



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

© Copyright by David Lamar Lewis (1989)

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

R. F. Fialant
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



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

	Page
LIST OF TABLES	vi
LIST OF FIGURES	vii
ABSTRACT	viii
1. INTRODUCTION	1
2. THE DISCRETE SYSTEM FOR THE SPATIAL DOMAIN	9
3. THE DISCRETE SYSTEM FOR THE TEMPORAL DOMAIN	20
4. ASSEMBLY OF THE ONE-DIMENSIONAL SYSTEMS	30
5. NUMERICAL IMPLEMENTATION	37
REFERENCES CITED	49

LIST OF TABLES

Table	Page
1. Numerical results for Example 1	41
2. Numerical results for Example 2	42
3. Numerical results for Example 3	43
4. Numerical results for Example 4	43
5. Numerical results for Example 5	44
6. Numerical results for Example 6	46
7. Error for Example 6 at time levels $t = 0.2079, 1$ and 4.8105 when $h = 0.7854$	47
8. Numerical results for Example 7	48

LIST OF FIGURES

Figure	Page
1. The domains \mathcal{D} and S_d	11
2. The domains \mathcal{D}_E and S_d	16
3. The domains \mathcal{D}_W and S_d	21

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

$$(1.1) \quad Lu = g$$

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

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

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

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

A standard weighted inner product will be used, given by

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

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

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

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,

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

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 $\text{sinc}(x)$, which is defined by

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

The approximation procedures herein are based on the following formal expansion

of a function f :

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

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 $O(\exp -\kappa\sqrt{N})$ 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 $O(\exp -\kappa\sqrt{N})$ 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

$$(1.8) \quad \mathcal{A}\vec{u} = \vec{b}$$

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

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

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] germane to the development are included for completeness. The chapter concludes with the symmetric Sinc-Galerkin method applied to the one-dimensional

problem

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

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

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

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

$$(2.1) \quad \begin{aligned} Pu(x,t) &\equiv u_t(x,t) - u_{xx}(x,t) = g(x,t) \quad ; \quad (x,t) \in D \\ u(x,t) &= 0 \quad , \quad (x,t) \in \partial D \end{aligned}$$

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

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

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

where

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

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

$$(2.4) \quad (Pu_{m_x, m_t} - g, S_k S_\ell^*) = 0 \quad , \quad \begin{aligned} k &= -M_x, \dots, 0, \dots, N_x \\ \ell &= -M_t, \dots, 0, \dots, N_t \end{aligned}$$

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

$$(2.5) \quad (u, v) = \int_0^\infty \int_0^1 u(x, t)v(x, t)w(x, t) dx dt$$

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

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

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

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

and

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

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.

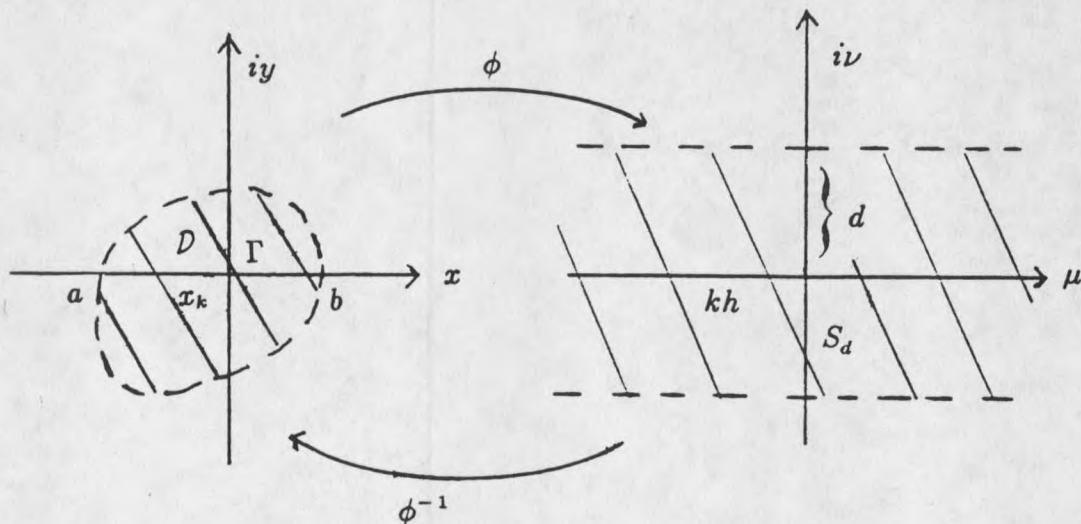
Before attacking either problem (2.6) or (2.7), the foundation and basic assumptions necessary for the Sinc-Galerkin method are presented. Since the domain of concern differs in problems (2.6) and (2.7), the preliminary results are developed on the general interval (a, b) with $a \neq b$.

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

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

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

Figure 1. The domains \mathcal{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 :

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

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

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

so that, in particular,

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

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

and

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

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

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

where

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

and

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

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

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

and

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

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

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

where

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

Furthermore, a short calculation produces the identity

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

and the inequality

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

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

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

where

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

and

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

Then

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

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

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

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

$$(2.25) \quad h = \sqrt{\frac{\pi d}{\alpha M}}$$

and

$$(2.26) \quad N = \left\lceil \frac{\alpha M}{\beta} + 1 \right\rceil$$

where $\lceil \cdot \rceil$ 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

$$(2.27) \quad (u, v) = \int_a^b u(x)v(x)(\phi'(x))^\sigma dx$$

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

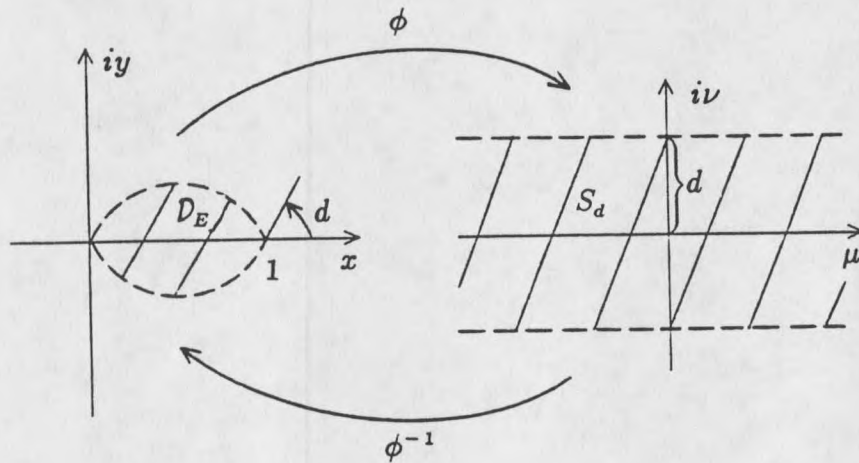
$$(2.28) \quad \phi(z) = \ln \left(\frac{z}{1-z} \right)$$

which carries the "eye"-shaped region

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

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

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

thus the sinc nodes are

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

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

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

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

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

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

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

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

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

where the boundary term B_T is given by

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

If the function u satisfies the inequality

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

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

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

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

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

The matrix $D[\phi']$ is an $m \times m$ diagonal matrix with elements $\phi'(x_j)$, $j = -M, \dots, N$ 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

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

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

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

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

and

$$(2.41) \quad N = \left\lceil \frac{\alpha}{\beta} M + 1 \right\rceil$$

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 $O\left(\exp(-(\pi d\alpha M)^{1/2})\right)$. As shown in [12], the matrix $I^{(2)}$ in (2.38) is a negative definite symmetric matrix with condition number of order $O\left((N + M + 2)/2\right)^2$.

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)

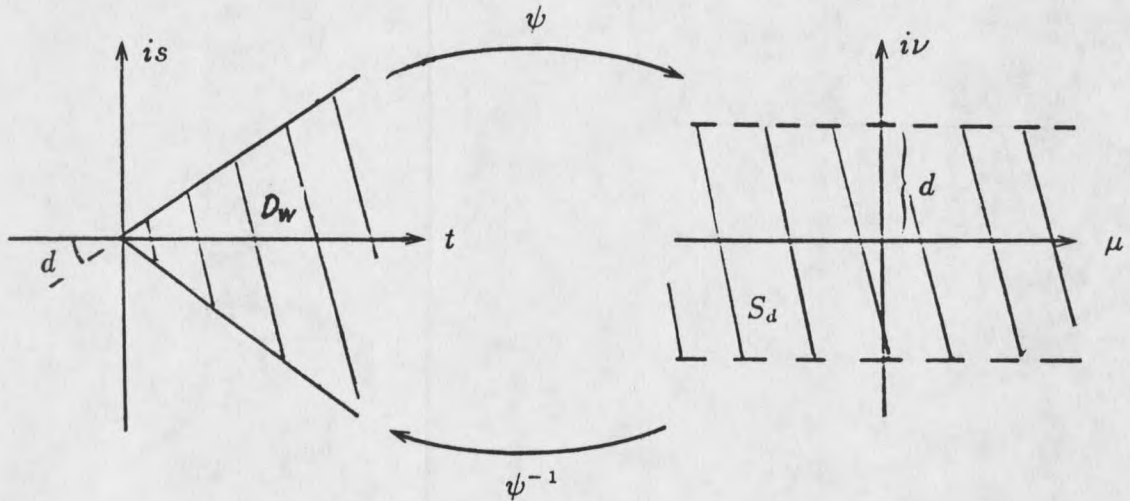
$$(3.1) \quad \begin{aligned} u'(t) &= g(t), \quad t > 0 \\ u(0) &= 0 \end{aligned}$$

Select the conformal map

$$(3.2) \quad \psi(z) = \ell n z$$

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

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

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

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

The weighted inner product for (3.1) assumes the form

$$(3.5) \quad (u, v) = \int_0^\infty u(t)v(t)(\psi'(t))^\sigma dt$$

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,

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

where

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

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

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

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

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

where the following notation has been introduced: $f_j^{(i)}$ represents the i th 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:

$$(3.10) \quad \left| \frac{S_k^*(w)}{\sin(\pi\psi(w)/h)} \right| \Big|_{w \in \partial D_w} \leq \frac{h}{\pi d} \equiv C_0(h, d)$$

and

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

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

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

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

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

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

$$\begin{aligned}
 & \left| -h \sum_{j=-M}^N \beta_{kj} u_j (\psi'_j)^\sigma - h g_k (\psi'_k)^{\sigma-1} \right| \\
 & \leq K \left\{ \frac{|\beta|}{\gamma} \exp(-\gamma M h) + \frac{|\beta|}{\delta} \exp(-\delta N h) \right\} \\
 (3.13) \quad & + \frac{1}{2} \exp\left(\frac{-\pi d}{h}\right) \left\{ C_1(h, d) \int_{\partial D_w} |u(\zeta) [\psi'(\zeta)]^{\sigma+1} d\zeta| \right. \\
 & + |\sigma| C_0(h, d) \int_{\partial D_w} |u(\zeta) [\psi'(\zeta)]^{\sigma-1} \psi''(\zeta) d\zeta| \\
 & \left. + C_0(h, d) \int_{\partial D_w} |g(\zeta) [\psi'(\zeta)]^\sigma d\zeta| \right\}.
 \end{aligned}$$

In (3.13) the quantity

$$\begin{aligned}
 (3.14) \quad \beta_{kj} & \equiv \frac{1}{h} \delta_{kj}^{(1)} + \frac{\sigma \psi_j''}{[\psi_j']^2} \delta_{kj}^{(0)} \\
 & = \frac{1}{h} \delta_{kj}^{(1)} - \sigma \delta_{kj}^{(0)}
 \end{aligned}$$

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

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

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

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

and

$$(3.16) \quad N = \left\lceil \frac{\gamma M}{\delta} + 1 \right\rceil.$$

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

$$(3.17) \quad u(t) = O(\exp(-\delta t)) \quad \text{as } t \rightarrow \infty$$

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

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

Upon noting that

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

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

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

In this case, the truncation errors are asymptotically balanced if

$$(3.20) \quad N = \left\lceil \left\lfloor \frac{\ln(\gamma M h / \delta)}{h} + 1 \right\rfloor \right\rceil$$

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

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

where

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

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

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

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

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

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

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 $\text{Re}(\lambda) < 0$ or $\text{Re}(\lambda) > 0$ 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 $\text{Re}(\beta) < 0$ if $\sigma > 0$ or $\text{Re}(\beta) > 0$ 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

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

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

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

where (3.17) is being assumed. The truncation error term leading to the selection (3.20) assumes the form $\exp(-(\gamma + \sigma)Mh)$. 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 preceding 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

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

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

$$\begin{aligned}
 \kappa(D((\psi')^\sigma)B) &= \|D((\psi')^{\sigma+1})C\|_2 \|D((\psi')^{-\sigma-1})C^{-1}\|_2 \\
 (3.27) \qquad &\leq \|C\|_2 \|C^{-1}\|_2 \|D((\psi')^{\sigma+1})\|_2 \|D((\psi')^{-\sigma-1})\|_2 \\
 &\leq \frac{1}{\sigma} \sqrt{\sigma^2 + 2\gamma M} \max \left\{ e^{2Nh(\sigma+1)}, e^{2Mh(\sigma+1)} \right\} .
 \end{aligned}$$

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:

$$(4.1) \quad \begin{aligned} Pu(x, t) &\equiv u_t(x, t) - u_{xx}(x, t) = g(x, t) ; (x, t) \in D \\ u(x, t) &= 0 \quad , \quad (x, t) \in \partial D \\ D &= \{(x, t) : 0 < x < 1, 0 < t < \infty\} \quad , \end{aligned}$$

and the associated Sinc-Galerkin approximation of u

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

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

$$(4.3) \quad 0 = \int_0^\infty \int_0^1 [u_t(x, t) - u_{xx}(x, t) - g(x, t)] S_k(x) S_l^*(t) [\phi'(x)]^{-1/2} [\psi'(t)]^{1/2} dx dt \quad .$$

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:

$$\begin{aligned}
(4.4) \quad 0 = & \sum_{j=-\infty}^{\infty} \left\{ \left[\frac{1}{h_t} \delta_{\ell j}^{(1)} - \frac{1}{2} \delta_{\ell j}^{(0)} \right] [\phi'_k]^{-3/2} u_{kj} [\psi'_j]^{1/2} \right\} \\
& + \sum_{i=-\infty}^{\infty} \left\{ \left[\frac{1}{h_x^2} \delta_{ki}^{(2)} - \frac{1}{4} \delta_{ki}^{(0)} \right] [\phi'_i]^{1/2} u_{i\ell} [\psi'_\ell]^{-1/2} \right\} \\
& + [\phi'_k]^{-3/2} g_{k\ell} [\psi'_\ell]^{-1/2} .
\end{aligned}$$

Assuming that u satisfies a growth bound of the form

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

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 $-M_x \leq k, i \leq N_x$ and $-M_t \leq j, \ell \leq N_t$. The discrete Sinc-Galerkin system can then be written as follows:

$$\begin{aligned}
(4.6) \quad & D[(\phi'_k)^{-3/2}] U D[(\psi'_\ell)^{1/2}] \left[\frac{1}{h_t} (I^{(1)})^T - \frac{1}{2} I \right] \\
& + \left[\frac{1}{h_x^2} I^{(2)} - \frac{1}{4} I \right] D[(\phi'_k)^{1/2}] U D[(\psi'_\ell)^{-1/2}] \\
& = -D[(\phi'_k)^{-3/2}] G D[(\psi'_\ell)^{-1/2}]
\end{aligned}$$

where

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

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

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

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

$$(4.8) \quad VB^T + AV = F$$

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

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

and

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

