

Space-Time Spectral Methods and Adaptive Regularization for Ill-Posed and Mixed-Type Partial Differential Equations

by

Mingze Li

A thesis submitted to the Faculty of Graduate and Postdoctoral Studies of
The University of Manitoba
in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Mathematics
University of Manitoba
Winnipeg

Copyright © 2026 by Mingze Li

Abstract

This thesis investigates the stability behavior of space-time spectral discretizations for ill-posed and mixed-type partial differential equations. Using space-time collocation methods, we identify a characteristic V-shaped error profile that reflects an intrinsic interaction between ill-posedness and spectral approximation. All baseline experiments are performed without additional stabilization in order to isolate this structural instability.

To mitigate the observed error growth, we examine two complementary stabilization strategies. The first improves conditioning at the discrete operator level through ultraspherical spectral methods, while the second introduces an external regularization mechanism via the proposed technique, called smooth extension embedding methods with two sets. Although neither approach eliminates instability, both decrease the error amplification due to roundoff errors.

Another central component of this study is the systematic investigation of optimal node selection in space-time collocation methods. The V-shaped error behavior reveals a delicate trade-off between resolution and conditioning, leading to the existence of an optimal number of nodes that minimizes total error.

The spectral methods are further validated on the full potential equation with Mach numbers up to 0.7, connecting our numerical methods with practical problems.

Acknowledgements

I would like to take this opportunity to express my sincere gratitude to all those who contributed to the successful completion of this project. In particular, I would like to extend my appreciation especially to the following.

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Shaun Lui, Department of Mathematics, University of Manitoba, for his unwavering guidance, encouragement, and insightful advice throughout this project. I greatly appreciate his enthusiasm, profound knowledge, and dedication, which have been invaluable in helping me develop a deep understanding of this research. I am especially thankful for the time and care he devoted to supervising and mentoring me. Without his support, this work would not have been possible, and I am sincerely grateful for his guidance.

I am also deeply grateful to the members of my advisory committee, Prof. Mikael Slevinsky and Prof. Ruppa Thulasiram, for their insightful feedback and guidance, which greatly enhanced the quality of this research.

I also wish to thank the financial support provided by UMGF, Department of Mathematics, and the Faculty of Graduate Studies at the University of Manitoba as well as my supervisor's grant throughout my years of study.

I further extend my appreciation to my family and friends for their endless support even after I moved across the world to pursue my studies. Last but not least, my deepest appreciation goes to my beloved wife, who has been a true pillar of strength, providing constant support and encouragement throughout my academic journey.

List of Figures

2.1	Convergence for the heat equation.	5
2.2	Convergence of different schemes for the model problem as a function of n for the exact solution $u(x) = (x + 1)(e^{2x} - e^2x - 2e^2)$	8
3.1	Comparison of convergence behavior between finite difference and spectral methods. In both cases, finite difference solutions diverge due to high-frequency amplification, while the spectral discretization exhibits the characteristic V-shaped error profile.	11
3.2	Computational domain with vertical boundaries Γ_4 , Γ_6 and upper boundary Γ_5 . The lower boundary consists of two horizontal segments Γ_1 and Γ_3 , and a curved part Γ_2	17
3.3	Numerical results obtained using space-time collocation methods. . .	18
4.1	Numerical results obtained using US method for well-posed PDEs in contrast to space-time collocation methods.	29
4.2	Numerical results obtained using US method for ill-posed PDEs in contrast to space-time collocation methods.	30
4.3	Convergence for the nonlinear wave equation using US methods combined with NGMRES(5), with solution magnitude parameter $a = 3.8$	32
5.1	Numerical results obtained using the SEEM-TWO. For each problem, the performance of SEEM-TWO is compared with that of space-time collocation methods.	42
5.2	The Vandeven filter function with different orders p	45
5.3	Comparison of numerical errors obtained by space-time collocation methods, US methods, and SEEM-TWO under the same noise perturbation.	48
7.1	Numerical solutions and errors for the subsonic case $M_\infty = 0.1$ under different boundary curvatures.	59
7.2	Numerical solutions and errors for the nonlinearity-dominant case $M_\infty = 0.6$ under different boundary curvatures.	60
7.3	Streamlines and Mach number contours for $M_\infty = 0.1$	61
7.4	Streamlines and Mach number contours for $M_\infty = 0.6$	62
7.5	Streamlines and Mach number contours for $M_\infty = 0.7$	62

List of Tables

6.1	Optimal values of n obtained by Algorithm 1 for different test problems.	53
6.2	Comparison of Algorithm 1 and Algorithm 2 for the oscillatory exact solution $u(x, t) = e^{xt} \sin(5(x^2 - 1))$ on the search interval $[4, 40]$. Entries marked by $*$ indicate that Algorithm 1 terminated prematurely and failed to detect the optimal n_* .	53
6.3	Comparison of Algorithm 1 and Algorithm 2 for the oscillatory exact solution $u(x, t) = e^{xt} \sin(5(x^2 - 1))$ on the search interval $[4, 40]$. Here $\#$ Solves denotes the number of linear systems solved. Cost represents the estimated total floating-point operations. Entries marked by $*$ indicate that Algorithm 1 terminated prematurely and failed to detect the correct n_* .	55

Contents

Abstract	i
Acknowledgements	ii
List of Figures	iii
List of Tables	iv
1 Introduction	1
2 Space-time and Ultraspherical Spectral Methods	3
2.1 Space-time collocation methods	3
2.2 Ultraspherical spectral methods	5
3 Ill-posed and Mixed-type Equations: Error Behavior of Space-time Collocation Methods	9
3.1 Backward heat equation	10
3.2 Initial-value problem for the Poisson equation	10
3.3 Tricomi equation	11
3.4 Keldysh equation	12
3.5 General mixed-type equation	13
3.6 Degenerate parabolic equation	14
3.7 Full potential equation	15
3.8 Numerical experiments	17
3.9 Implications for stabilization	18
4 Operator-level Stabilization: Ultraspherical Spectral Methods	20
4.1 Construction of ultraspherical scheme	20
4.1.1 Assembly of the differentiation matrix	21
4.1.2 Construction of the truncation matrix	23
4.1.3 Imposing boundary conditions	24
4.1.4 Final discrete system and recovery of the physical solution . .	27
4.1.5 Summary of the ultraspherical formulation	27

4.2	Numerical experiments	27
4.3	Ultraspherical methods with stabilization via Anderson acceleration and nonlinear-GMRES	31
5	Smooth Extension Embedding Method with Two Sets of Nodes	33
5.1	Description of the method	33
5.2	The smoother S	36
5.2.1	The artificial viscosity's perspective	36
5.2.2	The energy functional's perspective	37
5.3	SEEM-TWO and Tikhonov regularization	38
5.4	Numerical experiments	40
5.5	Remarks on other regularization strategies	43
5.5.1	Artificial viscosity	43
5.5.2	Filtering	44
5.6	Noise amplification test for regularization effects	46
6	Balancing Accuracy and Stability: Determination of the Optimal Node Number	49
6.1	Review: the behavior of spectral errors	49
6.2	Limitations of error estimation without exact solutions	49
6.3	Difference-based error indicators	50
6.4	An adaptive stopping criterion for node selection	50
6.5	Numerical experiments	52
7	Full Potential Equation	56
7.1	Spectral discretization	56
7.1.1	Spectral discretization in the interior	56
7.1.2	Irregular physical domain	57
7.1.3	Nonlinear iteration: Picard and Newton schemes	57
7.1.4	Mach number continuation strategy for larger M_∞	58
7.2	Numerical experiments	58
7.2.1	Subsonic regime: $M_\infty = 0.1$	59
7.2.2	Nonlinear-dominant subsonic regime: $M_\infty = 0.6$	60

7.2.3	Streamlines and local Mach contours	61
8	Discussion and Conclusion	63

1 Introduction

Spectral methods are well known for their high-order accuracy and excellent efficiency in solving partial differential equations (PDEs) with smooth solutions. By representing the solution in terms of global basis functions, such as Legendre, Chebyshev or Fourier polynomials, spectral methods can achieve exponential convergence with relatively few degrees of freedom. Excellent references on spectral methods include [23, 33, 36]. For time-dependent problems, the most common approach is to use a low-order finite difference approximation of the time derivative and spectral approximation of the spatial derivatives. This is not ideal since the time discretization error overwhelms the spatial discretization error. In contrast, space-time collocation methods, which discretize the spatial and temporal directions simultaneously, provide spectral accuracy and avoid the low-order convergence associated with time-marching schemes.

It is well known that numerical schemes suffer from severe numerical instabilities when applied to ill-posed and mixed-type problems. Typical examples include backward heat equations, initial-value problem for the Poisson equation (IVP), and equations of Tricomi and Keldysh type. The error often exhibits a characteristic V- or U-shaped behavior as a function of the number of nodes, n ; the error initially decreases as n is increased, but eventually rises with n getting larger due to the amplification of high-frequency components inherent to the ill-posedness of the problems. This phenomenon indicates a fundamental limitation of space-time spectral schemes for such equations and motivates the development of stabilization strategies.

The first stabilization strategy explored in this thesis focuses on improving the conditioning of the differential operators themselves. Ultraspherical spectral methods (US) are known to yield well-conditioned differentiation matrices by exploiting recurrence relations between Chebyshev polynomial bases. As a result, the growth of errors can be reduced, compared to classical spectral collocation methods. In the context of ill-posed and mixed-type problems, ultraspherical discretizations are shown to delay somewhat the onset of the error growth.

The second stabilization strategy is based on regularization. We propose a regularization technique termed smooth extension embedding method with two sets of nodes (SEEM-TWO), which introduces a smoothing mechanism into the space-time collocation methods. Rather than modifying differential operators directly, SEEM-TWO suppresses high-frequency modes through a regularized formulation. Numerical results demonstrate that this approach can slightly reduce the growth of numerical errors.

Although both US methods and SEEM-TWO can improve the numerical stability of space-time collocation methods, neither approach can completely eliminate the intrinsic ill-posedness of the underlying problems. In particular, the V/U-shaped error behavior persists, albeit in a weakened form. This observation leads to a key insight of this work: in ill-posed spectral discretizations, the number of spectral nodes itself acts as an implicit regularization parameter. Consequently, it is neither necessary nor desirable to increase the number of spectral nodes indefinitely. Instead, an optimal number, n_* , exists, at which the total numerical error is minimized. Based on this perspective, we propose practical strategies for identifying an optimal node number n_* that minimizes the spectral error.

Finally, we apply the proposed framework to the full potential equation. This non-linear equation changes type as the Mach number increases and becomes mixed-type at sufficiently high Mach numbers, e.g., $M_\infty = 0.8$. By full spectral discretizations, we are able to compute stable and accurate numerical solutions up to Mach numbers as high as $M_\infty = 0.7$. These results demonstrate the effectiveness of the proposed framework for challenging, physically relevant problems, while also illustrating its limitations at higher Mach numbers.

The remainder of this thesis is organized as follows. Chapter 2 reviews space-time spectral discretizations and presents benchmark results for well-posed problems. Chapter 3 investigates the V/U-shaped error behavior in various ill-posed and mixed-type equations. Chapters 4 and 5 introduce ultraspherical spectral methods and the SEEM-TWO regularization approach, respectively. The former enhances stability intrinsically at the operator level, while the latter acts as an external regularization strategy applied to the discretized system. Chapter 6 discusses the selection of an optimal number of spectral nodes. Finally, Chapter 7 applies the proposed methods to the full potential equation and presents numerical results.

2 Space-time and Ultraspherical Spectral Methods

Space-time collocation methods have been successfully used for solving time-dependent PDEs for several decades. Unlike finite difference or finite element methods, whose basis functions are typically local, spectral methods employ global smooth basis functions, leading to dense differentiation matrices. Their main advantage is that, for analytic solutions, the convergence rate is exponential with respect to the number of basis functions. This exponential convergence stems from the rapid decay of Fourier or polynomial expansion coefficients for smooth solutions. However, in many traditional approaches to time-dependent problems, spectral approximation is applied only in space, while low-order finite difference schemes are used for time discretization. This mismatch is undesirable, since the temporal discretization error may dominate the spatial error and destroy the overall spectral accuracy. To overcome this drawback, space-time collocation methods were developed, in which spectral discretizations are employed in both spatial and temporal directions. As a result, the numerical solution converges exponentially in both time and space, provided that the exact solution is analytic. Numerical experiments and condition number estimates for space-time collocation methods applied to a variety of linear and nonlinear PDEs can be found in [20, 21, 24–26].

In this chapter, we briefly review the formulation of space-time collocation methods based on Legendre and Chebyshev polynomial expansions. This chapter serves two purposes. First, it introduces the notation and basic discretization framework used throughout the thesis. Second, we use the heat equation as a representative example to demonstrate the spectral convergence properties of space-time collocation methods for well-posed time-dependent problems.

2.1 Space-time collocation methods

We now describe how the method works for the heat equation, which reads

$$\begin{cases} u_t - u_{xx} = f, & \text{in } \Omega = (-1, 1)^2, \\ u(\pm 1, t) = 0, \\ u(x, -1) = u_0(x). \end{cases}$$

Let $n \geq 2$ and denote by $u_h \in \mathbb{R}^{(n+1)^2}$ the vector of nodal values of the approximate solution evaluated at the Legendre–Gauss–Lobatto nodes in both space and time. We

seek a numerical solution $u_h(x, t)$ in the following form:

$$u_h(x, t) = \sum_{j=0}^n \sum_{i=0}^n u_{ij} \ell_i(x) \ell_j(t),$$

which is a linear combination of basis functions in both x and t . Here, ℓ_i is the Lagrange interpolation polynomial corresponding to the node x_i or t_i . The discrete matrix equation reads

$$(I_{n+1} \otimes D - D^2 \otimes I_{n+1})u_h = f_h,$$

where the Kronecker product structure reflects the tensor-product discretization in space and time, f_h is the long vector of the given right-hand side, f , evaluated at the collocation nodes of dimension $(n+1)^2$, D is the Legendre pseudospectral derivative matrix and I_{n+1} the identity matrix. Because of the given initial and boundary conditions, it suffices to solve for the solution, which is denoted by $\hat{u}_h$, deleting the components corresponding to those nodes associated with known initial and boundary values. Similarly, we do the same deletion to f_h . The final discrete equations are

$$(I_{n-1} \otimes [D] - [[D^2]] \otimes I_n)\hat{u}_h = \hat{f}_h - (u_{0h} \otimes d_h),$$

where, for any $n \times n$ matrix M , $[M]$ is the $(n-1) \times (n-1)$ matrix obtained from M by removing the last column and row, while $[[M]]$ denotes the $(n-2) \times (n-2)$ matrix obtained from M by deleting the first and last columns and rows. Also, u_{0h} corresponds to the initial discrete vector evaluated at the interior spatial collocation nodes and d_h denotes the first n entries of the last column of D : $d_h = D(0 : n-1, n)$. Note that $u_{0h} \otimes d_h$ denotes the contribution of the initial condition.

We implemented a basic space-time Legendre collocation method solving the heat equation in MATLAB. In our numerical experiments, we choose the exact solution to time-dependent PDEs to be $u(x, t) = (x^2 - 1) \ln(x + t + 5)$. The exponential convergence of space-time collocation methods is clearly illustrated in Figure 2.1, in contrast to the second-order convergence of a finite difference method. The error in this thesis is measured in discrete ℓ^∞ -norm at the collocation nodes.

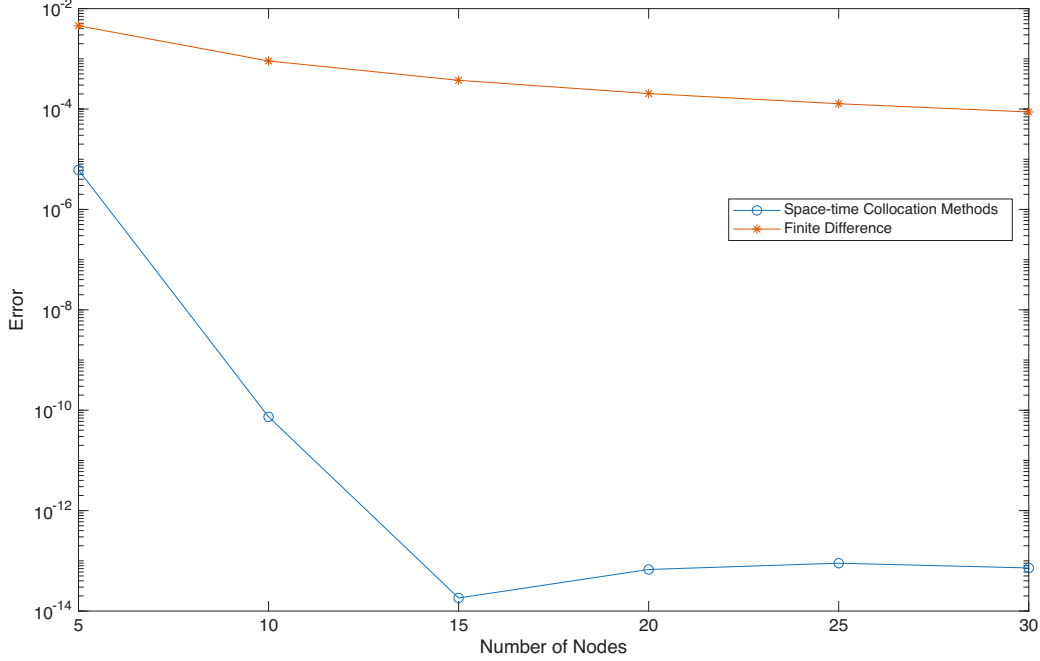


Fig. 2.1: Convergence for the heat equation.

2.2 Ultraspherical spectral methods

Although spectral collocation methods are straightforward to implement, they typically produce dense and severely ill-conditioned matrices. In contrast, ultraspherical spectral (US) methods represent each term of the differential equation using expansions in ultraspherical polynomials. This formulation results in matrices that are significantly better conditioned and exhibit an almost banded structure. We first give a description of the method for the following one-dimensional model problem. Time-dependent PDEs solved by US schemes will be discussed in Chapter 4. Also, more details can be found in [3, 8, 9, 28, 31, 35].

Consider

$$-u'' + u = f, \quad u(-1) = 0 = u'(1). \quad (1)$$

For each non-negative integer ℓ , the ultraspherical polynomials with degree n are denoted by $C_n^{(\ell)}$, $n \geq 0$. They satisfy the recurrence relation

$$C_n^{(\ell)}(x) = \frac{1}{n} [2x(n + \ell - 1)C_{n-1}^{(\ell)}(x) - (n + 2\ell - 2)C_{n-2}^{(\ell)}(x)], \quad n \geq 2,$$

with $C_0^{(\ell)} = 1$ and $C_1^{(\ell)} = 2\ell x$. The expression for $C_1^{(\ell)}$ holds only for $\ell > 0$. When $\ell = 0$, we instead have $C_1^{(0)} = T_1(x) = x$, where T_k denotes the k -th order Chebyshev

polynomial. The following properties are related to our discussion:

$$C_n^{(1)} = \begin{cases} T_0, & n = 0; \\ 2T_1, & n = 1; \\ C_{n-2}^{(1)} + 2T_n, & n \geq 2; \end{cases} \quad C_n^{(2)} = \begin{cases} C_0^{(1)}, & n = 0; \\ 2C_1^{(1)}, & n = 1; \\ C_{n-2}^{(2)} + (n+1)C_n^{(1)}, & n \geq 2; \end{cases}$$

and

$$\frac{dC_n^{(0)}}{dx} = \begin{cases} 0, & n = 0; \\ nC_{n-1}^{(1)}, & n \geq 1; \end{cases} \quad \frac{dC_n^{(1)}}{dx} = \begin{cases} 0, & n = 0; \\ 2C_{n-1}^{(2)}, & n \geq 1. \end{cases}$$

Assume that

$$u(x) = \sum_{k=0}^n \hat{u}_k T_k(x)$$

is the spectral approximate solution of (1) for some integer $n \geq 2$. The right-hand side can be expanded in a similar way:

$$f(x) = \sum_{k=0}^n \hat{f}_k T_k(x).$$

Let $u_h = [\hat{u}_0; \dots; \hat{u}_n]$ and $f_h = [\hat{f}_0; \dots; \hat{f}_n]$. By the above properties,

$$u''(x) = 2 \sum_{k=2}^n \hat{u}_k k C_{k-2}^{(2)}(x),$$

which can be written in terms of the spectral second derivative matrix in the basis $\{C_k^{(2)}\}$

$$D_2 = 2 \begin{bmatrix} 0 & 0 & 2 & & & \\ & 0 & 0 & 3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 0 & 0 & n \\ & & & & 0 & 0 \\ & & & & & 0 \end{bmatrix} \in \mathbb{R}^{(n+1) \times (n+1)}.$$

This matrix only has one non-zero diagonal. Note that the left-hand side of the problem has been converted into a linear combination with basis functions $\{C_k^{(2)}\}$ while the right-hand side is still based on $C_k^{(0)}$ polynomials. This incompatibility can be addressed by two conversion matrices that transform coefficients from $C^{(0)}$ to $C^{(1)}$

and $C^{(1)}$ to $C^{(2)}$ given by S_0 and S_1 , respectively, where

$$S_0 = \begin{bmatrix} 1 & 0 & -1/2 & & & \\ & 1/2 & 0 & -1/2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1/2 & 0 & -1/2 \\ & & & & 1/2 & 0 \\ & & & & & 1/2 \end{bmatrix} \in \mathbb{R}^{(n+1) \times (n+1)},$$

and

$$S_1 = \begin{bmatrix} 1 & 0 & -1/3 & & & \\ & 1/2 & 0 & -1/4 & & \\ & & 1/3 & 0 & -1/5 & \\ & & & \ddots & \ddots & \ddots \\ & & & & 1/(n-1) & 0 & -1/(n+1) \\ & & & & & 1/n & 0 \\ & & & & & & 1/(n+1) \end{bmatrix} \in \mathbb{R}^{(n+1) \times (n+1)}.$$

In this way, $S_1 S_0 f_h$ converts the coefficients of f expanded in $\{C_k^{(0)}\}$ to those expanded in $\{C_k^{(2)}\}$.

The last step is to impose the boundary conditions by τ -method:

$$u(-1) = \sum_{k=0}^n \hat{u}_k T_k(-1) = \sum_{k=0}^n \hat{u}_k (-1)^k = 0,$$

and

$$u'(1) = \sum_{k=0}^n \hat{u}_k T'_k(1) = \sum_{k=0}^n \hat{u}_k k^2 = 0.$$

Let $S = S_1 S_0$ with all entries in the last two rows replaced by 0. The discrete system reads $A_h u_h = \hat{f}_h$, where

$$A_h = \begin{bmatrix} 0 & 0 & -4 & & & \\ & 0 & 0 & -6 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 0 & 0 & -2n \\ 1 & -1 & 1 & -1 & \cdots & (-1)^n \\ 0 & 1 & 4 & 9 & \cdots & n^2 \end{bmatrix} + S$$

and $\hat{f}_h = [\tilde{f}_h; 0; 0]$, where $\tilde{f}_h$ is f_h removing the last 2 entries.

It can be observed from the following figure that the US method is more robust than the classical spectral methods. In contrast to the spectral collocation case, the error remains near machine precision for large values of n .

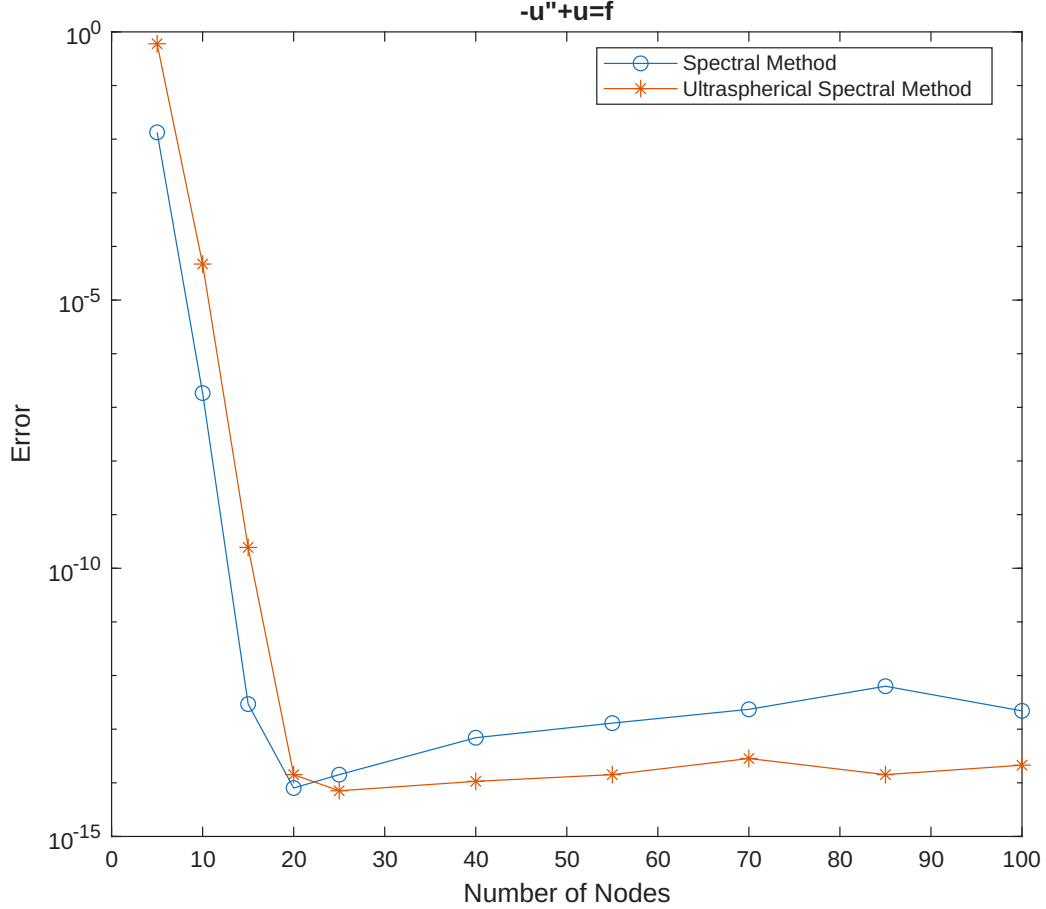


Fig. 2.2: Convergence of different schemes for the model problem as a function of n for the exact solution $u(x) = (x + 1)(e^{2x} - e^2x - 2e^2)$.

In summary, the US method reformulates differential operators in ultraspherical polynomial bases, leading to sparse and well-conditioned linear systems. This structure makes the method particularly attractive for high-order and time-dependent problems as we will see in this thesis.

3 Ill-posed and Mixed-type Equations: Error Behavior of Space-time Collocation Methods

In this chapter, we investigate the numerical behavior of space-time collocation methods when applied to ill-posed and mixed-type PDEs. Ill-posed problems, such as the backward heat equation and the initial-value problem for the Poisson equation, violate one or more of the Hadamard conditions for well-posedness. In particular, small perturbations in the data may lead to great instabilities in the solution, making stable numerical approximation extremely challenging. Mixed-type equations, including the Tricomi and Keldysh equations and full potential equation, exhibit a change of type within the computational domain, combining elliptic and hyperbolic behavior. This change of type introduces additional instability and complicates both analysis and numerical discretization. References on these equations and corresponding properties include [4, 7, 18, 27, 29, 30, 32, 40] and the references therein. Although spectral methods are well known for their high accuracy and exponential convergence for smooth and well-posed problems, their performance can deteriorate significantly once ill-posedness or mixed-type characteristics are present. Understanding this deterioration is a necessary first step before a stabilization strategy is proposed.

The focus of this chapter is not on proposing new numerical techniques, but on systematically demonstrating a characteristic error behavior that arises when space-time collocation methods are employed for such problems. Through a collection of representative examples, we show that the numerical error typically exhibits a non-monotone dependence on the number of nodes, displaying a pronounced V- or U-shaped curve. Specifically, the error initially decreases as the number of nodes, n , increases, reflecting the high-order accuracy of spectral approximations, but beyond a certain point, n_* , the numerical error begins to grow rapidly as n continues increasing. This phenomenon persists even when the exact solution is smooth and when no explicit noise is introduced in the data.

The equations considered in this chapter include classical ill-posed problems such as the backward heat equation and the initial-value problem for the Poisson equation, as well as mixed-type equations including Tricomi, Keldysh and general mixed-type equations. Also, we consider a class of degenerate parabolic problems. In the last part of this chapter, the background of the full potential equation is introduced.

3.1 Backward heat equation

Consider

$$u_t + u_{xx} = f, \text{ in } (-1, 1)^2,$$

with initial condition $u(x, -1) = u_0(x)$ and homogeneous Dirichlet boundary conditions. If $u_0(x) = \sin k\pi x$ for some positive integer k and $f \equiv 0$, then the solution is $u(x, t) = e^{k^2\pi^2(t+1)} \sin k\pi x$. This illustrates that high frequency terms grow exponentially in time and is the root cause of the ill-posedness of the PDE. Classical time-stepping schemes cannot solve this problem even if the solution is smooth and bounded in time because roundoff errors will introduce high frequency terms which cause the numerical scheme to blow up. In fact, standard finite difference discretizations suffer from the same instability mechanism. Even with small mesh sizes, numerical roundoff errors inevitably introduce high-frequency components, which are then exponentially amplified by the backward heat operator. As a result, the discrete solution typically diverges rapidly, reflecting the intrinsic ill-posedness of the problem rather than a deficiency of a particular numerical scheme. See Figure 3.1a for the divergence of the finite difference error compared to spectral methods. The spectral scheme is

$$(I_{n+1} \otimes D + D^2 \otimes I_{n+1})u_h = f_h,$$

where f_h is f evaluated at collocation points. Deleting the known boundary and initial values, the final scheme reads

$$(I_{n-1} \otimes [D] + [[D^2]] \otimes I_n)\hat{u}_h = \hat{f}_h - (u_{0h} \otimes d_h),$$

where u_{0h} is u_0 evaluated at the interior collocation points. As shown in Figure 3.3a, the spectral error exhibits a V-shape.

3.2 Initial-value problem for the Poisson equation

Consider

$$u_{tt} + u_{xx} = f, \text{ in } (-1, 1)^2,$$

with initial conditions $u(x, -1) = u_0(x)$ and $u_t(x, -1) = u_1(x)$ and homogeneous Dirichlet boundary conditions. Following an example due to Hadamard, suppose k is a positive integer. If $u_0(x) = 0$, $u_1(x) = k^{-1} \sin k\pi x$ and $f \equiv 0$, then the solution is $u(x, t) = k^{-2} \sinh(k\pi(t+1)) \sin k\pi x$. The ill-posedness is apparent as $t \rightarrow \infty$. Similarly, classical finite difference schemes for this initial-value formulation exhibit severe instability. Small perturbations in the initial data or discretization errors

can trigger rapidly growing high-frequency modes, causing the numerical solution to diverge even when the exact solution remains smooth. Figure 3.1b shows this divergent behavior of the error obtained using a finite difference scheme compared to spectral methods. The spectral scheme is

$$(I_{n+1} \otimes D^2 + D^2 \otimes I_{n+1})u_h = f_h.$$

Deleting the known boundary and initial values, the final scheme reads

$$(I_n \otimes [D]^2 + [[D^2]] \otimes I_n)\hat{u}_h = \hat{f}_h - (u_{0h} \otimes [D]d_h) - (u_{1h} \otimes d_h),$$

where u_{1h} is u_1 evaluated at the interior collocation points. The blow-up of solutions of space-time collocation methods are much slower than those of finite difference methods because the former employ far fewer unknowns. The V-shaped behavior of the spectral error can be observed in Figure 3.3b.

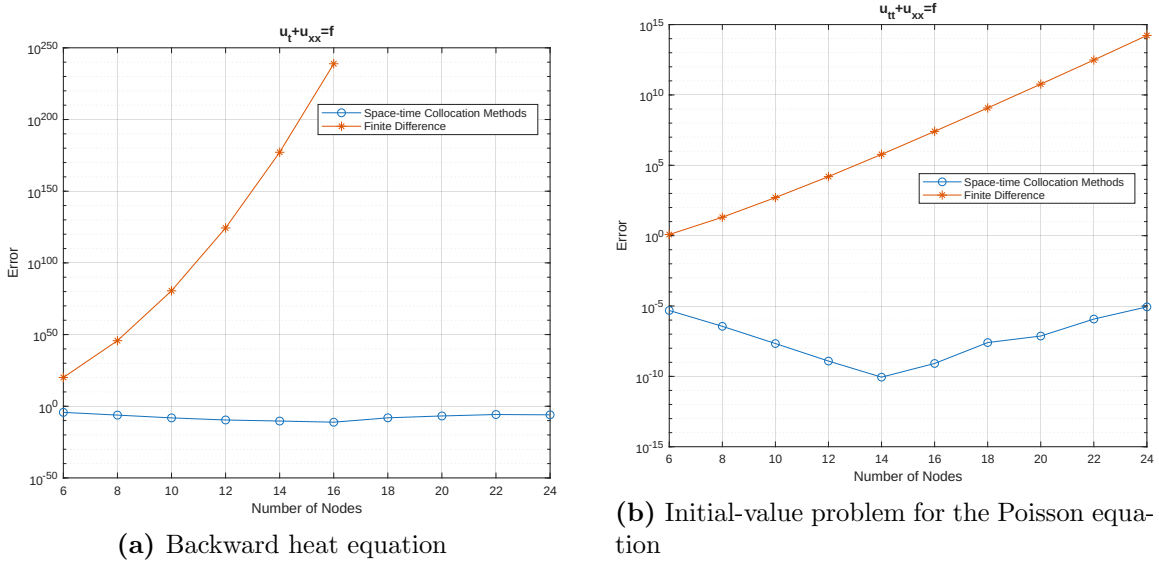


Fig. 3.1: Comparison of convergence behavior between finite difference and spectral methods. In both cases, finite difference solutions diverge due to high-frequency amplification, while the spectral discretization exhibits the characteristic V-shaped error profile.

3.3 Tricomi equation

The Tricomi equation is

$$u_{tt} + tu_{xx} = f, \text{ in } (-1, 1)^2,$$

with initial conditions $u(x, -1) = u_0(x)$ and $u_t(x, -1) = u_1(x)$ and homogeneous Dirichlet boundary conditions. The equation is of mixed-type since the sign of the principal part changes across $t = 0$: for $t < 0$, the terms u_{tt} and tu_{xx} have opposite signs and the equation is hyperbolic; for $t > 0$, they have the same sign and the equation becomes elliptic; the line $t = 0$ corresponds to a degeneracy of the principal symbol.

Let

$$B = I_{n+1} \otimes \text{diag}(t_h),$$

where t_h denote the vector in the temporal direction at the collocation nodes, the spectral scheme is

$$(I_{n+1} \otimes D^2 + B(D^2 \otimes I_{n+1}))u_h = f_h.$$

Deleting the known boundary and initial values, the final scheme is

$$(I_{n-1} \otimes [D]^2 + \hat{B}([D^2] \otimes I_n))\hat{u}_h = \hat{f}_h - (u_{0h} \otimes [D]d_h) - (u_{1h} \otimes d_h).$$

Here,

$$\hat{B} = I_{n-1} \otimes \text{diag}(t_{h0}),$$

where t_{h0} is obtained by removing the last entry of t_h . The V-shaped error behavior can be observed in Figure 3.3c.

3.4 Keldysh equation

The Keldysh equation reads

$$tu_{tt} + u_{xx} = f, \text{ in } (-1, 1)^2,$$

with initial conditions $u(x, -1) = u_0(x)$ and $u_t(x, -1) = u_1(x)$ and homogeneous Dirichlet boundary conditions. This equation is also mixed-type. In contrast to the Tricomi equation, the sign change now affects the coefficient of u_{tt} . For $t < 0$, the principal terms have opposite signs and the equation is hyperbolic; for $t > 0$, they share the same sign and the equation becomes elliptic. Again, the line $t = 0$ corresponds to a degeneracy of the principal symbol.

Under the same notation as in Chapter 3.3, the spectral scheme is

$$(B(I_{n+1} \otimes D^2) + D^2 \otimes I_{n+1})u_h = f_h.$$

Deleting the known boundary and initial values, the final scheme becomes

$$(\hat{B}(I_{n-1} \otimes [D]^2) + [[D^2]] \otimes I_n) \hat{u}_h = \hat{f}_h - (u_{0h} \otimes (t_{0h} \circ ([D]d_h)) - (u_{1h} \otimes (t_{0h} \circ d_h)),$$

where $\circ$ denotes the Hadamard (element-wise) product, i.e., $(u \circ v)_i = u_i v_i$. Figure 3.3d illustrates the V-shape behavior of the spectral error.

3.5 General mixed-type equation

Consider the following mixed-type equation

$$u_{tt} + (t - x^2 + \frac{1}{2}) u_{xx} = f, \text{ in } (-1, 1)^2,$$

with initial conditions $u(x, -1) = u_0(x)$ and $u_t(x, -1) = u_1(x)$ and homogeneous Dirichlet boundary conditions. The equation is hyperbolic in the region

$$t < x^2 - \frac{1}{2},$$

where the principal terms have opposite signs, and elliptic in the region

$$t > x^2 - \frac{1}{2},$$

where they share the same sign. The degeneracy occurs along the parabolic curve

$$t = x^2 - \frac{1}{2},$$

which separates the hyperbolic and elliptic regions.

Let

$$V = I_{n+1} \otimes \text{diag}(t_h) - \text{diag}(x_h^2) \otimes I_{n+1} + \frac{1}{2} I_{(n+1)^2},$$

where x_h denotes the vector in the spatial direction at the collocation nodes and other notation is the same as in Chapter 3.3, The spectral scheme is

$$(I_{n+1} \otimes D^2 + V(D^2 \otimes I_{n+1})) u_h = f_h.$$

Deleting the known boundary and initial values, the final scheme reads

$$(I_{n-1} \otimes [D]^2 + \hat{V}([D^2] \otimes I_n)) \hat{u}_h = \hat{f}_h - (u_{0h} \otimes [D]d_h) - (u_{1h} \otimes d_h).$$

Here,

$$\hat{V} = I_{n-1} \otimes \text{diag}(t_{h0}) - \text{diag}(x_{h0}^2) \otimes I_n + \frac{1}{2} I_{n(n-1)},$$

where x_{h0} is obtained by removing the first and last entries of x_h . Figure 3.3e illustrates the V-shape behavior of the spectral error.

3.6 Degenerate parabolic equation

We consider the following degenerate parabolic equation with the degeneracy $x = a$:

$$u_t - ((x - a)^2 u_x)_x = f, \text{ in } (-1, 1)^2,$$

with initial condition $u(x, -1) = u_0(x)$ and homogeneous Dirichlet boundary conditions. Here, $a \in (-1, 1)$ is a constant.

Let

$$B_1 = 2 \text{diag}(x_h - a) \otimes I_{n+1},$$

and

$$B_2 = \text{diag}(x_h - a)^2 \otimes I_{n+1},$$

the spectral scheme is

$$(I_{n+1} \otimes D - B_1(D \otimes I_{n+1}) - B_2(D^2 \otimes I_{n+1}))u_h = f_h.$$

Deleting the known boundary and initial values, the final scheme reads

$$(I_{n-1} \otimes [D] - \hat{B}_1([D] \otimes I_n) - \hat{B}_2([D^2] \otimes I_n))\hat{u}_h = \hat{f}_h - (u_{0h} \otimes d_h).$$

Here,

$$\hat{B}_1 = 2 \text{diag}(x_{h0} - a) \otimes I_n,$$

and

$$\hat{B}_2 = \text{diag}(x_{h0} - a)^2 \otimes I_n.$$

The diffusion coefficient vanishes at $x = a$ and it therefore constitutes a degeneracy in the equation, along which the parabolic character weakens due to the loss of uniform ellipticity in the spatial operator. From the perspective of stability analysis, the presence of degeneracy suggests a potential loss of regularity and amplification of numerical errors near $x = a$. One might therefore anticipate the similarly characteristic V- or U-shaped error behavior observed in ill-posed or mixed-type problems. Surprisingly, however, our numerical experiments indicate that the scheme remains

remarkably stable for this degenerate parabolic case, and no pronounced instability pattern is observed. In our experiment, the constant a is generated randomly by MATLAB. Figure 3.3f illustrates the stable convergence of the spectral error with $a = 0.6324$.

3.7 Full potential equation

This part is devoted to the mathematical background of the full potential equation. The construction of numerical schemes and the corresponding computational results will be presented in Chapter 7.

In aerodynamics, the full potential equation plays a fundamental role in the mathematical modeling of compressible flows. It is derived from the Euler equations under the assumption of irrotationality, where the velocity field can be expressed as the gradient of a scalar potential function. This formulation provides a significant simplification of the governing equations while still retaining the essential nonlinear features of compressible fluid dynamics. The full potential equation is capable of describing a wide range of flow regimes, from subsonic to supersonic, and is particularly useful in the study of full potential equations, where both subsonic and supersonic regions coexist within the same flow field. In this regime, the equation degenerates along the sonic line, changing its mathematical type from elliptic to hyperbolic, which introduces substantial analytical and numerical challenges. As a result, the full potential equation has served as a central model problem in both theoretical investigations of mixed-type PDEs and in the development of computational methods for aerodynamics.

Full potential equation occupies a central position in modern aerodynamics because it frequently occurs in practical engineering applications such as the design of aircraft wings, turbine blades, and jet engine inlets. In this regime, the local Mach number varies across the flow field, so that subsonic and supersonic regions coexist, separated by a sonic line or surface. The interaction of these regions often leads to the formation of shock waves, which significantly affect lift, drag, and overall aerodynamic performance. From the mathematical perspective, the governing equations—such as the full potential equation—become a mixed-type PDE, elliptic in the subsonic domain and hyperbolic in the supersonic domain. This change of type across the sonic boundary gives rise to inherent difficulties in both analysis and computation, as classical techniques designed for purely elliptic or purely hyperbolic problems are no longer directly applicable. Consequently, the full potential equation has long been regarded as a benchmark problem that not only advances the understanding of

nonlinear mixed-type PDEs but also drives the development of innovative numerical algorithms in computational fluid dynamics.

We give a concise summation of the full potential equation of aerodynamics. The equation of mass conservation in a steady state fluid flow can be written as

$$\nabla \cdot (\rho v) = 0,$$

where $v = (v_1, v_2)^T$ is the velocity and ρ the local density. In addition, the flow is assumed to be irrotational, which means that there exists a velocity potential ϕ such that $v = \nabla \phi$. Denote by $q = (v_1^2 + v_2^2)^{\frac{1}{2}} = \|\nabla \phi\|_2$ the local flow speed and a , the local sound speed. Then the Mach number is defined as $M = q/a$ and M_∞ is the freestream Mach number, which determines the type of equation. Assume q_∞ , the flow speed at infinity, equals 1. After some computation, the local density can be written as

$$\rho(\phi) = \left(1 + \frac{\gamma - 1}{2} M_\infty^2 (1 - \|\nabla \phi\|_2^2)\right)^{\frac{1}{\gamma - 1}},$$

where γ represents the specific heat ratio. Without loss of generality, we choose $\gamma = 1.6$ in our experiments. The equation can be expressed as

$$\nabla \cdot (\rho(\phi) \nabla \phi) = 0.$$

We consider a non-lifting symmetric wing so the flow is symmetric about the x-axis. The domain of the equation is $(-1, 1)^2 \setminus \{(x, y) \mid -\frac{1}{2} \leq x \leq \frac{1}{2}, -1 \leq y \leq f(x)\}$, where $f(x)$ is the shape function of the wing, which makes the domain irregular. See Figure 3.2. Such an issue can be solved by the spectral domain embedding method, see [22]. Furthermore, the equation requires boundary conditions. In this thesis, the boundary conditions considered are the following:

- On the farfield boundaries Γ_4 , Γ_5 and Γ_6 , we assume $\phi = \phi_\infty = x$;
- On Γ_2 , the homogeneous Neumann condition is imposed:

$$\frac{\partial \phi}{\partial n} = 0;$$

- On Γ_1 and Γ_3 , we impose for symmetry the no penetration condition:

$$\frac{\partial \phi}{\partial y} = 0.$$

There are three main difficulties in solving this problem:

- Nonlinearity of the equation;
- M_∞ : When M_∞ is small (e.g., $M_\infty < 0.8$), the equation is elliptic (subsonic); when M_∞ is large (e.g., $M_\infty > 0.8$), the equation becomes mixed-type (transonic);
- Irregular physical domain.

The numerical scheme of this problem will be discussed in Chapter 7.

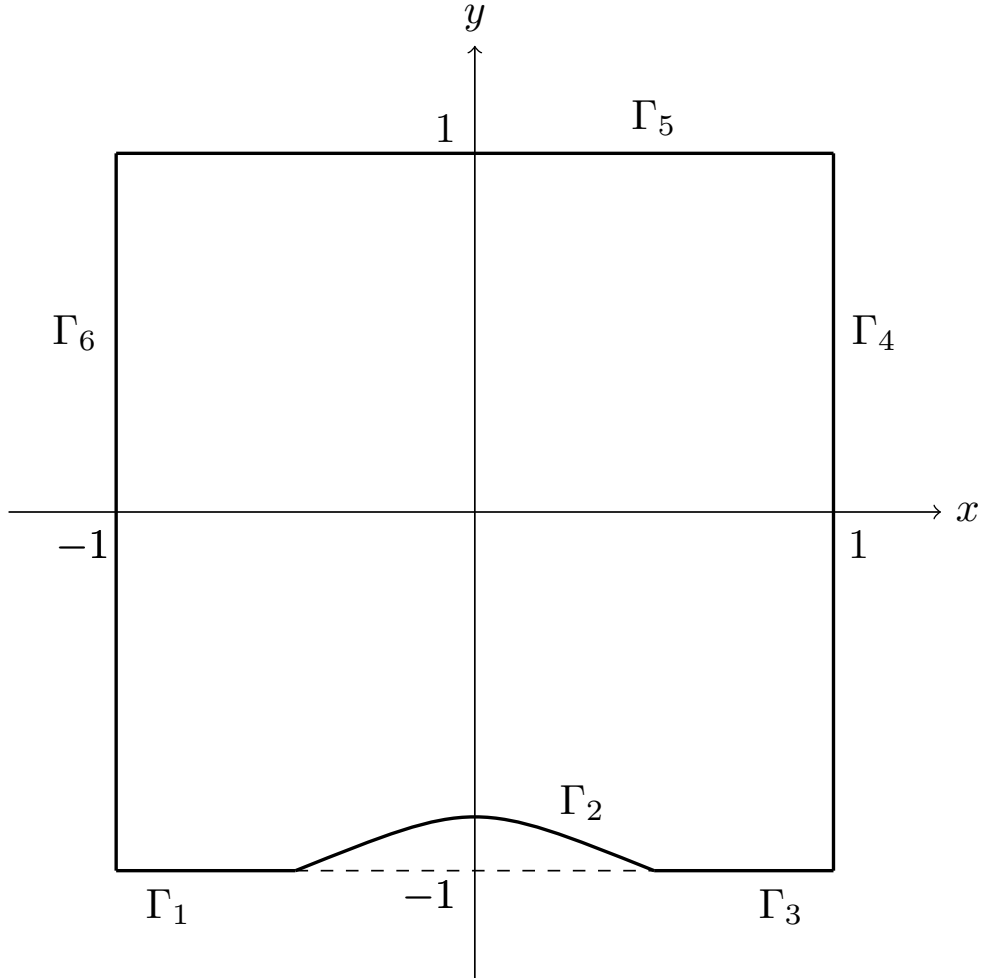
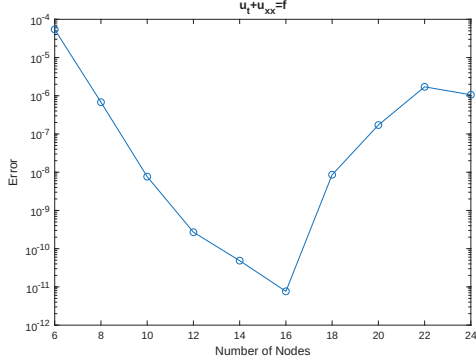


Fig. 3.2: Computational domain with vertical boundaries Γ_4 , Γ_6 and upper boundary Γ_5 . The lower boundary consists of two horizontal segments Γ_1 and Γ_3 , and a curved part Γ_2 .

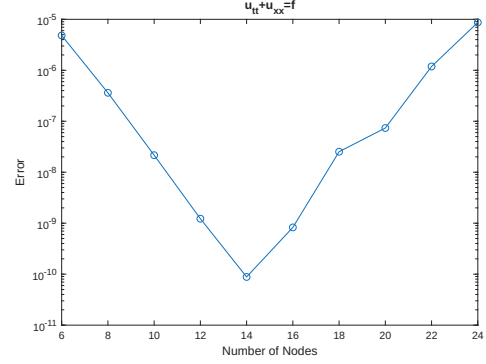
3.8 Numerical experiments

In this part, we present numerical results obtained using the space-time collocation methods for above PDEs. Clearly, V-shaped behaviors of errors can be observed

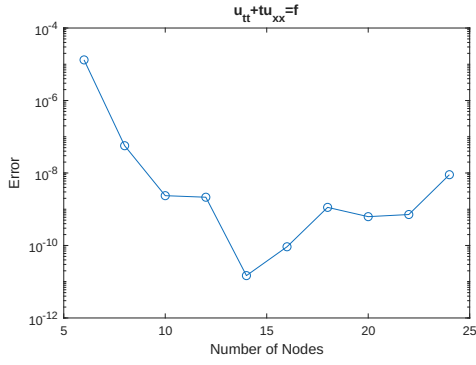
except for degenerate parabolic equations.



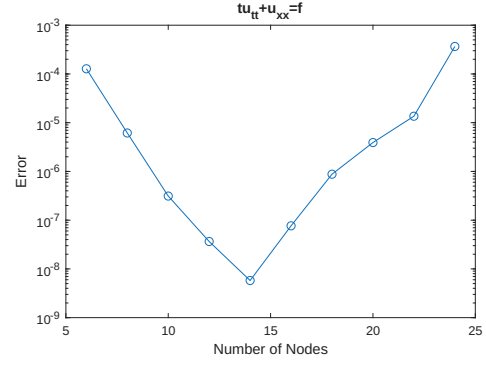
(a) Convergence for the backward heat equation



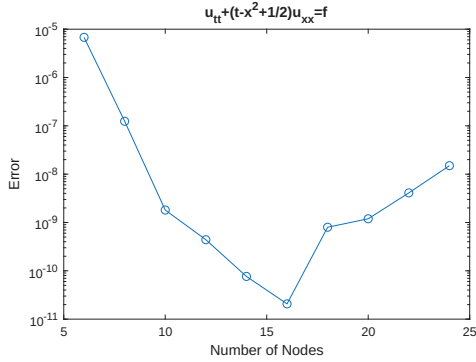
(b) Convergence for IVP for the Poisson equation



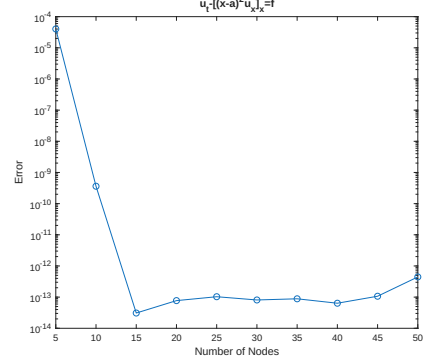
(c) Convergence for the Tricomi equation



(d) Convergence for the Keldysh equation



(e) Convergence for general mixed-type equation



(f) Convergence for the degenerate parabolic equation with $a = 0.6324$

Fig. 3.3: Numerical results obtained using space-time collocation methods.

3.9 Implications for stabilization

We emphasize that all numerical results in this chapter are obtained using space-time collocation methods without additional stabilization or regularization. This

restriction is deliberate: it isolates the intrinsic interaction between ill-posedness (and mixed-type behavior) and the collocation schemes under consideration. The results therefore serve as a baseline against which stabilization mechanisms can be assessed.

The observed V/U-shaped error behavior highlights the necessity of improved numerical strategies. In Chapters 4 and 5, we investigate two distinct approaches aimed at mitigating this instability: the first enhances conditioning at the operator level through US methods, while the second introduces an external regularization framework (SEEM-TWO). Even if these approaches do not eliminate the instability entirely, they do delay the onset of error growth and improve practical stability. Establishing the robustness and universality of the instability is thus a prerequisite for evaluating these remedies, and it constitutes the central objective of the present chapter.

4 Operator-level Stabilization: Ultraspherical Spectral Methods

As demonstrated in Chapter 3, when applied to ill-posed and mixed-type PDEs, space-time collocation methods often exhibit pronounced V- or U-shaped error curves, reflecting the intrinsic instability associated with such problems. To mitigate this phenomenon, one natural strategy is to improve the conditioning of the differential operators themselves.

In this chapter, we focus on more details about the US method. As observed in Chapter 2.2, US methods improve conditioning at the operator level, thereby delaying the growth of spectral error and enhancing the practical convergence behavior of spectral discretizations. Importantly, this improvement does not rely on any external regularization of the underlying PDEs; it arises solely from a reformulation of the differential operators in ultraspherical coefficient space.

While the primary motivation is to stabilize ill-posed or mixed-type problems, the US method can also improve the numerical behavior for well-posed equations. Therefore, in addition to ill-posed and mixed-type problems mentioned above, we also present numerical results for canonical well-posed equations such as the Helmholtz equation, the heat equation, the linear and nonlinear wave equations and biharmonic equation. These examples allow for direct comparisons between standard space-time collocation methods and their US counterparts, illustrating the broader applicability of operator-level stabilization.

4.1 Construction of ultraspherical scheme

We directly present the construction of the US methods for time-dependent problems. The model problem here is the Keldysh equation, which has the following form:

$$\begin{cases} tu_{tt} + u_{xx} = f(x, t), & \text{in } (-1, 1)^2, \\ u(\pm 1, t) = 0, \\ u(x, -1) = u_0(x), \\ u_t(x, -1) = u_1(x). \end{cases}$$

The ultraspherical formulation is constructed in three main steps: (i) construction of differentiation matrices D_n together with transition matrices S ; (ii) construction of the truncation matrix P ; and (iii) imposition of boundary conditions. These steps are discussed sequentially.

Differentiation properties of Chebyshev polynomials. Chebyshev polynomials satisfy the classical differentiation identity

$$\frac{d}{dx}T_n(x) = nU_{n-1}(x),$$

where $U_n(x)$ denotes the Chebyshev polynomials of the second kind. It is well known that $U_n(x)$ can be identified as a special case of ultraspherical polynomials with parameter $\lambda = 1$, namely

$$U_n(x) \equiv C_n^{(1)}(x).$$

More generally, for any integer $k \geq 1$, the k -th derivative of a Chebyshev polynomial belongs to the ultraspherical polynomial space with parameter $\lambda = k$:

$$\frac{d^k}{dx^k}T_n(x) \in \text{span} \left\{ C_{n-k}^{(k)}(x) \right\}.$$

This observation implies that each differentiation increases the ultraspherical parameter by one. Consequently, a k -th order derivative naturally belongs to the ultraspherical space with parameter $\lambda = k$.

4.1.1 Assembly of the differentiation matrix

Ultraspherical differentiation matrices D_k . Let $\hat{\mathbf{u}}^{(0)}$ denote the vector of expansion coefficients of a function $u(x)$ in the Chebyshev basis. We define the k -th order ultraspherical differentiation matrix

$$D_k : \quad \hat{\mathbf{u}}^{(0)} \longmapsto \hat{\mathbf{u}}^{(k)},$$

where $\hat{\mathbf{u}}^{(k)}$ represents the coefficients of the k -th derivative $\frac{d^k u}{dx^k}$ in the ultraspherical basis with parameter $\lambda = k$. Specifically, in the construction of the k -th order differentiation matrix D_k , we make use of the differentiation property of ultraspherical polynomials that relates spaces with successive parameters. Namely, ultraspherical polynomials satisfy

$$\frac{d}{dx}C_n^{(k)}(x) = 2k C_{n-1}^{(k+1)}(x).$$

By repeated application of this identity, an arbitrary k -th order derivative can be represented in coefficient space by a diagonal operator. Consequently, the k -th order differentiation matrix takes the form

$$(D_k)_{j,j+k} = 2^{k-1}(k-1)!(j+k), \quad j = 0, \dots, n-k,$$

and all other entries are zero. Note that $\hat{\mathbf{u}}^{(k)}$ now is in the $C^{(k)}$ space.

Transition matrices between ultraspherical spaces. Since different derivatives naturally live in ultraspherical spaces with different parameters, all differential terms must be mapped to a common coefficient space before they can be combined. This is achieved by introducing transition matrices between ultraspherical polynomial spaces.

Ultraspherical polynomials satisfy explicit recurrence relations that allow polynomials in $C^{(\lambda)}$ space to be expanded in terms of polynomials in $C^{(\lambda+1)}$ space for $\lambda \geq 1$. In the construction of the transition matrix S , we make use of the parameter-raising identity for ultraspherical polynomials. For $n \geq 2$, ultraspherical polynomials satisfy

$$C_n^{(\lambda)}(x) = \frac{\lambda}{\lambda + n} \left(C_n^{(\lambda+1)}(x) - C_{n-2}^{(\lambda+1)}(x) \right).$$

Accordingly, the transition matrix mapping coefficients from the $C^{(\lambda)}$ basis to the $C^{(\lambda+1)}$ basis is given by

$$S_\lambda = \text{diag} \left(\frac{\lambda}{[\lambda : n]} \right) + \text{diag} \left(-\frac{\lambda}{[\lambda + 2 : n + \lambda]}, 2 \right),$$

where $[a, b]$ represents the column vector $[a; a + 1; a + 2; \dots; b]$ for any integers a and b with $a \leq b$. In particular, the product

$$S_\lambda S_{\lambda-1} \dots S_1 S_0$$

maps coefficient vectors from the Chebyshev space $C^{(0)}$ to the ultraspherical space $C^{(\lambda+1)}$.

Handling variable coefficient terms. For the model equations under consideration, a key difficulty arises from the presence of variable coefficient terms, such as tu_{tt} . In the ultraspherical spectral framework, multiplication by the independent variable t can be represented as a sparse operator in the ultraspherical polynomial basis. Specifically, the ultraspherical polynomials satisfy the recurrence

$$t C_n^{(\lambda)}(t) = \frac{n+1}{2(n+\lambda)} C_{n+1}^{(\lambda)}(t) + \frac{n+2\lambda-1}{2(n+\lambda)} C_{n-1}^{(\lambda)}(t), \quad n \geq 1,$$

with the special case for $n = 0$ given by

$$t C_0^{(\lambda)}(t) = \frac{1}{2\lambda} C_1^{(\lambda)}(t).$$

For $\lambda = 2$, this recurrence yields a tridiagonal multiplication matrix M_t such that

$${}_t C_0^{(2)} = \frac{1}{4} C_1^{(2)}, \quad {}_t C_n^{(2)} = \frac{n+3}{2(n+2)} C_{n-1}^{(2)} + \frac{n+1}{2(n+2)} C_{n+1}^{(2)}, \quad n \geq 1. \quad (2)$$

This representation allows the variable coefficient term tu_{tt} to be implemented as the sparse operator composition

$$M_t D_2 \hat{\mathbf{u}}^{(0)},$$

preserving the banded structure of the ultraspherical discretization.

This is illustrated for the Keldysh equation below. Based on the constructions of the differentiation matrix, transition matrix, and coefficient multiplication matrix described above, the discrete Keldysh operator in the ultraspherical spectral formulation can be written as

$$A = D_2 \otimes S + S \otimes (M_t D_2),$$

where

$$S = S_1 S_0.$$

We notice that A takes functions in $C^{(0)}$ space to $C^{(2)}$ space. At this point, we have completed the construction of the differentiation matrix in the ultraspherical formulation

4.1.2 Construction of the truncation matrix

The second step of the ultraspherical formulation is the construction of the truncation matrix P . The introduction of P serves two essential purposes.

First, the truncation matrix is required to ensure that after imposing boundary conditions, the resulting discrete ultraspherical matrix is square. This is crucial for the formulation of a well-defined linear system and for the efficient application of standard solvers.

More importantly, the truncation matrix eliminates the zero rows introduced by differentiation in the Chebyshev basis. Each differentiation increases the number of leading zero rows in the corresponding differentiation matrix by one, reflecting the loss of low-order modes under differentiation. If left untreated, these zero rows enlarge the coefficient space unnecessarily and lead to rank-deficient systems.

As a consequence, the construction of the truncation matrix P is intrinsically related to the highest order of differentiation appearing in the governing equation.

For a differential operator involving k -th order derivatives, k zero rows must be removed in the corresponding spectral discretization.

For the Keldysh problem considered in this work, the operator involves second-order derivatives in both the spatial variable x and the temporal variable t . Therefore, two zero rows must be truncated in each direction. The resulting truncation matrix is given by

$$P = I(n-1, n+1) \otimes I(n-1, n+1),$$

where $I(m, n)$ denotes the $m \times n$ rectangular identity matrix.

We emphasize that the truncation matrix P defines the finite-dimensional subspace in which the discrete problem is solved. Consequently, P must be applied not only to the ultraspherical differential operator constructed in the previous step, but also to the right-hand side of the equation. The resulting truncated system therefore takes the form

$$PA\mathbf{u} = P\mathbf{f},$$

ensuring the consistency of the discretization and the stability of the resulting linear system.

4.1.3 Imposing boundary conditions

The final step of the ultraspherical formulation is the imposition of boundary conditions. We briefly describe the τ -method and then focus on the construction of homogeneous boundary solutions, which is the primary approach adopted in this work.

τ -method. The first classical approach for enforcing boundary conditions is τ -method, which sacrifices some equations to impose boundary conditions directly. It is sometimes argued that τ -method may reduce effective accuracy by replacing interior equations with boundary constraints. However, in the present ultraspherical formulation, this interpretation must be revisited. Due to the use of high-order differential operators, the discretized operators naturally generate zero rows associated with uncontrolled high-frequency degrees of freedom. These rows do not contribute any meaningful equations to the system.

Consequently, when τ -method is applied to such zero rows, no effective equations are discarded. In this setting, the boundary constraints merely occupy rows that would otherwise be inactive, therefore, do not induce any additional loss of accuracy. This observation explains why, for certain model equations (except for Keldysh

problem), τ -method performs comparably to, or in some cases even better than, the approach based on explicitly constructed homogeneous boundary solutions.

Homogeneous initial-and-boundary technique. For the Keldysh equation, we introduce a new unknown function

$$w(x, t) = u(x, t) - u_0(x) - (t + 1) u_1(x).$$

It can be verified directly that w satisfies homogeneous boundary and initial conditions while the PDE for w reads

$$\begin{cases} tw_{tt} + w_{xx} = \tilde{f}(x, t), & \text{in } (-1, 1)^2, \\ w(\pm 1, t) = 0, \\ w(x, -1) = 0, \\ w_t(x, -1) = 0, \end{cases}$$

where $\tilde{f}(x, t)$ is the modified right-hand side with respect to the function $w(x, t)$ and it is discussed in the following paragraph. As a consequence, the original problem is transformed into an equivalent problem for w with homogeneous boundary and initial conditions.

Modification of the right-hand side. The introduction of w requires a corresponding modification of the forcing term. For the Keldysh equation, the right-hand side becomes

$$\tilde{f}(x, t) = f(x, t) - (u_0)_{xx}(x) - (t + 1) (u_1)_{xx}(x).$$

The transformed equation for w is then solved under homogeneous boundary and initial conditions.

Basis construction in the spatial direction. Since $w(\pm 1, t) = 0$, the solution in the spatial direction is expanded using basis functions that vanish at both endpoints. A convenient option is

$$\phi_k(x) = T_k(x) - T_{k+2}(x), \quad k \geq 0,$$

and this choice induces a basis transformation from the Chebyshev coefficients to the homogeneous boundary basis, represented by a matrix B_x , and it can be expressed

as follows:

$$(B_x)_{i,j} = \begin{cases} 1, & i = j, \\ -1, & i = j + 2, \quad 1 \leq i \leq N_x, \quad 1 \leq j \leq N_x - 2, \\ 0, & \text{otherwise.} \end{cases}$$

Basis construction in the temporal direction. In the temporal direction, the homogeneous conditions are

$$w(x, -1) = 0, \quad w_t(x, -1) = 0$$

must be satisfied simultaneously. Accordingly, we seek basis functions of the form

$$\psi_k(t) = T_k(t) + a_k T_{k+1}(t) + b_k T_{k+2}(t), \quad k = 0, \dots, N_t - 3,$$

where the coefficients a_k and b_k are determined by enforcing the two homogeneous initial conditions. Solving the resulting linear system yields

$$a_k = \frac{4(k+1)}{2k+3}, \quad b_k = \frac{2k+1}{2k+3},$$

and

$$B_t = \begin{pmatrix} 1 & & & & \\ a_0 & 1 & & & \\ b_0 & a_1 & 1 & & \\ \ddots & \ddots & \ddots & \ddots & \\ & & b_{n_t-3} & a_{n_t-2} & 1 \\ & & & b_{n_t-2} & a_{n_t-1} \\ & & & & b_{n_t-1} \end{pmatrix}.$$

With these basis transformations, the solution automatically satisfies all homogeneous boundary and initial conditions since

$$\phi_k(\pm 1) = 0, \quad \psi_k(-1) = \psi'_k(-1) = 0, \quad k \geq 0.$$

The ultraspherical discretization can be applied to the transformed problem without further constraints.

4.1.4 Final discrete system and recovery of the physical solution

Combining differentiation matrices, transition matrices, truncation matrix, and the basis transformations associated with homogeneous boundary and initial conditions, the final discrete ultraspherical formulation for the Keldysh problem takes the form of

$$PA(B_x \otimes B_t)\hat{\mathbf{w}} = P\tilde{\mathbf{f}},$$

where $\hat{\mathbf{w}}$ denotes the coefficient vector of $w(x, t)$ in the homogeneous basis.

Since the numerical solution obtained from the above system corresponds to the variable w rather than the original unknown u , the real numerical solution is recovered through the inverse transformation

$$u(x, t) = w(x, t) + u_0(x) + (t + 1) u_1(x).$$

4.1.5 Summary of the ultraspherical formulation

Overall, the ultraspherical spectral formulation proceeds through three main steps. First, differentiation operators are represented by sparse ultraspherical differentiation matrices together with transition matrices that reconcile the parameter shifts induced by differentiation. Second, a truncation matrix is introduced to remove the zero rows generated by high-order derivatives and to ensure a well-defined square system. Finally, boundary and initial conditions are imposed through the construction of homogeneous basis functions, yielding a transformed problem that can be discretized without additional constraints.

This framework results in a sparse, better-conditioned linear system whose solution automatically satisfies all boundary and initial conditions. The physical solution is subsequently recovered by an explicit lifting transformation, completing the ultraspherical discretization of the Keldysh problem. Other PDEs discussed in this thesis can be treated similarly, such as Tricomi's equation, $u_{tt} + tu_{xx} = f$, where the term with variable coefficient, tu_{xx} , can be treated in the same way.

4.2 Numerical experiments

In this part, we present numerical results obtained using US methods for a variety of PDEs. In Figure 4.1, we show convergence results using US methods for well-posed PDEs, such as heat, wave and some other linear PDEs, and the nonlinear wave

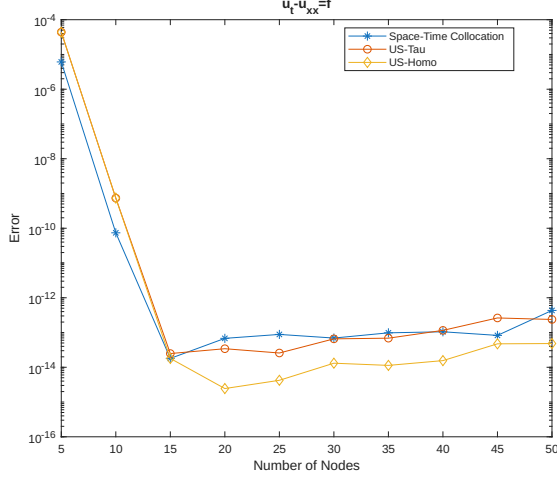
equation, which has the form

$$\left\{ \begin{array}{l} u_{tt} - \left(\frac{u_x}{\sqrt{1+u_x^2}} \right)_x = f(x, t), \quad \text{in } (-1, 1)^2, \\ u(\pm 1, t) = 0, \\ u(x, -1) = u_0(x), \\ u_t(x, -1) = u_1(x), \end{array} \right.$$

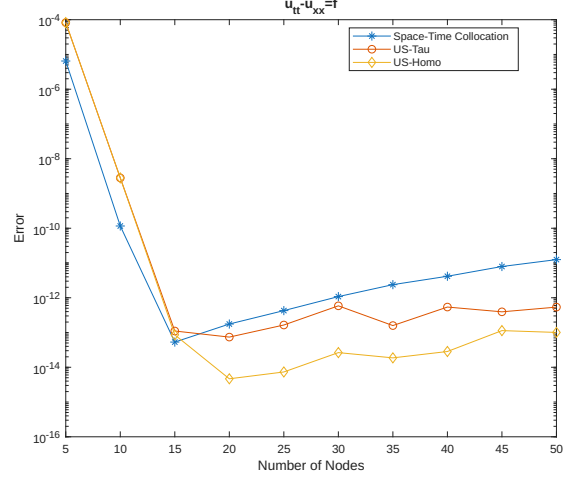
with the chosen exact solution $u(x, t) = (x^2 - 1) \ln(x + t + 5)$. The problem is solved by Picard and Newton's methods with line search and both can achieve a similar level of accuracy. The algorithm stops when the difference of two consecutive numerical solutions is less than $O(10^{-11})$ for space-time collocation methods when the number of nodes is 25 for each variable. The Picard iteration requires more than 800 steps, whereas Newton's method converges in only 19 iterations. In terms of the US methods, the stopping criterion is that the difference of two consecutive numerical solutions is less than $O(10^{-15})$ and only 79 iterations are needed for the Picard scheme with NGMRES(5), which will be discussed in subsection 4.3.

In Figure 4.2, we show convergence results using the same schemes for ill-posed and mixed-type PDEs. To demonstrate the effectiveness of operator-level stabilization, each example includes a comparison among three approaches:

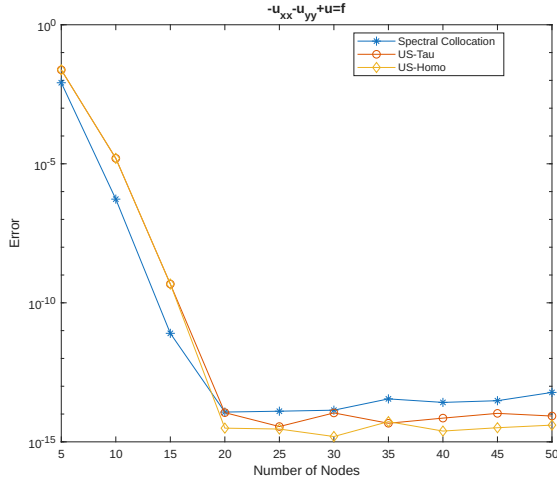
1. standard space-time collocation methods;
2. US method combined with τ -method;
3. US method with homogeneous boundary technique.



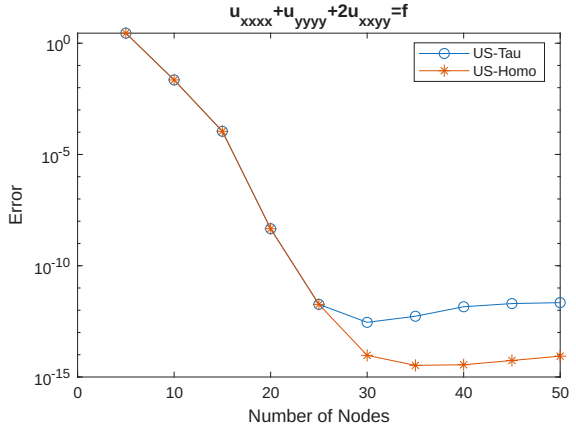
(a) Convergence for the heat equation



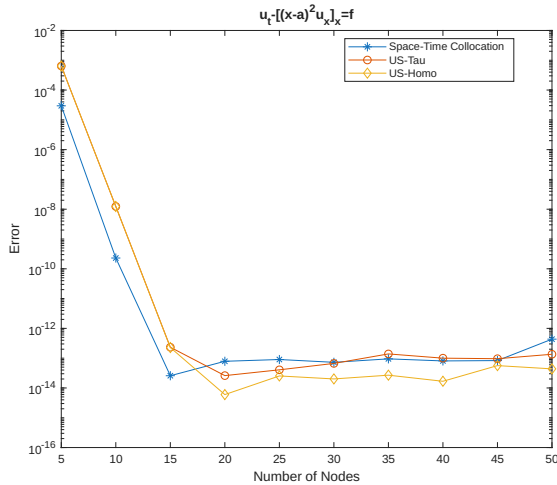
(b) Convergence for the wave equation



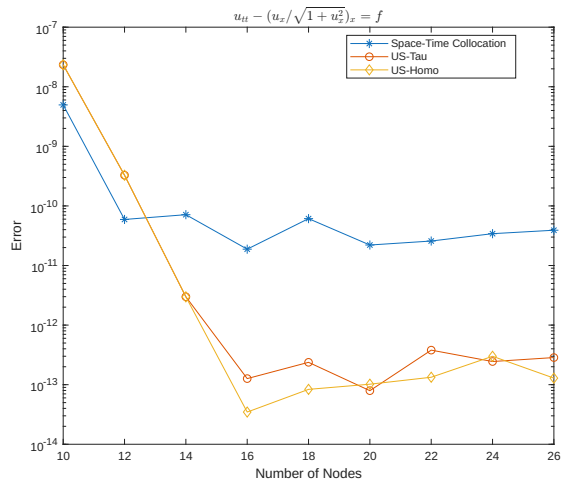
(c) Convergence for the Helmholtz equation



(d) Convergence for the biharmonic equation with homogeneous Dirichlet boundary conditions

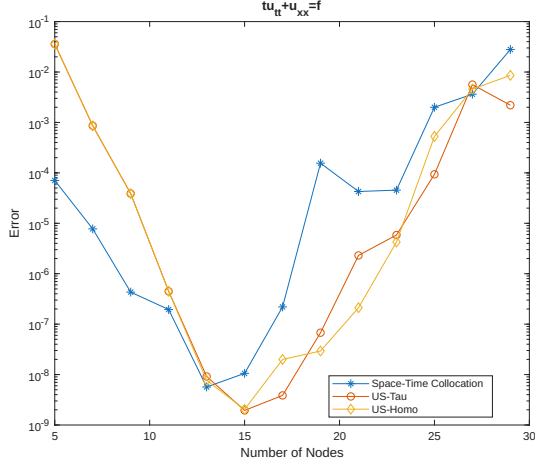


(e) Convergence for the degenerate parabolic equation with $a = 0.6324$

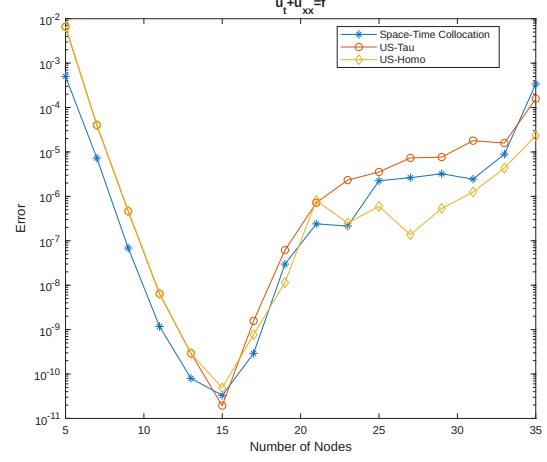


(f) Convergence for the nonlinear wave equation

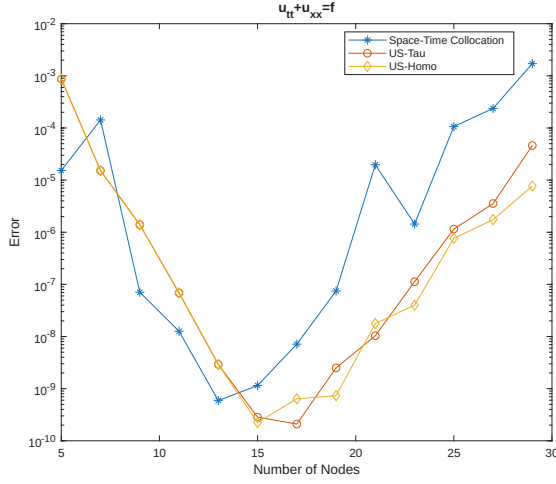
Fig. 4.1: Numerical results obtained using US method for well-posed PDEs in contrast to space-time collocation methods.



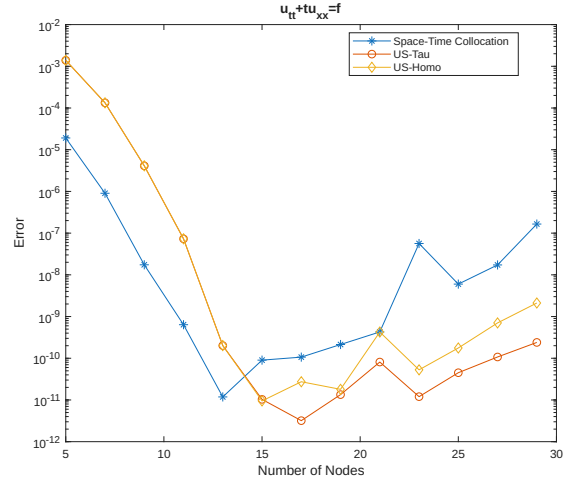
(a) Convergence for the Keldysh equation



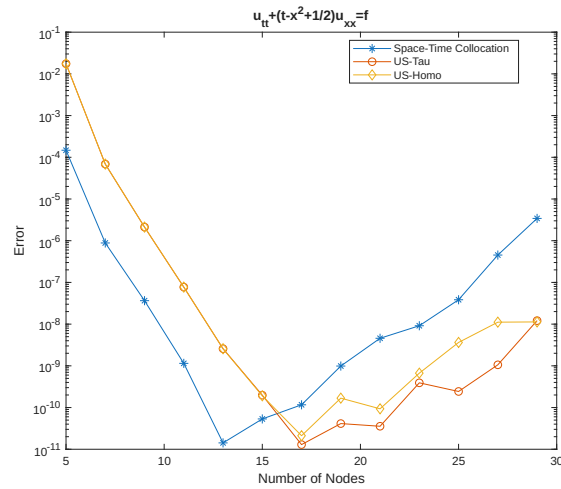
(b) Convergence for the backward heat equation



(c) Convergence for IVP for the Poisson equation



(d) Convergence for the Tricomi equation



(e) Convergence for a general mixed-type equation

Fig. 4.2: Numerical results obtained using US method for ill-posed PDEs in contrast to space-time collocation methods.

Figure 4.1f shows that the US scheme is more stable in solving the nonlinear wave equation than using space-time collocation methods. This result indicates that the US method can also improve the stability of spectral discretizations for nonlinear PDEs, even though the construction of differentiation matrices is more involved than in space-time collocation methods.

From the numerical experiments presented in this chapter, it is clear that the ultraspherical scheme can improve numerical stability when compared with the standard space-time collocation methods for all PDEs tested, except for the backward heat equation. However, while the growth of numerical error is reduced, the underlying ill-posedness or mixed-type nature of the problem is not eliminated. As a result, the ultraspherical approach should be viewed as a partial stabilization strategy rather than a complete remedy.

4.3 Ultraspherical methods with stabilization via Anderson acceleration and nonlinear-GMRES

In addition to operator-level stabilization, further numerical robustness and acceleration of nonlinear iteration can be achieved by employing Anderson acceleration (AA(m)) and nonlinear GMRES (NGMRES(m)) to enhance both the convergence speed and stability of the iterations arising from the ultraspherical discretization of the nonlinear wave equation. The parameter m denotes the number of previous iterates retained in the acceleration subspace and controls the amount of historical information used to construct the update.

AA(m) is a multiseant method that improves fixed-point iterations by forming linear combinations of previous iterates. Similarly, NGMRES(m) can be viewed as a nonlinear extension of the GMRES method, in which a Krylov-like subspace of dimension m is built from the most recent residuals. In practice, larger values of m generally improve robustness at the expense of increased computational cost. More details about the contents can be found in [34, 38].

To illustrate the stabilizing effect of these acceleration techniques, we consider a nonlinear wave equation with the exact solution

$$u(x, t) = a(x^2 - 1) \ln(x + t + 5),$$

where the parameter a controls the magnitude of the solution. When using space-time collocation methods combined with Picard iterations, the numerical scheme becomes increasingly sensitive to the initial guess as a increases. In particular, stable

convergence is observed only for $|a| \leq 1.3$.

In contrast, when acceleration techniques are applied, the robustness of the iteration is significantly improved. AA(5) is able to make the numerical scheme converge up to $|a| = 3.2$; particularly, using NGMRES(5), convergence can be achieved for values of a up to $|a| = 3.8$, even when the initial guess is not carefully tuned. This represents a substantial enlargement of the admissible solution magnitude and demonstrates that acceleration methods can play an important role in stabilizing nonlinear spectral discretizations. Figure 4.3 shows the error behavior obtained using the US method combined with NGMRES(5) for $a = 3.8$, demonstrating stable convergence well beyond the range attainable by standard Picard iteration.

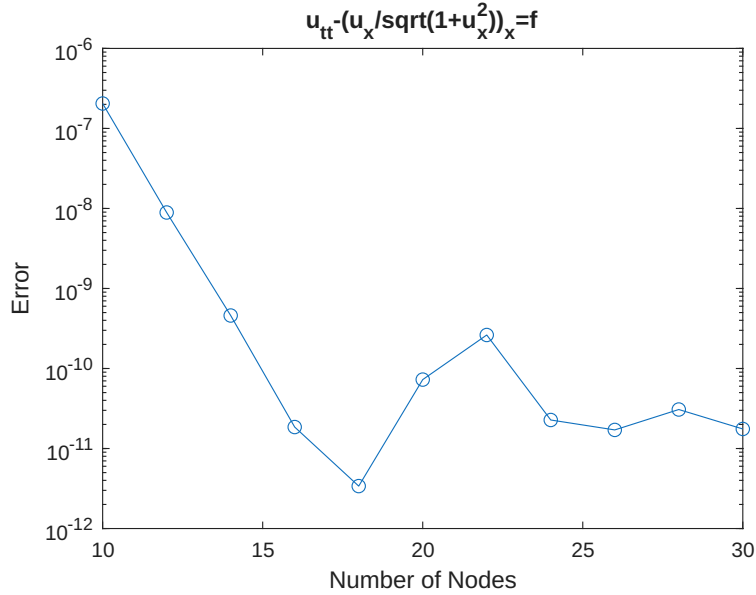


Fig. 4.3: Convergence for the nonlinear wave equation using US methods combined with NGMRES(5), with solution magnitude parameter $a = 3.8$.

5 Smooth Extension Embedding Method with Two Sets of Nodes

In this chapter, we propose a novel regularization-based approach, termed the smooth extension embedding method with two sets of nodes, to stabilize the numerical solution of ill-posed and mixed-type PDEs when using space-time collocation methods. Unlike US methods introduced in the previous chapter, which improve numerical stability at the operator level, SEEM-TWO achieves stabilization through a regularization mechanism.

As will be demonstrated, the proposed approach combines the high accuracy of spectral methods with the stabilizing effect of regularization, yielding numerical solutions that exhibit enhanced regularity compared with those produced by standard space-time collocation methods. From this viewpoint, SEEM-TWO can be interpreted as a particular form of regularization tailored to spectral discretizations. By extending the computational framework and introducing a smoother, the method can suppress the growth of high-frequency components that are responsible for numerical instability in ill-posed and mixed-type problems.

The SEEM-TWO method is developed based on the original smooth extension embedding method, proposed by Agress and Guidotti [1, 2], but differs from it in a fundamental way. In the original SEEM formulation, a single set of spectral collocation points is used to discretize both the physical domain and the fictitious domain, and the collocation points that lie inside the physical domain are simply treated as discretization points for the physical problem. In contrast, SEEM-TWO employs two distinct sets of spectral collocation points: one dedicated to the physical domain and another to the fictitious domain. This separation allows greater flexibility in controlling the regularity of the extended solution and leads to improved numerical stability for ill-posed and mixed-type equations. For additional background on SEEM and related embedding formulations, see [11, 39].

5.1 Description of the method

We begin by illustrating the method using a standard Dirichlet Poisson problem although the stabilizing effect becomes more apparent for ill-posed and mixed-type equations. Consider the following problem:

$$\begin{cases} -\Delta u = f, & \text{in } \Omega = (-1, 1)^2, \\ u = 0, & \text{on } \partial\Omega. \end{cases}$$

In SEEM-TWO, instead of directly solving the problem, we view the original problem as a constrained optimization problem in a fictitious domain, $\mathbb{B}$, for which $\overline{\Omega} \subset \mathbb{B}$. In other words, the physical domain, Ω is embedded in a larger one. Two distinct sets of nodes are introduced: one for the physical domain and one for the fictitious domain, discretizing the physical domain and the fictitious domain, respectively. Indeed, we assume that in the physical domain, both spatial and temporal directions are discretized by n nodes, while N nodes are used to discretize both directions of the fictitious domain, $\mathbb{B}$. We stress that $N > n$ always holds, that is, more nodes are used in the fictitious domain than in the physical one. When the underlying PDE is ill-posed or of mixed-type, high-frequency components may grow rapidly. By allocating additional degrees of freedom in the fictitious domain, these oscillatory components can be accommodated in the extended region, allowing the restriction to the physical domain to remain smoother. In this thesis, $N = 2n$ is used. However, since the original problem is only considered as a constraint on the physical domain, where fewer nodes are used, the problem is under-determined, which generally leads to non-uniqueness. This issue can be overcome by requiring additional regularity of the numerical solutions. Let $\|\cdot\|_S$ be the discretization of a high order norm, for instance, $\|\cdot\|_S = \|(1 - \Delta_{\mathbb{B}})^{p/2} \cdot\|_2$, where $-\Delta_{\mathbb{B}}$ denotes the Laplacian on $\mathbb{B}$ with homogeneous Dirichlet boundary conditions and $p \geq 1$. Now the problem becomes

$$\min_{u \in \mathbb{R}^N} \frac{1}{2} \|u\|_S^2 \quad \text{subject to } Cu = f.$$

where $C=AR$ consists of the discrete differential operator $-\Delta$ on Ω . Here A is the discrete differential operator of the given problem and R is the restriction matrix taking a function defined on nodes in the fictitious domain $\mathbb{B}$ onto those in the physical domain Ω defined by interpolation. Such constrained optimization can be reformulated as an unconstrained minimization

$$\min_{u \in \mathbb{R}^N, \lambda \in \mathbb{R}^n} \left(\frac{1}{2} \|u\|_S^2 + \lambda^T (Cu - f) \right),$$

where λ is the Lagrange multiplier for the constraint. A direct computation obtains the following saddle point system:

$$\begin{bmatrix} S & C^T \\ C & 0 \end{bmatrix} \begin{bmatrix} u \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ f \end{bmatrix},$$

where S is the invertible discretization matrix with respect to the norm $\|\cdot\|_S$. Such S is called a smoother and plays an important part in this approach. Assume C has

rank n and S is symmetric positive definite (SPD) ensuring the invertibility. Then the regularized solution is

$$u = S^{-1}C^T(CS^{-1}C^T)^{-1}f. \quad (3)$$

This saddle-point system can be solved using a pseudoinverse method. The solution satisfies

$$u = S^{-\frac{1}{2}}(CS^{-\frac{1}{2}})^\dagger f,$$

where $(CS^{-\frac{1}{2}})^\dagger$ is the pseudoinverse of $CS^{-\frac{1}{2}}$. To see this, we first need the following proposition.

Proposition 1. *Let $B \in \mathbb{R}^{n \times N}$ with $\text{rank}(B) = n \leq N$. Then*

$$B^\dagger = B^T(BB^T)^{-1}.$$

Proof. Since $\text{rank}(B) = n$, the matrix $BB^T \in \mathbb{R}^{n \times n}$ is SPD and hence invertible. Let

$$X := B^T(BB^T)^{-1}.$$

We verify that X satisfies the four Moore-Penrose conditions.

First,

$$BX = BB^T(BB^T)^{-1} = I_n$$

and

$$BXB = IB = B.$$

Second,

$$XBX = B^T(BB^T)^{-1}BB^T(BB^T)^{-1} = B^T(BB^T)^{-1} = X.$$

Third,

$$(BX)^T = I_n = BX.$$

Finally,

$$(XB)^T = (B^T(BB^T)^{-1}B)^T = B^T(BB^T)^{-1}B = XB,$$

since BB^T is symmetric. Thus X satisfies all Moore-Penrose conditions, and therefore $B^\dagger = B^T(BB^T)^{-1}$. $\square$

Let $B = CS^{-1/2} \in \mathbb{R}^{n \times N}$. Then

$$\begin{aligned} u &= S^{-1}C^T(CS^{-1}C^T)^{-1}f \\ &= S^{-1/2}S^{-1/2}C^T(CS^{-1/2}S^{-1/2}C^T)^{-1}f \\ &= S^{-1/2}B^T(BB^T)^{-1}f. \end{aligned}$$

By the above proposition, since $B^T(BB^T)^{-1} = B^\dagger$, we obtain

$$u = S^{-1/2}B^\dagger f = S^{-1/2}(CS^{-1/2})^\dagger f.$$

In experiments, to obtain the pseudoinverse, the singular value decomposition of the explicit matrix $CS^{-\frac{1}{2}}$ is computed first.

5.2 The smoother S

5.2.1 The artificial viscosity's perspective

The smoother S is the most characteristic feature of the method, and the choice of an appropriate smoother is crucial for the performance of the method. Inspired by the artificial viscosity method, which adds a high order spatial derivative term to the differential operator (e.g., εu_{xxxx} , where ε is called the viscous parameter and is usually small, is added to the backward heat equation), the first choice of the smoother is

$$S^{-1}u = (1 - \Delta_{\mathbb{B}})^{-p}u.$$

as mentioned previously. In the one-dimensional setting, a discrete SPD realization takes the form

$$S^{-1}u = (I + D^TWD \otimes W + W \otimes D^TWD)^{-p}u.$$

In addition, one needs to consider the order of the smoother, p . Recall that the purpose of the smoother is to enforce some level of regularity to the under-determined system. Therefore, on the one hand, we expect that a larger p gives a more regular approximation; on the other hand, however, a smoother of higher order may increase the condition number of the resultant matrices. Hence, the order of the smoother must be selected to balance the accuracy and stability. In our numerical experiments, $p = 2$ provides a favorable balance between regularization strength and conditioning.

5.2.2 The energy functional's perspective

We provide an alternative to smoothers. In particular, we use the Tricomi equation as an example. The problem is

$$\begin{cases} u_{tt} + tu_{xx} = 0, & \text{in } \Omega = (-1, 1)^2, \\ u(\pm 1, t) = 0, \\ u(x, -1) = u_0(x), \\ u_t(x, -1) = u_1(x). \end{cases}$$

We give concise steps in obtaining the discrete energy functional. For each fixed time t , we multiply the equation by u_t and integrate over $x \in (-1, 1)$:

$$\int_{-1}^1 u_{tt} u_t dx + \int_{-1}^1 t u_{xx} u_t dx = 0;$$

Integrating by parts of the left hand side and some direct computation gives

$$\frac{1}{2} \frac{\partial}{\partial t} \int_{-1}^1 u_t^2 dx - \frac{t}{2} \frac{\partial}{\partial t} \int_{-1}^1 u_x^2 dx = 0;$$

At this point, one would like to construct an energy functional $E(t)$ satisfying two natural requirements:

- (i) $E(t) \geq 0$ for all $t \in [-1, 1]$;
- (ii) $E'(t) \leq 0$ so the energy does not increase.

However, due to the mixed-type nature of the Tricomi operator, these two requirements cannot be achieved simultaneously. Indeed, the identity obtained above naturally suggests the energy

$$E_0(t) = \frac{1}{2} \int_{-1}^1 u_t^2 dx - \frac{t}{2} \int_{-1}^1 u_x^2 dx,$$

whose time derivative exactly matches the differential structure of the equation. Nevertheless, $E_0(t)$ fails to be nonnegative for $t > 0$ in the elliptic region.

For the purpose of numerical stability, especially in the SEEM-TWO framework where the smoothing operator S is required to be SPD, we prioritize the nonnegativity of the energy functional. Therefore, instead of enforcing exact structural consistency,

we modify the coefficient of the spatial term and define

$$E(t) = \frac{1}{2} \int_{-1}^1 u_t^2 dx + \frac{1-t}{2} \int_{-1}^1 u_x^2 dx,$$

which is nonnegative for all $t \in [-1, 1]$. Although this choice is no longer non-increase in the time derivative, it guarantees a positive semidefinite energy structure compatible with the stability requirements of the numerical scheme.

The resulting energy functional can be approximated by spectral quadrature by

$$E_h = \frac{1}{2} u_h^T D_t^T W D_t u_h + \frac{1}{2} u_h^T D_x^T T W D_x u_h,$$

where D_t and D_x are discrete spectral differential matrices with respect to time and space, respectively; W is the weight matrix and T is the diagonal matrix from the term $1 - t$ in the above energy functional. Numerical experiments further indicate that both this smoother and the artificial-viscosity-based smoothing strategy provide comparable stabilization effects.

5.3 SEEM-TWO and Tikhonov regularization

As mentioned above, regularization is necessary to solve ill-posed and mixed-type PDEs in a stable manner. Among the various regularization strategies, the Tikhonov regularization is one of the most fundamental and widely studied. The central idea is to stabilize an ill-posed problem by augmenting the original formulation with an additional penalty term that enforces smoothness of the solution. In this way, the influence of noise is reduced, while a well-posed problem is obtained that admits stable numerical solutions. Due to its clear variational interpretation and strong theoretical foundation, Tikhonov regularization has become a cornerstone in the analysis and numerical treatment of ill-posed problems. Excellent references on the theory include [13, 14].

We first give the definition within the framework of functional analysis:

Definition 1. *Let X and Y be Hilbert spaces, and $A, L: X \rightarrow Y$ be bounded linear operators, $f \in Y$ and $\alpha > 0$. The functional*

$$J_\alpha(u) = \|Au - f\|^2 + \alpha \|Lu\|^2, \quad u \in X,$$

is called the Tikhonov functional, where L is called the regularization operator and the second term of the right-hand side is called the penalty term, while α is the regu-

larization parameter.

The discrete Tikhonov regularization seeks to determine an approximation of u , denoted by u_h , by solving for the following minimization problem:

$$J_h(u_h) = \min_{u_h \in \mathbb{R}^n} \|A_h u_h - f_h\|^2 + \alpha \|L_h u_h\|^2$$

under some chosen norm, where f_h is the discrete vector of values of the function f ; A_h and L_h are matrices with respect to discretization of operators A and L , respectively.

Next we give our main result, which shows that SEEM-TWO is a regularization:

Theorem 1. *Assume $L_h^T L_h$ and $A_h^T A_h$ are non-singular. The SEEM-TWO solutions 3 can be interpreted as the limit as $\alpha \rightarrow 0$, of a Tikhonov regularized formulation with regularizer $L_h^T L_h = S$.*

Proof. By the above definition,

$$\begin{aligned} J_h(u_h) &= \|A_h u_h - f_h\|^2 + \alpha \|L_h u_h\|^2 \\ &= (A_h u_h - f_h)^T (A_h u_h - f_h) + \alpha (L_h u_h)^T (L_h u_h). \end{aligned}$$

Then at a minimum of J_h ,

$$\frac{\partial J_h}{\partial u_h} = 2A_h^T (A_h u_h - f_h) + 2\alpha L_h^T L_h u_h = 0,$$

it follows that

$$(A_h^T A_h + \alpha L_h^T L_h) u_h = A_h^T f_h,$$

or

$$u_h = (\alpha L_h^T L_h + A_h^T A_h)^{-1} A_h^T f_h.$$

By the Sherman-Morrison-Woodbury formula,

$$(W + UV)^{-1} = W^{-1} - W^{-1}UYVW^{-1}, \quad Y = I_p + VW^{-1}U,$$

where W is an $n \times n$ non-singular matrix, U and V are $n \times p$ and $p \times n$ matrices, respectively, with $p \leq n$. Assume $W = \alpha L_h^T L_h$, $U = A_h^T$, $V = A_h$ and $Y = I +$

$A_h(\alpha L_h^T L_h)^{-1} A_h^T$. It follows

$$\begin{aligned}
(\alpha L_h^T L_h + A_h^T A_h)^{-1} A_h^T &= (\alpha L_h^T L_h)^{-1} A_h^T - (\alpha L_h^T L_h)^{-1} A_h^T (I + A_h(\alpha L_h^T L_h)^{-1} A_h^T)^{-1} A_h(\alpha L_h^T L_h)^{-1} A_h^T \\
&= (\alpha L_h^T L_h)^{-1} A_h^T (I - (I + B)^{-1} B) \\
&= (\alpha L_h^T L_h)^{-1} A_h^T ((I + B)^{-1} (I + B - B)) \\
&= (\alpha L_h^T L_h)^{-1} A_h^T (I + B)^{-1} \\
&= (\alpha L_h^T L_h)^{-1} A_h^T (I + A_h(\alpha L_h^T L_h)^{-1} A_h^T)^{-1} \\
&= (L_h^T L_h)^{-1} A_h^T (\alpha I + A_h(L_h^T L_h)^{-1} A_h^T)^{-1},
\end{aligned}$$

where I denotes the identity matrix and $B = A_h(\alpha L_h^T L_h)^{-1} A_h^T$.

Then

$$u_h = (L_h^T L_h)^{-1} A_h^T (\alpha I + A_h(L_h^T L_h)^{-1} A_h^T)^{-1} f_h.$$

Identifying the regularization operator $L_h^T L_h$ with the smoother S , it follows

$$u_h = S^{-1} A_h^T (\alpha I + A_h S^{-1} A_h^T)^{-1} f_h.$$

Taking the limit as $\alpha \rightarrow 0$ yields precisely the SEEM-TWO formulation (3) with $A_h = C$. $\square$

The above result shows the connection between SEEM-TWO and the classical Tikhonov regularization. The main difference is that, in Tikhonov method, one can expect regular solutions by setting desired penalty term; while in SEEM-TWO, both smoother and more nodes in the fictitious domain can make solutions more regular.

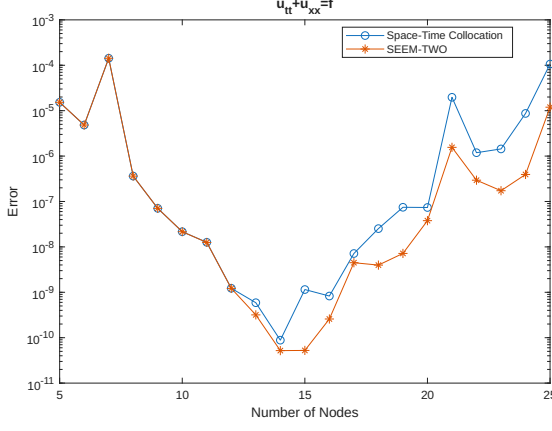
Since SEEM-TWO is formulated as a regularization-based method, it is designed to suppress instability arising from ill-posed and mixed-type PDEs. When applied to well-posed problems, the regularization does not degrade the numerical accuracy, and the resulting solutions are comparable to those obtained by space-time collocation methods as shown in the experiments.

5.4 Numerical experiments

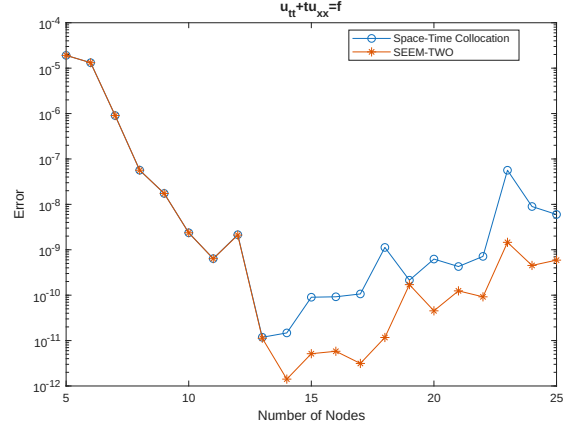
In the numerical experiments that follow, we first present results obtained by applying SEEM-TWO to a collection of ill-posed and mixed-type equations. For each example, the numerical performance of SEEM-TWO is compared with that of space-time collocation methods. As will be shown, SEEM-TWO mildly improves numerical stability. For completeness, we also include representative well-posed problems to demonstrate that the proposed regularization does not adversely affect accuracy in

stable settings.

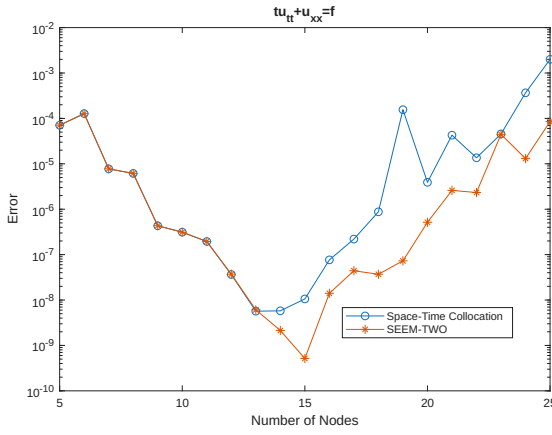
The numerical results in this chapter demonstrate that SEEM-TWO provides a regularization mechanism for stabilizing space-time collocation methods applied to ill-posed and mixed-type equations except for the backward heat equation. Similarly to the US scheme, the proposed method improves stability but does not fully remove the effects of ill-posedness. In general, achieving numerical behavior comparable to that of well-posed problems remains a challenging task, and complete stabilization of ill-posed and mixed-type equations by spectral methods is, in many cases, inherently difficult.



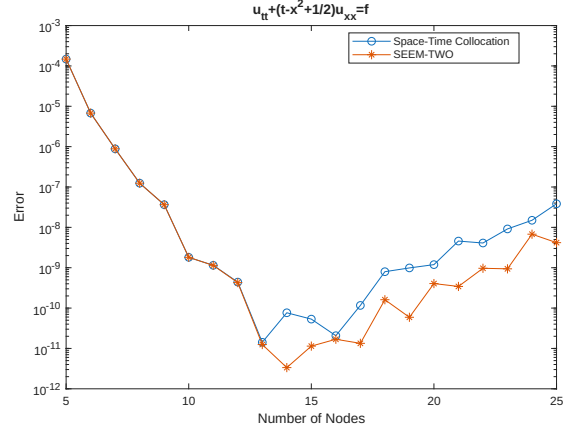
(a) Convergence for IVP for the Poisson equation



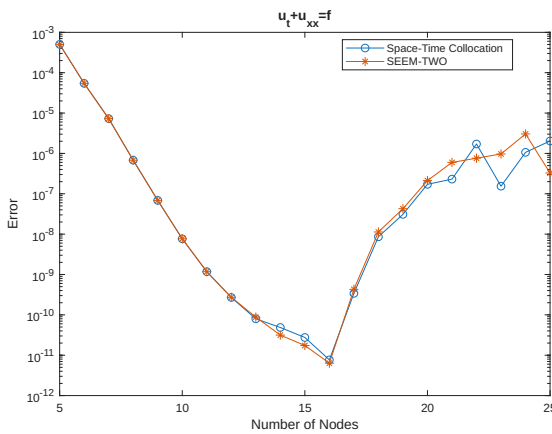
(b) Convergence for the Tricomi equation



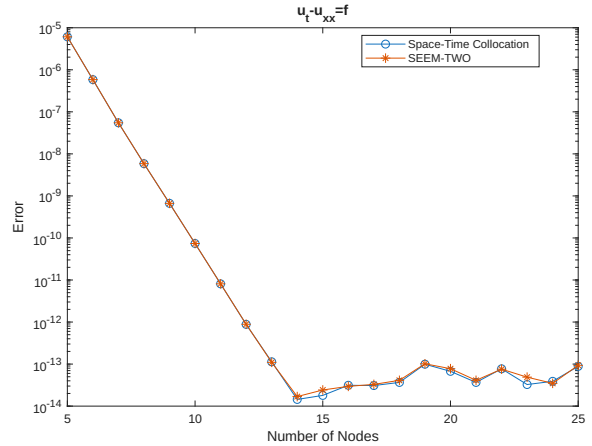
(c) Convergence for the Keldysh equation



(d) Convergence for a general mixed-type equation



(e) Convergence for the backward heat equation



(f) Convergence for the heat equation

Fig. 5.1: Numerical results obtained using the SEEM-TWO. For each problem, the performance of SEEM-TWO is compared with that of space-time collocation methods.

5.5 Remarks on other regularization strategies

While SEEM-TWO provides a structured regularization mechanism, it is by no means the only approach for stabilizing ill-posed or mixed-type problems. In fact, regularization has a long history in the numerical treatment of unstable differential equations, and numerous strategies have been developed in different contexts, including inverse problems, high-frequency instabilities, and under-resolved spectral discretizations.

To place SEEM-TWO in a broader perspective, we briefly review several classical regularization techniques that are commonly employed to suppress numerical instabilities. Our discussion here is at a survey level, aiming to highlight the underlying philosophy rather than provide an exhaustive treatment.

5.5.1 Artificial viscosity

One of the most classical stabilization mechanisms is the introduction of artificial viscosity. The central idea is to modify the original PDE by adding a higher-order dissipative term, typically of the form

$$u_t = \mathcal{L}u + \varepsilon u_{xxxx},$$

or more generally,

$$\mathcal{F}(u) + \varepsilon(-\Delta)^p u = 0,$$

where $\mathcal{F}$ denotes the discretization of the given PDEs, $\varepsilon > 0$ is a small parameter and $p \geq 1$ controls the order of dissipation.

The purpose of the added term is to damp high-frequency components of the numerical solution. In spectral discretizations, instabilities in ill-posed or mixed-type equations are often manifested through rapid growth of high-order modal coefficients. Since higher derivatives amplify high-frequency modes more strongly than low-frequency ones, the artificial viscosity term selectively penalizes oscillatory components, thereby restoring numerical stability.

From a functional-analytic viewpoint, artificial viscosity can be interpreted as a Tikhonov-type regularization in which smoothness is enforced by penalizing higher Sobolev norms. In practical computations, the parameter ε plays a crucial role: if chosen too small, the stabilization effect is insufficient; if chosen too large, the modified equation may deviate significantly from the original model and introduce excessive smoothing.

Artificial viscosity has been widely used in computational fluid dynamics, spectral

shock capture, and stabilization of hyperbolic or mixed-type systems. However, it alters the governing equation itself and therefore may introduce modeling bias. In contrast, SEEM-TWO aims to improve stability through structural reformulation and smooth embedding without explicitly modifying the differential operator.

5.5.2 Filtering

In addition to artificial viscosity, another important class of stabilization techniques is based on spectral filtering, which directly attenuates high-frequency coefficients. We consider polynomial expansions of the form

$$u_N(x) = \sum_{n=0}^N \hat{u}_n P_n(x),$$

where $P_n(x)$ denotes the n -th order of Legendre, Chebyshev or Fourier polynomial. If $u(x)$ is smooth, the spectral expansion is expected to converge rapidly. That is to say, only few terms are needed to generate an accurate approximation. However, if the approximated function is nonsmooth, the decay of the expansion coefficients, $\hat{u}_n$, is relevant to the accuracy. It is therefore natural to introduce a modification of the coefficients such that the accuracy of the truncated expansion can be improved. Of course, such modification should not destroy the convergence rate if u is already smooth.

We consider filtered expansion of the kind

$$\mathcal{F}_N u_N(x) = \sum_{n=0}^N \sigma\left(\frac{n}{N}\right) \hat{u}_n P_n(x),$$

where $\sigma(\eta)$ is a real filter function, which is the essential element in the strategy. We give the following definition.

Definition 2. *The filter function, $\sigma(\eta) \in C^p: \mathbb{R}^+ \rightarrow [0, 1]$, $p > 1$, has the following properties:*

$$\begin{cases} \sigma(0) = 1, \\ \sigma^{(k)}(0) = 0, \quad k = 1, \dots, p-1, \\ \sigma(\eta) = 0, \quad \eta \geq 1, \\ \sigma^{(k)}(1) = 0, \quad k = 1, \dots, p-1. \end{cases}$$

We call this a p -th order filter.

There are two common filter functions. The first one is called the Vandeven filter,

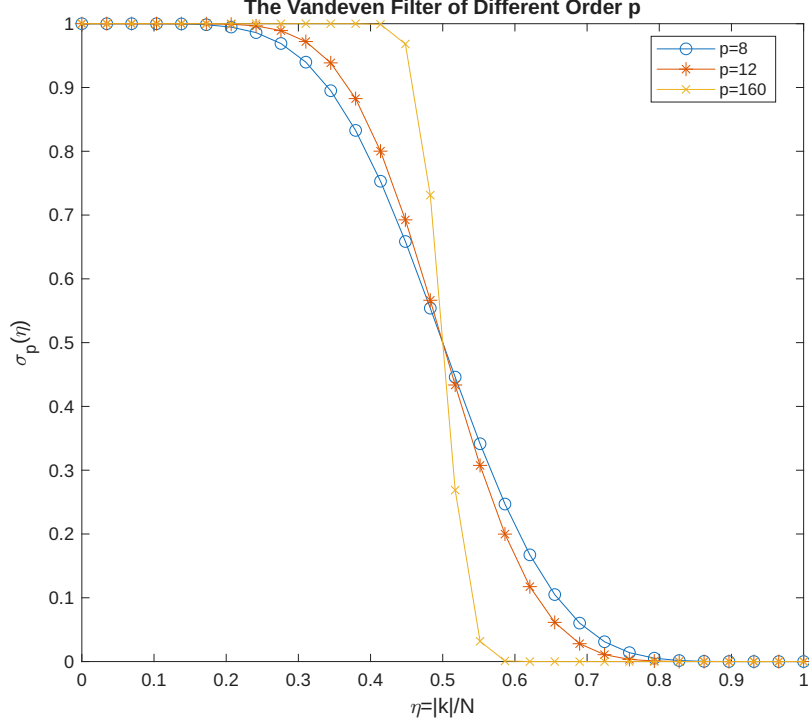


Fig. 5.2: The Vandeven filter function with different orders p .

which has the form

$$\sigma_O(\eta) = 1 - \frac{(2p-1)!}{(p-1)!^2} \int_0^\eta (t(1-t))^{p-1} dt,$$

with $p = cN^{\varepsilon/4}$, where c is a positive constant independent of N and $\varepsilon > 0$ is a real number. This is a $(2p-1)$ -th order Birkhoff-Hermite interpolating polynomial. It can be verified that $\sigma_O(\eta)$ satisfies all properties of Definition 2.

Another widespread filter is called the exponential filter, $\sigma_E(\eta)$, defined as

$$\sigma_E(\eta) = e^{-\alpha\eta^p}.$$

Here, the parameter α is a measure of the modification of the maximum mode, i.e., $\sigma_E(1) = e^{-\alpha}$. Typically, $\alpha = -\ln(\varepsilon_M)$, where ε_M is the machine accuracy. However, the exponential filter does not conform to the last property of Definition 2. In particular, we have that

$$|\sigma_E^{(k)}(1)| \simeq (\alpha p)^k e^{-\alpha},$$

which can be very far from zero for large values of k and p .

From the definitions of the two filter functions and Figure 5.2, which illustrates

the Vandeven filter functions, we observe that for larger values of p , a significant portion of the low-frequency modes remains essentially unchanged, while only the high-frequency components (for sufficiently large n) are attenuated by the filter. Since the exponential filter exhibits a qualitatively similar profile, its graph is omitted for brevity. For the class of problems considered in this thesis, neither approach was able to provide satisfactory stabilization. In particular, although these techniques may alleviate high-frequency oscillations to some extent, they do not fundamentally suppress the intrinsic instability mechanisms of the equations under consideration. Some references on filtering functions include [10, 15–17, 19, 37].

5.6 Noise amplification test for regularization effects

To further examine the stabilizing properties of the proposed schemes (US methods and SEEM-TWO), we design a numerical experiment based on the sensitivity of ill-conditioned systems to small perturbations. Consider the discrete system

$$Au = f,$$

where A denotes the discretization matrix obtained by numerical schemes. We perturb the right-hand side by adding a small random noise term,

$$Au^n = f + \alpha \text{Noise},$$

where α is a small parameter and Noise represents random perturbations. Subtracting the two systems yields

$$u - u^n = -\alpha A^{-1} \text{Noise},$$

and therefore

$$\|u - u^n\| \leq \alpha \|A^{-1}\| \|\text{Noise}\|.$$

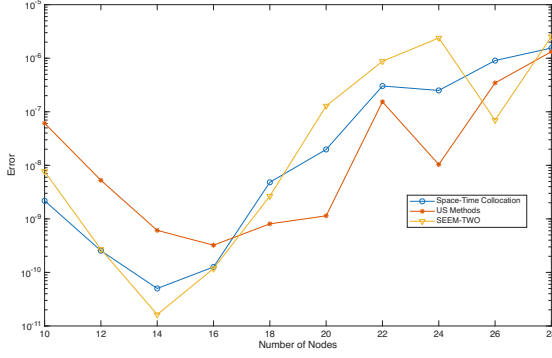
This relation reveals the basic principle of the test: if the discretization matrix A is ill-conditioned, the perturbation can be significantly amplified through the action of A^{-1} . In practice, such amplification is reflected in a rapid growth of numerical errors when small perturbations are introduced. Consequently, a more robust or regularized numerical scheme should reduce the sensitivity of the computed solution to these perturbations and exhibit improved numerical stability.

Numerical results and discussion. As shown in Figures 3.1a and 3.1b, the numerical errors produced by the finite difference method increase rapidly, even when

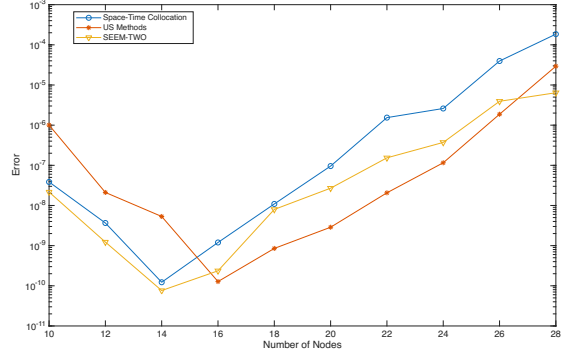
the right-hand side contains no artificial noise. Hence we mainly compare three discretization schemes here: standard space-time collocation methods, US methods and SEEM-TWO. For each problem, the same perturbation with magnitude $\alpha = 10^{-12}$ is added to the discrete right-hand side. Five representative problems are tested: the backward heat equation, IVP for the Poisson problem, the Tricomi equation, the Keldysh equation, and a general mixed-type equation. The resulting numerical errors are shown in Figures 5.3a–5.3e.

One main observation can be drawn from these experiments. Except for the backward heat equation, and for the cases $n = 22$ and $n = 26$ in the Tricomi equation, both US methods and SEEM-TWO consistently produce numerical solutions that are less sensitive to perturbations than those obtained by space-time collocation methods. This behavior indicates that the US schemes and SEEM-TWO can suppress the amplification of perturbations and therefore exhibit an intrinsic regularization effect.

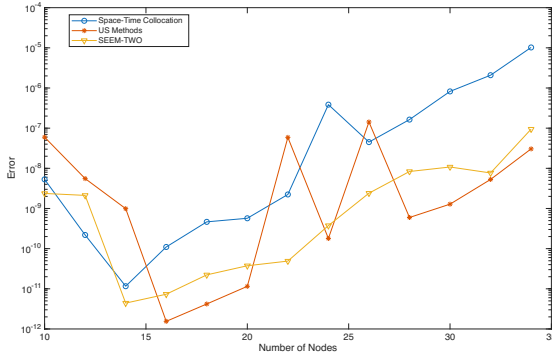
In contrast, for the backward heat equation, all three schemes display similar behavior, suggesting that the severe ill-posedness of this problem cannot be significantly mitigated by these discretization strategies alone. This observation is consistent with our earlier numerical findings that neither the US methods nor SEEM-TWO substantially improves the stability of the backward heat equation.



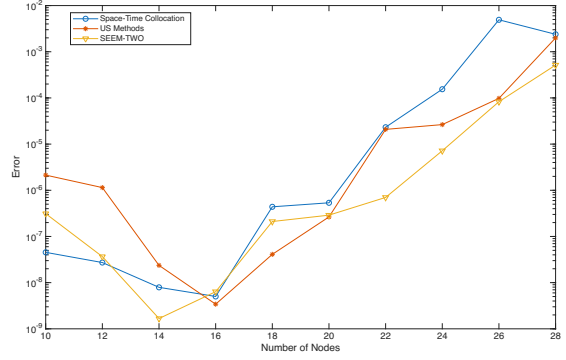
(a) The backward heat equation.



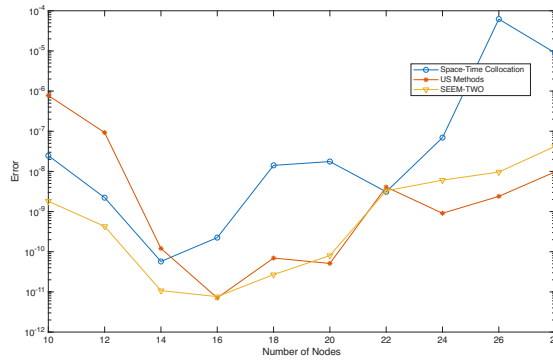
(b) IVP for the Poisson equation.



(c) The Tricomi equation.



(d) The Keldysh equation.



(e) The general mixed-type equation.

Fig. 5.3: Comparison of numerical errors obtained by space-time collocation methods, US methods, and SEEM-TWO under the same noise perturbation.

6 Balancing Accuracy and Stability: Determination of the Optimal Node Number

In this chapter, we address the problem of selecting the number of nodes that minimizes the numerical error for ill-posed and mixed-type PDEs considered above, without a priori knowledge of the exact solution.

6.1 Review: the behavior of spectral errors

Numerical experiments indicate that the spectral error exhibits a characteristic non-monotonic dependence on the number of nodes n . When n is below a critical value, denoted by n_* , the error decreases as n increases and attains a minimum near n_* . Beyond this point, the error increases as n grows further. Consequently, the error curve typically displays a V-shaped or U-shaped profile.

Therefore, a central question is how to determine the optimal number of nodes n_* that minimizes the numerical error. In the numerical experiments presented above, this optimal value can be readily identified since exact solutions are available and the true errors can be explicitly computed. In practical applications, however, exact solutions are unknown, rendering direct evaluation of the numerical error impossible.

To address this difficulty, we propose a numerical criterion that identifies suitable candidates corresponding to near-minimal errors, based solely on the computed numerical solutions without recourse to exact solutions.

6.2 Limitations of error estimation without exact solutions

When an exact solution is not available, estimating the numerical error becomes impossible. In particular, numerical solutions computed with different numbers of nodes are defined on different grids, making direct comparison between solutions not straightforward. Interpolation is therefore unavoidable, which may introduce additional numerical artifacts.

A natural idea is to compare numerical solutions obtained with different numbers of nodes and infer the optimal resolution from their differences. However, such an approach faces intrinsic difficulties for ill-posed problems. Small differences between two numerical solutions do not necessarily indicate that both solutions are close to the exact solution; instead, they may simply reflect that both solutions are dominated by similar unstable modes. Conversely, a large difference may arise even when one of

the solutions is already near-optimal.

These observations suggest that selecting the optimal number of nodes requires a criterion that is sensitive to the onset of numerical instability, rather than relying on absolute error estimates or global monotonicity assumptions.

6.3 Difference-based error indicators

Although the exact numerical error cannot be directly computed without knowledge of the exact solution, the difference between numerical solutions obtained with consecutive node numbers carries useful information. Let u_n denote the numerical solution computed using n nodes for each variable. We consider the quantity

$$\Delta_n = \|u_n - u_{n-1}\|_\infty,$$

where the solutions are compared after interpolation onto a common set of points.

Consistent with the previously observed error behavior, the difference Δ_n typically decreases within the stable regime and increases once instability-dominated modes begin to affect the numerical solution.

Importantly, the magnitude of Δ_n is largely determined by the less accurate of the two solutions u_n and u_{n-1} . As a consequence, Δ_n provides an indirect indicator of the transition from the error-decreasing regime to the error-increasing regime, without requiring access to the exact solution.

This observation motivates the use of Δ_n as a practical and computable quantity for detecting the onset of numerical instability and identifying suitable candidates for the optimal number of nodes.

6.4 An adaptive stopping criterion for node selection

Based on the difference indicator Δ_n , we propose an adaptive stopping criterion for selecting the number of nodes. The basic algorithm proceeds by increasing n sequentially and monitoring the behavior of Δ_n . The observed behavior of Δ_n provides a practical way to detect the transition from the stable regime to the instability-dominated regime. This behavior motivates the following stopping criteria:

- (i) **Consecutive increase criterion:** If $\Delta_n > \Delta_{n-1}$ and $\Delta_{n-1} > \Delta_{n-2}$, the algorithm terminates and selects $n - 2$ as a candidate for the optimal number of nodes.

- (ii) **Sudden growth criterion:** If $\Delta_n > 10\Delta_{n-1}$, indicating an increase of more than one order of magnitude, the algorithm terminates and selects $n - 1$ as a candidate.

These criteria are designed to detect both gradual and abrupt transitions from the stable regime to the unstable regime. The selected node numbers correspond to the last reliable resolutions before numerical instability dominates.

We emphasize that the proposed algorithm does not require any tolerance parameter related to the unknown exact error, nor does it assume any specific functional form of the error curve. The stopping decision is made solely based on the observed behavior of the computed numerical solutions.

The following pseudocode presents the logical structure of the proposed algorithm, omitting implementation details such as interpolation construction and linear solver choices.

Algorithm 1 Adaptive node selection based on consecutive solution differences

```

1: Choose an initial range of nodes  $\{n_0, n_0 + 1, \dots, n_{\max}\}$ 
2: for  $n = n_0 + 1$  to  $n_{\max}$  do
3:   Compute numerical solution  $u_n$ 
4:   Interpolate  $u_n$  and  $u_{n-1}$  onto a common set of points
5:   Compute  $\Delta_n = \|u_n - u_{n-1}\|_\infty$ 
6:   if  $\Delta_n > \Delta_{n-1}$  and  $\Delta_{n-1} > \Delta_{n-2}$  then
7:     stop and select  $n - 2$ 
8:   else if  $\Delta_n > 10\Delta_{n-1}$  then
9:     stop and select  $n - 1$ 
10:  end if
11: end for

```

Starting from a moderate number of nodes, typically $n_0 = 10$ and $n_{\max} = 20$, the algorithm incrementally increases the resolution and computes the corresponding numerical solutions. At each step, two consecutive solutions are interpolated onto a common set of points, and their difference in the maximum norm is evaluated.

As long as the discretization remains stable, this difference indicator decreases or stays at a low level. Once instabilities emerge, the indicator increases sharply, either through consecutive growth or via a sudden jump in magnitude. The algorithm terminates when such behavior is detected and selects a node number preceding the onset of instability.

To improve robustness of the algorithm in cases where the spectral error exhibits local fluctuations caused by low-frequency oscillatory exact solutions, we propose a

coarse-to-fine strategy (Algorithm 2). Unlike smooth solutions, for which the spectral error generally decreases monotonically and Algorithm 1 reliably identifies the optimal node number, low-frequency oscillatory solutions may produce local error spikes that can mislead Algorithm 1 into premature termination. For Algorithm 2, we first obtain a candidate with the step size 2, denoted by m_* ; then we find two candidates within the interval $[m_* - 1, m_* + 3]$, where the stopping criterion is that any of the difference of consecutive numerical solutions increases once. Assume m_* is known. The following pseudocode shows the structure:

Algorithm 2 Robust algorithm for solutions with oscillations

```

1: Obtain  $m_*$  from Algorithm 1 with the step size 2; store  $u_{m_*}$  and  $u_{m_*+2}$ 
2: for  $n = m_* - 1$  to  $m_* + 3$  do with a step size of 2
3:   Compute numerical solution  $u_n$ 
4:   Interpolate  $u_n$  and  $u_{n-1}$  onto a common set of points
5:   Compute  $\Delta_n = \|u_n - u_{n-1}\|_\infty$ 
6:   if  $\Delta_n > \Delta_{n-1}$  then
7:     stop and select  $n - 1$ 
8:   else Select  $m_* + 3$ 
9:   end if
10: end for

```

This two-stage coarse-to-fine strategy ensures that the algorithm is less sensitive to local error fluctuations caused by oscillatory solutions. Numerical experiments in the next subsection show that Algorithm 2 successfully identifies the optimal node number for all tested PDEs for oscillatory solutions.

6.5 Numerical experiments

To assess the performance of the proposed method, we apply it to several ill-posed and mixed-type PDEs discretized by space-time collocation methods. We first test Algorithm 1 on a relatively smooth exact solution

$$u(x, t) = (x^2 - 1) \ln(x + t + 5),$$

for which the optimal number of nodes n_* for all the equations considered lies within the interval $[10, 20]$. This corresponds to cases where the solution exhibits no oscillatory behavior and thus can be accurately resolved with a small number of spectral nodes. For solutions exhibiting low-frequency oscillatory behavior over more general intervals, Algorithm 2 is employed.

Since the search range is moderate and the error curve exhibits a clear minimum, Algorithm 1 is sufficient to identify n_* . The results are summarized in Table 6.1.

Equation	Ideal n_*	Algorithm 1
Backward heat	16	15, 16
IVP for Poisson	14	13, 14
Tricomi	13	13, 14
Keldysh	14	13, 14
General mixed-type	13	13, 14

Table 6.1: Optimal values of n obtained by Algorithm 1 for different test problems.

To further examine the performance of the two proposed search algorithms, we consider a more oscillatory exact solution

$$u(x, t) = e^{xt} \sin(a(x^2 - 1)),$$

where the parameter a controls the oscillation frequency. Increasing a leads to larger optimal numbers of nodes, allowing us to test the algorithms on a wider search interval $[4, 40]$. In our experiments, a is chosen to be 5 so that n_* lies in the above interval. We apply both Algorithm 1 (unit step search) and Algorithm 2 (step size two) to tested problems. The results are summarized in Table 6.2.

Equation	Ideal n_*	Algorithm 1	Algorithm 2
Backward heat	25	27, 28	25, 26
IVP for Poisson	24	23, 24	23, 24
Tricomi	29	*	29, 30
Keldysh	24	*	23, 24
General mixed-type	27	29, 30	27, 28

Table 6.2: Comparison of Algorithm 1 and Algorithm 2 for the oscillatory exact solution $u(x, t) = e^{xt} \sin(5(x^2 - 1))$ on the search interval $[4, 40]$. Entries marked by * indicate that Algorithm 1 terminated prematurely and failed to detect the optimal n_* .

For the backward heat equation, the optimal node number is $n_* = 25$, while Algorithm 1 returns $n = 27$ or 28 . A closer examination of the computed error values shows that the errors corresponding to n from 25 to 28 are of similar magnitudes and differ by less than 10 times. This indicates that the minimal error is attained over a relatively flat interval rather than at a sharply isolated point in the case considered. Therefore, although Algorithm 1 does not identify the exact minimizer, the selected node number still lies within the admissible near-optimal range and is numerically acceptable.

We also observe that for oscillatory solutions, Algorithm 1 may terminate prematurely. The stopping criterion of Algorithm 1 is triggered when the difference indicator increases twice consecutively, or increases once by a factor larger than ten. However, for oscillatory problems, the error curve may exhibit local fluctuations before reaching its global minimum. As a result, the stopping condition may be satisfied prematurely, causing Algorithm 1 to terminate before the optimal number of nodes is reached. This phenomenon is particularly evident for the Tricomi and Keldysh equations, where Algorithm 1 stops early and fails to identify the optimal n_* .

In contrast, Algorithm 2 increases the number of nodes with a step size of two, which effectively reduces sensitivity to local non-monotonicity in the error curve. By sampling the error at a coarser resolution in the initial stage, the algorithm captures the overall trend of the error decay rather than reacting to small fluctuations caused by discretization effects. Moreover, due to the larger step size, the stopping criterion requires multiple consecutive increases in the sampled error values, making premature termination significantly less likely. In practice, such sustained growth in the error is only observed when the number of nodes approaches or exceeds the optimal value n_* .

Next, we investigate the theoretical computational costs for two algorithms. Since linear systems arising from the space-time collocation discretization grow rapidly with n , a comparison based solely on the number of solves may be misleading. To obtain a more realistic measure of computational complexity, we assume that each Kronecker-structured matrix equation is solved by the Bartels–Stewart algorithm [5], which exploits the Sylvester structure directly. Instead of forming the full $n^2 \times n^2$ Kronecker system—whose solution via Gaussian elimination would require $O(n^6)$ operations—the Bartels–Stewart algorithm reduces the cost to $O(n^3)$. Therefore, the total computational cost of an algorithm is estimated by summing n_i^3 , where i is the index over all solves performed during the search process. The Bartels–Stewart algorithm is applied to all tested equations except for the general mixed-type equation. In this case, the variable coefficient leads to a discretized operator with splitting rank 3, which does not admit a Sylvester-type formulation. As a result, the Bartels–Stewart algorithm cannot be applied, and a direct Gaussian elimination is employed instead.

Equation	# Solves 1	# Solves 2	Ideal Cost	Cost 1	Cost 2
Backward heat	27	16	15625	216189	150500
IVP for Poisson	21	14	13824	89964	100359
Tricomi	*	18	24389	*	231336
Keldysh	*	14	13824	*	100359
General mixed-type	29	16	3.87×10^8	5.46×10^9	4.96×10^9

Table 6.3: Comparison of Algorithm 1 and Algorithm 2 for the oscillatory exact solution $u(x, t) = e^{xt} \sin(5(x^2 - 1))$ on the search interval $[4, 40]$. Here # Solves denotes the number of linear systems solved. Cost represents the estimated total floating-point operations. Entries marked by * indicate that Algorithm 1 terminated prematurely and failed to detect the correct n_* .

Below we briefly explain how the *Cost* entries in Table 6.3 are computed. For instance, consider Algorithm 1 applied to the backward heat equation. The algorithm starts from $n = 4$ and increments until $n = 30$. Assuming Bartels-Stewart algorithm for each linear system of size $N = n^2$, the cost of a single solve is approximated by n^3 floating-point operations. Thus, the total cost is estimated as

$$\text{Cost 1} = 4^3 + 5^3 + \cdots + 30^3 = 216189.$$

When Algorithm 2 is applied to this equation,

$$\text{Cost 2} = \underbrace{4^3 + 6^3 + \cdots + 30^3}_{\text{coarse-search stage}} + \underbrace{25^3 + 27^3}_{\text{fine-search stage}} = 150500.$$

Finally, Ideal Cost is defined as n_*^3 .

We note that the experiments presented here are restricted to relatively low-frequency regimes, where the solution can be accurately resolved with a modest number of spectral nodes. For highly oscillatory solutions, the required degrees of truncation grow proportionally to the frequency, and a more extensive investigation would be needed to assess the performance of the proposed algorithms in such settings. We emphasize that the purpose of the proposed algorithms is not to recover the exact minimizer of the error curve, but to identify a numerically robust and accurate resolution using only computed data. This feature makes the method suitable for ill-posed and mixed-type problems, where reliable error estimation is generally unavailable.

7 Full Potential Equation

The mathematical background of the problem has been introduced in Chapter 3.7. The present chapter focuses on the numerical discretization and computational experiments.

7.1 Spectral discretization

In this part, we present the numerical methodology used to solve the full potential equation. In [6], the authors successfully solved the problem with a high Mach number ($M_\infty = 0.8$) using finite element methods with overlapping Schwarz method as a preconditioner. Our approach combines a global spectral discretization in the interior of the computational domain with τ -method to impose boundary conditions, including those on the curved boundary. The resulting nonlinear system is solved using either Picard fixed-point methods or Newton methods. Another relevant reference on the full potential equation is [12].

7.1.1 Spectral discretization in the interior

We employ the full spectral collocation method for discretization. In each spatial direction x and y , $n + 1$ spectral nodes are used, resulting in a total of $(n + 1)^2$ degrees of freedom in the two-dimensional square computational domain.

Let D denote the one-dimensional spectral differentiation matrix of size $(n + 1) \times (n + 1)$. The two-dimensional first-order differentiation matrices with respect to x and y directions are constructed via Kronecker products as

$$D_x = D \otimes I_{n+1}, \quad D_y = I_{n+1} \otimes D.$$

The nonlinearity in the full potential equation enters through the density function $\rho(\nabla\phi)$. For illustration, taking the Mach number, $M_\infty = 0.1$, the density term is then discretized as

$$\rho = \text{diag}\left(1.003 - 0.003 \left((D_x\phi)^2 + (D_y\phi)^2\right)\right)^{5/3},$$

where the square and the power are taken componentwisely. This leads to a diagonal matrix ρ of size $(n + 1)^2 \times (n + 1)^2$.

With this notation, the discrete interior operator takes the form

$$A = D_x \rho D_x + D_y \rho D_y,$$

which represents a spectral approximation of the divergence-form operator $\nabla \cdot (\rho (\nabla \phi) \nabla \phi)$ in the interior of the computational domain.

7.1.2 Irregular physical domain

One of the main numerical challenges in solving the full potential equation using spectral methods lies in the treatment of irregular physical domains. In the present problem, the middle part of the lower boundary, Γ_2 , is a concave curve defined on the interval $x \in (-\frac{1}{2}, \frac{1}{2})$.

Specifically, in our experiment, the curved boundary Γ_2 is given by

$$\Gamma_2 : f = \frac{1}{a} \cos(\pi x) - 1,$$

where the parameter a controls the curvature. In our numerical experiments, we consider both $a = 1$ and $a = 100$ to investigate the influence of boundary curvature.

Boundary conditions on Γ_2 , as well as on the remaining boundaries, are imposed using τ -method. Rather than modifying the interior operator, τ -method enforces boundary conditions by replacing selected equations in the spectral system, allowing for a unified treatment of straight and curved boundaries within the same global discretization framework. Since there is no condition along the straight line segment directly underneath Γ_2 , the technique we use to incorporate spectral discretization in the irregular domain is the spectral embedding method [22], where the physical domain is embedded in the large square domain.

7.1.3 Nonlinear iteration: Picard and Newton schemes

Combining the interior spectral discretization with boundary conditions leads to a nonlinear algebraic system of the form

$$\mathcal{A}(\phi_h^k) \phi_h^{k+1} = F,$$

where ϕ_h^k denotes the numerical solution obtained by the k -th iteration and the operator $\mathcal{A}(\phi_h)$ depends nonlinearly on the solution through the density $\rho(\nabla \phi_h)$.

For the above discretization scheme, the primary solver is the Picard fixed-point

method, in which the density ρ is frozen in the previous iterate. This results in a sequence of linear problems with the same boundary treatment but updated coefficients. We note that in our experiments, the Newton method is also applied to the same formulation and typically achieves faster convergence; however, both approaches yield comparable accuracy once convergence is attained.

7.1.4 Mach number continuation strategy for larger M_∞

Numerical experiments show that both Picard and Newton methods converge robustly for Mach numbers in the range $M_\infty = 0.1$ to 0.3 . For higher Mach numbers, typically $M_\infty = 0.4$ to 0.7 , the increased strength of the nonlinearity renders the iteration highly sensitive to initial guess.

To overcome this difficulty, we adopt a Mach number continuation strategy, inspired by standard continuation (homotopy) techniques for nonlinear PDEs. Starting from a converged solution at a lower Mach number (e.g., $M_\infty = 0.3$), the resulting numerical solution is used as the initial guess for a slightly larger Mach number. By gradually increasing M_∞ in this manner, spectral convergence can be maintained up to $M_\infty = 0.7$.

This continuation approach significantly enlarges the range of Mach numbers for which the fully spectral method remains stable and accurate, without modifying the underlying discretization or boundary treatment.

In this thesis, we are able to solve the full potential equation when the flow is subsonic everywhere so that the equation remains elliptic. For a truly transonic flow, since the spectral method is a global high-order method with low dissipation, it encounters Gibbs phenomenon in the supersonic region. There we must, for instance, apply filtering techniques or artificial viscosity to damp the non-physical high frequency oscillations. Other approaches include using a spectral element method, where the supersonic region is localized in a spectral element. These techniques are very much in the forefront of research and beyond the scope of this thesis.

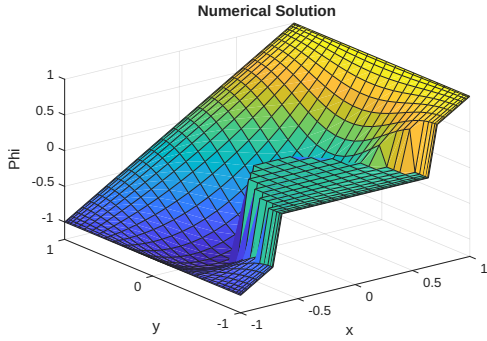
7.2 Numerical experiments

Numerical experiments are performed for two different curvature parameters imposed on the lower boundary Γ_2 (see Figure 3.2), namely $a = 1$ and $a = 100$. These two choices correspond to strongly and mildly curved boundary configurations, respectively, and allow us to examine the robustness of the method with respect to geometric variations. The smooth exact solution for the full potential

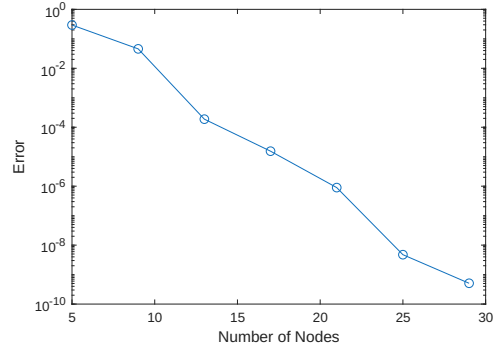
equation with a non-homogeneous right-hand side F in this thesis is chosen as $\phi(x, y) = x + (x^2 - 1) \sin^2 \frac{\pi(y+1)}{2}$. This analytical solution is introduced only for the purpose of verifying the stability and convergence of the numerical scheme. Consequently, the boundary data is determined from the exact solution, which leads to a non-homogeneous boundary condition on Γ_2 . Although this setting does not represent the non-penetration condition usually imposed on the airfoil, it provides a smooth manufactured solution that allows us to assess the stability and accuracy of the numerical method. For $M_\infty = 0.1$ and 0.6, standard spectral methods are robust enough to solve the problems; the algorithm stops when the difference of two consecutive numerical solutions is less than $O(10^{-13})$. We also demonstrate figures of local Mach numbers and streamlines for $M_\infty=0.1, 0.6$ and 0.7, respectively.

7.2.1 Subsonic regime: $M_\infty = 0.1$

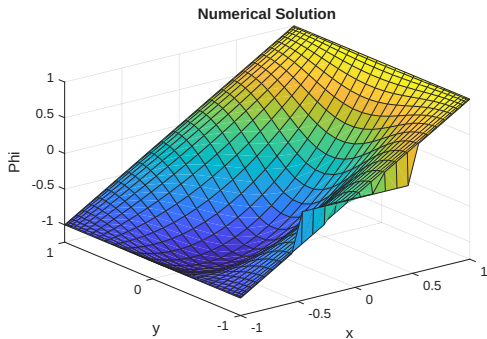
We first consider the subsonic regime with Mach number $M_\infty = 0.1$, for which the full potential equation remains elliptic throughout the computational domain.



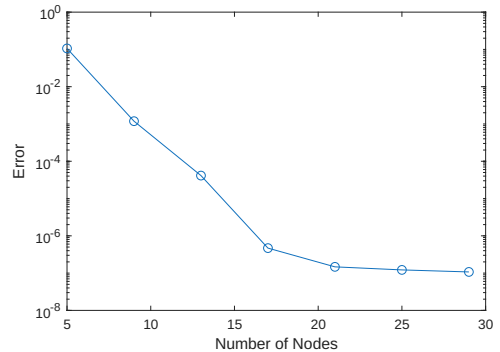
(a) Numerical solution for $M_\infty = 0.1$, $a = 1$.



(b) Spectral error for $M_\infty = 0.1$, $a = 1$.



(c) Numerical solution for $M_\infty = 0.1$, $a = 100$.



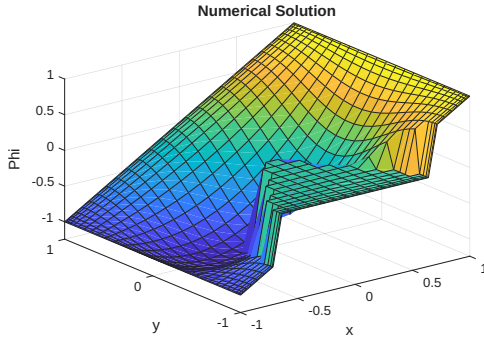
(d) Spectral error for $M_\infty = 0.1$, $a = 100$.

Fig. 7.1: Numerical solutions and errors for the subsonic case $M_\infty = 0.1$ under different boundary curvatures.

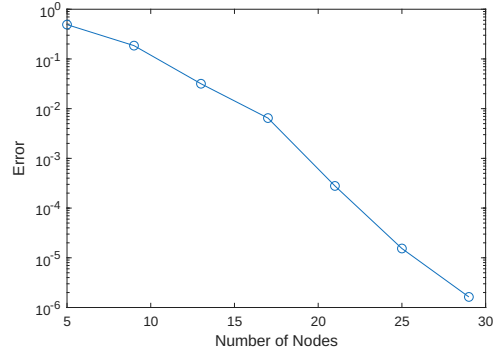
As shown in Figure 7.1, the numerical solutions remain smooth and accurate for both curvature parameters. In particular, no spurious oscillations or instability are observed even for the highly curved boundary case $a = 1$, indicating that the proposed method performs robustly in the elliptic regime.

7.2.2 Nonlinear-dominant subsonic regime: $M_\infty = 0.6$

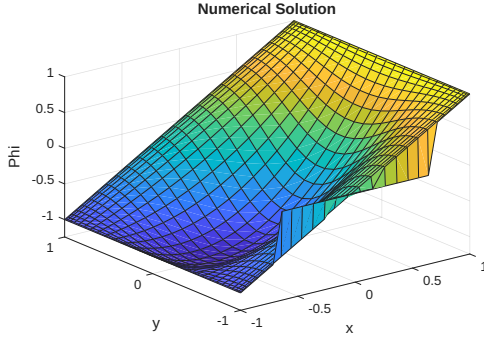
We next consider the subsonic regime with Mach number $M_\infty = 0.6$, for which the nonlinearity of the equation becomes more severe, making the scheme sensitive to the initial guess. As in the earlier case, two curvature parameters $a = 1$ and $a = 100$ are examined to investigate the influence of boundary geometry on the numerical solution.



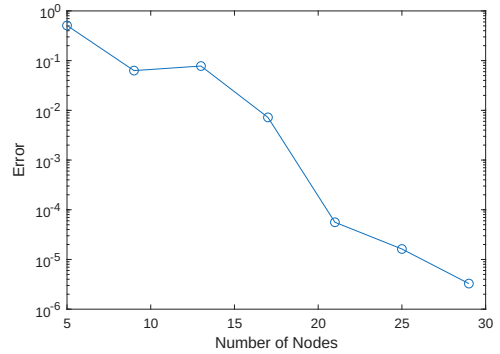
(a) Numerical solution for $M_\infty = 0.6$, $a = 1$.



(b) Spectral error for $M_\infty = 0.6$, $a = 1$.



(c) Numerical solution for $M_\infty = 0.6$, $a = 100$.



(d) Spectral error for $M_\infty = 0.6$, $a = 100$.

Fig. 7.2: Numerical solutions and errors for the nonlinearity-dominant case $M_\infty = 0.6$ under different boundary curvatures.

When the Mach number increases beyond this threshold (e.g., $M_\infty > 0.7$), the governing equation exhibits stronger mixed-type characteristics and enhanced instability. In this regime, neither US methods nor SEEM-TWO is able to provide sufficient stabilization, and the iterative solver fails to converge. Developing a struc-

turally stable spectral discretization for transonic regimes therefore remains an open research direction.

7.2.3 Streamlines and local Mach contours

To further investigate the qualitative behavior of the flow, we present streamlines together with contour plots of the local Mach number for $M_\infty=0.1$, 0.6 and 0.7. These visualizations help illustrate high-speed regions and the acceleration near the curved boundary.

The local Mach number is defined as

$$M = \frac{\|\nabla\phi\|}{a}, \quad a = \sqrt{\gamma}\rho^{(\gamma-1)/2},$$

where a is the local speed of sound. The streamlines are defined as the level curves of the stream function ψ , which satisfies

$$(\psi_y, -\psi_x) = \rho\nabla\phi.$$

Equivalently, the streamlines are tangent to the flux vector field $\rho\nabla\phi$. Fluid particles travel along streamlines.

In all experiments, we employ a Picard iteration accelerated by AA(5). The number of collocation points in each spatial direction is fixed at 40. The curved boundary Γ_2 is defined by the function $y = \frac{1}{12}\cos(\pi x) - 1$ with $a = 12$. The iteration is terminated when the difference between two consecutive iterates is less than $O(10^{-12})$ in the infinity norm. For each farfield Mach number considered, we plot both the streamline and the contour of the local Mach number.

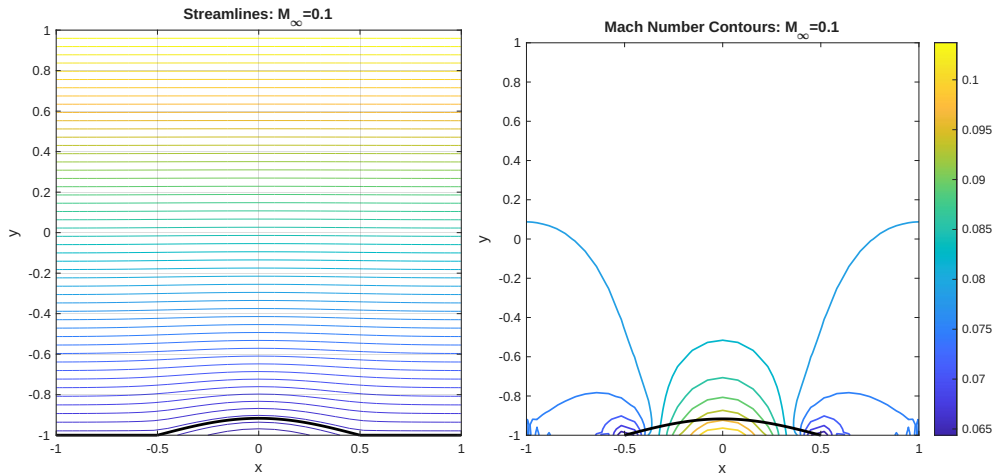


Fig. 7.3: Streamlines and Mach number contours for $M_\infty = 0.1$.

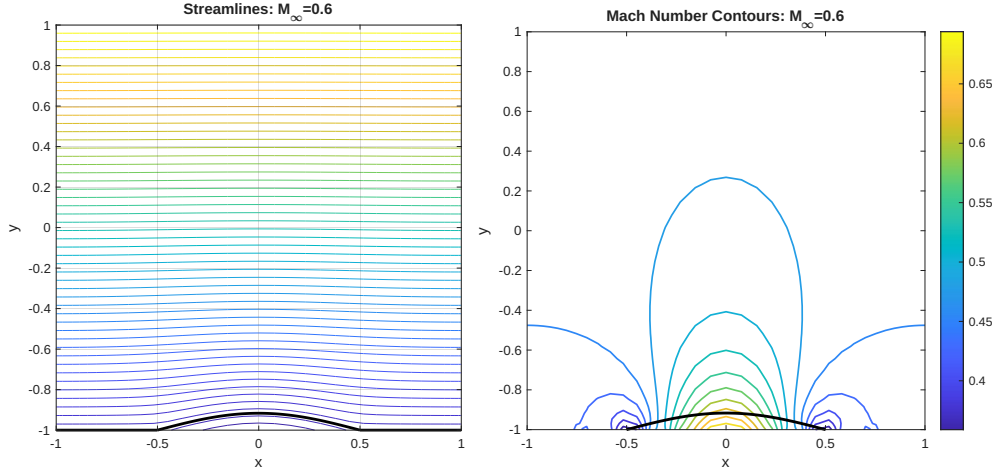


Fig. 7.4: Streamlines and Mach number contours for $M_\infty = 0.6$.

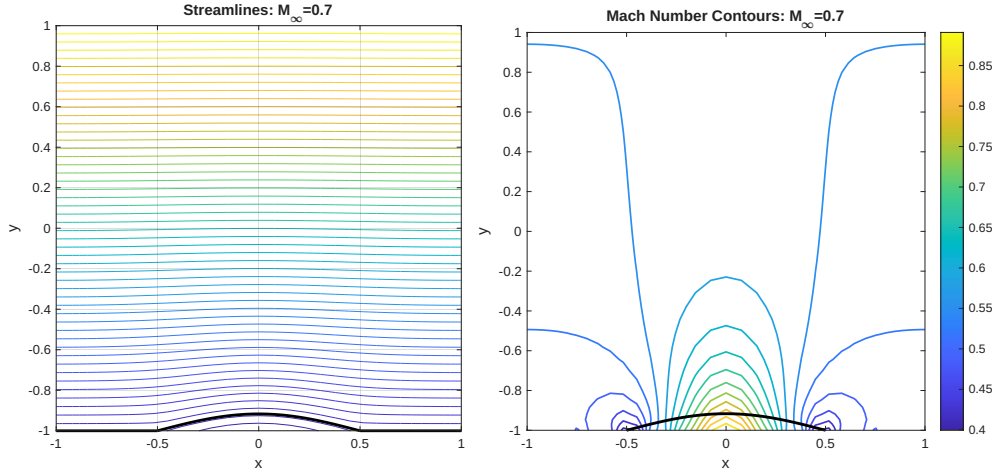


Fig. 7.5: Streamlines and Mach number contours for $M_\infty = 0.7$.

These results indicate that when $M_\infty \leq 0.7$, the flow remains subsonic everywhere, and no shocks are observed. The agreement between streamlines and local Mach number contours confirms that the full potential equation remains elliptic. This ensures that fully spectral methods can provide reliable numerical approximations.

8 Discussion and Conclusion

In this thesis, we investigate the use of space-time collocation methods for solving ill-posed and mixed-type PDEs. It is shown that when directly applied to such problems, space-time collocation discretizations may exhibit severe numerical instability, typically reflected in a V-shaped error profile as the number of nodes increases.

To address this difficulty, two stabilization strategies are proposed and studied. The first approach focuses on modifying the operator itself, leading to US methods. By reformulating the differential operators within the ultraspherical framework, the conditioning of the discrete system is improved, thereby reducing numerical instability so that more robust solutions can be obtained for a wider range of problems compared with standard space-time collocation methods.

The second approach is based on regularization. We introduce the SEEM-TWO, which incorporates a smoother into the original formulation. This regularization technique suppresses the growth of high-frequency components and further improves the numerical robustness of space-time collocation methods for ill-posed and mixed-type equations. Although neither approach completely eliminates instability in all cases, especially for the backward heat equation, both methods extend the practical applicability of spectral methods to challenging problems.

In addition, we establish a practical criterion for identifying an optimal number of spectral nodes, which balances accuracy and stability. This criterion provides practical guidance for selecting discretization parameters in numerical simulations of ill-posed and mixed-type problems.

Finally, a fully spectral scheme is applied to solve the full potential equation arising in problems on an irregular domain. Numerical results indicate that stable and accurate solutions can be obtained for Mach numbers up to $M_\infty = 0.7$.

Several directions deserve further investigation. First, for the SEEM-TWO regularization framework, a theory for the construction of appropriate smoothers remains to be developed. A rigorous analysis of how the order of the smoother, denoted by p , as well as other parameters such as the size of the fictitious domain, affect stability and accuracy would be highly valuable. Second, it is of interest to design more robust regularization strategies that enable spectral methods to stably solve a broader class of ill-posed and mixed-type equations. In the context of full potential equation, when the Mach number increases further (e.g., $M_\infty > 0.7$), the equation becomes strongly mixed-type, and the stabilization mechanisms considered in this work are no longer sufficient. Developing new regularization methods capable of handling such high Mach number regimes remains an important open problem. In addition, incorporat-

ing $AA(m)$ or $NGMRES(m)$ may provide an effective way to enhance robustness and convergence, and constitutes a promising direction for future investigation. Finally, filtering-based approaches offer a promising alternative direction. For instance, for the backward heat equation, it is known that high-frequency coefficients grow rapidly with increasing polynomial degree. This observation suggests the design of suitable spectral filters that selectively damp high-frequency components while preserving essential solution features. A carefully constructed filter may provide an effective and flexible stabilization mechanism for space-time collocation methods applied to ill-posed and mixed-type problems.

References

- [1] Daniel Agress, Patrick Guidotti, and Dong Yan. The smooth extension embedding method with Chebyshev polynomials. *Numerical Methods for Partial Differential Equations*, 39(3):2355–2377, 2023.
- [2] Daniel J Agress and Patrick Q Guidotti. The smooth extension embedding method. *SIAM Journal on Scientific Computing*, 43(1):A446–A471, 2021.
- [3] Jared Lee Aurentz and Richard Mikael Slevinsky. On symmetrizing the ultraspherical spectral method for self-adjoint problems. *Journal of Computational Physics*, 410:109383, 2020.
- [4] Anatoly Bakushinsky and A Goncharsky. *Ill-posed problems: theory and applications*, volume 301. Springer Science & Business Media, 2012.
- [5] Richard H. Bartels and George W Stewart. Solution of the matrix equation $AX+XB=C$. *Communications of the ACM*, 15(9):820–826, 1972.
- [6] Xiao-Chuan Cai, William D Gropp, David E Keyes, Robin G Melvin, and David P Young. Parallel Newton–Krylov–Schwarz algorithms for the transonic full potential equation. *SIAM Journal on Scientific Computing*, 19(1):246–265, 1998.
- [7] Shuxing Chen. A mixed equation of Tricomi–Keldysh type. *Journal of Hyperbolic Differential Equations*, 9(03):545–553, 2012.
- [8] Lu Cheng and Kuan Xu. Solving time-dependent PDEs with the ultraspherical spectral method. *Journal of Scientific Computing*, 96(3):70, 2023.
- [9] Lu Cheng and Kuan Xu. Understanding the ultraspherical spectral method. *arXiv preprint arXiv:2405.11847*, 2024.
- [10] Yana Di, Yuwei Fan, Zhenzhong Kou, Ruo Li, and Yanli Wang. Filtered hyperbolic moment method for the Vlasov equation. *Journal of Scientific Computing*, 79(2):969–991, 2019.
- [11] Benoit Fabrèges, Loïc Gouarin, and Bertrand Maury. A smooth extension method. *Comptes Rendus Mathématique*, 351(9-10):361–366, 2013.
- [12] Dalia Fishelov. Spectral methods for the small disturbance equation of transonic flows. *SIAM journal on scientific and statistical computing*, 9(2):232–251, 1988.

- [13] Martin Fuhry and Lothar Reichel. A new Tikhonov regularization method. *Numerical Algorithms*, 59(3):433–445, 2012.
- [14] Gene H Golub, Per Christian Hansen, and Dianne P O’Leary. Tikhonov regularization and total least squares. *SIAM journal on matrix analysis and applications*, 21(1):185–194, 1999.
- [15] David Gottlieb and Jan S Hesthaven. Spectral methods for hyperbolic problems. *Journal of Computational and Applied Mathematics*, 128(1-2):83–131, 2001.
- [16] David Gottlieb and Chi-Wang Shu. On the Gibbs phenomenon and its resolution. *SIAM review*, 39(4):644–668, 1997.
- [17] Jan Hesthaven and Robert Kirby. Filtering in Legendre spectral methods. *Mathematics of Computation*, 77(263):1425–1452, 2008.
- [18] Ilse CF Ipsen. Solution of least squares problems with randomized preconditioned normal equations. *arXiv preprint arXiv:2507.18466*, 2025.
- [19] Alex Kanevsky, Mark H Carpenter, and Jan S Hesthaven. Idempotent filtering in spectral and spectral element methods. *Journal of Computational Physics*, 220(1):41–58, 2006.
- [20] Avleen Kaur and SH Lui. Space-time spectral method for the Stokes problem. *Applied Numerical Mathematics*, 187:206–234, 2023.
- [21] Chandramali Piyasundara Wilegoda Liyanage. *Space-time Spectral Methods for PDEs on Irregular Domains*. PhD thesis, University of Manitoba, 2025.
- [22] SH Lui. Spectral domain embedding for elliptic PDEs in complex domains. *Journal of Computational and Applied Mathematics*, 225(2):541–557, 2009.
- [23] SH Lui. *Numerical analysis of partial differential equations*. John Wiley & Sons, 2012.
- [24] SH Lui. Legendre spectral collocation in space and time for PDEs. *Numerische Mathematik*, 136(1):75–99, 2017.
- [25] SH Lui and Sarah Nataj. Chebyshev spectral collocation in space and time for the heat equation. *Electronic Transactions on Numerical Analysis*, 52:295–319, 2020.
- [26] SH Lui and Sarah Nataj. Spectral collocation in space and time for linear PDEs. *Journal of Computational Physics*, 424:109843, 2021.

- [27] Willard L Miranker. A well-posed problem for the backward heat equation. *Proceedings of the American Mathematical Society*, 12(2):243–247, 1961.
- [28] Sheehan Olver and Alex Townsend. A fast and well-conditioned spectral method. *SIAM Review*, 55(3):462–489, 2013.
- [29] Thomas H Otway. *The Dirichlet problem for elliptic-hyperbolic equations of Keldysh type*, volume 2043. Springer, 2012.
- [30] Thomas H Otway. *Elliptic–Hyperbolic Partial Differential Equations: A Mini-Course in Geometric and Quasilinear Methods*. Springer, 2015.
- [31] Ouyuan Qin and Kuan Xu. Solving nonlinear ODEs with the ultraspherical spectral method. *IMA Journal of Numerical Analysis*, 44(6):3749–3779, 2024.
- [32] John Michael Rassias. Mixed-type partial differential equations with initial and boundary values in fluid mechanics. *Int. J. Appl. Math. Stat*, 13(J08):77–107, 2008.
- [33] Jie Shen, Tao Tang, and Li-Lian Wang. *Spectral methods: algorithms, analysis and applications*, volume 41. Springer Science & Business Media, 2011.
- [34] Hans De Sterck and Yunhui He. On the asymptotic linear convergence speed of Anderson acceleration, Nesterov acceleration, and nonlinear GMRES. *SIAM Journal on Scientific Computing*, 43(5):S21–S46, 2021.
- [35] Alex Townsend and Sheehan Olver. The automatic solution of partial differential equations using a global spectral method. *Journal of Computational Physics*, 299:106–123, 2015.
- [36] Lloyd N Trefethen. *Spectral methods in MATLAB*. SIAM, 2000.
- [37] Hervé Vandeven. Family of spectral filters for discontinuous problems. *Journal of Scientific Computing*, 6(2):159–192, 1991.
- [38] Dawei Wang, Yunhui He, and Hans De Sterck. On the asymptotic linear convergence speed of Anderson acceleration applied to ADMM. *Journal of Scientific Computing*, 88(2):38, 2021.
- [39] Dong Yan. *The Smooth Extension Embedding Methods for Free Boundary Problems*. PhD thesis, University of California, Irvine, 2024.

- [40] Paul Andries Zegeling and MWF van Spengler. A generalized midpoint-based boundary value method for unstable partial differential equations. *Journal of Computational and Applied Mathematics*, 424:115014, 2023.