A fully Galerkin method for parabolic problems
by David Lamar Lewis

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Mathematics
Montana State University

Abstract:
This thesis contains the formulation and development of a space-time Sinc-Galerkin method for the numerical solution of a parabolic partial differential equation in one space dimension. The space-time adjective means that a Galerkin technique is employed simultaneously over the spatial and temporal domains. Salient features of the method include: exponential convergence rate, an easily assembled discrete system, production of a global approximation to the solution, and the ability to handle singular problems. Methods of solution for the discrete system and implementation of the method for a number of test problems are included.

A FULLY GALERKIN METHOD

FOR PARABOLIC PROBLEMS

by

David Lamar Lewis

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

of

Doctor of Philosophy

in

Mathematics

MONTANA STATE UNIVERSITY
Bozeman, Montana

December 1989

APPROVAL

of a thesis submitted by

David Lamar Lewis

This thesis has been read by each member of the thesis 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 College of Graduate Studies.

11/27/89
Date

John Lund
Chairperson, Graduate Committee

Approved for the Major Department

11/30/89
Date

K. Tiabut
Head, Major Department

Approved for the College of Graduate Studies

December 7, 1989
Date

Henry L. Parsons
Graduate Dean

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this thesis is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this thesis should be referred to University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute copies of the dissertation in and from microfilm and the right to reproduce and distribute by abstract in any format."

Signature David Lamar Lewis

Date 29 November 1989

ACKNOWLEDGMENTS

Several people were of particular help in the various phases of this endeavor. Thank you.

To John Lund, whose knowledge, guidance and tenacity permeates each phase. This project may not have driven him to distraction, but it certainly got him to the outlying suburbs.

To Ken Bowers, who assured me that it was neither the computer nor the program.

To Rene' Tritz, whose unsurpassed skills and flying fingers enabled existent deadlines to be met.

To Sheri, whose support and counseling during several digressions put the train back on track.

To Mom and Dad for always being there.

And to serendipity.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This thesis contains the formulation and development of a space-time Sinc-Galerkin method for the numerical solution of a parabolic partial differential equation in one space dimension. The space-time adjective means that a Galerkin technique is employed simultaneously over the spatial and temporal domains. Salient features of the method include: exponential convergence rate, an easily assembled discrete system, production of a global approximation to the solution, and the ability to handle singular problems. Methods of solution for the discrete system and implementation of the method for a number of test problems are included.

CHAPTER 1

INTRODUCTION

A two-dimensional Galerkin scheme for the solution of the differential equation

$$Lu = g \tag{1.1}$$

on a domain $D \subseteq \mathbb{R}^2$ can be summarized as follows. Let u_N be an approximation to u defined by

$$u_N(x,t) = u_0(x,t) + \sum_{j=1}^{N} a_j \phi_j(x,t) \tag{1.2}$$

where the ϕ_j are known basis functions and u_0 is defined to satisfy the boundary conditions. Define the residual R by

$$R(a_1, a_2, \ldots, a_N, x, t) \equiv Lu_N - g \quad .$$

A standard weighted inner product will be used, given by

$$(f,g) = \int\int_D f(x,t)g(x,t)w(x,t)dxdt \quad . \tag{1.3}$$

A Galerkin scheme determines the unknown coefficients by orthogonalizing the residual with respect to the basis functions for $k = 1, \ldots, N$

$$(R, \phi_k) = 0 \quad . \tag{1.4}$$

The properties of the method are governed by the basis functions ϕ_j in (1.2), the inner product defined in (1.3) and the quadrature used to evaluate the integrals arising from (1.4).

The linear operator considered here is the constant coefficient parabolic operator with zero boundary conditions,

$$
\begin{aligned}
Pu(x,t) &\equiv u_t(x,t) - u_{xx}(x,t) = g(x,t) \ ; \ (x,t) \in D \\
u(x,t) &= 0 \ ; \ (x,t) \in \partial D \qquad (1.5) \\
D &= \{(x,t) : 0 < x < 1 \ , \ 0 < t < \infty\} \quad .
\end{aligned}
$$

The subject of this thesis is the numerical solution of (1.5) using a Galerkin method simultaneously in space and time with emphasis placed on the temporal approximation. Nonhomogeneous boundary data may be handled by transforming the problem to one of the form of (1.5). This procedure is illustrated in Examples 6 and 7 of Chapter 5. For spatial product regions contained in $\mathbb{R}^n$, $n \geq 2$, the technique of the present thesis is applicable via Kronecker products of the matrices developed in this thesis. This has been developed and discussed in [2]. For more general spatial domains in $\mathbb{R}^n$, $n \geq 2$, the procedure developed in [14] is applicable. As this approach is somewhat unconventional, a brief description of the standard approach used for constructing an approximation to (1.5) is provided for contrast.

Generally, construction of an approximate to the solution u of (1.5) begins with a discretization of the spatial domain using a finite difference, finite element, collocation or Galerkin method [4,6,7,10]. This approach yields a system of ordinary differential equations in the time variable. To complete the approximation, one uses a time-differencing scheme in which the approximation at a given time is dependent on the approximation constructed at the previous time level or in some cases on several previous time levels. Stability constraints on the discrete evolution operator result in low order methods that require "small" time steps to obtain accurate approximations [1]. When the approximate solution is desired at

moderate (or large) times, a large number of applications of the discrete evolution operator must be used to reach that time.

A great deal of effort has been expended on developing methods which improve the approximation of the time derivative. The survey paper [10] classifies methods of order greater than two for the time integration. In [4], methods which are up to fourth order accurate in time are developed by employing various extrapolations at each time level. While each of these approaches possess certain advantages they all result in methods that converge algebraically in time. When used in conjunction with an exponential order spatial approximation (e.g. Galerkin, spectral) the result is a method that is exponentially convergent in the spatial domain and algebraically convergent in the temporal domain. A type of space-time finite element technique was developed in [15]. When piecewise polynomials are used in the temporal domain, the resulting method remains an algebraically convergent method. In contrast the method presented here enjoys an exponential convergence rate in each coordinate direction, and appropriate parameter selections allow one to balance these rates.

The present method maintains its exponential convergence rate in the case of problem (1.5) whose solution has a singularity on the boundary of D. This is in marked contrast to polynomial based methods whose approximation qualities deteriorate in the presence of singularities.

The basis elements employed in the Sinc-Galerkin method are translates of the sinc function, denoted $sinc(x)$, which is defined by

$$sinc(x) = \frac{\sin(\pi x)}{\pi x} \quad , \quad x \in \mathbb{R} \quad . \tag{1.6}$$

The approximation procedures herein are based on the following formal expansion

of a function f:

$$(1.7) \qquad C(f,h,x) = \sum_{k=-\infty}^{\infty} f(kh) sinc\left[\frac{x-kh}{h}\right] \quad , \quad h>0 \quad .$$

The expansion (1.7) was developed and studied by E.T. Whittaker [16] who dubbed the function $C(f,h,x)$ the "cardinal function" of f (whenever the series converges). He conjectured that given a function with singularities and fluctuations, the "cardinal function" could be substituted for it "in all practical and some theoretical investigations since they are identical at an infinite number of values." He called $C(f,h,x)$ "a function of royal blood in the family of entire functions, whose distinguished properties separate it from its bourgeois brethren." Currently, the function $C(f,h,x)$ is called the Whittaker cardinal function. The cardinal function used as a tool in numerical analysis is of relatively recent vintage. The survey paper [13] provides both historical information on the cardinal function as well as an exhaustive list of its properties through 1981.

F. Stenger [11] showed that for functions defined on the real line, interpolation and integration procedures based on the Whittaker cardinal function have the exponential convergence rate $0(\exp -\kappa\sqrt{N})$ where $\kappa > 0$ is a constant independent of N and $2N+1$ basis elements are used in the procedure. By using conformal maps, he expanded these procedures to include functions defined on an arbitrary interval (a,b) and showed that the exponential convergence rate was maintained.

In [12] Stenger showed that the convergence rate $0(\exp -\kappa\sqrt{N})$ was optimal over the class of N-point approximations whether or not the function being approximated had singularities at the endpoints of the interval, (a,b). Included in [12] is an initial development of the Sinc-Galerkin method to approximate the solution of linear and nonlinear second order ordinary differential equations. The approximate solution of some elliptic and parabolic partial differential equations is

also discussed. The reduction of the differential equation to a system of algebraic equations was accomplished via explicit approximations of the weighted inner product defining the Sinc-Galerkin scheme. In each case the matrix formulation of the system of algebraic equations can be expressed in the form

$$\mathcal{A}\vec{u} = \vec{b} \tag{1.8}$$

where $\vec{u}$ is the vector of unknowns and the components of the matrix $\mathcal{A}$ and the vector $\vec{b}$ are known. Unlike the sparse matrices arising from polynomial approximations, the matrix $\mathcal{A}$ for the Sinc-Galerkin method is full.

Specifically, when applying the method to a second order ordinary differential equation Stenger showed that the discretization yields a matrix $\mathcal{A}$ represented by a sum of matrices

$$\mathcal{A} = I^{(2)} D_2 + I^{(1)} D_1 + D_0 \tag{1.9}$$

where $I^{(1)}$ is a full skew-symmetric matrix, $I^{(2)}$ is a full negative definite symmetric matrix and the matrices D_0, D_1 and D_2 are diagonal and dependent upon the domain of definition. In the case of a self-adjoint second order problem, the matrix $\mathcal{A}$ is not in general a symmetric matrix. However, the method is applicable to any second order problem including those with regular singular points. While it is possible to make assertions about the spectrum of each term of $\mathcal{A}$, a specific characterization of the spectral properties of $\mathcal{A}$ is domain dependent.

Lund [9] showed that for self-adjoint second order differential equations, the skew-symmetric matrix $I^{(1)}$ could be eliminated from $\mathcal{A}$ in (1.9) by selecting a weight function different from that used in [12]. The selected weight produced a symmetric discrete system (1.8). Numerical testing showed an unexpected increase in the accuracy of the approximation produced by the symmetrized method

over that produced by the method of [12]. In exchange for the symmetrization, a stronger assumption is required on the behavior of the true solution. As a consequence, the methodology of [9] applies to a more restrictive class of problems than the methodology in [12]. Due to the nature of the solution of problem (1.5), the method of [9] is applicable and is the method of choice for the spatial approximation in this thesis.

Initially the methodology outlined in [12] was used in the time domain. When this method was applied to a variety of test cases, the numerical results often failed to exhibit the expected exponential rate of convergence. Whereas the accuracy of the interpolation for the various derivative approximations can be guaranteed, a closer examination of the discrete system for (1.5) indicated that matrix conditioning was the source of these discrepancies. When the development was carried out using a general weight it was clear that the conditioning was dependent on the weight function used in the inner product. It is shown that both the eigenvalues of the discrete system as well as matrix conditioning are controlled by selecting a different weight than that proposed in [12]. Additionally, when the new weight was tested, the numerical results in each test case exhibited the anticipated rate of convergence. As in the symmetrization method of [9], this new selection requires a stronger assumption on the behavior of the true solution.

Chapter 2 contains an outline of the approach taken to produce a numerical approximation to the solution of (1.5) using the Sinc-Galerkin method. This technique is a product approximation and as such the spatial and temporal approximations are given separately. Various results from the work of Stenger [11,12,13] germaine to the development are included for completeness. The chapter concludes with the symmetric Sinc-Galerkin method applied to the one-dimensional

problem

$$
\begin{aligned}
u''(x) &= g(x) \quad , \quad x \in (0,1) \\
u(0) &= u(1) = 0 \quad .
\end{aligned}
\tag{1.10}
$$

Chapter 3 begins with a detailed development of the Sinc-Galerkin method using an arbitrary weight function in the inner product for the one-dimensional problem

$$
\begin{aligned}
u'(t) &= g(t) \quad , \quad t \in (0,\infty) \\
u(0) &= 0 \quad .
\end{aligned}
\tag{1.11}
$$

The focus here is toward identifying the impact of the weight function on the conditioning of the resulting discrete system. The analysis shows that in the one parameter family of weight functions identified, there is a range of values of the parameter for which the conditioning is controlled. Correspondingly there is a range of values of the parameter for which the exponential convergence rate of the method is enhanced. In the intersection of these two ranges is a parameter defining the weight function for which the conditioning of the system is controlled and the exponential convergence rate is maintained.

The approximations in Chapters 2 and 3 are assembled in Chapter 4 to produce the discrete system which determines the numerical approximation of the solution to (1.5). Three methods for solving the discrete system are discussed. The first method is in the form of (1.8) where the coefficient matrix is the Kronecker sum of the matrices developed in Chapters 2 and 3. This form is amenable to either direct or iterative solution methods. The second method is based on the diagonalizability of the matrices from the spatial and temporal discretizations. Due to the weight selection from [9] the former is orthogonally diagonalizable. The diagonalizability of the latter is a conjecture of this thesis. This

method is included both to numerically test the conjecture and to clearly reveal the eigenstructure of the system. The validity of the numerical results of this thesis do not depend on this diagonalizability conjecture. Indeed, the third method of solution presented, based on the Schur decomposition, produces identical numerical results without the assumption of diagonalizability.

Finally, Chapter 5 includes the numerical results for seven examples. From the many examples tested, these were selected to illustrate implementation aspects of the method. Both homogeneous and nonhomogeneous problems are included. Specifically, the first four examples were selected to numerically illustrate the exponential convergence rate for singular and nonsingular problems. In the former case, problems exhibiting algebraic or logarithmic singularities are included. The last three examples were taken from the literature and could provide an initial basis for a comparative study of this Sinc-Galerkin method and alternative methods of solution of parabolic problems.

CHAPTER 2

THE DISCRETE SYSTEM FOR THE SPATIAL DOMAIN

The development of the Sinc-Galerkin method for either of problems (1.10) or (1.11) depends on the accurate interpolation and quadrature formulas of [11,12,13] for the approximate inner products defining the method. Both to see why the two problems (1.10) and (1.11) arise from (1.5) as well as to review the results needed from [11,12,13], an outline of the space-time Sinc-Galerkin method is given in this chapter. Following this general discussion, the specific development for problem (1.10) is presented.

The space-time Sinc-Galerkin scheme to approximate the solution of the parabolic problem

$$Pu(x,t) \equiv u_t(x,t) - u_{xx}(x,t) = g(x,t) \;\; ; \;\; (x,t) \in D$$

$$u(x,t) = 0 \;\; , \;\; (x,t) \in \partial D \tag{2.1}$$

$$D = \{(x,t) : 0 < x < 1 \; ; \; 0 < t < \infty\}$$

may be summarized as follows: Select the basis functions $\{S_i(x), S_j^*(t)\}$ and define an approximate solution to (2.1) by way of the formula

$$u_{m_x,m_t}(x,t) \equiv \sum_{i=-M_x}^{N_x} \sum_{j=-M_t}^{N_t} u_{ij} S_i(x) S_j^*(t) \tag{2.2}$$

where

$$m_x = M_x + N_x + 1 \;\text{ and }\; m_t = M_t + N_t + 1 \quad . \tag{2.3}$$

The coefficients $\{u_{ij}\}$ in (2.2) are determined from the discrete system

$$(Pu_{m_x,m_t} - g \;,\; S_k S_\ell^*) = 0 \;\; , \;\; \begin{matrix} k = -M_x, \ldots, 0, \ldots, N_x \\ \ell = -M_t, \ldots, 0, \ldots, N_t \end{matrix} \tag{2.4}$$

where $(\cdot,\cdot)$ is an appropriate inner product. The choice of the inner product in (2.4) coupled with the choice of the basis functions in (2.2) determines the properties of the approximating method.

For the Sinc-Galerkin method, the basis functions are compositions of the sinc function (1.6) with conformal maps and will be defined later. The inner product indicated in (2.4) is defined by

$$(u,v) = \int_0^\infty \int_0^1 u(x,t)v(x,t)w(x,t)dxdt \tag{2.5}$$

where w is a weight function that will be selected as the development proceeds. To motivate the one-dimensional development that follows, assume that the weight $w(x,t) = w_1(x)w_2(t)$ so that the residual in (2.4) with (2.5) may be written

$$\begin{aligned} &\int_0^\infty \int_0^1 u_t(x,t)S_k(x)S_\ell^*(t)w_1(x)w_2(t)dxdt \\ &-\int_0^\infty \int_0^1 u_{xx}(x,t)S_k(x)S_\ell^*(t)w_1(x)w_2(t)dxdt \\ &=\int_0^\infty \int_0^1 g(x,t)S_k(x)S_\ell^*(t)w_1(x)w_2(t)dxdt \quad . \end{aligned}$$

While construction of the discrete system defined in (2.4) is the result of direct computation, it is tedious, obscures important parameter selections and fails to provide direction in selecting the weight function. A product-method of approach is achieved through developing the Sinc-Galerkin method for the one-dimensional problems

$$\begin{aligned} &u''(x) = g(x) \ , \ 0 < x < 1 \\ &u(0) = u(1) = 0 \end{aligned} \tag{2.6}$$

and

$$\begin{aligned} &u'(t) = g(t) \ , \ t > 0 \\ &u(0) = 0 \quad . \end{aligned} \tag{2.7}$$

The latter approach produces matrices that are identical to those derived from a direct discretization of (2.4) and illuminates parameter and weight function selection.

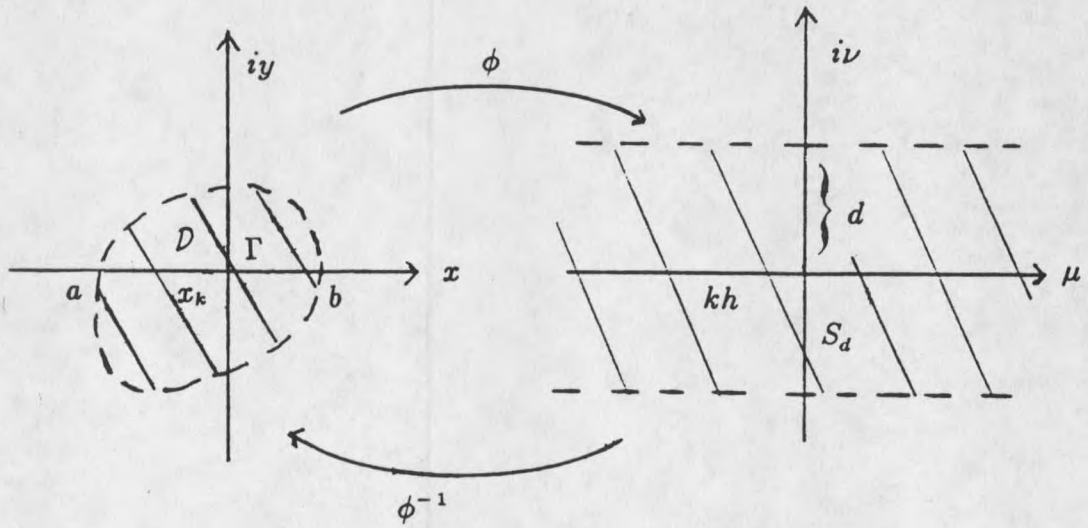
Before attacking either problem (2.6) or (2.7), the foundation and basic assumptions necessary for the Sinc-Galerkin method are presented. Since the domain of concern differs in problems (2.6) and (2.7), the preliminary results are developed on the general interval (a, b) with $a \neq b$.

Let a and b be real and let D be a simply connected domain in the complex plane such that $a, b \in \partial D$. Let $d > 0$ and $\phi : D \to S_d$ be a conformal single valued mapping onto the infinite strip

$$S_d = \{\zeta = \mu + i\nu : |\nu| < d\} \quad . \tag{2.8}$$

Furthermore, select ϕ such that $\phi(a) = -\infty$ and $\phi(b) = \infty$. Let ϕ^{-1} denote the inverse of ϕ and set $\Gamma = (a, b) = \phi^{-1}\big((-\infty, \infty)\big)$ and $x_k = \phi^{-1}(kh)$ where $h > 0$ and k is an integer. A graphic respresentation of the situation is supplied in Figure 1.

Figure 1. The domains D and S_d.



The basis functions used in the Sinc-Galerkin method are translates of the sinc function (1.6) composed with the conformal mapping ϕ:

$$\begin{aligned} S_k(x) &\equiv S(k,h) \circ \phi(x) \\ &\equiv sinc\left[\frac{\phi(x) - kh}{h}\right] \\ &= \frac{\sin\left[\frac{\pi}{h}\left(\phi(x) - kh\right)\right]}{\frac{\pi}{h}\left(\phi(x) - kh\right)} \end{aligned} \tag{2.9}$$

It is convenient to note a property of the basis functions and to introduce some notation at this time. Define

$$\delta_{kj}^{(n)} \equiv h^n \left(\frac{d}{d\phi}\right)^n S(k,h) \circ \phi(x)\Big|_{x=x_j} \tag{2.10}$$

so that, in particular,

$$\delta_{kj}^{(0)} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases}, \tag{2.11}$$

$$\delta_{kj}^{(1)} = \begin{cases} 0 & \text{if } j = k \\ \frac{(-1)^{j-k}}{j-k} & \text{if } j \neq k \end{cases}, \tag{2.12}$$

and

$$\delta_{kj}^{(2)} = \begin{cases} \frac{-\pi^2}{3} & \text{if } j = k \\ \frac{-2(-1)^{j-k}}{(j-k)^2} & \text{if } j \neq k \end{cases}. \tag{2.13}$$

Of primary importance to the Sinc-Galerkin method is the quadrature rule used to evaluate the integrals resulting from the various inner products in (2.4); in particular, the form of the error incurred by the quadrature rule is needed. The following theorems are given without proof (see [11,12] for the proofs) and embody the tools required to construct an approximation to the solution u of either (2.6) or (2.7) using a finite set of the basis elements. The theorems are stated for a contour Γ in the complex plane though they are used in this thesis on

subintervals of the real line. Theorem 2.1 identifies the class of functions for which sinc interpolation yields an exponentially accurate sinc interpolant. The result of Theorem 2.2 exhibits the exponential rate of convergence of the quadrature rule and Theorem 2.3 provides a technique for truncating the infinite sum resulting from the quadrature rule.

Let $B(\mathcal{D})$ denote the class of functions f that are analytic in $\mathcal{D}$ such that

$$\int_{\phi^{-1}(\mu+L)} |f(\zeta)d\zeta| \to 0 \quad \text{as} \quad \mu \to \pm\infty \tag{2.14}$$

where

$$L = \{i\nu : |\nu| < d\}$$

and

$$N(f,\mathcal{D}) = \int_{\partial\mathcal{D}} |f(z)dz| < \infty \quad . \tag{2.15}$$

Theorem 2.1: Let $z_k = \phi^{-1}(kh)$ and $f \in B(\mathcal{D})$. Then for every $z \in \Gamma$

$$\begin{aligned} \epsilon(z) &\equiv \frac{f(z)}{\phi'(z)} - \sum_{k=-\infty}^{\infty} \frac{f(z_k)}{\phi'(z_k)}\, S_k(z) \\ &= \frac{\sin\left(\frac{\pi}{h}\,\phi(z)\right)}{2\pi i} \int_{\partial\mathcal{D}} \frac{f(w)}{[\phi(w)-\phi(z)]\sin\left(\frac{\pi}{h}\,\phi(w)\right)}\, dw \end{aligned} \tag{2.16}$$

and

$$|\epsilon(z)| \le \frac{N(f,\mathcal{D})}{2\pi d\, sinh\left(\frac{\pi d}{h}\right)} \quad . \tag{2.17}$$

Theorem 2.2: Let $f \in B(\mathcal{D})$. Then

$$\begin{aligned} I_f &\equiv \int_{\Gamma} f(z)dz - h \sum_{k=-\infty}^{\infty} \frac{f(z_k)}{\phi'(z_k)} \\ &= \frac{i}{2} \int_{\partial\mathcal{D}} \frac{K(\phi,h)(w)f(w)}{\sin(\pi\phi(w)/h)}\, dw \end{aligned} \tag{2.18}$$

where

$$K(\phi,h)(w) = \exp\left\{i\pi\phi(w)sgn\left[Im(\phi(w))\right]/h\right\} \quad . \tag{2.19}$$

Furthermore, a short calculation produces the identity

$$\left|K(\phi,h)(w)\right|\Big|_{w\in\partial\mathcal{D}} = \exp(-\pi d/h) \tag{2.20}$$

and the inequality

$$|I_f| \leq \frac{e^{-\pi d/h}}{2sinh(\pi d/h)}\, N(f,\mathcal{D}) \quad . \tag{2.21}$$

Theorem 2.3: Suppose f is analytic in $\mathcal{D}$ and that there exist positive constants α, β and C such that

$$\left|\frac{f(z)}{\phi'(z)}\right| \leq C \begin{cases} \exp(-\alpha|\phi(z)| & , z\in\Gamma_L \\ \exp(-\beta|\phi(z)| & , z\in\Gamma_R \end{cases} \tag{2.22}$$

where

$$\Gamma_L = \left\{z : z\in\phi^{-1}\left((-\infty,0)\right)\right\}$$

and

$$\Gamma_R = \left\{z : z\in\phi^{-1}\left([0,\infty)\right)\right\} \quad .$$

Then

$$\begin{aligned} &\left|\sum_{k=-\infty}^{\infty}\frac{f(z_k)}{\phi'(z_k)} - \sum_{k=-M}^{N}\frac{f(z_k)}{\phi'(z_k)}\right| \\ &\leq \left|\sum_{k=-\infty}^{-M-1}\frac{f(z_k)}{\phi'(z_k)}\right| + \left|\sum_{k=N+1}^{\infty}\frac{f(z_k)}{\phi'(z_k)}\right| \\ &\leq C\left\{(1/\alpha)\exp(-\alpha Mh) + (1/\beta)\exp(-\beta Nh)\right\} \quad . \end{aligned} \tag{2.23}$$

Combining the result of Theorems 2.2 and 2.3, the error in the finite sinc quadrature can be expressed as

$$\left| \int_\Gamma f(z)dz - \sum_{k=-M}^{N} \frac{f(z_k)}{\phi'(z_k)} \right| \leq C \left\{ (1/\alpha) \exp(-\alpha M h) + (1/\beta) \exp(-\beta N h) \right\} + \frac{e^{-\pi d/h} N(f, \mathcal{D})}{2 sinh(\pi d/h)} \quad . \tag{2.24}$$

The boundary integral and truncation errors in the finite sinc quadrature (2.24) are asymptotically balanced when $\pi d/h = \alpha M h$ and $\alpha M h = \beta N h$ or when

$$h = \sqrt{\frac{\pi d}{\alpha M}} \tag{2.25}$$

and

$$N = \left[\!\left[\frac{\alpha M}{\beta} + 1 \right]\!\right] \tag{2.26}$$

where $[\![\cdot]\!]$ denotes the greatest integer function. The inequality in (2.24) is maintained if $\frac{\alpha}{\beta} M$ is an integer and $N = \frac{\alpha}{\beta} M$. The addition of 1 in (2.26) is required only when $\frac{\alpha}{\beta} M$ is not an integer.

It should be noted that the exponential rate of convergence in (2.21) and (2.23) does not depend on the boundedness of derivatives of f. It is this fact that preserves the exponential accuracy of the Sinc-Galerkin method when applied to problems with singular solutions.

The inner product employed for either of problems (2.6) or (2.7) is defined by

$$(u, v) = \int_a^b u(x) v(x) \left(\phi'(x)\right)^\sigma dx \quad . \tag{2.27}$$

The value of σ is problem dependent and will be identified during the ensuing development for each problem.

The remainder of this chapter is devoted to identifying the form of the discrete system which produces the Sinc-Galerkin approximation of (2.6). A detailed development of the approximation is provided in [9], thus only the results are included here. Select the conformal map

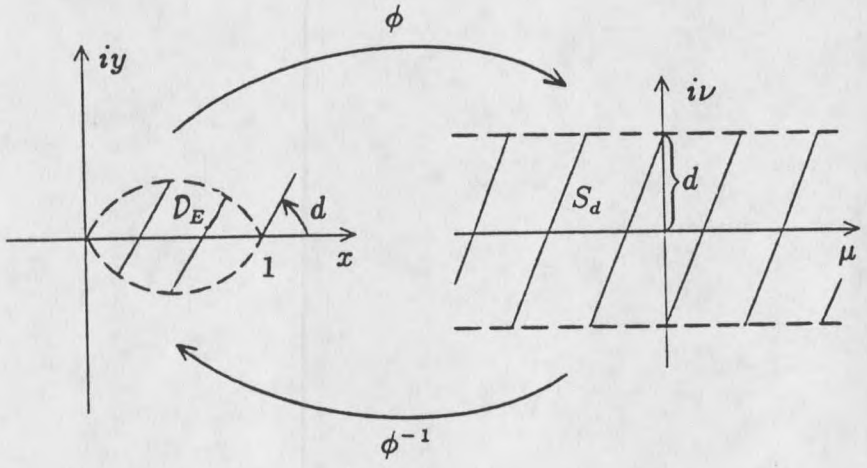
$$\phi(z) = \ell n\left(\frac{z}{1-z}\right) \tag{2.28}$$

which carries the "eye"-shaped region

$$\mathcal{D}_E = \{z = x + iy : |arg[z/(1-z)]| < d \leq \pi/2\} \tag{2.29}$$

onto the infinite strip S_d in (2.8). These domains are depicted in Figure 2.

Figure 2. The domains $\mathcal{D}_E$ and S_d.



The inverse map, ϕ^{-1}, is given by

$$\phi^{-1}(\zeta) = e^{\zeta}/(1+e^{\zeta}) \quad , \tag{2.30}$$

thus the sinc nodes are

$$x_k = \phi^{-1}(kh) = e^{kh}/(1+e^{kh}) \quad . \tag{2.31}$$

The basis elements, $S_k(x)$, are defined as in (2.9) and an approximate solution of (2.6) is defined by

$$u_m(x) = \sum_{j=-M}^{N} u_j S_j(x) \quad , \quad m = M + N + 1 \quad . \tag{2.32}$$

The discrete system produced by orthogonalizing the residual with respect to the basis functions $S_k(x)$,

$$(u'' - g, S_k) = 0 \quad , \quad -M \leq k \leq N \quad , \tag{2.33}$$

provides the defining relationship which determines the coefficients $\{u_j\}$ of (2.32). The use of u in (2.33) as opposed to u_m is a convenience which facilitates the derivation of the discrete system. The weighted inner product (2.27) with $\sigma = -1/2$ assumes the form

$$(u, v) = \int_0^1 u(x)v(x)\big(\phi'(x)\big)^{-1/2} dx \quad . \tag{2.34}$$

In [9], it was shown that this selection produces a symmetric discrete system. The initial step in constructing the discrete system is achieved by setting

$$0 = (u'' - g, S_k) = \int_0^1 [u''(x) - g(x)] S_k(x) [\phi'(x)]^{-1/2} dx$$

and is followed by two applications of integration by parts. This leads to the equality

$$\int_0^1 g(x) S_k(x) \big(\phi'(x)\big)^{-1/2} dx = B_T + \int_0^1 u(x) \left[S_k(x) \big(\phi'(x)\big)^{-1/2} \right]'' dx \tag{2.35}$$

where the boundary term B_T is given by

$$B_T \equiv u'(x) S_k(x) \big(\phi'(x)\big)^{-1/2} \Big|_0^1$$

$$- u(x) \left[\frac{d}{d\phi} \big(S_k(x)\big) \big(\phi'(x)\big)^{1/2} - \frac{1}{2} S_k(x) \big(\phi'(x)\big)^{-3/2} \big(\phi''(x)\big) \right] \Big|_0^1 \quad .$$

If the function u satisfies the inequality

$$\left|u(x)\left(\phi'(x)\right)^{1/2}\right| \leq L \begin{cases} x^{\alpha} & , \ 0 < x < 1/2 \\ (1-x)^{\beta} & , \ 1/2 \leq x < 1 \end{cases} \tag{2.36}$$

then with the weight function selected, $B_T = 0$ if $\alpha > 0$ and $\beta > 0$.

Applying the quadrature rule of Theorem 2.2 to the integrals in (2.35) and truncating the infinite sum according to Theorem 2.3 produces

$$0 = h \sum_{j=-M}^{N} \left\{ \left[\frac{1}{h^2}\, \delta_{kj}^{(2)} - \frac{1}{4}\, \delta_{kj}^{(0)} \right] u(x_j)[\phi'(x_j)]^{1/2} - \frac{g(x_j)}{[\phi'(x_j)]^{3/2}}\, \delta_{kj}^{(0)} \right\} \quad . \tag{2.37}$$

Let $I^{(2)}$ denote the $m \times m$ ($m = M + N + 1$) matrix whose kjth element is given by (2.13) and define

$$A \equiv D[\phi'] \left[\frac{1}{h^2}\, I^{(2)} - \frac{1}{4}\, I^{(0)} \right] D[\phi'] \quad . \tag{2.38}$$

The matrix $D[\phi']$ is an $m \times m$ diagonal matrix with elements $\phi'(x_j)$, $j = -M, \ldots, N$ and $I^{(0)}$ is the $m \times m$ identity matrix.

If an approximate solution of (2.6) is (2.32) and if one replaces $u(x_j)$ with u_j in (2.37) then the discrete Sinc-Galerkin system for (2.6) is

$$A\vec{v} = D[(\phi')^{-1/2} g]\vec{1} \tag{2.39}$$

where $\vec{v} = D[(\phi')^{-1/2}]\vec{u}$, $\vec{u} = (u_{-M}, \ldots, u_0, \ldots, u_N)^T$ and $\vec{1} = (1, \ldots, 1)^T$.

The specific parameter selections made in (2.39) are defined by

$$h = \left(\pi d/(\alpha M)\right)^{1/2} \tag{2.40}$$

and

$$N = \left\lceil \frac{\alpha}{\beta}\, M + 1 \right\rceil \quad . \tag{2.41}$$

It has been shown in [9] that, asymptotically, the error incurred by approximating the true solution of (2.6) by (2.32) has the order $0\Big(\exp\big(-(\pi d\alpha M)^{1/2}\big)\Big)$. As shown in [12], the matrix $I^{(2)}$ in (2.38) is a negative definite symmetric matrix with condition number of order $0\big((N+M+2)/2\big)^2$.

The weight $(\phi')^{-1/2}$ in (2.34) produces a symmetric system to determine the coefficients of the approximation to the solution of (2.6) unlike the selection $(\phi')^{-1}$ used in [12]. For a general ϕ, the convergence rate is less rapid when the weight function selected is $(\phi')^{-1/2}$. For the map ϕ in (2.28) it was shown in [9] that the convergence rate is the same whether $(\phi')^{-1/2}$ or $(\phi')^{-1}$ is used. This plays a role for the problem (2.1) in the parameter selection of Chapter 4.

CHAPTER 3

THE DISCRETE SYSTEM FOR THE TEMPORAL DOMAIN

The Sinc-Galerkin method for problems (2.6) and (2.7) differs in two major respects, the conformal map and the choice of the weight used in the inner product. The difference in conformal maps is due to the obvious difference in domains, $(0,1)$ versus $(0,\infty)$. The selection of the weight for problem (2.7) is not as obvious. This chapter is devoted to the development of the discrete system for (2.7) and is carried out with an arbitrary weight in the form indicated in (2.27). More detail will be devoted to the selection of σ both with respect to the elimination of the boundary terms in the integration by parts and in the conditioning of the discrete system.

Recall problem (2.7)

$$\begin{aligned} u'(t) &= g(t)\ ,\ t>0 \\ u(0) &= 0 \quad . \end{aligned} \tag{3.1}$$

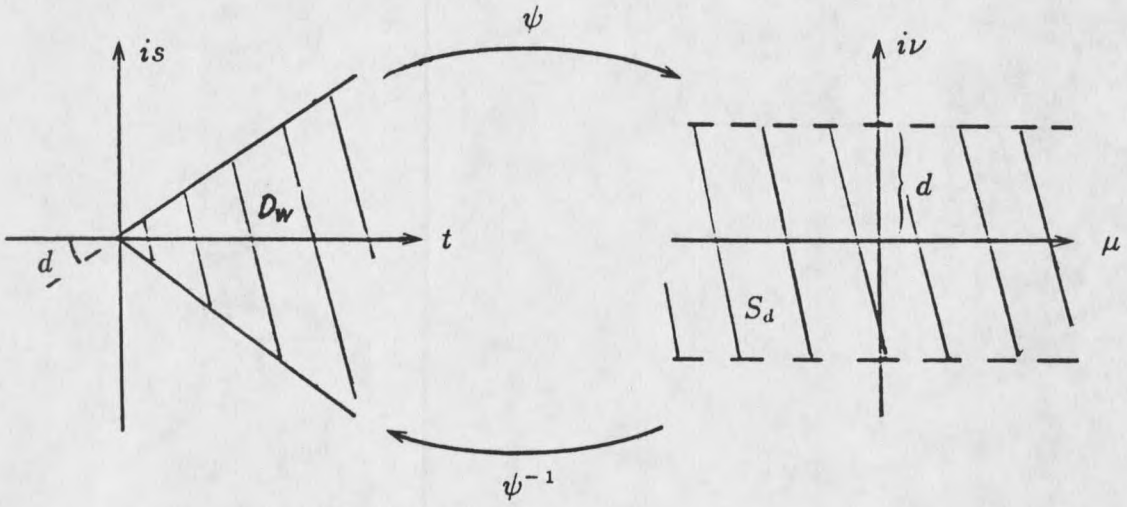
Select the conformal map

$$\psi(z) = \ell n\, z \tag{3.2}$$

which carries the wedge shaped region $D_W = \{z = t + is : |arg(z)| < d \leq \pi/2\}$ onto the strip S_d (2.8). The domains D_W and S_d are depicted graphically in Figure 3. The sinc nodes in D_W are defined by

$$t_k = \psi^{-1}(kh) = e^{kh} \quad . \tag{3.3}$$

Figure 3. The domains D_W and S_d.



The basis functions $S_k^*(t)$ are defined as in (2.9) with ϕ and x replaced by ψ and t respectively. An approximate solution of (3.1) is given by

$$u_m(t) = \sum_{j=-M}^{N} u_j S_j^*(t) \quad , \quad m = M + N + 1 \quad . \tag{3.4}$$

The weighted inner product for (3.1) assumes the form

$$(u, v) = \int_0^\infty u(t)v(t)\big(\psi'(t)\big)^\sigma \, dt \tag{3.5}$$

where σ is real and as yet unspecified. Orthogonalizing the residual using the inner product in (3.5) yields, after a single application of integration by parts,

$$\begin{aligned} 0 = (u' - g, S_k^*) &= \int_0^\infty [u'(t) - g(t)] S_k^*(t)\big(\psi'(t)\big)^\sigma \, dt \\ &= B_t - \int_0^\infty \left\{ u(t)\left[S_k^*(t)\big(\psi'(t)\big)^\sigma\right]' + g(t)S_k^*(t)\big(\psi'(t)\big)^\sigma \right\} dt \end{aligned} \tag{3.6}$$

where

$$B_t = u(t)S_k^*(t)\big(\psi'(t)\big)^\sigma\Big|_0^\infty \quad . \tag{3.7}$$

Again, u is used in (3.6) instead of u_m to derive the discrete system.

If it is assumed that u vanishes at a sufficiently rapid rate at infinity and that u is Lip γ at $t = 0$ then, coupled with the initial condition in (3.1), $B_t = 0$ for all $\sigma < \gamma$. The assumed vanishing of u at infinity is not a restriction for most parabolic problems and will be discussed in what follows. Taking the derivative of the bracketed term in (3.6) yields the equality

$$\begin{aligned} 0 &= \int_0^\infty -\Bigg\{ u(t)\left[\frac{d}{d\psi}\left(S_k^*(t)\right)[\psi'(t)]^{\sigma+1} + S_k^*(t)\sigma\big(\psi'(t)\big)^{\sigma-1}\psi''(t)\right] \\ &\qquad +g(t)S_k^*(t)[\psi'(t)]^\sigma \Bigg\} dt \\ &= \int_0^\infty -\Bigg\{ \left[\frac{d}{d\psi}\left(S_k^*(t)\right)\psi'(t) + \frac{\sigma\psi''(t)}{\psi'(t)}\,S_k^*(t)\right] u(t)[\psi'(t)]^\sigma \\ &\qquad +g(t)S_k^*(t)[\psi'(t)]^\sigma \Bigg\} dt \quad . \end{aligned} \tag{3.8}$$

Applying the sinc quadrature (Theorem 2.2) to the integral in (3.8) yields

$$0 = -h\sum_{j=-\infty}^{\infty}\left\{\left[\frac{1}{h}\,\delta_{kj}^{(1)} + \frac{\sigma\psi_j''}{(\psi_j')^2}\,\delta_{kj}^{(0)}\right] u_j(\psi_j')^\sigma\right\} - hg_k(\psi_k')^{\sigma-1} + I_u + I_g \tag{3.9}$$

where the following notation has been introduced: $f_j^{(i)}$ represents the ith derivative of f evaluated at the sinc node t_j. The elements $\delta_{kj}^{(1)}$ and $\delta_{kj}^{(0)}$ are as defined in (2.12) and (2.11), respectively. The integrals I_u and I_g are defined by (2.18) with f replaced by $u[S_k^*(\psi')^\sigma]'$ and $gS_k^*(\psi')^\sigma$, respectively.

The boundary integrals I_u and I_g are bounded using (2.20) and the following inequalities:

$$\left|\frac{S_k^*(w)}{\sin\big(\pi\psi(w)/h\big)}\right|\Bigg|_{w\in\partial\mathcal{D}_w} \le \frac{h}{\pi d} \equiv C_0(h,d) \tag{3.10}$$

and

$$\left|\frac{\frac{d}{d\psi}\left(S_k^*(w)\right)}{\sin\left(\pi\psi(w)/h\right)}\right|\Bigg|_{w\in\partial\mathcal{D}_w} \leq \frac{\pi d+\tanh(\pi d/h)}{\pi d^2\tanh(\pi d/h)} \equiv C_1(h,d) \quad . \tag{3.11}$$

In the interest of completeness, the bound on the equality $|I_u|$ proceeds as follows:

$$\begin{aligned}
|I_u| &= \left|\frac{i}{2}\int_{\partial\mathcal{D}_w} \frac{K(\psi,h)(\zeta)u(\zeta)\{S_k^*(\zeta)[\psi'(\zeta)]^\sigma\}'}{\sin\left(\pi\psi(\zeta)/h\right)}\,d\zeta\right| \\
&\leq \frac{1}{2}\left|\int_{\partial\mathcal{D}_w}\left\{\frac{K(\psi,h)(\zeta)u(\zeta)\frac{d}{d\psi}\left(S_k^*(\zeta)\right)[\psi'(\zeta)]^{\sigma+1}}{\sin\left(\pi\psi(\zeta)/h\right)}\right.\right. \\
&\qquad \left.\left.+\frac{K(\psi,h)(\zeta)u(\zeta)S_k^*(\zeta)\sigma[\psi'(\zeta)]^{\sigma-1}\psi''(\zeta)}{\sin\left(\pi\psi(\zeta)/h\right)}\right\}\,d\zeta\right| \\
&\leq \frac{1}{2}\exp\left(\frac{-\pi d}{h}\right)\left\{C_1(h,d)\int_{\partial\mathcal{D}_w}\left|u(\zeta)[\psi'(\zeta)]^{\sigma+1}d\zeta\right|\right. \\
&\qquad \left.+|\sigma|C_0(h,d)\int_{\partial\mathcal{D}_w}\left|u(\zeta)[\psi'(\zeta)]^{\sigma-1}\psi''(\zeta)d\zeta\right|\right\} \quad ,
\end{aligned}$$

where it is assumed that the final integrals are bounded as in Theorem 2.1.

To truncate the infinite sum in (3.9) suppose that $u(\psi')^\sigma$ is analytic in the domain $\mathcal{D}$ and satisfies

$$|u(t)[\psi'(t)]^\sigma| \leq K\begin{cases} t^\gamma & , \ 0<t<1 \\ t^{-\delta} & , \ t\geq 1 \end{cases} \tag{3.12}$$

for positive constants K, γ and δ (assumption (2.22) with f/ϕ' replaced by $u(\psi')^{1+\sigma}$). Employing the results of Theorem 2.3 on the infinite sum in (3.9)

yields the inequality

$$\left| -h \sum_{j=-M}^{N} \beta_{kj} u_j (\psi_j')^{\sigma} - h g_k (\psi_k')^{\sigma-1} \right|$$

$$\leq K \left\{ \frac{|\beta|}{\gamma} \exp(-\gamma M h) + \frac{|\beta|}{\delta} \exp(-\delta N h) \right\}$$

$$+ \frac{1}{2} \exp\left(\frac{-\pi d}{h}\right) \left\{ C_1(h,d) \int_{\partial D_w} \left| u(\zeta) [\psi'(\zeta)]^{\sigma+1} d\zeta \right| \right. \tag{3.13}$$

$$+ |\sigma| C_0(h,d) \int_{\partial D_w} \left| u(\zeta) [\psi'(\zeta)]^{\sigma-1} \psi''(\zeta) d\zeta \right|$$

$$\left. + C_0(h,d) \int_{\partial D_w} |g(\zeta) [\psi'(\zeta)]^{\sigma} d\zeta| \right\} \quad .$$

In (3.13) the quantity

$$\beta_{kj} \equiv \frac{1}{h} \delta_{kj}^{(1)} + \frac{\sigma \psi_j''}{[\psi_j']^2} \delta_{kj}^{(0)} \tag{3.14}$$

$$= \frac{1}{h} \delta_{kj}^{(1)} - \sigma \delta_{kj}^{(0)}$$

since, from (3.2), $\psi_j''/[\psi_j']^2 = -1$ for all j. The quantity $|\beta|$ on the right-hand side of (3.13) is, from (2.12) and (3.14), bounded by

$$|\beta| \leq \left(\frac{1}{h} + |\sigma| \right) \quad .$$

Assuming the boundary integrals on the extreme right of (3.13) are bounded, the exponential error terms in (3.13) are asymptotically balanced when the parameters h, M, N, γ and δ satisfy $\pi d/h = \gamma M h = \delta N h$. This yields the selections

$$h = \left(\pi d/(\gamma M)\right)^{1/2} \tag{3.15}$$

and

$$N = \left[\!\left[\frac{\gamma M}{\delta} + 1 \right]\!\right] \quad . \tag{3.16}$$

Before identifying the final form of the discrete system given by the left-hand side of (3.13), it should be noted that the calculation of N in (3.16) only assumes that the solution u decays algebraically at infinity, (3.12). For many parabolic problems the solution u satisfies

$$u(t) = 0\big(\exp(-\delta t)\big) \quad \text{as} \quad t \to \infty \tag{3.17}$$

for some $\delta > 0$. In this case, the upper sum is bounded by

$$\sum_{j=N+1}^{\infty} |\beta_{kj} u_j (\psi_j')^{\sigma}| \le \left(\frac{1}{h} + |\sigma|\right) \sum_{j=N+1}^{\infty} \exp(-\delta t_j) \quad . \tag{3.18}$$

Upon noting that

$$\begin{aligned} h \sum_{j=N+1}^{\infty} \exp(-\delta t_j) &\le \int_{Nh}^{\infty} e^{-\delta e^t} dt \\ &= \int_{e^{Nh}}^{\infty} e^{-\delta u} \frac{du}{u} \\ &\le \int_{e^{Nh}}^{\infty} e^{-\delta u} du \\ &= \frac{1}{\delta} e^{-\delta e^{Nh}} \quad , \end{aligned}$$

the truncation of the sum on the right-hand side of (3.18) is bounded by

$$\left(\frac{1}{h} + |\sigma|\right) \frac{1}{h\delta} \exp(-\delta e^{Nh}) \quad . \tag{3.19}$$

In this case, the truncation errors are asymptotically balanced if

$$N = \left[\!\!\left[\frac{\ell n(\gamma M h/\delta)}{h} + 1 \right]\!\!\right] \tag{3.20}$$

as opposed to (3.16). As the examples in Chapter 5 will show, this choice of N significantly reduces the dimension of the discrete system while maintaining the exponential convergence rate. If (3.4) is assumed to be an approximate solution of

(3.1) then, since $u_m(t_j) = u_j$, the discrete Sinc-Galerkin system which determines the $\{u_j\}$ is given by

$$B\vec{v} = -D[(\psi')^{\sigma} g]\vec{1} \tag{3.21}$$

where

$$B = D[\psi']\big((1/h)I^{(1)} - \sigma I^{(0)}\big) \quad . \tag{3.22}$$

$I^{(1)}$ is the $m \times m$ matrix whose kjth element is defined by (2.12) and $I^{(0)}$ is the $m \times m$ identity matrix. The vector $\vec{v}$ is defined by

$$\vec{v} = D[(\psi')^{\sigma}]\vec{u} \tag{3.23}$$

where $\vec{u} = (u_{-M}, u_{-M+1}, \ldots, u_N)^T$ and the matrices $D[f]$ are $m \times m$ diagonal matrices with diagonal elements $f(t_j)$, $j = -M, \ldots, N$.

A unique solution to (3.21) is guaranteed provided B is nonsingular. This may be seen as follows: The matrix $I^{(1)}$ is a real skew-symmetric matrix and therefore has spectrum that is pure imaginary. Hence for all non-zero σ the matrix $(1/h)I^{(1)} - \sigma I^{(0)}$ is invertible. Since ψ is conformal the matrix $D[\psi']$ is invertible so that B, as the product of invertible matrices, is invertible.

As important as the invertibility of B in (3.22) is its distribution of eigenvalues. This plays a role both in the conditioning of B as well as the method of solution employed in the two-dimensional problem (detailed in Chapter 4). Indeed this author believes that the eigenvalues of B are distinct. There is ample numerical evidence indicating that B is diagonalizable but an analytic proof remains elusive. Another concern which surfaces in Chapter 4 is how the eigenvalues of A in (2.38) and B intermingle. As will be seen, the solution requires a division by

the sum of the eigenvalues of B with the eigenvalues of A (in all possible combinations). Recalling that A is a negative symmetric definite matrix, it is required that no real eigenvalue of B is the additive inverse of an eigenvalue of A.

The approach taken is to show that the real part of each eigenvalue of B is negative provided an adroit selection of σ is made. To this end, let C be defined by

$$C = \frac{1}{h} I^{(1)} - \sigma I \quad . \tag{3.24}$$

The matrix $I^{(1)}$ is a skew-symmetric Toeplitz matrix with eigenvalues in the interval $(-i\pi, i\pi)$, thus the eigenvalues of C are translates of those of $I^{(1)}$ where the translation is dependent on the sign of σ. Notice that the matrix $D(1/\sqrt{\psi'})BD(\sqrt{\psi'})$ has the same spectrum as $D(\sqrt{\psi'})CD(\sqrt{\psi'})$. It follows from the definition of C that if $\lambda \in \sigma(C)$ then $\mathrm{Re}(\lambda) < 0$ or $\mathrm{Re}(\lambda) > 0$ as $\sigma > 0$ or $\sigma < 0$, respectively. Hence, the same is true of any spectral point of $D(1/\sqrt{\psi'})BD(\sqrt{\psi'})$ and since the latter is a similarity transformation of B it follows that if $\beta \in \sigma(B)$ then $\mathrm{Re}(\beta) < 0$ if $\sigma > 0$ or $\mathrm{Re}(\beta) > 0$ if $\sigma < 0$. Hence, selecting $\sigma > 0$ guarantees that the real part of the eigenvalues of A and B have the same sign, so that no eigenvalue of A is the additive inverse of an eigenvalue of B. The following theorem has been shown.

Theorem 3.1: Let the matrices A and B be defined as in (2.38) and (3.22) respectively, and let $\beta \in \sigma(B)$ and $\alpha \in \sigma(A)$. If $\sigma > 0$ then

$$\mathrm{Re}(\alpha + \beta) < 0 \quad .$$

In light of the discussion following (3.7), if u is Lip γ at $t = 0$ then $B_t = 0$ for all σ where $0 < \sigma < \gamma$. Consequently, one should select a reasonably small σ. The

selection of σ also impacts the convergence rate of the method. Rewriting (3.12) in terms of a bound on u yields

$$|u(t)| \leq K \begin{cases} t^{\gamma+\sigma} & , \ 0 < t < 1 \\ \exp(-\delta t) & , \ t \geq 1 \end{cases} , \tag{3.25}$$

where (3.17) is being assumed. The truncation error term leading to the selection (3.20) assumes the form $\exp\big(-(\gamma+\sigma)Mh\big)$. Hence, to speed the convergence of the method one should select σ "close to" γ. As will be seen in the examples of Chapter 5, $\gamma = 1$ is a common case so that the selection $\sigma = 1 - \epsilon$, $\epsilon > 0$, is optimal with respect to the preceeding sentence. However, besides optimizing the convergence rate, the selection of σ plays a competing role in the conditioning of the matrix B in (3.22).

To begin this investigation note that the matrix C in (3.24) is normal and the computation

$$\begin{aligned} C^T C &= \left(-\frac{1}{h}\, I^{(1)} - \sigma I\right)\left(\frac{1}{h}\, I^{(1)} - \sigma I\right) \\ &= -\frac{1}{h^2}\,(I^{(1)})^2 + \sigma^2 I \end{aligned}$$

shows that the eigenvalues of $C^T C$ are real and lie in the interval $(\sigma^2, \sigma^2 + \pi^2/h^2)$. The condition number of C is then

$$\begin{aligned} \kappa(C) &\equiv \|C\|_2 \|C^{-1}\|_2 \\ &= \sqrt{\rho(C^T C)} \cdot \sqrt{\rho\big((C^{-1})^T C^{-1}\big)} \\ &\leq \sqrt{\sigma^2 + \frac{\pi^2}{h^2}}\ \sqrt{\frac{1}{\sigma^2}} \\ &= \frac{1}{\sigma}\sqrt{\sigma^2 + 2\gamma M} \end{aligned} \tag{3.26}$$

where the last equality follows from (3.15) with $d = \pi/2$. Using (3.26) and the

identity $\psi'(t_j) = e^{-jh}$ there follows

$$\kappa\big(D((\psi')^\sigma)B\big) = \|D\big((\psi')^{\sigma+1}\big)C\|_2 \|D\big((\psi')^{-\sigma-1}\big)C^{-1}\|_2$$

$$\leq \|C\|_2 \|C^{-1}\|_2 \|D\big((\psi')^{\sigma+1}\big)\|_2 \|D\big((\psi')^{-\sigma-1}\big)\|_2 \tag{3.27}$$

$$\leq \frac{1}{\sigma}\sqrt{\sigma^2 + 2\gamma M}\,\max\left\{e^{2Nh(\sigma+1)},\ e^{2Mh(\sigma+1)}\right\} \quad .$$

Hence for positive σ the right-hand side of (3.27) is minimized for σ "near" zero (as opposed to $\sigma = \gamma$ as in the discussion following (3.25)). In the interest of balancing the competing inequalities (3.25) and (3.27) governing the selection of σ this author, with the common case $\gamma = 1$ in mind, selects their average value $\sigma = 1/2$.

CHAPTER 4

ASSEMBLY OF THE ONE-DIMENSIONAL SYSTEMS

Recall the parabolic partial differential equation:

$$Pu(x,t) \equiv u_t(x,t) - u_{xx}(x,t) = g(x,t) \ ; \ (x,t) \in D$$

$$u(x,t) = 0 \ , \quad (x,t) \in \partial D \tag{4.1}$$

$$D = \{(x,t) : 0 < x < 1 \ , \ 0 < t < \infty\} \quad ,$$

and the associated Sinc-Galerkin approximation of u

$$u_{m_x,m_t}(x,t) = \sum_{i=-M_x}^{N_x} \sum_{j=-M_t}^{N_t} u_{ij} S_i(x) S_j^*(t) \quad . \tag{4.2}$$

The subscripts x and t are used to identify parameters of the discrete system that are similar in nature. For example, the dimension of the coefficient matrix and the mesh size h will be subscripted since they depend on the independent variables x and t.

The weighted inner product (2.5) applied to (2.4) yields

(4.3)

$$0 = \int_0^\infty \int_0^1 [u_t(x,t) - u_{xx}(x,t) - g(x,t)] S_k(x) S_\ell^*(t) [\phi'(x)]^{-1/2} [\psi'(x)]^{1/2} dx dt \quad .$$

As in the one-dimensional problems, u is used in (2.4) instead of u_{m_x,m_t} as a matter of convenience. Proceeding as in the one-dimensional problems by performing the necessary integration by parts and applying the quadrature rule (Theorem 2.2) to the result, the following discretization for (4.3) is obtained:

$$
\begin{aligned}
0 = & \sum_{j=-\infty}^{\infty} \left\{ \left[\frac{1}{h_t}\, \delta_{\ell j}^{(1)} - \frac{1}{2}\, \delta_{\ell j}^{(0)} \right] [\phi_k']^{-3/2} u_{kj} [\psi_j']^{1/2} \right\} \\
& + \sum_{i=-\infty}^{\infty} \left\{ \left[\frac{1}{h_x^2}\, \delta_{ki}^{(2)} - \frac{1}{4}\, \delta_{ki}^{(0)} \right] [\phi_i']^{1/2} u_{i\ell} [\psi_\ell']^{-1/2} \right\} \\
& + [\phi_k']^{-3/2} g_{k\ell} [\psi_\ell']^{-1/2} \quad .
\end{aligned} \tag{4.4}
$$

Assuming that u satisfies a growth bound of the form

$$
|u(x,t)| \le K\, t^{\gamma+1/2} e^{-\delta t} x^{\alpha+1/2} (1-x)^{\beta+1/2} \quad , \tag{4.5}
$$

the infinite sums in (4.4) can be truncated without impingement to the convergence rate of the method. To transform (4.4) into a finite dimensional matrix equation, suppose that $-M_x \le k, i \le N_x$ and $-M_t \le j, \ell \le N_t$. The discrete Sinc-Galerkin system can then be written as follows:

$$
\begin{aligned}
& D[(\phi_k')^{-3/2}] U D[(\psi_\ell')^{1/2}] \left[\frac{1}{h_t}\, (I^{(1)})^T - \frac{1}{2}\, I \right] \\
& + \left[\frac{1}{h_x^2}\, I^{(2)} - \frac{1}{4}\, I \right] D[(\phi_k')^{1/2}] U\, D[(\psi_\ell')^{-1/2}] \\
& = -D[(\phi_k')^{-3/2}] G\, D[(\psi_\ell')^{-1/2}]
\end{aligned} \tag{4.6}
$$

where

$$
U = \begin{bmatrix}
u_{-M_x,-M_t} & u_{-M_x,-M_t+1} & \cdots & u_{-M_x,N_t} \\
u_{-M_x+1,-M_t} & u_{-M_x+1,-M_t+1} & \cdots & u_{-M_x+1,N_t} \\
\vdots & \vdots & & \vdots \\
\vdots & \vdots & \ddots & \vdots \\
u_{N_x,-M_t} & u_{N_x,-M_t+1} & \cdots & u_{N_x,N_t}
\end{bmatrix} .
$$

The order of the coefficients, u_{ij}, in the first column of U results from fixing t at the first time node, t_{-M_t}, and letting k and i assume each value from $-M_x$ to N_x in (4.4). Subsequent columns of U are generated by advancing to the next time node and repeating the process. The order of the elements of G in (4.4) is precisely

that of U. This particular ordering was precipitated by mimicking the approach used in a standard time-differencing scheme. It should be noted that the Sinc-Galerkin method is not bound to any specific ordering of the grid. Additionally, it should be noted that the concept of stability, inherent in a time-differencing scheme, does not apply to the Sinc-Galerkin method in the standard form of the eigenvalue growth of an iteration matrix. In the present setting the approximate solution u_{m_x,m_t} is computed at all times t_j. The stability is "controlled" via the selection of σ outlined at the end of Chapter 3.

To obtain a more convenient form for the discrete system (4.6), multiply (4.6) on the left and right by the diagonal matrices $D[(\phi')]$ and $D[(\psi')]$ respectively and rewrite the matrix $D[(\phi')^{1/2}]$ in the second term on the left of (4.6) as the product $D[(\phi')]D[(\phi')^{-1/2}]$. The final form of the discrete system is then

$$
\begin{aligned}
& D[(\phi'_k)^{-1/2}]UD[(\psi'_\ell)^{1/2}]\left\{\left[-\frac{1}{h_t}\, I^{(1)} - \frac{1}{2}\, I\right] D[(\psi'_\ell)]\right\} \\
& + \left\{D[(\phi'_k)]\left[\frac{1}{h_x^2}\, I^{(2)} - \frac{1}{4}\, I\right] D[(\phi'_k)]\right\} D[(\phi'_k)^{-1/2}]U\, D[(\psi'_\ell)^{1/2}] \\
& = -D[(\phi'_k)^{-1/2}]G\, D[(\psi'_\ell)^{1/2}] \quad .
\end{aligned} \tag{4.7}
$$

One could argue the purported "convenience" of (4.7) over (4.6), however, the dispersement of the derivative of the conformal maps and associated weights exhibited in (4.7) provide a sense of "symmetry" to the Sinc-Galerkin method. The matrix equation (4.7) can be written as

$$VB^T + AV = F \tag{4.8}$$

where B is defined in (3.22), A is defined in (2.38),

$$V = D[(\phi'_k)^{-1/2}]U\, D[(\psi'_\ell)^{1/2}] \tag{4.9}$$

and

$$F = D[(\phi'_k)^{-1/2}]G\, D[(\psi'_\ell)^{1/2}] \quad . \tag{4.10}$$

The solvability of the system in (4.8) can be characterized in terms of the intermixing of the eigenvalues of A with those of B. This characterization is a well known result from linear algebra but is revealed in the first method of solution presented for solving the system in (4.8).

Algebraically, (4.8) is equivalent to the system

$$\mathcal{A}\vec{v} \equiv \{I \otimes A + B \otimes I\} \overrightarrow{\mathrm{co}(V)} = \overrightarrow{\mathrm{co}(F)} \tag{4.11}$$

where the tensor product of an $m \times n$ matrix E with a $p \times q$ matrix G is defined by

$$E \otimes G \equiv [e_{ij} G]_{mp \times nq} \quad .$$

The vector $\vec{v} \equiv \overrightarrow{\mathrm{co}(V)}$ is the concatenation of the $m_x \times m_t$ matrix V obtained by successively "stacking" the columns of V, one upon another, to obtain an $m_x m_t \times 1$ vector. It is known [8] that the eigenvalues of $\mathcal{A}$ in (4.11) are given by the sums

$$\{\alpha_i + \beta_j\}_{\substack{-M_x \le i \le N_x \\ -M_t \le j \le N_t}}$$

so that the system (4.8) has a unique solution if and only if no eigenvalue of A is the negative of an eigenvalue of B. Theorem 3.1 implies then that (4.8) is uniquely solvable.

A direct method of solving the system (4.11) consists of the implementation of any of the decomposition methods available for linear systems. From a computational point of view, a more economical method of solution proceeds as follows: Since A is orthogonally diagonalizable there exists an orthogonal matrix Q so that

$$Q^T A Q = \Lambda_\alpha \tag{4.12}$$

where Λ_α is a diagonal matrix containing the eigenvalues α_i of A. Define the $m_x m_t \times m_x m_t$ matrix

$$S = I \otimes Q$$

(note I is $m_t \times m_t$). Factoring S and S^T from the matrix $\mathcal{A}$ in (4.11) and pre-multiplying the result by S^T produces

$$(4.13) \qquad (I \otimes \Lambda_\alpha + B \otimes I)(I \otimes Q^T)\,\overrightarrow{\mathrm{co}(V)} = (I \otimes Q^T)\,\overrightarrow{\mathrm{co}(F)} \quad .$$

The system in (4.12) may be solved either by a block Gauss elimination method or via a block Gauss Jordan routine.

A second solution method for the system (4.8) is obtained by a diagonalization procedure. It is the method of solution originally used by Stenger [12] to solve system (4.8). Specifically, assuming that B in (4.8) is diagonalizable then there is an invertible matrix P so that

$$(4.14) \qquad P^{-1} B^T P = \Lambda_\beta$$

where Λ_β is a diagonal matrix containing the eigenvalues β_j of B^T. If the change of variables

$$(4.15) \qquad W = Q^T V P$$

and

$$(4.16) \qquad H = Q^T F P$$

is made in (4.8) the result is

$$(4.17) \qquad QWP^{-1}B^T + AQWP^{-1} = QHP^{-1} \quad .$$

Multiplying (4.17) on the left by Q^T and right by P yields

$$(4.18) \qquad W\Lambda_\beta + \Lambda_\alpha W = H \quad .$$

The solution of (4.18) in component form is

$$w_{ij} = \frac{h_{ij}}{\beta_j + \alpha_i} \tag{4.19}$$

for $-M_x \leq i \leq N_x$ and $-M_t \leq j \leq N_t$. Once W is determined, V is recovered via the relationship defined in (4.15) and U follows from (4.9).

Due to Theorem 3.1 the w_{ij} in (4.19) are well defined and the solution U of (4.7) is obtained via the outlined procedure, the only analytic constraint to the method being the assumed diagonalizability of B in (4.14). In the interest of supporting the conjectured diagonalizability of B, the examples of the next chapter are all computed via the procedure outlined in (4.14) through (4.19). To motivate the conjectured diagonalization of B consider the following intuitive discussion: The matrix hB is the product of a positive diagonal matrix $\left(D(\psi')\right)$ and a normal matrix $\left(I^{(1)} - \frac{h}{2} I\right)$. The matrix $D(\psi')I^{(1)}$ may be written as $D(\sqrt{\psi'})\left[D\sqrt{\psi'}\, I^{(1)} D(\sqrt{\psi'})\right] D(1/\sqrt{\psi'})$ hence the eigenvalues of $D(\psi')I^{(1)}$ are the same as the eigenvalues of the normal matrix $D(\sqrt{\psi'})I^{(1)}D(\sqrt{\psi'})$. As an $0(h)$ perturbation of a diagonalizable matrix, it seems reasonable that for h sufficiently small, hB and hence B is diagonalizable.

An alternative procedure, based on no diagonalizability assumptions, consists of replacing (4.14) by the Schur decomposition of B [5]. That is, replace (4.14) by

$$B^T = ZTZ^T \tag{4.20}$$

where Z is orthogonal and T is upper triangular. The discrete system (4.8) is now equivalent to

$$VZT + AVZ = FZ \quad . \tag{4.21}$$

In (4.21) use the diagonalization of A in (4.12) to transform (4.21) to

$$WT + \Lambda W = Q^T FZ = M \tag{4.22}$$

where

$$W = Q^T V Z \quad .$$

The components of w_{ij} are given recursively by

$$\sum_{p=1}^{n} w_{ip} t_{pj} + \alpha_i w_{ij} = m_{ij} \tag{4.23}$$

where, for convenience, it is assumed that all matrices are $n \times n$ and indexed from 1 to n. Also $t_{pj} = 0$ for all $p > j$. Notice that (4.23) yields the same solution as the diagonalization procedure beginning with (4.14) since (4.8) has a unique solution.

CHAPTER 5

NUMERICAL IMPLEMENTATION

The Sinc-Galerkin method for approximating the solution of the parabolic problem (4.1) was tested on several examples with known solutions. The numerical results for seven distinct problems are presented here. Problem dependent parameters are identified and the results for each problem are listed in a tabular format which reveals the exponential rate of convergence of the method. Example 1, whose solution is analytic, is included to compare with Examples 2 and 3 which have singular solutions. The latter two examples converge at the same or a faster rate than that of Example 1. As emphasized in Chapters 2 and 3, it is the Lipschitz character of the solution that governs the rate of convergence of the Sinc-Galerkin method. Indeed, Examples 2, 3 and 4 demonstrate the method's ability to approximate singular solutions whether the singularity is logarithmic or algebraic (in space and/or time) or both. Examples 5 and 6 are taken from the literature. The last two examples demonstrate a technique for transforming a parabolic problem with non-zero boundary conditions to a problem in the form of (4.1).

Recall that the approximate solution of (4.1) is defined by

$$u_{m_x,m_t}(x,t) = \sum_{i=-M_x}^{N_x} \sum_{j=-M_t}^{N_t} u_{ij} S_i(x) S_j^*(t) \tag{5.1}$$

where the basis elements $S_i(x)$ and $S_j^*(t)$ are defined in (2.9) with ϕ replaced by (2.28) and (3.2), respectively. The coefficients, $\{u_{ij}\}$, are determined by solving

the discrete system

$$VB^T + AV = F \tag{5.2}$$

for V then recovering the u_{ij} from the v_{ij} by way of the relationship (4.9). The matrices V and F are of dimension $m_x \times m_t$ while B^T is $m_t \times m_t$ and A is $m_x \times m_x$. The technique used to solve (5.2) is outlined in Chapter 4. Based on the discussion following (4.19) the solution technique is based on the diagonalization procedure in (4.14)–(4.19).

The stepsizes h_x and h_t and limits of summation M_x, N_x, M_t and N_t are selected so that the errors in each coordinate direction are asymptotically balanced. For solutions $u(\cdot,t)$ analytic in the right half-plane ($\mathcal{D}_W$ with $d = \pi/2$) the contour integral error I_u in (3.9) is $0\big(\exp(-\pi^2/(2h_t))\big)$. The contour integral error for solutions $u(x,\cdot)$ analytic in the disc ($\mathcal{D}_E$ with $d = \pi/2$) is $0\big(\exp(-\pi^2/(2h_x))\big)$. Consequently the contour integral errors are asymptotically balanced provided

$$h_t = h_x \quad . \tag{5.3}$$

The conditions (2.36) and (3.25) translate to the bound

$$|u(x,t)| \leq Kt^{\gamma+1/2}e^{-\delta t}x^{\alpha+1/2}(1-x)^{\beta+1/2} \tag{5.4}$$

on the solution u of (4.1). If the selections

$$h_x = \left(\frac{\pi^2}{2\alpha M_x}\right)^{1/2} \quad , \tag{5.5}$$

$$N_x = \left[\!\left[\frac{\alpha M_x}{\beta} + 1\right]\!\right] \quad , \tag{5.6}$$

$$M_t = \left[\!\left[\frac{\alpha M_x}{\gamma} + 1\right]\!\right] \tag{5.7}$$

and

$$N_t = \left[\!\left[\frac{1}{h_t}\, \ell n \left(\frac{\gamma M_t h_t}{\delta} \right) + 1 \right]\!\right] \tag{5.8}$$

are made, then the errors incurred by truncating the infinite sums are asymptotically balanced with those arising from the contour integrals. The balance is maintained whenever $\frac{\alpha}{\beta} M_x$ and $\frac{\alpha}{\gamma} M_x$ are integers. In practice, the addition of 1 in (5.6) and/or (5.7) is required only if $\frac{\alpha}{\beta} M_x$ and/or $\frac{\alpha}{\gamma} M_x$ is not an integer. The six terms contributing to the error all have the form

$$ACR \equiv \exp\left(-(\pi\sqrt{\alpha M_x/2})\right) \quad . \tag{5.9}$$

This asymptotic convergence rate provides a measure of the accuracy of the computed solution (4.1) to the true solution of (4.1).

The methodology used in each example was to first determine the parameters α, β, γ and δ from (5.4), then for a given M_x, determine h, N_x, M_t and N_t from the relationships (5.5), (5.6), (5.7) and (5.8), respectively. Then the approximation (5.1) is constructed using the technique of Chapter 4.

To demonstrate the exponential convergence rate of the method, approximations were constructed for a sequence of values of M_x on each problem. Each approximation was compared to the true solution at the sinc nodes

$$(x_i, t_j) = \left(e^{ih}/(e^{ih}+1)\,,\, e^{jh}\right) \tag{5.10}$$

and the maximum absolute error was recorded. Each table identifies the mesh size used, $h \equiv h_x = h_t$, the lower and upper summation limits M_x, N_x, M_t and N_t and the maximum error

$$E_u \equiv \underset{\substack{-M_x \le i \le N_x \\ -M_t \le j \le N_t}}{\text{maximum}} |u_{ij} - u(x_i, t_j)|$$

at the sinc nodes (5.10).

In order to gage the sharpness of the asymptotic convergence rate in (5.9) a subset of the examples include in their display a column corresponding to

$$NAC \equiv E_u / ACR \quad . \tag{5.11}$$

This "numerical asymptotic constant" should, as M_x increases, remain "relatively" constant if the asymptotic convergence rate in (5.10) is an accurate expression for the convergence rate of the method. The displays in Examples 1, 3 and 5 numerically verify this accuracy. These three examples correspond to the cases when $\alpha = 1/2$, 1 and $3/2$, respectively. In particular, it is of interest to note that although the solution in Example 1 is analytic whereas the solution in Example 2 is singular, the latter converges more rapidly. This corresponds to the fact that the α in (5.10) is larger for Example 2 than it is for Example 1. Similar remarks regarding increased accuracy apply in comparing the results of Example 3 with Example 5. These three examples highlight the remarks from the introduction about the Sinc-Galerkin method's "permanence" of convergence in the presence of singularities.

All computations were performed in double precision using the ANS FORTRAN-77 compiler on a Honeywell Level 66 computer. The notation $.xy(-d)$ denotes $.xy \times 10^{-d}$ in all examples.

Example 1. If one sets

$$g(x,t) = e^{-t}[x(1-t)(1-x) + 2t]$$

in (4.1), the analytic solution is

$$u(x,t) = te^{-t}x(1-x) \quad . \tag{5.12}$$

From (5.4) one determines that $\alpha = \beta = \gamma = 1/2$ and $\delta = 1$. The numerical results for runs with $M_x = 4, 8, 16$ and 32 are listed in Table 1. The values of N_t listed in Table 1 were determined by (5.8). The dimension of matrices V and F in (5.2) is 9×6, 17×11, 33×20 and 65×37 for the given sequence of M_x. Had N_t been determined from relation (3.16), the dimension would have been 9×8, 17×14, 33×26 and 65×50.

Table 1. Numerical results for Example 1.

h	M_x	N_x	M_t	N_t	E_u	AEC
1.5708	4	4	4	1	0.92(-3)	.21(-1)
1.1107	8	8	8	2	0.27(-3)	.23(-1)
0.7854	16	16	16	3	0.84(-4)	.45(-1)
0.5554	32	32	32	4	0.71(-5)	.51(-1)

Example 2. For this problem, set

$$g(x,t) = t^{1/2}e^{-t}\left[\left(\frac{3}{2} - t\right)x(1-x) + 2t\right] \quad .$$

The singular solution of (4.1) is

$$u(x,t) = t^{3/2}e^{-t}x(1-x) \quad .$$

Here, u has an algebraic singularity at $t = 0$ yet the numerical results listed in Table 2 exhibit the anticipated exponential rate of convergence. For this problem, the values of the parameters α, β, γ and δ are 1/2, 1/2, 1 and 1 respectively.

Table 2. Numerical results for Example 2.

h	M_x	N_x	M_t	N_t	E_u
1.5708	4	4	2	2	0.17(-2)
1.1107	8	8	4	3	0.45(-3)
0.7854	16	16	8	4	0.40(-4)
0.5554	32	32	16	7	0.11(-5)

Example 3. If one sets

$$g(x,t) = t^{1/2}e^{-t}x^{-1/2}(1-x)^{-1/2}\left[\left(\frac{3}{2}-t\right)x^2(1-x)^2 - \frac{3}{4}\,t(8x^2-8x+1)\right]$$

then the solution of (4.1) is

$$u(x,t) = t^{3/2}e^{-t}[x(1-x)]^{3/2} \quad .$$

In this problem, u has algebraic singularities at each of $t = 0$, $x = 0$ and $x = 1$. The parameters α, β, γ and δ are readily identified to be $\alpha = \beta = \gamma = 1$ and $\delta = 1$. Table 3 contains the numerical results for each of $M_x = 4, 8, 16$ and 32. Examining the maximum error for each run shows that in spite of the singularities, the exponential convergence rate is maintained. Note that the convergence rate for this example is greater than it was for Example 1 as per the comments following (5.11).

Table 3. Numerical results for Example 3.

h	M_x	N_x	M_t	N_t	E_u	AEC
1.1107	4	4	4	3	0.36(-3)	.31(-1)
0.7854	8	8	8	5	0.66(-4)	.35(-1)
0.5554	16	16	16	9	0.81(-5)	.58(-1)
0.3927	32	32	32	17	0.22(-6)	.63(-1)

Example 4. If in (4.1)

$$g(x,t) = t^{3/2}e^{-t}\left[(3/(2t)-1)x\,\ell n\,x - \frac{1}{x}\right]$$

then the singular solution is given by

$$u(x,t) = t^{3/2}e^{-t}x\,\ell n\,x \quad .$$

Note the algebraic singularity at $t = 0$ and the logarithmic singularity at $x = 0$. The parameters used are $\alpha = \beta = 1/2$ and $\gamma = \delta = 1$. The results for this problem are listed in Table 4.

Table 4. Numerical results for Example 4.

h	M_x	N_x	M_t	N_t	E_u
1.5708	4	4	2	1	0.40(-2)
1.1107	8	8	4	2	0.11(-2)
0.7854	16	16	8	3	0.21(-3)
0.5554	32	32	16	4	0.26(-4)

The results presented in Tables 2 through 4 emphasize the fact that the exponential convergence rate of the Sinc-Galerkin method is maintained in spite of a singular solution. It is the parameters α, β, γ and δ that govern the rate of convergence, not the existence and boundedness of higher order derivatives.

Example 5. In this problem

$$g(x,t) = (1-t)e^{-t}x^2(1-x)^2 - 2te^{-t}(1-6x+6x^2)$$

the analytic solution of (4.1) is then

$$u(x,t) = te^{-t}x^2(1-x)^2 \quad .$$

A slight variation of this problem was considered by Hlava'cek [6] and by Fairweather and Saylor [4]. The values of α, β, γ and δ are readily identified in the solution: they are $\alpha = \beta = 3/2$, $\gamma = 1/2$ and $\delta = 1$. The numerical results for $M_x = 2, 4, 8$ and 16 are listed in Table 5.

Table 5. Numerical results for Example 5.

h	M_x	N_x	M_t	N_t	E_u	AEC
1.2826	2	2	6	2	0.22(-3)	.10(-1)
0.9069	4	4	12	2	0.57(-4)	.13(-1)
0.6413	8	8	24	4	0.50(-5)	.11(-1)
0.4535	16	16	48	6	0.21(-6)	.11(-1)

Example 6. Evans and Abdullah [3] considered the problem

$$w_t = w_{xx} - w_x$$

$$w(0,t) = w(1,t) = 0 \tag{5.13}$$

$$w(x,0) = \exp(x/2)x(1-x) \quad .$$

Two simple transformations are required to obtain a problem in the form of (4.1). The change of variables $v(x,t) = \exp(t/4 - x/2)w(x,t)$ yields the problem

$$\begin{aligned} v_t - v_{xx} &= 0 \\ v(0,t) = v(1,t) &= 0 \\ v(x,0) &= x(1-x) \end{aligned} \tag{5.14}$$

which has the solution

$$v(x,t) = (8/\pi^3)\sum_{k=0}^{\infty}(2k+1)^{-3}\exp\bigl(-(2k+1)^2\pi^2 t\bigr)\sin[(2k+1)\pi x] \quad .$$

This problem was also considered by Joubert [7] (replace v by $4v$). To get zero boundary conditions set

$$u(x,t) = v(x,t) - e^{-2t}x(1-x) \quad .$$

This yields a problem in the form of (4.1) with

$$g(x,t) = -2e^{-2t}(x^2 - x + 1) \quad .$$

From the Sinc-Galerkin approximation to this problem, one can approximate the solutions of (5.13) and (5.14) by back substitution. For example

$$v(x_i,t_j) \approx v_{ij} = u_{ij} + \exp(-2t_j)x_i(1-x_i)$$

where (x_i,t_j) are the sinc nodes found in (2.31) and (3.3), respectively.

The parameters used for the Sinc-Galerkin method on this problem are $\delta = 2$ and $\alpha = \beta = \gamma = 1/2$. The results are listed in Table 6.

Table 6. Numerical results for Example 6.

h	M_x	N_x	M_t	N_t	E_u
1.5708	4	4	4	1	0.14(-1)
1.1107	8	8	8	1	0.22(-2)
0.7854	16	16	16	2	0.29(-3)
0.5554	32	32	32	3	0.29(-4)

A table to illustrate the error distribution at different time levels has been included for the run with $h = 0.7854$. Table 7 shows the maximum absolute errors at $t = 0.2079$, 1.0 and 4.8105 for a sample of the x grid points as listed. It should be remembered that the numerical solution (5.1) is actually global and these data are given only as an indication of the actual time levels generated by the Sinc-Galerkin method and the observed error at these time levels. Note the maximum absolute error of 0.29(-3) occurs at $x_0 = 0.5$ and $t_0 = 1$ which is the "center" sinc node in the grid. Also the error is symmetric about $x = 0.5$ at each time level. This may be seen in the last three columns of Table 7.

Table 7. Error for Example 6 at time levels $t = 0.2079$, 1 and 4.8105 when $h = 0.7854$.

		E_u		
i	x_i	$t_{-2} = 0.2079$	$t_0 = 1$	$t_2 = 4.8105$
-16	0.35(-5)	0.39(-5)	0.10(-5)	0.10(-7)
-12	0.81(-4)	0.39(-5)	0.10(-5)	0.10(-7)
-8	0.18(-2)	0.43(-5)	0.27(-5)	0.51(-6)
-4	0.41(-1)	0.13(-4)	0.38(-4)	0.11(-4)
0	0.5	0.94(-4)	0.29(-3)	0.91(-4)
4	0.958576	0.13(-4)	0.38(-4)	0.11(-4)
8	0.998136	0.43(-5)	0.27(-5)	0.51(-6)
12	0.999919	0.39(-5)	0.10(-5)	0.10(-7)
16	0.999996	0.39(-5)	0.10(-5)	0.10(-7)

Example 7. Consider the problem

$$w_t = w_{xx}$$

$$w(0,t) = w(1,t) = 0 \tag{5.15}$$

$$w(x,0) = \sin \pi x \quad .$$

To transform (5.15) to a problem in the form of (4.1), set

$$u(x,t) = w(x,t) - e^{-4t} \sin \pi x \tag{5.16}$$

yielding the problem

$$u_t - u_{xx} = (4 - \pi^2) e^{-4t} \sin \pi x$$

$$u(0,t) = u(1,t) = u(0,x) = 0 \tag{5.17}$$

which has the solution

$$u(x,t) = \left(e^{-\pi^2 t} - e^{-4t}\right) \sin \pi x \quad .$$

The parameters α, β and γ can be obtained by examining the Taylor series of u expanded about $t = 0$, $x = 0$ and $x = 1$, respectively. One arrives at the values $\alpha = \beta = \gamma = 1$ and $\delta = 4$. The numerical results are listed in Table 8.

Table 8. Numerical results for Example 7.

h	M_x	N_x	M_t	N_t	E_u
1.5708	4	4	4	0	0.10(-1)
1.1107	8	8	8	1	0.33(-2)
0.7854	16	16	16	1	0.23(-3)
0.5554	32	32	32	2	0.13(-4)

REFERENCES CITED

[1] J.F. Botha and G.F. Pinder, *Fundamental Concepts in the Numerical Solution of Differential Equations*, Wiley, New York, 1983.

[2] K. Bowers, J. Lund and K. McArthur, "The sinc method in multiple space dimensions: Model problems," to appear in *Numerische Mathematik*.

[3] D.J. Evans and A.R. Abdullah, "Boundary value techniques for finite element solutions of the diffusion-convection equation," *International Journal for Numerical Methods in Engineering*, 23, 1145–1159 (1986).

[4] G. Fairweather and A.V. Saylor, "On the application of extrapolation, deferred correction and defect correction to discrete-time Galerkin methods for parabolic problems," *IMA Journal of Numerical Analysis*, 3, 173–192, (1983).

[5] G. Golub and C. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, 1983.

[6] I. Hlava'cek, "On a semi-variational method for parabolic equations," *Institute for Applied Mathematics*, 17, 327–351 (1972).

[7] G.R. Joubert, "Explicit difference approximations of the one-dimensional diffusion equation, using a smoothing technique," *Numerische Mathematik*, 17, 409–430 (1971).

[8] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Academic Press, Orlando, Florida, 1985.

[9] J. Lund, "Symmetrization of the Sinc-Galerkin method for boundary value problems," *Mathematics of Computation*, 47, 571–588 (1986).

[10] W.L. Seward, G. Fairweather and R.L. Johnson, "A survey of higher-order methods for the numerical integration of semidiscrete parabolic problems," *IMA Journal of Numerical Analysis*, 4, 375–425 (1984).

[11] F. Stenger, "Approximations via Whittaker's cardinal function," *Journal of Approximation Theory*, 17, 222–240 (1976).

[12] F. Stenger, "A 'Sinc-Galerkin' method of solution of boundary value problems," *Mathematics of Computation*, 33, 85–109 (1979).

[13] F. Stenger, "Numerical methods based on Whittaker cardinal, or sinc functions," *SIAM Review*, 23, 165–224 (1981).

[14] F. Stenger, "Sinc methods of approximate solution of partial differential equations," *IMACS Proceedings on Advances in Computer Methods for Partial Differential Equations*, 244–251 (1984).

[15] J.R. Yu and T.R. Hsu, "On the solution of diffusion-convection equations by the space-time finite element method," *International Journal for Numerical Methods in Engineering*, 23, 737–750 (1986).

[16] E.T. Whittaker, "On the functions which are represented by the expansions of the interpolation theory," *Proceedings of the Royal Society of Edinburgh*, 35, 181–194 (1915).