

Extensions to the development of the Sinc-Galerkin method for parabolic problems
by Randy Ross Doyle

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Mathematics
Montana State University
© Copyright by Randy Ross Doyle (1990)

Abstract:
A Galerkin method in both spatial and temporal domains is developed for the parabolic problem. The development is carried out in detail in the case of one spatial dimension. The discretization in the case of two spatial dimensions along with numerical implementation is also carried out. The basis functions for the Galerkin methods are the sinc functions (composed with conformal maps) which, in conjunction with the highly accurate sinc function quadrature rules, form the foundation of a very powerful numerical method for the parabolic problem. The spectral analysis of the discrete sinc system receives close attention and, in particular, is shown to be uniquely solvable. This is the case for either of the conformal mappings used in the temporal domain. A consequence of this spectral analysis provides the motivation for the iterative method of solution of the discrete system. This iterative solution method is numerically tested not only on linear problems, but also on Burgers' (a nonlinear) problem.

EXTENSIONS TO THE DEVELOPMENT OF THE SINC-GALERKIN METHOD FOR PARABOLIC PROBLEMS

by

Randy Ross Doyle

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

of

Doctor of Philosophy

in

Mathematics

MONTANA STATE UNIVERSITY
Bozeman, Montana

May 1990

D378
D777

APPROVAL

of a thesis submitted by

Randy Ross Doyle

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.

May 21, 1990
Date

John Lund
Chairperson, Graduate Committee

Approved for the Major Department

May 21, 1990
Date

K J Tiahrt
Head, Major Department

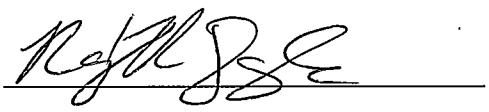
Approved for the College of Graduate Studies

May 24, 1990
Date

Henry J 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 May 15, 1990

ACKNOWLEDGMENTS

The author would like to acknowledge gratefully:

Dr. Norman Eggert, Dr. Kenneth Bowers and Dr. John Lund, whose combined talents in editing have improved the content of this thesis beyond measure.

Rene' Tritz, for her competent skill in producing lovely final documents from handwritten scribbles.

The entire staff and my fellow students of the Montana State University Department of Mathematical Sciences, who were sources of help, encouragement and stimulus.

Jo Pierson, whose continual support and understanding led directly to the completion of this thesis.

Dr. Robert Engle, who extended much needed motivation during hard times.

And especially Dr. John Lund, whose patience has been strained to the breaking point in the process of the creation of this thesis, and yet he never let that interfere with his fairness and good advice. He has taught me a love of mathematics that transcends all of the coursework I have done.

TABLE OF CONTENTS

Page

LIST OF TABLES vi

LIST OF FIGURES vii

ABSTRACT viii

1. INTRODUCTION AND SUMMARY 1

2. SINC INTERPOLATION AND QUADRATURE 7

3. INNER PRODUCT APPROXIMATIONS 19

4. THE SINC-GALERKIN MATRIX SYSTEM 35

5. SPECTRAL ANALYSIS OF THE SINC-GALERKIN SYSTEM 50

6. THE NUMERICAL PROCEDURE AND EXAMPLES 57

REFERENCES CITED 73

LIST OF TABLES

Table — Page

1. Formulation of Definition 2.12 for the conformal maps of Chapter 3 32
2. Formulation of Definition 2.12 for the non-homogeneous terms in (3.1) 33
3. The spatial matrix $A(w)$ in (4.10) corresponding to $w = 1/\sqrt{\phi'}$ and $w = 1/\phi'$ 43
4. The temporal matrix $B(\hat{w})$ in (4.9) corresponding to $w = \sqrt{\hat{\phi}'}$ for both $\hat{\phi}(t) = \ell n(t)$ and $\hat{\phi}(t) = \ell n(\sinh(t))$ 44
5. Numerical Results for Example 1 61
6. Asymptotic and Iterative Error Analysis for Example 2 62
7. Numerical Results for Example 2 62
8. Numerical Results for Example 3 63
9. Illustration of Misuse of Method 64
10. Numerical Results for Example 4 64
11. Numerical Results for Example 5 using $\hat{\phi}(t) = \ell n(t)$ 66
12. Numerical Results for Example 5 using $\hat{\phi}(t) = \ell n(\sinh(t))$ 67
13. Numerical Results for Example 6 69
14. Numerical Results for Example 7 70

LIST OF FIGURES

Figure Page

1. The spatial map ϕ 33
2. The temporal map $\hat{\phi}$ 34
3. Iterative Convergence Rate for Example 4 65
4. Mesh Plot of the Solution of Example 5 Looking Back in Time 67
5. Mesh Plot of the Solution of Example 5 Looking Outward in Time . . . 68
6. Mesh Plot of the Computed Solution of Example 7 71
7. Contour Plot Taken from the Mesh Plot of Figure 6 72

ABSTRACT

A Galerkin method in both spatial and temporal domains is developed for the parabolic problem. The development is carried out in detail in the case of one spatial dimension. The discretization in the case of two spatial dimensions along with numerical implementation is also carried out. The basis functions for the Galerkin methods are the sinc functions (composed with conformal maps) which, in conjunction with the highly accurate sinc function quadrature rules, form the foundation of a very powerful numerical method for the parabolic problem. The spectral analysis of the discrete sinc system receives close attention and, in particular, is shown to be uniquely solvable. This is the case for either of the conformal mappings used in the temporal domain. A consequence of this spectral analysis provides the motivation for the iterative method of solution of the discrete system. This iterative solution method is numerically tested not only on linear problems, but also on Burgers' (a nonlinear) problem.

CHAPTER 1

INTRODUCTION AND SUMMARY

Parabolic partial differential equations arise in many applications of the physical and biological sciences. The equations for fluid flow in the boundary layer along a wall and for axi-symmetric flow in a channel give rise to parabolic partial differential equations. A description of bacterial population is given by Fisher's equation, a nonlinear parabolic partial differential equation, while another such equation models neutron population in a nuclear reactor. Both the advection-diffusion equation and Burgers' equation, used as model equations of the vorticity transport arising from the vorticity-stream function formulation of the incompressible Navier-Stokes equations, are nonlinear parabolic differential equations.

All of these parabolic problems in one spatial dimension are encompassed in the formulation

$$u_t(x,t) - \mu u_{xx}(x,t) = f(x,t,u) \quad , \tag{1.1}$$

where $\mu > 0$.

A standard approach in obtaining a numerical solution of the linear parabolic partial differential equation

$$\begin{aligned} &V_t(x,t) = \mu V_{xx}(x,t) + f(x,t,V) \ ; \ (x,t) \in \mathcal{D} \\ &V(0,t) = g_0(t) \\ &V(1,t) = g_1(t) \\ &V(x,0) = V_0(x) \\ &\mathcal{D} = \{(x,t) : 0 < x < 1; t > 0\} \end{aligned} \tag{1.2}$$

where

$$V_0(0) = g_0(0)$$

$$V_0(1) = g_1(0)$$

begins with a discretization of the spatial domain using a finite difference, finite element, collocation, or Galerkin method [1], [3], [4], [5] and [6]. The result of this approach is a system of ordinary differential equations in the time variable. One generally uses a time differencing scheme to complete the approximation. Low order methods with small time steps are required to obtain accurate approximations due to stability constraints on the discrete time evolution operator [6].

A great deal of effort has been expended in the development of methods which increase the accuracy of the temporal approximation. Seward, Fairweather, and Johnson [15] have classified methods of order greater than two for the time integration. Fairweather and Saylor [5] give methods which have up to fourth order accuracy in time. Each of these techniques possess certain advantages. However, each of these methods has a finite order of accuracy in time and so, when used with an infinite order spatial approximation (spectral, for example), the overall result is a method unbalanced in total order. Space-time finite element methods have been developed [4] which avoid time differencing, and yet when piecewise polynomial elements are used, again a finite order method results.

A method that is infinite in order in the time domain has been developed by Lewis, Lund, and Bowers [9] for the linear parabolic partial differential equation. There the fully Sinc-Galerkin discretization of (1.1) has the matrix form

$$\mathcal{A}\vec{u} = \vec{f} \tag{1.3}$$

where $\mathcal{A}$ is a tensor product of the spatial and temporal matrices associated with a Sinc-Galerkin approach to the one-dimensional problems taken from (1.1) by fixing t and x, respectively. Due to the manner in which the discrete system in

(1.3) was developed in [9], the authors there could only numerically justify the invertibility of the matrix $\mathcal{A}$. The work in this thesis removes this defect from the fully Sinc-Galerkin method. By rewriting the system (1.3) in a slightly different form, the invertibility of the coefficient matrix $\mathcal{A}$ will be established in Corollary 5.5.

The present work extends the results of Lewis, Lund, and Bowers in two other complementary directions. First is the development of a methodology to handle the more general problem (1.1) when f represents heat loss or gain

$$f = -cu + g$$

or convection

$$f = -cu_x \quad .$$

Secondly, it is shown in the linear case

$$f(x,t,u) = g(x,t) - cu(x,t) \quad ,$$

that the method, which is iterative, must converge, and an upper bound for the rate of convergence is exhibited. Finally, motivated by the convergence results for the iterative method on the linear problem, a first approach to nonlinear problems is developed using Burgers' equation as the prototype.

In the present work an infinite order, multiple space dimension, fully Galerkin method using sinc basis functions in all dimensions is developed. The problem is transformed as follows: If $V(x,t)$ solves system (1.2), let $U(x,\tau)$ be defined by

$$U(x,\tau) = V(x,t) - G(x,t) \tag{1.4}$$

where

$$G(x,t) = (1-x)g_0(t) + xg_1(t) + (V_0(x) - (1-x)V_0(0) - xV_0(1))\exp(-\tau)$$

and

$$\tau = \mu t \quad .$$

Then $U(x,\tau)$ solves the differential equation

$$U_\tau(x,\tau) + G_t(x,t)/\mu = U_{xx}(x,\tau) + G_{xx}(x,t) + f\left(x,t,U(x,\tau) + G(x,t)\right)/\mu \quad ,$$

where

$$t = \tau/\mu \quad .$$

This reduces to

$$\begin{aligned} & U_\tau(x,\tau) = U_{xx}(x,\tau) + F(x,\tau,U) \\ & U(0,\tau) = U(1,\tau) = U(x,0) = 0 \\ & 0 < x < 1 \quad , \quad \tau > 0 \end{aligned} \tag{1.5}$$

where

$$F(x,\tau,U) = G_{xx}(x,t) - G_t(x,t)/\mu + f\left(x,t,U(x,\tau) + G(x,t)\right)/\mu \quad .$$

The function $U(x,\tau)$ solving (1.5) is equivalent to the function $V(x,t)$ solving (1.2) via the change of variable in (1.4). This transformation of the problem allows for the use of sinc basis functions without the need to add special functions to handle initial and boundary conditions.

The organization of this thesis is summarized in the following remarks. Chapter 2 contains a brief review of some analytic and numerical properties of the sinc function. The former is included so as to assemble in one convenient source various results that are scattered in the literature. The approximate properties of the sinc function are largely taken from the extensive work in [17] but have here been limited and restricted to include the results germaine to the present thesis. Included, in particular, is the development of the sinc quadrature rules which

play such an important role in the approximation of the inner products for the Sinc-Galerkin method.

These inner product approximations are developed in Chapter 3. Not only is the time domain mapping function which was used in [9] examined, but also a different choice of mapping function is considered. This latter choice has provided a complementary alternative to numerical procedures for Sturm-Liouville eigenvalue problems [3], quadratures for various integral transforms [12] and the numerical solution of a (scalar) nonlinear ordinary differential equation [16]. In the context of the present thesis, this choice of mapping in the temporal domain also gives rise to a discrete system of the form (1.3) which is shown to be uniquely solveable.

In Chapter 4 the assembly of the inner product development for the problem (1.1) is carried out where the function f can represent simple non-homogeneous, heat loss (gain), or linear or nonlinear convection. In all cases the discrete system takes the form

$$\mathcal{A}\vec{u} = \vec{f} + m(\vec{u}) \tag{1.6}$$

where m is a linear or nonlinear (in the case of Burgers' equation) function of the vector $\vec{u}$. Whereas the former may be solved by direct methods an iterative procedure is defined and carried out in this thesis.

The foundation for this iterative procedure is a consequence of the spectral study of the matrix $\mathcal{A}$ which is the subject of Chapter 5. This chapter not only puts the fully Sinc-Galerkin method on a firm analytic footing (unique solvability of the discrete system), but also points (via the above mentioned iterative scheme) to a future direction for an efficient computational procedure in the case of nonlinear problems.

The final Chapter 6 gives the numerical results for seven examples. The examples are selected to illustrate the choice of the temporal domain mapping function, the various parameter selections and the exponential convergence rate of the method. Specifically, examples are included which are computed directly (a direct method is used to solve (1.6)) or iteratively (an iterative method is used to solve (1.6)). To illustrate the method's applicability in the case of more than one spatial variable, a two-dimensional example is also included. The final example is Burgers' well-known sine initial condition problem.

CHAPTER 2

SINC INTERPOLATION AND QUADRATURE

This thesis addresses the development and extension of the fully Sinc-Galerkin method which began in [16] and was refined in [9]. The building blocks of the method for (1.1) consist of the Galerkin approach coupled with the sinc function. Numerical sinc function methods are based on E.T. Whittaker's work [19] concerning interpolation of functions at the integers. The foundation of this work is the sinc function

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

The resulting formal cardinal expansion for a function $f(x)$ is defined by

$$\sum_{k=-\infty}^{\infty} f(k)\operatorname{sinc}(x-k) \quad .$$

In order to generalize the expansion to handle interpolation at any evenly spaced grid, define for $h > 0$

$$S(k,h)(x) = \operatorname{sinc}\left[\frac{x-kh}{h}\right] \tag{2.1}$$

and denote the Whittaker cardinal function of f by

$$C(f,h)(x) = \sum_{k=-\infty}^{\infty} f(kh)S(k,h)(x) \tag{2.2}$$

whenever this series converges.

With regard to using (2.2) as an interpolation tool, it is important to know for which functions the right-hand side of (2.2) converges, and when convergent,

to what it converges. In this direction, two important classes of functions will be identified. The first is the somewhat restrictive class for which (2.2) is exact. The second is the class of functions where the difference between $f(x)$ and $C(f,h)(x)$ is "small". J.M. Whittaker [20] and McNamee, Stenger and Whitney [13] identified these classes by displaying a natural link between the Whittaker series of f and the Fourier transform of f.

The Fourier transform of a function g is defined by

$$\mathcal{F}(g)(x) = \int_{-\infty}^{\infty} g(t)\exp(ixt)dt \quad . \tag{2.3}$$

A fundamental result of Fourier analysis is the inversion theorem. Specifically, if $g \in L^2(\mathbb{R})$ then there exists $\mathcal{F}(g) \in L^2(\mathbb{R})$ whereby g can be recovered from $\mathcal{F}(g)$ by the Fourier inversion integral

$$g(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathcal{F}(g)(x)\exp(-ixt)dx \quad .$$

If the analytic structure of g is taken into account, it is to be expected that more can be said about the Fourier transform of g. When g is an entire function this is a fundamental result due to Paley and Wiener [14, page 375].

Theorem 2.1 (The Paley-Wiener Theorem): *If $g \in L^2(\mathbb{R})$, is entire, and there are positive constants A and C so that for all $w \in \mathbb{C}$*

$$|g(w)| \leq C\exp(A|w|) \quad ,$$

then

$$g(w) = \frac{1}{2\pi}\int_{-A}^{A} \mathcal{F}(g)(x)\exp(-iwx)dx \quad .$$

To show that the sinc function obeys the hypotheses of the Paley-Wiener Theorem with $A = \pi$ and $C = 1$ is straightforward. A simple integration yields

the identity

$$\begin{aligned}\operatorname{sinc}(t) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp(-ixt)dx \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \chi_{(-\pi,\pi)}(x)\exp(-ixt)dx \quad .\end{aligned}$$

That is, the characteristic function $\chi_{(-\pi,\pi)}(x)$ for the interval $(-\pi,\pi)$ is the inverse Fourier transform of $\operatorname{sinc}(t)$. From this, a consequence of the inversion theorem is

$$\mathcal{F}(\operatorname{sinc})(x) = \chi_{(-\pi,\pi)}(x) = \int_{-\infty}^{\infty} \operatorname{sinc}(t)e^{ixt}dt \quad . \tag{2.4}$$

Hence, using (2.1), a change of variables, and (2.4) gives

$$\begin{aligned}\mathcal{F}(S(k,h))(x) &= \int_{-\infty}^{\infty} \operatorname{sinc}\left[\frac{t-kh}{h}\right]\exp(ixt)dt \\ &= h\exp(ixkh)\chi_{(-\pi/h,\pi/h)}(x) \quad .\end{aligned} \tag{2.5}$$

Hence, using the inversion theorem yields the useful identity

$$S(k,h)(x) = \frac{h}{2\pi}\int_{-\pi/h}^{\pi/h} \exp[-i(x-kh)t]dt \quad . \tag{2.6}$$

Definition 2.2 (The Paley-Wiener Class of Functions): *Let $B(h)$ be the set of functions g such that g is entire, $g \in L^2(\mathbb{R})$, and for all $w \in \mathbb{C}$,*

$$|g(w)| \leq C\exp\left(\frac{\pi}{h}|w|\right) \quad .$$

This is precisely the class where the representation of g by its cardinal series (2.2) is exact. To see this a few preliminary results are needed.

Theorem 2.3: *If $g \in B(h)$ then*

$$g(w) = \frac{1}{h}\int_{-\infty}^{\infty} g(t)\operatorname{sinc}\left[\frac{t-w}{h}\right]dt \quad .$$

Theorem 2.4: *If* $g \in L^2(\mathbb{R})$ *then*

$$k(w) = \frac{1}{h}\int_{-\infty}^{\infty} g(t)\text{sinc}\left[\frac{t-w}{h}\right]dt \quad .$$

is in $B(h)$.

A proof of Theorems 2.3 and 2.4 can be found in [17]. The above can now be combined in the following theorem.

Theorem 2.5: *If* $g \in B(h)$ *then*

$$g(w) = \sum_{k=-\infty}^{\infty} \alpha_k S(k,h)(w)$$

where

$$\alpha_k = \frac{1}{h}\int_{-\infty}^{\infty} g(t)\text{sinc}\left[\frac{t-kh}{h}\right]dt = g(kh) \quad .$$

The result can be stated in terms of the class of functions where the Whittaker cardinal expansion (2.2) is exact: If $g \in B(h)$ then $g(w) = C(f,h)(w)$ for all $w \in \mathbb{C}$. This result is stronger than that which was sought, and can be strengthened further by a result derived from Theorem 2.4 and the identity

$$\frac{1}{h}\int_{-\infty}^{\infty} S(k,h)(t)S(p,h)(t)dt = \delta_{pk} = \begin{cases} 1 & , \ p = k \\ 0 & , \ p \neq k \end{cases} \quad .$$

This identity follows upon replacing g in Theorem 2.3 by $S(k,h)(t)$ and recalling that $S(k,h)$ is in $B(h)$.

Theorem 2.6: *The set*

$$\left\{\frac{1}{\sqrt{h}}\, S(k,h)\right\}_{k=-\infty}^{\infty}$$

is a complete orthonormal set in $B(h)$ *([17]).*

The class of functions $B(h)$ is quite restrictive in the sense that entirety cannot be expected of a function arising in applications. Hence, a larger class of functions is sought whereby the Whittaker cardinal expansion (2.2), while not exact, provides an accurate interpolatory series. A class of functions with this property is identified in [17].

Definition 2.7: *Let*

$$D_S = \{z \in \mathbb{C} : |\mathrm{Im}(z)| < d,\ 0 < d \leq \pi/2\}$$

and denote by $B^p(D_S)$ *the set of functions* f *such that* f *is analytic in* D_S, *and*

$$\int_{-d}^{d} |f(t+iy)|dy = O(|t|^{\alpha}) \ \text{ as } \ t \to \pm\infty$$

for some α, $0 \leq \alpha < 1$, *and for* $p = 1$ *or* 2

$$N_p(f, D_S) = \lim_{y \to d^-} \left[\left\{ \int_{-\infty}^{\infty} |f(t+iy)|^p dt \right\}^{1/p} + \left\{ \int_{-\infty}^{\infty} |f(t-iy)|^p dt \right\}^{1/p} \right] < \infty \quad .$$

For $p = 1$ *let* $N(f, D_S) = N_1(f, D_S)$ *and* $B(D_S) = B^1(D_S)$.

In this set of functions the error between f and $C(f,h)$ is given by the following theorem.

Theorem 2.8: *If for* $p = 1$ *or* 2, $f \in B^p(D_S)$ *and* $C(f,h)(x)$ *is convergent then*

$$\begin{aligned} E(f)(x) &:= f(x) - C(f,h)(x) \\ &= \frac{\sin(\pi x/h)}{2\pi i} \int_{-\infty}^{\infty} \left[\frac{f(t-id)}{(t-x-id)\sin[\pi(t-id)/h]} \right. \\ &\qquad \left. - \frac{f(t+id)}{(t-x+id)\sin[\pi(t+id)/h]} \right] dt \quad . \end{aligned} \tag{2.7}$$

Further, if $f \in B(D_S)$ *then*

$$\|E(f)\|_\infty \leq \frac{N_1(f, D_S)}{2\pi d \sinh(\pi d/h)} \tag{2.8}$$

and if $f \in B^2(D_S)$ *then*

$$\|E(f)\|_\infty \leq \frac{N_2(f, D_S)}{2\sqrt{\pi d}\, \sinh(\pi d/h)} \quad . \tag{2.9}$$

A proof of the above theorem can be found in Stenger [17]. A corollary to Theorem 2.8 is that if f is in $B^p(D_S)$ and g is a function analytic in D_S that is bounded there independent of h, then the product function fg is also in $B^p(D_S)$.

The error statement in Theorem 2.8 is valid only on the real line. While there are extensions to regions in $\mathbb{C}$, the goal of this work is to approximate on subsets of the real line and so the theorem is sufficiently general for the purposes of the thesis.

An order statement can be made using (2.8) and (2.9). As $h \to 0$, $\sinh(\pi d/h) \to \infty$. Hence $\|E(f)\|_{2,\infty} \to 0$ for either $f \in B(D_S)$ or $f \in B^2(D_S)$, and the rate of this convergence is dependent on $\sinh(\pi d/h)$. Note that $1/\sinh(\pi d/h)$ is $O\left(\exp(-\pi d/h)\right)$ as $h \to 0$.

For the purpose of practical approximation the infinite sum defining $C(f,h)$ in (2.2) is truncated to

$$C_{M,N}(f,h)(x) = \sum_{k=-M}^{N} f(kh)\mathrm{sinc}\left[\frac{x-kh}{h}\right] \quad .$$

It cannot, in general, be expected that the error introduced by truncation will maintain the exponential rate of convergence of the previous paragraph. However, in the case where f decreases sufficiently rapidly the exponential error rate (as a function of the number of retained interpolation points) is maintained.

Theorem 2.9: *If $f \in B^p(D_S)$ for $p = 1$ or 2, $d > 0$, and there exist positive constants α, β and L so that*

$$|f(x)| \leq L \begin{cases} \exp(\alpha x) & , \ x \in (-\infty, 0) \\ \exp(-\beta x) & , \ x \in [0, \infty) \end{cases} \tag{2.10}$$

then given a positive integer M, choose N and h by the relations

$$N = \left[\left| \frac{\alpha}{\beta} M + 1 \right|\right] \tag{2.11}$$

and

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

This leads to

$$\| f - C_{M,N}(f,h) \|_\infty \leq \kappa_1 \sqrt{M} \exp\left(-\sqrt{\pi d \alpha M} \right)$$

where κ_1 is a constant dependent only on f.

Proof: From Theorem 2.8 there exists a constant κ_2 so that

$$|f(x) - C(f,h)(x)| \leq \kappa_2 \exp(-\pi d/h)$$

for all $x \in \mathbb{R}$. By the triangle inequality and the geometric series it follows that

$$\begin{aligned} &|f(x) - C_{M,N}(f,h)(x)| \\ &\quad \leq \kappa_2 \exp(-\pi d/h) + \sum_{j=M+1}^{\infty} |f(-jh)| + \sum_{j=N+1}^{\infty} |f(jh)| \\ &\quad \leq \kappa_2 \exp(-\pi d/h) + L \left\{ \sum_{j=M+1}^{\infty} \exp(-\alpha j h) + \sum_{j=N+1}^{\infty} \exp(-\beta j h) \right\} \\ &\quad \leq \kappa_2 \exp(-\pi d/h) + L \left\{ \frac{\exp(-\alpha M h)}{\alpha h} + \frac{\exp(-\beta N h)}{\beta h} \right\} \\ &\quad \leq \left\{ \kappa_2 + L \left[\frac{1}{\alpha} + \frac{1}{\beta} \right] \frac{\sqrt{\alpha \pi d M}}{\pi d} \right\} \exp\left(-\sqrt{\alpha \pi d M} \right) \\ &\quad \leq \left\{ \kappa_2 + L \left(\frac{\alpha + \beta}{\alpha \beta} \right) \frac{\sqrt{\alpha \pi d}}{\pi d} \right\} \sqrt{M} \exp\left(-\sqrt{\alpha \pi d M} \right) . \end{aligned}$$

The proof is completed by defining the constant κ_1

$$\kappa_1 = \left\{ \kappa_2 + L \frac{\alpha+\beta}{\alpha\beta} \frac{\sqrt{\alpha\pi d}}{\pi d} \right\} \quad .$$

The interpolation results above show that the use of sinc functions as an interpolatory basis results in very accurate approximations of a function f on the real line if f is in $B^p(D_S)$ and satisfies the decay condition (2.10).

The primary use of Theorems 2.8 and 2.9 in this thesis is to develop the sinc quadrature rule. A few preliminary results are needed to develop this integration rule.

Substituting $x = 0$ in (2.5) shows that

$$\frac{1}{h} \int_{-\infty}^{\infty} S(k,h)(x)dx = 1$$

for all integers k. Hence if the Whittaker cardinal expansion of f in (2.2) is convergent for all x then

$$\int_{-\infty}^{\infty} C(f,h)(x)dx = h \sum_{k=-\infty}^{\infty} f(kh) \quad .$$

Combining this identity with the result of Theorem 2.8 leads to the quadrature rule

$$\int_{-\infty}^{\infty} f(x)dx = h \sum_{k=-\infty}^{\infty} f(kh) + \Theta(f) \quad , \tag{2.12}$$

where

$$\Theta(f) = \int_{-\infty}^{\infty} E(f)(x)dx \quad .$$

For real positive α, the Laplace integral

$$\int_{-\infty}^{\infty} \frac{\exp(i\alpha x)}{x-z}\, dx = \begin{cases} 2\pi i \exp(i\alpha z) & , \ \mathrm{Im}\{z\} > 0 \\ 0 & , \ \mathrm{Im}\{z\} < 0 \end{cases}$$

may be used in $\Theta(f)$ along with the definition of $N(f, D_S)$ to show that

$$|\Theta(f)| \leq \frac{\exp(-\pi d/h) N(f, D_S)}{2 \sinh(\pi d/h)} \quad .$$

This result shows that the sinc quadrature on $\mathbb{R}$ is identical to the standard trapezoidal rule, but that the restriction of f to $B(D_S)$ gives an error of the order of $O\left(\exp(-2\pi d/h)\right)$ as opposed to the trapezoidal error of order $O(h^2)$ when f'' is bounded.

A restriction on the growth rate of f must be imposed to guarantee an exponential rate of convergence for (2.12) when the sum in (2.12) is truncated. This is similar to the restrictions outlined in Theorem 2.9 for interpolation.

Theorem 2.10: *Assume $f \in B(D_S)$, $d > 0$, and for positive constants α, β and L, f satisfies (2.10). Given an integer M, choose N by (2.11) and h by*

$$h = \left(\frac{2\pi d}{\alpha M}\right)^{1/2} \quad . \tag{2.13}$$

Then

$$\left| \int_{-\infty}^{\infty} f(x)dx - h \sum_{k=-M}^{N} f(kh) \right| \leq \kappa_3 \exp\left(-\sqrt{2\alpha\pi d M}\right) \quad , \tag{2.14}$$

where κ_3 depends only on f.

The domains of the present thesis are different than the entire real line so that the development of an accurate quadrature rule based on (2.14) requires the following definitions.

Definition 2.11: *Let D be a simply connected domain and D_S as in Definition 2.7. Given distinct a, b on the boundary of D, let χ be a conformal map from D onto D_S satisfying*

$$\Gamma = \chi^{-1}(\mathbb{R})$$

and

$$\lim_{\substack{z \to a \\ z \in \Gamma}} \chi(z) = -\infty \quad , \quad \lim_{\substack{z \to b \\ z \in \Gamma}} \chi(z) = +\infty \quad .$$

Let $B(D)$ denote the family of functions analytic in D that satisfy

$$\int_{\chi^{-1}(x+L)} |F(z)dz| \to 0 \text{ as } x \to \pm\infty$$

$$L = \{iy : |y| < d\}$$

and

$$N(F, D) = \lim_{\substack{C \subset D \\ C \to \partial D}} \int_C |F(z)dz| < \infty$$

where C is a simple closed curve in D.

If $f \in B(D)$ then g defined by

$$g(x) = f\left(\chi^{-1}(x)\right)\left(\chi^{-1}\right)'(x)$$

is in $B(D_S)$ of Definition 2.7. The growth restriction in (2.10) for an arbitrary simply connected domain D is characterized by the following definition.

Definition 2.12: *A function $f \in B(D)$ is said to decay exponentially with respect to the conformal mapping χ of D onto D_S if there exist positive constants L, α and β so that*

$$\left|\frac{f(x)}{\chi'(x)}\right| \leq L \begin{cases} \exp\left(-\alpha|\chi(x)|\right) & , \ x \in \Gamma_L \\ \exp\left(-\beta|\chi(x)|\right) & , \ x \in \Gamma_R \end{cases} \tag{2.15}$$

where

$$\Gamma_L = \{z : \chi(z) \in (-\infty, 0)\} \quad , \quad \Gamma_R = \{z : \chi(z) \in [0, \infty)\} \quad . \tag{2.16}$$

With this definition, the following theorems are the direct analogues of Theorems 2.9 and 2.10 for an arc $\Gamma \subset D$.

Theorem 2.13: *Assume $f \in B(D)$ and f decays exponentially with respect to the conformal map $\chi : D \to D_S$. Given a positive integer M select*

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

and

$$h = \sqrt{\pi d/(\alpha M)} \quad .$$

Then upon setting

$$z_k = \chi^{-1}(kh)$$

it follows that for all $z \in \Gamma$

$$\begin{aligned} &\left|\frac{f(z)}{\chi'(z)} - C_{M,N}\left[\frac{f}{\chi'}\,,\, h\right](\chi(z))\right| \\ &= \left|\frac{f(z)}{\chi'(z)} - \sum_{k=-M}^{N} \frac{f(z_k)}{\chi'(z_k)}\, S(k,h)\,(\chi(z))\right| \\ &\leq \kappa_4 \exp\left(-\sqrt{\alpha\pi d M}\right) \quad , \end{aligned}$$

where the constant κ_4 depends only on f.

Theorem 2.14: *Assume $f \in B(D)$ and f decays exponentially with respect to the conformal map $\chi : D \to D_S$. Given a positive integer M define*

(2.17)
$$N = \left[\left|\frac{\alpha}{\beta}\, M + 1\right|\right]$$

and

(2.18)
$$h = \sqrt{2\pi d/(\alpha M)} \quad .$$

Upon recalling $\Gamma = \chi^{-1}(\mathbb{R})$ *and setting*

$$z_k = \chi^{-1}(kh)$$

it follows that

$$\left| \int_\Gamma f(z)dz - h \sum_{k=-M}^{N} \frac{f(z_k)}{\chi'(z_k)} \right| \leq \kappa_5 \exp\left(-\sqrt{2\pi d\alpha M}\right) \tag{2.19}$$

where κ_5 *depends only on* f.

There is an important special case of the previous result that plays a fundamental role in the construction and evaluation of the inner products defined in the next chapter. This special case uses the interpolatory identity

$$S(j,h)\left(\chi(z_k)\right) = \delta_{jk}$$

and takes the form for $f \in B(D)$

$$\int_\Gamma f(z)S(k,h)(\chi(z))dz = h\,\frac{f(z_k)}{\chi'(z_k)} + O\left(\exp(-\pi d/h)\right) \quad . \tag{2.20}$$

That is, the reproduction property of the sinc kernel in Theorem 2.3 remains "approximately" true for $f \in B(D)$.

CHAPTER 3

INNER PRODUCT APPROXIMATIONS

The quadrature rule (2.19) is readily adaptable to the numerical approximation of the inner products arising in the Sinc-Galerkin discretization of (1.1). In this chapter, the focus is on the discretization of

$$Pu(x,t) \equiv u_t(x,t) - u_{xx}(x,t) = f(x,t,u) \tag{3.1}$$

where

$$(x,t) \in (a,b) \times (0,\infty) \quad .$$

The emphasis will be on the discretization of Pu but also included is the discretization of $f(x,t,u)$ where f takes any one of the forms

$$f(x,t,u) = f(x,t) \quad , \tag{3.2}$$

$$f(x,t,u) = u(x,t) \quad , \tag{3.3}$$

$$f(x,t,u) = u_x(x,t) \tag{3.4}$$

and

$$f(x,t,u) = u(x,t)u_x(x,t) = \frac{1}{2}\,\frac{\partial}{\partial x}\left[(u(x,t))^2\right] \quad . \tag{3.5}$$

The following definition, which specializes Definition 2.11, establishes the notation that will be adhered to throughout the remainder of this thesis.

Definition 3.1: *Let ϕ be an invertible conformal map in the simply connected domain D with the property that $\phi : (a,b) \to (-\infty,\infty)$. Let $\{x_k\}$ be implicitly defined by the relationship $kh = \phi(x_k)$, for all $k \in \mathbb{Z}$. Let $\hat{\phi}$ be an invertible conformal map in the simply connected domain D_t with the property that $\hat{\phi}$: $(0,\infty) \to (-\infty,\infty)$. Let $\{t_j\}$ be implicitly defined by the relationship $j\hat{h} = \hat{\phi}(t_j)$, for all $j \in \mathbb{Z}$.*

For the following, assume that the parameters h, $\hat{h}$, M_t, N_t, M_x and N_x have been assigned. In the ensuing work, error terms are developed as functions of these assignments. These error terms are used to determine optimal choices for the parameters, and this is all sorted out in detail in Chapter 6. Towards simplifying notation and the convenient delineation of spatial versus temporal domain variables the following definition is useful.

Definition 3.2: *Let $w(x)$ and $\hat{w}(t)$ represent positive weight functions on (a,b) and $(0,\infty)$, respectively. The Galerkin inner product of f is defined by*

$$
\begin{aligned}
< f >_{kj} &\equiv \frac{1}{h\hat{h}} \int_a^b \int_0^\infty \left\{ f(x,t) S(k,h) \circ \phi(x) S(j,\hat{h}) \circ \hat{\phi}(t) w(x) w(t) \right\} dt dx \\
&\equiv \frac{1}{h\hat{h}} \int_a^b \int_0^\infty \left\{ f(x,t) S_k(x) \hat{S}_j(t) w(x) w(t) \right\} dx dt \quad .
\end{aligned} \tag{3.6}
$$

In (3.6) the sinc function (2.1) when composed with $\phi(x)$ will be denoted by $S_k(x)$ (similarly for $\hat{S}_j(t) \equiv S(j,\hat{h}) \circ \hat{\phi}(t)$). The convention of a hat denoting a function in the temporal domain will be adhered to throughout the remainder of the thesis. In like manner this applies to parameters where appropriate, e.g. $\hat{h}$ is the "mesh size" in the temporal domain. As a final convention, it is typographically convenient to set

$$
\frac{d}{d\phi} S_k(x) \equiv S_k'(x) \quad , \quad \frac{d^2}{d\phi^2} S_k(x) = S_k''(x) \tag{3.7}
$$

and

$$\frac{d}{d\hat{\phi}}\,\hat{S}_j(t) \equiv \hat{S}_j'(t) \quad . \tag{3.8}$$

This replaces the somewhat unwieldly notation $S_{k\phi}(x)$, $S_{k\phi\phi}(x)$ *or* $S_{j\hat{\phi}}(t)$.

For $u_t(x,t)$, the Galerkin inner product with sinc basis elements in both the space and time domains is given by

$$< u_t >_{kj} \equiv \frac{1}{h\hat{h}} \int_a^b \int_0^\infty \left[u_t(x,t) S_k(x) w(x) \hat{S}_j(t) \hat{w}(t)\right] dt dx \quad .$$

This expression contains the t derivative of u, but the desired result is interpolation of the variable u with no derivatives. Theorem 2.13 guarantees an accurate interpolation of $u(x,t)$ but does not guarantee an accurate interpolation of $u_t(x,t)$. Integrating by parts to remove the t derivative from the dependent variable u leads to the equality

$$\begin{aligned} < u_t >_{kj} \; &= \frac{1}{h\hat{h}} \int_a^b \left[u(x,t) S_k(x) w(x) \hat{S}_j(t) \hat{w}(t)\right] \Big|_{t=0}^{t=\infty} dx \\ &\quad - \frac{1}{h\hat{h}} \int_a^b \int_0^\infty \left[u(x,t) S_k(x) w(x) \left(\hat{S}_j(t) \hat{w}(t)\right)_t\right] dt dx \quad . \end{aligned} \tag{3.9}$$

While there are several possible methods for dealing with the first integral in (3.9) (the boundary integral), for the present it is convenient to assume that for each fixed x,

$$\lim_{t \to 0^+} u(x,t) \hat{w}(t) / \hat{\phi}(t) = 0 \tag{3.10}$$

and

$$\lim_{t \to \infty} u(x,t) \hat{w}(t) / \hat{\phi}(t) = 0 \quad . \tag{3.11}$$

The immediate consequence of the assumptions in (3.10) and (3.11), coupled with the fact that

$$\hat{S}_j(t) = \frac{\sin\left[\frac{\pi}{\hat{h}}\left(\hat{\phi}(t) - j\hat{h}\right)\right]}{\frac{\pi}{\hat{h}}\left(\hat{\phi}(t) - j\hat{h}\right)} \quad ,$$

is that the boundary integral in (3.9) vanishes. The rationale of these assumptions is related to the boundary behaviors of $u(*,t)$ and $\hat{w}(t)$, and partially motivates the selection of $\hat{w}(t)$ in the present development. Under the assumption that (3.10) and (3.11) are satisfied, then (3.9) may be written as

$$\begin{aligned} < u_t >_{kj} &= -\frac{1}{h\hat{h}} \int_a^b \int_0^\infty \left[u(x,t) S_k(x) w(x) \left(\hat{S}_j(t)\hat{w}(t) \right)_t \right] dxdt \\ &= -\frac{1}{h\hat{h}} \int_a^b \int_0^\infty u(x,t) S_k(x) w(x) \left(\hat{S}_j'(t)\hat{\phi}'(t)\hat{w}(t) + \hat{S}_j(t)\hat{w}'(t) \right) dtdx \quad , \end{aligned} \tag{3.12}$$

where the primes denote differentiation with respect to t and the convention in (3.8) is used.

In order to apply the quadrature rule (2.19) to the right-hand side of (3.12) assume that for each fixed x

$$u(x,t)\hat{w}'(t) \in B(D_t)$$

and

$$u(x,t)\hat{w}(t)\hat{\phi}'(t) \in B(D_t) \quad ,$$

where $B(D_t)$ is defined in Definition 2.11 with $D \equiv D_t$. If each of the two above functions decay exponentially with respect to $\hat{\phi}$ (Definition 2.12), then the application of the sinc quadrature rule (2.19) in the variable t gives

$$\begin{aligned} < u_t >_{kj} = -\frac{1}{h} \int_a^b S_k(x) w(x) \Bigg\{ & u(x,t_j)\hat{w}'(t_j)/\hat{\phi}'(t_j) \\ & + \sum_{i=-M_t}^{N_t} u(x,t_i)\hat{w}(t_i)\hat{S}_j'(t_i) \Bigg\} dx \quad . \end{aligned} \tag{3.13}$$

The error in this approximation is of the order $O\left(\exp\left(-\gamma_t\sqrt{M_t}\right)\right)$ where γ_t depends on the exponential decay of u and the region D_t. If in addition it is assumed that for each fixed t

$$u(x,t)w(x) \in B(D)$$

then the special case of the sinc quadrature (2.20) in the x-variable can be applied to (3.13) resulting in the expression

$$\begin{aligned} < u_t >_{kj} = &- \sum_{i=-M_t}^{N_t} \left[u(x_k,t_i)\hat{w}(t_i)\hat{S}_j'(t_i)w(x_k)/\phi'(x_k)\right] \\ &- u(x_k,t_j)w(x_k)\hat{w}'(t_j)/\left(\hat{\phi}'(t_j)\phi'(x_k)\right) \end{aligned} \tag{3.14}$$

with the addition of an error term of the order $O\left(\exp(-\pi d/h)\right)$.

The Galerkin inner product for $u_{xx}(x,t)$ may be handled in a similar manner upon integrating by parts twice. This gives the computation

$$\begin{aligned} < u_{xx} >_{kj} &= \frac{1}{h\hat{h}}\int_a^b\int_0^\infty u_{xx}(x,t)w(x)S_k(x)\hat{w}(t)\hat{S}_j(t)dtdx \\ &= \frac{1}{h\hat{h}}\int_0^\infty \left[u_x(x,t)w(x)S_k(x)\right]\Big|_{x=a}^{x=b}\hat{w}(t)\hat{S}_j(t)dt \\ &\quad - \frac{1}{h\hat{h}}\int_0^\infty \left[u(x,t)\{w(x)S_k(x)\}'\right]\Big|_{x=a}^{x=b}\hat{w}(t)\hat{S}_j(t)dt \\ &\quad + \frac{1}{h\hat{h}}\int_a^b\int_0^\infty u(x,t)\{w(x)S_k(x)\}''\hat{w}(t)\hat{S}_j(t)dtdx \ . \end{aligned} \tag{3.15}$$

In the above calculation the primes denote differentiation with respect to x. As before, the boundary integrals in (3.15) will be assumed to be zero; that is, for each fixed t, assume

$$\lim_{x\to b^-} u_x(x,t)w(x)/\phi(x) = 0$$

and

$$\lim_{x\to a^+} u_x(x,t)w(x)/\phi(x) = 0$$

as well as

$$\lim_{x \to b^-} u(x,t)\{\phi'(x)w(x) + w'(x)\}/\phi(x) = 0$$

and

$$\lim_{x \to a^+} u(x,t)\{\phi'(x)w(x) + w'(x)\}/\phi(x) = 0 \quad .$$

Continuing with the only remaining integral in (3.15) and expanding the derivative results in

$$\begin{aligned} < u_{xx} >_{kj} &= \frac{1}{h\hat{h}} \int_a^b \int_0^\infty u(x,t)\{w(x)S_k(x)\}''\hat{w}(t)\hat{S}_j(t)dtdx \\ &= \frac{1}{h\hat{h}} \int_a^b \int_0^\infty u(x,t)\hat{w}(t)\hat{S}_j(t)S_k(x)w''(x)dtdx \\ &+ \frac{1}{h\hat{h}} \int_a^b \int_0^\infty u(x,t)\hat{w}(t)\hat{S}_j(t)S_k'(x)\{\phi''(x)w(x) + 2\phi'(x)w'(x)\}dtdx \\ &+ \frac{1}{h\hat{h}} \int_a^b \int_0^\infty u(x,t)\hat{w}(t)\hat{S}_j(t)S_k''(x)[\phi'(x)]^2 w(x)dtdx \quad . \end{aligned}$$

Under the assumptions that for each fixed t

$$u(x,t)w''(x) \quad ,$$

$$u(x,t)\{\phi''(x)w(x) + 2w'(x)\phi'(x)\}$$

and

$$u(x,t)w(x)[\phi'(x)]^2$$

are each in $B(D)$ and each decays exponentially with respect to ϕ, then the sinc quadrature rule (2.19) in the variable x applied to the above integral yields the

approximation

$$< u_{xx} >_{kj} = \frac{1}{\hat{h}} \int_0^\infty u(x_k,t)\hat{w}(t)\hat{S}_j(t)w''(x_k)/\phi'(x_k)dt$$

$$+ \frac{1}{\hat{h}} \int_0^\infty \sum_{i=-M_x}^{N_x} u(x_i,t)\hat{w}(t)\hat{S}_j(t)S_k'(x_i)\{\phi''(x_i)w(x_i)/\phi'(x_i)$$

$$+ 2w'(x_i)\}\, dt$$

$$+ \frac{1}{\hat{h}} \int_0^\infty \sum_{i=-M_x}^{N_x} u(x_i,t)\hat{w}(t)\hat{S}_j(t)S_k''(x_i)\phi'(x_i)w(x_i)dt$$

If in addition, for each fixed x, the function

$$u(x,t)\hat{w}(t) \in B(D_t)$$

then the sinc quadrature rule (2.20) in the variable t results in the approximation

$$< u_{xx} >_{kj} = \frac{\hat{w}(t_j)u(x_k,t_j)w''(x_k)}{\hat{\phi}'(t_j)\phi'(x_k)}$$

$$+ \hat{w}(t_j) \sum_{i=-M_x}^{N_x} u(x_i,t_j)S_k'(x_i)\{\phi''(x_i)w(x_i)/\phi'(x_i) + 2w'(x_i)\}/\hat{\phi}'(t_j) \tag{3.16}$$

$$+ \hat{w}(t_j) \sum_{i=-M_x}^{N_x} u(x_i,t_j)S_k''(x_i)\phi'(x_i)w(x_i)/\hat{\phi}'(t_j) \quad .$$

As in the development from (3.13) to (3.14) the error incurred in the approximation in (3.16) is of the order $O\left(\exp\left(-\gamma_x\sqrt{M_x}\right) + \exp(-\pi\hat{d}/\hat{h})\right)$. Before specifying the weight functions as well as the conformal maps, it is convenient to collect the above development into a theorem.

Theorem 3.3: *Let $u(x,t)$ be a function of two variables, and let*

$$D_E = \{z : -d < \arg[(z-a)/(b-z)] < d\} \quad , \tag{3.17}$$

$$D_W = \{z : -d < \arg(z) < d\} \tag{3.18}$$

and

$$D_B = \{z : -d < \arg(\sinh(z)) < d\} \quad . \tag{3.19}$$

Given w, $\hat{w}$, ϕ, and $\hat{\phi}$, and for $q = W$ or B, assume that

$$u\hat{w}\,,\ u\hat{w}\hat{\phi}'\,,\ u\hat{w}' \in B(D_q) \tag{3.20}$$

and each decays exponentially with respect to $\hat{\phi}$ (Definition 2.12) and that

$$uw(\phi')^2\,,\ uw\,,\ uw''\,,\ u(\phi''w + 2w'\phi') \in B(D_E) \tag{3.21}$$

and each decays exponentially with respect to ϕ. Assume further that

$$\lim_{\substack{x \to b^- \\ x \to a^+}} u_x(x,*)w(x)/\phi(x) = 0 \quad , \tag{3.22}$$

$$\lim_{\substack{x \to b^- \\ x \to a^+}} u(x,*)\{\phi'(x)w(x) + w'(x)\}/\phi(x) = 0 \tag{3.23}$$

and

$$\lim_{\substack{t \to 0^+ \\ t \to \infty}} u(*,t)\hat{w}(t)/\hat{\phi}(t) = 0 \quad . \tag{3.24}$$

Then the approximation

$$\begin{aligned} < u_t >_{kj} = -\sum_{i=-M_t}^{N_t} & \left[u(x_k,t_i)\hat{w}(t_i)\hat{S}'_j(t_i)w(x_k)/\phi'(x_k)\right] \\ & - u(x_k,t_j)w(x_k)\hat{w}'(t_j)/\left(\hat{\phi}'(t_j)\phi'(x_k)\right) \end{aligned} \tag{3.25}$$

is of order $O\left(\exp\left(-\gamma_t\sqrt{M_t}\right) + \exp(-\pi d/h)\right)$ and the approximation

$$< u_{xx} >_{kj} = \frac{\hat{w}(t_j)u(x_k,t_j)w''(x_k)}{\hat{\phi}'(t_j)\phi'(x_k)}$$

$$+ \hat{w}(t_j) \sum_{i=-M_x}^{N_x} u(x_i,t_j)S_k'(x_i)\{\phi''(x_i)w(x_i)/\phi'(x_i) + 2w'(x_i)\}/\hat{\phi}'(t_j) \tag{3.26}$$

$$+ \hat{w}(t_j) \sum_{i=-M_x}^{N_x} u(x_i,t_j)S_k''(x_i)\phi'(x_i)w(x_i)/\hat{\phi}'(t_j)$$

is of the order $O\left(\exp\left(-\gamma_x\sqrt{M_x}\right) + \exp(-\pi\hat{d}/\hat{h})\right)$.

In a similar fashion, the Galerkin inner products for the function $f(x,t,u)$ on the right-hand side of (3.1) may be derived. The inner products of each of these will be dealt with separately, but in little detail as the development is for the most part a reproduction of the above arguments. For f in (3.2), the inner product is, from (3.6) with (2.20)

$$\begin{aligned} < f >_{kj} &= \frac{1}{h\hat{h}} \int_a^b \int_0^\infty \left\{ f(x,t)w(x)S_k(x)\hat{w}(t)\hat{S}_j(t) \right\} dtdx \\ &= \frac{\hat{w}(t_j)f(x_k,t_j)w(x_k)}{\hat{\phi}'(t_j)\phi'(x_k)} \end{aligned} \tag{3.27}$$

with error order $O\left(\exp(-\pi d/h) + \exp(-\pi\hat{d}/\hat{h})\right)$ under the assumption that

$$f(x,*)w(x)$$

is in $B(D_E)$ and

$$f(*,t)\hat{w}(t)$$

is in $B(D_q)$, $q = W$ or B. In a similar fashion the inner product for (3.3) is

$$\begin{aligned} < u >_{kj} &= \frac{1}{h\hat{h}} \int_a^b \int_0^\infty \left\{ u(x,t)w(x)S_k(x)\hat{w}(t)\hat{S}_j(t) \right\} dtdx \\ &= \frac{\hat{w}(t_j)f(x_k,t_j)w(x_k)}{\hat{\phi}'(t_j)\phi'(x_k)} \end{aligned} \tag{3.28}$$

with error order the same as for f, under the assumption that

$$u(x,*)w(x)$$

is in $B(D_E)$, and

$$u(*,t)\hat{w}(t)$$

is in $B(D_q)$, $q = W$ or B. Notice that these two conditions are already included in Theorem 3.3.

The u_x development reproduces the u_t development in (3.14), and so, without proof,

$$\begin{aligned} < u_x >_{kj} = & - \sum_{i=-M_x}^{N_x} \left[u(x_i,t_j)w(x_i)S_k'(x_i)\hat{w}(t_j)/\hat{\phi}'(t_j) \right] \\ & - u(x_k,t_j)w'(x_k)\hat{w}(t_j)/\left(\hat{\phi}'(t_j)\phi'(x_k)\right) \end{aligned} \tag{3.29}$$

under the assumptions that the functions

$$u(x,*)w'(x) \quad , \quad u(x,*)w(x)\phi'(x) \tag{3.30}$$

are in $B(D_E)$ and decay exponentially with respect to ϕ while

$$u(*,t)\hat{w}(t)$$

is in $B(D_q)$, $q = W$ or B. The last condition is part of the hypotheses of Theorem 3.3, while (3.30) is new.

In the case of (3.5), the identity $uu_x = \frac{1}{2}\frac{\partial}{\partial x}(u^2)$, shows that the development in (3.29) is applicable. That is, at every step of (3.29) u is replaced by u^2. Hence, for Burgers' equation the approximation reads

$$\begin{aligned} 2 < uu_x >_{kj} &= < \frac{\partial}{\partial x}(u^2) >_{kj} \\ &= - \sum_{i=-M_x}^{N_x} \left[u^2(x_i,t_j)w(x_i)S_k'(x_i)\hat{w}(t_j)/\hat{\phi}'(t_j) \right] \\ &\quad - u^2(x_k,t_j)w'(x_k)\hat{w}(t_j)/\left(\hat{\phi}'(t_j)\phi'(x_k)\right) \end{aligned} \tag{3.31}$$

under the assumptions that

$$u^2(x,*)w'(x) \quad , \quad u^2(x,*)w(x)\phi'(x)$$

are in $B(D_E)$ and decay exponentially with respect to ϕ while

$$u^2(*,t)\hat{w}(t)$$

is in $B(D_q)$, $q = W$ or B. Also the boundary conditions

$$\lim_{\substack{x \to a^+ \\ x \to b^-}} u^2(x,*)w(x)/\phi(x) = 0$$

must be met.

A choice of weight and mapping functions will be made for the present discussion, although there are many possible choices. Let (a,b) be $(0,1)$, and define

$$\phi(x) = \ell n[x/(1-x)]$$

and

$$w(x) = 1/\sqrt{\phi'(x)} \quad .$$

As for the assumptions in (3.21)–(3.23), the boundary condition (3.23) is trivially satisfied. The assumptions in (3.21) take the following form. Assume the functions

$$(3.32) \qquad \frac{u}{(x(1-x))^{3/2}} \quad , \quad u\sqrt{x(1-x)} \quad , \quad \frac{-u}{4(x(1-x))^{3/2}} \quad , \quad 0$$

are in $B(D_E)$ and each decays exponentially with respect to ϕ (Definition 2.12). These assumptions on

$$(3.33) \qquad u(x,*)/(x(1-x))^{3/2}$$

imply the conditions in (3.32) since $x(1-x)$ is bounded on $[0,1]$. The function $u/(x(1-x))^{3/2}$ decays exponentially with respect to ϕ. The limiting value of

$\sqrt{x(1-x)}/\ell n[x/(1-x)]$ is 0 as $x \to 0^+$ or as $x \to 1^-$. These two conditions combined show that condition (3.22) is satisfied as well. This condition also implies condition (3.30). Thus (3.33) is a sufficient condition to meet all the required conditions of Theorem 3.3 in the variable x.

The mapping function used in [9] is defined by

$$\hat{\phi}(t) = \ell n(t)$$

with the weight function

$$\hat{w}(t) = \sqrt{\hat{\phi}'(t)} \quad .$$

A short calculation gives

$$\hat{w}(t) = 1/\sqrt{t}$$

$$\hat{w}'(t) = -1/(2t^{3/2})$$

and

$$\hat{w}(t)\hat{\phi}'(t) = 1/t^{3/2} \quad .$$

The assumptions in (3.20) reduce to one of (the latter two are the same)

(3.34) $$u(*,t)/t^{3/2}$$

or

(3.35) $$u(*,t)/\sqrt{t}$$

are in $B(D_W)$ and decay exponentially with respect to $\hat{\phi}$. Since $1/\hat{\phi}(t) = 1/\ell n(t) \to 0$ as $t \to 0^+$ or as $t \to \infty$ this verifies that condition (3.24) is satisfied. So the two conditions (3.34) and (3.35) are sufficient to meet all the required conditions of Theorem 3.3 in the variable t.

An alternative choice for the t domain mapping is defined by

$$\hat{\phi}(t) = \ell n(\sinh(t)) \quad .$$

Here the weight function selected is

$$\hat{w}(t) = \sqrt{\hat{\phi}'(t)}$$

as before. Then

$$\hat{w}(t) = \sqrt{\coth(t)} \quad ,$$

$$\hat{w}'(t) = -(\coth(t))^{3/2} / \left(2\cosh^2(t)\right)$$

and

$$\hat{w}(t)\hat{\phi}'(t) = (\coth(t))^{3/2} \quad ,$$

and condition (3.20) reduces to

(3.36) $$u(*,t)(\coth(t))^{3/2}\mathrm{sech}^2(t) \quad ,$$

(3.37) $$u(*,t)(\coth(t))^{3/2}$$

and

(3.38) $$u(*,t)\sqrt{\coth(t)}$$

are in $B(D_B)$ and decay exponentially with respect to $\hat{\phi}$. Now (3.37) implies (3.36) and (3.38) as $\mathrm{sech}(t)$ and $\tanh(t)$ are bounded on $(0,\infty)$. Also the exponential decay with respect to $\hat{\phi}$ implies that $u(x,t)\sqrt{\coth(t)}$ is bounded on $(0,\infty)$, and so as before condition (3.24) is necessarily satisfied. So the condition (3.37) is sufficient in the variable t for this choice of mapping function. The conditions of Theorem 3.3 for the two choices of mapping function are collected succinctly in Table 1. The constraints on the functions f in (3.4)–(3.5) are concisely tabulated in Table 2.

The particular choices of the parameters M_x, M_t, N_x, N_t, h and $\hat{h}$ which give rise to the exponential convergence rates of the above approximations will be spelled out in the examples of Chapter 6. The basic philosophy is to balance (with respect to order) the various error contributions arising from the different inner product approximations. The domains for the different variables and mapping functions are shown explicitly in Figures 1 and 2.

Variable and Subdomain	Growth Constraints
Mapping Function	Mesh points in the Domain
$x \in (0, .5)$	$\lvert u(x,*)\rvert < K\, x^{\alpha+1/2}$
$x \in [.5, 1)$	$\lvert u(x,*)\rvert < K(1-x)^{\beta+1/2}$
$\phi(x) = \ell n\left(\frac{x}{1-x}\right)$	$x_k = \frac{\exp(kh)}{1+\exp(kh)}$
$t \in (0,1)$	$\lvert u(*,t)\rvert < C\, t^{\delta+1/2}$
$t \in [1, \infty)$	$\lvert u(*,t)\rvert < C\, t^{-(\mu+1/2)}$
$\hat{\phi}(t) = \ell n(t)$	$t_j = \exp(j\hat{h})$
$t \in \left(0, \ell n(1+\sqrt{2})\right)$	$\lvert u(*,t)\rvert < C\, t^{\delta+1/2}$
$t \in \left[\ell n(1+\sqrt{2}), \infty\right)$	$\lvert u(*,t)\rvert < C\exp(-\mu t)$
$\hat{\phi}(t) = \ell n(\sinh(t))$	$t_j = j\hat{h} + \ell n\left[1+\sqrt{1+\exp(-2j\hat{h})}\right]$

Table 1. Formulation of Definition 2.12 for the conformal maps of Chapter 3.

Conditions on the Function f in (3.4)–(3.5)	
$f(x,t,u)$ and Subdomain	Growth Constraints
$u_x(x,t)$	
$x \in (0,.5)$	$\lvert u(x,*)\rvert < K\, x^{\alpha+1/2}$
$x \in [.5,1)$	$\lvert u(x,*)\rvert < K(1-x)^{\beta+1/2}$
$\frac{1}{2}\left(u^2(x,t)\right)_x$	
$x \in (0,.5)$	$\lvert u^2(x,*)\rvert < K\, x^{\alpha+1/2}$
$x \in [.5,1)$	$\lvert u^2(x,*)\rvert < K(1-x)^{\beta+1/2}$

Table 2. Formulation of Definition 2.12 for the non-homogeneous terms in (3.1).

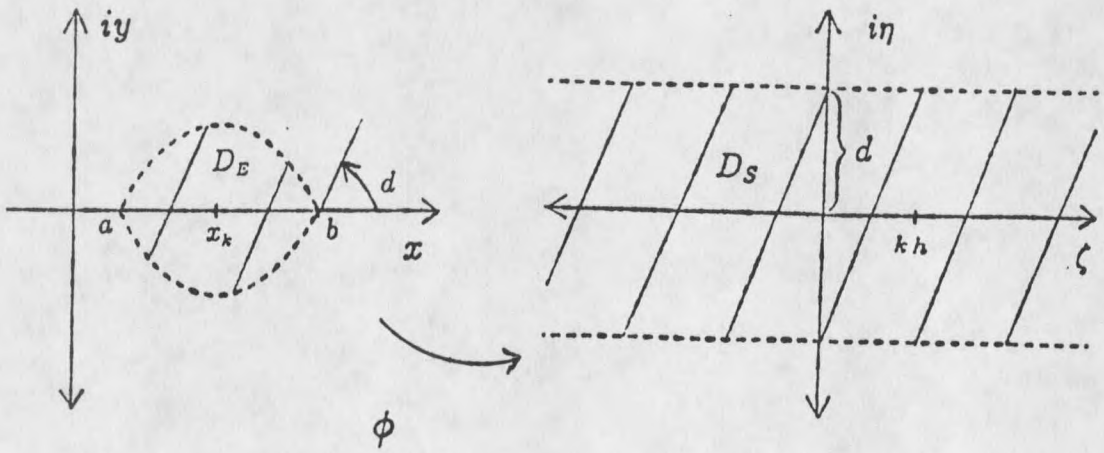


Figure 1. The spatial map ϕ.

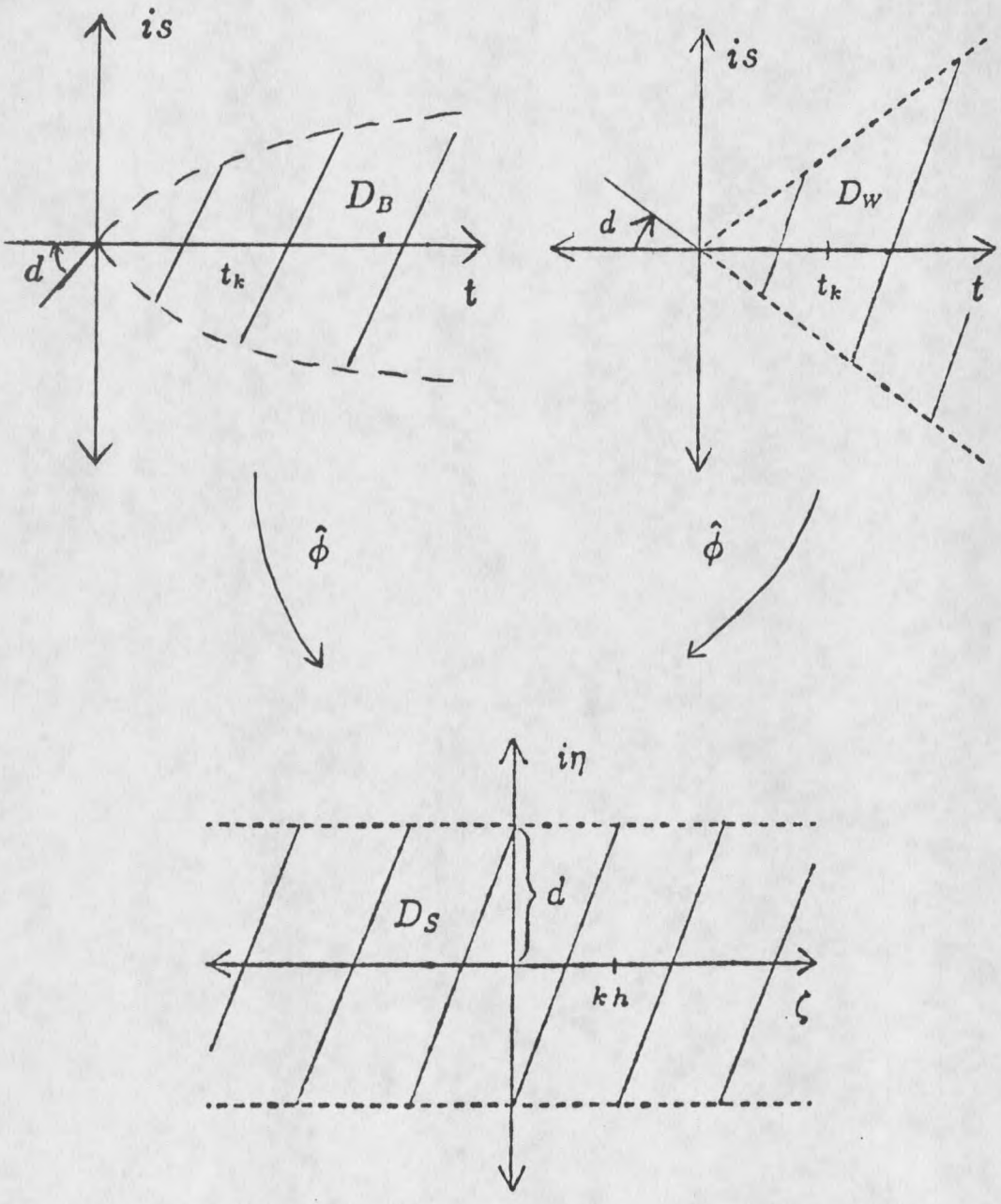


Figure 2. The temporal map $\hat{\phi}$.

CHAPTER 4

THE SINC-GALERKIN MATRIX SYSTEM

The matrix representation of the discrete system for the problem

$$Pu(x,t) \equiv u_t(x,t) - u_{xx}(x,t) = f(x,t,u) \tag{4.1}$$

is obtained via the orthogonalization of the residual with respect to the basis elements $\{S_k \hat{S}_j\}$, that is

$$< Pu_A - f >_{k,j} = 0 \quad , \tag{4.2}$$

where the inner product is given in Definition 3.2. The assumed approximate solution u_A to the true solution u of (4.1) takes the form

$$u_A(x,t) \equiv \sum_{\ell=-M_x}^{N_x} \sum_{p=-M_t}^{N_t} u_{\ell p} S_\ell(x) \hat{S}_p(t) \tag{4.3}$$

where the conventions laid down in Definition 3.2 of setting $S_p(x) \equiv S(p,h) \circ \phi(x)$ and hats for functions (or parameters) in the time domain remain in force. Also, throughout this chapter the integers $m_q = M_q + N_q + 1$, $q = x$ or t. That is, there are $m_x \cdot m_t$ unknown coefficients to determine in (4.3).

The basic methodology of the chapter is to use the inner product approximations of Chapter 3 to replace (4.2) by the Sinc-Galerkin system. In the course of this development it is notationally simpler to record the system as a tensor or Kronecker product. Besides a simplification, the tensor form has a twofold benefit. The first is in a convenient method to write down the discretization of (4.1)

based on the discretization of related one-dimensional problems. This in turn yields an easy method to write down the discretization of (4.1) in the case that the independent variable x is not scalar; i.e. for multi-dimensional problems. A second benefit will emerge in Chapter 5 in the spectral analysis of the components comprising the tensor product.

The description of this matrix system is facilitated upon recalling that for a continuous function f on $[-\pi,\pi]$ the Fourier coefficients of f are given by the sequence

$$w_p = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\exp(-ipx)dx \quad , \quad p = 0,\pm 1,\dots \quad . \tag{4.4}$$

Lemma 4.1: *Let* $\{w_{k-j}\}$ *and* $\{v_{k-j}\}$ *denote the Fourier coefficients of* $w(x) = -ix$ *and* $v(x) = -x^2$, *respectively. Then*

$$h\,\frac{d}{d\chi}\left(S(j,h)\circ\chi(x)\right)\Big|_{x=x_k} = w_{k-j} = \begin{cases} 0 & , \ j = k \\ \frac{(-1)^{k-j}}{k-j} & , \ j \neq k \end{cases} \tag{4.5}$$

and

$$h^2\,\frac{d^2}{d\chi^2}\left(S(j,h)\circ\chi(x)\right)\Big|_{x=x_k} = v_{k-j} = \begin{cases} \frac{-\pi^2}{3} & , \ j = k \\ \frac{(-1)^{k-j}}{(k-j)^2} & , \ j \neq k \end{cases} \tag{4.6}$$

Proof: The representation

$$S(j,h)(x) = \frac{h}{2\pi}\int_{-\pi/h}^{\pi/h} \exp(-i(x-jh)t)dt$$

in (2.6) leads to the following derivative calculations

$$S'(j,h)(x) = \frac{h}{2\pi}\int_{-\pi/h}^{\pi/h} -it\exp(-i(x-jh)t)dt$$

and

$$S''(j,h)(x) = \frac{h}{2\pi}\int_{-\pi/h}^{\pi/h} -t^2\exp(-i(x-jh)t)dt \quad .$$

The substitutions $x = kh$ in the above derivatives, together with (4.4) lead to the identities (4.5) and (4.6), respectively.

Definition 4.2: *The matrix* $I^{(p)}$, $p = 0, 1, 2$ *is the* $m \times m$ *Toeplitz matrix of elements*

$$\delta_{j,k}^{(p)} = h^p S^{(p)}(j,h)(kh) \quad .$$

These are the $(k-j)$*-th Fourier coefficients of the function* $\omega(x) = (-ix)^p$, $x \in [-\pi, \pi]$. *In particular, the matrix* $I^{(0)}$ *is the* $m \times m$ *identity matrix. The matrix* $I^{(1)}$ *is the* $m \times m$, *Toeplitz, skew symmetric matrix which has as its* jk*-th element* $h\,S'(j,h)(kh)$. *Explicitly,*

$$I^{(1)} \equiv \begin{bmatrix} 0 & -1 & 1/2 & -1/3 & \cdots & \frac{(-1)^{m-1}}{m-1} \\ 1 & 0 & -1 & 1/2 & & \\ -1/2 & 1 & 0 & -1 & & \\ 1/3 & -1/2 & 1 & 0 & & \vdots \\ \vdots & & & & \ddots & -1 \\ \frac{(-1)^m}{m-1} & & \cdots & \cdots & 1 & 0 \end{bmatrix} .$$

Finally, the matrix $I^{(2)}$ *is the* $m \times m$, *Toeplitz, symmetric matrix which has as its* jk*-th element* $h^2 S''(j,h)(kh)$. *Explicitly,*

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

Besides the matrices in Definition 4.2, the following notational conventions will be adopted for the remainder of this thesis. Let f be a function defined on $(0,1)\times(0,\infty)$, and denote the points in $(0,1)$ and $(0,\infty)$ by $\{x_k\}$, $k = -M_x \ldots N_x$, and $\{t_j\}$, $j = -M_t \ldots N_t$ respectively. Recall the definitions of these points from Definition 3.1. For a function f let $[f]$ be the matrix which has as elements the pointwise evaluation of the function f, i.e. $f\left(x_{k-M_x-1}, t_{j-M_t-1}\right)$, $k = 1 \ldots m_x$, $j = 1 \ldots m_t$, where $m_x = M_x + N_x + 1$ and $m_t = M_t + N_t + 1$. Note that $[f]$ is $m_x \times m_t$. Similarly, let $[< f >]$ represent the matrix of dimension $m_x \times m_t$ of inner products $< f >_{kj}$, where $< f >_{kj}$ is given in Definition 3.2. If g is a function of one variable, let

$$D[g] = [\delta_{ij} g(q_i)] \quad , \quad 1 < i, j < m_q$$

where $q = x$ or t. The matrix $D[g]$ is a $m_q \times m_q$ diagonal matrix whose diagonal entries consist of the node evaluations of the function g.

With the notation of the above paragraph, the inner product in (3.25) has the matrix representation

$$[< u_t >] = -D\left[\frac{w}{\phi'}\right][u] D[\hat{w}] \left\{ \frac{1}{\hat{h}} I^{(1)} + D\left[\frac{\hat{w}'}{\hat{\phi}'\hat{w}}\right] \right\}^t \quad . \tag{4.7}$$

The four matrices on the right-hand side of (4.7) have dimension $m_x \times m_x$, $m_x \times m_t$, $m_t \times m_t$, and $m_t \times m_t$, respectively. Similarly, the matrix representation of (3.26)

is given by

$$
\begin{aligned}
[< u_{xx} >] \;=\; & \left\{ \frac{1}{h^2}\, I^{(2)} + \frac{1}{h}\, I^{(1)} D \left[\frac{\phi''}{(\phi')^2} + \frac{2w'}{\phi' w} \right] \right. \\
& \left. + D \left[\frac{w''}{w(\phi')^2} \right] \right\} D[\phi' w][u] D \left[\frac{\hat{w}}{\hat{\phi}'} \right]
\end{aligned} \tag{4.8}
$$

where the first five matrices on the right-hand side of (4.8) have dimension $m_x \times m_x$, and the remaining two matrices are of dimension $m_x \times m_t$, and $m_t \times m_t$, respectively.

For reasons which will become apparent in the compact representation of the discrete sinc system for the residual in (4.2), it is convenient to define the temporal matrix

$$
B(\hat{w}) = D \left[\sqrt{\hat{\phi}'} \right] \left\{ \frac{1}{\hat{h}}\, I^{(1)} + D \left[\frac{\hat{w}'}{\hat{\phi}' \hat{w}} \right] \right\} D \left[\sqrt{\hat{\phi}'} \right] \tag{4.9}
$$

and the spatial matrix

$$
\begin{aligned}
A(w) = D[\phi'] & \left\{ \frac{1}{h^2}\, I^{(2)} + \frac{1}{h}\, I^{(1)} D \left[\frac{\phi''}{(\phi')^2} + \frac{2w'}{\phi' w} \right] \right. \\
& \left. + D \left[\frac{w''}{w(\phi')^2} \right] \right\} D[\phi'] \quad .
\end{aligned} \tag{4.10}
$$

With this notation the discrete Galerkin representation of Pu_A in (4.1) is given by the $m_x \times m_t$ matrix

$$
\begin{aligned}
[< Pu_A >] \;=\; & -D[w/\phi'][u] D \left[\hat{w} / \sqrt{\hat{\phi}'} \right] B^t(\hat{w}) D \left[1/\sqrt{\hat{\phi}'} \right] \\
& - D[1/\phi'] A(w) D(w)[u] D[\hat{w}/\hat{\phi}'] \\
= \; & -D \left[\frac{1}{\phi'} \right] \left\{ D(w)[u] D \left[\frac{\hat{w}}{\sqrt{\hat{\phi}'}} \right] \right\} B^t(\hat{w}) D \left[\frac{1}{\sqrt{\hat{\phi}'}} \right] \\
& - D \left[\frac{1}{\phi'} \right] A(w) \left\{ D[w][u] D \left[\frac{\hat{w}}{\sqrt{\hat{\phi}'}} \right] \right\} D \left[\frac{1}{\sqrt{\hat{\phi}'}} \right] \\
= \; & -D \left[\frac{1}{\phi'} \right] \left\{ V(w, \hat{w}) B^t(\hat{w}) + A(w) V(w, \hat{w}) \right\} D \left[\frac{1}{\sqrt{\hat{\phi}'}} \right]
\end{aligned} \tag{4.11}
$$

where the "weighted" coefficients for (4.3) are given by

$$V(w,\hat{w}) \equiv D[w][u]D\left[\frac{\hat{w}}{\sqrt{\hat{\phi}'}}\right] \quad . \tag{4.12}$$

When the function on the right-hand side of (4.1) is independent of u, the approximate inner product in (3.27) shows that

$$[< f >] \; = D\left[\frac{1}{\phi'}\right] F(w,\hat{w}) D\left[\frac{1}{\sqrt{\hat{\phi}'}}\right] \tag{4.13}$$

where

$$F(w,\hat{w}) \equiv D[w][f(x_k,t_j)]D\left[\frac{\hat{w}}{\sqrt{\hat{\phi}'}}\right] \tag{4.14}$$

is the $m_x \times m_t$ matrix of "weighted" point evaluations of f at the spatial nodes x_k and the temporal nodes t_j. Upon setting the right-hand side of (4.13) equal to the right-hand side of (4.11) followed by a pre and post multiplication by $D[\phi']$ and $D\left[\sqrt{\hat{\phi}'}\right]$, respectively, the Sinc-Galerkin system for (4.1) with (3.2) takes the form

$$A(w)V(w,\hat{w}) + V(w,\hat{w})B^t(\hat{w}) = -F(w,\hat{w}) \quad . \tag{4.15}$$

This is a diagonal perturbation of the system in [9]. The only difference is in the definition of $B(\hat{w})$ in (4.9). Indeed the $B(\hat{w})$ in (4.9) is related to the temporal matrix in [9] via the product $D\left[\sqrt{\hat{\phi}'}\right] B(\hat{w}) D\left[\left(\sqrt{\hat{\phi}'}\right)^{-1}\right]$. The advantages of the definition in (4.9) over that in [9] will become more transparent in the spectral analysis of Chapter 5. Briefly, it will be shown that not only is $B(\hat{w})$ invertible but also, if $\beta \in \sigma(B(\hat{w}))$ then $\mathrm{Re}(\beta) < 0$. This mathematical detail was only numerically settled in [9].

Before incorporating the specific maps given in Table 1 in the system (4.15) the matrix representations of the discretizations (3.28), (3.29) and (3.31) take the respective forms

$$[< u >] = D\left[\frac{w}{\phi'}\right][u]D\left[\frac{\hat{w}}{\hat{\phi}'}\right] \quad , \tag{4.16}$$

$$[< u_x >] = -\left\{\frac{1}{h}\, I^{(1)} D[w] + D\left[\frac{w'}{\phi'}\right]\right\}[u]D\left[\frac{\hat{w}}{\hat{\phi}'}\right] \tag{4.17}$$

and

$$\left[< \frac{\partial}{\partial x}(u)^2 >\right] = -\left\{\frac{1}{h}\, I^{(1)} D[w] + D\left[\frac{w'}{\phi'}\right]\right\}[u^2]D\left[\frac{\hat{w}}{\hat{\phi}'}\right] \quad . \tag{4.18}$$

The special cases of (4.15) resulting from the use of the weighting functions discussed at the close of Chapter 3 for the spatial and temporal matrices $A(w)$ and $B(\hat{w})$ take on particularly simple forms. It is observed that the matrix $A(w)$ in (4.10) is symmetric if w is chosen to be $1/\sqrt{\phi'}$. Lund [10] has shown that there are cases where the choice of weight $w = 1/\phi'$ is preferable, but for the problems of this thesis this choice plays a subordinate role. It is recorded in Table 1 in light of the remarks in the paragraph opening Chapter 5. Hence, define the spatial matrix corresponding to $w = 1/\sqrt{\phi'}$

$$A \equiv D[\phi']\left\{\frac{1}{h^2}\, I^{(2)} - \frac{1}{4}\, I^{(0)}\right\} D[\phi'] \quad . \tag{4.19}$$

The matrix $B(\hat{w})$ in (4.9) takes the form

$$B \equiv D\left[\sqrt{\hat{\phi}'}\right]\left\{\frac{1}{\hat{h}}\, I^{(1)} + \frac{1}{2}\, D\left[\hat{\phi}''/\left(\hat{\phi}'\right)^2\right]\right\} D\left[\sqrt{\hat{\phi}'}\right] \tag{4.20}$$

since, whether one selects the temporal map $\hat{\phi}(t) = \ell n(t)$ or $\hat{\phi}(t) = \ell n(\sinh(t))$, the weight $\hat{w}(t) = \sqrt{\hat{\phi}'(t)}$ shows that the diagonal matrix $D\left[\hat{w}(t)/\sqrt{\hat{\phi}'(t)}\right]$ is the

$m_t \times m_t$ identity matrix. The matrix of weighted coefficients $V(w, \hat{w})$ in (4.12) then takes the form

$$V \equiv D\left[1/\sqrt{\phi'}\right][u] \quad . \tag{4.21}$$

The symbols A, B and V will be reserved for the spatial matrix in (4.19), the temporal matrix in (4.20) and the coefficient matrix in (4.21), respectively. That is, these are the matrices corresponding to the choices of weight functions of this thesis. The final matrix form taken by the terms on the right-hand side of (4.1) corresponding to the discretizations in (4.14), (4.16), (4.17) and (4.18) (recalling for the latter three that there is a pre and post multiplication of (4.11) by $D[\phi']$ and $D\left[\sqrt{\hat{\phi}'}\right]$, respectively), reads

$$[< f >]_S \equiv F\left(\frac{1}{\sqrt{\phi'}}, \sqrt{\hat{\phi}'}\right) = D\left[1/\sqrt{\phi'}\right][f] \quad , \tag{4.22}$$

$$[< u >]_S = D\left[1/\sqrt{\phi'}\right][u] \quad , \tag{4.23}$$

$$[< u_x >]_S = -\left\{\frac{1}{h} D[\phi'] I^{(1)} D\left[1/\sqrt{\phi'}\right] + D\left[-\phi''/\left(2(\phi')^{3/2}\right)\right]\right\}[u] \tag{4.24}$$

and

$$\begin{aligned}[< u_x^2 >]_S = -\Bigg\{&\frac{1}{h} D[\phi'] I^{(1)} D\left[1/\sqrt{\phi'}\right] \\ &+ D\left[-\phi''/\left(2(\phi')^{3/2}\right)\right]\Bigg\}[u^2] \quad .\end{aligned} \tag{4.25}$$

The notation $[< \cdot >]_S$ denotes the system entries as, for example, in (4.26). The matrices A and B corresponding to these particular choices of the weights with the mapping functions specified are summarized in Tables 3 and 4.

Spatial Matrices

$\phi(x) = \ell n \left[\frac{x-a}{b-x}\right]$ $\qquad$ $\phi'(x) = \frac{(b-a)}{(x-a)(b-x)}$

$$x_k = \frac{b\exp(kh)+a}{\exp(kh)+1}$$

for $w = 1/\sqrt{\phi'}$:

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

for $w = 1/\phi'$:

$$A(1/\phi') = D[\phi']\left\{\frac{1}{h^2}\ I^{(2)} + \frac{1}{h}\ I^{(1)} D\left(\frac{a+b-2x}{b-a}\right) - 2D\left(\frac{(x-a)(b-x)}{(b-a)^2}\right)\right\} D[\phi']$$

Table 3. The spatial matrix $A(w)$ in (4.10) corresponding to $w = 1/\sqrt{\phi'}$ and $w = 1/\phi'$.

Temporal Matrices		
$\hat{w} = \sqrt{\hat{\phi}'}$		
$\hat{\phi}(t) = \ell n(t)$	$\hat{\phi}'(t) = 1/t$	$t_j = \exp(j\hat{h})$
$B\left(\sqrt{\hat{\phi}'}\right) = D\left[\sqrt{\hat{\phi}'}\right]\left\{\frac{1}{\hat{h}}\, I^{(1)} - \frac{1}{2}\, I^{(0)}\right\} D\left[\sqrt{\hat{\phi}'}\right]$		
$\hat{\phi}(t) = \ell n(\sinh(t))$		$\hat{\phi}'(t) = \coth(t)$
$t_j = j\hat{h} + \ell n\left(1 + \sqrt{1 + \exp(-2j\hat{h})}\right)$		
$B = D\left[\sqrt{\hat{\phi}'}\right]\left\{\frac{1}{\hat{h}}\, I^{(1)} - \frac{1}{2}\, D[\mathrm{sech}^2]\right\} D\left[\sqrt{\hat{\phi}'}\right]$		

Table 4. The temporal matrix $B(\hat{w})$ in (4.9) corresponding to $w = \sqrt{\hat{\phi}'}$ for both $\hat{\phi}(t) = \ell n(t)$ and $\hat{\phi}(t) = \ell n(\sinh(t))$.

As an example here, to solve (4.1) when

$$f(x,t,u) = u(x,t) + g(x,t)$$

one combines (4.19), (4.20) and (4.21) in the general system (4.15) with (4.22) and (4.23) to write

$$\begin{aligned} AV + VB^t &= -[< f >]_S \; - \; [< u >]_S \\ &= -D\left[1/\sqrt{\phi'}\right]([f] + [u]) \quad . \end{aligned} \tag{4.26}$$

This is precisely the system used in each of Examples 2 and 4 of Chapter 6.

The matrix on the left-hand side of (4.26) has a particularly nice form in terms of tensor products. Towards this end it is helpful to review some matrix algebra. If A and B are $m \times p$ and $q \times n$ matrices respectively, then the Kronecker product of A and B is given by

$$A \otimes B = [a_{ij}B]_{mq \times pn} \quad .$$

The linear operator 'co' acting on a matrix A is called the concatenation of A and is defined by

$$\mathrm{co}(A) = \begin{bmatrix} a_{i1} \\ a_{i2} \\ a_{i3} \\ \vdots \\ a_{ip} \end{bmatrix}_{mp \times 1}$$

where a_{ik} is the k-th column of A. A fundamental identity relating tensor products and concatenation is given by the identity

$$\mathrm{co}(AVB^t) = (B \otimes A)\mathrm{co}(V) \tag{4.27}$$

where the matrices A, V and B^t are conformable for multiplication.

As an immediate application of the concatenation operator and the tensor product, denote the right-hand side of (4.26) by $F\ (m_x \times m_t)$ and rewrite (4.26) in the equivalent algebraic form

$$\text{co}(F) = \{I_{m_t} \otimes A) + (B \otimes I_{m_x})\}\ \text{co}(V) \quad . \tag{4.28}$$

There are a number of useful results concerning the system (4.26) that are more directly discernable via using (4.28) (Chapter 5 exploits this connection). The item of interest here is the connection between the problem (4.1) and certain one-dimensional problems. To see this connection, it is helpful to think of the left-hand side of (4.28) as the Kronecker sum of the discretization of the two one-dimensional problems

$$u''(x) = f(x) \quad , \quad u(a) = u(b) = 0 \tag{4.29}$$

and

$$u'(t) = g(t) \quad , \quad u(0) = 0 \quad . \tag{4.30}$$

It is assumed that (4.30) has a unique solution with $u(\infty) = 0$. That is, if the integrations by parts performed in Chapter 3 for the functions of two variables are applied to the problems (4.29) and (4.30), the results are the approximations

$$A_x[v^x] = D\left(1/\sqrt{\phi'}\right)[f] \tag{4.31}$$

and

$$B[u^t] = [g] \quad , \tag{4.32}$$

respectively. The matrices A_x and B are the same as the matrices A and B in (4.19) and (4.20). The subscripted x on the spatial matrix A will be clarified

shortly. The vector $[v^x]$ in (4.31) is given by $D\left(1/\sqrt{\phi'}\right)[u^x]$. In this case $[u^x]$ as well as $[u^t]$ in (4.32) are the point evaluations of the Sinc-Galerkin approximation to the solution of (4.29) and (4.30) by (4.31) and (4.32), respectively. The tensor product of $[u^x]$ with $[u^t]$ defines the approximate solution (4.3) and the Kronecker sum of the discrete systems in (4.29) and (4.30) defines the discrete system for the problem (4.1); i.e. (4.26).

With this connection between one-dimensional and higher dimensional problems it is quite direct to write down the discretization of the two spatial dimension problem

$$u_t(x,y,t) - \Delta^2 u(x,y,t) = f(x,y,t) \tag{4.33}$$

with the homogeneous conditions

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

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

$$u(x,y,0) = 0 \quad .$$

Based on Definition 3.2, the inner product used in orthogonalizing the residual is

$$(u,v) = \int_0^\infty \int_0^1 \int_0^1 u(x,y,t)v(x,y,t)w(x,y,t)dxdydt \quad ,$$

and the weight w is defined by

$$w(x,y,t) = \left[\hat{\phi}'(t)/(\phi'(x)\phi'(y))\right]^{1/2} \quad .$$

The analogue of (4.3) takes the form

$$(4.34) \qquad u_A(x,y,t) = \sum_{i=-M_x}^{N_x} \sum_{j=-M_y}^{N_y} \sum_{k=-M_t}^{N_t} u_{ijk} S_i(x) S_j(y) \hat{S}_k(t)$$

for the assumed approximate solution of (4.33).

A direct development of the sinc discrete system defining the pointwise approximate solution $[u]_{ijk}$ of (4.33) that is based on the various integrations by parts leading to Theorem 3.3 is both tedious and lengthy. The connection between one-dimensional problems and a multiple dimension one, discussed in the paragraph following (4.28), yields the same discrete system, that is

$$(4.35) \quad \mathrm{co}(F) = \{(I_{m_t} \otimes I_{m_y} \otimes A_x) + (I_{m_t} \otimes A_y \otimes I_{m_x}) + (B \otimes I_{m_y} \otimes I_{m_x}\}\mathrm{co}(V)$$

where

$$(4.36) \qquad \begin{aligned} \mathrm{co}(V) &= \mathrm{co}\left([v_{ijk}]\right) \\ &= \begin{bmatrix} \mathrm{co}\left([V_{ij,-M_t}]\right) \\ \vdots \\ \mathrm{co}\left([V_{ij,0}]\right) \\ \vdots \\ \mathrm{co}\left([V_{ij,N_t}]\right) \end{bmatrix} . \end{aligned}$$

In (4.35), I_q is a $q \times q$ identity matrix ($q = m_x$, m_y or m_t), A_q is the $q \times q$ matrix in (4.19) ($q = x$ or y depending on whether the mapping ϕ_x or ϕ_y is used) and B is the matrix (4.20). To recover the coefficients $[u_{ijk}]$ in (4.34) from (4.36) the transformation

$$(4.37) \qquad \mathrm{co}(V) = \left\{ I_{m_t} \otimes D\left[1/\sqrt{\phi_y'}\right] \otimes D\left[1/\sqrt{\phi_x'}\right] \right\} \mathrm{co}(U)$$

connects V to U (two-dimensional analogue of (4.21)).

With regard to the error in approximating the true solution of (4.33) by (4.34), an iteration of the argument leading to Theorem 3.3 is all that is required.

The necessary assumptions on $u(x,y,t)$ to obtain exponential convergence rates are exactly those listed in the hypothesis of Theorem 3.3 (with the addition of the corresponding assumptions in the y-variable). A succinct development of this sort of error analysis may be found in [18]. A numerical illustration of this convergence rate is the content of Example 6 in Chapter 6.

CHAPTER 5

SPECTRAL ANALYSIS OF THE SINC-GALERKIN SYSTEM

This chapter addresses the solvability of the system in (4.26). The specific system (4.26), as opposed to the more general system (4.15), is concentrated upon as it is the system implemented in Chapter 6 for the numerical results. This is not to discount the potential use of (4.15); e.g. different maps ϕ or $\hat{\phi}$ or (the more likely event) different choices of the weighting functions w or $\hat{w}$. For example, it was mentioned in the paragraph following (4.18) that the spatial weight of choice in this thesis is given by $w(x) = (\phi'(x))^{-1/2}$. This, as already noted, is due to the resultant symmetry of the matrix A in (4.19). In the original development of the Sinc-Galerkin method [16] as well as in the study in both [10] and [11] the choice $w(x) = (\phi'(x))^{-1}$ plays a prominent role. Indeed in [11] it is shown that the luxury of symmetry is not free. The class of problems to which the weight $w(x) = (\phi'(x))^{-1}$ is applicable (and maintains exponential convergence) is wider than the class for which one can guarantee exponential convergence when using the weight $w(x) = (\phi'(x))^{-1/2}$. When the problem being solved can be handled using either weight (which is the case for the problems of this thesis), symmetry is too nice a property to sacrifice. In particular, the methods of analysis carried out in this chapter do not apply to the matrix $A(1/\phi')$ in Table 3.

To begin this analysis, recall the definitions

$$A = D[\phi'] \left\{ \frac{1}{h^2} I^{(2)} - \frac{1}{4} I^{(0)} \right\} D[\phi'] \quad , \tag{5.1}$$

and

$$B = D\left[\sqrt{\hat{\phi}'}\right]\left\{\frac{1}{\hat{h}}\, I^{(1)} + \frac{1}{2}\, D\left[\frac{\hat{\phi}''}{(\hat{\phi}')^2}\right]\right\} D\left[\sqrt{\hat{\phi}'}\right] \quad , \tag{5.2}$$

which are the matrices in (4.19) and (4.20), respectively. These matrices define the system (4.26); i.e.

$$AV + VB^t = F \quad . \tag{5.3}$$

The matrices A, B and V are $m_x \times m_x$, $m_t \times m_t$ and $m_x \times m_t$, respectively. The matrix

$$V = D\left[1/\sqrt{\phi'}\right][u]$$

is given in (4.21) and F is a generic $m_x \times m_t$ matrix which can take any one of the forms in (4.22) through (4.25) or linear combinations of these matrices.

To show that the system (5.3) is always solvable, the most convenient approach is to use the tensor form of (5.3)

$$(I \otimes A + B \otimes I)\mathrm{co}(V) = \mathrm{co}(F) \quad . \tag{5.4}$$

In this form it can be seen that a unique solution to (5.3) depends on the invertability of the matrix $I \otimes A + B \otimes I$. This in turn depends on the comingling of the eigenvalues of the matrices A and B. Specifically, a unique solution of (5.4) exists only if no eigenvalue of A is the negative of an eigenvalue of B since the eigenvalues of the Kronecker sum on the left-hand side of (5.4) are the sums of the eigenvalues of A and B [8]. To prove that this is the case requires an investigation of the eigenvalues of the matrices A and B. The next theorem gives an initial handle on this eigenvalue problem.

Theorem 5.1: *If H is a Hermitian $m \times m$ matrix then for any $x \in \mathbb{C}^N$,*

$$\lambda_1 \leq \frac{x^* H x}{x^* x} \leq \lambda_m \quad .$$

where the real eigenvalues of H have been ordered $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m$.

Proof: Since H is Hermitian, there is an orthonormal basis of eigenvectors of H for $\mathbb{C}^N$. Let (λ_i, v_i) be the set of eigenpairs of H so that

$$Hv_i = \lambda_i v_i$$

for each $i = 1, \ldots, m$ and $v_i^* v_j = \delta_{ij}$. Then for any x in $\mathbb{C}^N$ there is a unique set of $\alpha_i \in \mathbb{C}$, $i = 1, \ldots, m$ so that $x = \sum_{i=1}^m \alpha_i v_i$. Since $x^* x = \sum_{i=1}^m \overline{\alpha}_i \alpha_i$ and $Hx = \sum_{i=1}^m \lambda_i \alpha_i v_i$ it follows that

$$\begin{aligned}
\lambda_1 x^* x &= \lambda_1 \sum_{i=1}^m \overline{\alpha}_i \alpha_i \\
&\leq \sum_{i=1}^m \lambda_i \overline{\alpha}_i \alpha_i \\
&= x^* H x \\
&\leq \lambda_m \sum_{i=1}^m \overline{\alpha}_i \alpha_i \\
&= \lambda_m x^* x \quad .
\end{aligned}$$

The result of Theorem 5.1 can be used to explore the properties of the matrix A defined in (5.1) directly since A is a Hermitian (symmetric) matrix.

Theorem 5.2: *The matrix A defined in (5.1) is negative definite. If $\alpha \in \sigma(A)$ then*

$$\alpha \leq -4/(b-a)^2 \quad .$$

Proof: $I^{(2)}$ is Toeplitz and constructed from the first m Fourier coefficients of the function $f(x) = -x^2$ (Lemma 4.1). The Szego-Grenander Theorem [7] guarantees that the $\sigma(I^{(2)}) \subset [-\pi^2, 0]$. Since $h > 0$ the matrix $C = \frac{1}{h^2} I^{(2)} - \frac{1}{4} I^{(0)}$ is negative definite with largest eigenvalue bounded above by $-1/4$. Since $D[\phi']$ has diagonal entries given by the point evaluations of $\phi'(x)$ and $\phi'(x) \geq 4/(b-a)$ it follows that $D[\phi']CD[\phi'] = A$ is also negative definite. If α is the largest eigenvalue of A then by Theorem 5.1

$$\begin{aligned}
\alpha &= \max_{x \neq 0} \left(\frac{x^* A x}{x^* x} \right) \\
&= \max_{x \neq 0} \left(\frac{x^* D[\phi'] C D[\phi'] x}{x^* x} \right) \\
&= \max_{x \neq 0} \left(\frac{x^* D[\phi'] C D[\phi'] x}{x^* D[\phi'] D[\phi'] x} \cdot \frac{x^* D[\phi'] D[\phi'] x}{x^* x} \right) \\
&\leq \max_{x \neq 0} \left(\frac{x^* D[\phi'] C D[\phi'] x}{x^* D[\phi'] D[\phi'] x} \right) \cdot \min_{x \neq 0} \left(\frac{x^* D[\phi'] D[\phi'] x}{x^* x} \right) \\
&= \max_{w \neq 0} \left(\frac{w^* C w}{w^* w} \right) \cdot \min_{x \neq 0} \left(\frac{x^* D[\phi'] D[\phi'] x}{x^* x} \right) \\
&\leq (-1/4) \cdot (4/(b-a))^2 \\
&= -4/(b-a)^2 \quad .
\end{aligned}$$

The matrix B in (5.2) requires a more delicate analysis as B is not Hermitian. However, B is easily divided into Hermitian and skew Hermitian parts since

$$D\left[\sqrt{\hat{\phi}'}\right] I^{(1)} D\left[\sqrt{\hat{\phi}'}\right]$$

is skewsymmetric and real, and

$$D\left[\sqrt{\hat{\phi}'}\right] D\left(\frac{\hat{\phi}''}{(\hat{\phi}')^2}\right) D\left[\sqrt{\hat{\phi}'}\right] = D\left(\frac{\hat{\phi}''}{\hat{\phi}'}\right)$$

is real and symmetric. Suppose (β, v) is an eigenpair for B. Then

$$v^* B v = \beta v^* v$$

and

$$v^* B^* v = \overline{\beta} v^* v \quad .$$

From this it can be seen that

$$v^*(B + B^*)v/2 = \mathrm{Re}\{\beta\} v^* v$$

and that

$$v^*(B - B^*)v/2 = i\,\mathrm{Im}\{\beta\} v^* v \quad .$$

This together with the following lemma completes the analysis of the spectrum of B.

Lemma 5.3: *The real part of the spectrum of the matrix B in (5.2) is nonpositive.*

Proof: The matrix B in (5.2) corresponds to the selection of the weight function $\sqrt{\hat{\phi}'}$ where the two choices for $\hat{\phi}'(t)$ are $\ell n(t)$ and $\ell n(\sinh(t))$. In either case the diagonal matrix $D\left[\hat{\phi}''/(2\hat{\phi}')\right]$ is negative (Table 4). In the computation below, the first line corresponds to $\hat{\phi}(t) = \ell n(t)$ and the second line to $\hat{\phi}(t) = \ell n(\sinh(t))$, where for $\ell n(t)$, $D\left(\hat{\phi}''/(\hat{\phi}')^2\right) = -I^{(0)}$ and for $\ell n(\sinh(t))$, $D\left(\hat{\phi}''/(\hat{\phi}')^2\right) = -D(\mathrm{sech}^2)$. Let (β, v) be an eigenpair for B then

$$
\begin{aligned}
\mathrm{Re}\{\beta\}v^*v &= v^*(B+B^*)v/2 \\
&= \frac{1}{2}\, v^* D\left[\sqrt{\hat{\phi}'}\right] D\left(\frac{\hat{\phi}''}{(\hat{\phi}')^2}\right) D\left[\sqrt{\hat{\phi}'}\right] \\
&= \begin{cases} \frac{-1}{2}\, v^* D\left[\sqrt{\hat{\phi}'}\right] I^{(0)} D\left[\sqrt{\hat{\phi}'}\right] v \\ \frac{-1}{2}\, v^* D\left[\sqrt{\hat{\phi}'}\right] D[\mathrm{sech}^2] D\left[\sqrt{\hat{\phi}'}\right] v \end{cases} \\
&= \begin{cases} \frac{-1}{2}\left(v^* D[\hat{\phi}']v\right) \\ \frac{-1}{2}\left(v^* D[\hat{\phi}'\,\mathrm{sech}^2]v\right) \end{cases} \\
&\leq \begin{cases} \frac{-1}{2}\left(\min \hat{\phi}'(t)\right) v^*v \\ \frac{-1}{2}\left(\min \hat{\phi}'(t)\mathrm{sech}^2(t)\right) v^*v \end{cases} \\
&\leq 0\, v^*v
\end{aligned}
$$

since, in the case of either map, $\hat{\phi}'(t) > 0$ on $(0, \infty)$.

It is now possible to analyze $I \otimes A + B \otimes I$ using the results of Theorem 5.2 and Lemma 5.3. This analysis gives the following theorem.

Theorem 5.4: *Let the matrices A and B be as defined in (5.1) and (5.2), respectively. Then if λ is in the spectrum of $I \otimes A + B \otimes I$,*

$$\mathrm{Re}\{\lambda\} \leq -4/(b-a)^2 \quad .$$

Proof: If $\lambda \in \sigma(I \otimes A + B \otimes I)$ then there are elements α and β of the spectrums of A and B, respectively, so that $\lambda = \alpha + \beta$ where $\alpha \leq -4/(b-a)^2$ and $\mathrm{Re}\{\beta\} \leq 0$. Hence

$$
\begin{aligned}
\mathrm{Re}\{\lambda\} &= \mathrm{Re}\{\alpha+\beta\} = \alpha + \mathrm{Re}\{\beta\} \\
&\leq -4/(b-a)^2 \quad .
\end{aligned}
$$

Corollary 5.5: *The matrix*

$$\mathcal{A} \equiv I \otimes A + B \otimes I$$

where A and B are given by (5.1) and (5.2), respectively, is invertible and if $\lambda \in \sigma(\mathcal{A}^{-1})$ then

$$|\lambda| \leq \frac{(b-a)^2}{4} \quad .$$

The extension of Corollary 5.5 to two or higher dimensions is straightforward. Recall that the system in (4.35) for the two-dimensional problem (4.33) is a tensor sum of A_x, A_y and B where each of A_q $(q = x \text{ or } y)$ is the matrix A of Theorem 5.2 and B is the same matrix as the matrix in Lemma 5.3. Hence the real part of any eigenvalue of the coefficient matrix in (4.35) for the two-dimensional problem is bounded above by $-4\left((b_1 - a_1)^{-2} + (b_2 - a_2)^{-2}\right)$ where the spatial domain of (4.33) is $(a_1, b_1) \times (a_2, b_2)$.

The spectral bound in Corollary 5.5 is somewhat gross in the sense that numerical evidence indicates the spectral minimum is near the original differential operator's spectral minimum of $-\pi^2$ $(a = 0, b = 1)$. Corollary 5.5 says that the inverse matrix for this case will have a spectral radius less than $1/4$, and in practice about $1/\pi^2$. This advertises the potential for using an iterative method to solve (4.26). Numerical results of this application are tested in three of the examples in the next chapter.

CHAPTER 6

THE NUMERICAL PROCEDURE AND EXAMPLES

In this chapter the performance of the numerical algorithm defined by the system

$$AV + VB^t = F \tag{6.1}$$

where

$$V = D\left[1/\sqrt{\phi'}\right]U \tag{6.2}$$

and F is any one of the terms in (4.22)–(4.25) is tested in the construction of the approximate solution

$$u_A(x,t) = \sum_{\ell=-M_x}^{N_x} \sum_{p=-M_t}^{N_t} u_{\ell p} S(\ell,h) \circ \phi(x) S(p,\hat{h}) \circ \hat{\phi}(t) \tag{6.3}$$

to the true solution of

$$Pu(x,t) \equiv u_t(x,t) - u_{xx}(x,t) = f(x,t,u(x,t))$$

$$u(0,t) = u(1,t) = 0 \tag{6.4}$$

$$u(x,0) = g(x) \quad .$$

To fully define the system in (6.1) selections of the parameters M_x, N_x, h, M_t, N_t and $\hat{h}$ need to be specified. It is assumed that $d = \hat{d} = \pi/2$ (Figures 1 and 2) so that the line following (3.27) dictates

$$h = \hat{h} \quad . \tag{6.5}$$

From the lines following the approximations in (3.13) and (3.16), the exponential errors are different. To balance the errors with respect to order, the following selections, in conjunction with (6.5), balance all of the error contributions. Given M_x select

$$h = \left(\frac{\pi d}{\alpha M_x}\right)^{1/2} = \pi/\sqrt{2\alpha M_x} \quad , \quad N_x = \left[\left|\frac{\alpha}{\beta}\, M_x + 1\right|\right] \tag{6.6}$$

and

$$M_t = \left[\left|\frac{\alpha}{\delta}\, M_x + 1\right|\right] \quad , \quad N_t = \left[\left|\frac{\alpha}{\mu}\, M_x + 1\right|\right] \quad . \tag{6.7}$$

The parameters α and β are the decay constants for $u(x,*)$ given in Definition 2.12, while δ and μ are the analogues for $u(*,t)$. For the particular maps ϕ and $\hat{\phi}$ these constants are identified in Table 1.

The actual method of solution for (6.1) used for this thesis involves the diagonalization of the matrices A and B. The procedure is as follows: Assume that Q and P^t are invertible matrices that diagonalize A and B, respectively, i.e.

$$A = Q\Lambda_\alpha Q^{-1} \tag{6.8}$$

and

$$B = P^t \Lambda_\beta (P^t)^{-1} \tag{6.9}$$

where Λ_q ($q = \alpha$ or β) is the diagonal matrix of eigenvalues of A or B. Substitution of (6.8) and (6.9) into (6.1) leads to

$$\begin{aligned} F &= AV + VB^t \\ &= Q\Lambda_\alpha Q^{-1}V + VP^{-1}\Lambda_\beta P \\ &= Q\left\{\Lambda_\alpha (Q^{-1}VP^{-1}) + (Q^{-1}VP^{-1})\Lambda_\beta\right\}P \\ &\equiv Q\left\{E \circ (Q^{-1}VP^{-1})\right\}P \quad . \end{aligned} \tag{6.10}$$

In (6.10) the eigenvalue matrix E has the entries

$$[E]_{ij} \equiv [\alpha_i + \beta_j]_{m_x \times m_t} \tag{6.11}$$

consisting of the sums of the eigenvalues of A, $\{\alpha_1, \ldots, \alpha_{m_x}\}$, $m_x = M_x + N_x + 1$ and the eigenvalues of B, $\{\beta_1, \ldots, \beta_{m_t}\}$, $m_t = M_t + N_t + 1$. The Hadamard product $E \circ (Q^{-1}VP^{-1})$ is the element by element product of the matrix E with $Q^{-1}VP^{-1}$. The matrix

$$E_H \equiv \left[\frac{1}{\alpha_i + \beta_j}\right] \tag{6.12}$$

is the Hadamard inverse of E $(E \circ E_H = I^{(0)})$ and is well defined by Theorem 5.4. Using (6.12) and (6.10) leads to the equation

$$V = Q\left\{E_H \circ (Q^{-1}FP^{-1})\right\}P \tag{6.13}$$

for the coefficients V in (6.2).

Solving (6.13) can be done directly in the case where $u(x,t)$ does not appear on the right-hand side of (6.4), or it can be used iteratively in the case where $u(x,t)$ or its derivatives do appear on the right-hand side of (6.4). While other methods may be used effectively to solve problems of the latter sort, the above iterative process has the advantage of straightforward implementation and a relatively rapid convergence rate based on Corollary 5.5. In all of the examples, (6.13) is used to solve problems where both maps $\hat{\phi}(t) = \ell n(t)$ and $\hat{\phi}(t) = \ell n(\sinh(t))$ are applicable (Example 5) and problems for which only the former map is applicable (Examples 1 and 3). Examples 2 and 4 use the equation (6.13) to define the iterative method of solution in the case that F in (6.1) depends on the dependent variable $u(x,t)$. Example 6 depends on two spatial variables and the analogue of (6.13) (from (4.35)) is implemented to solve this problem. The final Example 7 is Burgers' sine

equation and the encouraging convergence rates there displayed strongly suggest that the iterative method from (6.13) deserves further study.

Each example is computed for the selections $M_x = 4$, 8, 16 and 32 in conjunction with the defining relations in (6.6) and (6.7). The entries in the tables labeled E_u denote the errors

$$E_u \equiv \max_{\substack{-M_x \le k \le N_x \\ -M_t \le j \le N_t}} |u(x_k, t_j) - u_A(x_k, t_j)| \tag{6.14}$$

where u is the true solution of (6.4), u_A is the Sinc-Galerkin solution in (6.3) and the spatial nodes $\{x_k\}$ and temporal nodes $\{t_j\}$ are defined in Tables 3 and 4, respectively. The entries in the tables labeled ICR are taken from the successive iterates

$$E^n \equiv \max_{\substack{-M_x \le k \le N_x \\ -M_t \le j \le N_t}} |u_{ij}^n - u_{ij}^{n-1}| \tag{6.15}$$

where the computed iterates of (6.13) are defined by

$$V^n \equiv Q\left\{E_H \circ (Q^{-1} F(V^{n-1}) P^{-1}\right\} \quad , \quad n \ge 1 \tag{6.16}$$

and

$$U^n \equiv D\left[\sqrt{\phi'}\right] V^n \quad . \tag{6.17}$$

The iteration in (6.16) is started with V^0 taken to be the zero matrix. The iterative convergence rate is defined by

$$\text{ICR} \equiv \frac{E^n}{E^{n-1}} \quad . \tag{6.18}$$

This ratio should, in theory, remain approximately constant (increasing n) and the numerical results substantiate this statement. Indeed, the theory from Corollary

5.5 says the constant is bounded by $1/4$ ($a = 0$, $b = 1$). The numerical results are much more encouraging as indicated in the closing paragraph of Chapter 5.

Example 1:

$$u_t - u_{xx} = \frac{x(1-x)(1-t^2) + 2t(1+t^2)}{(1+t^2)^2}$$

$$u(0,t) = u(1,t) = u(x,0) = 0$$

has the solution

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

Referring to Table 1 of Chapter 3, the map $\hat{\phi}(t) = \ell n(t)$ satisfies the requisite growth constraints. This is not the case for the map $\hat{\phi}(t) = \ell n(\sinh(t))$, as this demands exponential decay, rather than rational decay as $t \to \infty$. The true solution decays as $1/t$, so one cannot guarantee exponential convergence if the map $\hat{\phi}(t) = \ell n(\sinh(t))$ is used in defining the matrix B in (4.20). The parameters for the maps $\hat{\phi}(t) = \ell n(t)$ and $\phi(x) = \ell n(x/(1-x))$ are, from Table 1, given by $\alpha = \beta = \delta = \mu = 1/2$.

h	M_x	N_x	M_t	N_t	E_u
1.5708	4	4	4	4	$1.79\,e-3$
1.1107	8	8	8	8	$7.41\,e-4$
0.7854	16	16	16	16	$1.54\,e-4$
0.5554	32	32	32	32	$4.06\,e-5$

Table 5. Numerical Results for Example 1.

Example 2:

$$u_t - u_{xx} = \frac{x(1-x)(1-t-t^2-t^3) + 2t(1+t^2)}{(1+t^2)^2} - u$$

$$u(0,t) = u(1,t) = u(x,0) = 0$$

has the solution

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

As in Example 1, the map $\hat{\phi}(t) = \ell n(t)$ is applicable, so that $\alpha = \beta = \delta = \mu = 1/2$ are still suitable decay constants. The predicted asymptotic rate of convergence is $O\left(e^{-\frac{\pi}{2}\sqrt{M_x}}\right)$. This is tabulated under the heading PARC in Table 5. Iterating with $\mathcal{A}$ guarantees a convergence rate for iterative error of $E^n \leq \left(\frac{1}{4}\right)^n E^1$. To balance iterative error with the asymptotic error one selects n so that

$$\left(\frac{1}{4}\right)^n E^1 \cong e^{-\frac{\pi}{2}\sqrt{M_x}}$$

or

$$n \cong \left[\left|\frac{\pi\sqrt{M_x}}{2\,\ell n\,4}\right|\right] \quad .$$

This predicted iteration number (PIN) is tabulated in the column labeled PIN in Table 6.

M_x	PARC	PIN
4	$4.32\,e-2$	2
8	$1.18\,e-2$	3
16	$1.87\,e-3$	4
32	$1.38\,e-4$	6

Table 6. Asymptotic and Iterative Error Analysis for Example 2.

h	M_x	N_x	M_t	N_t	E_u	ICR
1.5708	4	4	4	4	$2.12\,e-3$	.1050
1.1107	8	8	8	8	$8.25\,e-4$	.1022
0.7854	16	16	16	16	$1.68\,e-4$	.0999
0.5554	32	32	32	32	$1.56\,e-5$	.1011

Table 7. Numerical Results for Example 2.

Example 3:

$$u_t - u_{xx} = \frac{[x(1-x)(1-t)+2t]\left(\cosh\frac{\pi}{2}+\cos t\right)+x(1-x)t\sin t}{\left(\cosh\frac{\pi}{2}+\cos t\right)^2}\, e^{-t}$$

$$u(0,t) = u(1,t) = u(x,0) = 0$$

has the solution

$$u(x,t) = \frac{x(1-x)t}{\cosh\frac{\pi}{2}+\cos t}\, e^{-t} \quad .$$

Due to the poles of $u(x,t)$, $t = (2\ell+1)\pi + i\pi/2$, $\ell = 0,1,2,\ldots$ in the right half-plane, there is no wedge (Figure 2) in which $u(x,t)$ is analytic in t. However, $u(\cdot,t)$ is analytic in the domain D_B in Figure 2. Also, the growth constraints of Table 1 are satisfied for the map $\hat{\phi}(t) = \ell n(\sinh(t))$ with the selection of $\alpha = \beta = \delta = \mu = 1/2$ as decay constants.

h	M_x	N_x	M_t	N_t	E_u
1.5708	4	4	4	4	$2.28\,e-4$
1.1107	8	8	8	8	$1.18\,e-4$
0.7854	16	16	16	16	$2.94\,e-5$
0.5554	32	32	32	32	$2.05\,e-6$

Table 8. Numerical Results for Example 3.

Using $\hat{\phi}(t) = \ell n(t)$ in this case violates the conditions for convergence and cannot be expected to maintain the exponential rate of convergence. This is illustrated in Table 9 with the choice $\alpha = \beta = \delta = 1/2$, N_t chosen by the process outlined in Example 5.

h	M_x	N_x	M_t	N_t	E_u
1.5708	4	4	4	1	$4.53\,e-4$
1.1107	8	8	8	2	$8.20\,e-4$
0.7854	16	16	16	3	$3.48\,e-4$
0.5554	32	32	32	4	$3.36\,e-4$

Table 9. Illustration of Misuse of Method.

Example 4:

$$u_t - u_{xx} = \frac{x(1-x)(1-2t)+2t}{\cosh\frac{\pi}{2}+\cos t}\,e^{-t} + \frac{x(1-x)t\sin t}{\left(\cosh\frac{\pi}{2}+\cos t\right)^2}\,e^{-t} - u$$

$$u(0,t) = u(1,t) = u(x,0) = 0$$

has the solution

$$u(x,t) = \frac{x(1-x)t}{\cosh\frac{\pi}{2}+\cos t}\,e^{-t} \quad .$$

As in Example 3, $\hat{\phi}(t) = \ell n(\sinh(t))$ is the appropriate temporal map, with $\alpha = \beta = \delta = \mu = 1/2$ as the decay constants. Again, the iteration defined in (6.16) is implemented. Numerical results for this case are tabulated in Table 10.

h	M_x	N_x	M_t	N_t	E_u	ICR
1.5708	4	4	4	4	$3.17\,e-4$	.0988
1.1107	8	8	8	8	$1.30\,e-4$	.1027
0.7854	16	16	16	16	$3.22\,e-5$	.1010
0.5554	32	32	32	32	$2.28\,e-6$	.1101

Table 10. Numerical Results for Example 4.

Notice that the iterative convergence rate is about one decimal place per iterate; i.e. $E^n \simeq \left(\frac{1}{10}\right) E^{n-1}$, $n \geq 1$ as opposed to $E^n \simeq \left(\frac{1}{4}\right) E^{n-1}$ as predicted by Corollary 5.5. Figure 3 is a plot of $\log(E^n)$ versus n for n iterates of this example with $M_x = 8$. The slope is approximately one which shows the constancy of this iterative convergence rate.

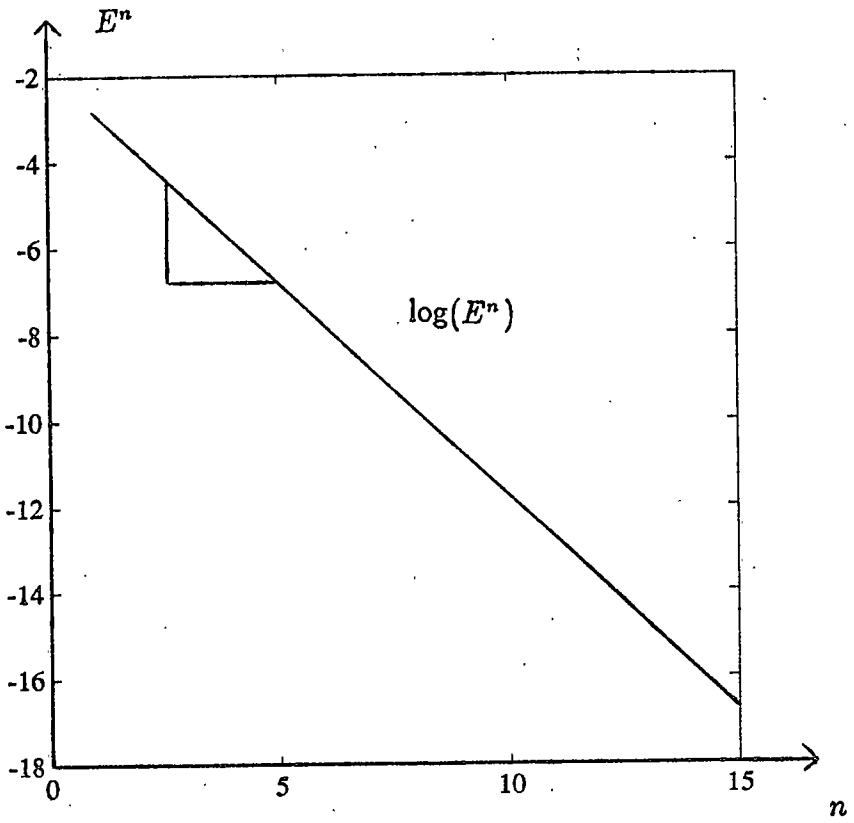


Figure 3. Iterative Convergence Rate for Example 4.

Example 5:

$$u_t - u_{xx} = [x(1-x)(1-t) + 2t]e^{-t}$$

$$u(0,t) = u(1,t) = u(x,0) = 0$$

has the solution

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

Since the solution decays exponentially as $t \to \infty$, either of the maps $\hat{\phi}(t) = \ell n(t)$ or $\hat{\phi}(t) = \ell n(\sinh(t))$ can be used in the construction of the matrix B in (4.20). In the case of $\hat{\phi}(t) = \ell n(\sinh(t))$, Table 1 dictates the choice of $\alpha = \beta = \delta = \mu = 1/2$. Similarly, $\alpha = \beta = \delta = 1/2$ are suitable decay constants for $\hat{\phi}(t) = \ell n(t)$. However, as shown in [10], for exponential decay in t one can select

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

instead of the selection in (6.7). Although the matrix system size is smaller when the map $\hat{\phi}(t) = \ell n(t)$ is used, the accuracy is not as good as when $\hat{\phi}(t) = \ell n(\sinh(t))$ is the map. A potential source for this discrepancy in accuracy is in the conditioning of the matrix B in (4.20). Numerical results for these cases are tabulated in Tables 11 and 12 for $\ell n(t)$ and $\ell n(\sinh(t))$, respectively. The additional column labeled C is the condition number $C = \|B\|_2 \|B^{-1}\|_2$. Mesh plots of the interpolated solution ($M_x = 8$) using $\ell n(\sinh(t))$ for $t \in [0,5]$ are illustrated from two different viewpoints in Figures 4 and 5.

h	M_x	N_x	M_t	N_t	E_u	C
1.5708	4	4	4	1	$9.25\,e-4$	$8.2\,e+3$
1.1107	8	8	8	2	$2.73\,e-4$	$4.3\,e+5$
0.7854	16	16	16	3	$8.42\,e-5$	$1.4\,e+8$
0.5554	32	32	32	4	$7.04\,e-6$	$4.7\,e+11$

Table 11. Numerical Results for Example 5 using $\hat{\phi}(t) = \ell n(t)$.

h	M_x	N_x	M_t	N_t	E_u	C
1.5708	4	4	4	4	$3.30\,e-4$	$1.1\,e+3$
1.1107	8	8	8	8	$5.27\,e-5$	$2.1\,e+4$
0.7854	16	16	16	16	$3.44\,e-6$	$1.4\,e+6$
0.5554	32	32	32	32	$2.07\,e-7$	$5.9\,e+8$

Table 12. Numerical Results for Example 5 using $\hat{\phi}(t) = \ell n(\sinh(t))$.

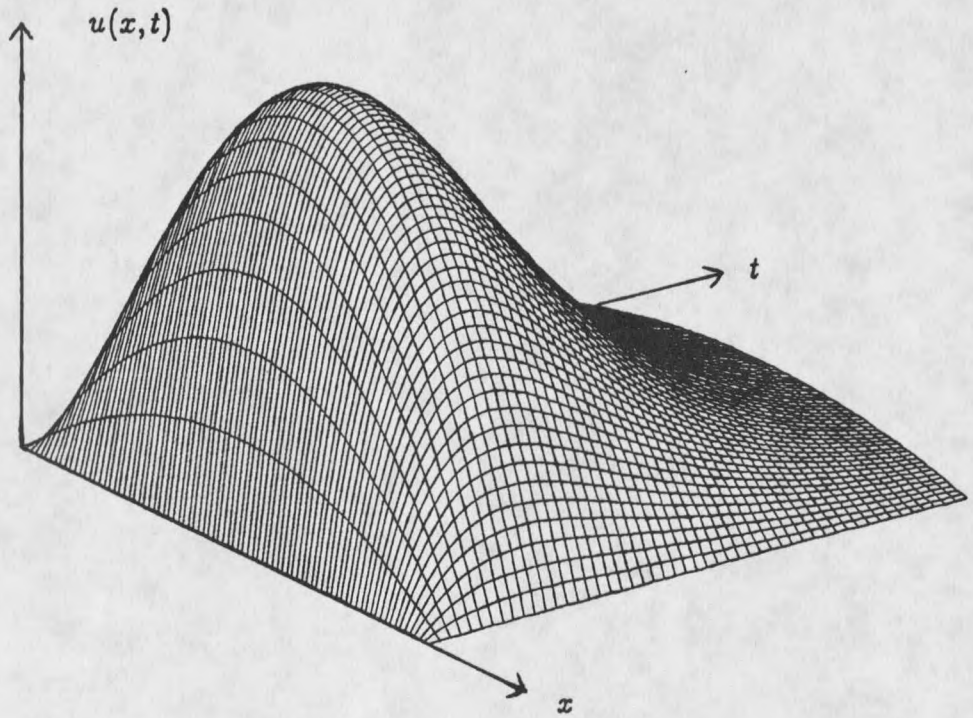


Figure 4. Mesh Plot of the Solution of Example 5 Looking Back in Time.

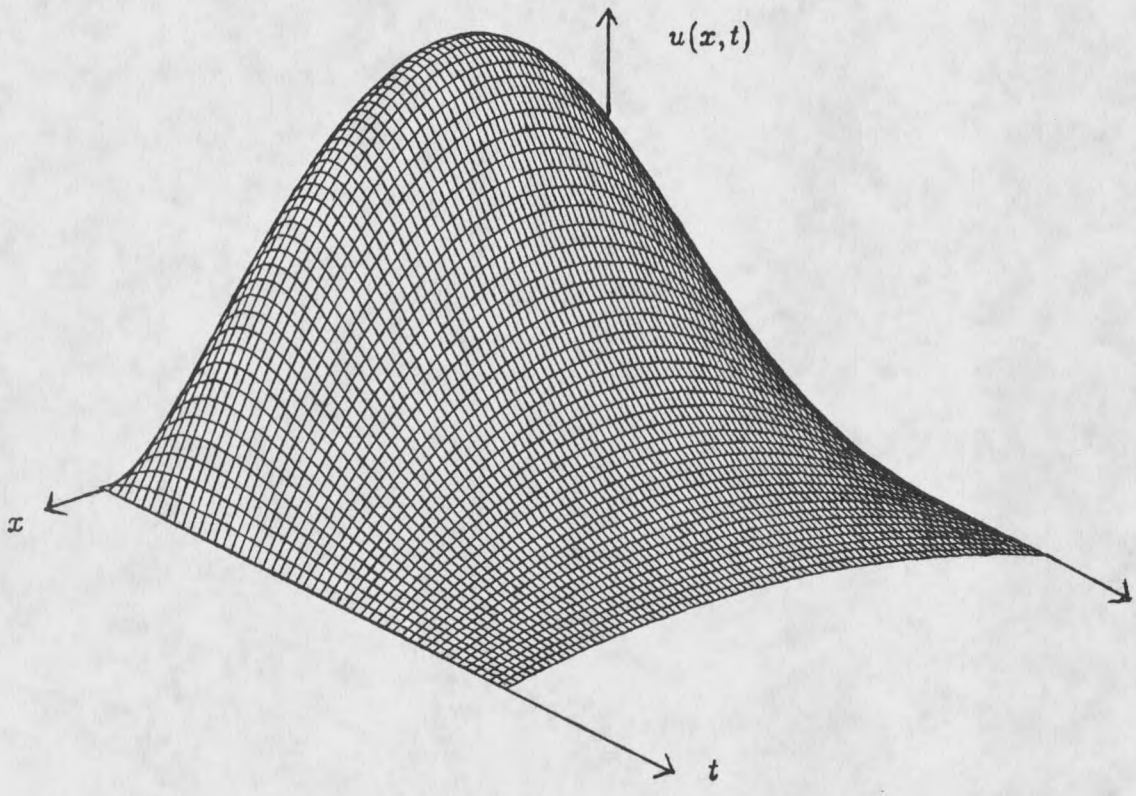


Figure 5. Mesh Plot of the Solution of Example 5 Looking Outward in Time.

Example 6:

$$u_t - \nabla^2 u = [xy(1-x)(1-y)(1-2t) + 2(x(1-x) + y(1-y))t]e^{-t}$$

$$u(0,y,t) = u(1,y,t) = u(x,0,t) = u(x,1,t) = u(x,y,0) = 0$$

has the solution

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

For this two-dimensional example, the y spatial dimension adds additional considerations for asymptotic error balancing. The procedure is the same as outlined in (6.6) and (6.7). The result for this case is that h_y, M_y and N_y are identical to their corresponding values for the parameter choices (6.6) in the x-dimension.

Referring to Table 1 it can be seen that the choice of $\hat{\phi}(t) = \ell n(\sinh(t))$ is appropriate, and that $\alpha_x = \beta_x = \alpha_y = \beta_y = \delta = \mu = 1/2$ are suitable decay constants. The final column in Table 13, marked $50 \cdot E_u$, is a more faithful representation of the error of the method. That is, $50 \cdot E_u$ closely approximates the relative error of the method.

h	M_x	N_x	M_y	N_y	M_t	N_t	$50 \cdot E_u$
1.5708	4	4	4	4	4	4	$1.17\,e-2$
1.1107	8	8	8	8	8	8	$1.27\,e-3$
0.7854	16	16	16	16	16	16	$1.46\,e-4$
0.5554	32	32	32	32	32	32	$1.86\,e-5$

Table 13. Numerical Results for Example 6.

Example 7 (Burgers' Equation): This example is used to illustrate the iterative approach in (6.16) to solve a nonlinear problem. As this is to illustrate the iterative results, rather than asymptotic convergence, no comparisons with the true solution are given. However, it should be noted that the asymptotic convergence rates still hold.

Burger's equation with sine boundary condition is

$$u_t - \epsilon u_{xx} = u u_x$$

$$u(x,0) = \sin(\pi x)$$

$$u(0,t) = u(1,t) = 0$$

$$\epsilon > 0$$

whose analytic solution is an infinite series involving Bessel functions [2]. The choice of mapping function is $\hat{\phi}(t) = \ell n(t)$ for this example, and with $\epsilon = .05$ the decay constants are $\alpha = \beta = \delta = 1/2$, $\mu = 1$, $\epsilon = .05$. Numerical results are listed in Table 14. A mesh plot of the computed solution (for $M_x = 8$) and a contour plot are given in Figures 6 and 7, respectively.

h	M_x	N_x	M_t	N_t	ICR
1.5708	4	4	4	1	.3857
1.1107	8	8	8	2	.6090
0.7854	16	16	16	3	.6256
0.5554	32	32	32	4	.6137

Table 14. Numerical Results for Example 7.

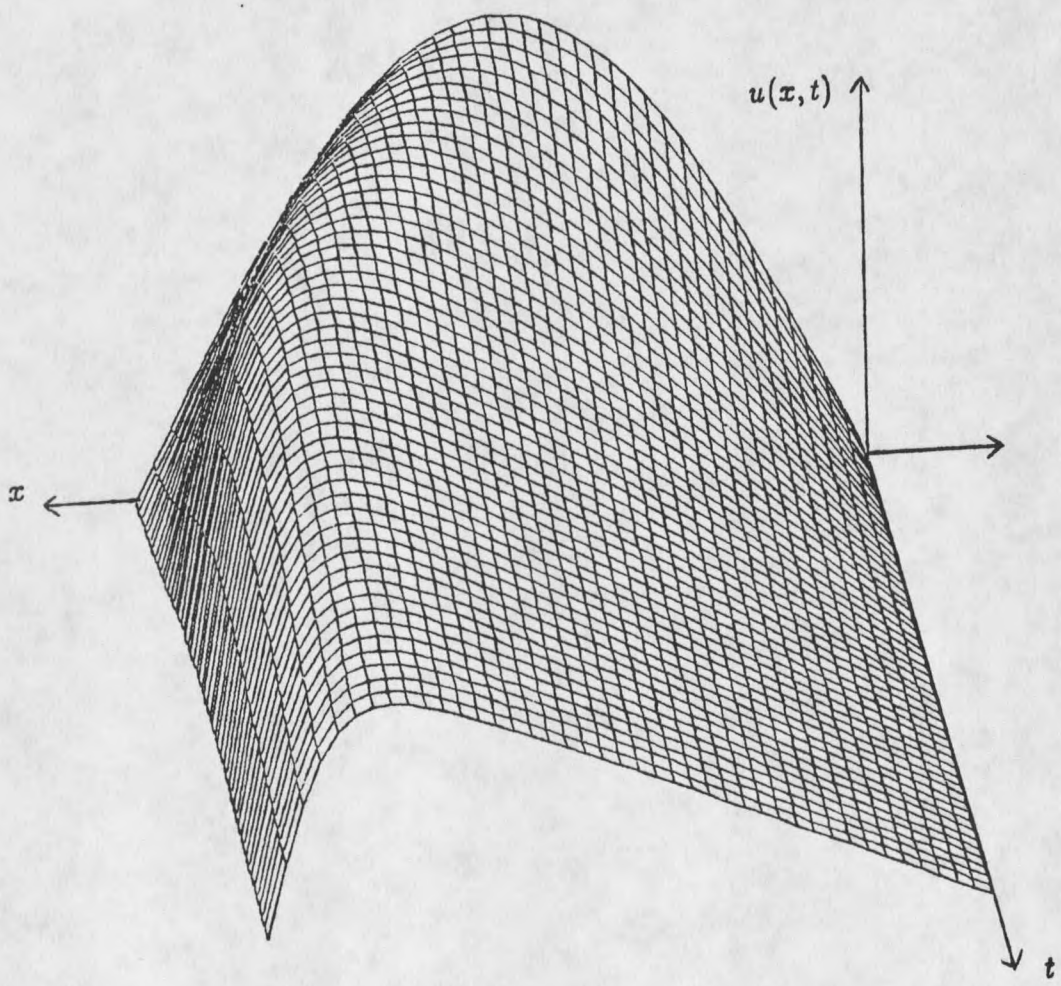


Figure 6. Mesh Plot of the Computed Solution of Example 7.

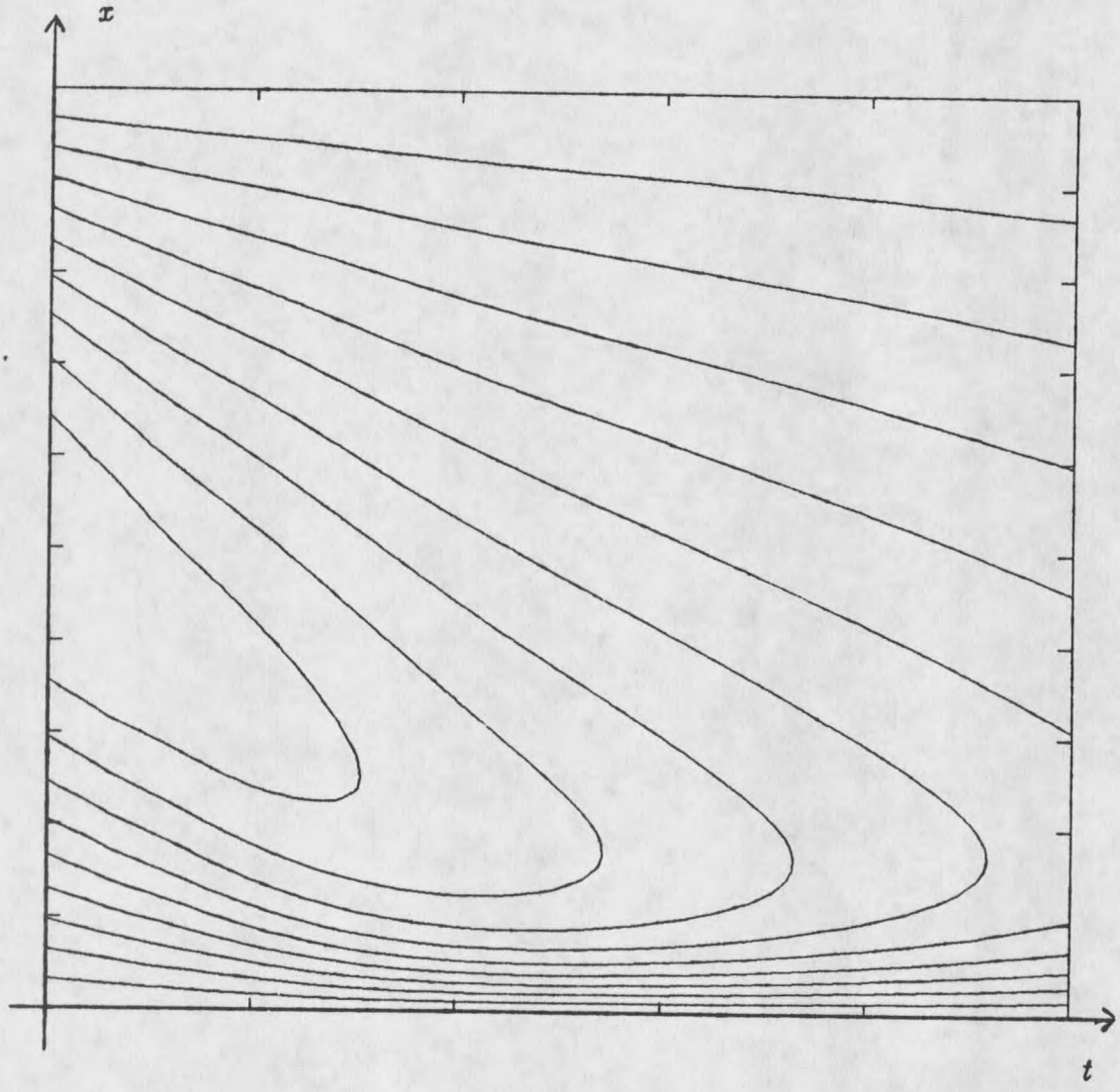


Figure 7. Contour Plot Taken from the Mesh Plot of Figure 6.

REFERENCES CITED

[1] K.L. Bowers and J. Lund, "Numerical solution of singular Poisson problems via the Sinc-Galerkin method," *SIAM J. Numer. Anal.*, 24, 36–51 (1987).

[2] J.D. Cole, "On a quasi-linear parabolic equation occuring in Aerodynamics," *Quarterly Appl. Math.*, 9, 225–236 (1951).

[3] N. Eggert, M. Jarratt and J. Lund, "Sinc function computation of the eigenvalues of Sturm-Liouville problems," *J. Comp. Phys.*, 69, 209–229 (1987).

[4] D.J. Evans and A.R. Abdullah, "Boundary value techniques for finite element solutions of the diffusion-convection equation," *Int. J. Numer. Methods Eng.*, 23, 1145–1159 (1986).

[5] 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 J. Numer. Anal.*, 3, 173–192 (1983).

[6] D. Gottlieb and S.A. Orszag, *Numerical Analysis of Spectral Methods: Theory and Applications*, SIAM, Philadelphia (1977).

[7] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*, University of California Press, Berkeley (1958).

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

[9] D.L. Lewis, J. Lund and K.L. Bowers, "The space-time Sinc-Galerkin method for parabolic problems," *Int. J. Numer. Methods Eng.*, 24, 1629–1644 (1987).

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

[11] J. Lund, K.L. Bowers and K.M. McArthur, "Symmetrization of the Sinc-Galerkin method with block techniques for elliptic equations," *IMA J. Numer. Anal.*, 9, 29–46 (1989).

[12] J. Lund, "Sinc function quadrature rules for the Fourier integral," *Math. Comp.*, 41, 103–113 (1983).

[13] J. McNamee, F. Stenger and E.L. Whitney, "Whittaker's cardinal function in retrospect," *Math. Comp.*, 25, 141–154 (1971).

[14] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York (1966).

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

[16] F. Stenger, "A Sinc-Galerkin method of solution of boundary value problems," *Math. Comp.*, 33, 85–109 (1979).

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

[18] F. Stenger, "Sinc methods of approximate solution of partial differential equations," *IMACS Proc. Adv. Comput. Methods Partial Diff. Equations*, 244-251 (1984).

[19] E.T. Whittaker, "On the functions which are represented by the expansions of the interpolation theory," *Proc. Roy. Soc. Edinburgh*, 35, 181–194 (1915).

[20] J.M. Whittaker, "On the cardinal function of interpolation theory," *Proc. Edinburgh Math. Soc.*, 2 (Series 1), 41–46 (1927).

MONTANA STATE UNIVERSITY LIBRARIES
3 1762 10071956 4