_4 is a constant depending on F, d, ϕ, and D.

The Sinc-Galerkin Method

In this section we develop the Sinc-Galerkin method for the two-point boundary-value problem

$$\begin{aligned} \mathcal{L}u(x) \equiv -u''(x) + p(x)u'(x) \quad &+ \quad q(x)u(x) = f(x), \quad a < x < b \\ u(a) \quad &= \quad u(b) = 0. \end{aligned} \tag{2.28}$$

Sinc methods for these problems are thoroughly discussed in [9], [11], [24], and [25]. While the development below is given for a finite interval (a,b), the modifications necessary for problems posed on an infinite interval are minimal and involve little more than an appropriate selection of the conformal map, $\phi(z)$.

Begin by choosing $\phi(z)$. Let $\psi(z) = \phi^{-1}(z)$ and suppose $h > 0$ and the positive integers M and N are given. For $j = -M, ..., N$ define the sinc basis functions

$$S_j(x) \equiv S(j,h) \circ \phi(x) = \text{sinc}\left[\frac{\phi(x) - jh}{h}\right]. \tag{2.29}$$

The approximate solution to (2.28) is given by

$$u_m(x) = \sum_{k=-M}^{N} u_k S_k(x), \quad m = M + N + 1 \tag{2.30}$$

where the coefficients u_k are to be determined. The Sinc-Galerkin method requires orthogonalizing the residual $\mathcal{L}u_m - f$ against each sinc basis function, S_j, using a weighted inner product. If we define the inner product of f and g, denoted (f, g), as follows

$$(f, g) = \int_a^b f(x)g(x)w(x)\,dx, \tag{2.31}$$

where $w(x)$ is a weight function, then for $j = -M, ..., N$ we require that

$$(\mathcal{L}u_m - f, S_j) = 0.$$

To simplify the notation we orthogonalize $\mathcal{L}u - f$ against each basis function and obtain for $j = -M, ..., N$,

$$\int_a^b (-u''(x) + p(x)u'(x) + q(x)u(x) - f(x))S_j(x)w(x)\,dx = 0.$$

Integrating by parts to remove all derivatives of u gives

$$\begin{aligned} \int_a^b f(x)S_j(x)w(x)\,dx &= -\int_a^b u(x)(S_jw)''(x)\,dx \qquad (2.32)\\ -\int_a^b u(x)(pS_jw)'(x)\,dx &+ \int_a^b u(x)q(x)S_j(x)w(x)\,dx + BT \end{aligned}$$

where

$$BT = (upS_jw)(x)\Big|_a^b - (u'S_jw)(x)\Big|_a^b + (u(S_jw)')(x)\Big|_a^b. \tag{2.33}$$

Note that the weight function w may be chosen so that the boundary term BT vanishes. To record the approximations to the integrals above it is convenient to introduce the notation $\delta_{jk}^{(p)}$ where

$$\delta_{jk}^{(p)} \equiv h^p \frac{d^p}{d\phi^p}[S(j,h)\circ\phi(x)]\bigg|_{x=x_k}, \quad p=0,1,2,... \tag{2.34}$$

and $x_k = \phi^{-1}(kh) = \psi(kh)$. These quantities are given explicitly in the following theorem which is proven in [11, p. 106].

Theorem 5. *Let ϕ be a conformal one-to-one map of the simply connected domain D onto D_S. Then*

$$\delta_{jk}^{(0)} = \begin{cases} 1, & j=k \\ 0, & j\neq k \end{cases}$$

$$\delta_{jk}^{(1)} = \begin{cases} 0, & j=k \\ \frac{(-1)^{k-j}}{k-j}, & j\neq k \end{cases}$$

$$\delta_{jk}^{(2)} = \begin{cases} \frac{-\pi^2}{3}, & j=k \\ \frac{-2(-1)^{k-j}}{(k-j)^2}, & j\neq k. \end{cases}$$

The following theorem whose proof may be found in [11, pp. 108, 110], gives the approximations to the inner products necessary for forming the Sinc-Galerkin linear system.

Theorem 6. *Let ϕ be a conformal one-to-one map of the simply connected domain D onto D_S. Assume $\phi(a) = -\infty$ and $\phi(b) = \infty$ and let $x_k = \phi^{-1}(kh)$. Further assume that there exist positive constants α, β, and K such that*

$$|F(x)| \leq K \begin{cases} e^{-\alpha|\phi(x)|}, & x\in\Gamma_a \\ e^{-\beta|\phi(x)|}, & x\in\Gamma_b \end{cases} \tag{2.35}$$

where $F = upw$, $u\phi' w$, or $u\left[\frac{\phi''}{\phi'}w + 2w'\right]$ and Γ_a and Γ_b are as defined in (2.22) *and* (2.23)*, respectively. Choose the positive integer M arbitrarily, let*

$$N = \left[\left|\frac{\alpha}{\beta}M + 1\right|\right] \tag{2.36}$$

and let

$$h = \sqrt{\frac{\pi d}{\alpha M}}. \tag{2.37}$$

(a) Let $vw \in B(D)$ *for* $v = f$ *or* qu*. Then*

$$\left| \int_a^b (vS_jw)(x)dx - h\left(\frac{vw}{\phi'}\right)(x_j) \right| \le L_0 M^{-1/2} e^{-\sqrt{\pi d\alpha M}} \tag{2.38}$$

where L_0 *is a constant depending on* v*,* w*, and* d*.*

(b) Let $u(pS_jw)' \in B(D)$ *and let the boundary term* $(upS_jw)(x)\Big|_a^b$ *be zero. Then*

$$\begin{aligned} &\left| \int_a^b (pu'S_jw)(x)dx + h\sum_{k=-M}^{N}(upw)(x_k)\frac{\delta_{jk}^{(1)}}{h} + h\left(\frac{u(pw)'}{\phi'}\right)(x_j)\right| \\ &\le L_1 M^{1/2} e^{-\sqrt{\pi d\alpha M}} \end{aligned} \tag{2.39}$$

where L_1 *is a constant depending on* u*,* p*,* w*,* ϕ*, and* d*.*

(c) Let $u(S_jw)'' \in B(D)$ *and let the boundary term* $-(u'S_jw)(x)\Big|_a^b + (u(S_jw)')(x)\Big|_a^b$ *be zero. Then*

$$\begin{aligned} &\left| \int_a^b (u''S_jw)(x)dx - h\sum_{k=-M}^{N} u(x_k)\left[\frac{\delta_{jk}^{(2)}}{h^2}\,(\phi'w)(x_k) \right.\right. \\ &\qquad \left.\left. + \frac{\delta_{jk}^{(1)}}{h}\left(\frac{\phi''}{\phi'}\,w + 2w'\right)(x_k)\right] - h\left(\frac{w''u}{\phi'}\right)(x_j)\right| \\ &\le \; L_2 M e^{-\sqrt{\pi d\alpha M}} \end{aligned} \tag{2.40}$$

where L_2 *depends on* u*,* w*,* ϕ*, and* d*.*

Now we apply Theorem 6 to the integrals in (2.32). Choose w so that the boundary terms, BT, vanish. Use the approximation

$$u_m(x) = \sum_{k=-M}^{N} u_k S_k(x), \quad m = M + N + 1$$

given in (2.30) in place of u, so that u_k replaces $u(x_k)$, and delete the error term of order $\mathcal{O}\left(Me^{-\sqrt{\pi d\alpha M}}\right)$ to obtain the following set of m equations where $-M \leq j \leq N$,

$$\sum_{k=-M}^{N}\left[-\frac{1}{h^2}\delta_{jk}^{(2)}\phi'(x_k)w(x_k) - \frac{1}{h}\delta_{jk}^{(1)}\left(\frac{\phi''(x_k)w(x_k)}{\phi'(x_k)} + 2w'(x_k)\right)\right]u_k$$
$$-\frac{w''(x_j)}{\phi'(x_j)}u_j - \sum_{k=-M}^{N}\frac{1}{h}\delta_{jk}^{(1)}p(x_k)w(x_k)u_k - \frac{(pw)'(x_j)}{\phi'(x_j)}u_j$$
$$+\frac{q(x_j)w(x_j)}{\phi'(x_j)}u_j = \frac{f(x_j)w(x_j)}{\phi'(x_j)}. \tag{2.41}$$

This system of equations may be expressed more conveniently in matrix form as

$$\left\{\frac{-1}{h^2}I^{(2)}D(\phi'w) - \frac{1}{h}I^{(1)}D\left(\frac{\phi''w}{\phi'} + 2w' + pw\right)\right. \tag{2.42}$$
$$\left.-D\left(\frac{w'' + (pw)' - qw}{\phi'}\right)\right\}\mathbf{u} = D\left(\frac{w}{\phi'}\right)\mathbf{f}$$

where the $m \times m$ diagonal matrix $D(g)$ is given by

$$D(g)=\begin{pmatrix} g(x_{-M}) & & \\ & \ddots & \\ & & g(x_N)\end{pmatrix},$$

the vectors $\mathbf{u}$ and $\mathbf{f}$ are defined by

$$\mathbf{u} = [u_{-M}, ..., u_N]^T$$

and

$$\mathbf{f} = [f(x_{-M}), ..., f(x_N)]^T$$

and the $m \times m$ Toeplitz matrices $I^{(p)}$, $p = 0, 1, 2$, are given by $I^{(p)} \equiv [\delta_{jk}^{(p)}]$. Note that $I^{(0)}$ is the identity matrix and $I^{(1)}$ and $I^{(2)}$ are the skew-symmetric and symmetric matrices, respectively, given by

$$I^{(1)}=\begin{pmatrix} 0 & -1 & \frac{1}{2} & \cdots & \frac{(-1)^{m-1}}{m-1} \\ 1 & \ddots & \ddots & \ddots & \vdots \\ -\frac{1}{2} & \ddots & \ddots & \ddots & \frac{1}{2} \\ \vdots & \ddots & \ddots & \ddots & -1 \\ \frac{(-1)^{m}}{m-1} & \cdots & -\frac{1}{2} & 1 & 0 \end{pmatrix}$$

and

$$I^{(2)}=\begin{pmatrix} -\frac{\pi^2}{3} & 2 & -\frac{2}{2^2} & \cdots & \frac{-2(-1)^{m-1}}{(m-1)^2} \\ 2 & \ddots & \ddots & \ddots & \vdots \\ -\frac{2}{2^2} & \ddots & \ddots & \ddots & -\frac{2}{2^2} \\ \vdots & \ddots & \ddots & \ddots & 2 \\ \frac{-2(-1)^{m-1}}{(m-1)^2} & \cdots & -\frac{2}{2^2} & 2 & -\frac{\pi^2}{3} \end{pmatrix}.$$

The choice of weight function w can have a significant impact on the difficulty of solving the system (2.42). The choice $w(x) = \dfrac{1}{\phi'(x)}$ for the weight function is thoroughly reviewed in [25] and the following theorem specifying the convergence rate for that choice of weight function is proven in [24]. In all examples below we assume solutions exist and are unique.

Theorem 7. *Assume that the coefficients* p*,* q*, and* f *in*

$$\begin{aligned} \mathcal{L}u(x) \equiv -u''(x) + p(x)u'(x) &+ q(x)u(x) = f(x), \quad a < x < b \\ u(a) &= u(b) = 0 \end{aligned} \tag{2.43}$$

and the solution u *are analytic in the simply connected domain* D*. Let* ϕ *be a conformal one-to-one map of* D *onto* D_S*. Assume also that* $f/\phi' \in B(D)$ *and* $uF \in B(D)$ *where*

$$F = (1/\phi')'',\ \phi',\ (\phi''/\phi'),\ (p/\phi')',\ p,\ (q/\phi').$$

Suppose there are positive constants C, α, *and* β *so that*

$$|u(x)| \leq C \left\{ \begin{array}{ll} e^{-\alpha|\phi(x)|}, & x \in \Gamma_a \\ e^{-\beta|\phi(x)|}, & x \in \Gamma_b. \end{array} \right. \tag{2.44}$$

If the Sinc-Galerkin solution is defined by

$$u_m(x) = \sum_{k=-M}^{N} u_k S_k(x) \tag{2.45}$$

where the positive integer M *is arbitrary,*

$$N = \left[\left| \frac{\alpha}{\beta} M + 1 \right|\right], \quad h = \sqrt{\frac{\pi d}{\alpha M}} \tag{2.46}$$

and the $\{u_k\}_{k=-M}^{N}$ *are found by solving the discrete system given in* (2.42)*, which for the choice* $w(x) = \dfrac{1}{\phi'(x)}$ *becomes*

$$\begin{aligned} &\left\{ \frac{-1}{h^2} I^{(2)} - \frac{1}{h} I^{(1)} D\left(\frac{p}{\phi'} - \frac{\phi''}{(\phi')^2} \right) \right. \\ &\left. - D\left(\frac{2(\phi'')^2}{(\phi')^4} - \frac{\phi''' + p\phi''}{(\phi')^3} + \frac{p' - q}{(\phi')^2} \right) \right\} \mathbf{u} = D\left(\frac{1}{(\phi')^2} \right) \mathbf{f}, \end{aligned} \tag{2.47}$$

then

$$||u - u_m||_\infty \leq K M^2 e^{-\sqrt{\pi d \alpha M}}. \tag{2.48}$$

We note, however, that if $p \equiv 0$ in (2.28) then the resulting differential operator is self-adjoint, and it is proven in [9] that choosing the weight function $w(x) = \dfrac{1}{\sqrt{\phi'(x)}}$ implies $\left(\dfrac{\phi'' w}{\phi'} + 2w' \right)(x) = \dfrac{\phi''(x)}{(\phi'(x))^{3/2}} + 2\left(\dfrac{1}{\sqrt{\phi'(x)}} \right)' \equiv 0$, so that the system in (2.42) becomes

$$\left\{\frac{-1}{h^2}I^{(2)} + D\left(\frac{\phi'''}{2(\phi')^3} - \frac{3(\phi'')^2}{4(\phi')^4} + \frac{q}{(\phi')^2}\right)\right\}\mathbf{y} = D\left(\frac{1}{(\phi')^{3/2}}\right)\mathbf{f} \tag{2.49}$$

where $\mathbf{y} \equiv D(\sqrt{\phi'})\mathbf{u}$. Since $I^{(2)}$ is symmetric, the Sinc-Galerkin system in (2.49) is a symmetric discrete system for computing the coefficients y_k in

$$y_m(x) = \sum_{k=-M}^{N} y_k S_k(x). \tag{2.50}$$

The identity $y_k = u_k\sqrt{\phi'(x_k)}$ allows us to compute the coefficients u_k and thus the approximate solution $u_m(x)$ of the self-adjoint problem. The resulting method is sometimes referred to as the *Symmetric Sinc-Galerkin Method.* The use of a symmetric system matrix enhances computational efficiency and plays a valuable role in the development of the Alternating-Direction Sinc-Galerkin method below.

The following theorem, proven in [11], states that exponential convergence is attainable with the symmetric Sinc-Galerkin method for self-adjoint two-point boundary-value problems.

Theorem 8. *Assume* $\sqrt{\phi'}(1/\sqrt{\phi'})'$, $\sqrt{\phi'}f$ *and* $(\sqrt{\phi'}(1/\sqrt{\phi'})'' - q)$ *and* y *are analytic in the simply connected domain* D. *Let* ϕ *be a conformal one-to-one map of* D *onto* D_S. *Assume also that* $f/\sqrt{\phi'} \in B(D)$ *and* $yF \in B(D)$, *where*

$$F = (1/\phi')'',\ \phi',\ \phi''/\phi',\ 1/\sqrt{\phi'}(1/\sqrt{\phi'})'',\ q/\phi'.$$

Suppose there are positive constants C_s, α_s, *and* β_s *so that*

$$|y(x)| = |u(x)(\phi'(x))^{1/2}| \le C_s \begin{cases} e^{-\alpha_s|\phi(x)|}, & x \in \Gamma_a \\ e^{-\beta_s|\phi(x)|}, & x \in \Gamma_b. \end{cases} \tag{2.51}$$

If

$$y_m(x) = \sum_{k=-M}^{N} y_k S_k(x) \tag{2.52}$$

is the assumed Sinc-Galerkin solution and the coefficients $\{y_k\}_{k=-M}^{N}$ *are determined from the discrete system in* (2.49) *with* M *an arbitrary positive integer,*

$$N = \left[\left|\frac{\alpha_s}{\beta_s}M + 1\right|\right], \quad \textit{and} \quad h_s = \sqrt{\frac{\pi d}{\alpha_s M}} \tag{2.53}$$

then

$$||y - y_m||_\infty \leq K_s M^2 e^{-\sqrt{\pi d \alpha_s M}}. \tag{2.54}$$

Next, we establish the relationship between α and α_s and between β and β_s for a finite interval (a, b). Let Γ be the finite interval (a, b) and suppose $x \in \Gamma$. We note that if $x \in \Gamma_a = \left(a, \dfrac{a+b}{2}\right)$ then $0 < \dfrac{x-a}{b-x} < 1$ and if $x \in \Gamma_b = \left[\dfrac{a+b}{2}, b\right)$ then $\dfrac{x-a}{b-x} \geq 1$. Since $\phi(x) = \ln\left(\dfrac{x-a}{b-x}\right)$, it follows that $\phi(x) < 0$ if $x \in \Gamma_a$ and $\phi(x) \geq 0$ if $x \in \Gamma_b$. From (2.44) we deduce that

$$|u(x)| \leq C \left\{ \begin{array}{ll} \left(\frac{x-a}{b-x}\right)^{\alpha}, & x \in \Gamma_a \\ \left(\frac{x-a}{b-x}\right)^{-\beta}, & x \in \Gamma_b. \end{array} \right. \tag{2.55}$$

Similarly, from (2.51) we deduce that

$$|u(x)(\phi'(x))^{1/2}| \leq C_s \left\{ \begin{array}{ll} \left(\frac{x-a}{b-x}\right)^{\alpha_s}, & x \in \Gamma_a \\ \left(\frac{x-a}{b-x}\right)^{-\beta_s}, & x \in \Gamma_b. \end{array} \right. \tag{2.56}$$

Note that $\phi'(x) = \dfrac{b-a}{(x-a)(b-x)}$. The condition (2.55) leads to

$$|u(x)| \leq K \left\{ \begin{array}{ll} (x-a)^{\alpha}, & x \in \Gamma_a \\ (b-x)^{\beta}, & x \in \Gamma_b \end{array} \right. \tag{2.57}$$

where K is constant, and the condition (2.56) leads to

$$|u(x)| \le K_s \begin{cases} (x-a)^{\alpha_s+1/2}, & x \in \Gamma_a \\ (b-x)^{\beta_s+1/2}, & x \in \Gamma_b, \end{cases} \tag{2.58}$$

where K_s is constant. Comparing (2.57) and (2.58) yields

$$\alpha_s = \alpha - 1/2 \tag{2.59}$$

$$\beta_s = \beta - 1/2. \tag{2.60}$$

In a similar manner, for problems on $\Gamma = [0, \infty)$, if we choose $\phi(x) = \ln(x)$ then

$$\alpha_s = \alpha - 1/2 \tag{2.61}$$

$$\beta_s = \beta + 1/2 \tag{2.62}$$

and if we choose $\phi(x) = \ln(\sinh(x)$ then

$$\alpha_s = \alpha - 1/2 \tag{2.63}$$

$$\beta_s = \beta. \tag{2.64}$$

Examples on the Unit Interval

In this section we employ Sinc-Galerkin methods to solve a number of two-point boundary-value problems to illustrate exponential convergence even in the presence of boundary singularities. Some of these examples will lead to a set of two-dimensional examples to be used later in this chapter and in Chapter 4. Since the problems in the examples considered are posed on the unit interval, we let $\phi(x) = \ln\left(\frac{x}{1-x}\right)$. Furthermore, we let $d = \pi/2$, $N = [|\alpha M/\beta + 1|]$ and $h = \sqrt{\pi d/\alpha M}$. Errors over the set of sinc grid points

$$\mathcal{S} \quad = \quad \{x_{-M}, ..., x_{-1}, x_0, x_1, ..., x_N\} \tag{2.65}$$

where

$$x_k \quad = \quad \frac{e^{kh}}{e^{kh}+1} \tag{2.66}$$

are reported as

$$||ES(h)|| = \max_{-M \le k \le N} |u(x_k) - u_k|. \tag{2.67}$$

Errors are also reported on a uniform grid

$$U \quad = \quad \{z_0, z_1, ..., z_{100}\} \tag{2.68}$$

where

$$h_u = 0.01 \quad \text{and} \quad z_k = kh_u \quad \text{for} \quad k = 0, 1, ..., 100. \tag{2.69}$$

These errors are given as

$$||EU(h_u)|| = \max_{0 \le k \le 100} |u(z_k) - u_m(z_k)|. \tag{2.70}$$

Example 4. *Consider the two-point boundary-value problem*

$$\begin{aligned} \mathcal{L}u(x) \equiv -u''(x) \quad &+ \quad \frac{u(x)}{x(1-x)} = -13x^2 + 13x - 2, \quad 0 < x < 1 \\ u(0) \quad &= \quad u(1) = 0 \end{aligned} \tag{2.71}$$

whose exact solution, $u(x) = x^2(1-x)^2$*, is known. Let the weight function be given by* $w(x) = \dfrac{1}{\phi'(x)}$ *and choose the positive integer,* M*, arbitrarily. From the form of the exact solution we note that we should select* $\alpha = \beta = 2$*. While in general we may choose* $N = [|\alpha M/\beta + 1|]$*, in problems where* $\alpha M/\beta$ *is an integer, one may choose* $N = \alpha M/\beta$ ([11, pp. 40-41])*. So here* $N = M$ *and from* (2.46) $h = \dfrac{\pi}{2\sqrt{M}}$*. The resulting discrete system from* (2.47)

$$\left\{\frac{-1}{h^2}I^{(2)} + \frac{1}{h}I^{(1)}D\left(2x-1\right) + D\left(3x-3x^2\right)\right\}\mathbf{u} = D\left(x^2(1-x)^2\right)\mathbf{f} \qquad (2.72)$$

allows the construction of the following set of errors in Table 2*.*

Table 2. Errors in the Sinc-Galerkin solution of (2.71).

M	N	m	h	$\|\|ES(h)\|\|$	$\|\|EU(h_u)\|\|$
2	2	5	1.11072	1.801e-003	1.917e-003
4	4	9	0.78540	4.318e-004	4.322e-004
8	8	17	0.55536	4.979e-005	5.079e-005
16	16	33	0.39270	1.733e-006	1.650e-006
32	32	65	0.27768	1.176e-008	1.116e-008

Note that we have exponential convergence even though $q(x) = \dfrac{1}{x(1-x)}$ *is singular at* $x = 0$ *and at* $x = 1$*.*

The next problem demonstrates that exponential convergence is attainable even when the exact solution, $u(x)$, is singular at a boundary point of the domain on which the problem is posed.

Example 5. *Consider the two-point boundary-value problem*

$$\begin{aligned} \mathcal{L}u(x) \equiv -u''(x) \quad &+ \quad \frac{u(x)}{x^2} = \frac{\ln(x)-1}{x}, \quad 0 < x < 1 \\ && (2.73) \\ u(0) \quad &= \quad u(1) = 0. \end{aligned}$$

Once again we solve the system using the weight function $w(x) = \dfrac{1}{\phi'(x)}$. *Since the exact solution,* $u(x) = x\ln(x)$, *is known we use it to determine* α *and* β. *We note that* $x\ln(x) \leq 1 - x$ *when* x *is close to 1 and* $x\ln(x) \leq x^{1-\epsilon}$ *for* $1 - \epsilon > 0$ *when* x *is close to 0. Thus we may select* $\beta = 1$ *and* $\alpha = 1 - \epsilon$. *For convenience we choose* $\alpha = \beta = 1$. *Since* $\alpha M/\beta$ *is an integer we choose* $N = M$ *and* $h = \pi/\sqrt{2M}$. *The resulting discrete system*

$$\left\{ \frac{-1}{h^2} I^{(2)} + \frac{1}{h} I^{(1)} D\left(2x - 1\right) + D\left(1 - x^2\right) \right\} \mathbf{u} = D\left(x^2(1-x)^2\right) \mathbf{f} \qquad (2.74)$$

gives the following errors in Table 3.

Table 3. Errors in the Sinc-Galerkin solution of (2.73).

M	N	m	h	$\|\|ES(h)\|\|$	$\|\|EU(h_u)\|\|$
2	2	5	1.57080	3.023e-002	4.333e-002
4	4	9	1.11072	1.497e-002	1.681e-002
8	8	17	0.78540	4.515e-003	1.544e-003
16	16	33	0.55536	6.076e-004	8.327e-005
32	32	65	0.39270	2.624e-005	2.445e-006

Example 6. *Now consider the following strongly singular problem*

$$\mathcal{L}u(x) \equiv -u''(x) \quad + \quad \frac{u(x)}{x^2} = f(x), \quad 0 < x < 1 \tag{2.75}$$

$$u(0) \quad = \quad u(1) = 0,$$

where $f(x) = \dfrac{4-\sqrt{x}}{4x}$. *For reference, the exact solution is* $u(x) = x - x\sqrt{x}$. *Assume* $\alpha = \beta = 1$. *For an arbitrary positive integer* M *we have* $N = M$ *and for* $d = \pi/2$ *we have* $h = \pi/\sqrt{2M}$. *The resulting discrete system*

$$\left\{\frac{-1}{h^2}I^{(2)} + \frac{1}{h}I^{(1)}D\left(2x-1\right) + D\left(1-x^2\right)\right\}\mathbf{u} = D\left(x^2(1-x)^2\right)\mathbf{f} \tag{2.76}$$

leads to the following errors in Table 4.

Table 4. Errors in the Sinc-Galerkin solution of (2.75).

M	N	m	h	$\|\|ES(h)\|\|$	$\|\|EU(h_u)\|\|$
2	2	5	1.57080	6.954e-003	8.888e-003
4	4	9	1.11072	2.690e-003	3.010e-003
8	8	17	0.78540	6.310e-004	4.628e-004
16	16	33	0.55536	6.423e-005	4.175e-005
32	32	65	0.39270	2.029e-006	1.215e-006

Note that even though $q(x)$, $f(x)$, and $u(x)$ are singular, exponential convergence is still attained.

In the next four examples we employ the symmetric Sinc-Galerkin method. We use $d = \pi/2$ and reiterate that $\alpha_s = \alpha - 1/2$, $\beta_s = \beta - 1/2$,

$$N = \left[\left|\frac{\alpha_s}{\beta_s}M + 1\right|\right], \quad \text{and} \quad h_s = \sqrt{\frac{\pi d}{\alpha_s M}}. \tag{2.77}$$

Tables 5, 6, 7, and 8 below, exhibit the errors for the symmetric Sinc-Galerkin method, where errors on the sinc grid are reported as

$$||ES^s(h_s)|| = \max_{-M\le k\le N} |u(x_k) - u_k^s| \tag{2.78}$$

and errors on the uniform grid are reported as

$$||EU^s(h_u)|| = \max_{0\le k\le 100} |u(z_k) - u_m^s(z_k)|. \tag{2.79}$$

Example 7. *Consider the problem given in* (2.71)*, which we restate here for reference.*

$$\begin{aligned} \mathcal{L}u(x) \equiv -u''(x) \quad &+ \quad \frac{u(x)}{x(1-x)} = -13x^2 + 13x - 2, \quad 0 < x < 1 \\ u(0) \quad &= \quad u(1) = 0. \end{aligned}$$

Note that the exact solution is $u(x) = x^2(1-x)^2$. *Choose the positive integer,* M*, arbitrarily and assume* $w(x) = \dfrac{1}{\sqrt{\phi'(x)}}$. *Recall that* $\alpha_s = \alpha - 1/2$ *and* $\beta_s = \beta - 1/2$. *Since in this example* $\alpha = \beta = 2$*, then we may choose* $\alpha_s = \beta_s = 3/2$. *Although in general* $N = [|\alpha_s M/\beta_s + 1|]$*, in this case* $\alpha_s M/\beta_s$ *is an integer and we can choose* $N = M$. *Since* $d = \pi/2$ *then from* (2.53) *we have that* $h_s = \sqrt{\pi d/\alpha_s M} = \pi/\sqrt{3M}$. *The resulting discrete system* (2.49) *for the coefficients* $\{u_k^s\}_{k=-M}^N$ *in the approximate solution*

$$u_m^s(x) = \sum_{k=-M}^{N} u_k^s S_k(x) \tag{2.80}$$

of (2.71) *is*

$$\left\{\frac{-1}{h^2} I^{(2)} + D\left(1/4 + x - x^2\right)\right\} D((x(1-x))^{-1/2})\mathbf{u^s} = D\left((x(1-x))^{3/2}\right)\mathbf{f}. \tag{2.81}$$

Solving this system we obtain the errors reported in Table 5*. Note that the symmetric Sinc-Galerkin method yields results even more accurate than the standard Sinc-Galerkin method, as can be seen by comparing* Table 5 *to* Table 2*.*

Table 5. Errors in the symmetric Sinc-Galerkin solution of (2.71).

M	N	m	h_s	$\|\|ES^s(h_s)\|\|$	$\|\|EU^s(h_u)\|\|$
2	2	5	1.28255	5.801e-004	1.916e-003
4	4	9	0.90690	1.307e-004	2.386e-004
8	8	17	0.64127	1.098e-005	1.293e-005
16	16	33	0.45345	2.237e-007	2.109e-007
32	32	65	0.32064	6.976e-010	6.581e-010

Example 8. *Next, we use the symmetric Sinc-Galerkin method to solve the problem given in* (2.73) *which we restate as*

$$\begin{aligned} \mathcal{L}u(x) \equiv -u''(x) \quad &+ \quad \frac{u(x)}{x^2} = \frac{\ln(x)-1}{x}, \quad 0 < x < 1 \\ u(0) \quad &= \quad u(1) = 0. \end{aligned}$$

The exact solution is $u(x) = x\ln(x)$*. Once again, let* $w(x) = \dfrac{1}{\sqrt{\phi'(x)}}$*. Since we chose* $\alpha = \beta = 1$ *in our analysis of* (2.73) *above, we now choose* $\alpha_s = \beta_s = 1/2$*, so that* $N = M$*, and since* $d = \pi/2$ *we have that* $h_s = \sqrt{\pi d/\alpha_s M} = \pi/\sqrt{M}$*. The resulting discrete system for the coefficients* $\{u_k^s\}_{k=-M}^{N}$ *in the approximate solution*

$$u_m^s(x) = \sum_{k=-M}^{N} u_k^s S_k(x) \tag{2.82}$$

of problem (2.73) *is*

$$\left\{\frac{-1}{h^2} I^{(2)} + D\left(1/4 + (1-x)^2\right)\right\} D((x(1-x))^{-1/2})\mathbf{u^s} = D\left((x(1-x))^{3/2}\right)\mathbf{f} \tag{2.83}$$

and it generates the errors in Table 6*. Again, by comparing to* Table 3 *we see that the symmetric Sinc-Galerkin method is more accurate.*

Table 6. Errors in the symmetric Sinc-Galerkin solution of (2.73).

M	N	m	h_s	$\|\|ES^s(h_s)\|\|$	$\|\|EU^s(h_u)\|\|$
2	2	5	2.22144	4.853e-003	2.109e-002
4	4	9	1.57080	1.920e-003	3.862e-003
8	8	17	1.11072	3.164e-004	3.639e-004
16	16	33	0.78540	1.611e-005	1.062e-005
32	32	65	0.55536	1.642e-007	6.218e-008

Example 9. *Next we consider the strongly singular problem given in* (2.75) *which we restate below for reference,*

$$\begin{aligned} \mathcal{L}u(x) \equiv -u''(x) &+ \frac{u(x)}{x^2} = f(x), \quad 0 < x < 1 \\ u(0) &= u(1) = 0 \end{aligned}$$

where $f(x) = \dfrac{4-\sqrt{x}}{4x}$. *Recall that the exact solution is* $u(x) = x - x\sqrt{x}$. *As before, assume* $\alpha = \beta = 1$ *so that* $\alpha_s = \beta_s = 1/2$. *For an arbitrary positive integer* M *we have that* $N = M$ *and for* $d = \pi/2$ *we compute* $h_s = \pi/\sqrt{M}$. *The resulting discrete system*

$$\left\{\frac{-1}{h^2} I^{(2)} + D\left(1/4 + (1-x)^2\right)\right\} D((x(1-x))^{-1/2})\mathbf{u}^{\mathrm{s}} = D\left((x(1-x))^{3/2}\right)\mathbf{f} \quad (2.84)$$

leads to the following errors in Table 7. *Comparing these results to those in* Table 4 *we see, as noted before, that the symmetric Sinc-Galerkin method is more accurate. Note once again that, despite having several singularities, excellent accuracy is still obtained.*

Table 7. Errors in the symmetric Sinc-Galerkin solution of (2.75).

M	N	m	h_s	$\|\|ES^s(h_s)\|\|$	$\|\|EU^s(h_u)\|\|$
2	2	5	2.22144	1.577e-003	9.326e-003
4	4	9	1.57080	2.584e-004	1.997e-003
8	8	17	1.11072	3.178e-005	1.943e-004
16	16	33	0.78540	1.213e-006	6.021e-006
32	32	65	0.55536	8.991e-009	4.147e-008

Example 10. *Lastly, we use the symmetric Sinc-Galerkin method to solve a problem in which* α_s *and* β_s *are different. Consider*

$$\begin{aligned} \mathcal{L}u(x) \equiv -u''(x) \quad &+ \quad \frac{u(x)}{x^2} = f(x), \quad 0 < x < 1 \\ & \\ u(0) \quad &= \quad u(1) = 0 \end{aligned} \tag{2.85}$$

where

$$f(x) = \frac{-3}{4\sqrt{x}} + \frac{15\sqrt{x}}{2} - \frac{35x\sqrt{x}}{4} + \frac{(1-x)^2}{\sqrt{x}}.$$

Since the exact solution, $u(x) = x^{3/2}(1-x)^2$, *is known, we see that* $\alpha = 3/2$ *and* $\beta = 2$ *so that* $\alpha_s = \alpha - 1/2 = 1$ *and* $\beta_s = \beta - 1/2 = 3/2$. *For an arbitrary positive integer* M *we have* $N = \alpha_s M/\beta_s = \frac{2}{3}M$ *if* M *is a multiple of* 3 *or* $N = [|\alpha_s M/\beta_s + 1|] = [|\frac{2}{3}M + 1|]$ *otherwise. Since* $d = \pi/2$ *we have, from* (2.53), *that* $h_s = \pi/\sqrt{2M}$. *The resulting discrete system*

$$\left\{ \frac{-1}{h^2} I^{(2)} + D\left(1/4 + (1-x)^2\right) \right\} D((x(1-x))^{-1/2})\mathbf{u^s} = D\left((x(1-x))^{3/2}\right)\mathbf{f} \tag{2.86}$$

leads to the errors reported in Table 8.

Examples on the Nonnegative Real Line

As it is pointed out in [11, p. 124], the convergence theorems above hold in the case where $b = +\infty$ with only minimal changes in the integrability requirements on the functions. Below we employ Sinc-Galerkin methods to solve ordinary differential equation boundary-value problems posed on $[0, \infty)$. For these problems we select

Table 8. Errors in the symmetric Sinc-Galerkin solution of (2.85).

M	N	m	h_s	$\|\|ES^s(h_s)\|\|$	$\|\|EU^s(h_u)\|\|$
2	2	5	1.57080	2.376e-003	7.204e-003
4	3	8	1.11072	1.758e-004	9.427e-004
9	7	17	0.74048	1.201e-005	2.938e-005
19	13	33	0.50963	7.128e-007	7.695e-007
38	26	65	0.36037	3.862e-009	3.718e-009

either $\phi(x) = \ln(x)$ or $\phi(x) = \ln(\sinh(x))$, where use of the latter requires exponential decay of the solution to guarantee an exponential convergence rate. Once again we let $d = \pi/2$, choose M arbitrarily, select appropriate values for α and β and define $h = \sqrt{\pi d/\alpha M}$. If $\alpha\, M/\beta$ is an integer then $N = \alpha\, M/\beta$ otherwise $N = [|\alpha M/\beta + 1|]$. We note that the set of sinc grid points is

$$\mathcal{S} \;=\; \{x_{-M}, ..., x_{-1}, x_0, x_1, ..., x_N\} \tag{2.87}$$

where from Table 1, $x_k = e^{kh}$ when $\phi(x) = \ln(x)$ and $x_k = \ln(e^{kh} + \sqrt{e^{2kh} + 1})$ when $\phi(x) = \ln(\sinh(x))$. Errors on the sinc grid are reported as

$$||ES(h)|| = \max_{-M \le k \le N} |u(x_k) - u_k|. \tag{2.88}$$

Errors are also reported on a uniform grid

$$U \;=\; \{z_0, z_1, ..., z_{100}\} \tag{2.89}$$

where if we let $h_u = 0.02$ and $h_v = 0.2$ then for $0 \le k \le 50,\ z_k = kh_u$ and for $51 \le k \le 100,\ z_k = 1 + (k - 50)h_v$. Since $\lim_{x\to\infty} u(x) = 0$, we restrict examination of the uniform error to values of x no larger than 11. These errors are given as

$$||EU(h_u)|| = \max_{0\le k\le 100} |u(z_k) - u_m(z_k)|. \tag{2.90}$$

Consider the following example.

Example 11.

$$\begin{aligned} \mathcal{L}u(x) \equiv -u''(x) \quad &+ \quad \frac{u(x)}{x^2+1} = f(x), \quad 0 < x < \infty \\ u(0) \quad &= \quad \lim_{x\to\infty} u(x) = 0 \end{aligned} \tag{2.91}$$

where $f(x) = \dfrac{x(7-x^2)}{(x^2+1)^3}$*. Since the exact solution,* $u(x) = \dfrac{x}{x^2+1}$*, is known and decays algebraically, we use* $\phi(x) = \ln(x)$ *and note that* $\alpha = \beta = 1$*. For an arbitrary positive integer* M *we have* $N = [|\alpha M/\beta|] = M$ *and again assuming* $d = \pi/2$ *we use* (2.46) *to compute* $h = \pi/\sqrt{2M}$*. The nodes are given by* $x_k = e^{kh}$*. Using the weight function* $w(x) = \dfrac{1}{\phi'(x)} = x$ *and solving the resulting discrete system* (2.47)

$$\left\{\frac{-1}{h^2}I^{(2)} - \frac{1}{h}I^{(1)} + D\left(\frac{x^2}{x^2+1}\right)\right\}\mathbf{u} \quad = \quad D\left(x^2\right)\mathbf{f} \tag{2.92}$$

gives us the following errors in Table 9*.*

Table 9. Errors in the Sinc-Galerkin solution of (2.91).

M	N	m	h	$\|\|ES(h)\|\|$	$\|\|EU(h_u)\|\|$
2	2	5	1.57080	7.269e-002	1.173e-001
4	4	9	1.11072	1.390e-002	2.516e-002
8	8	17	0.78540	1.455e-003	2.332e-003
16	16	33	0.55536	1.065e-004	1.600e-004
32	32	65	0.39270	2.457e-006	4.311e-006

Example 12. *For the following problem, in which the coefficient* $q(x) = \dfrac{1}{x^2}$ *of* $u(x)$ *contains a singularity on the boundary of the domain, we again employ a Sinc-Galerkin method. Consider*

$$\begin{aligned} \mathcal{L}u(x) \equiv -u''(x) + \frac{u(x)}{x^2} &= e^{-x}(2 - x + \frac{1}{x}), \quad 0 < x < \infty \\ &\qquad\qquad\qquad\qquad\qquad\qquad (2.93) \\ u(0) &= \lim_{x\to\infty} u(x) = 0. \end{aligned}$$

Since the exact solution, $u(x) = xe^{-x}$ *is known, and decays exponentially, we use* $\phi(x) = \ln(\sinh(x))$ *and note from* (2.44) *that* $\alpha = \beta = 1$. *For an arbitrary positive integer* M *we have* $N = M$ *and since* $d = \pi/2$ *we have* $h = \pi/\sqrt{2M}$. *The nodes are given by* $x_k = \ln(e^{kh} + \sqrt{e^{2kh} + 1})$. *Using the weight function* $w(x) = \dfrac{1}{\phi'(x)} = \tanh(x)$ *and* (2.47), *we solve the resulting discrete system*

$$\begin{aligned} \left\{ \frac{-1}{h^2} I^{(2)} - \frac{1}{h} I^{(1)} D(\operatorname{sech}^2(x)) \right. &- \left. D\left(\tanh^2(x)\left(2\operatorname{sech}^2(x) + \frac{1}{x^2}\right)\right)\right\} \mathbf{u} \\ &= D\left(\tanh^2(x)\right)\mathbf{f} \qquad (2.94) \end{aligned}$$

to obtain the following errors in Table 10.

Table 10. Errors in the Sinc-Galerkin solution of (2.93).

M	N	m	h	$\|\|ES(h)\|\|$	$\|\|EU(h_u)\|\|$
2	2	5	1.57080	2.398e-002	2.739e-002
4	4	9	1.11072	1.125e-002	1.319e-002
8	8	17	0.78540	3.156e-003	3.525e-003
16	16	33	0.55536	3.968e-004	4.205e-004
32	32	65	0.39270	1.615e-005	1.194e-005

In the next two examples we employ the symmetric Sinc-Galerkin method to solve the same two problems posed on the nonnegative real line. We let $d = \pi/2$, choose M arbitrarily, select appropriate values for α_s and β_s, and define $h_s = \sqrt{\pi d/\alpha_s M}$. If $\alpha_s M/\beta_s$ is an integer then we let $N = \alpha_s M/\beta_s$, otherwise we let $N = [|\alpha_s M/\beta_s + 1|]$. Tables 11 and 12 below, exhibit the errors for the symmetric Sinc-Galerkin method, where errors on the sinc grid are reported as

$$||ES^s(h_s)|| = \max_{-M \le k \le N} |u(x_k) - u_k^s| \tag{2.95}$$

and errors on the uniform grid are reported as

$$||EU^s(h_u)|| = \max_{0 \le k \le 100} |u(z_k) - u_m^s(z_k)|. \tag{2.96}$$

Example 13. *Consider the problem given in* (2.91) *which we rewrite below for reference.*

$$\begin{aligned} \mathcal{L}u(x) \equiv -u''(x) \quad &+ \quad \frac{u(x)}{x^2+1} = f(x), \quad 0 < x < \infty \\ u(0) \quad &= \quad \lim_{x \to \infty} u(x) = 0 \end{aligned}$$

where $f(x) = \dfrac{x(7-x^2)}{(x^2+1)^3}$ *and the exact solution is given by* $u(x) = \dfrac{x}{x^2+1}$. *We use* $\phi(x) = \ln(x)$ *and note that since* $\alpha = \beta = 1$ *then* $\alpha_s = \alpha - 1/2 = 1/2$ *and* $\beta_s = \beta + 1/2 = 3/2$. *For an arbitrary positive integer* M *we have that* $N = M/3$ *if* M *is divisible by* 3 *and* $N = [|M/3+1|]$ *otherwise. Using* $d = \pi/2$ *we compute* $h_s = \pi/\sqrt{M}$. *The nodes are given by* $x_k = e^{kh}$. *Using the weight function* $w(x) = \dfrac{1}{\sqrt{\phi'(x)}}$ *and* (2.49), *the resulting discrete system*

$$\left\{\frac{-1}{h^2}I^{(2)} + D\left(\frac{1}{4} + \frac{x^2}{x^2+1}\right)\right\} D\left(\frac{1}{\sqrt{x}}\right)\mathbf{u}^{\mathrm{s}} \quad = \quad D\left(x^{3/2}\right)\mathbf{f} \qquad (2.97)$$

gives us the errors in Table 11. *Note that here the symmetric Sinc-Galerkin method does not perform quite as well as the standard Sinc-Galerkin method, which can be verified by comparing the results here with those in* Table 9.

Table 11. Errors in the symmetric Sinc-Galerkin solution of (2.91).

M	N	m	h_s	$\|\|ES^s(h_s)\|\|$	$\|\|EU^s(h_u)\|\|$
3	1	5	1.81380	1.308e-001	1.888e-001
6	2	9	1.28255	3.524e-002	5.176e-002
12	4	17	0.90690	5.836e-003	7.459e-003
24	8	33	0.64127	1.645e-003	4.572e-004
48	16	65	0.45345	2.657e-004	1.605e-005

Example 14. *Now we employ the symmetric Sinc-Galerkin method to solve the problem in* (2.93) *which we rewrite below for reference.*

$$\mathcal{L}u(x) \equiv -u''(x) + \frac{u(x)}{x^2} \quad = \quad e^{-x}(2 - x + \frac{1}{x}), \quad 0 < x < \infty$$

$$u(0) \quad = \quad \lim_{x\to\infty} u(x) = 0$$

where we recall that the exact solution is $u(x) = xe^{-x}$. *We use* $\phi(x) = \ln(\sinh(x))$ *and note that since* $\alpha = \beta = 1$ *then* $\alpha_s = \alpha - 1/2 = 1/2$ *and* $\beta_s = \beta = 1$. *For any even positive integer* M *we have* $N = M/2$, *otherwise* $N = [|M/2+1|]$. *Using* $d = \pi/2$ *we have* $h_s = \pi/\sqrt{M}$. *The nodes are given by* $x_k = \ln(e^{kh} + \sqrt{e^{2kh}+1})$. *Using the weight function* $w(x) = \dfrac{1}{\sqrt{\phi'(x)}} = \sqrt{\tanh(x)}$ *and* (2.49), *we derive the discrete system*

$$\left\{\frac{-1}{h^2}I^{(2)} + D\left(\operatorname{sech}^2(x)\left(1-\frac{3}{4}\operatorname{sech}^2(x)\right) + \frac{\tanh^2(x)}{x^2}\right)\right\} D\left(\sqrt{\coth(x)}\right)\mathbf{u}^{\mathbf{s}} = D\left(\tanh^{\frac{3}{2}}(x)\right)\mathbf{f}, \qquad (2.98)$$

which we solve to obtain the errors in Table 12. *Comparing the results here with those in* Table 10, *we see again that the standard Sinc-Galerkin method performs slightly better than the symmetric Sinc-Galerkin method.*

Table 12. Errors in the symmetric Sinc-Galerkin solution of (2.93).

M	N	m	h_s	$\|\|ES^s(h_s)\|\|$	$\|\|EU^s(h_u)\|\|$
2	1	4	2.22144	1.189e-002	4.216e-002
5	3	9	1.40496	8.480e-003	1.435e-002
12	6	19	0.90690	5.524e-003	6.261e-003
22	11	34	0.66979	1.348e-003	1.486e-003
43	22	66	0.47909	9.526e-005	9.078e-005

Elliptic Problems on the Unit Square

In this section we apply the symmetric Sinc-Galerkin method to the Poisson problem

$$\begin{aligned} -\triangle u(x,y) &\equiv -(u_{xx}+u_{yy}) = f(x,y), \quad (x,y)\in\Omega=(0,1)\times(0,1) \\ & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (2.99) \\ u(x,y) &= 0, \quad (x,y)\in\partial\Omega. \end{aligned}$$

Let M_x, N_x, M_y and N_y be positive integers and assume the approximate solution to (2.99) takes the form

$$u^s_{m_x,m_y} = \sum_{j=-M_y}^{N_y}\sum_{i=-M_x}^{N_x} u^s_{ij}S_{ij}(x,y) \tag{2.100}$$

where

$$m_x = M_x+N_x+1, \quad m_y = M_y+N_y+1$$

and the basis functions $\{S_{ij}(x,y)\}$, $-M_x\le i\le N_x$, $-M_y\le j\le N_y$ are given by

$$S_{ij}(x,y) = [S(i,h_x)\circ\phi_x(x)][S(j,h_y)\circ\phi_y(y)]$$

for $\phi_z(z)=\ln\left(\dfrac{z}{1-z}\right)$, where $z=x$ or y. Define the inner product by

$$(f,g) = \int_0^1\int_0^1 f(x,y)g(x,y)v(x)w(y)dxdy$$

where the product $v(x)w(y)$ plays the role of a weight function. Assume that product is given by

$$v(x)w(y) = \frac{1}{\sqrt{\phi_x'(x)\phi_y'(y)}}.$$

Orthogonalizing the residual via

$$(-\triangle u^s_{m_x,m_y} - f, S_{kl}) = 0$$

for $-M_x \le k \le N_x$, $-M_y \le l \le N_y$ yields $m_x m_y$ equations. In a manner analogous to the development for ordinary differential equations above, the specifics of which are detailed in [11, p. 196], we apply Green's identity to remove the derivatives from u and choose the positive integer M_x arbitrarily. Assuming

$$|u(x,y)| \le C x^{\alpha_s+1/2}(1-x)^{\beta_s+1/2} y^{\zeta_s+1/2}(1-y)^{\eta_s+1/2}$$

for some constant C, we determine parameters α_s, β_s, ζ_s and η_s and then use these to compute $h_s \equiv h_x = h_y = \sqrt{\pi d/\alpha_s M_x}$ and the positive integers $N_x = [|\alpha_s M_x/\beta_s + 1|]$, $M_y = [|\alpha_s M_x/\zeta_s + 1|]$, and $N_y = [|\alpha_s M_x/\eta_s + 1|]$. (The discussion on how to choose N when $\alpha M/\beta$ is an integer also applies here.) Finally, use quadrature rules (deleting error terms) to approximate the iterated integrals and replace $u(x_p, y_q)$ with u^s_{pq} to derive the discrete sinc system (which, for generality, we have left in terms of $v(x)$, $w(y)$, h_x and h_y)

$$\begin{aligned}
\frac{w(y_l)}{\phi'_y(y_l)} \sum_{p=-M_x}^{N_x} \Bigg[-\frac{1}{h_x^2}\delta^{(2)}_{kp}\phi'_x(x_p)v(x_p) - \frac{1}{h_x}\delta^{(1)}_{kp}\left(\frac{\phi''_x(x_p)v(x_p)}{\phi'_x(x_p)} + 2v'(x_p)\right) \\
-\delta^{(0)}_{kp}\frac{v''(x_p)}{\phi'_x(x_p)}\Bigg] u^s_{pl} + \frac{v(x_k)}{\phi'_x(x_k)} \sum_{q=-M_y}^{N_y} \Bigg[-\frac{1}{h_y^2}\delta^{(2)}_{lq}\phi'_y(y_q)w(y_q) \\
-\frac{1}{h_y}\delta^{(1)}_{lq}\left(\frac{\phi''_y(y_q)w(y_q)}{\phi'_y(y_q)} + 2w'(y_q)\right) - \delta^{(0)}_{lq}\frac{w''(y_q)}{\phi'_y(y_q)}\Bigg] u^s_{kq} \\
= \frac{f(x_k,y_l)v(x_k)w(y_l)}{\phi'_x(x_k)\phi'_y(y_l)}.
\end{aligned} \tag{2.101}$$

For $p = 0$, 1 and 2, we define $I^{(p)}_{m_x}$ to be the $m_x \times m_x$ Toeplitz matrix with jk-th entry $\delta^{(p)}_{jk}$ and we define $I^{(p)}_{m_y}$ to be the $m_y \times m_y$ Toeplitz matrix with jk-th entry $\delta^{(p)}_{jk}$ (see Theorem 5). Let $D_{m_x}(g)$ denote the $m_x \times m_x$ diagonal matrix defined by

$$D_{m_x}(g) = \begin{bmatrix} g(x_{-M_x}) & & \\ & \ddots & \\ & & g(x_{N_x}) \end{bmatrix}$$

and define the $m_y \times m_y$ matrix $D_{m_y}(g)$ similarly. Let U and F be the $m_x \times m_y$ matrices with jk-th entries given by u^s_{jk} and $f(x_j, y_k) = f\left(\frac{e^{jh_x}}{e^{jh_x}+1}, \frac{e^{kh_y}}{e^{kh_y}+1}\right)$, respectively.

The discrete sinc system may be put into matrix form as the *Sylvester equation*

$$AX + XB = C \tag{2.102}$$

where noting that for $z = x$ or y,

$$\frac{-1}{(\phi'_z)^{3/2}} \left(\frac{1}{(\phi'_z)^{1/2}}\right)'' = \frac{1}{4}$$

and

$$\frac{\phi''_z}{(\phi'_z)^{3/2}} + 2\left((\phi'_z)^{-1/2}\right)' = 0,$$

we have

$$\begin{aligned} A &= D_{m_x}(\phi'_x)\left[-\frac{1}{h_x^2} I^{(2)}_{m_x} + \frac{1}{4} I^{(0)}_{m_x}\right] D_{m_x}(\phi'_x) \\ B &= D_{m_y}(\phi'_y)\left[-\frac{1}{h_y^2} I^{(2)}_{m_y} + \frac{1}{4} I^{(0)}_{m_y}\right] D_{m_y}(\phi'_y) \end{aligned} \tag{2.103}$$

$$\begin{aligned} C &= D_{m_x}((\phi'_x)^{-1/2})\, F\, D_{m_y}((\phi'_y)^{-1/2}) \\ X &= D_{m_x}((\phi'_x)^{-1/2})\, U\, D_{m_y}((\phi'_y)^{-1/2}). \end{aligned}$$

If the matrices A and B are diagonalizable via the matrices P and Q, respectively, so that $P^{-1}AP = \Lambda_a$ and $Q^{-1}BQ = \Lambda_b$ where Λ_a is a diagonal matrix (consisting of the eigenvalues $\{\alpha_i\}_{i=1}^{m_x}$ of A) and Λ_b is a diagonal matrix (consisting of the eigenvalues $\{\beta_j\}_{j=1}^{m_y}$ of B) then the Sylvester equation, (2.102), may be written as

$$P\Lambda_a P^{-1}X + XQ\Lambda_b Q^{-1} = C. \tag{2.104}$$

Multiplying (2.104) by P^{-1} on the left and by Q on the right, we obtain

$$\Lambda_a Y + Y\Lambda_b = Z \tag{2.105}$$

where $Y = P^{-1}XQ$ and $Z = P^{-1}CQ$. It follows that (2.105) has the component solution

$$y_{ij} = \frac{z_{ij}}{\alpha_i + \beta_j} \tag{2.106}$$

for $1 \leq i \leq m_x$, $1 \leq j \leq m_y$. X is recovered from Y via $X = PYQ^{-1}$.

Alternatively, we can solve the Sylvester equation using the Kronecker product, Kronecker sum, and the concatenation operator. According to [4, p.30], the Kronecker sum of the $m_x \times m_x$ matrix A and the $m_y \times m_y$ matrix B is denoted $A \oplus B$ and defined by

$$A \oplus B \equiv A \otimes I_{m_y}^{(0)} + I_{m_x}^{(0)} \otimes B \tag{2.107}$$

where $\otimes$ denotes the Kronecker product. We see in [4, p. 31] that (2.102) is equivalent to

$$(I_{m_y}^{(0)} \otimes A + B^T \otimes I_{m_x}^{(0)})co(X) = co(C) \tag{2.108}$$

and using the Kronecker sum we may rewrite (2.108) as

$$(B^T \oplus A)co(X) = co(C). \tag{2.109}$$

According to [4, p. 30], if $\{\lambda_i\}$ and $\{\mu_j\}$ are the eigenvalues of B^T and A, respectively, then $\{\lambda_i + \mu_j\}$ are the eigenvalues of $B^T \oplus A$. Since the eigenvalues of B are the same as the eigenvalues of B^T it follows from (2.109) and (2.108) that (2.102), has a unique solution if and only if no eigenvalue of A is the negative of an eigenvalue of B. (Clearly (2.109) may be solved using Gaussian elimination.) We also note that if $B = A^T$ then the Sylvester equation takes the form

$$AX + XA^T = C \tag{2.110}$$

and is referred to as a *Lyapunov equation.* We can transform this Lyapunov equation using (2.109) to obtain an equivalent matrix equation

$$(A \oplus A)co(X) = co(C). \tag{2.111}$$

Based on results in [11, p. 203], we again expect exponential accuracy as

$$||u - u^s_{m_x,m_y}||_\infty \le C_s M_x^2 e^{-\sqrt{\pi d \alpha_s M_x}} \tag{2.112}$$

for some constant C_s. For the examples below we set $d = \pi/2$, choose M_x arbitrarily and select parameters α_s, β_s, ζ_s and η_s. If $\alpha_s M_x/\beta_s$ is an integer then we let $N_x = \alpha_s M_x/\beta_s$ otherwise we let $N_x = [|\alpha_s M_x/\beta_s + 1|]$. If $\alpha_s M_x/\zeta_s$ is an integer then we let $M_y = \alpha_s M_x/\zeta_s$ otherwise we let $M_y = [|\alpha_s M_x/\zeta_s + 1|]$. If $\alpha_s M_x/\eta_s$ is an integer then we let $N_y = \alpha_s M_x/\eta_s$ otherwise we let $N_y = [|\alpha_s M_x/\eta_s + 1|]$. Note that balancing the different error terms in the inner product approximations dictates that $h_x = h_y$, so we define $h_s \equiv h_x = h_y = \sqrt{\dfrac{\pi d}{\alpha_s M_x}}$. We report the maximum errors over the set of sinc grid points

$$S = \{x_i\}_{i=-M_x}^{N_x} \times \{y_j\}_{j=-M_y}^{N_y} \tag{2.113}$$

and the set of uniform gridpoints (with stepsize $h_u = 0.01$)

$$U = \{w_i\}_{i=0}^{100} \times \{z_j\}_{j=0}^{100} \tag{2.114}$$

given $w_i = ih_u$ and $z_j = jh_u$ as

$$||ES^s(h_s)(x,y)|| = \max_{\substack{-M_x \le i \le N_x \\ -M_y \le j \le N_y}} |u(x_i, y_j) - u_{ij}^s|$$

and

$$||EU^s(h_u)(x,y)|| = \max_{\substack{0 \le i \le 100 \\ 0 \le j \le 100}} |u(w_i, z_j) - u_{m_x,m_y}^s(w_i, z_j)|,$$

respectively.

Examples on the Unit Square

Example 15. *Consider the elliptic problem*

$$\begin{aligned} -\triangle u(x,y) &= f(x,y), \quad (x,y) \in \Omega = (0,1) \times (0,1) \\ u(x,y) &= 0, \quad (x,y) \in \partial\Omega \end{aligned} \tag{2.115}$$

where $f(x,y) = -2x^2(1-x)^2(1-6y+6y^2) - 2y^2(1-y)^2(1-6x+6x^2)$ *and the exact solution is given by* $u(x,y) = x^2y^2(1-x)^2(1-y)^2$. *Note that this problem is a higher dimensional analogue of the problem in* Example 7. *Recall that* $\phi_z(z) = \ln\left(\dfrac{z}{1-z}\right)$ *for* $z = x$ *or* y. *Choose parameters* $\alpha_s = \beta_s = \zeta_s = \eta_s = 3/2$ *so that* $M \equiv M_x = N_x = M_y = N_y$ *and* $h_s \equiv h_x = h_y = \pi/\sqrt{3M_x}$ *for* $d = \pi/2$. *Hence* $m \equiv m_x = m_y = 2M+1$ *and the system size is* $m^2 \times m^2$. *Since the nodes* $x_i = e^{ih}/(1+e^{ih})$ *and* $y_j = e^{jh}/(1+e^{jh})$ *are also the same, then* $A = B$ *in the Sylvester equation* (2.102). *Since* $I^{(2)}_{m_x}$ *is symmetric, then* A *is symmetric and the resulting discrete system for the elliptic problem* (2.115) *is a Lyapunov equation,* (2.110)*, which we represent as a matrix equation of the form given in* (2.111). *This system is solved using Gaussian elimination to generate the errors in* Table 13. *Note the exponential convergence and the similar degree of accuracy when compared to the analogous one-dimensional problem whose results are exhibited in* Table 5.

Table 13. Errors in the symmetric Sinc-Galerkin solution of (2.115).

M	m	m^2	h_s	$\|\|ES^s(h_s)(x,y)\|\|$	$\|\|EU^s(h_u)(x,y)\|\|$
2	5	25	1.28255	5.084e-005	1.117e-004
4	9	81	0.90690	8.211e-006	1.495e-005
8	17	289	0.64127	6.893e-007	8.260e-007
16	33	1089	0.45345	1.400e-008	1.367e-008
32	65	4225	0.32064	4.361e-011	4.207e-011

Example 16. *Consider the elliptic problem*

$$-\triangle u(x,y) = -\left(\frac{x\ln(x)}{y}+\frac{y\ln(y)}{x}\right), \quad (x,y)\in\Omega=(0,1)\times(0,1) \tag{2.116}$$

$$u(x,y) = 0, \quad (x,y)\in\partial\Omega$$

whose exact solution is $u(x,y)=xy\ln(x)\ln(y)$. *As in* Example 8 *we see that one may choose the parameters* $\alpha_s=\beta_s=\zeta_s=\eta_s=1/2$ *so that* $M\equiv M_x=N_x=M_y=N_y$ *and* $h_s\equiv h_x=h_y=\pi/\sqrt{M}$ *when* $d=\pi/2$. *Since the nodes* $x_i=e^{ih}/(1+e^{ih})$ *and* $y_j=e^{jh}/(1+e^{jh})$ *are the same and as before* $I^{(2)}_{m_x}$ *is symmetric, then the resulting discrete system for the elliptic problem* (2.116) *is a Lyapunov equation which we may express in matrix form as*

$$(A\oplus A)co(X)\equiv(I^{(0)}_{m_x}\otimes A+A\otimes I^{(0)}_{m_x})co(X)=co(C) \tag{2.117}$$

where $I^{(0)}_{m_x}$ *is the* $m_x\times m_x$ *identity matrix and* $m\equiv m_x=m_y=2M+1$. *Note that the coefficient matrix in* (2.117) *is* $m^2\times m^2$. *Solving this system using Gaussian elimination we obtain the errors reported in* Table 14 *which are analogous to the errors reported in* Table 6. *These results illustrate the expected exponential convergence.*

Table 14. Errors in the symmetric Sinc-Galerkin solution of (2.116).

M	m	m^2	h_s	$\|\|ES^s(h_s)(x,y)\|\|$	$\|\|EU^s(h_u)(x,y)\|\|$
2	5	25	2.22144	3.952e-003	1.055e-002
4	9	81	1.57080	1.305e-003	2.788e-003
8	17	289	1.11072	1.830e-004	3.075e-004
16	33	1089	0.78540	8.564e-006	1.081e-005
32	65	4225	0.55536	7.969e-008	8.851e-008

Example 17. *Consider the elliptic problem*

$$-\triangle u(x,y) = -\frac{3}{4}\left[\frac{x\sqrt{x}-x}{\sqrt{y}}+\frac{y\sqrt{y}-y}{\sqrt{x}}\right], \quad (x,y)\in\Omega=(0,1)\times(0,1) \tag{2.118}$$

$$u(x,y) = 0, \quad (x,y)\in\partial\Omega$$

whose exact solution is $u(x,y)=xy(1-\sqrt{x})(1-\sqrt{y})$*. Let* $\alpha_s=\beta_s=\zeta_s=\eta_s=1/2$ *and let* $M\equiv M_x=N_x=M_y=N_y$*. For* $d=\pi/2$ *compute* $h_s\equiv h_x=h_y=\pi/\sqrt{M_x}$*. The resulting discrete system for the elliptic problem* (2.118) *can be expressed in matrix form as*

$$(A\oplus A)co(X)=co(C) \tag{2.119}$$

where as before $m\equiv m_x=m_y=2M+1$*, the* $m\times m$ *matrix* A *is given in* (2.103) *and the coefficient matrix* $A\oplus A$ *in* (2.119) *is* $m^2\times m^2$*. Solving this system using Gaussian elimination we obtain the errors reported in* Table 15*. Note that the method performs quite well and the errors here compare favorably with the errors reported for the one-dimensional analogue of this problem. (See* Table 7 *above.)*

Table 15. Errors in the symmetric Sinc-Galerkin solution of (2.118).

M	m	m^2	h_s	$\|\|ES^s(h_s)(x,y)\|\|$	$\|\|EU^s(h_u)(x,y)\|\|$
2	5	25	2.22144	2.531e-004	1.911e-003
4	9	81	1.57080	7.006e-005	3.691e-004
8	17	289	1.11072	7.765e-006	3.433e-005
16	33	1089	0.78540	2.596e-007	1.201e-006
32	65	4225	0.55536	1.746e-009	8.954e-009

We conclude this section with an example in which M_x, N_x, M_y, and N_y may all be different.

Example 18. *Consider the following elliptic problem.*

$$\begin{aligned} -\triangle u(x,y) &= f(x,y), \quad (x,y) \in \Omega = (0,1)\times(0,1) \\ & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (2.120) \\ u(x,y) &= 0, \quad (x,y) \in \partial\Omega, \end{aligned}$$

where $f(x,y) = 2xy^3\left[3y^2(-1+y)^2(-1+2x) + x^2(x-1)(10-30y+21y^2)\right]$. *Note that the exact solution is* $u(x,y) = x^3(1-x)y^5(1-y)^2$. *Let* $\alpha_s = 5/2$, $\beta_s = 1/2$, $\zeta_s = 9/2$, *and* $\eta_s = 3/2$. *Choose* M_x *arbitrarily and compute* $N_x = 5M_x$. *Note that* $M_y = \frac{5}{9}M_x$ *when* M_x *is a multiple of* 9 *and* $M_y = \left[\left|\frac{5}{9}M_x + 1\right|\right]$ *otherwise. Similarly,* $N_y = \frac{5}{3}M_x$ *when* M_x *is a multiple of* 3 *and* $N_y = \left[\left|\frac{5}{3}M_x + 1\right|\right]$ *otherwise. Compute* $m_x = M_x + N_x + 1$, $m_y = M_y + N_y + 1$ *and for* $d = \pi/2$ *note that* $h_s \equiv h_x = h_y = \pi/\sqrt{5M_x}$. *Determine the* $m_x \times m_x$ *matrix* A, *the* $m_y \times m_y$ *matrix* B, *and the* $m_x \times m_y$ *matrix* C *using* (2.103). *Use the Kronecker sum to express the resulting Sylvester equation* *for the solution of* (2.120) *as*

$$(B^T \oplus A)co(X) = co(C). \tag{2.121}$$

Solve (2.121) *using Gaussian elimination to obtain the errors in* Table 16.

Table 16. Errors in the symmetric Sinc-Galerkin solution of (2.120).

M_x	N_x	m_x	M_y	N_y	m_y	h_s	$\|\|ES^s(h_s)(x,y)\|\|$	$\|\|EU^s(h_u)(x,y)\|\|$
1	5	7	1	2	4	1.40496	1.267e-004	3.870e-004
2	10	13	2	4	7	0.99346	1.451e-005	4.679e-005
3	15	19	2	5	8	0.81116	6.743e-006	1.230e-005
5	25	31	3	9	13	0.62832	4.142e-007	8.439e-007
11	55	67	7	19	27	0.42361	5.044e-009	6.168e-009

Elliptic Problems on the First Quadrant

We now apply the symmetric Sinc-Galerkin method to the Poisson problem on the first quadrant given by

$$\begin{aligned} -\triangle u(x,y) &= f(x,y), \quad (x,y) \in \Omega = (0,\infty)\times(0,\infty) \\ u(x,0) &= 0, \quad x \in (0,\infty) \\ u(0,y) &= 0, \quad y \in (0,\infty) \\ u(x,y) &\to 0 \quad as \quad x \to \infty \\ u(x,y) &\to 0 \quad as \quad y \to \infty \end{aligned} \tag{2.122}$$

Though little work has been done on using the symmetric Sinc-Galerkin method on this class of problems, the technique is developed here to provide yet another setting in which to test the ADSG method presented in Chapter 4. Mimicking the development for the elliptic problem on the unit square above, we let M_x, N_x, M_y and N_y be positive integers and assume the approximate solution to (2.122) takes the form

$$u^s_{m_x,m_y} = \sum_{j=-M_y}^{N_y} \sum_{i=-M_x}^{N_x} u^s_{ij} S_{ij}(x,y) \tag{2.123}$$

where

$$m_x = M_x + N_x + 1, \quad m_y = M_y + N_y + 1.$$

While the basis functions $\{S_{ij}(x,y)\}$, $-M_x \le i \le N_x$, $-M_y \le j \le N_y$ are still given by

$$S_{ij}(x,y) = [S(i,h_x) \circ \phi_x(x)][S(j,h_y) \circ \phi_y(y)],$$

we note that for problems on the first quadrant we may choose either $\phi_z(z) = \ln(z)$, or $\phi_z(z) = \ln(\sinh(z))$ where $z = x$ or y. We note further that the choice of ϕ for $z = x$ may differ from the choice of ϕ for $z = y$. The choice of ϕ will determine the nodes. Define the inner product by

$$(f,g) = \int_0^\infty \int_0^\infty f(x,y)g(x,y)v(x)w(y)dxdy$$

where as before the product $v(x)w(y)$ is given by

$$v(x)w(y) = \frac{1}{\sqrt{\phi'_x(x)\phi'_y(y)}}.$$

Orthogonalize the residual via

$$(-\triangle u^s_{m_x,m_y} - f, S_{kl}) = 0.$$

In a manner analogous to the development for problems on the unit square above, we remove the derivatives from u and choose the positive integer M_x arbitrarily. Assuming that either

$$|u(x,y)| \leq Cx^{\alpha_s+1/2}x^{-\beta_s+1/2}y^{\zeta_s+1/2}y^{-\eta_s+1/2} \tag{2.124}$$

or

$$|u(x,y)| \leq Cx^{\alpha_s+1/2}e^{-\beta_s x}y^{\zeta_s+1/2}e^{-\eta_s y} \tag{2.125}$$

holds for some constant C, we determine parameters α_s, β_s, ζ_s and η_s and then use these to compute $h_s \equiv h_x = h_y = \sqrt{\pi d/\alpha_s M_x}$ and the positive integers $N_x = [|\alpha_s M_x/\beta_s+1|]$, $M_y = [|\alpha_s M_x/\zeta_s+1|]$, and $N_y = [|\alpha_s M_x/\eta_s+1|]$. (Again, the discussion on how to choose N when $\alpha M/\beta$ is an integer also applies here.) Finally, we use quadrature rules (deleting error terms) to approximate the iterated integrals and derive the discrete sinc system given in (2.101), where $\phi_x(x)$ and $\phi_y(y)$ have been suitably replaced. We define $I^{(p)}_{m_x}$, $I^{(p)}_{m_y}$, $D_{m_x}(g)$, $D_{m_y}(g)$ and U as before. We also define the $m_x \times m_y$ matrix F according to our choices for $\phi_x(x)$ and $\phi_y(y)$ as follows. Suppose $\phi_z(z)$ represents either $\phi_x(x)$ or $\phi_y(y)$. Let $\psi_z(z) = \phi_z^{-1}(z)$, the inverse of $\phi_z(z)$. If $\phi_z(z) = \ln(z)$ then $\psi_z(z) = e^z$. If $\phi_z(z) = \ln(\sinh(z))$ then $\psi_z(z) = \ln(e^z + \sqrt{e^{2z}+1})$. Now let F be the $m_x \times m_y$ matrix with jk-th entry given by $f(x_j, y_k) = f\left(\psi_x(jh_x), \psi_y(kh_y)\right)$.

The discrete sinc system may be put into matrix form as the Sylvester equation

$$AX + XB \;=\; C \tag{2.126}$$

and we note that in this case if $\phi_z(z) = \ln(z)$ then we still have that

$$\frac{-1}{(\phi_z')^{3/2}} \left(\frac{1}{(\phi_z')^{1/2}} \right)'' = \frac{1}{4}.$$

However, if $\phi_z(z) = \ln(\sinh(z))$ then we have

$$\frac{-1}{(\phi_z')^{3/2}} \left(\frac{1}{(\phi_z')^{1/2}} \right)'' = \frac{4\cosh^2(z) - 3}{4\cosh^4(z)}$$

and the matrix A and/or the matrix B must be modified accordingly. In either case

$$\frac{\phi_z''}{(\phi_z')^{3/2}} + 2\left((\phi_z')^{-1/2}\right)' = 0.$$

We report the maximum errors over the set of sinc grid points

$$S = \{x_i\}_{i=-M_x}^{N_x} \times \{y_j\}_{j=-M_y}^{N_y} \tag{2.127}$$

and the set of uniform gridpoints

$$U = \{w_i\}_{i=0}^{100} \times \{z_j\}_{j=0}^{100} \tag{2.128}$$

as

$$||ES^s(h_s)(x,y)|| = \max_{\substack{-M_x \le i \le N_x \\ -M_y \le j \le N_y}} |u(x_i, y_j) - u_{ij}^s|$$

and

$$||EU^s(h_u)(x,y)|| = \max_{\substack{0 \le i \le 100 \\ 0 \le j \le 100}} |u(w_i, z_j) - u_{m_x,m_y}^s(w_i, z_j)|,$$

respectively. Let $h_u = 0.02$ and $h_v = 0.2$. Note that if $0 \le i \le 50$ then $w_i = ih_u$, whereas if $51 \le i \le 100$ then $w_i = 1 + (i - 50)h_v$. Similarly, if $0 \le j \le 50$ then $z_j = jh_u$, whereas if $51 \le j \le 100$ then $z_j = 1 + (j - 50)h_v$.

Examples on the First Quadrant

Example 19. *Consider the elliptic problem*

$$\begin{aligned} -\triangle u(x,y) &= f(x,y), \quad (x,y) \in \Omega = (0,\infty)\times(0,\infty) \\ u(x,0) &= 0, \quad x \in (0,\infty) \\ u(0,y) &= 0, \quad y \in (0,\infty) \qquad (2.129) \\ u(x,y) &\to 0 \quad as \quad x \to \infty \\ u(x,y) &\to 0 \quad as \quad y \to \infty \end{aligned}$$

where $f(x,y) = \dfrac{2xy}{(x^2+1)(y^2+1)}\left[\dfrac{3-x^2}{(x^2+1)^2} + \dfrac{3-y^2}{(y^2+1)^2}\right]$ *and the exact solution is given by* $u(x,y) = \dfrac{xy}{(x^2+1)(y^2+1)}$. *This is analogous to the problem in* Example 13. *Since the exact solution decays algebraically with respect to* x *or* y, *then for* $z = x$ *or* y *we choose* $\phi_z(z) = \ln(z)$. *We also note that the exact solution satisfies* (2.124) *and hence we can choose the parameters* $\alpha_s = \zeta_s = 1/2$ *and* $\beta_s = \eta_s = 3/2$ *so that* $M \equiv M_x = M_y$, $N \equiv N_x = N_y = [|\frac{1}{3}M + 1|]$ *and* $h_s \equiv h_x = h_y = \pi/\sqrt{M}$ *when* $d = \pi/2$. *(Note that if* M *is divisible by* 3 *then* $N = \dfrac{M}{3}$.*) Let* $m \equiv m_x = m_y = M+N+1$. *Since* $I^{(2)}_{m_x}$ *is symmetric, then the matrix* A *in the Sylvester equation* (2.102) *is symmetric. Since the nodes* $x_i = e^{ih}$ *and* $y_j = e^{jh}$ *are the same, then* $A = B$ *and the resulting discrete system for the elliptic problem* (2.129) *is a Lyapunov equation which we represent as a matrix equation of the form given in* (2.117) *and solve to generate the errors in* Table 17. *Note that the uniform errors in* Table 17 *are analogous to those reported in* Table 11.

Table 17. Errors in the symmetric Sinc-Galerkin solution of (2.129).

M	N	m	h_s	$\|\|ES^s(h_s)(x,y)\|\|$	$\|\|EU^s(h_u)(x,y)\|\|$
3	1	5	1.81380	6.507e-002	1.074e-001
6	2	9	1.28255	2.368e-002	3.310e-002
12	4	17	0.90690	3.396e-003	5.345e-003
24	8	33	0.64127	1.912e-004	3.226e-004
48	16	65	0.45345	4.832e-006	1.253e-005

Example 20. *Consider the elliptic problem*

$$\begin{aligned} -\triangle u(x,y) &= f(x,y), \quad (x,y)\in\Omega=(0,\infty)\times(0,\infty) \\ u(x,0) &= 0, \quad x\in(0,\infty) \\ u(0,y) &= 0, \quad y\in(0,\infty) \\ u(x,y) &\rightarrow 0 \quad as \quad x\rightarrow\infty \\ u(x,y) &\rightarrow 0 \quad as \quad y\rightarrow\infty \end{aligned} \tag{2.130}$$

where $f(x,y) = 2(x-xy+y)e^{-(x+y)}$ *and the exact solution is given by* $u(x,y) = xye^{-(x+y)}$. *We note that the exact solution has the form given in* (2.125) *and thus we can choose the parameters* $\alpha_s = \zeta_s = 1/2$ *and* $\beta_s = \eta_s = 1$. *Once again it follows that* $M \equiv M_x = M_y$ *and either* $N \equiv N_x = N_y = \dfrac{M}{2}$ *if* M *is even or* $N \equiv N_x = N_y = [|\frac{1}{2}M+1|]$ *otherwise. Let* $m \equiv m_x = m_y = M+N+1$. *We also have that* $h_s \equiv h_x = h_y = \pi/\sqrt{M}$ *when* $d = \pi/2$. *Since the exact solution decays exponentially with respect to* x *or* y *then for* $z = x$ *or* y, *we choose* $\phi_z(z) = \ln(\sinh(z))$ *and compute the nodes* $x_i = \ln(e^{ih}+\sqrt{e^{2ih}+1})$ *and* $y_j = \ln(e^{jh}+\sqrt{e^{2jh}+1})$. *We again derive a Lyapunov equation which we transform into an equation of the form given* *in* (2.117) *and solve using Gaussian elimination. We thus obtain the errors in* Table 18. *Note that the errors in* Table 18 *are roughly comparable to those in* Table 12.

Table 18. Errors in the symmetric Sinc-Galerkin solution of (2.130).

M	N	m	h_s	$\|\|ES^s(h_s)(x,y)\|\|$	$\|\|EU^s(h_u)(x,y)\|\|$
2	1	4	2.22144	8.104e-003	1.861e-002
5	3	9	1.40496	1.718e-003	4.745e-003
12	6	19	0.90690	1.295e-003	2.258e-003
22	11	34	0.66979	3.359e-004	5.175e-004
43	22	66	0.47909	2.553e-005	1.762e-005

CHAPTER 3

ALTERNATING-DIRECTION FINITE DIFFERENCE METHODS

Introduction

When applied to a parabolic partial differential equation such as the diffusion equation, explicit finite difference schemes require a large number of small time steps to ensure stability (and convergence) whereas implicit finite difference schemes require the solution of a large system of equations at each time step. The Crank-Nicolson scheme combines both explicit and implicit schemes. The Alternating-Direction Implicit (ADI) method, introduced in [20], is a splitting of the Crank-Nicolson scheme that combines both explicit and implicit schemes.

The ADI method, now part of a class known as operator splitting methods, led to the development of more accurate and efficient algorithms for the numerical approximation of solutions to partial differential equations. This is mainly because it reduced the large matrix systems necessary for such approximations to simple tridiagonal subsystems.

In an attempt to develop easily implemented algorithms with a great increase in accuracy over standard finite difference methods, ADI and Galerkin methods were combined and extensively analyzed in [3]. More recently, in [18] and [19], banded preconditioners and the conjugate gradient method have been employed to obtain fast accurate solutions to matrix systems such as that found in (2.108). In the Alternating-Direction Sinc-Galerkin (ADSG) method [1], which is described in detail in Chapter 4 below, an attempt is made to avoid both the very large matrices (and the corresponding extensive storage requirements) that result when a system matrix is constructed as a Kronecker sum and the computation of spectral bounds (which are typically employed in ADI schemes to determine optimum iteration parameters).

Finite Difference Methods

To more comprehensively illustrate the development that leads to the Alternating-Direction Sinc-Galerkin method we begin with the following parabolic problem.

Example 21. *Consider*

$$\frac{\partial u}{\partial t} = \triangle u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}, \qquad (x,y) \in \Omega = (0,1)\times(0,1), \quad t > 0 \tag{3.1}$$

with Dirichlet boundary conditions

$$\begin{aligned} u(0,y,t) &= 0, \quad y \in (0,1), \quad t > 0 \\ u(1,y,t) &= 0, \quad y \in (0,1), \quad t > 0 \\ u(x,0,t) &= 0, \quad x \in (0,1), \quad t > 0 \\ u(x,1,t) &= 0, \quad x \in (0,1), \quad t > 0 \end{aligned} \tag{3.2}$$

and the initial condition

$$u(x,y,0) = \sin(\pi x)\,\sin(\pi y), \qquad (x,y) \in \Omega. \tag{3.3}$$

For reference, the exact solution to this problem is given by

$$u(x,y,t) = e^{-2\pi^2 t} \sin(\pi x) \sin(\pi y).$$

We attempt to find a numerical solution to this problem by proceeding as follows.

Choose positive integers, M and N, arbitrarily. Define the spatial increments

$$\Delta x = \Delta y = \frac{1}{M+1}. \tag{3.4}$$

Choose the time increment Δt and set

$$\mu \equiv \frac{\Delta t}{(\Delta x)^2} = \frac{\Delta t}{(\Delta y)^2}. \tag{3.5}$$

Discretize the computational domain so that

$$\begin{aligned} x_i &= i\Delta x \quad \text{for } i = 0,1,2,...,M+1, \\ y_j &= j\Delta y \quad \text{for } j = 0,1,2,...,M+1, \quad \text{and} \\ t_k &= k\Delta t \quad \text{for } k = 0,1,2,...,N. \end{aligned}$$

Assume u_{ij}^k approximates $u(x_i, y_j, t_k)$.

An Explicit Finite Difference Scheme

Consider the following explicit finite difference scheme for solving (3.1)-(3.3). Approximate the partial derivatives using

$$\frac{u_{ij}^{k+1} - u_{ij}^k}{\Delta t} = \frac{u_{i-1j}^k - 2u_{ij}^k + u_{i+1j}^k}{(\Delta x)^2} + \frac{u_{ij-1}^k - 2u_{ij}^k + u_{ij+1}^k}{(\Delta y)^2} \tag{3.6}$$

and rearrange terms to obtain the iterative scheme (for $i, j = 1, ..., M$)

$$u_{ij}^{k+1} = (1 - 4\mu)u_{ij}^k + \mu(u_{i-1j}^k + u_{i+1j}^k + u_{ij-1}^k + u_{ij+1}^k). \tag{3.7}$$

From (3.3) we can start this iteration with

$$u_{ij}^0 \quad = \quad \sin(\pi x_i)\,\sin(\pi y_j) \tag{3.8}$$

We use lexicographic ordering of the grid points (left-to-right, bottom-to-top) along with (3.7) to approximate the solution to the parabolic problem in (3.1)-(3.3). Then for $k = 0, 1, 2, ..., N - 1$, the scheme may be represented using the matrix equation

$$\mathbf{u}^{k+1} = A_E\mathbf{u}^k \tag{3.9}$$

where the $M^2 \times 1$ vector $\mathbf{u}^k$ is defined by

$$\mathbf{u}^k = [u_{11}^k, u_{21}^k, ..., u_{M\ 1}^k, u_{12}^k, u_{22}^k, ..., u_{M\ 2}^k, ..., u_{1\ M}^k, u_{2\ M}^k \cdots, u_{M\ M}^k]^T, \tag{3.10}$$

A_E is the $M^2 \times M^2$ block Toeplitz, symmetric, block symmetric, block tridiagonal matrix given by

$$A_E = \begin{pmatrix} P & \mu I & O & \cdots & O \\ \mu I & \ddots & \ddots & \ddots & \vdots \\ O & \ddots & \ddots & \ddots & O \\ \vdots & \ddots & \ddots & \ddots & \mu I \\ O & \cdots & O & \mu I & P \end{pmatrix},$$

I and O are the $M \times M$ identity and zero matrices, respectively, and P is the $M \times M$ Toeplitz, symmetric and tridiagonal (TST) matrix given by

$$P=\begin{pmatrix} 1-4\mu & \mu & 0 & \cdots & 0 \\ \mu & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \mu \\ 0 & \cdots & 0 & \mu & 1-4\mu \end{pmatrix}.$$

Note that $\mathbf{u}^0$ is defined by (3.8). To analyze the stability, consistency and convergence of this and other finite difference schemes, it is often helpful to determine the eigenvalues and eigenvectors of the matrices used in their construction. We note from [22, pp. 113-115] that a common $M \times M$, Toeplitz, tridiagonal matrix T of the form

$$T=\begin{pmatrix} a & b & 0 & \cdots & 0 \\ c & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & b \\ 0 & \cdots & 0 & c & a \end{pmatrix}$$

has eigenvalues, λ_k, for $k = 1, ..., M$ given by

$$\lambda_k = a + 2\sqrt{bc}\,\cos\left(\frac{\pi k}{M+1}\right) \tag{3.11}$$

and eigenvectors, $\mathbf{q}_k$, for $k = 1, ..., M$ given by

$$\begin{aligned} \mathbf{q}_k \;\; = \;\; & \left[\left(\frac{c}{b}\right)^{\frac{1}{2}} \sin\left(\frac{\pi k}{M+1}\right), \frac{c}{b} \sin\left(\frac{2\pi k}{M+1}\right), \left(\frac{c}{b}\right)^{\frac{3}{2}} \sin\left(\frac{3\pi k}{M+1}\right),\right. \\ & \left. ..., \left(\frac{c}{b}\right)^{\frac{M}{2}} \sin\left(\frac{M\pi k}{M+1}\right)\right]^T . \end{aligned} \tag{3.12}$$

The explicit scheme above is of *order* $\mathcal{O}(\Delta t + (\Delta x)^2)$. (The order of a scheme is also called the *local truncation error.*) However, von Neumann stability analysis reveals it is stable only if $\mu \leq \frac{1}{4}$ (see [22, p. 38]) so that $\frac{\Delta t}{(\Delta x)^2} \leq \frac{1}{4}$ or $\Delta t \leq \frac{1}{4(M+1)^2}$.

Thus for $\Delta x = 0.01$ we must have $\Delta t \leq \frac{1}{40000}$. We note that stabiity requires an enormous computational cost.

An Implicit Finite Difference Scheme

Instead of the explicit scheme above we may construct an implicit finite difference scheme to solve (3.1)-(3.3) by approximating the partial derivatives as follows

$$\frac{u_{ij}^{k+1} - u_{ij}^{k}}{\Delta t} = \frac{u_{i-1j}^{k+1} - 2u_{ij}^{k+1} + u_{i+1j}^{k+1}}{(\Delta x)^2} + \frac{u_{ij-1}^{k+1} - 2u_{ij}^{k+1} + u_{ij+1}^{k+1}}{(\Delta y)^2} \tag{3.13}$$

and then rearranging terms to obtain (for $i, j = 1, ..., M$)

$$\mu(u_{i-1j}^{k+1} + u_{i+1j}^{k+1} + u_{ij-1}^{k+1} + u_{ij+1}^{k+1}) - (4\mu + 1)u_{ij}^{k+1} = -u_{ij}^{k}. \tag{3.14}$$

If we define the $M^2 \times 1$ vector $\mathbf{u}^k$ as in (3.10) (which corresponds to using the lexicographic ordering of the grid points), then the scheme in (3.14) may be represented in matrix form as

$$A_I \mathbf{u}^{k+1} = \mathbf{u}^k \tag{3.15}$$

for $k = 0, 1, 2, ..., N-1$, where A_I is given by

$$A_I = \begin{pmatrix} Q & -\mu I & O & \cdots & O \\ -\mu I & \ddots & \ddots & \ddots & \vdots \\ O & \ddots & \ddots & \ddots & O \\ \vdots & \ddots & \ddots & \ddots & -\mu I \\ O & \cdots & O & -\mu I & Q \end{pmatrix},$$

and Q is the $M \times M$, TST, positive definite matrix given by

$$Q=\begin{pmatrix} 1+4\mu & -\mu & 0 & \cdots & 0 \\ -\mu & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -\mu \\ 0 & \cdots & 0 & -\mu & 1+4\mu \end{pmatrix}.$$

Again $\mathbf{u}^0$ is defined by (3.8). The $M^2 \times M^2$ matrix A_I is block Toeplitz, symmetric, block symmetric, block tridiagonal, strictly diagonally dominant and positive definite.

The numerical solution of (3.1)-(3.3) using this implicit scheme involves solving the matrix system (3.15) for $k = 0, 1, 2, ..., N-1$. This implicit scheme is of order $\mathcal{O}(\Delta t + (\Delta x)^2)$ and is unconditionally stable. However, it represents a system of M^2 equations in M^2 unknowns for each time step. For $\Delta x = 0.01$ we must solve a system of 9801 equations in 9801 unknowns at each time step. Once again, this represents an enormous computational cost.

The Crank-Nicolson Scheme

If we approximate the partial derivatives in (3.1) using the average of a centered difference at the k-th time step and a centered difference at the $(k+1)$-st time step we obtain the Crank-Nicolson scheme given by

$$\begin{aligned} \frac{u_{ij}^{k+1} - u_{ij}^{k}}{\Delta t} &= \frac{1}{2}\left(\frac{u_{i-1j}^{k+1} - 2u_{ij}^{k+1} + u_{i+1j}^{k+1}}{(\Delta x)^2} + \frac{u_{ij-1}^{k+1} - 2u_{ij}^{k+1} + u_{ij+1}^{k+1}}{(\Delta y)^2} \right. \\ &+ \left. \frac{u_{i-1j}^{k} - 2u_{ij}^{k} + u_{i+1j}^{k}}{(\Delta x)^2} + \frac{u_{ij-1}^{k} - 2u_{ij}^{k} + u_{ij+1}^{k}}{(\Delta y)^2} \right). \end{aligned} \quad (3.16)$$

Rearranging terms we obtain (for $i, j = 1, ..., M$)

$$\begin{aligned}(1+2\mu)u_{ij}^{k+1} - \frac{\mu}{2}\left(u_{i-1j}^{k+1} + u_{ij-1}^{k+1} + u_{i+1j}^{k+1} + u_{ij+1}^{k+1}\right) \qquad (3.17)\\ = (1-2\mu)u_{ij}^{k} + \frac{\mu}{2}\left(u_{i-1j}^{k} + u_{ij-1}^{k} + u_{i+1j}^{k} + u_{ij+1}^{k}\right).\end{aligned}$$

The matrix form of the Crank-Nicolson scheme for the parabolic problem (3.1)-(3.3) is then given by

$$A_C\mathbf{u}^{k+1} = B_C\mathbf{u}^k. \qquad (3.18)$$

The vector $\mathbf{u}^k$ is defined in (3.10) and the matrices A and B are given by

$$A_C = \begin{pmatrix} R & -\frac{\mu}{2}I & O & \cdots & O \\ -\frac{\mu}{2}I & \ddots & \ddots & \ddots & \vdots \\ O & \ddots & \ddots & \ddots & O \\ \vdots & \ddots & \ddots & \ddots & -\frac{\mu}{2}I \\ O & \cdots & O & -\frac{\mu}{2}I & R \end{pmatrix}$$

and

$$B_C = \begin{pmatrix} S & \frac{\mu}{2}I & O & \cdots & O \\ \frac{\mu}{2}I & \ddots & \ddots & \ddots & \vdots \\ O & \ddots & \ddots & \ddots & O \\ \vdots & \ddots & \ddots & \ddots & \frac{\mu}{2}I \\ O & \cdots & O & \frac{\mu}{2}I & S \end{pmatrix},$$

where R and S are the $M \times M$ TST matrices given by

$$R = \begin{pmatrix} 1+2\mu & -\frac{\mu}{2} & 0 & \cdots & 0 \\ -\frac{\mu}{2} & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -\frac{\mu}{2} \\ 0 & \cdots & 0 & -\frac{\mu}{2} & 1+2\mu \end{pmatrix}$$

and

$$S = \begin{pmatrix} 1-2\mu & \frac{\mu}{2} & 0 & \cdots & 0 \\ \frac{\mu}{2} & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \frac{\mu}{2} \\ 0 & \cdots & 0 & \frac{\mu}{2} & 1-2\mu \end{pmatrix}.$$

Again $\mathbf{u}^0$ is defined by (3.8).

The Crank-Nicolson scheme is of the higher order $\mathcal{O}((\Delta t)^2 + (\Delta x)^2)$ and is unconditionally stable. However, it still requires the solution of M^2 equations in M^2 unknowns along with an additional matrix-vector multiplication at each time step. Next we consider the classical Peaceman-Rachford Alternating-Direction Implicit (ADI) scheme and apply it to approximating the solution of (3.1)-(3.3).

An ADI Scheme for Parabolic Problems

In 1955, Donald Peaceman and Henry Rachford, Jr., [20], published a new scheme that employed the idea of splitting each time step into two half-steps, and combined both implicit and explicit schemes on each half-step. This Alternating-Direction Implicit (ADI) scheme is a perturbation (or factorization) of the Crank-Nicolson scheme. Let v_{ij}^{2k} represent u_{ij}^{k} and let v_{ij}^{2k+1} represent intermediate values of the dependent variable used only as an aid to obtaining the values at the new time level.

Construct the classical Peaceman-Rachford ADI scheme as

$$\frac{v_{ij}^{2k+1} - v_{ij}^{2k}}{(\Delta t/2)} = \frac{v_{i-1j}^{2k+1} - 2v_{ij}^{2k+1} + v_{i+1j}^{2k+1}}{(\Delta x)^2} + \frac{v_{ij-1}^{2k} - 2v_{ij}^{2k} + v_{ij+1}^{2k}}{(\Delta y)^2} \tag{3.19}$$

$$\frac{v_{ij}^{2k+2} - v_{ij}^{2k+1}}{(\Delta t/2)} = \frac{v_{i-1j}^{2k+1} - 2v_{ij}^{2k+1} + v_{i+1j}^{2k+1}}{(\Delta x)^2} + \frac{v_{ij-1}^{2k+2} - 2v_{ij}^{2k+2} + v_{ij+1}^{2k+2}}{(\Delta y)^2}. \tag{3.20}$$

Assuming, as before, that $\Delta x = \Delta y$, (3.19) and (3.20) may be rearranged to form the system

$$-v_{i-1j}^{2k+1} + (2 + \frac{2}{\mu})v_{ij}^{2k+1} - v_{i+1j}^{2k+1} = v_{ij-1}^{2k} - (2 - \frac{2}{\mu})v_{ij}^{2k} + v_{ij+1}^{2k} \tag{3.21}$$

$$-v_{ij-1}^{2k+2} + (2 + \frac{2}{\mu})v_{ij}^{2k+2} - v_{ij+1}^{2k+2} = v_{i-1j}^{2k+1} - (2 - \frac{2}{\mu})v_{ij}^{2k+1} + v_{i+1j}^{2k+1} \tag{3.22}$$

which is unconditionally stable even though each of (3.21) and (3.22) is separately unstable. These equations may be written as a system of matrix equations

$$\begin{aligned} A_A\mathbf{v}^{2k+1} &= \mathbf{b}^{2k} \\ A_A\mathbf{w}^{2k+2} &= \mathbf{c}^{2k+1} \end{aligned} \tag{3.23}$$

where the vectors $\mathbf{v}$ (corresponding to grid points ordered left-to-right, bottom-to-top) and $\mathbf{w}$ (corresponding to grid points ordered bottom-to-top, left-to-right) are defined by

$$\mathbf{v}^{2k+1} = [v_{11}^{2k+1}, v_{21}^{2k+1}, ..., v_{M\,1}^{2k+1}, v_{12}^{2k+1}, v_{22}^{2k+1}, ..., v_{M\,2}^{2k+1}, ..., v_{1\,M}^{2k+1}, ..., v_{M\,M}^{2k+1}]^T \tag{3.24}$$

and

$$\mathbf{w}^{2k+2} = [v_{11}^{2k+2}, v_{12}^{2k+2}, ..., v_{1\,M}^{2k+2}, v_{21}^{2k+2}, v_{22}^{2k+2}, ..., v_{2\,M}^{2k+2}, ..., v_{M\,1}^{2k+2}, ..., v_{M\,M}^{2k+2}]^T, \tag{3.25}$$

the vector

$$\mathbf{b}^{2k} = [b_1, ..., b_{M^2}] \tag{3.26}$$

satisfies

$$b_{(j-1)M+i} = v_{ij-1}^{2k} - (2 - \frac{2}{\mu})v_{ij}^{2k} + v_{ij+1}^{2k} \tag{3.27}$$

where for each $j \in \{1, ..., M\}$, i runs from 1 to M, and the vector

$$\mathbf{c}^{2k+1} = [c_1, ..., c_{M^2}] \tag{3.28}$$

satisfies

$$c_{(i-1)M+j} = v_{i-1j}^{2k+1} - (2 - \frac{2}{\mu})v_{ij}^{2k+1} + v_{i+1j}^{2k+1} \tag{3.29}$$

where for each $i \in \{1, ..., M\}$, j runs from 1 to M. Furthermore, if T is the $M \times M$, TST, strictly diagonally dominant, positive definite matrix

$$T = \begin{pmatrix} 2+\frac{2}{\mu} & -1 & 0 & \cdots & 0 \\ -1 & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \cdots & 0 & -1 & 2+\frac{2}{\mu} \end{pmatrix},$$

then A_A is the $M^2 \times M^2$ tridiagonal matrix given by

$$A_A = \begin{pmatrix} T & O & \cdots & O \\ O & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & O \\ O & \cdots & O & T \end{pmatrix}. \tag{3.30}$$

The system (3.23) is of order $\mathcal{O}((\Delta t)^2+(\Delta x)^2)$. It can be solved more cheaply than the explicit scheme because there is no stability constraint and it can be solved more efficiently than both the implicit scheme and the Crank-Nicolson scheme through the use of the Thomas algorithm for tridiagonal matrices. The system may be solved even more efficiently by first solving the M matrix equations (each only $M \times M$) given by

$$T\mathbf{v}_j^{2k+1} = \mathbf{b}_j^{2k},$$

where

$$\mathbf{v}_j^{2k+1} = [v_{1j}^{2k+1}, ..., v_{Mj}^{2k+1}]$$

and

$$\mathbf{b}_j^{2k} = [b_{(j-1)M+1}^{2k}, ..., b_{(j-1)M+M}^{2k}]$$

for $j = 1, ..., M$ ($b_{(j-1)M+i}$ is defined in (3.27)) and then solving the similarly defined M matrix equations (each only $M \times M$)

$$T\mathbf{w}_i^{2k+2} = \mathbf{c}_i^{2k+1},$$

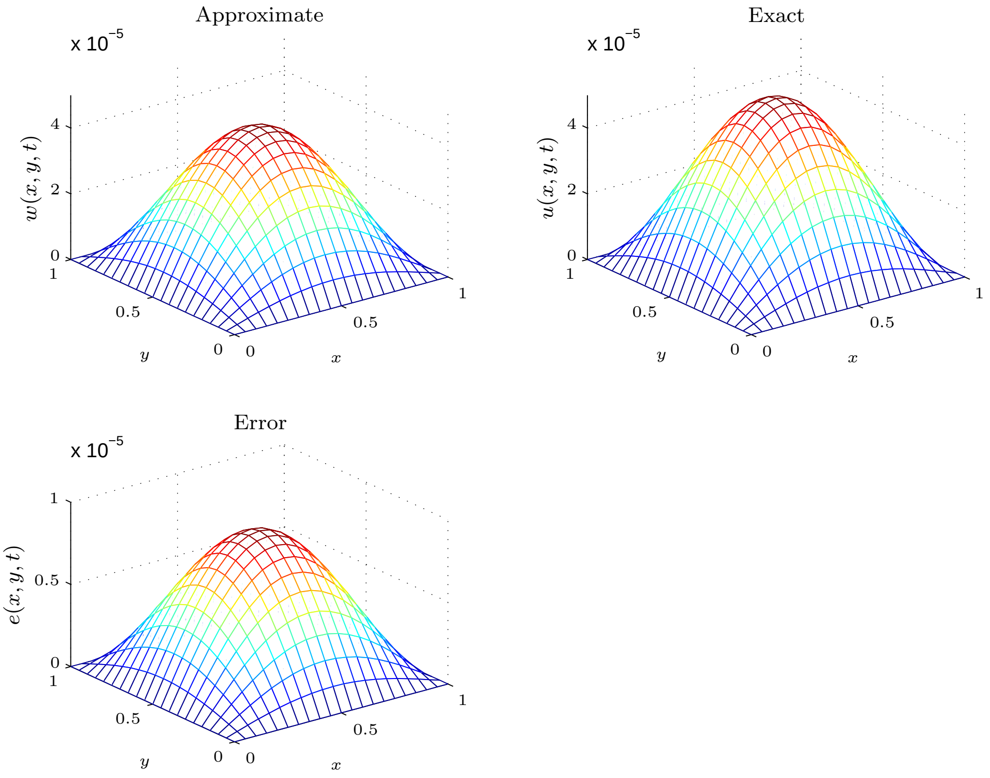
where

$$\mathbf{w}_i^{2k+2} = [v_{i1}^{2k+2}, ..., v_{iM}^{2k+2}]$$

and

$$\mathbf{c}_i^{2k+1} = [c_{(i-1)M+1}^{2k+1}, ..., c_{(i-1)M+M}^{2k+1}]$$

for $i = 1, ..., M$ ($c_{(i-1)M+j}$ is defined in (3.29)). A graph of the ADI solution to (3.1)-(3.3) at the final time step, $t = \frac{1}{2}$, is given in Figure 6 below, along with graphs of the exact solution and the error.


Approximate

Exact

Error

Figure 6. 3D plots of the approximate solution, $w(x,y,t)$, the exact solution, $u(x,y,t)$ and the error, $e(x,y,t)$ for the parabolic problem (3.1)-(3.3) for $\Delta x = \Delta y = \Delta t = \frac{1}{20}$ at $t = \frac{1}{2}$. Note that for this example $||w - u||_\infty = 8.7939 \times 10^{-6}$.

An ADI Scheme for Elliptic Problems

It is noted in [27] that the alternating-direction implicit (ADI) method can be applied to steady-state (elliptic) problems by using the first estimate as the initial condition, considering iterations instead of time steps, and replacing the increment ratio $\mu \equiv (\Delta x)^2/(\Delta t)$ with an iteration parameter. To be specific, we apply the Peaceman-Rachford alternating-direction scheme to approximate the solution to the following elliptic problem.

Example 22. *Consider the following problem on the unit square.*

$$-\triangle u \equiv -\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right) = f(x,y), \qquad (x,y)\in\Omega=(0,1)\times(0,1), \quad (3.31)$$

with Dirichlet boundary conditions

$$\begin{aligned} u(0,y) &= 0, \quad y\in(0,1)\\ u(1,y) &= 0, \quad y\in(0,1)\\ u(x,0) &= 0, \quad x\in(0,1)\\ u(x,1) &= 0. \quad x\in(0,1) \end{aligned} \qquad (3.32)$$

Assume the source term is

$$f(x,y)=3xye^{x+y}[(x+3)(1-y)+(y+3)(1-x)], \quad (x,y)\in\Omega \qquad (3.33)$$

and note for reference that the exact solution is given by

$$u(x,y)=3e^{x+y}xy(1-x)(1-y). \qquad (3.34)$$

Letting v_{ij}^{2k} represent u_{ij}^{k} and assuming $\Delta x = \Delta y$ as before, we construct the Peaceman-Rachford ADI scheme for the given problem as

$$-v_{i-1j}^{2k+1}+(2+\frac{1}{r})v_{ij}^{2k+1}-v_{i+1j}^{2k+1}=v_{ij-1}^{2k}-(2-\frac{1}{r})v_{ij}^{2k}+v_{ij+1}^{2k}+(\Delta x)^2 f_{ij}$$

$$-v_{ij-1}^{2k+2}+(2+\frac{1}{r})v_{ij}^{2k+2}-v_{ij+1}^{2k+2}=v_{i-1j}^{2k+1}-(2-\frac{1}{r})v_{ij}^{2k+1}+v_{i+1j}^{2k+1}+(\Delta x)^2 f_{ij}.$$

The system above is analogous to the system given in (3.21)-(3.22). The iteration parameter $r > 0$ is the analog of $\mu/2$ dictated by the choice of time step Δt in an initial-value problem. In this context, k represents successive iterates rather than successive time steps and $f_{ij} = f(x_i, y_j)$. This scheme is of order $\mathcal{O}((\Delta x)^2)$. The system above may be written as the system of matrix equations

$$\begin{aligned} A_A \mathbf{v}^{2k+1} &= \mathbf{b}^{2k} + (\Delta x)^2 \mathbf{f} \\ A_A \mathbf{w}^{2k+2} &= \mathbf{c}^{2k+1} + (\Delta x)^2 \mathbf{g}, \end{aligned} \tag{3.35}$$

which is analogous to the system given in (3.23). The vectors $\mathbf{v}^{2k+1}, \mathbf{w}^{2k+2}, \mathbf{b}^{2k}$, and $\mathbf{c}^{2k+1}$ are defined in (3.24)-(3.29) and the vectors $\mathbf{f}$ and $\mathbf{g}$ are defined by

$$\mathbf{f} = [f_{11}, f_{21}, ..., f_{M\ 1}, f_{12}, f_{22}, ..., f_{M\ 2}, ..., f_{1\ M}, f_{2\ M}, ..., f_{M\ M}]^T \tag{3.36}$$

$$\mathbf{g} = [f_{11}, f_{12}, ..., f_{1\ M}, f_{21}, f_{22}, ..., f_{2\ M}, ..., f_{M\ 1}, f_{M\ 2}, ..., f_{M\ M}]^T. \tag{3.37}$$

To solve this system we define the $M \times M$, TST, strictly diagonally dominant, positive definite matrix

$$T = \begin{pmatrix} 2+\frac{1}{r} & -1 & 0 & \cdots & 0 \\ -1 & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \cdots & 0 & -1 & 2+\frac{1}{r} \end{pmatrix},$$

and note that the system matrix A_A has the same form as the $M^2 \times M^2$, tridiagonal matrix given in (3.30).

Before solving the resulting system we discuss the problem of computing iteration parameters. We note from [21, p. 662] that the determination of iteration parameters for an arbitrary elliptic problem is an unsolved problem. However, in special cases, such as for the *continuous* ADI *model problem* given in (3.31)-(3.32), we can find an optimal sequence of iteration parameters using the following procedure.

Let $M-1$ denote the number of interior mesh points in the row or column of greatest length. Define $a = 4\sin^2(\frac{\pi}{2M})$, $b = 4\cos^2(\frac{\pi}{2M})$, and let $\sigma = a/b$. Suppose $z = \sigma^{\frac{1}{m-1}}$ where m is still to be determined. In 1961, D. Young, [31], discussed a variation of a procedure due to E. Wachspress specifying that after m "double sweeps" (iterations) the factor of reduction of the error is approximately

$$P_m = \left(\frac{1 - z^{1/2}}{1 + z^{1/2}} e^{-z^{3/2}/(1-z)} \right)^4 .$$

Since the average factor of reduction per double sweep is defined to be $S_m = (P_m)^{1/m}$, we choose m so as to minimize S_m by evaluating S_m for several values of m. Finally the iteration parameters are chosen to be $r_k = 1/(bz^{k-1})$ for $k = 1, 2, ..., m$, and we use them in the cyclic order $r_1, r_2, ..., r_m, r_1, r_2, ...$.

For the elliptic problem given in (3.31)-(3.33), we choose $M = 20$ so that $a = 4\sin^2(\frac{\pi}{40}) \approx 0.02462$, $b = 4\cos^2(\frac{\pi}{40}) \approx 3.97538$, and $\sigma = a/b \approx 0.00619$. Computing S_m for several values of m we obtain

$$S_6 \approx 0.3152, \quad S_7 \approx 0.3086, \quad S_8 \approx 0.3060, \quad S_9 \approx 0.3056, \quad \text{and} \quad S_{10} \approx 0.3064.$$

The minimum value of S_m is taken as 0.3056, which occurs for $m = 9$. It follows that $z = \sigma^{\frac{1}{m-1}} = \sigma^{\frac{1}{8}} \approx 0.52969$. The corresponding iteration parameters are given specifically by

$$r_1 \approx 0.25155, \quad r_2 \approx 0.47493, \quad r_3 \approx 0.89666,$$

$$r_4 \approx 1.69291, \quad r_5 \approx 3.19623, \quad r_6 \approx 6.03450,$$

$$r_7 \approx 11.39319, \quad r_8 \approx 21.51045, \quad \text{and} \quad r_9 \approx 40.61191.$$

Below in Table 19 we display the results obtained using the ADI scheme above in (3.35) and three different sets of iteration parameters given by (1) $r_1,\ r_2,\ \ldots, r_9$, (2) $r_1 = 0.25$ and $r_{k+1} = 2r_k$ for $k = 1, 2, ..., 8$, and (3) $r_k = 17$ for $k = 1, 2, ..., 9$. Note the excellent convergence obtained with the third choice of a constant iteration parameter of $r_k = 17$. This result motivated the choice of a constant iteration parameter for the Alternating-Direction Sinc-Galerkin (ADSG) scheme introduced in Chapter 4 and thus greatly facilitated the solution of the resulting matrix systems. A graph of the ADI solution to the elliptic problem in (3.31)-(3.33) obtained after nine iterations using the constant iteration parameter $r_k = 17$ is also given in Figure 7, along with graphs of the exact solution and the error. The column headed *iter* gives the number of complete double sweeps (iterations) and the columns headed *par* give the parameters r_k used each iteration. The maximum absolute error over all grid points is denoted by $||u - w||_\infty$.

Table 19. Errors in the ADI solution of (3.31)-(3.33) for three sets of parameters.

iter	*par*	$\|\|u-w\|\|_\infty$	*par*	$\|\|u-w\|\|_\infty$	*par*	$\|\|u-w\|\|_\infty$
1	0.25155	5.608e-001	0.25	5.608e-001	17	1.069e-001
2	0.47492	5.364e-001	0.50	5.352e-001	17	1.416e-002
3	0.89666	4.913e-001	1.00	4.851e-001	17	3.963e-003
4	1.69291	4.122e-001	2.00	3.954e-001	17	1.033e-003
5	3.19623	2.998e-001	4.00	2.642e-001	17	8.187e-004
6	6.03450	1.630e-001	8.00	1.183e-001	17	7.009e-004
7	11.39320	5.181e-002	16.00	2.285e-002	17	7.000e-004
8	21.51045	5.512e-003	32.00	1.002e-003	17	6.953e-004
9	40.61191	6.952e-004	64.00	7.105e-004	17	6.953e-004

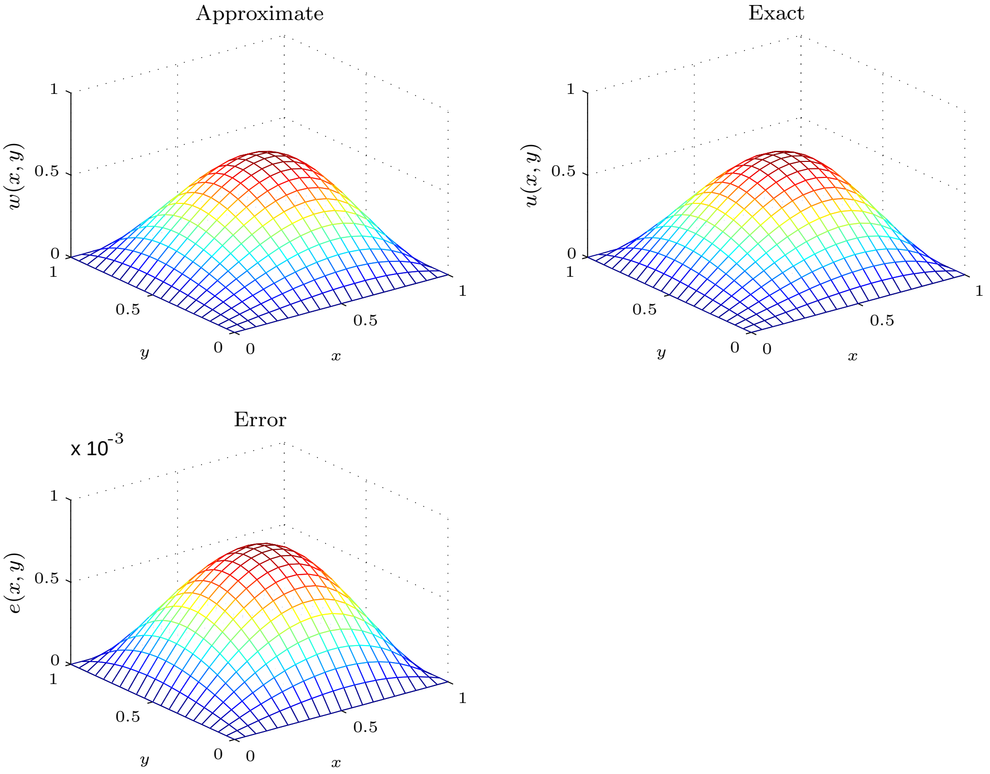


Figure 7. 3D plots of the approximate and exact solutions and the error for the elliptic problem (3.31)-(3.33) in Example 22, solved using the constant parameter $r_k = 17$.

CHAPTER 4

AN ALTERNATING-DIRECTION SINC-GALERKIN METHOD

Introduction

The Galerkin method is a variational method for approximating the solution of differential equations based on the weak form of the differential equation. Let $\{\chi_i\}_{i=1}^{\infty}$ be a dense linearly independent set in a Hilbert space H and suppose that H^N=span $\{\chi_1, ..., \chi_N\}$. In H^N, define the approximation

$$u_N(x) = \sum_{k=1}^{N} c_k \chi_k(x)$$

to the solution $u \in H$ of the differential equation $Lu = f$ where the coefficients $\{c_k\}_{k=1}^{N}$ are to be determined. Galerkin's idea is to require the *residual*

$$R_N(x) \equiv (Lu_N - f)(x)$$

to be orthogonal to each $\chi_j \in H^N$ with respect to the inner product on H.

The Sinc-Galerkin method, employed throughout Chapter 2 of this thesis, used a linearly independent set of sinc basis functions to construct the approximations to the solutions of a variety of differential equations. Alternating-Direction Galerkin (ADG) methods were first introduced by Jim Douglas, Jr. and Todd Dupont in 1971. The Douglas-Dupont ADG method is described in detail in [3]. Below we develop an alternating-direction Galerkin method that differs significantly from the Douglas-Dupont ADG method. By applying alternating-direction iteration to the discrete system in (2.102) we obtain the Alternating-Direction Sinc-Galerkin (ADSG) method for the Sylvester equation that is the main goal of this thesis.

An Alternating-Direction Scheme for a Linear System

As described in [8], alternating-direction iteration may be used to solve the linear system

$$M\mathbf{y} = \mathbf{b} \tag{4.1}$$

when the matrix M is symmetric positive definite (SPD), M can be expressed as a sum of SPD matrices H and V, $M = H + V$, and numerical implementation of the following algorithm

$$\begin{aligned} \mathbf{y}_0 &= 0, \\ (H + p_j I)\mathbf{y}_{j-\frac{1}{2}} &= \mathbf{b} - (V - p_j I)\mathbf{y}_{j-1}, \\ (V + q_j I)\mathbf{y}_j &= \mathbf{b} - (H - q_j I)\mathbf{y}_{j-\frac{1}{2}}, \quad j = 1, 2, ..., J \end{aligned} \tag{4.2}$$

is efficient. When the eigenvalues of M are in the positive-real half-plane (which is certainly the case for an SPD matrix M) and, in addition, H and V commute, the system (4.1) is classified as a *discrete* ADI *model problem* and (4.2) is very efficient. (See [28].)

While the method in (4.2) shows that different iteration parameters p_j and q_j may be used in each iteration, the original Peaceman-Rachford method used $q_j = p_j$ and, according to [8], this choice is now known to be appropriate when H and V have the same eigenvalue bounds. Now if we suppose that this common eigenvalue interval is $[a, b]$ with $\frac{a}{b} \ll 1$, then a relative error norm of $\frac{||Y_j - Y||}{||Y||} \leq \epsilon$ is attained when

$$J > \frac{1}{\pi^2} \ln\left(\frac{\epsilon}{4}\right) \ln\left(\frac{a}{4b}\right) \tag{4.3}$$

with iteration parameters

$$p_j = b \,\mathrm{dn}\left[\frac{2j-1}{2J}K, \sqrt{1 - \frac{a^2}{b^2}}\right], \tag{4.4}$$

where $\mathrm{dn}(z, k)$ is the Jacobian elliptic function of argument z and modulus k, and K is the complete elliptic integral, $K(k)$. (See [8].)

For our purposes we focus on the Sylvester equation

$$AX + XB = C. \tag{4.5}$$

We note once again that when $B = A^T$ the resulting linear system,

$$AX + XA^T = C, \tag{4.6}$$

is referred to as a Lyapunov equation. The relevance of ADI iteration to the solution of (4.6) was revealed in [5], in 1984, when it was observed that (4.6) represents a discrete ADI model problem. Below we establish a slightly more general result. We show that sometimes even if $B \neq A^T$, indeed even if A and B do not have the same size, we may still have that (4.5) represents an ADI model problem.

For the $m_x \times m_x$ matrix A, the $m_y \times m_y$ matrix B, and the $m_x \times m_y$ matrices X and C, the Sylvester equation (4.5) is equivalent to

$$(I_{m_y}^{(0)} \otimes A + B^T \otimes I_{m_x}^{(0)}) co(X) = co(C). \tag{4.7}$$

If $H = I_{m_y}^{(0)} \otimes A$ and $V = B^T \otimes I_{m_x}^{(0)}$, then by the 'mixed product rule' [4, p. 25], we have that $HV = (I_{m_y}^{(0)} \otimes A)(B^T \otimes I_{m_x}^{(0)}) = B^T \otimes A = (B^T \otimes I_{m_x}^{(0)})(I_{m_y}^{(0)} \otimes A) = VH$. Thus H and V commute. If A and B are SPD then H and V are SPD and since $H+V$ is the Kronecker sum of A and B it follows that $H+V$ is also SPD. Thus (4.7) represents a discrete ADI model problem. We conclude that if A and B are SPD then the Sylvester equation in (4.5) represents a discrete ADI model problem.

The Alternating-Direction Sinc-Galerkin (ADSG) method proposed below is an iterative scheme that may be used to solve the Sylvester equation that results when one employs a symmetric Sinc-Galerkin method to solve the Poisson equation (2.99). In that case, the $m_x \times m_x$ matrix A, the $m_y \times m_y$ matrix B, and the $m_x \times m_y$ matrices C and X are given by

$$\begin{aligned}
A &= D_{m_x}(\phi_x') \left[-\frac{1}{h_x^2} I_{m_x}^{(2)} + D_{m_x}\left(\frac{-1}{(\phi_x')^{3/2}} \left(\frac{1}{(\phi_x')^{1/2}} \right)'' \right) \right] D_{m_x}(\phi_x') \\
B &= D_{m_y}(\phi_y') \left[-\frac{1}{h_y^2} I_{m_y}^{(2)} + D_{m_y}\left(\frac{-1}{(\phi_y')^{3/2}} \left(\frac{1}{(\phi_y')^{1/2}} \right)'' \right) \right] D_{m_y}(\phi_y')
\end{aligned} \tag{4.8}$$

$$\begin{aligned}
C &= D_{m_x}((\phi_x')^{-1/2})\, F\, D_{m_y}((\phi_y')^{-1/2}) \\
X &= D_{m_x}((\phi_x')^{-1/2})\, U\, D_{m_y}((\phi_y')^{-1/2})
\end{aligned}$$

using the definitions specified in Chapter 2. We note that since the matrices $I_{m_x}^{(2)}$ and $I_{m_y}^{(2)}$ are symmetric then the matrices A and B are symmetric. Moreover, we learn in [11, pp. 216-218] that the matrices A and B given in (4.8) are also positive definite. It follows that when a symmetric Sinc-Galerkin method is used to solve a Poisson problem posed on any finite domain, the resulting Sylvester equation is a discrete ADI model problem. The ADSG method below results from a direct application of the ADI equations in (4.2) to the discrete system in (4.7).

An Alternating-Direction Sinc-Galerkin Method

In [4, p.25] we find that for the $m \times n$ matrix P, the $n \times r$ matrix Q, and the $r \times s$ matrix R we have

$$co(PQR) = (R^T \otimes P)co(Q). \tag{4.9}$$

Applying (4.2) to the discrete system in (4.7) yields

$$\left[I_{m_y}^{(0)} \otimes A + p_j I_{m_y}^{(0)} \otimes I_{m_x}^{(0)}\right] co(X_{j-\frac{1}{2}}) = co(C) - \left[B^T \otimes I_{m_x}^{(0)} - p_j I_{m_y}^{(0)} \otimes I_{m_x}^{(0)}\right] co(X_{j-1})$$

and

$$\left[B^T \otimes I_{m_x}^{(0)} + q_j I_{m_y}^{(0)} \otimes I_{m_x}^{(0)}\right] co(X_j) = co(C) - \left[I_{m_y}^{(0)} \otimes A - q_j I_{m_y}^{(0)} \otimes I_{m_x}^{(0)}\right] co(X_{j-\frac{1}{2}}).$$

Distributing the column vectors and applying (4.9) we find that

$$co(AX_{j-\frac{1}{2}}) + p_j co(X_{j-\frac{1}{2}}) = co(C) - [co(X_{j-1}B) - p_j co(X_{j-1})]$$

and

$$co(X_jB) + q_jco(X_j) = co(C) - \left[co(AX_{j-\frac{1}{2}}) - q_jco(X_{j-\frac{1}{2}})\right].$$

It follows that

$$AX_{j-\frac{1}{2}} + p_jX_{j-\frac{1}{2}} = C - (X_{j-1}B - p_jX_{j-1})$$

and

$$X_jB + q_jX_j = C - (AX_{j-\frac{1}{2}} - q_jX_{j-\frac{1}{2}}).$$

Factoring the common factors and rewriting both equations, taking the transpose of the second, we conclude that

$$(A + p_jI_{m_x}^{(0)})X_{j-\frac{1}{2}} = C - X_{j-1}(B - p_jI_{m_y}^{(0)})$$

and

$$(B^T + q_jI_{m_y}^{(0)})X_j^T = C^T - X_{j-\frac{1}{2}}^T(A^T - q_jI_{m_x}^{(0)}).$$

We construct an ADSG iterative scheme for solving (4.5) as follows: Assume the $m_x \times m_x$ matrix A, the $m_y \times m_y$ matrix B, and the $m_x \times m_y$ matrix C are given by (4.8). Let X_0 be the $m_x \times m_y$ zero matrix and compute X_j for $j = 1, 2, ...$, using $X_{j-\frac{1}{2}}$ as temporary place holders, i.e., let

$$X_0 = O, \tag{4.10}$$

and solve successively

$$(A + p_j I_{m_x}^{(0)}) X_{j-\frac{1}{2}} = C - X_{j-1}(B - p_j I_{m_y}^{(0)})$$

$$(B^T + q_j I_{m_y}^{(0)}) X_j^T = C^T - X_{j-\frac{1}{2}}^T (A^T - q_j I_{m_x}^{(0)}). \tag{4.11}$$

where the positive real numbers p_j and q_j may be the typical iteration parameters associated with the ADI method. In order to avoid the often costly computation of eigenvalues, however, we include a simplifying stipulation in this ADSG scheme, i.e., based upon numerical experiments, we use a constant iteration parameter $p \equiv p_j = q_j = 17$ for all j. This constant was derived numerically using a variety of test problems. We solve each system in (4.11) via Gaussian elimination.

Since we do not compute spectral bounds for the matrices A and B, we do not use the typical iteration parameters associated with ADI. We stop iterating when either

(1) the maximum norm of the difference between consecutive iterates from X_{j-1} to X_j becomes smaller than a specified tolerance, which we take as $c(M_x) M_x^4 e^{-\pi\sqrt{2\alpha_s M_x}}$, or

(2) the number of iterations exceeds $\dfrac{m_x^2 m_y^2}{7(m_x + m_y)}$.

Note that M_x is an arbitrary positive integer used in the construction of the matrices A and B. Based on numerical experiments, if $M_x = 2^k$ we use $c(M_x) = 100^{1-k}$. For simplicity, we also use $c(M_x) = 100^{1-k}$ in the examples in which $M_x = 6k$ below. Recalling the equation for X in (2.103) we note that the approximate solution, U, to the given elliptic problem may be computed using the final iterate X_{f} as

$$U = D_{m_x}((\phi_x')^{1/2})X_{\mathrm{f}}D_{m_y}((\phi_y')^{1/2}).$$

Examples on the Unit Square

We now revisit the examples on the unit square first presented in Chapter 2. We recall that in each case we chose parameters α_s, β_s, ζ_s and η_s based on the asymptotic behavior of the exact solution. We then chose M_x arbitrarily and computed N_x, M_y and N_y as follows: $N_x = \frac{\alpha_s}{\beta_s}M_x$ if $\frac{\alpha_s}{\beta_s}M_x$ is an integer and $N_x = \left[\left|\frac{\alpha_s}{\beta_s}M_x + 1\right|\right]$ otherwise; $M_y = \frac{\alpha_s}{\zeta_s}M_x$ if $\frac{\alpha_s}{\zeta_s}M_x$ is an integer and $M_y = \left[\left|\frac{\alpha_s}{\zeta_s}M_x + 1\right|\right]$ otherwise; $N_y = \frac{\alpha_s}{\eta_s}M_x$ if $\frac{\alpha_s}{\eta_s}M_x$ is an integer and $N_y = \left[\left|\frac{\alpha_s}{\eta_s}M_x + 1\right|\right]$ otherwise. We let $m_x = M_x + N_x + 1$ and $m_y = M_y + N_y + 1$ and for $d = \pi/2$, defined $h_s = \sqrt{\frac{\pi d}{\alpha_s M_x}}$. Since the problems were posed on the unit square, we employed the mapping $\phi_z(z) = \ln\left(\frac{z}{1-z}\right)$ for $z = x$ or y. In each case the resulting matrix equation was a discrete ADI model problem expressed as a Sylvester equation or a Lyapunov equation. As we did in Chapter 2, we report the maximum errors over the set of sinc grid points

$$S = \{x_i\}_{i=-M}^{M} \times \{y_j\}_{j=-M}^{M} \tag{4.12}$$

and the set of uniform gridpoints

$$U = \{w_i\}_{i=0}^{100} \times \{z_j\}_{j=0}^{100} \tag{4.13}$$

as

$$||ES^s_{ADSG}(h_s)|| = \max_{\substack{-M\le i\le M \\ -M\le j\le M}} |u(x_i, y_j) - u^s_{ij}|$$

and

$$||EU^s_{ADSG}(h_u)|| = \max_{\substack{0\le i\le 100\\ 0\le j\le 100}} |u(w_i, z_j) - u^s_{m,m}(w_i, z_j)|,$$

respectively. For $h_u = 0.01$ we define $w_i = ih_u$, for $1 \le i \le 100$ and $z_j = jh_u$, for $1 \le j \le 100$.

Example 23. *Revisit the elliptic problem in* Example 15 *which we rewrite below for reference,*

$$\begin{aligned} -\triangle u(x,y) &= f(x,y), \quad (x,y) \in \Omega = (0,1)\times(0,1) \\ u(x,y) &= 0, \quad (x,y) \in \partial\Omega \end{aligned} \tag{4.14}$$

where $f(x,y) = -2x^2(1-x)^2(1-6y+6y^2) - 2y^2(1-y)^2(1-6x+6x^2)$ *and the exact solution is given by* $u(x,y) = x^2y^2(1-x)^2(1-y)^2$. *As in* Example 15 *we choose the parameter values* $\alpha_s = \beta_s = \zeta_s = \eta_s = 3/2$ *so that for* $d = \pi/2$ *and* $M \equiv M_x = N_x = M_y = N_y$, *we have* $h_s = \dfrac{\pi}{\sqrt{3M}}$ *and* $m \equiv m_x = m_y = 2M+1$. *Now the* $m \times m$ *matrix* A *may be computed using* (4.8)*. Since the nodes, given by* $x_i = e^{ih_s}/(1+e^{ih_s})$ *and* $y_j = e^{jh_s}/(1+e^{jh_s})$, *are the same, then* $A = B$ *in the Sylvester equation* (4.5)*. Since the matrix* A *(and similarly* B*) is symmetric, then the resulting discrete system for the elliptic problem* (4.14) *is a Lyapunov equation. Solving this equation using the Alternating-Direction Sinc-Galerkin method* (4.10)-(4.11)*, generates the errors reported in* Table 20*. The number of complete iterations* *used to solve the system for each value of M is given in the column headed* Iter. *We note that the uniform errors here are of the same order of magnitude as those obtained in* Example 15 *(where the Kronecker product, Kronecker sum, concatenation operator, and Gaussian elimination were employed) and reported in* Table 13.

Table 20. Errors in the ADSG solution of (4.14).

M	m	m^2	h_s	$\|\|ES^s_{ADSG}(h_s)\|\|$	$\|\|EU^s_{ADSG}(h_u)\|\|$	Iter
2	5	25	1.28255	2.861e-005	9.752e-005	2
4	9	81	0.90690	1.534e-005	1.647e-005	3
8	17	289	0.64127	7.032e-007	8.829e-007	15
16	33	1089	0.45345	1.400e-008	1.368e-008	110
32	65	4225	0.32064	4.363e-011	4.207e-011	1914

Example 24. *Revisit the elliptic problem in* Example 16 *which we rewrite below for reference,*

$$\begin{aligned} -\triangle u(x,y) &= -\left(\frac{x\ln(x)}{y}+\frac{y\ln(y)}{x}\right), \quad (x,y)\in\Omega=(0,1)\times(0,1) \\ u(x,y) &= 0, \quad (x,y)\in\partial\Omega \end{aligned} \tag{4.15}$$

where the exact solution is $u(x,y)=xy\ln(x)\ln(y)$. As in Example 16 *we choose the parameter values $\alpha_s=\beta_s=\zeta_s=\eta_s=1/2$ so that for $d=\pi/2$ and $M\equiv M_x=N_x=M_y=N_y$, we have $h_s=\dfrac{\pi}{\sqrt{M}}$ and $m=2M+1$. Again we note that the nodes, $x_i=e^{ih_s}/(1+e^{ih_s})$ and $y_j=e^{jh_s}/(1+e^{jh_s})$, are the same so A=B. Since A is also symmetric, the resulting discrete system for the elliptic problem* (4.15) *is a Lyapunov* *equation. Solving that equation using the Alternating-Direction Sinc-Galerkin method* (4.10)-(4.11), *we obtain the errors in* Table 21 *along with the number of iterations required to achieve the indicated accuracy. Note that the uniform errors here are of the same order of magnitude as those reported in* Table 14.

Table 21. Errors in the ADSG solution of (4.15).

M	m	m^2	h_s	$\|\|ES^s_{ADSG}(h_s)\|\|$	$\|\|EU^s_{ADSG}(h_u)\|\|$	Iter
2	5	25	2.22144	7.585e-003	1.302e-002	2
4	9	81	1.57080	2.365e-003	2.879e-003	3
8	17	289	1.11072	5.629e-004	5.868e-004	12
16	33	1089	0.78540	3.521e-005	3.393e-005	177
32	65	4225	0.55536	2.944e-006	8.829e-008	2112

Example 25. *Revisit the elliptic problem in* Example 17 *which we rewrite below for reference,*

$$\begin{aligned} -\triangle u(x,y) &= -\frac{3}{4}\left[\frac{x\sqrt{x}-x}{\sqrt{y}}+\frac{y\sqrt{y}-y}{\sqrt{x}}\right], \quad (x,y)\in\Omega=(0,1)\times(0,1) \\ u(x,y) &= 0, \quad (x,y)\in\partial\Omega. \end{aligned} \tag{4.16}$$

Recall that the exact solution is $u(x,y)=xy(1-\sqrt{x})(1-\sqrt{y})$. *Here as in* Example 17 *we let* $\alpha_s=\beta_s=\zeta_s=\eta_s=1/2$ *and choose the positive integer* M_x *arbitrarily. Since for the parameters chosen we have that* $M\equiv M_x=N_x=M_y=N_y$ *then* $m\equiv m_x=m_y=2M+1$ *and* $h_s=\pi/\sqrt{M}$. *Since the nodes are the same then*

$A = B$. *Computing the symmetric matrix* A *using* (4.8)*, and applying the Alternating-Direction Sinc-Galerkin method to the resulting Lyapunov equation yields the errors reported in* Table 22 *along with the number of iterations used to achieve the indicated accuracy. Note that the method performs well, in that the number of iterations required for large values of* M *is relatively small and the errors reported here are roughly comparable to those reported in* Table 15.

Table 22. Errors in the ADSG solution of (4.16).

M	m	m^2	h_s	$\|\|ES^s_{ADSG}(h_s)\|\|$	$\|\|EU^s_{ADSG}(h_u)\|\|$	Iter
2	5	25	2.22144	4.522e-004	1.804e-003	2
4	9	81	1.57080	2.238e-004	3.684e-004	2
8	17	289	1.11072	3.286e-005	3.443e-005	4
16	33	1089	0.78540	2.387e-006	2.612e-006	21
32	65	4225	0.55536	4.490e-008	4.041e-008	318

Example 26. *Revisit the elliptic problem in* Example 18 *which involved using different values for* M_x*,* N_x*,* M_y*, and* N_y*. We rewrite the problem below for reference,*

$$\begin{aligned} -\triangle u(x,y) &= f(x,y), \quad (x,y) \in \Omega = (0,1) \times (0,1) \\ u(x,y) &= 0, \quad (x,y) \in \partial\Omega, \end{aligned} \tag{4.17}$$

where $f(x,y) = 2xy^3\left[3y^2(-1+y)^2(-1+2x) + x^2(x-1)(10-30y+21y^2)\right]$ *when the exact solution is* $u(x,y) = x^3(1-x)y^5(1-y)^2$*. Recall that solving this problem using the symmetric Sinc-Galerkin method required solving a Sylvester equation. Let* $\alpha_s =$

5/2, $\beta_s = 1/2$, $\zeta_s = 9/2$, and $\eta_s = 3/2$. Choose M_x arbitrarily and compute $N_x = 5M_x$. Note that $M_y = \frac{5}{9}M_x$ when M_x is a multiple of 9 and $M_y = \left[\left|\frac{5}{9}M_x + 1\right|\right]$ otherwise. Similarly, $N_y = \frac{5}{3}M_x$ when M_x is a multiple of 3 and $N_y = \left[\left|\frac{5}{3}M_x + 1\right|\right]$ otherwise. Compute $m_x = M_x + N_x + 1$, $m_y = M_y + N_y + 1$ and for $d = \pi/2$ note that $h_s = \pi/\sqrt{5M_x}$. Determine the matrices A and B using (4.8) and apply the Alternating-Direction Sinc-Galerkin method to the resulting Sylvester equation to obtain the errors reported in Table 23*. As before, the column headed* Iter *shows the number of iterations used to achieve the indicated accuracy. Note that the errors reported below in* Table 23 *are comparable to those reported in* Table 16*.*

Table 23. Errors in the ADSG solution of (4.17).

M_x	N_x	m_x	M_y	N_y	m_y	h_s	$\|\|ES^s_{ADSG}(h_s)\|\|$	$\|\|EU^s_{ADSG}(h_u)\|\|$	Iter
1	5	7	1	2	4	1.40496	1.268e-004	4.006e-004	2
2	10	13	2	4	7	0.99346	1.446e-005	4.527e-005	4
3	15	19	2	5	8	0.81116	7.289e-006	1.289e-005	6
5	25	31	3	9	13	0.62832	4.136e-007	8.399e-007	47
11	55	67	7	19	27	0.42361	4.568e-008	6.206e-009	627

Examples on Infinite Domains

Now we turn our attention to problems posed on the first quadrant and consider an elliptic problem of the following form

$$\begin{aligned}
-\triangle u(x,y) &= f(x,y), \quad (x,y) \in \Omega = (0,\infty) \times (0,\infty) \\
u(x,0) &= 0, \quad x \in (0,\infty) \\
u(0,y) &= 0, \quad y \in (0,\infty) \\
u(x,y) &\to 0 \quad as \quad x \to \infty \\
u(x,y) &\to 0 \quad as \quad y \to \infty.
\end{aligned} \tag{4.18}$$

Once again we determine the parameters $\alpha_s, \beta_s, \zeta_s$, and η_s according to the asymptotic behavior of the exact solution (see (2.124) and (2.125)). We then choose M_x arbitrarily and compute the positive integers $N_x = [|\alpha_s M_x/\beta_s + 1|]$, $M_y = [|\alpha_s M_x/\zeta_s + 1|]$, and $N_y = [|\alpha_s M_x/\eta_s + 1|]$. (The discussion on how to choose N when $\alpha M/\beta$ is an integer also applies here.) We let $d = \pi/2$ and define $h_s \equiv h_x = h_y = \sqrt{\pi d/\alpha_s M_x}$.

We noted in Chapter 2 that assuming an approximate solution to (4.18) has the form

$$u^s_{m_x,m_y} = \sum_{j=-M_y}^{N_y} \sum_{i=-M_x}^{N_x} u^s_{ij} S_{ij}(x,y) \tag{4.19}$$

where $m_x = M_x + N_x + 1$ and $m_y = M_y + N_y + 1$, the resulting matrix system for the cofficients u^s_{ij} is a Sylvester equation $AX + XB = C$ where the matrices A, B, C, and X are given in (4.8). For the examples below, we note that the mapping $\phi_z(z) = \ln(z)$ is used when the exact solution decays algebraically with respect to z

and $\phi_z(z) = \ln(\sinh(z))$ is used when the exact solution decays exponentially with respect to z, where $z = x$ or y. We note further that if $\phi_z(z) = \ln(z)$ then we have that

$$\frac{-1}{(\phi_z')^{3/2}}\left(\frac{1}{(\phi_z')^{1/2}}\right)'' = \frac{1}{4}.$$

However, if $\phi_z(z) = \ln(\sinh(z))$ then

$$\frac{-1}{(\phi_z')^{3/2}}\left(\frac{1}{(\phi_z')^{1/2}}\right)'' = \frac{4\cosh^2(z) - 3}{4\cosh^4(z)}$$

and the matrix A and/or the matrix B must be modified accordingly.

We report the maximum errors over the set of sinc grid points

$$S = \{x_i\}_{i=-M_x}^{N_x} \times \{y_j\}_{j=-M_y}^{N_y} \tag{4.20}$$

and the set of uniform gridpoints

$$U = \{w_i\}_{i=0}^{100} \times \{z_j\}_{j=0}^{100} \tag{4.21}$$

as

$$||ES_{ADSG}^s(h_s)|| = \max_{\substack{-M_x \le i \le N_x \\ -M_y \le j \le N_y}} |u(x_i, y_j) - u_{ij}^s|$$

and

$$||EU_{ADSG}^s(h_u)|| = \max_{\substack{0 \le i \le 100 \\ 0 \le j \le 100}} |u(w_i, z_j) - u_{m_x,m_y}^s(w_i, z_j)|,$$

respectively. Let $h_u = 0.02$ and $h_v = 0.2$. Note that if $0 \le i \le 50$ then $w_i = ih_u$, whereas if $51 \le i \le 100$ then $w_i = 1 + (i - 50)h_v$. Similarly, if $0 \le j \le 50$ then $z_j = jh_u$, whereas if $51 \le j \le 100$ then $z_j = 1 + (j - 50)h_v$.

Example 27. *Revisit the elliptic problem in* Example 19 *which we rewrite below for reference,*

$$\begin{aligned} -\triangle u(x,y) &= f(x,y), \quad (x,y) \in \Omega = (0,\infty)\times(0,\infty) \\ u(x,0) &= 0, \quad x \in (0,\infty) \\ u(0,y) &= 0, \quad y \in (0,\infty) \\ u(x,y) &\to 0 \quad as \quad x \to \infty \\ u(x,y) &\to 0 \quad as \quad y \to \infty. \end{aligned} \tag{4.22}$$

where $f(x,y) = \dfrac{2xy}{(x^2+1)(y^2+1)}\left[\dfrac{3-x^2}{(x^2+1)^2} + \dfrac{3-y^2}{(y^2+1)^2}\right]$ *and the exact solution is given by* $u(x,y) = \dfrac{xy}{(x^2+1)(y^2+1)}$. *Choosing parameters as in* Example 19, $\alpha_s = \zeta_s = 1/2$ *and* $\beta_s = \eta_s = 3/2$. *For arbitrary* M_x *we note that* $M \equiv M_x = M_y$ *and* $N \equiv N_x = N_y$, *and thus* $m \equiv m_x = m_y$ *and for* $d = \pi/2$ *we have* $h_s = \pi/\sqrt{M}$. *If* M *is divisible by* 3 *then* $N = \dfrac{M}{3}$ *otherwise* $N = [|\frac{1}{3}M+1|]$. *Since the exact solution decays algebraically with respect to either* x *or* y *we use* $\phi_z(z) = \ln(z)$ *for* $z = x$ *or* y. *Since the nodes* $x_i = e^{ih}$ *and* $y_j = e^{jh}$ *are the same, then* $A = B$ *and the resulting discrete system for the elliptic problem is the Lyapunov equation,* $AX + XA = C$, *which we solve using the ADSG method to generate the errors in* Table 24. *We note that while the number of iterations is large, the errors are of the same order of magnitude as those in* Table 17.

Table 24. Errors in the ADSG solution of (4.22).

M	N	m	h_s	$\|\|ES^s_{ADSG}(h_s)\|\|$	$\|\|EU^s_{ADSG}(h_u)\|\|$	Iter
3	1	5	1.81380	1.484e-001	1.484e-001	2
6	2	9	1.28255	6.843e-002	7.352e-002	6
12	4	17	0.90690	5.908e-003	6.840e-003	239
24	8	33	0.64127	7.756e-004	3.182e-004	2567
48	16	65	0.45345	4.964e-004	1.255e-005	14587

Example 28. *Revisit the elliptic problem in* Example 20 *which we rewrite below for reference,*

$$\begin{aligned} -\triangle u(x,y) &= f(x,y), \quad (x,y) \in \Omega = (0,\infty) \times (0,\infty) \\ u(x,0) &= 0, \quad x \in (0,\infty) \\ u(0,y) &= 0, \quad y \in (0,\infty) \\ u(x,y) &\to 0 \quad as \quad x \to \infty \\ u(x,y) &\to 0 \quad as \quad y \to \infty \end{aligned} \tag{4.23}$$

where $f(x,y) = 2(x - xy + y)e^{-(x+y)}$ *and the exact solution is given by* $u(x,y) = xye^{-(x+y)}$. *We note that the exact solution has the form given in* (2.125) *and thus we can choose the parameters* $\alpha_s = \zeta_s = 1/2$ *and* $\beta_s = \eta_s = 1$ *so that* $M \equiv M_x = M_y$ *and* $N = N_x = N_y$. *If* M *is divisible by* 2 *then* $N = \dfrac{M}{2}$ *otherwise* $N = [|\frac{1}{2}M + 1|]$. *Note* $m = M + N + 1$, *and* $h_s \equiv h_x = h_y = \pi/\sqrt{M}$ *when* $d = \pi/2$. *Since the exact solution decays exponentially with respect to* x *or* y *then for* $z = x$ *or* y, *we choose* $\phi_z(z) = \ln(\sinh(z))$ *and compute the nodes* $x_i = \ln(e^{ih} + \sqrt{e^{2ih} + 1})$ *and* $y_j = \ln(e^{jh} + \sqrt{e^{2jh} + 1})$. *We again derive a Lyapunov equation which we solve using the*

ADSG method and obtain the errors in Table 25. *We note that the errors here are of the same order of magnitude as those in* Table 18.

Table 25. Errors in the ADSG solution of (4.23).

M	N	m	h_s	$\|\|ES^s_{ADSG}(h_s)\|\|$	$\|\|EU^s_{ADSG}(h_u)\|\|$	Iter
2	1	4	2.22144	6.982e-002	7.372e-002	2
5	3	9	1.40496	5.395e-002	5.694e-002	4
12	6	19	0.90690	1.538e-003	2.262e-003	71
22	11	34	0.66979	3.389e-004	5.177e-004	206
43	22	66	0.47909	2.553e-005	1.763e-005	678

It should be noted that program execution times for the ADSG method are much smaller than those for the algorithm that employs the Kronecker product, the Kronecker sum, and the concatenation operator. In that sense ADSG easily outperforms that standard method, avoiding the large and often poorly conditioned matrices that result from computing the Kronecker product, as well as the ensuing extensive storage requirements. The ADSG method is also very competitive with the diagonalization technique discussed in Chapter 2 (see (2.104), (2.105), (2.106)), but does not require the often extensive work necessary to compute matrix eigenvalues and eigenvectors. Thus the Alternating-Direction Sinc-Galerkin method is a fast and accurate practical alternative to those standard methods and is versatile enough to be used to solve singular and nonsingular problems in one or in multiple dimensions, on finite or on infinite domains.

REFERENCES CITED

[1] N. Alonso III and K. L. Bowers. An Alternating-Direction Sinc-Galerkin Method for Elliptic Problems. *Journal of Complexity* (2009), doi:10.1016/j.jco.2009.02.006.

[2] K. L. Bowers and J. Lund. Numerical Solution of Singular Poisson Problems via the Sinc-Galerkin Method. *SIAM Journal on Numerical Analysis*, 24(1):36-51, 1987.

[3] J. Douglas Jr. and T. Dupont. Alternating-Direction Galerkin Methods on Rectangles, in *Numerical Solution of Partial Differential Equations-II.* B. Hubbard, ed., Academic Press, New York, 1971, 133-214.

[4] A. Graham. *Kronecker Products and Matrix Calculus with Applications.* Ellis Horwood Limited: Market Cross House, Cooper Street, Chichester, West Sussex, PO19 1EB, England, 1981.

[5] H. Hurwitz. Infinitesimal Scaling: A New Procedure for Modeling Exterior Field Problems. *IEEE Transactions on Magnetics*, 20(5):1918-1923, 1984.

[6] S. Koonprasert. *The Sinc-Galerkin Method for Problems in Oceanography.* PhD thesis, Montana State University, Bozeman, Montana, 2003.

[7] D. L. Lewis, J. Lund and K. L. Bowers. The Space-Time Sinc-Galerkin Method for Parabolic Problems. *International Journal for Numerical Methods in Engineering*, 24(9):1629-1644, 1987.

[8] A. Lu and E. L. Wachspress. Solution of Lyapunov Equations by Alternating-Direction Implicit Iteration. *Computers and Mathematics with Applications*, 21(9):43-58, 1991.

[9] J. Lund. Symmetrization of the Sinc-Galerkin Method for Boundary Value Problems. *Mathematics of Computation*, 47(176):571-588,1986.

[10] J. Lund, K. L. Bowers, and K. M. McArthur. Symmetrization of the Sinc-Galerkin Method with Block Techniques for Elliptic Equations. *IMA Journal of Numerical Analysis*, 9(1):29-46, 1989.

[11] J. Lund and K. L. Bowers. *Sinc Methods for Quadrature and Differential Equations.* SIAM: Philadelphia, 1992.

[12] N. J. Lybeck. *Sinc Domain Decomposition Methods for Elliptic Problems.* PhD thesis, Montana State University, Bozeman, Montana, 1994.

[13] N. J. Lybeck and K. L. Bowers. Domain Decomposition in Conjunction with Sinc Methods for Poisson's Equation. *Numerical Methods for Partial Differential Equations*, 12(4):461-487, 1996.

[14] K. M. McArthur, K. L. Bowers, and J. Lund. Numerical Implementation of the Sinc-Galerkin Method for Second-Order Hyperbolic Equations. *Numerical Methods for Partial Differential Equations*, 3(3):169-185, 1987.

[15] K. M. McArthur, K. L. Bowers, and J. Lund. The Sinc Method in Multiple Space Dimensions: Model Problems. *Numerische Mathematik*, 56(8):789-816, 1990.

[16] A. C. Morlet, N. J. Lybeck, and K. L. Bowers. The Schwarz Alternating Sinc Domain Decomposition Method. *Applied Numerical Mathematics*, 25(4):461-483, 1997.

[17] A. C. Morlet, N. J. Lybeck, and K. L. Bowers. Convergence of the Sinc Overlapping Domain Decomposition Method. *Applied Mathematics and Computation*, 98(2-3):209-227, 1999.

[18] M. Ng. Fast Iterative Methods for Symmetric Sinc-Galerkin Systems. *IMA Journal of Numerical Analysis*, 19:357-373, 1999.

[19] M. Ng and Z. Bai. A Hybrid Preconditioner of Banded Matrix Approximation and Alternating-Direction Implicit Iteration for Symmetric Sinc-Galerkin Linear Systems. *Linear Algebra and its Applications*, 366:317-335, 2003.

[20] D. W. Peaceman and H. H. Rachford, Jr. The Numerical Solution of Parabolic and Elliptic Differential Equations. *Journal of the Society for Industrial and Applied Mathematics*, 3:28-41, 1955.

[21] W. H. Press, B. P. Flannery, S. A. Teukolsky and W.T. Vetterling. *Numerical Recipes: The Art of Scientific Computing.* Cambridge University Press, 1986.

[22] G. D. Smith. *Numerical Solution of Partial Differential Equations: Finite Difference Methods (2^{nd} ed).* Clarendon Press, Oxford, 1978.

[23] F. Stenger. Approximations via Whittaker's Cardinal Function. *Journal of Approximation Theory*, 17(3):222-240, 1976.

[24] F. Stenger. A Sinc-Galerkin Method of Solution of Boundary Value Problems. *Mathematics of Computation*, 33(145):85-109, 1979.

[25] F. Stenger. Numerical Methods Based on Whittaker Cardinal, or Sinc Functions. *SIAM Review*, 23(2):165-224, 1981.

[26] F. Stenger. *Numerical Methods Based on Sinc and Analytic Functions*. Springer-Verlag, New York, 1993.

[27] D. U. von Rosenberg. Methods for the Numerical Solution of Partial Differential Equations in *Modern Analytic and Computational Methods in Science and Mathematics*. R. Bellman, ed., American Elsevier Publishing Company, Inc, New York, 1969.

[28] E. L. Wachspress. Trail to a Lyapunov Equation Solver. *Computers and Mathematics with Applications*, 55:1653-1659, 2008.

[29] E. T. Whittaker. On the Functions Which are Represented by the Expansions of the Interpolation Theory. *Proceedings of the Royal Society of Edinburg*, 35:181-194, 1915.

[30] J. M. Whittaker. *Interpolation Function Theory*. Cambridge Tracts in Mathematics and Mathematical Physics. No. 33, Cambridge University Press, London, 1935.

[31] D. Young, Section in *Modern Mathematics for the Engineer*. Second Series, edited by E.F. Beckenbach, McGraw-Hill, New York, 1961.