EXPLICIT CHARACTERISTIC–BASED HIGH–RESOLUTION ALGORITHMS FOR HYPERBOLIC CONSERVATION LAWS WITH STIFF SOURCE TERMS

by

Mohit Arora

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Aerospace Engineering and Scientific Computing)
in The University of Michigan
1996

Doctoral Committee:

Professor Philip L. Roe, Chairman
Associate Professor Eduard Harabetian
Associate Professor Kenneth G. Powell
Professor Martin Sichel
Professor Bram van Leer

To my parents

ACKNOWLEDGEMENTS

I'd like to thank my Mom and Dad for being the best possible parents (being more like friends than parents), who have gone through a lot but have ensured a good education for us, and my Mom, for some excellent quotes[1].

And I am fortunate to have the most loving, protective and helpful brother possible, whom I have happily followed everywhere. For a while, it appeared that we may be going in different directions; however, our paths have converged again (which is very nice). Without his help, I would not even *be* here at Michigan today.

At Michigan, I owe the most to my advisor, Professor Phil Roe. His unique approach of "answering" each question of mine by at least ten questions of his own taught me the power of asking the right questions and it is from observing him that I have *really* learned how to think. He gave me complete freedom to explore all my ideas, and taught me to exploit his magic drawer concept[2]. Our discussions are always generously sprinkled with jokes and anecdotes, and I firmly believe that this "lighter fare" helped me cope with the pressures of graduate school.

I thank Professor Van Leer for the opportunity to TA a class while simultaneously taking it for credit. More importantly, for being an extremely valuable mentor, taking special care to introduce me to his colleagues. Further, it was he who provided the initial impetus for the work in Chapter V.

---

[1] Whatever you do, do well. Even if you become a crook, just make sure you're a good one.

[2] When one of my projects is going nowhere, I leave it (in the magic drawer) and work on a totally different project. When I return and start over, the answers "magically" jump out.

Thanks to Professors Sichel, Van Leer and Powell for inviting me to Michigan. I have never regretted my decision – hopefully, they have never done so either.

My sincere gratitude to Professors Roe, Van Leer, Sichel, Powell and Harabetian, for serving on my committee, for reading through my dissertation at very short notice, and for their valuable insights and comments.

Special thanks go, first, to Rosemary, who quickly became a very good friend and helped me tremendously in my first year; to Char, a very close friend whose company I have enjoyed very much, and to my colleagues in the CFD Lab for all the discussions and the friendly, productive and fun atmosphere. Thanks especially to Lisa, who was the friend I counted on most for advice large and small, and with whom I had many discussions over long lunches (since I eat incredibly fast).

Thanks also to Bill, my office–mate of many years, whose house I have often crashed at when I was *****, and who introduced me to the SASP, where I made many friends – to them all, I say thank you for a really fun time (specially Amy).

I also thank Jeff, Jim, Geoff and Laurent at Chippewa for all the tennis and fun, and to Brian, my badminton sparring partner, for great workouts.

Many thanks to Bev, Karen, Sheila and Adrienne at UMMC Volunteer Services, for making sure my volunteer experience was a very fun one.

Thanks to Anand for being a most supportive and helpful friend at IIT–K, and to CCS Reddy, my mentor at CAD–P, for teaching me all I know about programming.

There is surely not enough room to individually thank all the people who have touched my life and made it so enjoyable – but this big THANKS is for them.

TABLE OF CONTENTS

CHAPTER

LIST OF FIGURES

Figure

LIST OF APPENDICES

CHAPTER I

INTRODUCTION

Non–equilibrium flows are often described by inhomogeneous hyperbolic systems of conservation laws, where the source terms cause a "relaxation" towards equilibrium. Such systems describe a wide variety of physical phenomena, such as gas flows with relaxation[1] [37, 36, 8, 78, 20, 81, 82, 21, 65, 47, 48, 64, 42, 73, 5, 77, 12, 53, 52], multiphase flow[2] [32, 51, 28, 75] and phase transitions [68], turbulence [30, 45], water waves[3] [66], viscoelasticity[4] [60] and reactive flows[5] [9, 29, 50, 23]. Further, they govern models for the kinetic theory of gases [17, 14, 7, 11, 33], and applications of these methods to kinetic–based schemes for compressible flow [27, 58, 56, 57] and rarefied–gas flows (including MHD and re–entry hypersonics) [25, 10, 13].

As is clear, such systems form an extremely important class of problems, and are typically of the form

$$\mathbf{w}_t + \mathbf{f}_x = -\frac{1}{\tau}\mathbf{s} \quad , \tag{1.1}$$

where $\mathbf{w}$, $\mathbf{f}$, $\mathbf{s}$ are state, flux and source vectors respectively, and τ is the relaxation time of the problem (one can also look at a typical relaxation distance $\lambda = a\tau$, where

---

[1] including thermo–chemical nonequilibrium such as vibrational relaxation, dissociation and recombination effects

[2] which includes dusty gas flows

[3] where gravity balances river–bed friction

[4] memory effects

[5] including computational combustion

a is some characteristic velocity). This dissertation will be restricted to systems of conservation laws with relaxation in the sense of Whitham [78, 79] and Liu [49], in which an $N \times N$ hyperbolic system relaxes in the limit of $\tau \to 0$ to an $M \times M$ system ($M < N$). An additional restriction is that the equilibrium eigenvalues satisfy Liu's sub–characteristic condition [49, 79] (*i.e.*, the equilibrium eigenvalues interlace with the frozen ones). This is the stability condition [49, 79], a requirement to obtain decaying (rather than growing) solutions. Alternatively, this condition can be understood in terms of causality[6].

Whitham's work [78, 79] on the theory of linear relaxation systems has since been extended to non–linear systems, using a prototypical 2×2 system. The stability of this system (under the sub–characteristic condition) and that of its zero relaxation limit (the time–asymptotic solution[7]) were proven in [49] and [18], respectively.

Relaxation systems are characterized by dispersive wave behavior[8] and the presence of multiple scales in the problem. The flow starts out at the frozen limit ($t/\tau \to 0$) and *relaxes* to the equilibrium limit ($t/\tau \to \infty$). When at least one of these scales is much smaller than the others, the problem is said to be stiff.

It is usually impossible to separate physical problems into non–stiff and stiff regimes, especially for multiple relaxation times, mandating a solution to the full relaxation system in all cases. For example, if there are two relaxation times, a numerical time–step that fully resolves one of these relaxation times (not stiff) may be inadequate to resolve the other (stiff). If the goal is to obtain microscale features (reaction fronts, shock structures *etc*), then there is no substitute for resolution

---

[6] Initial disturbances travel with the propagation speeds of the full relaxation system, while later, they travel with those of the reduced system; hence, a natural requirement is that the equilibrium speeds be bounded by the frozen ones (for each pair), resulting in the interlace condition.

[7] The knowledge of the long–time asymptotics for the relaxation system have been recently incorporated into a very successful numerical algorithm [41, 38, 15].

[8] Dispersive waves are those whose wavespeeds are functions of the wavenumber.

(*i.e.*, time steps that fully resolve all time scales of the problem), and the resulting computation would be necessarily in the non–stiff regime. If on the other hand, macroscopic quantities are of primary interest, then it would be prudent to use an *under–resolved* method that resolves such quantities but not relaxation zones or layers. However, such under–resolved (stiff) methods, while potentially economical, are hard to compute.

There are several approaches one could take. The first is to use a fully explicit scheme, which has the advantages of simplicity and local data–dependencies (hence parallelizable); however, since stability constrains the time step to be smaller than the relaxation time, the computational cost is tremendous. The second option is to use a fully implicit scheme, which eliminates the time step restrictions. But implicitness causes global data–dependencies, rendering such schemes very hard to parallelize, potentially a severe disadvantage when solving for flows around complex geometries. Further, the gain in stability is partially offset by the loss of accuracy, a situation that does not arise in ODE's but does in the context of wave–propagation problems. A third, and frequently used alternative, is operator splitting [67], in which the hydrodynamic and non–equilibrium equations are solved alternately. While this method appears reasonable, it is probably not the best choice (simply because two physically simultaneous effects are artificially converted into sequential effects). In fact, it has been demonstrated in [55, 38] that such fractional–step methods reduce from second– to first–order accuracy in the fluid dynamic limit ($\tau \to 0$), since the numerical approximation becomes an $O(\Delta t)$ approximation to the equilibrium equation [38]. It was also shown there that neither higher–order Runge–Kutta nor higher–order reconstruction alleviates this problem of reduced accuracy for split operator methods. Thus, it is not the stability condition alone that makes this problem hard.

Since each of the above options has advantages as well as drawbacks, it is quite likely that the ideal scheme would have a hybrid character – being local (hence parallelizable) and yet not having the time step constrained by stability to be smaller than the relaxation time. A promising variant that has such properties is the point–implicit method [83], where the source term is treated implicitly (locally).

Studies focusing on numerical algorithm development for relaxation systems have been many [31, 54, 55, 71], but only recently have good results been obtained for the difficult fluid dynamic limit by Jin and Levermore [41], Jin [38] and Caflisch, Jin and Russo [15]. It was shown there that poor results are achieved for under–resolved methods if the physical asymptotics are not matched by the numerical schemes, *e.g.*, in stiff cases, the numerical dissipation far overwhelms the physical dissipation. Further, some schemes may even converge to non–physical solutions [38, 15]. Hence, in [41], the physical asymptotics governing the relaxation systems are incorporated into the numerical algorithm in a semi–discrete manner[9]. This is extended to the fully–discrete case using Runge–Kutta methods coupled with a good integrator for the stiff ODE's, employing a method of lines approach [38, 15]. While the results obtained by this approach are very good, the method development is quite involved, requires some *a priori* knowledge of the asymptotics and appears to be computationally expensive.

So what are the goals of this dissertation? The first aim is to establish a physically–based systematic procedure to solve stiff relaxation systems. In contrast to earlier work, the approach will be to devise fully–discrete, second–order accurate schemes, that are reasonably accurate for a large range of $\Delta t/\tau$, $\Delta x/\lambda$, with results allowed to degrade gracefully as the stiffness (under–resolution) increases. How-

---

[9] This is not the first instance of using the asymptotics in numerical schemes, these having been applied in [44, 39, 43, 40].

ever, a natural requirement is that the method result in physical solutions (captured shocks with correct propagation speeds and jumps), even for severely under–resolved cases. The non–relaxable (pun intended) design constraints are to keep the scheme simple, economical and parallelizable. Additionally, it would be desirable to have stability governed by the CFL[10] condition based on the homogeneous system, and sharp captured shocks.

The first step was to select an analytically tractable linear model that exhibited most of the difficulties of realistic stiff relaxation systems. It was hoped that this model would provide valuable analytical insight into such systems, perhaps also clarifying the advantages and disadvantages of some of the current solution methodologies. To start with, the study was restricted to a shock–fitting framework with a unity CFL number (Chapter II). This helped isolate errors related to discretization and highlighted the fundamental problems associated with stiff source terms. Chapter III describes how the method development of Chapter II was extended to non–characteristic meshes and shock capturing, using one of the most successful methods for solving homogeneous non–linear conservation laws – the finite–volume method. However, at this stage, the problem was still restricted to the linear case, and an attempt was made to benchmark the results. Once satisfactory results were achieved, the method was modified to account for non–linear effects (Chapter IV); several test cases were attempted, which helped put the present work in perspective. In Chapter V, a new method for the solution of viscous conservation laws is proposed, that results from an inversion of the analysis performed and algorithms developed in Chapters II to IV. More realistic relaxation systems are solved in Chapter VI, followed by some concluding remarks and suggestions for future work (Chapter VII).

---

[10] The CFL, or Courant–Freidrichs–Levy, condition is the stability condition – physically, it demands that the numerical domain of dependence include the entire physical domain of dependence.

CHAPTER II

A MODEL FOR DISPERSIVE WAVES – THE HYPERBOLIC HEAT EQUATIONS

The first step towards establishing a systematic methodology for solving relaxation systems was the selection of an appropriate model problem that exhibited the difficulties of realistic systems of this type and yet was simple enough to permit a detailed analysis. The aim during this stage was merely to determine whether methods currently in use were adequate. If they weren't, to determine what modifications (if any) would work, or else to design a new method to solve this class of problems. Besides giving rise to dispersive wave behavior, the source terms present may also be stiff[1], which leads to a trade–off between spatial/temporal under–resolution and prohibitive computational expense.

There are several approaches one could take as discussed in Chapter I. However, all the schemes appear to represent tradeoffs, and this work represents a systematic quest for a method that combines all the advantages without carrying around the disadvantages. In the past, there have been several attempts, based on physical principles or mathematical analysis of (often scalar) model problems, to incorporate source terms into numerical procedures. However, the behavior uncovered in this

---

[1] Stiffness arises when the equilibration time of a non–equilibrium process is much smaller than the flow residence time.

chapter – that of the coupling between different wave families caused by the source term – can never be captured by scalar models. A brief description of the chosen model – the Hyperbolic Heat Equations [62, 63] (also referred to as the Maxwell–Cattaneo equations) – is given, followed by the design of numerical methods for solving this system in the smooth region alone (*i.e.*, shocks are not being captured in this chapter). Further, in order to isolate errors due to the source alone, only characteristic meshes are considered (*i.e.*, all the cases are computed for a unity CFL number based on the frozen wavespeed).

A brief derivation of the model, a linear 2×2 relaxation system, and its analytical properties, is given in Section 2.1. A Riemann problem is presented in Section 2.2, for which an exact solution exists (in closed form) for the heat flux in the smooth region between the characteristics and for the exponential decay of the discontinuities involved. The two simplest discretizations are shown in Section 2.3; however, counter–intuitive numerical results obtained for these schemes motivated attempts to improve on them. Lack of correlation between expectations of further experiments and numerical results (Section 2.3.1), based on a large number of experiments, suggested a fresh start using additional analytical information. In the process of devising an optimal higher–order scheme, the reasons for the poor performance of the pure method of characteristics have been uncovered (Section 2.4). Finally, more practical higher–order schemes are presented (Section 2.5), wherein the key features of the optimal scheme are incorporated in a much simpler manner, and Section 2.6 provides the reader with a brief summary of this chapter.

The work presented in this chapter has appeared in [62, 63]; while a comprehensive overview follows, the interested reader is referred to [63] for more details.

2.1 The Hyperbolic Heat Equations

Consider the flow of heat in a uniform conducting bar. Conservation of energy may be stated as

$$\theta_t + q_x = 0 \quad ,$$

where θ, q are the temperature and heat flux respectively, and units have been chosen to give a heat capacity of unity. Usually, one would now invoke Fourier's law $q = -\theta_x$ (assuming units where the proportionality constant, the heat conductivity, is unity), which would lead to the heat equation

$$\theta_t = \theta_{xx} \quad .$$

However, this gives the paradoxical result of an infinite propagation speed for a diffusion problem. A convenient way around this is to assume a relaxation model [16, 72, 47], the simplest version of which is

$$\tau q_t + \theta_x = -q \quad ,$$

and where τ is a relaxation time. This leads to the Hyperbolic Heat Equations (HHE's), which are[2]

$$\begin{aligned} \theta_t + q_x &= 0 \quad , & (2.3) \\ q_t + \frac{1}{\tau}\theta_x &= -\frac{q}{\tau} \quad . & (2.4) \end{aligned}$$

---

[2] The general form of the HHE's is

$$\begin{aligned} u_t + v_x &= 0 \quad , & (2.1) \\ v_t + a_F^2 u_x &= -\frac{1}{\tau}(v - a_E u) \quad , & (2.2) \end{aligned}$$

where a_F (a_E) are the *frozen* (*equilibrium*) wavespeeds [at which high (low) frequency waves travel]. In Section 5.1, a generalization of this system is analyzed, as well as its nonlinear analogue.

2.1.1 Dispersion

The dispersive character of these equations may be seen by considering solutions of the form

$$\begin{pmatrix} \theta \\ q \end{pmatrix} = \mathcal{R}\left[\begin{pmatrix} T \\ Q \end{pmatrix} \exp[i(\omega t - \xi x)]\right] \quad . \tag{2.5}$$

which leads to the dispersion relationship

$$\tau\omega^2 - \xi^2 = i\omega \quad . \tag{2.6}$$

For an initial value problem (which is of interest here), ξ is a real wave number and ω is a complex frequency ($\omega = \omega_R + I\omega_I$, $I^2 = -1$), whose real part is a frequency and imaginary part is a damping rate. These relations can be easily solved to obtain

$$\frac{\omega_R}{\xi} = \tau^{-\frac{1}{2}}\left(1 - \frac{1}{4\tau\xi^2}\right)^{\frac{1}{2}} \equiv a(\xi) \quad , \tag{2.7}$$

$$\omega_I\tau = \frac{1}{2}\left[1 \pm \left(1 - 4\xi^2\tau\right)^{\frac{1}{2}}\right] \quad , \tag{2.8}$$

where $a(\xi)$ is the wave speed and $\omega_I\tau$ is the damping rate (the waves being damped at a rate $e^{-\omega_I\tau}$). In Figure 2.1, the wavespeed $a(\xi)$ [upper] and the damping rate $e^{-\omega_I\tau}$ [lower] are plotted against the non-dimensional wavenumber $\tau^{\frac{1}{2}}\xi$.

For very high wavenumbers ξ, the propagation speed is the frozen characteristic speed $\tau^{-\frac{1}{2}}$. For decreasing wavenumbers, the propagation speed is reduced, becoming zero when $\xi = \frac{1}{2}\tau^{-\frac{1}{2}}$, *i.e.*, a vanishing equilibrium wavespeed at this wavenumber (called the bifurcation point). For all wavenumbers in the range $[\frac{1}{2}\tau^{-\frac{1}{2}}, \infty]$, the waves are damped like $e^{-t/2\tau}$.

When $\xi = 0$, the problem reduces to $\theta_t = 0$, $\tau q_t + q = 0$ [hence, $\theta(x)$ is an arbitrary function of x alone, and $q = q(x)e^{-t/\tau}$]. Since these have solutions corresponding to $\omega_I\tau = 0$, 1 respectively, both branches of Equation 2.8 are relevant. The upper

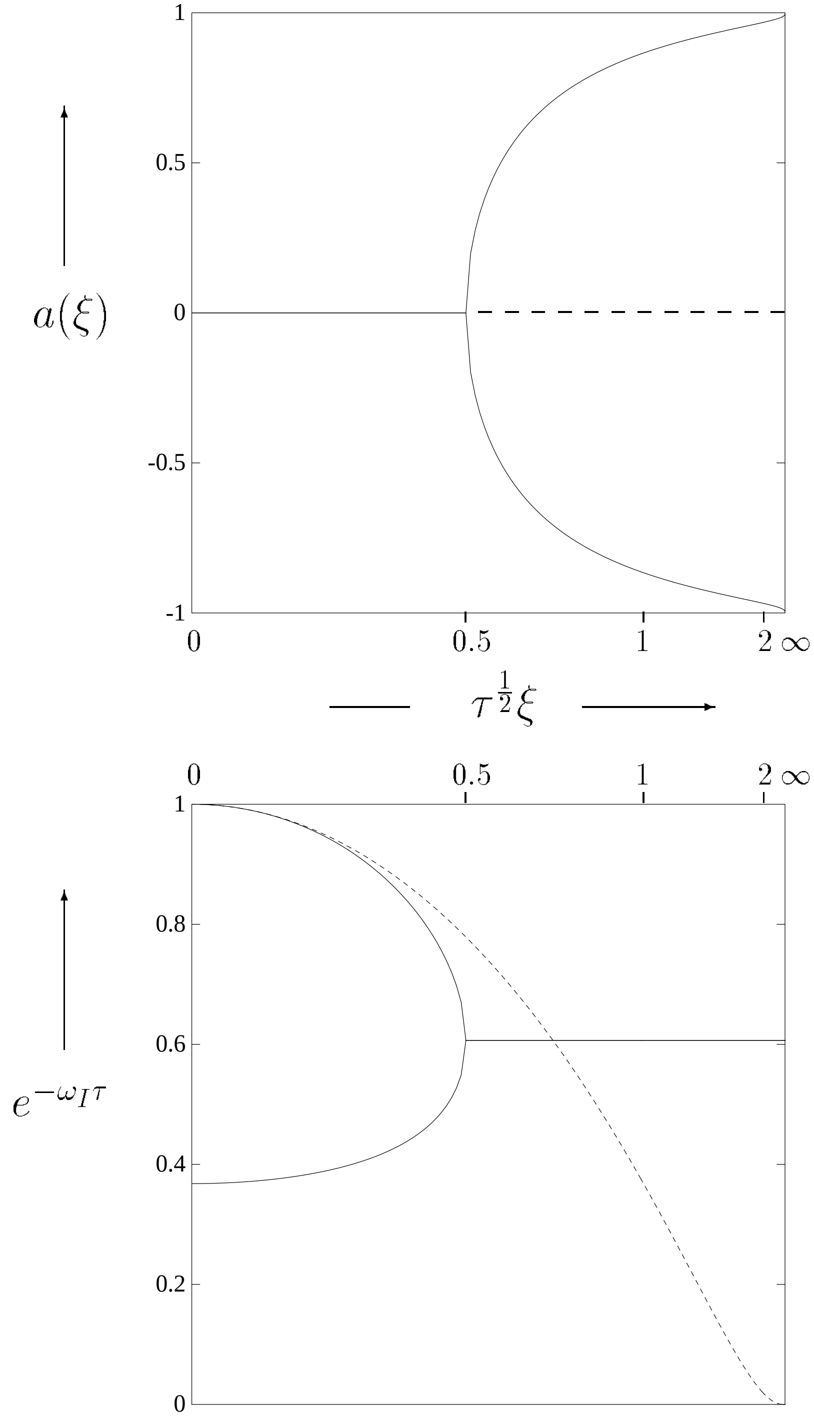


Figure 2.1: Analytic Dispersion Diagrams [wavespeed (Equation 2.7) and damping rate (Equation 2.8)] for the Hyperbolic (solid lines) and Parabolic (dashed lines) Heat Equations.

branch makes second-order contact with the dispersion relationship for the regular heat equation, which is

$$\omega_R = 0 \quad ; \quad \omega_I \tau = \xi^2 \tau \quad , \tag{2.9}$$

shown as dotted lines in Figure 2.1.

2.1.2 Characteristic Equations

Introducing characteristic coordinates

$$\zeta = t + \tau^{\frac{1}{2}} x \quad ; \quad \eta = t - \tau^{\frac{1}{2}} x \quad , \tag{2.10}$$

transforms the HHE's to

$$\theta_\zeta + \tau^{\frac{1}{2}} \left(q_\zeta + \frac{q}{2\tau} \right) = 0 \qquad \text{along} \qquad \frac{dx}{dt} = \tau^{-1/2} \quad , \tag{2.11}$$

$$\theta_\eta - \tau^{\frac{1}{2}} \left(q_\eta + \frac{q}{2\tau} \right) = 0 \qquad \text{along} \qquad \frac{dx}{dt} = -\tau^{-1/2} \quad , \tag{2.12}$$

which are the characteristic equations for the HHE's. Unfortunately, it is not possible to integrate these equations and obtain Riemann invariants, as can be done with linear homogeneous systems. Thus, a numerical method of characteristics is no longer an exact method.

As usual, the solution admits discontinuities that lie along characteristic paths, and it is easy to show that the jump relationships are those of the homogeneous problem (*i.e.*, across a jump in the ζ-direction, $\Delta\theta = \tau^{\frac{1}{2}} \Delta q$, while across a jump in the η-direction, $\Delta\theta = -\tau^{\frac{1}{2}} \Delta q$).

2.2 A Riemann Problem and its Exact Solution

A natural problem to pose in connection with the hyperbolic model of heat conduction is that of two semi-infinite rods, having temperatures θ_L, θ_R, brought into

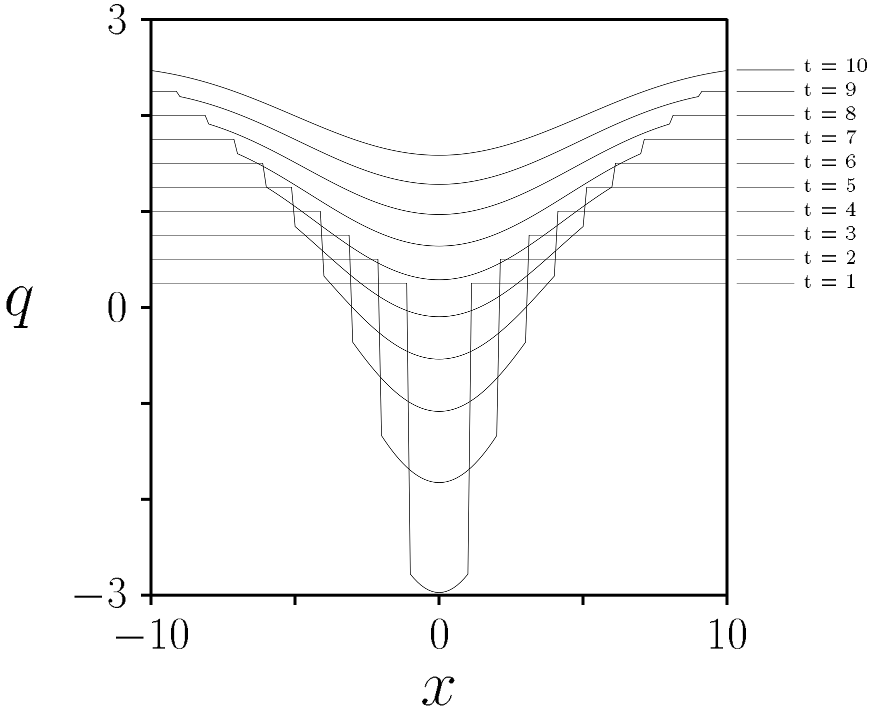


Figure 2.2: Exact solution to the Riemann Problem for the heat flow q, with the smooth region given by Equation 2.13, shown for several times. The initial condition is a jump in θ and $q = 0$.

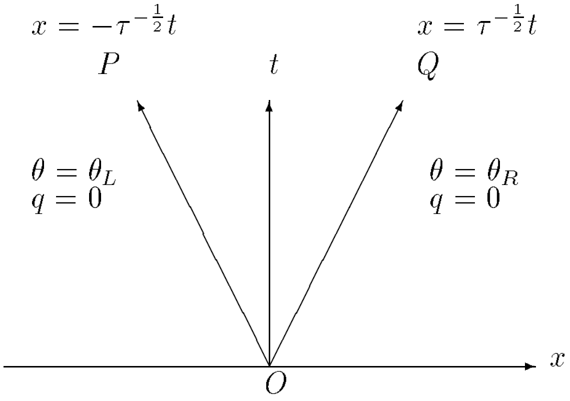


Figure 2.3: Schematic of the Riemann problem of Section 2.2, where OP and OQ are the frozen characteristics along which the (decaying) jumps propagate, and POQ is the region in which a solution is sought.

contact at $t = 0$. The analytical solution is shown in Figure 2.2, where the exponential decay of the jumps along the characteristics is clearly observed.

The solution for q in the region POQ (see Figure 2.3), is [63, 74]

$$\begin{aligned} q(\zeta, \eta) &= \left(\frac{\theta_L - \theta_R}{2\tau^{\frac{1}{2}}}\right) \exp\left(-\frac{\zeta + \eta}{4\tau}\right) I_0\left(\sqrt{\frac{\zeta\eta}{4\tau^2}}\right) \\ &= \left(\frac{\theta_L - \theta_R}{2\tau^{\frac{1}{2}}}\right) \exp\left(-\frac{t}{2\tau}\right) I_0\left(\sqrt{\frac{t^2 - \tau x^2}{4\tau^2}}\right) \quad , \end{aligned} \tag{2.13}$$

where I_0 is the modified Bessel function of order zero. Note that the discontinuities decay at a rate $e^{-t/2\tau}$.

2.3 Simple Discretizations and Numerical Results

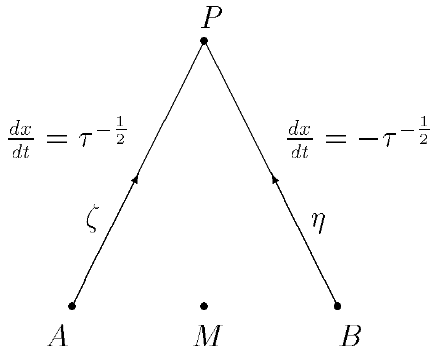


Figure 2.4: Stencil for the Method of Characteristics, where AP, BP are the characteristics and P is the point being computed. Note that no information from the point between A, B is used, resulting in odd–even *decoupling.*

The two simplest discretizations are the Point Implicit (PI) and the Operator Splitting (OPS) methods. Based on the characteristic stencil shown in Figure 2.4, both result in solutions of the form

$$\theta_P = \left(\frac{\theta_A + \theta_B}{2}\right) + \frac{\tau^{\frac{1}{2}}}{2} X(k)(q_A - q_B) \quad , \tag{2.14}$$

$$q_P = Z(k)\left(\frac{q_A + q_B}{2}\right) + \frac{\tau^{-\frac{1}{2}}}{2} Y(k)(\theta_A - \theta_B) \quad , \tag{2.15}$$

where the functional parameter k is the stiffness factor defined by

□— — —□	$t/\tau = 3$
◷ - - - - ◷	$t/\tau = 6$
△— - — △	$t/\tau = 10$
+— — —+	$t/\tau = 30$
× - - - - ×	$t/\tau = 60$
◇— - — ◇	$t/\tau = 100$
↑— - - —↑	$t/\tau = 300$
⧖--------- ⧖	$t/\tau = 600$
⋝— — —⋝	$t/\tau = 1000$

Table 2.1: This table describes the symbols and line styles used for plotting numerical results. They correspond to different times at which the solution errors in q and θ have been plotted versus the stiffness factor k.

$$k = \frac{1}{2}\frac{\Delta t}{\tau} \quad . \tag{2.16}$$

To obtain the PI scheme, the Equations 2.11 and 2.12 are integrated along the characteristics. The source term is evaluated along each characteristic separately, by taking the value halfway between the ends of the characteristic. These midpoint values are obtained via a linear interpolation, *i.e.*,

$$q_{\zeta,\,\eta} = \frac{1}{2}\left(q_P + q_{A,B}\right) \quad ,$$

which leads to coefficients

$$X = 1 - k \quad ; \quad Y = (1+k)^{-1} \quad ; \quad Z = (1-k)(1+k)^{-1} \quad . \tag{2.17}$$

This scheme has been applied to practical problems in [80], although $k < 1$ for all cases there (not stiff). For $k > 1$, however, the coefficient Z does not look promising.

On the other hand, the operator splitting method circumvents the problem of the source term by splitting the solution procedure into two parts – a damping stage (L_1) due to the source term alone, followed by a solution of the homogeneous problem (L_2), using the sequence of operations L_1 L_2 L_2 L_1 [67] for second–order accuracy[3]. This leads to coefficients

$$X = e^{-k} \quad ; \quad Y = e^{-k} \quad ; \quad Z = e^{-2k} \quad , \tag{2.18}$$

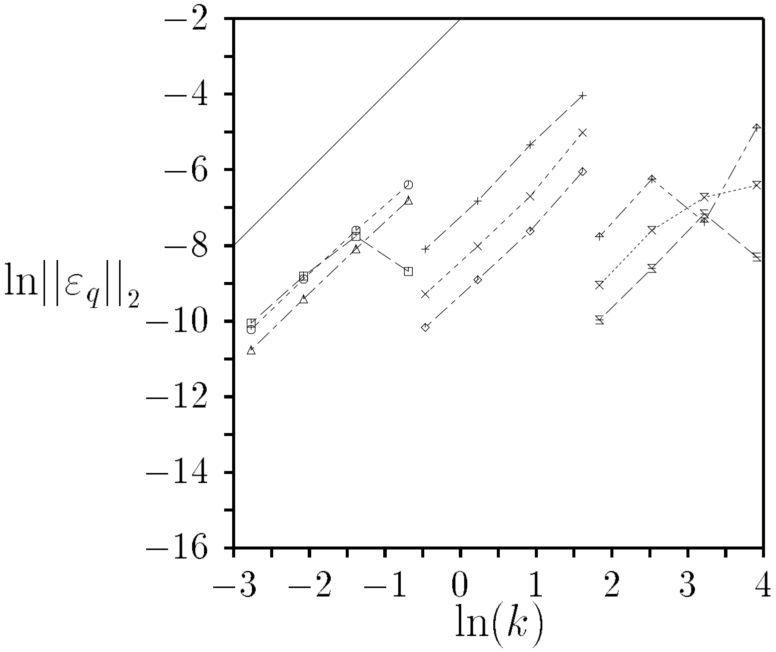
which look far more promising, because the coefficients appear to reflect appropriate decay rates. Indeed, $Z(k) = e^{-2k}$ corresponds to the decay rate of the ordinary differential equation, and $X(k) = Y(k) = e^{-k}$ corresponds to the decay rate of the characteristic discontinuities (Equation 2.13, [79][4]). However, contrary to this reasoning, observe the results in Figures 2.5 and 2.6, where the PI scheme clearly outperforms the OPS scheme by up to 4 orders of magnitude when k is large. Note that both schemes are second–order accurate (and roughly equivalent) if k is small. For k of order unity, the PI scheme is still second–order. For k large, the PI scheme is a little less than second–order accurate, but the OPS scheme degrades to O(1) (similar to the observations in [55, 38]).

Next, several attempts were made to improve on these methods, based on rational constraints. One which was thought to be important was "conservation" (in the sense described in Appendix A). The PI scheme, although it convincingly out-

---

[3] While L_2 L_1 L_1 L_2 is equivalent for second–order accuracy [67], it has been shown in [38] that the L_1 L_2 L_2 L_1 ordering is more appropriate when solving relaxation systems. This is because use of an under–resolved method (where spatial or temporal scales are such that $\Delta x > \lambda$ or $\Delta t > \tau$) does not resolve the initial layer. The effect of this layer is to project the data onto a local equilibrium. Since this layer is not resolved by the numerical scheme, its effect much somehow be built into the algorithm. This is important, because when such a step is missing, numerical methods can result in spurious solutions. An implicit step projects the initial data onto a local equilibrium [38]; thus, the damping step must be the first one in split–operator methods. Note, however, that the results of [55, 38] show that Strang splitting [67] reverts to first order accuracy in the fluid dynamic limit ($\tau \rightarrow 0$), as the scheme becomes an $O(\Delta t)$ approximation to the equilibrium system.

[4] In the general (linear) case, the decay rate of the discontinuities is $e^{-\alpha k}$, where α is a constant dependent on the frozen and equilibrium wavespeeds.



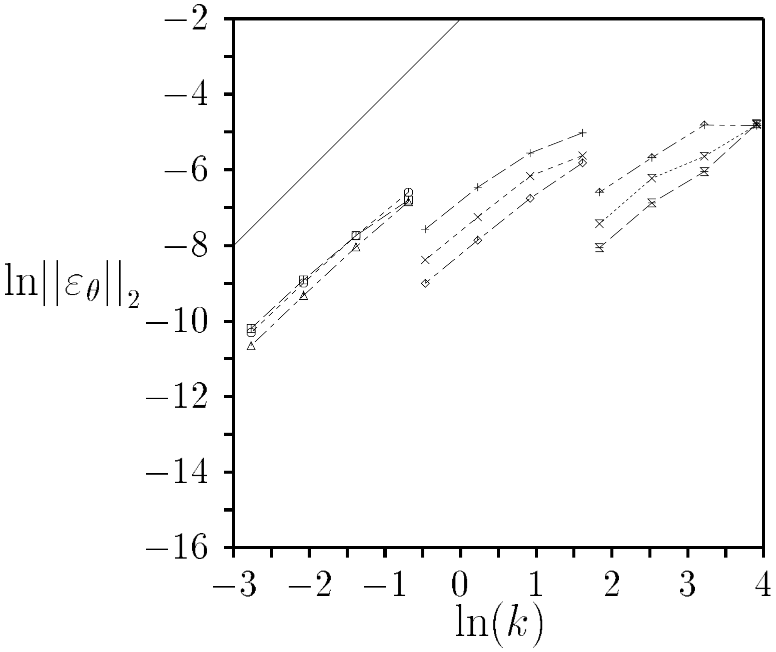


Figure 2.5: L_2 norm of the solution error in q and θ vs the stiffness factor k for the PI (Point Implicit) scheme (see Table 2.1 for a key to lines and symbols). The solid line has a slope of 2.

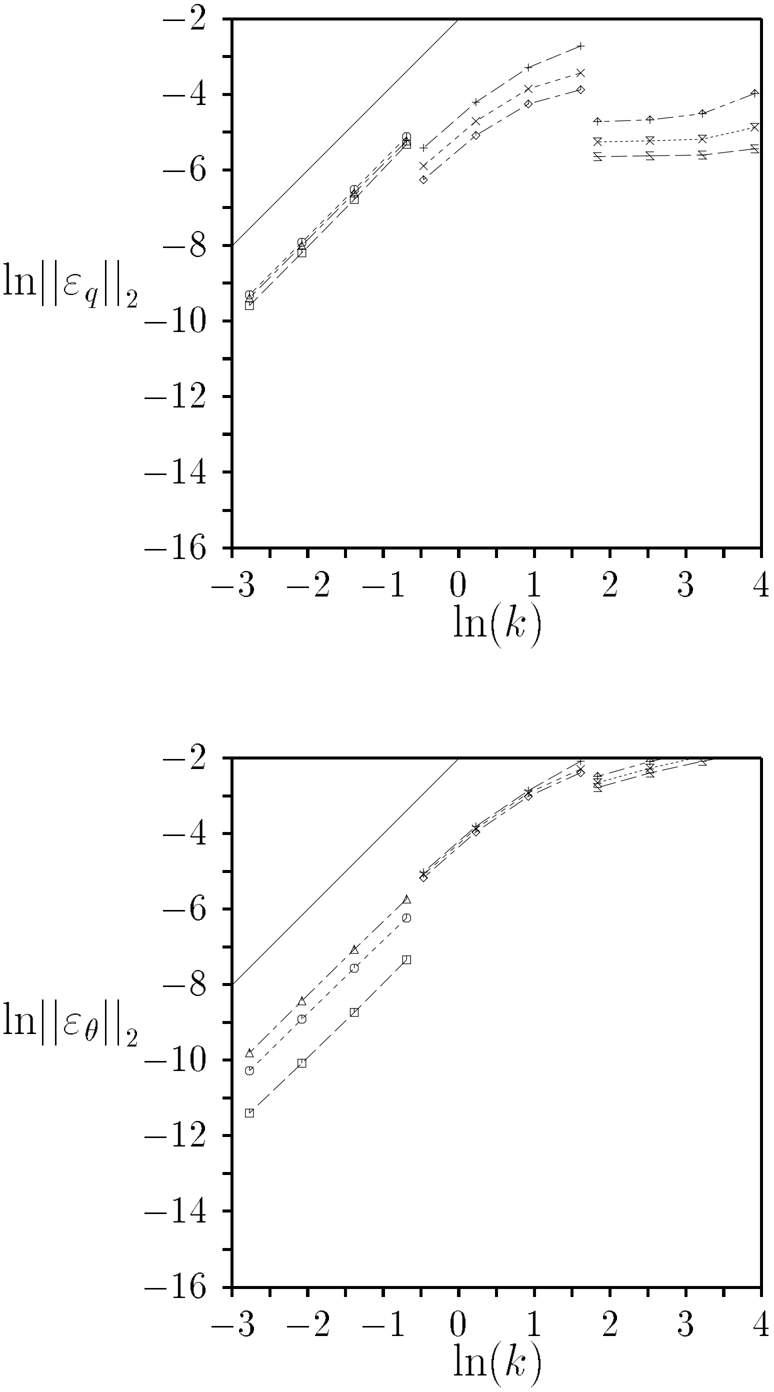


Figure 2.6: L_2 norm of the solution error in q and θ vs the stiffness factor k for the OPS (Operator Splitting) scheme (see Table 2.1 for a key to lines and symbols). The solid line has a slope of 2.

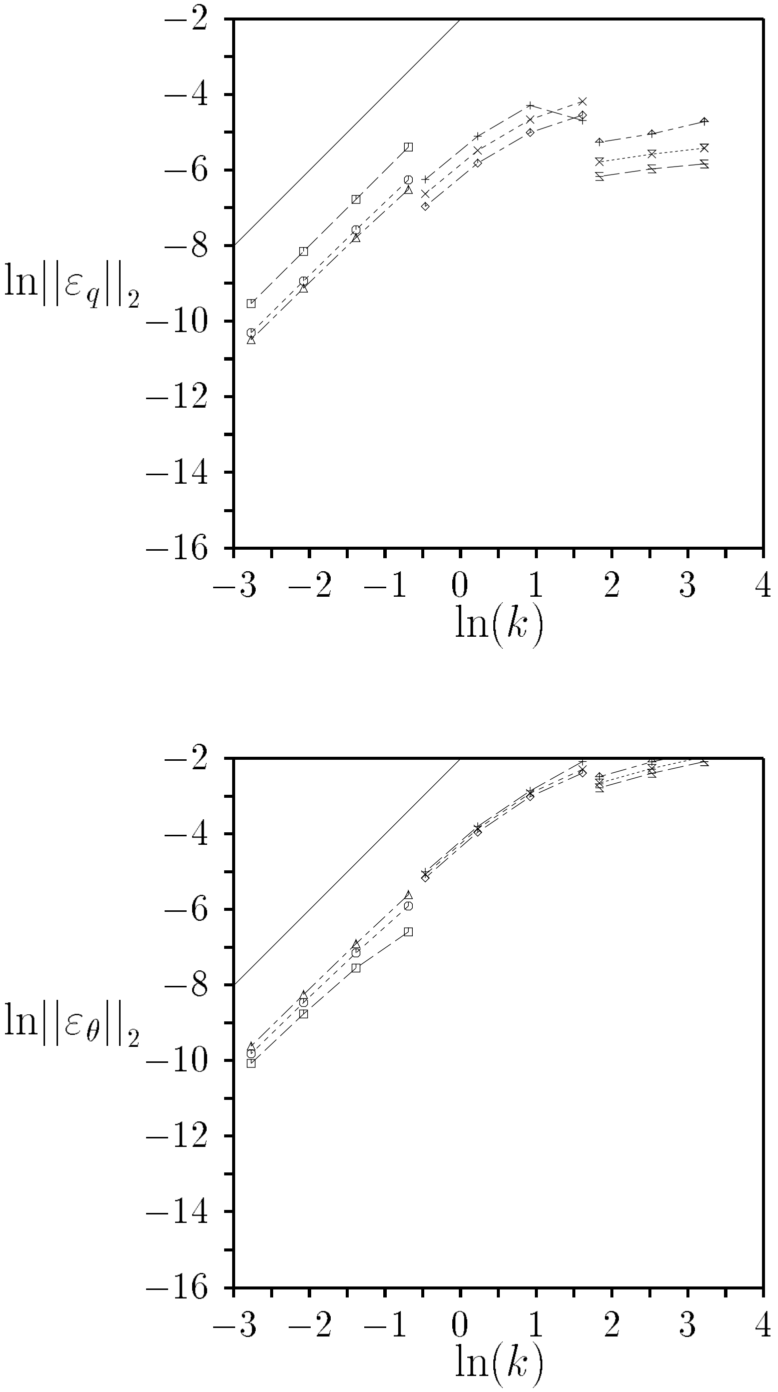


Figure 2.7: L_2 norm of the solution error in q and θ vs the stiffness factor k for the PI_μ (Point Implicit with implicitness factor) scheme (see Table 2.1 for a key to lines and symbols). The solid line has a slope of 2.

performs the OPS scheme, does not have this property. Is a modification to the PI scheme possible that would satisfy Equation A.1? This line of inquiry led to a generalization of the PI scheme, which will be denoted as the PI_μ scheme, where μ is the degree of implicitness[5]. Again, Equations 2.11 and 2.12 are integrated along the characteristics. However, the source term is now approximated as[6]

$$q_{\zeta,\ \eta} = \mu q_P + (1-\mu) q_{A,B} \quad ,$$

which leads to

$$X = 1 - 2k(1-\mu) \quad ; \quad Y = (1+2k\mu)^{-1} \quad ; \quad Z = \left[\frac{1-2k(1-\mu)}{1+2k\mu}\right] \quad . \tag{2.19}$$

The choice of μ remains. This is determined by enforcing "conservation" and is derived in Appendix A to be

$$\mu = \frac{1}{1-e^{-2k}} - \frac{1}{2k} \quad . \tag{2.20}$$

Note that in the non–stiff limit ($k \to 0$), $\mu = \frac{1}{2}$ (which is simply the PI scheme). However, in the limit of $k \to \infty, \mu \uparrow 1$ (which is fully implicit). Numerical results for this scheme are shown in Figure 2.7. Surprisingly, the PI scheme performs better than the PI_μ scheme[7], although the latter is somewhat better than the OPS scheme, and approaches the OPS scheme for large k.

Based on numerical experiments, the success or failure of these methods seemed to correlate with the discrete dispersion relationships, which are discussed next.

---

[5] By degree of implicitness, we mean that the scheme has a hybrid character, being partly explicit and partly implicit. In one limit ($\mu = 0$), it is fully explicit while in the other ($\mu = 1$) it is fully implicit. For all $0 < \mu < 1$, it has some "degree" of implicitness. Note that in the context of the method of characteristics, the work required to solve the pair of characteristic equations remains constant (in the linear case).

[6] The source term is being evaluated at some point $t_n + \mu\Delta t$, where $\mu \in [0,1]$ along the characteristics. Since μ is normalized, it can be considered to represent a point along the time axis, since the normalized distance along the characteristic is the same as that along the time axis.

[7] It will be shown in later chapters that conservation becomes far more important in the context of under–resolved methods for nonlinear relaxation systems.

2.3.1 Discrete Dispersion Relationships

Writing Equation 2.5 in discrete form as

$$\begin{pmatrix} \theta \\ q \end{pmatrix}_j^n = \mathcal{R}\left[\begin{pmatrix} T \\ Q \end{pmatrix} \exp[i(\omega n\Delta t - \xi j \Delta x)]\right] \quad , \tag{2.21}$$

and using Equations 2.14 and 2.15 gives

$$\det \begin{vmatrix} \exp(i\omega\Delta t) - \cos(\xi\Delta x) & -i\tau^{\frac{1}{2}}X(k)\sin(\xi\Delta x) \\ -i\tau^{-\frac{1}{2}}Y(k)\sin(\xi\Delta x) & \exp(i\omega\Delta t) - Z(k)\cos(\xi\Delta x) \end{vmatrix} = 0 \quad ,$$

which can be simplified to give

$$\exp(i\omega\Delta t) = \frac{(1+Z(k))\cos(\xi\Delta x) \pm [(1-Z(k))^2\cos^2(\xi\Delta x) - 4X(k)Y(k)\sin^2(\xi\Delta x)]^{\frac{1}{2}}}{2} \quad , \tag{2.22}$$

which is the discrete dispersion relationship for the method of characteristics. Note that only those wavenumbers for which $\xi\Delta x \leq \frac{1}{2}\pi$ can be resolved. The factor $\frac{1}{2}$ arises because the stencil for the method of characteristics decouples odd and even points (see Figure 2.4). For this maximum frequency,

$$\exp(i\omega\Delta t) = \pm i[X(k)Y(k)]^{\frac{1}{2}}$$

and hence, if XY is positive,

$$\omega_R\Delta t = \pm\frac{1}{2}\pi \quad ,$$

which gives the wavespeed $a(\xi_{max}) = \pm\frac{\Delta x}{\Delta t} = \pm\tau^{-\frac{1}{2}}$. If XY is negative, then

$$\omega_R = 0 \quad .$$

Thus, for any method of characteristics, the highest wavenumber observable on the mesh is either stationary or propagated at the frozen speed. Figure 2.8 shows the Discrete Dispersion plots for the PI scheme for several values of k. Each plot is

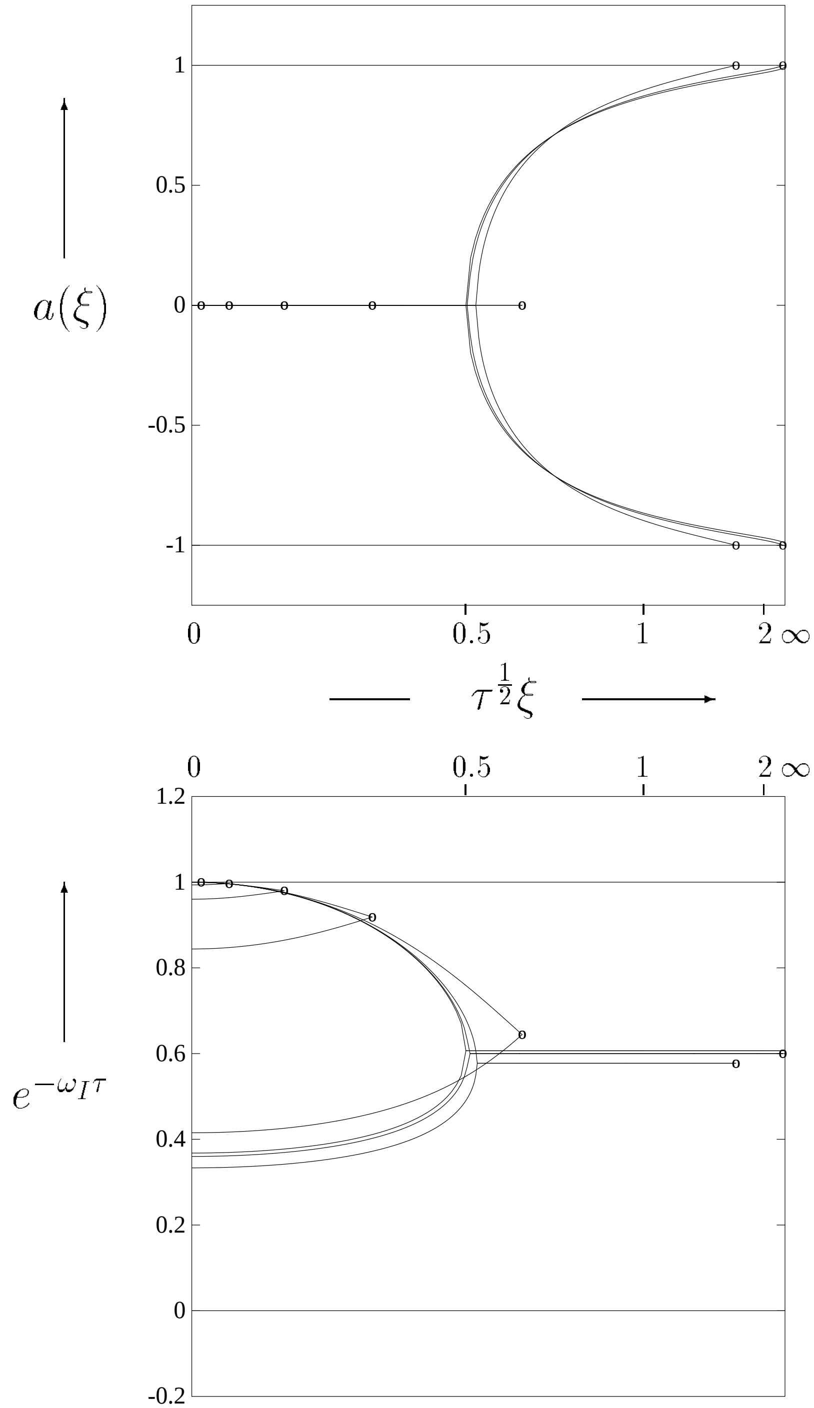


Figure 2.8: Propagating speeds and damping ratios for the PI (Point Implicit) method. $k = 0.25,\ 0.5,\ 1.25,\ 2.5,\ 5.0,\ 12.5,\ 50$

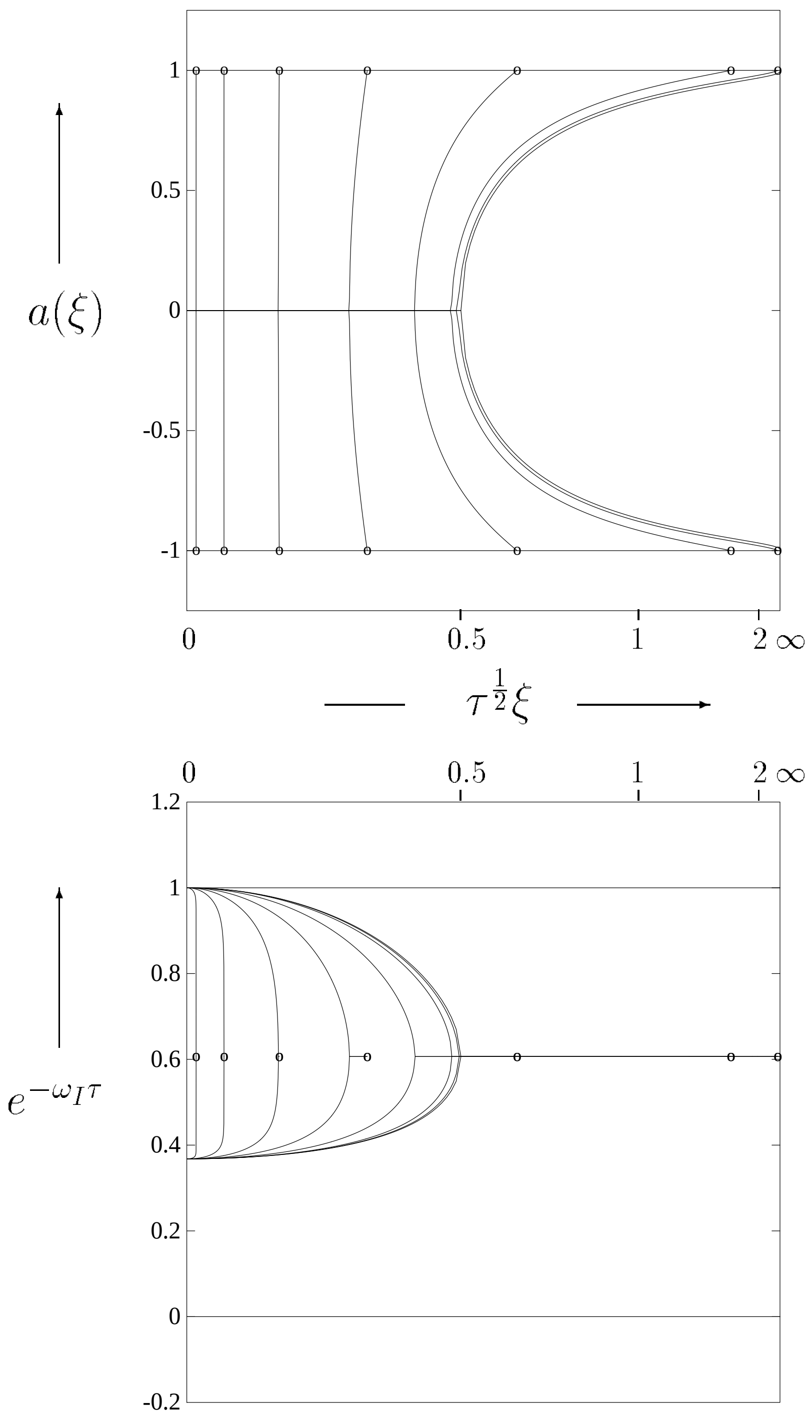


Figure 2.9: Propagating speeds and damping ratios for the OPS (Operator Split) scheme. $k = 0.25,\ 0.5,\ 1.25,\ 2.5,\ 5.0,\ 12.5,\ 50$

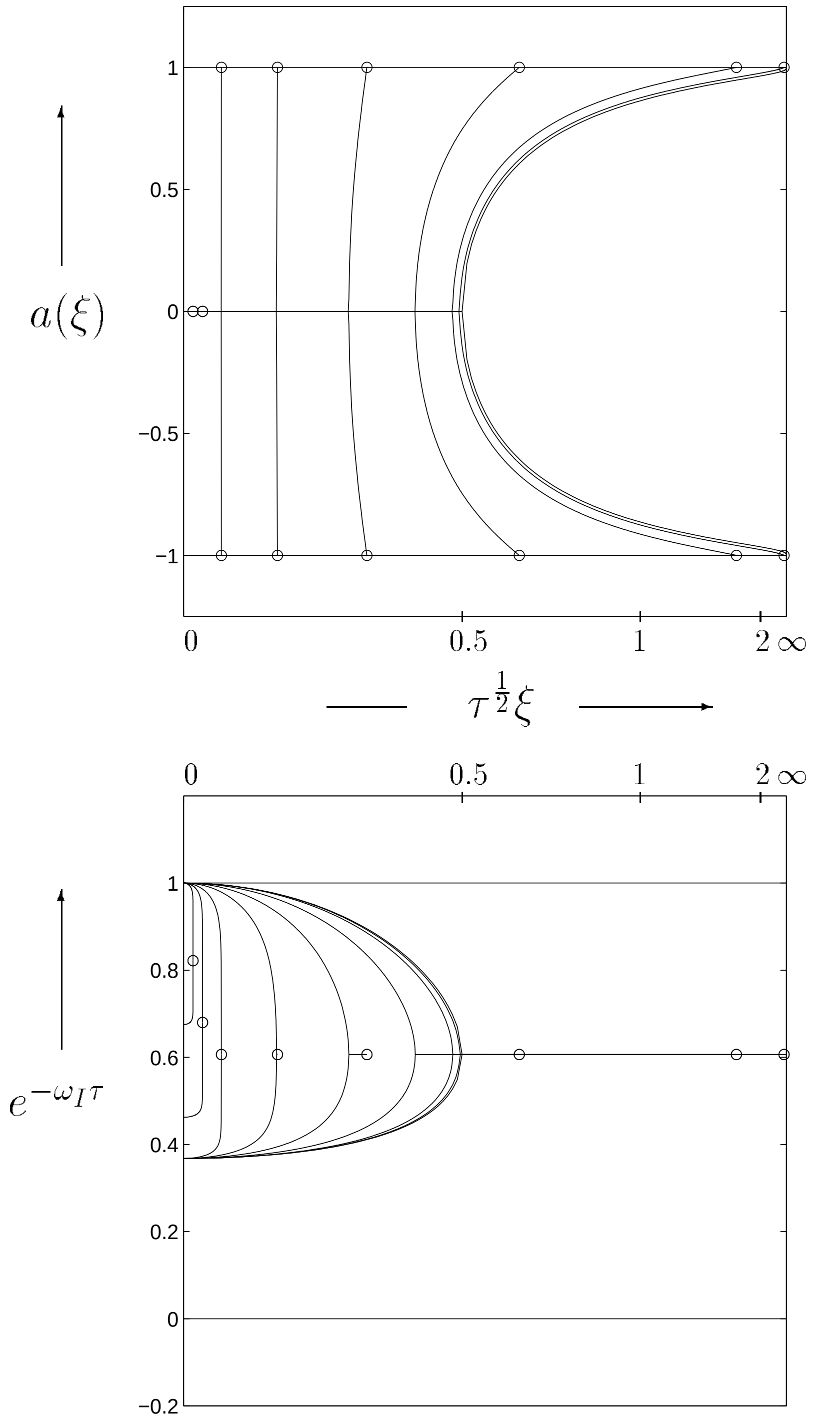


Figure 2.10: Propagating speeds and damping ratios for the PI$_\mu$ scheme. $k = 0.25$, 0.5, 1.25, 2.5, 5.0, 12.5, 25, 50

terminated at the right by a symbol located at the maximum wavenumber for that value of k, derived from

$$\begin{aligned} \frac{1}{2}\pi &= \xi_{max}\Delta x \\ &= \xi_{max}\frac{\Delta t}{\tau^{\frac{1}{2}}} \\ &= 2k\tau^{\frac{1}{2}}\xi_{max} \quad , \\ \tau^{\frac{1}{2}}\xi_{max} &= \frac{\pi}{4k} \quad . \end{aligned}$$

In the upper plot of Figure 2.8 (wavespeed), it is seen that these symbols lie either on the axis or on the upper and lower limits as predicted. The diagrams as a whole closely follow the analytical behavior. In the lower plot (damping) there is good agreement only for the upper lobe at large k but moderate agreement everywhere else.

In Figure 2.9, which shows the analogous plots for the OPS scheme, the situation is reversed. For small k, damping levels are very good; however, as k increases, the high wavenumbers are rather heavily damped, which may manifest itself as dissipation in a numerical scheme. But the wavespeeds are poorly approximated by this scheme, with stationary waves being incorrectly propagated at the grid speed $\Delta x/\Delta t$. Given this erroneous wave propagation, it is natural to expect these waves to be highly damped, as is clearly observable from Figure 2.9.

Interestingly enough, the PI_μ scheme follows in the footsteps of the OPS scheme for moderate stiffness (Figure 2.10), propagating all waves (including stationary ones) at the grid speed while damping them quite heavily. However, for large stiffness, it does capture the stationary wave behavior and reduces the damping to modest levels. Thus, the effects of incorrect wave propagation and damping levels are quite closely linked.

The problem of false wave propagation can be dealt with, to a large extent, by enforcing that transition between propagating and stationary waves occur at the correct wavenumber (called the bifurcation point). From Equation 2.22, the bifurcation occurs for

$$\tan^2(\xi\Delta x) = \frac{(1 - Z(k))^2}{4X(k)Y(k)} \quad . \tag{2.23}$$

Since bifurcation should actually take place at $\xi\Delta x = \frac{1}{2}\tau^{-\frac{1}{2}}\Delta x = k$ (see Section 2.1.1), enforcing the condition

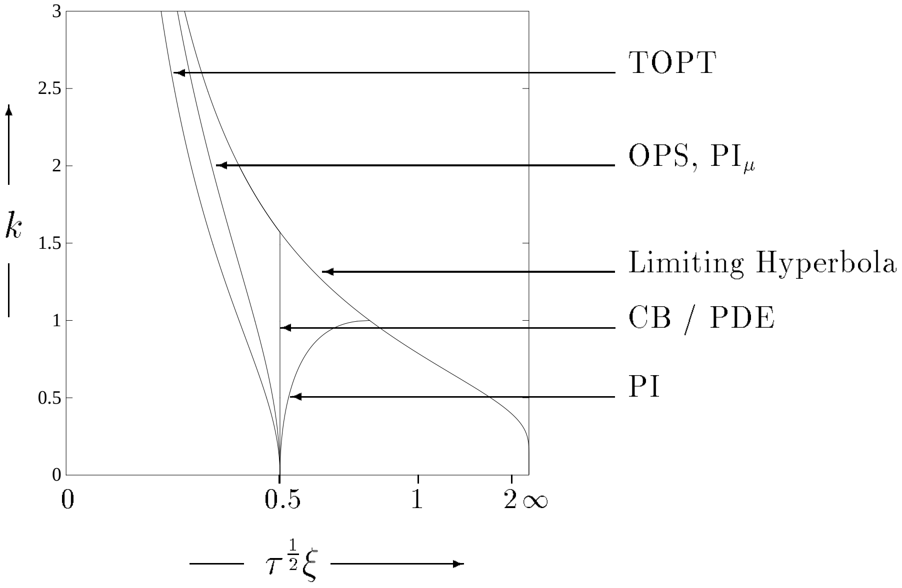


Figure 2.11: Splitting into damped and propagating modes according to the analytic solution, or the method of characteristics for the PI, OPS, PI$_\mu$, CB and TOPT schemes

$$\frac{(1 - Z(k))^2}{4X(k)Y(k)} = \tan^2(k) \tag{2.24}$$

ensures bifurcation at the correct wavenumber. The actual wavenumbers at which bifurcation occurs (for the schemes discussed so far) is plotted in Figure 2.11.

Several other "reasonable" constraints are possible, such as enforcing correct damping rates for the highest wavenumbers ($Z = e^{-2k}$) or requiring that the scheme be derivable from some pair of characteristic equations ($Z = X \cdot Y$). Imposing some pair of constraints then leads to different schemes. One such pair that looked promising was what will be called the CB scheme (CB enforces conservation and correct bifurcation), given by

$$Z(k) = e^{-2k} \quad , \tag{2.25}$$

$$X(k)Y(k) = \frac{(1-e^{-2k})^2}{4\tan^2 k} \qquad \text{for} \qquad k < \pi/2 \qquad (\text{assume } \; X = Y) \quad . \tag{2.26}$$

At the end of the day, it was found, surprisingly enough, that the simple PI scheme was awfully hard to beat [see Figure 2.12, which also includes the CB scheme]. Further, this figure shows that the various schemes can differ by more than two orders of accuracy in their root mean square errors. That is, even though all the schemes are formally second–order accurate, the error constants differ enormously. This lack of correlation between expectations based on desirable properties enforced and numerical results obtained led to the belief that there was more to this problem than met the eye.

Why is it that none of the method of characteristic schemes are a significant improvement over the simple PI scheme? And is the method of characteristics, in its pure form, even appropriate for dispersive wave problems? These lines of questioning led to the work in the next section, where the availability of an exact solution in integral form spotlights the missing feature in all the schemes tried so far.

2.4 Coupled Characteristics and the Optimum Scheme

Integrating around the circuit given in Figure 2.13[left] results in the governing integral equations (derived in [63]), which are

$$\theta_P = \frac{1}{2}e^{-k}(\theta_A + \theta_B) + \frac{\tau^{\frac{1}{2}}}{2}\int_A^B (\frac{\Omega}{\tau} - \Omega_t)\theta dx - \frac{\tau^{\frac{1}{2}}}{2}\int_A^B \Omega q_x dx, \tag{2.27}$$

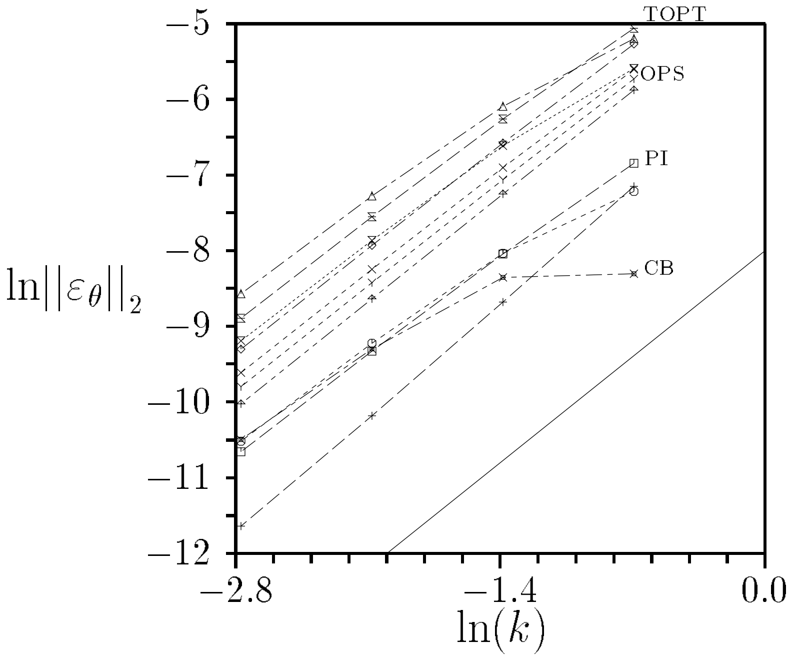


Figure 2.12: L_2 norm of the solution error in θ vs the stiffness factor k for several decoupled schemes at $t/\tau = 10$. The solid line has a slope of 2.

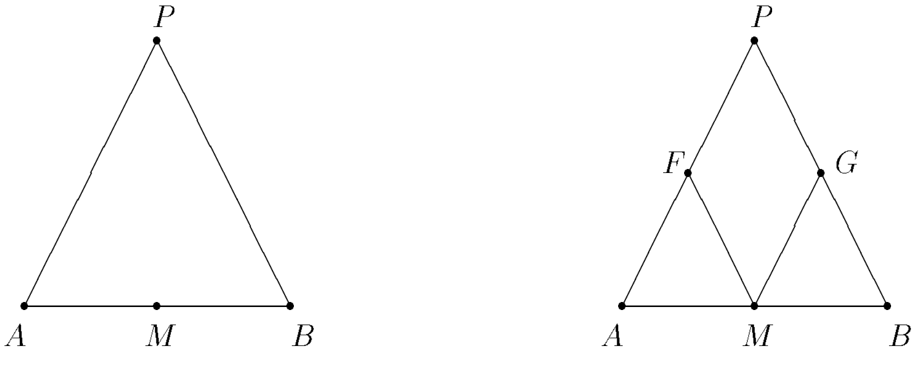


Figure 2.13: Stencil for the Coupled Method of Characteristics (OPT) [left] and that for the Simple Coupled Schemes (PC(S)) [right].

$$q_P = \frac{1}{2}e^{-k}(q_A + q_B) - \frac{\tau^{\frac{1}{2}}}{2}\int_A^B \Omega_t q dx - \frac{\tau^{-\frac{1}{2}}}{2}\int_A^B \Omega \theta_x dx, \tag{2.28}$$

where Ω is the Riemann function [74]

$$\Omega(\zeta, \eta) = \exp\left(\frac{\zeta - \zeta_1 + \eta - \eta_1}{4\tau}\right) I_0\left(\sqrt{\frac{(\zeta - \zeta_1)(\eta - \eta_1)}{4\tau^2}}\right). \tag{2.29}$$

To create a numerical method requires evaluation of the integrals. The function Ω (and hence its derivative Ω_t) is of course known exactly, but the functions θ and q need to be approximated, and will be represented as polynomials given by

$$u(x) = \frac{1}{2}(u_A + u_B) + \frac{x}{4k\tau^{\frac{1}{2}}}(u_B - u_A) + \frac{x^2 - 4k^2\tau}{8k^2\tau}(u_A - 2u_M + u_B) + O(\Delta x^3), \tag{2.30}$$

$$u_x(x) = \frac{(u_B - u_A)}{4k\tau^{\frac{1}{2}}} + \frac{x}{4k^2\tau}(u_A - 2u_M + u_B) + O(\Delta x^2), \tag{2.31}$$

where a quadratic variation in u has been assumed, and u is either q or θ. The fully discrete solution is [63]

$$\begin{aligned}\theta_P &= \frac{1}{2}(\theta_A + \theta_B) + \frac{\tau^{\frac{1}{2}}}{4k}(1 - e^{-2k})(q_A - q_B) \\ &+ \frac{1}{4k^2}(e^{-2k} - 1 + 2k - 2k^2)(\theta_A - 2\theta_M + \theta_B) \quad ,\end{aligned} \tag{2.32}$$

$$\begin{aligned}q_P &= \frac{e^{-2k}}{2}(q_A + q_B) + \frac{\tau^{-\frac{1}{2}}}{4k}(1 - e^{-2k})(\theta_A - \theta_B) \\ &- \frac{1}{4k^2}[e^{-2k}(1 + 2k + 2k^2) - 1](q_A - 2q_M + q_B) \quad .\end{aligned} \tag{2.33}$$

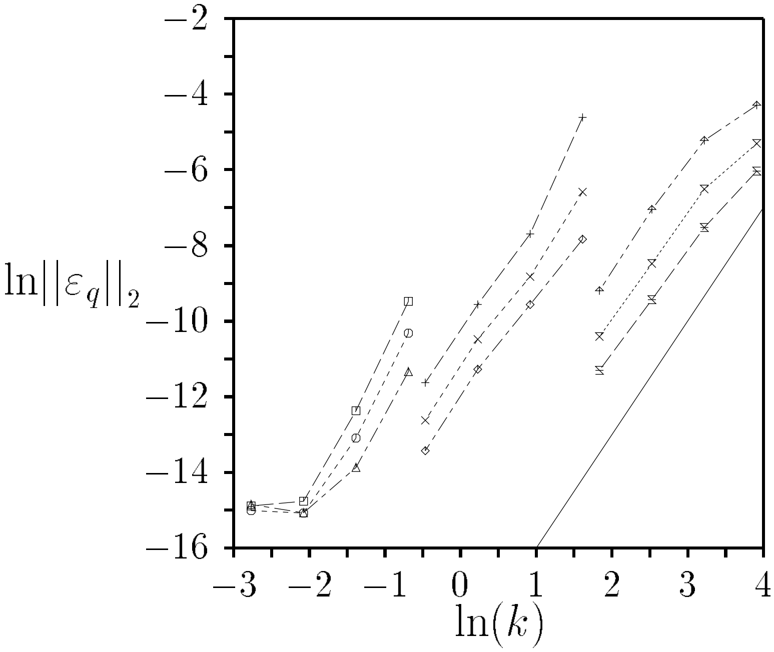
Note that Equations 2.32 and 2.33 are not of the form of Equations 2.14 and 2.15, having an extra term that involves the middle point M. This is the key observation here – this method (which is near–optimal) *couples the characteristics* via this middle point. What this means is that the source term provides a mechanism for the interaction of the wave families, which shows up in the discrete solution as the term involving the middle point M (see Equations 2.32 and 2.33). In short, the solution at P depends on the entire length of the initial line between the characteristics, *i.e.*, between A and B, including M). The entire spatio–temporal region $PAMB$ (Figure 2.13) is a region of wave interaction (of the two wave families) due to the source term.

The second–difference terms are in each case of order $k\Delta x^2$. An analysis of the truncation error indicates that their inclusion yields a scheme with third–order accuracy, whereas a scheme that excludes them is second-order at best.

Numerical solutions shown in Figure 2.14 are clearly far superior to those achieved by either the PI or the OPS schemes (and also much better than their second-order cousins shown in Figure 2.12). This scheme, which will be designated as the Optimum (OPT) scheme, is third–order accurate, and demonstrates that this coupling of characteristics is crucial. As a test, this system was "decoupled" by neglecting the last term in each of the above equations. This then gives the appearance of a method of characteristics similar to the PI and OPS schemes (we call the resulting scheme the TOPT or Truncated Optimum scheme). However, these cannot be decomposed into a characteristic pair ($XY \neq Z$), indicating some coupling of the characteristics is already involved. But the results (Figure 2.15) clearly show that the scheme has reverted to being a second–order one, similar to OPS. This similarity is also observed in the discrete dispersion plot (Figure 2.17, although there is a slight improvement for large k in the damping rates).

2.5 Simplified Coupled Schemes

In Section 2.4, it was shown that superior results could be achieved for the HHE's; however, that scheme was built upon the knowledge of an exact solution, which is obviously not a practical solution approach for general problems. In this section, a new method (based on the stencil shown in Figure 2.13[right]) is described, that



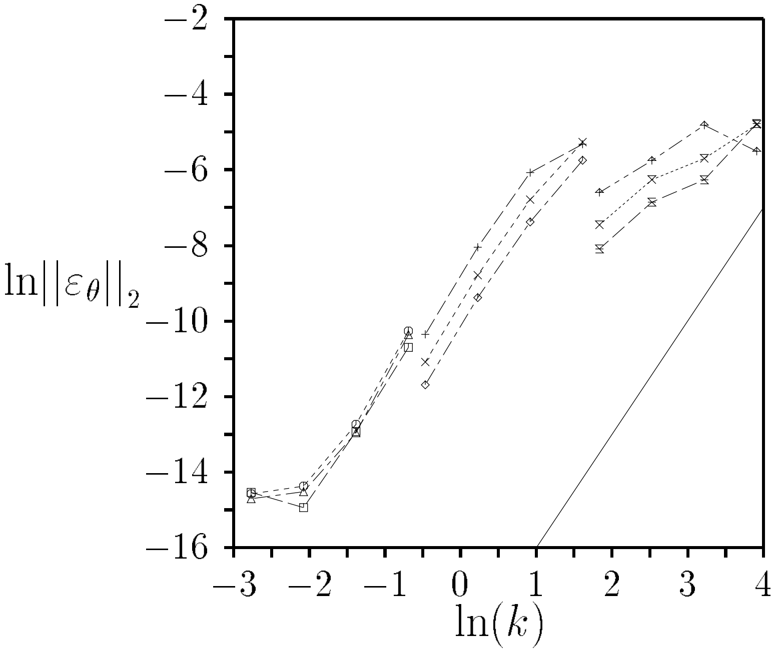
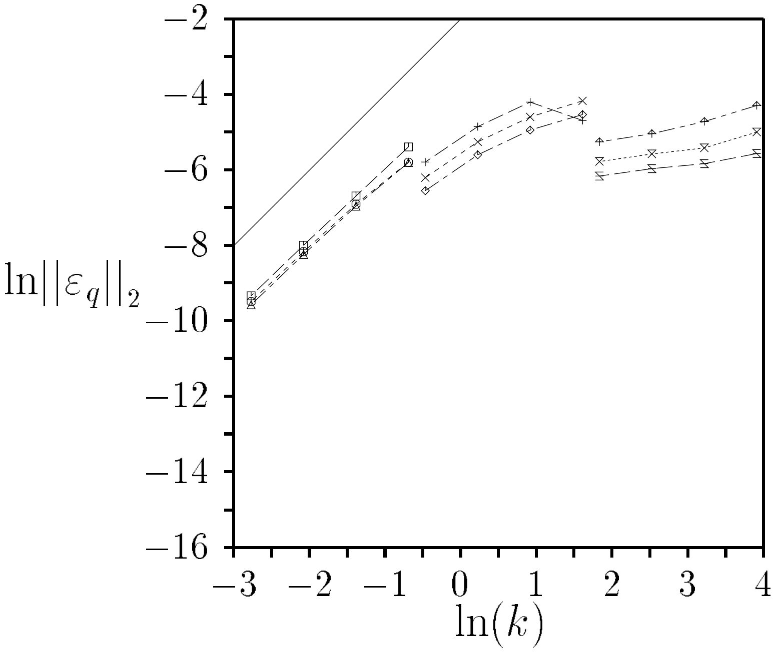


Figure 2.14: L_2 norm of the solution error in q and θ vs the stiffness factor k for the OPT (Optimum) scheme (see Table 2.1 for a key to lines and symbols). The solid line has a slope of 3.



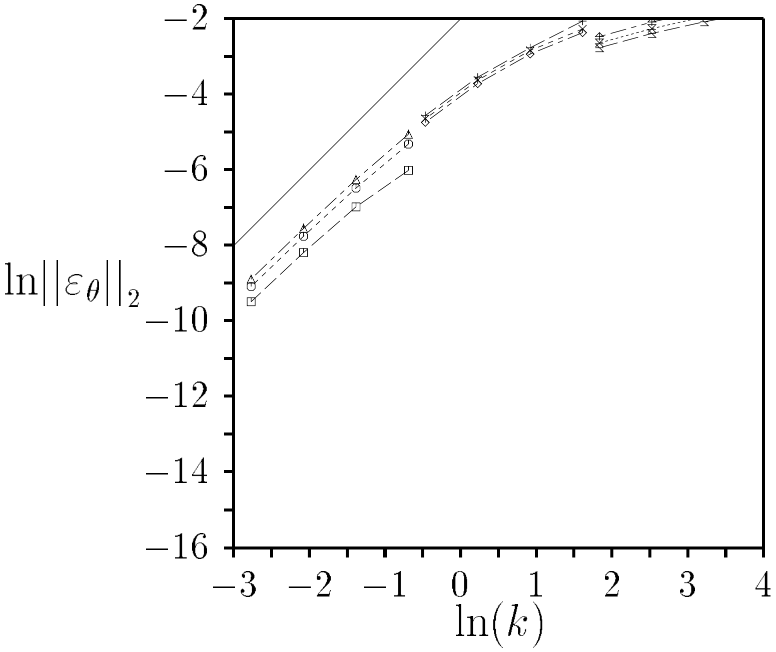


Figure 2.15: L_2 norm of the solution error in q and θ vs the stiffness factor k for the TOPT (Truncated Optimum) scheme (see Table 2.1 for a key to lines and symbols). The solid line has a slope of 2.

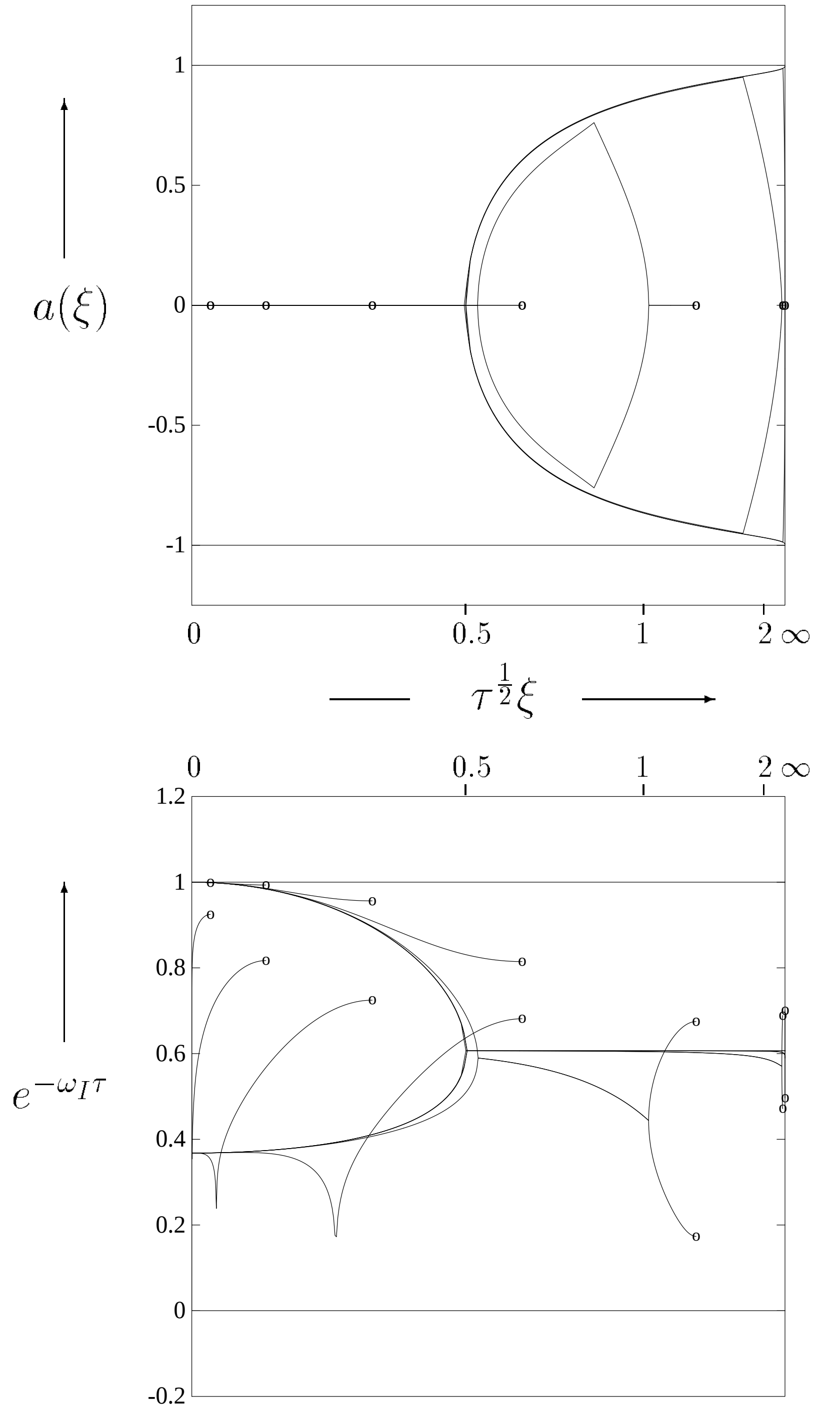


Figure 2.16: Propagating speeds and damping ratios for the OPT (Optimum) scheme. $k = 0.25$, 0.5, 1.25, 2.5, 5.0, 12.5, 50

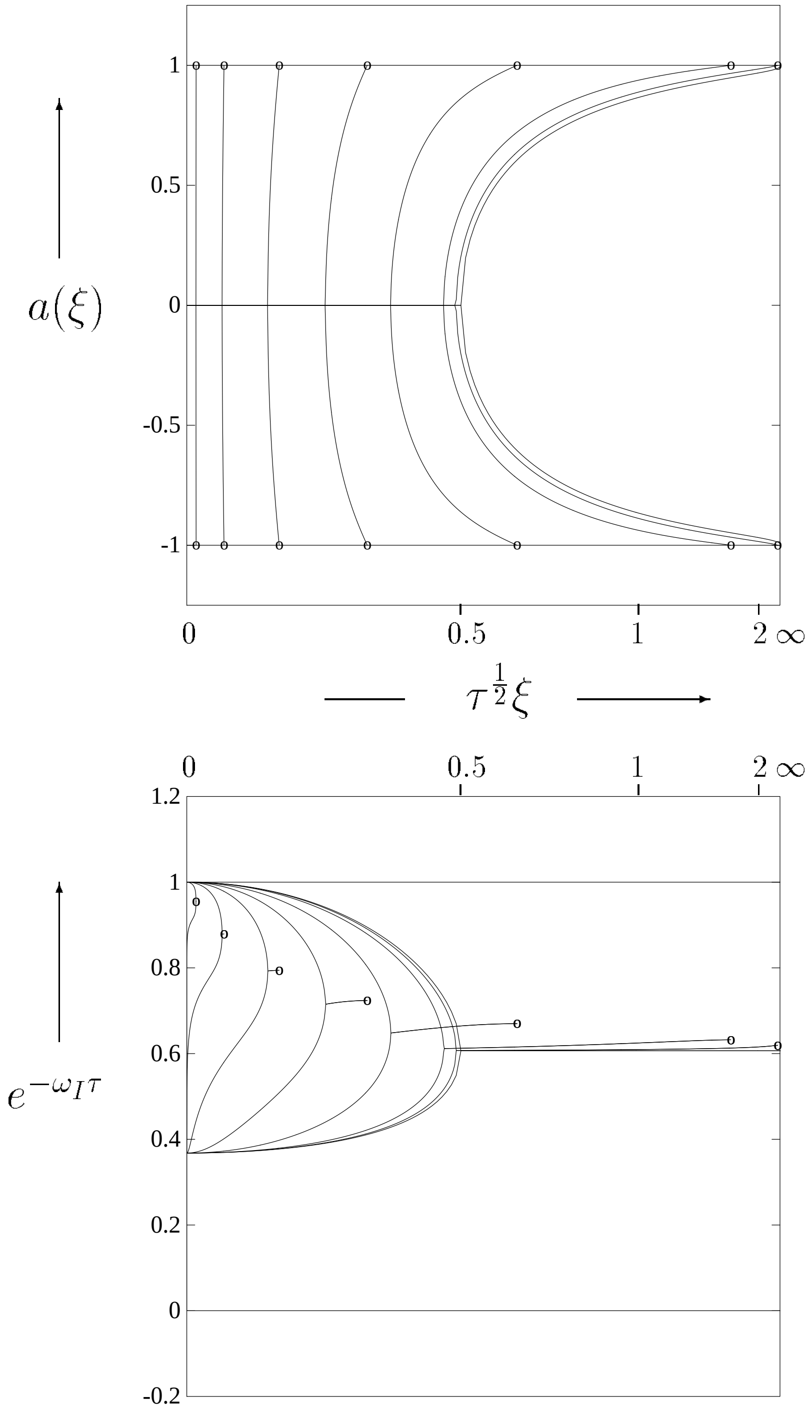


Figure 2.17: Propagating speeds and damping ratios for the TOPT (Truncated Optimum) scheme. $k = 0.25,\ 0.5,\ 1.25,\ 2.5,\ 5.0,\ 12.5,\ 50$

is close to optimal and yet is not dependent on any such knowledge (see [63] for a detailed derivation of this method).

Begin with the characteristic equations in conservation form

$$e^{\zeta/2\tau}\theta_\zeta + \tau^{\frac{1}{2}}\left[q \cdot e^{\zeta/2\tau}\right]_\zeta = 0 \quad , \tag{2.34}$$

$$e^{\eta/2\tau}\theta_\eta - \tau^{\frac{1}{2}}\left[q \cdot e^{\eta/2\tau}\right]_\eta = 0 \quad , \tag{2.35}$$

and then integrate these by parts along the characteristics to get

$$[\theta_P - \theta_A e^{-2k}] + \tau^{\frac{1}{2}}[q_P - q_A e^{-2k}] = \frac{1}{\tau}\int_{-\Delta t}^{0} e^{t/\tau}\theta(t)dt \quad , \tag{2.36}$$

$$[\theta_P - \theta_B e^{-2k}] - \tau^{\frac{1}{2}}[q_P - q_B e^{-2k}] = \frac{1}{\tau}\int_{-\Delta t}^{0} e^{t/\tau}\theta(t)dt \quad . \tag{2.37}$$

These are still exact, but now the integrals on the right hand side are approximated by assuming a polynomial variation in $\theta(t)$ (along the characteristics). For the case of a quadratic (the only case detailed here), three points are needed. From Figure 2.13[right], it is observed that the states at the feet of the characteristics (points A, B) are known. The values at the head are being computed (point P). However, solutions for the points midway along each characteristic (points F, G) are still needed; these are estimated by the PI method (note that AMF, MBG form characteristic stencils), which now provides a simple mechanism to couple the characteristics [since the values at F and G (and hence at P as well) involve the middle point M]. The hope was that merely incorporating this coupling, even in this simplistic fashion, would lead to second–order accuracy in both space and time. When solving relaxation systems, a point to note is that the information propagating along the characteristics is also exponentially damped [79] (see Equation 2.13), leading to rapid temporal variation. It is this feature that must be captured by successful higher–order schemes.

The final, fully discrete solution for this two–step method may be rewritten in the form

$$\begin{aligned}
\theta_P &= \left(\frac{\theta_A+\theta_B}{2}\right) + \frac{\tau^{\frac{1}{2}}}{2}(q_A-q_B)\left[\frac{(2+k+k^2)e^{-2k}-2+3k-k^2}{3k-2+(2+k)e^{-2k}}\right] \\
&\quad + \left[\frac{1-k-(1+k)e^{-2k}}{3k-2+(2+k)e^{-2k}}\right](\theta_A-2\theta_M+\theta_B) \quad , \qquad (2.38)\\
q_P &= \left(\frac{q_A+q_B}{2}\right)e^{-2k} + \frac{\tau^{\frac{1}{2}}}{2}(\theta_A-\theta_B)\left[\frac{1-e^{-2k}}{2k}\right] \\
&\quad + \frac{(k-2)}{4k^2}\left[1-k-(1+k)e^{-2k}\right](q_A-2q_M+q_B) \quad , \qquad (2.39)
\end{aligned}$$

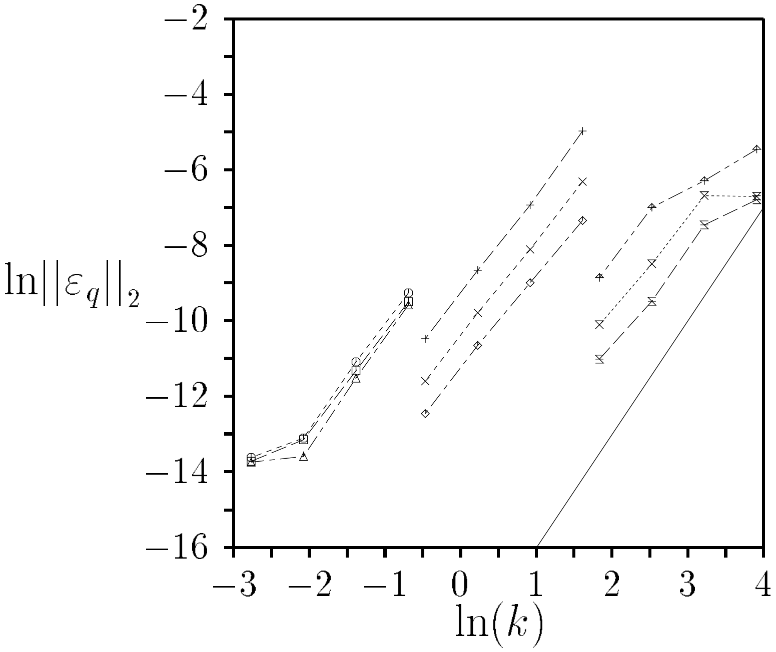
which is a (clumsy) rearrangement that facilitates a comparison with the OPT scheme. Like the OPT scheme, this method is third–order accurate, and the coefficients in the equations are comparable to the OPT scheme (in terms of Taylor series expansions for small k as well as asymptotic limits). The solution plots shown in Figure 2.18 demonstrate the accuracy (and viability) of this scheme. Note that when the OPS scheme was used instead of the PI scheme (in the predictor stage), the results were found to degrade in accuracy (they became more diffusive), as was hypothesized in Section 2.3. This led to the conclusion that the OPS scheme is inferior to the PI scheme, both on its own as well as when used in conjunction with another scheme, in the sense that it has larger dissipation[8] (even though the formal order of accuracy is not affected in either case).

2.6 Discussion of Results

A simple model has been studied in which dispersive wave behavior is caused by a source term. Computing cases where this term is also stiff has been the focus here, and the solution method was restricted to characteristic meshes ($\nu = 1$). Quite surprisingly, the popular operator splitting (OPS) method was found to be inferior to

---

[8] Recall the results of [55, 38] and the reduction of accuracy of Strang splitting for stiff problems.



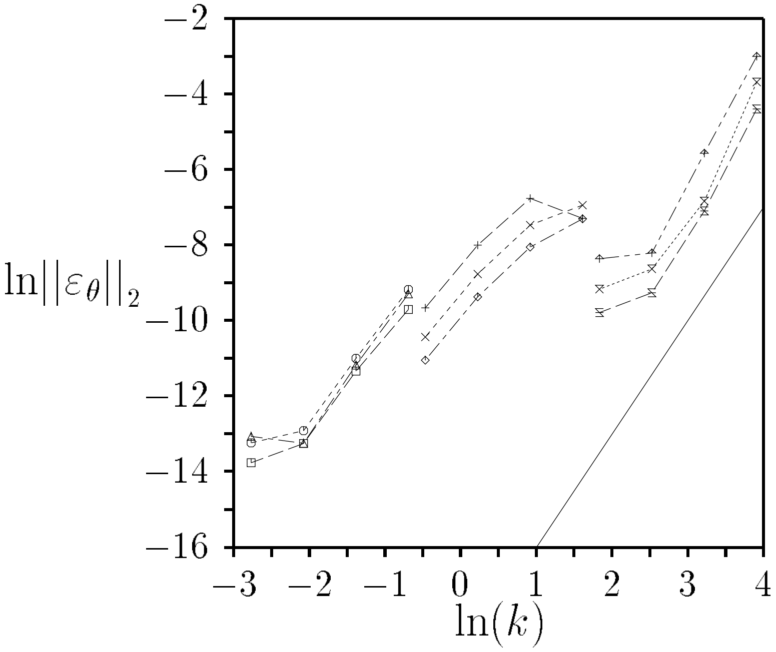


Figure 2.18: L_2 norm of the solution error in q and θ vs the stiffness factor k for the PC(S) (Predictor-Corrector (Simpson)) scheme (see Table 2.1 for a key to lines and symbols). The solid line has a slope of 3.

a simple point–implicit (PI) method. A modified PI scheme, the PI_{μ} scheme, was also developed, which had the "conservation" properties discussed in Appendix A [which will be seen to be crucial when solving nonlinear relaxation systems (Chapter IV)]. Further, this PI scheme was very hard to improve upon. It was also shown that any pure method of characteristics is at best second order accurate (for smooth flow and non–stiff problems on characteristic meshes) because it fails to account for the coupling of the wave families; moreover, the root–mean–square errors for different (method of characteristic) schemes (all of which are second–order accurate) were found to vary by more than two orders of magnitude (Figure 2.12).

The work described in this chapter has had a fairly limited aim – to "merely" design and implement an accurate solution algorithm for the smooth part of the model problem, even when the source term is stiff. However, the solution process proved to be non–trivial and quite illuminating. Note that since the model problem was linear, the method of characteristics would have been exact in the absence of the source term. Likewise, if only the source term had been present, it would have had led to a (trivial) ordinary differential equation, with an exponential solution and the (unimpressive) OPS scheme solves both these sub–problems exactly. Thus the source (no pun intended) of difficulties that arose were primarily due to the interaction of these two simple problems. In addition, by restricting this study to the smooth part of the flow on characteristic meshes, errors attributable to the discretization alone have been successfully isolated.

Similar to the observations of [55, 38], it was observed that the OPS scheme became less accurate as the fluid dynamic limit was approached ($\tau \rightarrow 0$). However, an advantage of the OPS scheme is that if the damping step is the first step, the method naturally projects the solution into equilibrium at each time step, which mimics the initial layer for under–resolved methods and pushes the solution to the correct asymptotic limit (see [38] for details).

Finally, these results clearly demonstrate that even if coupling of wave families is incorporated in a simple manner, the results improve dramatically, and accurate solutions can be obtained even in cases where the time step and mesh spacing are two orders larger than the relaxation time. This provides good impetus for the next stage – the development of finite volume shock capturing schemes on non–characteristic meshes (that have this coupling mechanism built in) to solve non–equilibrium flows – the subject of the next chapter.

CHAPTER III

CHARACTERISTIC–BASED FINITE–VOLUME SCHEMES FOR LINEAR 2×2 RELAXATION SYSTEMS

In Chapter II, it was shown that a scheme analogous to the optimum (OPT) scheme [63] probably represents the best that can be achieved (without expanding the stencil); however, since it relied heavily on analytical information (which is unlikely to be obtainable except in very special cases), it is not a practical approach. But a class of schemes that showed promise were the 2–stage Predictor–Corrector formulations (Section 2.5), producing results rivaling those obtained by the OPT scheme. Construction of such schemes is fairly simple and does not require any analytical information. However, the PCS scheme does require an estimate of the solution halfway through the time step. This predictor stage serves to couple[1] the characteristics in space and time, and facilitates a second–order integration along the characteristics. Generalization of the PI scheme to include a parameter μ (the degree of implicitness) led to the PI_μ scheme[2], where the value of the source term is computed at some point $(t_n + \mu\Delta t)$ in time (using a linear interpolation between

---

[1] The predictor stage involves the middle point, therefore influencing the update and incorporating coupling.

[2] Note that this virtual point does not cause coupling by itself; it simply represents an effort to compute the source term more accurately than the PI scheme by building in the effect of stiffness.

the endpoint values), and μ is a function of the stiffness factor k and determined by conservation (Appendix A).

Looking ahead, the eventual goal of this project is to solve non–linear systems with stiff relaxation terms. A natural framework for solving non–linear homogeneous systems is the finite–volume method [46]. Thus, in this chapter, an attempt will be made to come up with a good fully–discrete finite–volume formulation for solving the general linear 2×2 system shown in Chapter II.

The discretization of the domain is detailed in Section 3.1. This is followed by the reconstruction stage described in Section 3.2, where our choice is the Piecewise Parabolic Method (PPM) of Colella and Woodward [24]. Recall that use of a higher–order reconstruction does not in itself lead to higher accuracy for stiff problems (in the fluid dynamic limit) [38]. Rather, this choice of reconstruction was thought to be an economical way to improve results while using the simple PI_μ scheme to obtain predictor estimates in a characteristic tracing step (Section 3.3). The crucial flux computation step is described in Section 3.4, where a simple strategy is presented to accurately compute the flux, followed by a brief description of the update (Section 3.5). For simplicity of form as well as economy of computation, a point implicit update is chosen. Finally, the implementation of boundary conditions is described in Section 3.6, followed by the presentation of some numerical results in Section 3.7.

3.1 The Discretization

Spatially, the domain is divided into cells, which are restricted in this dissertation to be of uniform width Δx (Figure 3.1). However, the formulation can be extended quite readily to non–uniform meshes, such as those that arise in adaptive–grid computations [6, 26, 59]. In the finite–volume formulation, the cell averaged value in cell

j at time level t_n ($\mathbf{w}_j^n$), defined by

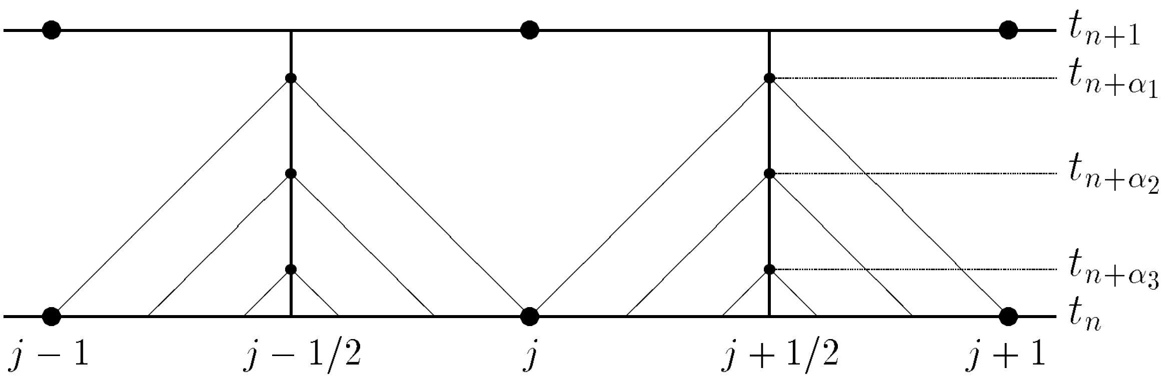


Figure 3.1: Schematic of the discretization, showing the cells (integers), the interfaces (half–integers), and the Gauss points ($t_{n+\alpha_k}$) along the interface. The mesh size is $\Delta x = x_{j+1/2} - x_{j-1/2}$ and $\Delta t = t_{n+1} - t_n$.

$$\mathbf{w}_j^n = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{w}\,(x, t_n) \cdot dx \quad ,$$

is known. The aim is to compute these solution vectors at the new time level t_{n+1}, (*i.e.*, $\mathbf{w}_j^{n+1}$, which will be denoted by $\mathbf{w}^j$).

The general form of the governing equations, written in conservation form, are

$$\mathbf{w}_t + \mathbf{f}_x = -\frac{1}{\tau}\mathbf{s} \quad ,$$

where $\mathbf{w} = (u, v)^T$, $\mathbf{f} = [v, p(u)]^T$ and $\mathbf{s} = [0, v - g(u)]^T$, and where it is assumed (in this chapter) that $p(u) = a_F^2 u$ and $g(u) = a_E u$. Integrating this system of equations over the cell j (Figure 3.1) gives

$$\mathbf{w}^j = \mathbf{w}_j - \frac{\Delta t}{\Delta x}\left[\mathbf{F}_{j+1/2} - \mathbf{F}_{j-1/2}\right] - \frac{1}{\tau}\int_{t_n}^{t_{n+1}} \mathbf{s}_j(t)dt \quad ,$$

$$\begin{aligned} \mathbf{F}_{j+1/2} &= \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}} \mathbf{f}_{j+1/2}(t)dt \quad , \\ \mathbf{s}_j &= \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{s}\,(x, t_n) \cdot dx \quad . \end{aligned}$$

Here, $\mathbf{F}_{j+1/2}$ is the time–averaged numerical flux through the $(j+1/2)$ interface and $\mathbf{s}_j$ is the cell–averaged value of the source vector.

3.2 The Reconstruction

The Piecewise Parabolic Method (PPM) of Colella and Woodward [24], based on one of the advection schemes (Scheme V) in [69], is used to generate the quadratic reconstruction in each cell, under the assumption of a uniform mesh. To start with, the cell–averaged values $\mathbf{w}_j$ for each cell are known. Let the (quadratic) variation within the cell (Figure 3.1) be given by

$$\mathbf{w}(x) = \mathbf{w}_{L_j} + x^* \left[c_{1_j} + c_{2_j}(1 - x^*)\right] \quad , \tag{3.1}$$

$$x^* = \frac{x - x_{j-1/2}}{\Delta x} \quad , \qquad x_{j-1/2} \le x \le x_{j+1/2} \quad . \tag{3.2}$$

Here, x^* is a normalized distance within the cell, such that $x^* \in [0,1]$, while x is the spatial running coordinate for the complete domain. The cell j is bounded by the $x_{j\mp1/2}$ interfaces, with $\mathbf{w}_{L_j}, \mathbf{w}_{R_j}$ being the reconstructed values at the left $(j-1/2)$ and right $(j+1/2)$ interfaces (of cell j) as $x^* \to \downarrow 0, \uparrow 1$, respectively. Define

$$\begin{aligned} c_{1_j} &= \mathbf{w}_{R_j} - \mathbf{w}_{L_j} \quad , & (3.3)\\ c_{2_j} &= 6\left[\mathbf{w}_j - \frac{1}{2}\left(\mathbf{w}_{L_j} + \mathbf{w}_{R_j}\right)\right] \quad , & (3.4) \end{aligned}$$

where these coefficients are related to the slope and the curvature of the reconstruction within the cell, being fully determined once $\mathbf{w}_{L_j}, \mathbf{w}_{R_j}$ are known. Further, define

$$\begin{aligned} \mathbf{w}_{j+1/2} &= \frac{1}{2}\left(\mathbf{w}_j + \mathbf{w}_{j+1}\right) + \frac{1}{6}\left(\delta\mathbf{w}_j - \delta\mathbf{w}_{j+1}\right) \quad , & (3.5)\\ \delta\mathbf{w}_j &= \frac{1}{2}\left(\mathbf{w}_{j+1} - \mathbf{w}_{j-1}\right) \quad , & (3.6) \end{aligned}$$

which is modified by the "discontinuity sharpening" mechanism of [24] using

$$\left\{ \begin{array}{l} \text{if} \\ \qquad (\mathbf{w}_{j+1} - \mathbf{w}_j)\,(\mathbf{w}_j - \mathbf{w}_{j-1}) > 0 \\ \text{then} \\ \qquad \delta\mathbf{w}_j \quad = \quad \min\left(|\delta\mathbf{w}_j|\,,\ 2|\mathbf{w}_j - \mathbf{w}_{j-1}|\,,\ 2|\mathbf{w}_j - \mathbf{w}_{j+1}|\,\right) \\ \text{else} \\ \qquad \delta\mathbf{w}_j \quad = \quad 0 \end{array} \right. \quad . \tag{3.7}$$

Now, assign this value of $\mathbf{w}_{j+1/2}$ (with the modified $\delta\mathbf{w}_j$) to $\mathbf{w}_{R_j}$ and $\mathbf{w}_{L_{j+1}}$, *i.e.*,

$$\mathbf{w}_{R_j} = \mathbf{w}_{j+1/2} \qquad ; \qquad \mathbf{w}_{L_{j+1}} = \mathbf{w}_{j+1/2} \quad . \tag{3.8}$$

To ensure monotonicity of this reconstruction, modify $\mathbf{w}_{R_j}, \mathbf{w}_{L_j}$ as

$$\left\{ \begin{array}{l} \text{if} \\ \qquad \left(\mathbf{w}_{R_j} - \mathbf{w}_j\right)\left(\mathbf{w}_j - \mathbf{w}_{L_j}\right) \quad \leq 0 \\ \text{then} \\ \qquad \mathbf{w}_{L_j} = \mathbf{w}_j \quad , \quad \mathbf{w}_{R_j} = \mathbf{w}_j \\ \\ \text{else if} \\ \qquad \left(\mathbf{w}_{R_j} - \mathbf{w}_{L_j}\right)\left[\mathbf{w}_j - \frac{1}{2}\left(\mathbf{w}_{L_j} + \mathbf{w}_{R_j}\right)\right] \quad > \quad \frac{1}{6}\left(\mathbf{w}_{R_j} - \mathbf{w}_{L_j}\right)^2 \\ \text{then} \\ \qquad \mathbf{w}_{L_j} = 3\mathbf{w}_j - 2\mathbf{w}_{R_j} \\ \\ \text{else if} \\ \qquad -\frac{1}{6}\left(\mathbf{w}_{R_j} - \mathbf{w}_{L_j}\right)^2 \quad > \quad \left(\mathbf{w}_{R_j} - \mathbf{w}_{L_j}\right)\left[\mathbf{w}_j - \frac{1}{2}\left(\mathbf{w}_{L_j} + \mathbf{w}_{R_j}\right)\right] \\ \text{then} \\ \qquad \mathbf{w}_{R_j} = 3\mathbf{w}_j - 2\mathbf{w}_{L_j} \end{array} \right. \quad . \tag{3.9}$$

In the actual implementation, boundary conditions are enforced on each of these intermediate variables at each step. The sequence of computation is: compute $\delta \mathbf{w}_j$ (Equation 3.6), modify it using Equation 3.7, obtain the face values $\mathbf{w}_{j+1/2}$ (Equation 3.5) and assign these to $\mathbf{w}_{R_j}$ and $\mathbf{w}_{L_{j+1}}$, enforce monotonicity (Equation 3.9) and finally compute the coefficients c_{1_j} and c_{2_j} (Equations 3.3, 3.4). At the end of this stage, the desired quadratic (PPM) reconstruction is fully determined (since $\mathbf{w}_{L_j}$, $\mathbf{w}_{R_j}$, c_{1_j}, c_{2_j} are now known in each cell).

3.3 The Characteristic Tracing Step

To evaluate the fluxes at the interfaces, trace the characteristics back from the point to be computed (on the interface) to the initial line (t_n), where the reconstructed data is now available (Figure 3.1). These values are input to the PI_μ formulation, which uses the method of characteristics to estimate the state quantities at the interface.

Since the study in this chapter is limited to the linear case, there is no ambiguity regarding the wavespeed because a_F is constant (note that a_F is the only sensible wavespeed on which to base a method of characteristics). The locations in the cell, *i.e.*, the x^*'s at which the reconstructed values are needed (for the 2 × 2 case) are

$$x^*_{L_j} = 1 - \frac{a_F \Delta t}{\Delta x} \Delta t^* = 1 - \nu \Delta t^* \quad , \tag{3.10}$$

$$x^*_{R_j} = 1 - x^*_{L_j} \quad . \tag{3.11}$$

Here, Δt^* is the time at which the solution on the interface is desired, normalized by the time step, *i.e.*,

$$\Delta t^* = \frac{\delta t}{\Delta t} \quad ; \quad \delta t \in [0, \Delta t] \quad ; \quad \Delta t^* \in [0, 1] \quad ,$$

(*e.g.*, if the solution at $t_{n+1/2}$ is needed, then $\Delta t^* = \frac{1}{2}$). Note that in order to compute a solution at the $(j+1/2)$ interface, the values at $x^*_{L_j}$ and $x^*_{R_{j+1}}$ are needed.

The implementation is fairly simple: for each cell, compute $x^*_{L_j}$ and $x^*_{R_j}$, knowing which the state vector at these points, *i.e.*, $\mathbf{w}(x^*_{L_j})$, $\mathbf{w}(x^*_{R_j})$, can be obtained from the PPM reconstruction. Since the points $x^*_{L_j}$, $x^*_{R_{j+1}}$ and $x^{n+\Delta t^*}_{j+1/2}$ form a characteristic stencil, the value $\mathbf{w}^{n+\Delta t^*}_{j+1/2}$ can be easily computed using the PI$_\mu$ scheme, where the degree of implicitness is given by

$$\mu = \frac{1}{1-e^{-2k^*}} - \frac{1}{2k^*} \quad , \quad \mu \in [\frac{1}{2}, 1] \quad ,$$

and $k^* = \frac{\Delta t^*}{2\tau}$ is the stiffness factor for this fraction of the time step.

One can now integrate along the (frozen) characteristics and solve for u_P, v_P to get

$$\begin{aligned} u_P &= \tfrac{1}{2}\left[u^*_A + u^*_B + (v^*_A - v^*_B)/a_F - 2\,(1-\mu)\,k^*\,(v^*_A - g^*_A - v^*_B + g^*_B)/a_F\right] \quad , \\ g^*_P &= g(u_P) \quad , \\ v_P &= \tfrac{1}{2(1+2k^*\mu)}\left[v^*_A + v^*_B + a_F\,(u^*_A - u^*_B) - 2k^*\,(1-\mu)\,(v^*_A - g^*_A + v^*_B - g^*_B) + 4k^*\mu g^*_P\right] \quad . \end{aligned}$$

Note that the source terms have been point–implicitly incorporated in the above expressions.

3.4 The Flux Computation

The simplest approximation to the flux is to compute the values u_P, v_P (as defined in Section 3.3) at the $(n+1/2)$ time level and define

$$\mathbf{F}_{j+1/2} = \mathbf{f}\left(\mathbf{w}^{n+1/2}_{j+1/2}\right) \quad . \qquad (3.12)$$

However, this proves to be too diffusive even for very slightly non–linear problems. Recall that, especially in the case where the problem is stiff, the source term causes very rapid (exponential) variation of the solution along the interface. Thus, it is very likely that an estimate that merely takes the value obtained at the $(n+1/2)$ time–level to represent the average value on the interface will be quite inaccurate.

At this stage, a mechanism to compute an approximate value for $\mathbf{w}$ at any temporal location along the interface was available, which led to the following obvious question: given this computing mechanism, is there a simple and economical way to obtain a more accurate solution for the interface flux? This is strongly desired, since the (simple) PI_μ scheme is used to obtain estimates along the interface. A strategy that suggested itself was to calculate the flux integral using Gaussian quadratures. Some testing on a function composed of a product of a polynomial and an exponential led to the empirical conclusion that a three point quadrature is sufficient, which is what will be used in all the examples that follow.

3.4.1 The Gaussian Quadrature

In order to use Gaussian quadratures to evaluate the interface flux

$$\mathbf{F}_{j+1/2} = \frac{1}{\Delta t}\int_0^{\Delta t} \mathbf{f}_{j+1/2}(t)dt \quad ,$$

the limits of integration must be renormalized to be $[-1,1]$, which leads to

$$\mathbf{F}_{j+1/2} = \frac{1}{2}\int_{-1}^{1} \mathbf{f}_{j+1/2}(\alpha)d\alpha \quad = \quad \frac{1}{2}\sum_{k=1}^{k=K} w_k \mathbf{f}_{j+1/2}(\alpha_k) \quad ,$$

$$\alpha_k = \frac{1}{2}\left(\Delta t \xi_k + \Delta t\right) = \frac{\Delta t}{2}\left(1+\xi_k\right) \quad .$$

Here, ξ_k and w_k are the location and weight of the k^{th} Gauss point, for a K point quadrature (these are given for the 3 point quadrature in Appendix B; also see [1]), while the α_k are the temporal locations at which the state variables $\mathbf{w}_{j+1/2}^{n+\alpha_k}$ are computed in order to obtain the fluxes $\mathbf{f}_{j+1/2}(\alpha_k) = \mathbf{f}\left(\mathbf{w}_{j+1/2}^{n+\alpha_k}\right)$.

3.5 The Update Step

For stability, this step must be at least partially implicit. However, if it had a predetermined (fixed) degree of implicitness, then one could choose a sufficiently large value of the stiffness factor, which could result in the scheme becoming unstable. Thus, for simplicity, this step is assumed to be point implicit (but with $\mu = 1$). Interestingly enough, a more sophisticated update using the values at t_n, $t_{n+1/2}$ and t_{n+1} (to form a quadratic) failed as the fluid dynamic limit was approached. It is quite plausible that the reason was insufficient "implicitness" (as reasoned above) for the time step chosen. Thus, the choice of update is primarily for simplicity and unconditional stability (at least for this particular step).

The equations studied so far have had the form

$$\begin{aligned} u_t + h(u)_x &= 0 \quad , \\ v_t + p(u)_x &= -\frac{1}{\tau}[v - g(u)] \quad , \end{aligned}$$

where the second equation was the only non–homogeneous one and linear in v. The update, then, is simply

$$\begin{aligned} u^j &= u_j - \frac{\Delta t}{\Delta x}\left(H_{j+1/2,m} - H_{j-1/2,m}\right) \quad , \\ v^j &= \frac{1}{1+2k}\left[v_j - \frac{\Delta t}{\Delta x}\left(P_{j+1/2} - P_{j-1/2}\right) + 2kg(u^j)\right] \quad , \end{aligned}$$

where H, P are the numerically computed values of the fluxes h, p.

For equations of this form, even if h, g, p are nonlinear, it doesn't affect the update procedure, since v is linear and the updated value of u is already available for use in the computation of the source term g. However, there are conceivable cases where such a form is not available; it will be seen in later chapters that such cases revive the notion that CFD is often more an art than an exact science.

3.6 Boundary Conditions

The computational domain consists of cells 1 through J in the interior, with $0, J+1$ being the boundary (ghost) cells. The boundary conditions used in this chapter are easily imposed on a variable z as follows.

3.6.1 Reflecting Boundary

$$z_0 = z_1 \qquad ; \qquad z_{J+1} = z_J \quad .$$

3.6.2 Periodic Boundary

$$z_0 = z_J \qquad ; \qquad z_{J+1} = z_1 \quad .$$

3.7 Numerical Results

In order to test this algorithm, a case recently presented by Jin and Levermore [41] is attempted. The system being solved is

$$\begin{array}{rcl} u_t + v_x & = & 0 \\ v_t + u_x & = & -\frac{1}{\tau}[v - g(u)] \end{array} \quad , \tag{3.13}$$

where $g(u) = 0$, with initial conditions

$$\begin{array}{rcl} u(t=0) & = & 2 + \sin(\pi x) \\ v(t=0) & = & -0.1 \end{array} \quad , \quad x \in [0,2] \quad , \tag{3.14}$$

and a relaxation time $\tau = 0.01$. The boundary conditions are periodic. Solutions are shown in Figure 3.2 for discretizations in the intermediate[3] $[dx = 0.0125 \sim O(\tau)]$ and thick[4] $[dx = 0.125 \gg O(\tau)]$ regimes, for a Courant (CFL) number $\nu = 0.5$, at $t = 2$ (note that the results in [41] were for semi–discrete schemes and used a

---

[3] The intermediate regime is where the spatial mesh size is of the same order as the relaxation length $[dx \sim O(\tau u_{\mathrm{char}})]$.

[4] The thick regime is one where the spatial mesh size fails to resolve the relaxation length scales $[dx \gg O(\tau u_{\mathrm{char}})]$.

time–step $\Delta t \ll \tau$). The "exact" solution is a fine–mesh solution ($dx = 0.001$), that fully resolves the relaxation time, computed at a Courant number of unity. The numerical results for the current method are clearly seen to be excellent, and are competitive with those shown in [41] (even though the present results are temporally more severely under–resolved).

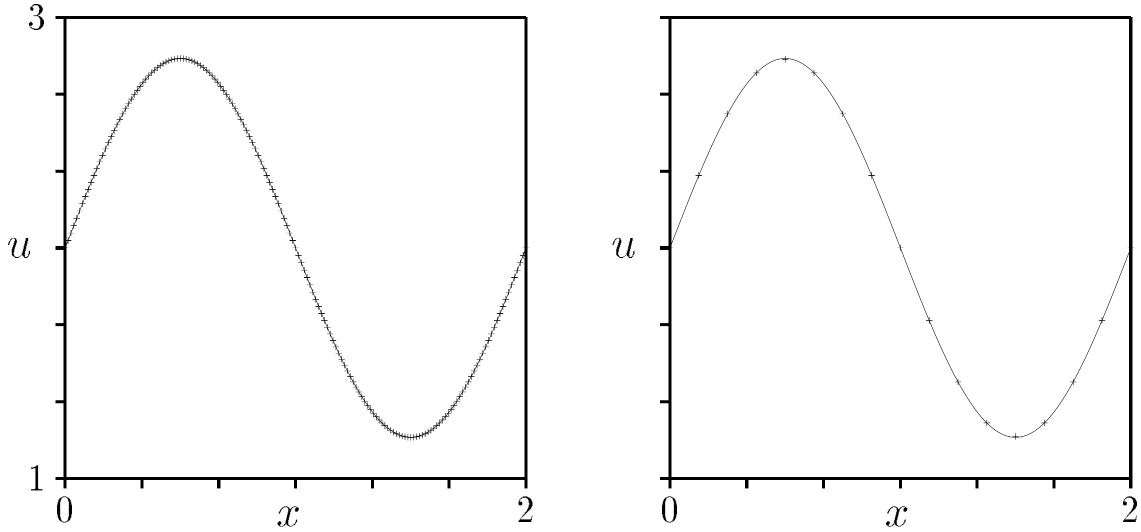


Figure 3.2: Solution to the Jin-Levermore linear case (Equation 3.13), for discretizations in the intermediate [$dx = 0.0125 \sim O(\tau)$, left figure (+)] and thick [$dx = 0.125 >> O(\tau)$, right figure (+)] regimes at $\nu = 0.5$ superposed onto the "exact" solution [$dx = 0.001$, $\nu = 1$ (solid line)].

3.8 The Next Steps

In Chapter II, several key features influencing the success of numerical algorithms were identified. First, the numerical scheme must couple the wave families. Second, operator–split schemes lose accuracy for large stiffness (see [55, 38]). Third, the under–resolved scheme must project the initial data onto a local equilibrium at each time step [38]. This mimics the (unresolved) initial layer and ensures passage from the relaxation system to the equilibrium system. Since the PI_μ scheme approaches a fully point implicit scheme ($\mu = 1$) in the fluid dynamic limit, it asymptotes to the correct behavior (an implicit step automatically enforces this projection condition).

These insights were incorporated into the finite volume method developed in this chapter. Specifically, use of a predictor stage similar to that used in Chapter II leads to estimates along the interface. These estimated states are then used in Gaussian quadratures and help to capture the temporal variation of the flux, leading to accurate (flux) computations. In this predictor stage, the PI_μ scheme is used, rather than an OPS based scheme [which was shown to be less accurate for the linear case in Chapter II]. One may well question the use of the PI_μ scheme over the PI scheme (since the PI scheme was seen in Chapter II to be far superior). For the case tested here, there were no visible differences between the results from the PI and PI_μ schemes (to eyeball norm[5]) for relaxation times up to 10^{-6}. It will be shown in Chapter IV, however, that for the nonlinear cases, there *is* a substantial difference between the two.

The reconstruction within each cell is quadratic (similar to the OPT scheme), based on the PPM formulation [24]. The update is taken to be point implicit, but with $\mu = 1$ (its attraction being its simplicity and unconditional stability). The resulting scheme is clearly seen to perform very well. Having accomplished the goals for this chapter, the next logical step is to extend this method to solve non–linear relaxation systems, which is the subject of the next chapter.

---

[5] The *eyeball norm* is a comparison of two plots by holding them together against a light source.

CHAPTER IV

THE TRANSITION TO FINITE-VOLUME SCHEMES FOR NON–LINEAR 2×2 RELAXATION SYSTEMS

A successful method was designed in Chapter III to solve stiff relaxation systems, but with one severe limitation – all the cases considered so far have been linear. In this chapter, the challenging “non–linear” journey is begun, but again limited to 2×2 relaxation systems of the form

$$\mathbf{w}_t + \mathbf{f}_x = \frac{1}{\tau}\mathbf{s} \quad ,$$

where $\mathbf{w},\ \mathbf{f},\ \mathbf{s}$ are state, flux and source vectors respectively, and τ is the relaxation time of the problem.

The first complication is that the characteristics are no longer straight lines, since wavespeeds are now functions of state variables. Further, such systems admit shocks and rarefactions, which need to be captured. Since the goal is to develop a systematic methodology, the aim is to solve successively harder problems, incorporating previous insights into solution procedures developed at each subsequent stage.

In this chapter, therefore, it is demonstrated how the algorithm detailed in Chapter III can be modified to account for the non–linearity of the problem. Specifically, the discretization, reconstruction, flux computation and update stages are unchanged, while the characteristic tracing step needs to be modified, and the changes are detailed in Section 4.1. Finally, numerical results for the proposed scheme are presented in Section 4.2, and are compared against the best results available in the literature at this time [41, 38, 15].

4.1 A Modified Characteristic Tracing Step

The first question that arises is: how does one determine the frozen wavespeeds at the interface? Since it must be some average value based on the states on either side of the interface, three simple averages were considered. These were the arithmetic $[u_{\text{avg}} = \frac{1}{2}(u_L + u_R)]$, geometric $[u_{\text{avg}} = (u_L \cdot u_R)^{1/2}]$, and Roe [61] $(u_{\text{avg}} = \hat{u}_{\text{Roe}})$ averages[1]. However, the numerical results were unchanged (to eyeball norm); hence, for simplicity and economy, the arithmetic mean has been used for all subsequent cases in this dissertation.

For the system

$$\begin{aligned} u_t + v_x &= 0 \quad , \\ v_t + p(u)_x &= -\frac{1}{\tau}[v - g(u)] \quad , \end{aligned}$$

the eigenvalues are $\lambda_{1,2} = \mp\,[p_u(u)]^{1/2}$, resulting in an averaged wavespeed $\hat{a}_F$

$$\hat{a}_F = \frac{1}{2}(a_L + a_R) \quad ; \quad a_* = [p'(u_*)]^{1/2} \quad ; \quad * = L, R \quad .$$

Following Section 3.3, the traced–back characteristics intersect the initial line at

$$x^*_{L_j} \;=\; 1 - \frac{\hat{a}_F \Delta t}{\Delta x}\Delta t^* \;=\; 1 - \nu \Delta t^* \quad , \tag{4.1}$$

$$x^*_{R_{j+1}} \;=\; 1 - x^*_{L_j} \quad . \tag{4.2}$$

---

[1] Note that the Roe–average is based on data at time t_n. Thus, Roe–averaging in the conventional sense is not adequate for *solving* stiff relaxation systems, since it assumes the flux doesn't vary temporally; in standard procedures, the source term is evaluated as $\mathbf{s}_j = \mathbf{s}(\mathbf{w}_j)$ in the update.

Given these locations, the reconstructed values are easily obtained, and now the PI_μ scheme is used to obtain the (method of characteristics) solution at the Gauss points along the interface.

The next modification is that since the wavespeed is no longer a constant, the time step must be computed at each time iteration to ensure that the CFL condition (based on the convection terms alone) is satisfied. This is a simple function call accomplished by first computing the maximum eigenvalue in the entire computational domain $\left\{[p_u(u)]^{1/2}\right\}_{\max}$, and then setting the time step as

$$\Delta t = \frac{\nu \Delta x}{\left\{[p_u(u)]^{1/2}\right\}_{\max}} \quad .$$

Note that in the Gaussian quadrature (flux computation) loop, one could base the frozen wavespeed on the "average" state computed at the previous Gauss point [*i.e.*, use $\hat{a}_F(\mathbf{w}_L,\ \mathbf{w}_R)$ for the first Gauss point, $\hat{a}_F(\mathbf{w}_{\mathrm{Gauss}_1})$ for the second Gauss point, and so on], but this was found to make no difference (to eyeball norm) and hence was not pursued further.

4.2 Test Cases

Recently, results have been presented for a wide range of problems with stiff source terms [41, 38], some of which are used here as benchmark tests for the proposed algorithm. The novelty of the approach in [41, 38] is that the physical asymptotics have been built into the numerical scheme (see Section 5.2), essentially using semi–discrete schemes (in the fully discrete context, Runge–Kutta schemes with a good temporal integrator for stiff ODE's are used, coupled with the semi–discrete formulation). In contrast, the method proposed here is a fully–discrete one that relies primarily on accurate flux computation. The system being solved is

$$\mathbf{w}_t + \mathbf{f}_x = -\frac{1}{\tau}\mathbf{s} \quad , \tag{4.3}$$

where $\mathbf{w} = (u,v)^T$, $\mathbf{f} = [v, p(u)]^T$ and $\mathbf{s} = [0, v - g(u)]^T$.

Note that a further complication can arise depending on the choice of initial data: if $v = g(u)$ at $t = 0$, the initial data is initially in local equilibrium; however, if this is not true (as in most of the cases presented here), an initial layer forms and under–resolving this layer can cause spurious or incorrect solutions [15].

4.2.1 The Weakly Non–Linear Test Case

Here, the variables of Equation 4.3 are $p(u) = -u^{-2}$ $\left(\text{hence } a_F = \left[\frac{2}{u^3}\right]^{1/2}\right)$, $g(u) = 0.01\,(u-2)^2$, and $\tau = 0.01$. The (non–equilibrium) initial conditions are

$$u(0,x) = 2 + \sin(\pi x) \quad , \quad v(0,x) = -0.01 \quad ,$$

and the boundary conditions are periodic.

Results were presented in [41] in the thick regime [$dx = 0.1$]. The standard scheme[2] (with no built–in information about the asymptotics) was shown to be highly diffusive, but the scheme proposed in [41] (with the asymptotic information incorporated) gave very good results. However, in the intermediate regime ($dx = 0.01$), the results of [41] were actually worse than those for the unmodified schemes[3].

Numerical results from the scheme proposed here are presented at a time $t = 10$ for the thick regime [$dx = 0.1$, Figure 4.1 (left)] as well as for the intermediate regime [$dx = 0.01$, Figure 4.1 (right)], and superposed onto the “exact” (fine–mesh) solution

---

[2] In a standard scheme, the source term is approximated by $\mathbf{s}(\mathbf{w}_j)$ and some higher–order reconstruction is used to evaluate the fluxes, *e.g.*, MUSCL [70] or PPM [24]. These do not incorporate asymptotic information, and the Riemann solver does not account for the source term. In [71], the source term actually influences the input states of the Riemann solver.

[3] In a more recent paper [15], results with uniform convergence have been achieved, but a special initial layer fix is required for the under–resolved scheme.

(solid line, $dx = 0.001$). The present results (Figure 4.1) are seen to be superior to those of [41], with no visible shortcoming in the intermediate regime.

Note that when this test case was run in the thick regime using the PI scheme (PI_μ with $\mu = \frac{1}{2}$), the results deteriorated as the fluid dynamic limit was approached, and blew up around $\tau = 5.5\ 10^{-5}$. This was the first indication that conservation was critical, and our hunch from Chapter II paid off.

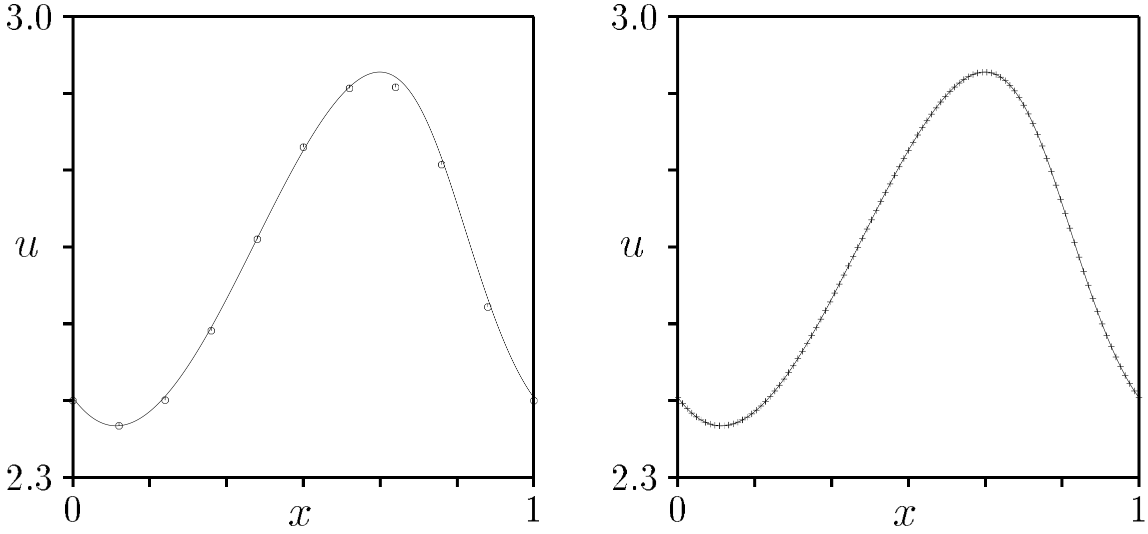


Figure 4.1: Solution to the Weakly Non–linear case [41] (also see Section 4.2.1), for discretizations in the thick [$dx = 0.1 >> O(\tau)$, left figure (o)] and intermediate [$dx = 0.01 \sim O(\tau)$, right figure (+)] regimes at $\nu = 0.9$ superposed onto the exact solution [$dx = 0.001$ (solid line)].

4.2.2 The Strongly Non–Linear Test Case

Here, the variables of Equation 4.3 are $p(u) = u + \frac{1}{2}u^2 \left[a_F = (1+u)^{1/2}\right]$, $g(u) = \frac{1}{2}u^2$, and $\tau = 0.01$. The (non–equilibrium) initial conditions are

$$u(0,x) \quad = \quad \begin{cases} 1 & -1 < x < 0.2 \\ 0.2 & 0.2 < x < 1 \end{cases} \quad , \quad v(0,x) = 0 \quad ,$$

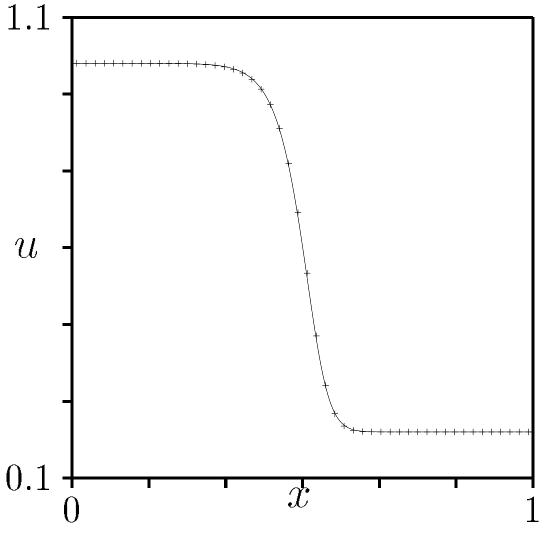


Figure 4.2: Solution to the Strongly Non–linear case [41] (also see Section 4.2.2) in the intermediate regime [$dx = 0.02 \sim O(\tau)$, (+)] at $\nu = 0.5$ superposed onto the "exact" solution [$dx = 0.001$ (solid line)].

with reflecting boundary conditions. The solution is plotted in Figure 4.2 at $t = 0.5$.

The results in [41] in the intermediate regime ($dx = 0.02$) were slightly worse than those of the unmodified schemes. Results from the proposed algorithm in the intermediate regime ($dx = 0.02$), superposed onto the "exact" solution (solid line, $dx = 0.001$), are shown in Figure 4.2, and are better than those in [41].

Again, the PI scheme was run on this case with a relaxation time $\tau = 10^{-4}$, while varying the mesh size Δx. With increasing mesh size, the results deteriorated rapidly, blowing up for $\Delta x > 0.0175$ (approximately). This demonstrated clearly that as the fluid dynamic limit is approached, the simple PI scheme which was so successful on the linear case, degrades rapidly, and actually blows up. The results presented in Section 4.2.3 for severely under–resolved flows clearly demonstrate the robustness of the PI_μ scheme (which enforces conservation in the sense of Appendix A).

4.2.3 Two Cases with Very Large Stiffness

The examples in Section 4.2.1 and 4.2.2 were for non–linear relaxation systems, but were not severely under–resolved for the cases run. Here, two cases are attempted where the temporal (spatial) under–resolution is of the order of one (fifty) million.

Initial Data in Local Equilibrium

For this case, the variables of Equation 4.3 are $p(u) = u + \frac{1}{2}u^2$ $\left[a_F = (1+u)^{1/2}\right]$, $g(u) = \frac{1}{2}u^2$. Smooth, equilibrium initial data

$$u(0,x) = 1 + 0.2\sin(8\pi x) \quad , \quad v(0,x) = \frac{1}{2}u(0,x)^2$$

is specified, so as not to generate an initial layer. The numerical solution (+) is plotted in Figure 4.3 (left) at a time $t = 0.3$ with periodic boundary conditions, and a relaxation time $\tau = 10^{-10}$ [$dx = 0.01$, $\nu = 0.03$, $\Delta t/\tau \sim 2 \cdot 10^6$] and superposed onto the "exact" solution [(solid line with $dx = 0.001$), which is nothing but the solution to Burger's equation with this initial data]. The results [Figure 4.3 (left)] are clearly seen to be very good, with better peak recovery and sharper shocks than obtained in [38]. However, there is some visible distortion of the profiles near the peaks, due to the severity of the under–resolution. Also note that the relaxation time in the current study is two orders of magnitude smaller than that used in [38].

Recall that the design philosophy being followed here is to keep the method simple and economical, but to allow the results to degrade smoothly for large stiffness; however, any method designed was still required (naturally) to yield the correct physical solutions (correct shock jumps and speeds), even for severely under–resolved flow. It is estimated that the current method is at least an order of magnitude cheaper to compute than the scheme developed in [15][4].

---

[4] The algorithm proposed in [15] is based on the method of lines, and consists of a second–order

Initial Data Not in Local Equilibrium

In Figure 4.3 [right], the strongly non–linear case from Section 4.2.2 is presented again, but with $\tau = 10^{-10}$, $dx = 0.01, \nu = 0.03, \Delta t/\tau \sim 2 \cdot 10^6$. The numerical solution (+) is superposed onto the exact (analytical) solution (solid line), which is a single shock moving to the right with a speed of 0.6. There is a small overshoot (less than 2% of the jump); however, the solution is clearly the physical one having the correct shock speed and jump. And for such drastic under–resolution, the (slight) overshoot is a small price to pay for the (large) gain in computational economy.

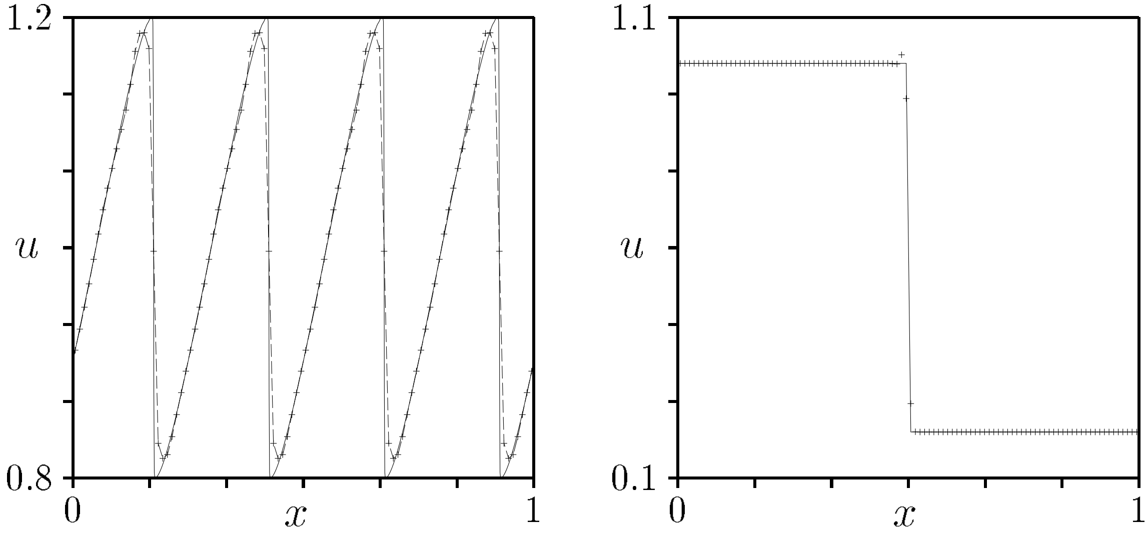


Figure 4.3: Solution to the very stiff (and strongly non–linear) cases [38] (also see Section 4.2.3). Here, $dx = 0.01, \tau = 10^{-10}, \nu = 0.03, \Delta t/\tau > 10^6$

Left Figure: Initial data in local equilibrium, numerical solution (+) superposed onto the "exact" solution [solid line with $dx = 0.001$], $t = 0.3$.

Right Figure: Initial data not in local equilibrium, numerical solution (+) superposed onto the exact (propagating shock) solution [solid line], $t = 0.5$.

Runge–Kutta Godunov splitting scheme. This combines two explicit steps for the convection terms (using a second–order MUSCL formulation) and three implicit steps for the source terms. If these implicit steps are done iteratively, the expense would increase proportional to the stiffness.

4.3 A Brief Status Report

In Chapter III, a finite volume algorithm to solve stiff linear 2×2 relaxation systems was developed, which incorporated the key features of successful numerical schemes (which were obtained in Chapter II). Surprisingly, the method required very minor changes to account for nonlinearity. The most significant modification was due to the wavespeed no longer being a constant, which led to the need for an "averaged" wavespeed at the interface. Seeing no difference between arithmetic, geometric and Roe averaging for the cases shown here, simplicity and economy dictated the selection of arithmetic averaging.

Further, this variation in wavespeed modified the time step computation in order to satisfy the CFL condition (based on the convection terms alone), which was a simple modification. Now, the algorithm was ready for benchmarking. Fortunately, there were several test cases available [41, 38]; being the best results currently available, they were the natural choice for any comparisons. Note that the results in [41, 38] were compared against fine–mesh computations and hence are assumed to be correct and accurate solutions.

The current results are found to be at least competitive with, and often better than, those shown in [41, 38]. Significantly, this has been achieved at a small fraction of the cost (as mentioned already in Section 4.2.3).

In this Chapter, the practical importance of the PI_μ scheme, and the conservation condition it enforces, was shown for the first time. For the linear cases, there was no difference between the PI and PI_μ schemes to eyeball norm; however, from the nonlinear cases attempted in this chapter, it is clear that the unmodified PI scheme is unsuitable for solving nonlinear relaxation systems.

Another interesting study was on how coarsening or refining in space or time affects the results. This was done for the two very stiff cases studied in Section 4.2.3, centered around the conditions used there. In the case of the single shock shown in Figure 4.3 [right], temporal refinement was more critical than spatial refinement (*i.e.*, spatial refinement or coarsening affected the overshoot only slightly, while temporal refinement caused much larger changes). However, in the case of the smooth initial conditions (case shown in Figure 4.3 [left], the reverse was true (*i.e.*, spatial under–resolution caused much more severe peak loss than temporal under–resolution). This indicates that the macroscale features must first be spatially resolved; after this, temporal resolution becomes critical.

The next step is to extend this algorithm to more complex problems and demonstrate the feasibility of the method. However, a brief excursion will be taken here. The results of the previous chapters coupled with the analysis shown in [41, 38] suggested a new application of the work done so far – a reinterpretation of viscous conservation laws as (possibly stiff) relaxation systems, which is described in the next chapter.

CHAPTER V

SOLUTIONS TO VISCOUS CONSERVATION LAWS VIA EQUIVALENT RELAXATION SYSTEMS

In this chapter, a new approach to solve viscous conservation laws is proposed (preliminary results on this subject appeared in [3]). While this may not seem to be relevant in the context of this thesis, it can be viewed as an interesting application of the methods developed in earlier chapters.

But first, what are the difficulties associated with viscous conservation laws? To use the parlance of relaxation systems, the primary issue is that of stiffness. Consider the simplest possible case, the scalar advection–diffusion equation

$$u_t + au_x = \epsilon u_{xx} \quad , \tag{5.1}$$

where a is the advection speed. As the coefficient of viscosity (ϵ) approaches zero, the equation becomes stiff. This stiffness is usually quantified by the numerical Peclet number $Pe_\Delta = a\Delta x/\epsilon$ (where Δx is the mesh size). Now if we consider the simplest nonlinear example, the viscous Burger's equation

$$u_t + \left(\frac{u^2}{2}\right)_x = \epsilon u_{xx} \quad , \tag{5.2}$$

the problem becomes harder. Our eventual goal is to develop an algorithm to tackle nonlinear systems of viscous conservation laws (*e.g.*, the Navier–Stokes equations).

In typical numerical computations of such systems of equations, a first step is often a Roe–type linearization for the convection terms; thus, the building blocks of the solution algorithm are scalar equations (since the linearized system is diagonalizable), and this chapter will be limited to the successful solution of Equations 5.1 and 5.2.

There are two primary difficulties with viscous conservation laws. First, the issue of stiffness, as mentioned above. For example, the standard numerical scheme for Equation 5.1[1] has a Pe_Δ limit of 2, which is extremely restrictive. A frequent remedy is the use of implicit schemes, but that leads to non–local algorithms, which are difficult to parallelize (an issue that gains importance as we move on to systems of equations and realistic problems). Second, a crucial difficulty arises in the context of adaptive Cartesian–grid methods for the Navier–Stokes equations [22]. These methods are being increasingly used to solve complex flows (such as Euler [6, 26, 59] and MHD [35] flows, among others), and are thought to hold immense promise even for viscous flows (once these problem issues are resolved). Hence, they provide a strong motivation for this work. It was shown in [22] that for such methods, the higher–order (viscous) terms led to an unavoidable loss of accuracy at high Reynolds numbers when the grid was non–smooth, a problem that became particularly acute at cut cells. This led to an impasse, with no obvious remedy. The hope is that these problems may be circumvented by the use of the algorithm proposed here.

A brief description of the general form of nonlinear 2 × 2 relaxation systems is given in Section 5.1, followed by an asymptotic analysis in Section 5.2, showing that the relaxation system is asymptotically equivalent to a scalar conservation law with viscosity. The idea is then to invert this analysis by finding, and using our algorithms to solve, a relaxation system whose solutions are asymptotically close to those of the

---

[1] upwind (central) differencing for the convection (diffusion) terms

conservation law (Section 5.3). Finally, in Section 5.4, several linear and nonlinear test cases are presented, compared against a rival method [76] and analytic solutions.

5.1 The General Form of the 2×2 Relaxation System

All the cases run in Chapters III and IV fell under a general canonical form. This was simply the $p-$system with a source term

$$\begin{aligned} u_t + v_x &= 0 \quad , && (5.3)\\ v_t + [p(u)]_x &= -\frac{1}{\tau}\left[v - g(u)\right] \quad , && (5.4) \end{aligned}$$

which has eigenvalues

$$a_F = \mp \left[p'(u)\right]^{1/2} \quad , \quad a_E = g'(u) \quad .$$

In Section 5.2.1, we will see that the system is stable under Liu's sub–characteristic condition [49]

$$-a_F \leq a_E \leq a_F \qquad \Rightarrow \qquad |a_E| \leq a_F \quad . \tag{5.5}$$

5.1.1 The Linear Case

Equations 5.3 and 5.4 can easily be specialized to the linear case, in which $p(u) = a_F^2 u$, $g(u) = a_E u$, leading to the system

$$\begin{aligned} u_t + v_x &= 0 \quad , && (5.6)\\ v_t + [a_F^2 u]_x &= -\frac{1}{\tau}\left[v - a_E u\right] \quad , && (5.7) \end{aligned}$$

with frozen (equilibrium) eigenvalues a_F (a_E) satisfying Liu's sub–characteristic condition (Equation 5.5). Setting $a_F = \tau^{-1/2}$, $a_E = 0$ recovers the HHE's (of Chapter II).

5.2 Asymptotic Analysis of the Nonlinear 2×2 System

Start with the general 2×2 system (Equations 5.3 and 5.4). Define $x^* = x/\tau$, $t^* = t/\tau$, which eliminates the relaxation time from explicitly appearing in these equations. The resulting equations, dropping the (*) superscript, are

$$u_t + v_x = 0 \quad , \tag{5.8}$$

$$v_t + [p(u)]_x = -[v - g(u)] \quad . \tag{5.9}$$

Now, let

$$v = g(u) + v_1 \quad , \tag{5.10}$$

where it is assumed that v_1 (a low frequency component) is small and that its derivatives (*i.e.*, v_{1t}, v_{1x}) are even smaller [19]. Then,

$$v_t = [g(u)]_t = g'(u) \cdot u_t \quad , \tag{5.11}$$

$$v_x = [g(u)]_x = g'(u) \cdot u_x \quad , \tag{5.12}$$

and Equations 5.9 and 5.10 give

$$v_1 = -v_t - [p(u)]_x \quad , \quad \text{where} \quad [p(u)]_x = p'(u) \cdot u_x \quad . \tag{5.13}$$

Substituting Equation 5.12 into Equation 5.8 gives

$$u_t = -v_x = -g'(u) \cdot u_x \quad ,$$

which is substituted into the expression for v_t (Equation 5.11) to get

$$v_t = g'(u) \cdot u_t = -[g'(u)]^2 u_x \quad .$$

This can be substituted into Equation 5.13 to give an expression for $v_1(u)$, which is

$$v_1 = -v_t - [p(u)]_x = [g'(u)]^2 u_x - [p(u)]_x = \left\{[g'(u)]^2 - p'(u)\right\} u_x \quad .$$

Thus, Equation 5.10 becomes

$$v = g(u) + v_1 \quad = \quad g(u) + \left\{[g'(u)]^2 - p'(u)\right\} u_x \quad , \tag{5.14}$$

which can be substituted into Equation 5.8 to give

$$u_t + v_x \quad = \quad 0 \quad \Rightarrow \quad u_t + [g(u)]_x = \left[\left\{p'(u) - [g'(u)]^2\right\} u_x\right]_x \quad . \tag{5.15}$$

Equations 5.14 and 5.15 are the desired asymptotic equations.

Recall that these equations are in transformed variable space $t/\tau, x/\tau$; however, we can (trivially) back-transform them into into the original coordinates to obtain

$$v \quad = \quad g(u) + \tau \left\{[g'(u)]^2 - p'(u)\right\} u_x \quad , \tag{5.16}$$

$$u_t + [g(u)]_x \quad = \quad \tau \left[\left\{p'(u) - [g'(u)]^2\right\} u_x\right]_x \quad . \tag{5.17}$$

5.2.1 The Linear Case

This asymptotic analysis is easily specialized to the general linear case. Setting $p = a_F^2 u, \; g = a_E u$ in Equations 5.16 and 5.17 gives

$$v \quad = \quad a_E u + \tau \left[a_E^2 - a_F^2\right] u_x \quad , \tag{5.18}$$

$$u_t + a_E u_x \quad = \quad \tau \left[a_F^2 - a_E^2\right] u_{xx} \quad . \tag{5.19}$$

Note that Equation 5.19 has stable solutions only if Equation 5.5 is satisfied (this can be interpreted as the RHS representing positive dissipation which leads to decaying solutions, or alternatively, as satisfying causality [79]).

The Hyperbolic Heat Equations

For the case of the HHE's ($a_F = \tau^{-1/2}, \; a_E = 0$), the result is a little different, because the frozen wavespeeds are a function of τ. Proceeding as in Section 5.2 gives

$$v \quad = \quad -u_x + \tau u_{xxx} \quad , \tag{5.20}$$

$$u_t \quad = \quad u_{xx} - \tau u_{xxxx} \quad . \tag{5.21}$$

Recently [41, 38, 15], schemes have been developed that solve relaxation systems by building in this asymptotic information. Here, however, these relations are being used to better design methods to solve viscous conservation laws.

5.3 A Solution Algorithm for Viscous Conservation Laws

Notice that for the general 2 × 2 case (Equations 5.3 and 5.4), the long–time asymptotics were given by Equations 5.16 and 5.17. Equation 5.17 is clearly seen to resemble the viscous Burger's equation, for appropriate choices of g and p.

Recall the derivation of the HHE's, as seen from the following (slightly different) perspective. Start with the heat equation $\theta_t = \theta_{xx}$. To get rid of the higher–order derivative from this equation, set

$$q = -\theta_x \quad , \tag{5.22}$$

resulting in

$$\theta_t = -q_x \quad \Rightarrow \quad \theta_t + q_x = 0 \quad ,$$

which is nothing but Equation 2.3. Next, write a relaxation equation for q as

$$q_t = -\frac{1}{\tau}\left(q - q_{\mathrm{eq}}\right) \quad .$$

But, $q_{\mathrm{eq}} = -\theta_x$ (from Equation 5.22), which gives

$$q_t = -\frac{1}{\tau}\left(q + \theta_x\right) \quad \Rightarrow \quad q_t + \frac{1}{\tau}\theta_x = -\frac{q}{\tau} \quad ,$$

which is precisely Equation 2.4.

A two–step method is proposed here that uses the two preceding ideas (asymptotics and removal of higher–order derivatives) as building blocks. Consider a general[2]

---

[2] the flux $g(u)$ is not specified as yet

scalar viscous conservation law given by

$$u_t + [g(u)]_x = \epsilon u_{xx} \quad . \tag{5.23}$$

The first step is to transform this equation into a 2×2 relaxation system, similar to the derivation of the HHE's. Define a new variable v such that

$$u_t + v_x = 0 \qquad \Rightarrow \qquad v = g(u) - \epsilon u_x \quad . \tag{5.24}$$

Now, write a relaxation equation for v (as done in [63] and above, for the HHE's) as

$$v_t = -\frac{1}{\tau}(v - v_{\mathrm{eq}}) \quad ,$$

where $v_{\mathrm{eq}} = g(u) - \epsilon u_x$ and τ is a parameter that will be determined by stability. This gives

$$v_t + \frac{\epsilon}{\tau} u_x = -\frac{1}{\tau}(v - g) \quad . \tag{5.25}$$

Thus, our transformed system can be written in the form of Equations 5.3 and 5.4, with $p = \epsilon u/\tau$. We know (from Section 5.2) that the asymptotics for this relaxation system are given by Equation 5.17, which needs to be matched to the conservation law (Equation 5.23). Thus,

$$\tau \left\{ p'(u) - [g'(u)]^2 \right\} = \epsilon \qquad \Rightarrow \qquad p = \int \left[\frac{\epsilon}{\tau} + (g')^2 \right] du \quad . \tag{5.26}$$

Finally, the relaxation system that solves the desired conservation law (Equation 5.23) is given by Equations 5.3 and 5.4, with $p(u)$ given by Equation 5.26.

For the viscous Burger's equation,

$$g(u) = \frac{1}{2}u^2 \qquad ; \qquad p(u) = \frac{\epsilon}{\tau}u + \frac{1}{3}u^3 \quad ,$$

while for the advection–diffusion equation

$$g(u) = au \qquad ; \qquad p(u) = \left[\frac{\epsilon}{\tau} + a^2 \right] u \quad .$$

Note that for stability,

$$[g']^2 \leq \epsilon/\tau \qquad \Rightarrow \qquad \tau \leq \frac{\epsilon}{[g']^2} \quad ,$$

which is Liu's sub–characteristic condition [49] and gives the restriction on τ. In the actual implementation, τ is chosen to be a little smaller (approximately 70%) than this maximum allowable value.

Next, some cases are presented for the advection–diffusion equation and the viscous Burger's equation, and are compared against analytical solutions.

5.4 Test Cases

5.4.1 A Linear Case: The Advection–Diffusion Equation

Some results for the advection–diffusion equation were recently presented in [76], using viscous traveling waves as a building block in the solution algorithm. This is an expensive algorithm; however, the results achieved are very good. The present formulation, which uses a transformation of the viscous conservation law to a relaxation system, results in a very efficient algorithm. Solutions for four test cases are presented in Figure 5.1, where the solid line is the analytical solution and the symbols (+) are the computed results. In each case, the boundary conditions are $[u_{-\infty} = 4, u_{+\infty} = 2]$. The equation specific (c, ϵ) and computation specific (τ, dx, ν, t) data are given on the plots themselves, as are the stiffness factors and the Peclet numbers.

The exact solution with these boundary conditions is well known, being

$$u(x,t) = A + B \cdot \mathrm{erf}\left[\frac{x}{2(\nu\ t)^{1/2}}\right] \quad ,$$

$$A = \left(\frac{u_{-\infty} + u_{+\infty}}{2}\right) \quad ; \quad B = \left(\frac{u_{-\infty} - u_{+\infty}}{2}\right) \quad .$$

The error function has been evaluated using a polynomial approximation [1].

In the first three cases (presented in [76]), stiffness factors and Peclet numbers are quite modest and the numerical results are seen to be in good agreement with analytical ones and competitive with those in [76]. The last case has a large Peclet number ($Pe_\Delta = 1000$): although there is a (barely perceptible) overshoot, the results are very good. Notice that the case with $Pe_\Delta = 0$ is nothing but the heat equation ($\theta_t = \epsilon\theta_{xx}$), being solved after the transformation to the relaxation system (though this system is not the same as the HHE's), which helps to tie up all the work done so far (in this dissertation). It shows that the heat equation could have been solved in this manner, since the proposed finite–volume schemes work equally well on the original problem of heat conduction.

5.4.2 A Nonlinear Case: The Viscous Burger's Equation

Here, a nonlinear viscous conservation law – the viscous Burger's equation – is transformed into a relaxation system and solved. Once again, comparative results are available from [76], where viscous traveling waves were used.

Numerical results (+) are presented in Figure 5.2, superposed onto the analytical solution (solid line). The boundary conditions are $[u_{-\infty} = 2, u_{+\infty} = -1]$, which results in a shock moving to the right at a speed $s = \frac{1}{2}$.

The analytical solution for this shock case is well known to be

$$u(x,t) = A - B \cdot \tanh\left[\frac{B(x-At)}{2\nu}\right] \quad ,$$

$$A = \left(\frac{u_{-\infty} + u_{+\infty}}{2}\right) \quad ; \quad B = \left(\frac{u_{-\infty} - u_{+\infty}}{2}\right) \quad .$$

The case on Figure 5.2 [left], which has a moderate stiffness factor and Peclet number, is competitive with [76]. On Figure 5.2 [right], a case with a large Peclet number ($Pe_\Delta = 5000$) is solved. In both cases, the numerical solutions match the analytical result very well.

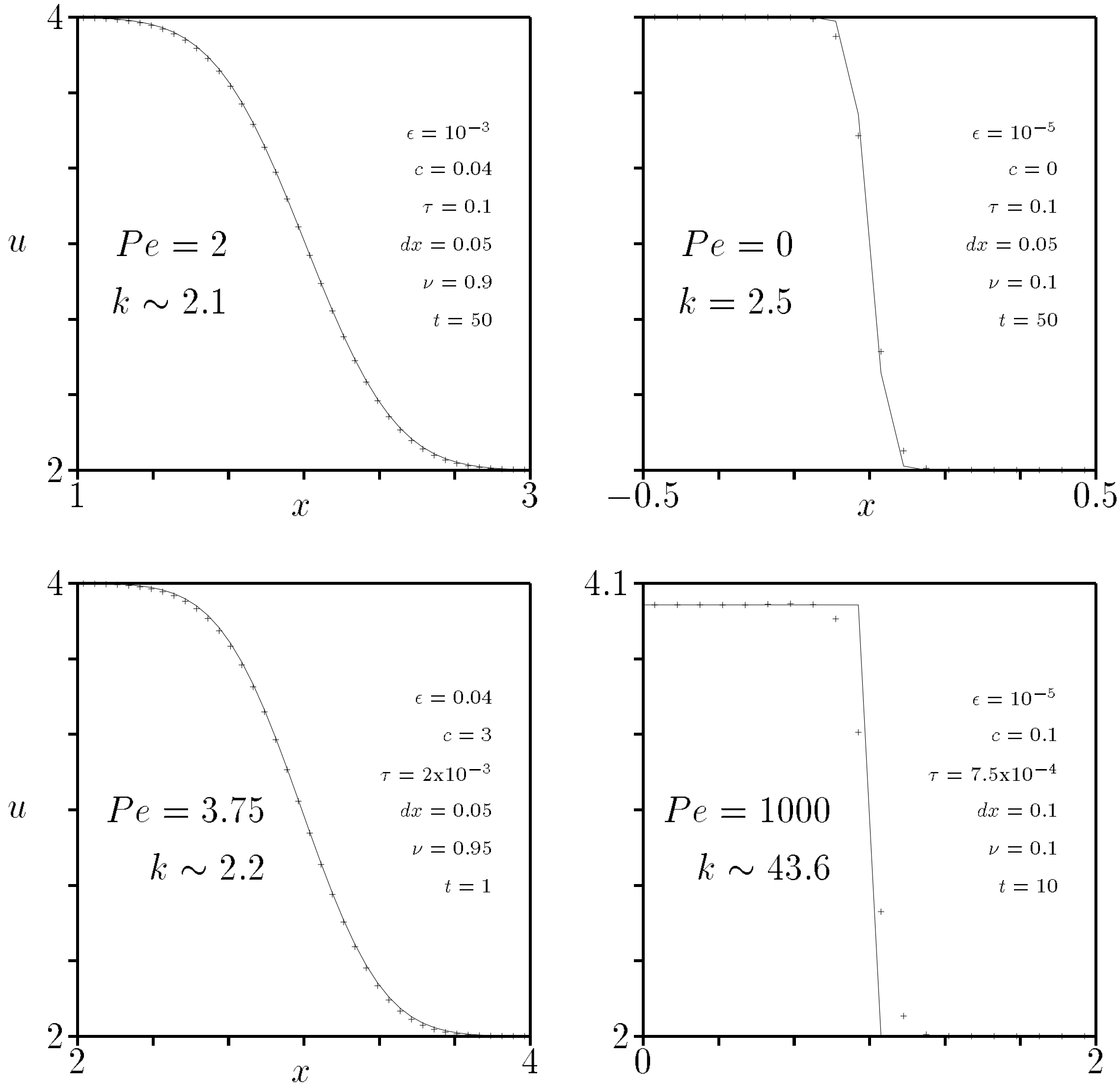


Figure 5.1: Solution (+) to the test cases presented in [76] [$u_{-\infty} = 4, u_{+\infty} = 2$], superposed onto the analytical solution (solid line). In each of these, we're solving Equation 5.1 after converting it into a relaxation system.

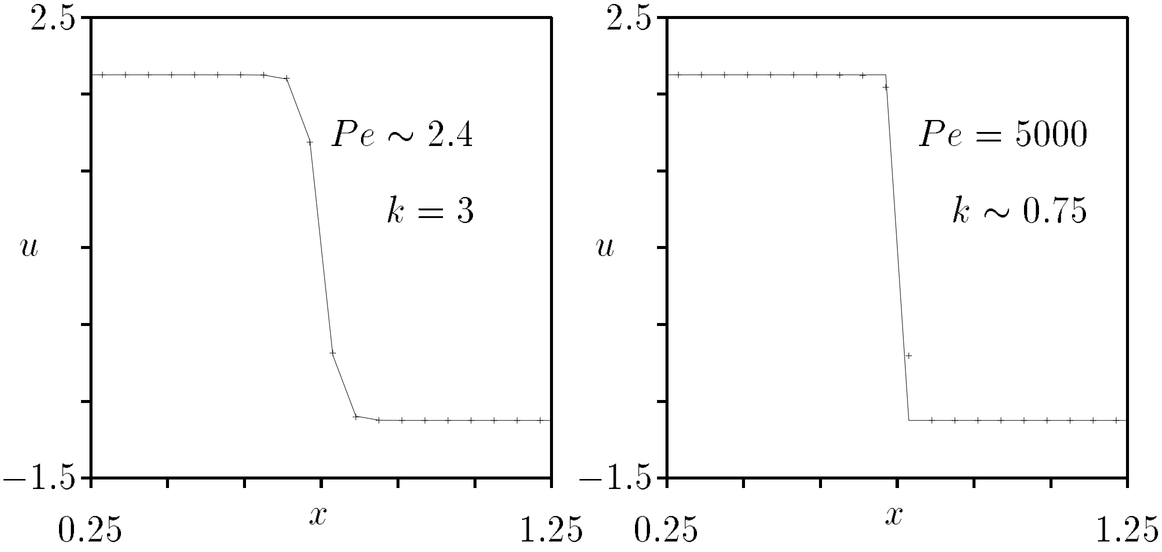


Figure 5.2: Solution (+) to the test cases presented in [76] [$u_{-\infty} = 2, u_{+\infty} = -1$], superposed onto the analytical solution (solid line). In each of these, we're solving Equation 5.2 after converting it into a relaxation system.

5.5 Discussion of Results

The goal of this chapter was to develop methods to solve linear and nonlinear viscous conservation laws. It was shown how a scalar viscous conservation law could be transformed into an equivalent relaxation system, and how this system could then be solved by the methods developed in earlier chapters. Once the convection terms are diagonalized (a standard procedure), the equations decouple and the scalar laws become the canonical forms. In terms of obtaining the relaxation system, a little care was needed, making sure that the asymptotics of the system matched those of the conservation law being solved.

Of course, this entire procedure assumed that stiff relaxation systems (such as Equations 5.3 and 5.4) could be solved accurately. This was justified by the promising results shown in Chapters II to IV, as well as those in [41, 38, 15]. But would this method even work? Several cases were shown in Section 5.4 for the advection–diffusion equation and for the viscous Burger's equation. In each case, the numerical computations were in good agreement with analytical solutions (and competitive with a rival method [76]), even for relatively large Peclet numbers (10^3 to 10^4). Moreover, this accuracy has been achieved at a very low cost, since the solver for the relaxation system is a very simple one (as developed so far in Chapters II to IV).

One might ask what has been gained in this exercise. At first sight, it appears that the viscosity stiffness has merely been replaced by the relaxation stiffness. However, there are several advantages to this approach. First, a second–order system has been transformed into a first–order one, which eliminates the problem of evaluating second derivatives on irregular grids [22]. Second, the stiffness is solely due to the source term, which is local; thus, the algorithm is only locally implicit and therefore parallelizable. Further, a practical advantage is that the solver for relaxation systems has not been specified – it can be chosen by the user. Finally, note that the time–step is now governed by the convection term alone. These advantages, coupled with the results shown already, indicate that the proposed approach is a viable one for solving scalar viscous conservation laws, and should extend quite readily to the case of practical systems, both for steady and unsteady applications.

CHAPTER VI

SOME REALISTIC APPLICATIONS

In this chapter, the methodology developed in earlier chapters will be extended to solve (model) problems that are closer to reality. The Broadwell model is considered first (Section 6.1), followed by the Euler equations with Heat Transfer (Section 6.2). Finally, a simple example from higher–order moment–closure formulations (which are currently of interest for rarefied–gas dynamics calculations) – the one–dimensional 10–moment model – is attempted (Section 6.3). The present method is demonstrated to be a viable alternative, even when the the shock structure is not resolved spatially.

6.1 The Broadwell Model

The Broadwell model [11] is a simplified model of the Boltzmann equations, and is a first step towards more complex discrete velocity kinetic gas models of fluid flow. It describes the gas composed of discrete velocities (four and six in two– and three–dimensions), governed by a binary collision law and limited to spatial variations in one direction only. The evolution equation for the one dimensional solutions is [38]

$$\mathbf{w}_t + \mathbf{f}_x = -\frac{1}{\tau}\mathbf{s} \quad ,$$

where $\mathbf{w} = (h_+, h_0, h_-)^T$, $\mathbf{f} = (h_+, 0, h_-)^T$, $\mathbf{s} = (q, -q, q)^T$ are the state, flux and source vectors respectively, τ is the relaxation distance (mean free path), and

$q = h_+h_- - (h_0)^2$ is a measure of departure from equilibrium (in characteristic coordinates). The state vector $\mathbf{w}$ denotes the mass densities of the gas particles with speeds $(1, 0, -1)$ respectively, in space and time (x, t). These equations are already in diagonal (characteristic) form.

The fluid dynamic (conserved) variables density (ρ) and momentum (m), as well as a third (new) variable z are defined in terms of these mass densities as

$$\rho = h_+ + 2h_0 + h_- \quad , \quad m = h_+ - h_- \quad , \quad z = h_+ + h_- \quad , \tag{6.1}$$

resulting in the conserved variable form of the Broadwell model

$$\mathbf{w}_t + \mathbf{f}_x = -\frac{1}{\tau}\mathbf{s} \quad , \tag{6.2}$$

where $\mathbf{w} = (\rho, m, z)^T$, $\mathbf{f} = (m, z, m)^T$, $\mathbf{s} = (0, 0, Q)^T$ are the state, flux and source vectors respectively, τ is the relaxation distance (mean free path), and $Q = \frac{1}{2}(2\rho z - \rho^2 - m^2)$ is the departure from equilibrium (in conserved variable space)[1].

The reverse transformation, from fluid dynamic variables to the mass density variables, is trivial, given by

$$h_+ = \frac{1}{2}(m + z) \quad , \quad h_0 = \frac{1}{2}(\rho - z) \quad , \quad h_- = \frac{1}{2}(z - m) \quad .$$

In equilibrium,

$$Q \equiv 0 \quad \Rightarrow \quad z = \frac{\rho^2 + m^2}{2\rho} \quad = \quad \frac{\rho}{2}\left(1 + u^2\right) \quad .$$

The characteristic equations can be integrated using the PI_μ scheme to give

$$(m + z)_P - (m + z)_A = -2k\left[\mu Q_P + (1 - \mu)Q_A\right] \quad , \tag{6.3}$$

$$(\rho - z)_P - (\rho - z)_M = k\left[\mu Q_P + (1 - \mu)Q_M\right] \quad , \tag{6.4}$$

$$(z - m)_P - (z - m)_B = -2k\left[\mu Q_P + (1 - \mu)Q_B\right] \quad . \tag{6.5}$$

---

[1] Note that q and Q are (naturally) related, with $Q = 2q$.

Solving the acoustic equations (6.3 and 6.5) simultaneously gives

$$m_P = \frac{1}{2}(m_A + z_A) - \frac{1}{2}(z_B - m_B) + k(1-\mu)(Q_B - Q_A) \quad , \tag{6.6}$$

$$z_P = \frac{1}{2}(m_A + z_A) + \frac{1}{2}(z_B - m_B) - 2k\mu Q_P - k(1-\mu)(Q_A + Q_B) \quad . \tag{6.7}$$

The crucial change from previous examples is that there are now three characteristics, which complicates the solution process. Note that in order to compute the interface flux , m_P and z_P are the only variables needed. While m_P has been explicitly obtained above (in terms of available data), z_P requires (in addition to the values at t_n) information at the interface at the previous time level (which itself would need to be approximated using the values on either side). Further, elimination of the nonlinear dependencies on the values at the new time level would require an expensive and messy iterative solution procedure, contrary to the simple and economical approximation desired. Recall that it was shown in [38] that the numerical scheme must project the solution onto a local Maxwellian (equilibrium) at each time step[2], in order to achieve physical results; further, this must be the first step in a multi–step process (to have the same effect in under–resolved methods as the initial layer).

Observing the difficulty in obtaining an explicit solution, and noting this equilibrium requirement, a natural approximation to make is to assume that equilibrium is achieved at the newly computed point (*i.e.*, $Q_P = 0$). This provides a convenient way to circumvent the difficulty while enforcing the equilibrium condition. Thus, this approximation serves two purposes: first, it avoids the iterative procedure that would have been necessary to solve the pair of nonlinear simultaneous equations, and second, it projects the solution to a local equilibrium as has been shown to be necessary to obtain physical solutions.

---

[2] In the case of under–resolved schemes, this mimics the initial layer. Further, the projection onto equilibrium at each time step ensures passage from the relaxation system to the equilibrium system.

After incorporating these minor changes, the current scheme was ready for benchmarking. Several test cases for the Broadwell model have been presented in [38, 15]. Figure 6.1 shows the results for the first case (initial data as well as all parameters are given on the figure itself), where the plots show the conserved variables (ρ, m, z) as well as $(z - z_{\mathrm{eq}})$. The results obtained here are clearly very good, and the captured shocks are crisp (narrower than those in [38]). The quantity $(z - z_{\mathrm{eq}})$, which measures the departure from a true equilibrium flow, is non–zero in the non–uniform parts of the solution. In [38]), it was proved that for that scheme, the departure would be $O(\tau)$, which appears to be true for the present method also, although the numerical values here are even smaller.

Figure 6.2 shows four more test cases, which are compared to those in [15], with each plot showing the conserved variables (ρ, m, z). Figure 6.2 [left, top] is a case in the frozen regime $(dx, dt \ll \tau)$, where the pure method of characteristics would be expected to do well (and does [15]). Figure 6.2 [right, top], the initial layer is severely under–resolved (thick regime), and schemes that do not mimic this layer (by projecting the solution onto a Maxwellian at each time step) are likely to have a glitch and may even converge to spurious solutions. Recall that the current scheme does not use any sort of initial–layer fix (unlike the method in [15][3]); however, it clearly gives the correct solution although with a small initial layer effect. Figure 6.2 [left, bottom] shows a case in the intermediate regime $(dx, dt \sim \tau)$, resulting in an initial layer which is physical. Unfortunately, the numerical results exaggerate the layer a little. Finally, Figure 6.2 [right, bottom] shows a Riemann problem in the thick regime with a shock and a rarefaction, both of which are very nicely captured.

---

[3] A fix was required in [15] since their first step did not project the data onto a local equilibrium. The fix is to use, for the first time step, a first–order scheme that does possess this projection property and use a Richardson extrapolation to recover second–order accuracy.

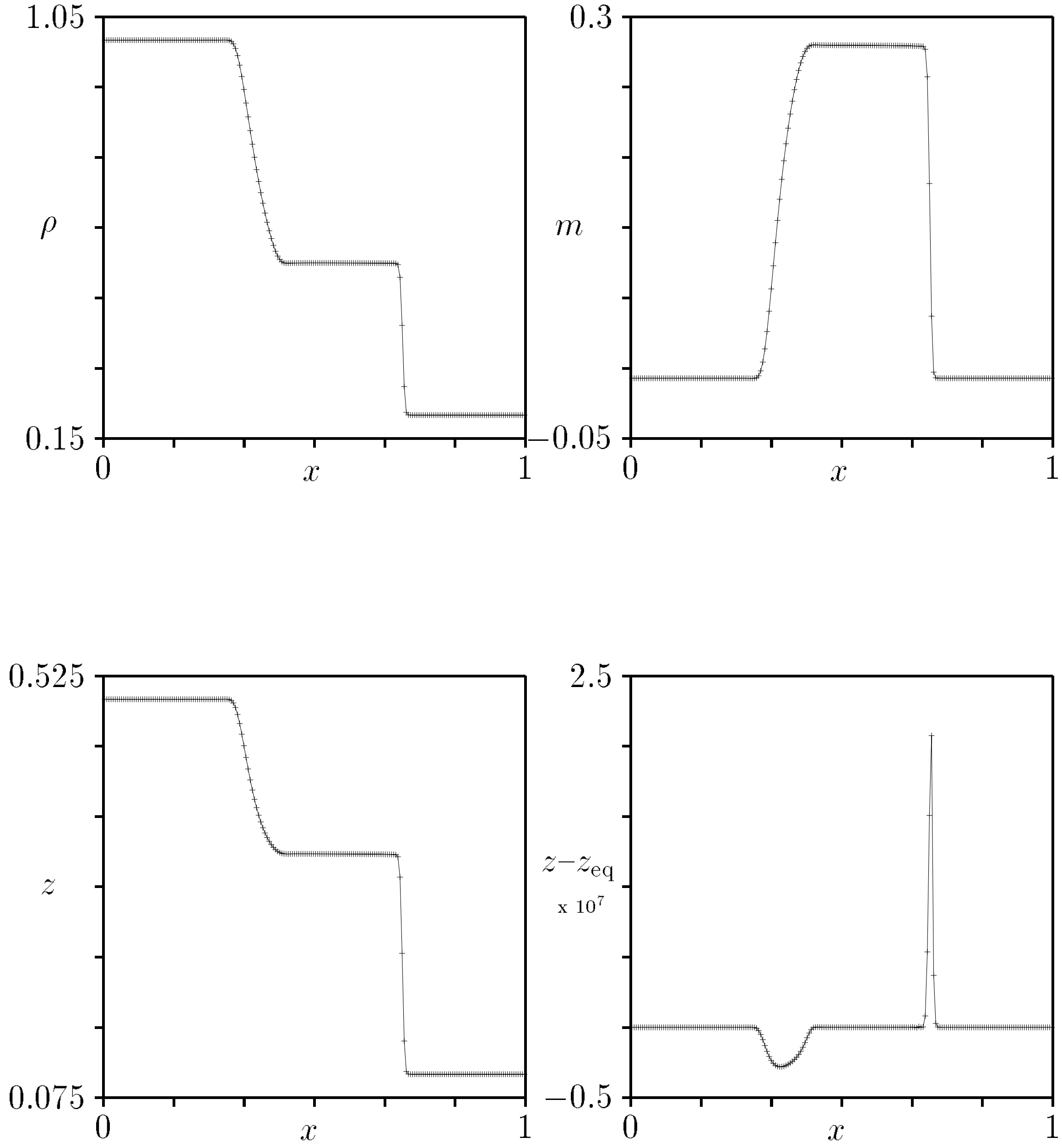


Figure 6.1: Solutions for the conserved variables density ρ, momentum m, z, as well as a departure from equilibrium parameter $z - z_{\text{eq}}$ are plotted for the Broadwell model (Equation 6.2), which may be compared to the results in [38]. Initial data for the states $\mathbf{w} = (\rho, m, z)^T$ are $[\mathbf{w}_L = (1,0,1)^T$, $\mathbf{w}_R = (0.2,0,1)^T]$, the relaxation distance is $\tau = 10^{-8}$, and the solutions are computed using a mesh size of $dx = 0.005$ and a CFL number of $\nu = 0.5$, and shown at a final time $t = 0.25$.

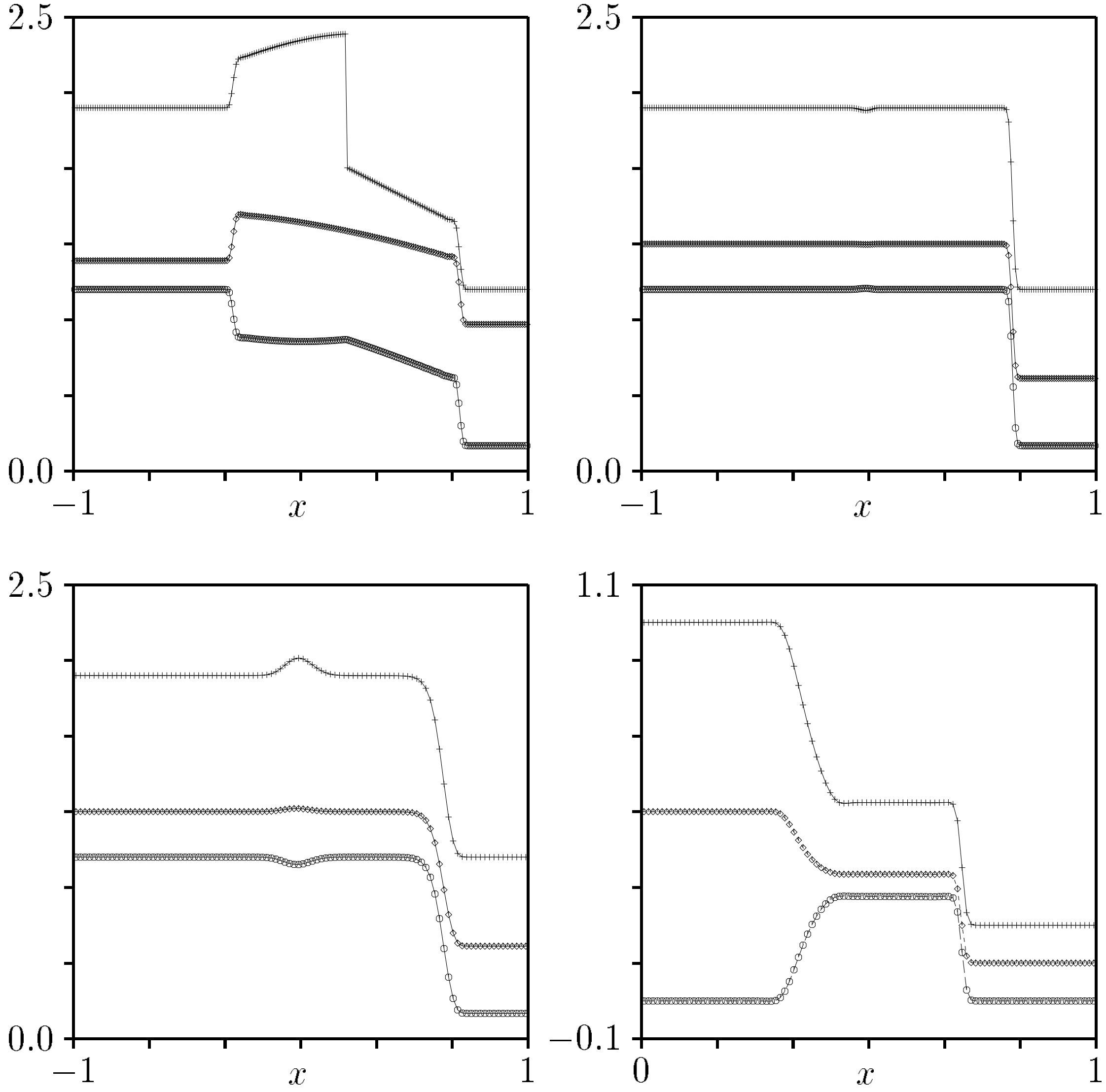


Figure 6.2: Solutions for the conserved variables density ρ (+), momentum m (o) and z ($\diamond$) are plotted for the Broadwell model (Equation 6.2) for four different initial conditions, all of which may be compared to the results in [15]. Each solution is plotted at a final time $t = 0.5$, and computed with a CFL number $\nu = 0.5$. The data used, the relaxation distance τ and the mesh size dx are given below.

Initial data A: $\mathbf{w}_L = (1,1,1)^T$, $\mathbf{w}_R = (1,0.13962,1)^T$, $x < 0.2$.
Initial data B: $\mathbf{w}_L = (1,0,1)^T$, $\mathbf{w}_R = (0.2,0,1)^T$, $x > 0.2$.

Left Top: Initial data A, $\tau = 1$, $dx = 0.01$.
Right Top: Initial data A, $\tau = 10^{-8}$, $dx = 0.01$.
Left Bottom: Initial data A, $\tau = 0.02$, $dx = 0.02$.
Right Bottom: Initial data B, $\tau = 10^{-8}$, $dx = 0.01$.

In fact, the present method results in sharper shocks than the scheme presented in [15].

6.2 The Euler Equations with Heat Transfer

In this model, the Euler equations for gas dynamics are coupled with a simplified heat transfer rate equation [55, 38], where the one dimensional flow of gas is assumed to be in contact with a constant temperature bath. The evolution equations are

$$\mathbf{w}_t + \mathbf{f}_x = -\frac{1}{\tau}\mathbf{s} \quad , \tag{6.8}$$

where $\mathbf{w} = (\rho, \rho u, \rho E)^T$, $\mathbf{f} = (\rho u, \rho u^2 + p, \rho u H)^T$, $\mathbf{s} = [0, 0, \rho(T - T_0)]^T$ are the state, flux and source vectors respectively, τ is the relaxation time, $E = e + \frac{1}{2}u^2$ is the total specific energy, e is the specific internal energy, and $H = E + p/\rho$ is the total specific enthalpy. The closure equation is the usual $\gamma-$gas law

$$p = (\gamma - 1)\rho e \quad ,$$

where p is the pressure of the gas. The specific internal energy is defined as

$$e = c_v T \quad ; \quad c_v = \frac{R}{\gamma - 1} \quad ,$$

where R is the gas constant, $\gamma = c_p/c_v = 1.4$ is the ratio of specific heats, and c_p, c_v are the specific heats of the gas. The equilibrium limit for Equation 6.8 is simply the isothermal Euler equations.

Here, a special case

$$c_v = 1 \quad \Rightarrow \quad R = \gamma - 1 = 0.4 \quad ,$$

is chosen, and the temperature of the bath is set at

$$T_0 = 1 \quad .$$

The simplest manner in which to obtain the characteristics equations is to work with the primitive variable formulation. The transformed equations can be easily obtained (with a little algebra), being of the form of Equation 6.8, but with $\mathbf{w} = (\rho, u, p)^T$, $\mathbf{s} = [0, 0, (\gamma - 1)\rho(T - T_0)]^T$ and

$$\mathbf{A} = \frac{\partial \mathbf{F}}{\partial \mathbf{w}} = \begin{pmatrix} u & \rho & 0 \\ 0 & u & 1/\rho \\ 0 & \gamma p & u \end{pmatrix} .$$

The (frozen) eigenvalues of this system are

$$\lambda_F = u - c,\ u,\ u + c \quad ; \quad c = \left(\frac{\gamma p}{\rho}\right)^{1/2} ,$$

and the matrix of left eigenvectors is

$$\mathbf{L} = \begin{pmatrix} 0 & \rho c & -1 \\ c^2 & 0 & -1 \\ 0 & \rho c & 1 \end{pmatrix} .$$

The characteristic equations, obtained by pre–multiplying Equation 6.8 by **L**, are

$$\hat{\rho}\hat{c}du - dp = -4k(\gamma - 1)\rho(T - T_0) \quad \text{along} \quad \lambda_1, \tag{6.9}$$

$$\hat{c}^2 d\rho - dp = -4k(\gamma - 1)\rho(T - T_0) \quad \text{along} \quad \lambda_2 \ , \tag{6.10}$$

$$\hat{\rho}\hat{c}du + dp = 4k(\gamma - 1)\rho(T - T_0) \quad \text{along} \quad \lambda_3 \ , \tag{6.11}$$

where the eigenvalues are numbered from smallest to largest.

Solving the acoustic equations (6.9 and 6.11) simultaneously gives

$$u_P = \left(\frac{u_A + u_B}{2}\right) + \left(\frac{p_A - p_B}{2\hat{\rho}\hat{c}}\right) + \frac{2k(1-\mu)(Q_A - Q_B)}{\hat{\rho}\hat{c}} \ , \tag{6.12}$$

$$p_P = \left(\frac{p_A + p_B}{2}\right) + \frac{\hat{\rho}\hat{c}(u_A - u_B)}{2} + 4k\mu Q_P + 2k(1-\mu)(Q_A + Q_B) \ . \tag{6.13}$$

Again, the projection onto equilibrium is enforced by assuming that $Q_P = 0$ in the above equation (as was done for the Broadwell model). The density can now be computed as

$$\rho = \rho_M - \left[\frac{\hat{\rho}(u_P - u_A)}{\hat{a}}\right] + \left[\frac{(p_A - p_M)}{\hat{a}^2}\right] + \left[\frac{4k(1-\mu)(Q_A - Q_M)}{\hat{a}^2}\right] \quad . \tag{6.14}$$

Note that for the Euler equations, we cannot say *a priori* which characteristics are positive or negative. Thus, the characteristic tracing step now includes the additional computation of the characteristic velocities (and directions). Now, the characteristics are traced back to the initial line, but take into account (via simple bookkeeping) whether the characteristic intersects the cell to the left or right of the interface (at the initial time level t_n).

To test the accuracy of the present method, the case shown in [38] is attempted, with numerical results shown in Figure 6.3 for the primitive variables (ρ, u, p) and for the departure from equilibrium: $(E - E_0)$. The initial data as well as other relevant parameters are shown on the figure itself. The solution has a left rarefaction and a right shock. The computation is in the thick regime (severe spatial and temporal under–resolution). A comparison of the solutions obtained here against those shown in [38] clearly demonstrates the accuracy of the present method. There are no overshoots at the tail of the rarefaction and the captured shock is extremely sharp (3 cells, as opposed to about 6 in [38], which is easily seen from the plot of $E - E_0$). Further, this departure from equilibrium is again $O(\tau)$ in the current method, far smaller than in [38].

The experience gained from the last two sections provided the confidence to attempt a case of particular interest to the CFD group at Michigan[4] – that of rarefied–

---

[4] The group at the W. M. Keck Foundation Laboratory for CFD at The University of Michigan–Ann Arbor.

gas dynamics and the higher–order moment–closure models [34, 13]. The simplest (practical) model is the 1–D 10–moment model, which is attempted next.

6.3 Higher–Order Moment–Closure Models: An Example from the One–Dimensional 10–Moment Equations

The 10–Moment model [13], specialized to the one–dimensional case, is

$$\mathbf{w}_t + \mathbf{f}_x = -\frac{1}{\tau}\mathbf{s} \quad . \tag{6.15}$$

$\mathbf{w} = (\rho, \rho u, \rho u^2 + p, z)^T$, $\mathbf{f} = (\rho u, \rho u^2 + p, \rho u^3 + 3up, uz)^T$, $\mathbf{s} = [0, 0, \frac{2}{3}(p - z), \frac{1}{3}(z - p)]^T$ being the state, flux and source vectors respectively, and τ being the relaxation time. The notation here is slightly different notation from [13], the correspondence being

$$u \equiv u_x \quad ; \quad p \equiv p_{xx} \quad ; \quad z \equiv p_{nn} \quad .$$

Observe that the equations in this form do not reduce to the correct equilibrium limit as $\tau \rightarrow 0$ (the one–dimensional Euler equations). This is because although there are several ways to write the energy equation (via linear transformations of Equation 6.15), only one reduces to the Euler limit as $\tau \rightarrow 0$. This form is obtained by adding the third and twice the fourth equations (of Equation 6.15), which eliminates the source term in the resulting equation. The modified equations are of the form of Equation 6.15, but with $\mathbf{w} = (\rho, \rho u, \rho u^2 + p + 2z, z)^T$, $\mathbf{f} = (\rho u, \rho u^2 + p, \rho u^3 + 3up + 2uz, uz)^T$, $\mathbf{s} = [0, 0, 0, \frac{1}{3}(z - p)]^T$. The system is now in the form of Liu's [49] relaxation system, and does go to the correct Euler limit as $\tau \rightarrow 0$.

The equations in primitive variable form are given by Equation 6.15, but with

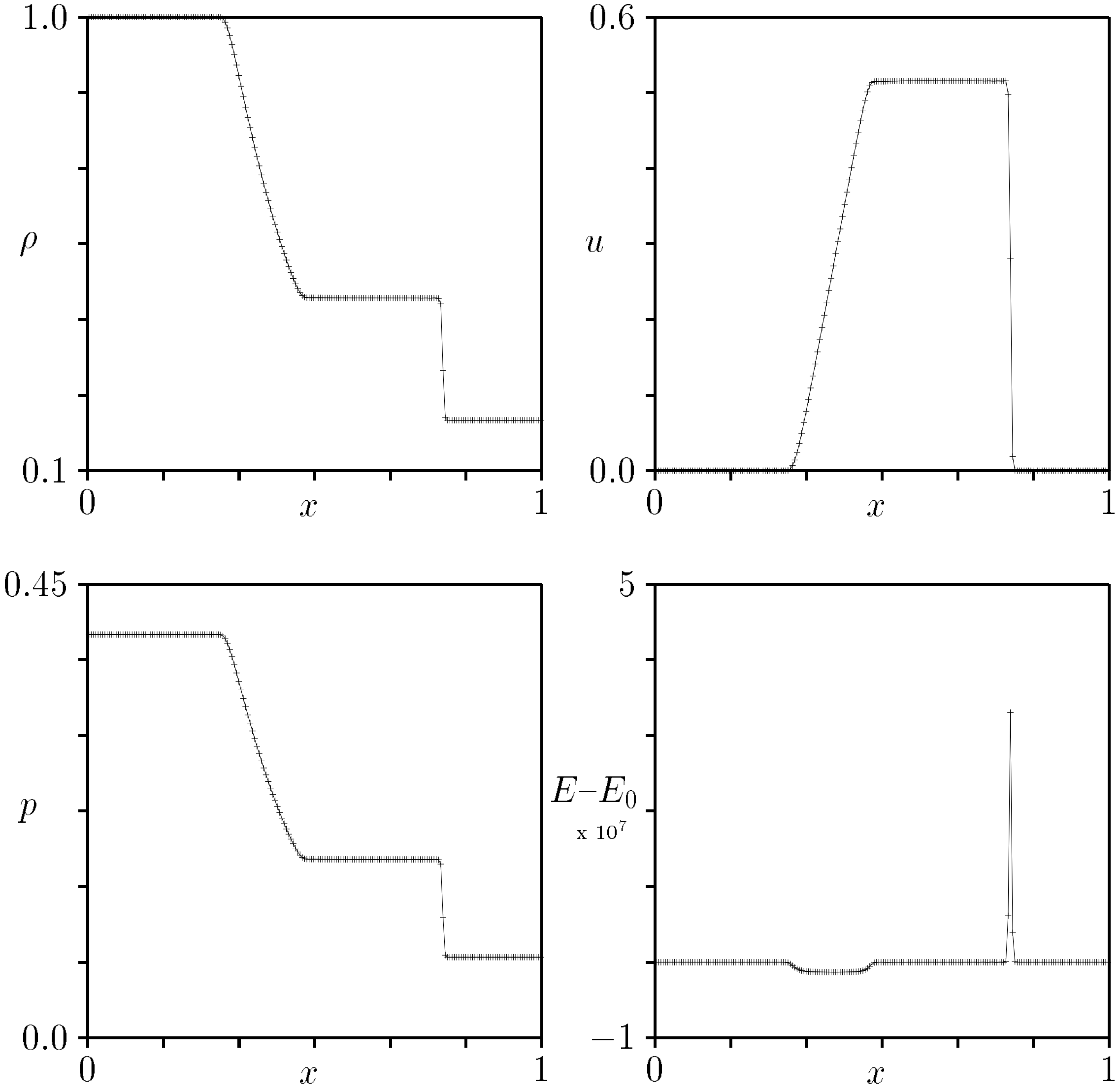


Figure 6.3: Solutions for the primitive variables density ρ, velocity u, pressure p, as well as departure from equilibrium $E - E_0$ are plotted for the Euler equations with heat transfer (Equation 6.8), which may be compared to the results in [38]. Initial data for the states $\mathbf{w} = (\rho, \rho u, \rho E)^T$ are $[\mathbf{w}_L = (1, 0, 1)^T,\ \mathbf{w}_R = (0.2, 0, 0.2)^T]$, the relaxation time is $\tau = 10^{-8}$, and the solutions are computed using a mesh size of $dx = 0.005$ and a CFL number of $\nu = 0.5$, and shown at a final time $t = 0.3$.

$\mathbf{w} = (\rho, u, p, z)^T$, $\mathbf{s} = [0, 0, \frac{2}{3}(p-z), \frac{1}{3}(z-p)]^T$ and

$$\mathbf{A} = \frac{\partial \mathbf{F}}{\partial \mathbf{w}} = \begin{pmatrix} u & \rho & 0 & 0 \\ 0 & u & 1/\rho & 0 \\ 0 & 3\rho a^2 & u & 0 \\ 0 & 3\rho b^2 & 0 & u \end{pmatrix} \quad ,$$

where the sound speeds and the (frozen) eigenvalues are

$$a = \left(\frac{p}{\rho}\right)^{1/2} \quad ; \quad b = \left(\frac{z}{\rho}\right)^{1/2} \quad ; \quad \lambda_k = u - \sqrt{3}a,\ u,\ u,\ u + \sqrt{3}a \quad .$$

The matrix of left eigenvectors is

$$\mathbf{A} = \frac{\partial \mathbf{F}}{\partial \mathbf{w}} = \begin{pmatrix} 0 & -\frac{\sqrt{3}\rho}{6a} & \frac{1}{6a^2} & 0 \\ 1 & 0 & -\frac{1}{3a^2} & 0 \\ 0 & 0 & -\frac{b^2}{3a^2} & 1 \\ 0 & \frac{\sqrt{3}\rho}{6a} & \frac{1}{6a^2} & 0 \end{pmatrix} \quad ,$$

leading to the characteristic equations

$$\begin{aligned} dp - \sqrt{3}\rho a du &= -\frac{2dt}{3\tau}(p-z) \quad , \\ dp - 3a^2 d\rho &= -\frac{2dt}{3\tau}(p-z) \quad , \\ b^2 dp - 3a^2 dz &= -\frac{dt}{\tau}\left[\left(\frac{2b^2+3a^2}{3}\right)(p-z)\right] \quad , \\ dp - \sqrt{3}\rho a du &= -\frac{2dt}{3\tau}(p-z) \quad , \end{aligned}$$

which are integrated along the characteristics using the PI_μ scheme to give

$$p_P - p_B - \sqrt{3}\hat{\rho}\hat{a}(u_P - u_B) = -\frac{8}{3}k\left[\mu Q_P + (1-\mu)Q_B\right] \quad , \tag{6.16}$$

$$p_P - p_M - 3\hat{a}^2(\rho_P - \rho_M) = -\frac{8}{3}k\left[\mu Q_P + (1-\mu)Q_M\right] \quad , \tag{6.17}$$

$$\hat{b}^2(p_P - p_M) - 3\hat{a}^2(z_P - z_M) = -\frac{4}{3}k\left(2b^2 + 3a^2\right)\left[\mu Q_P + (1-\mu)Q_M\right] \tag{6.18}$$

$$p_P - p_A + \sqrt{3}\hat{\rho}\hat{a}(u_P - u_A) = -\frac{8}{3}k\left[\mu Q_P + (1-\mu)Q_A\right] \quad . \tag{6.19}$$

The acoustic equations (6.16 and 6.19) can be trivially solved for u_P, p_P giving

$$u_P = \left(\frac{u_A + u_B}{2}\right) + \frac{1}{2\sqrt{3}\hat{\rho}\hat{a}}\left[p_A - p_B - \frac{8}{3}k(1-\mu)(Q_A - Q_B)\right] \quad ,$$

$$p_P = \frac{1}{2}\left[p_A + p_B + \sqrt{3}\hat{\rho}\hat{a}(u_A - u_B) - \frac{8}{3}k\left\{2\mu Q_P + (1-\mu)(Q_A + Q_B)\right\}\right] \quad .$$

As was done in Sections 6.1 and 6.2, the term involving Q_P is dropped to enforce equilibrium. Knowing u_P, p_P, ρ and z can be computed at the new time level (P) as

$$\begin{aligned}
\rho_P &= \rho_M + \frac{1}{3\hat{a}^2}\left[p_B - p_M + \sqrt{3}\hat{\rho}\hat{a}(u_P - u_B) + \frac{8}{3}k(1-\mu)(Q_M - Q_B)\right] \quad , \\
z_P &= z_M + \frac{1}{6\hat{a}^2}\Big\{2\hat{b}^2(p_P - p_M) - (2\hat{b}^2 + 3\hat{a}^2)\left[p_P - p_A + \sqrt{3}\hat{\rho}\hat{a}(u_P - u_A)\right] - \\
&\qquad \frac{8}{3}k(2\hat{b}^2 + 3\hat{a}^2)(1-\mu)(Q_A - Q_M)\Big] \quad .
\end{aligned}$$

Since this is a rather new area of research, there are very few test cases available. The only one attempted here is a steady solution of the shock structure [13]. Although this is not the best case for comparison (since it is a steady problem), it should provide some idea of how the current method would perform on relatively unknown equations. The initial conditions correspond to a steady (equilibrium) shock, with the data given on the figure itself.

Two cases are considered: in the first, conditions similar to those in [13] are used (non–stiff regime). In the second, a very small relaxation time is used (fluid dynamic limit). Both cases are run with absorbing boundary conditions.

The results should be analogous to a comparison of Navier Stokes to Euler solutions respectively. In Figures 6.4 and 6.5, we present the results of our numerical experiments, and see that in the non–stiff regime, our solutions essentially reproduce those of [13] with less spatial resolution. However, in the stiff case, the results look remarkably like those observed in [2, 4], where slowly moving discontinuities left a trail of oscillations. In this case, one would not expect this phenomenon, since the

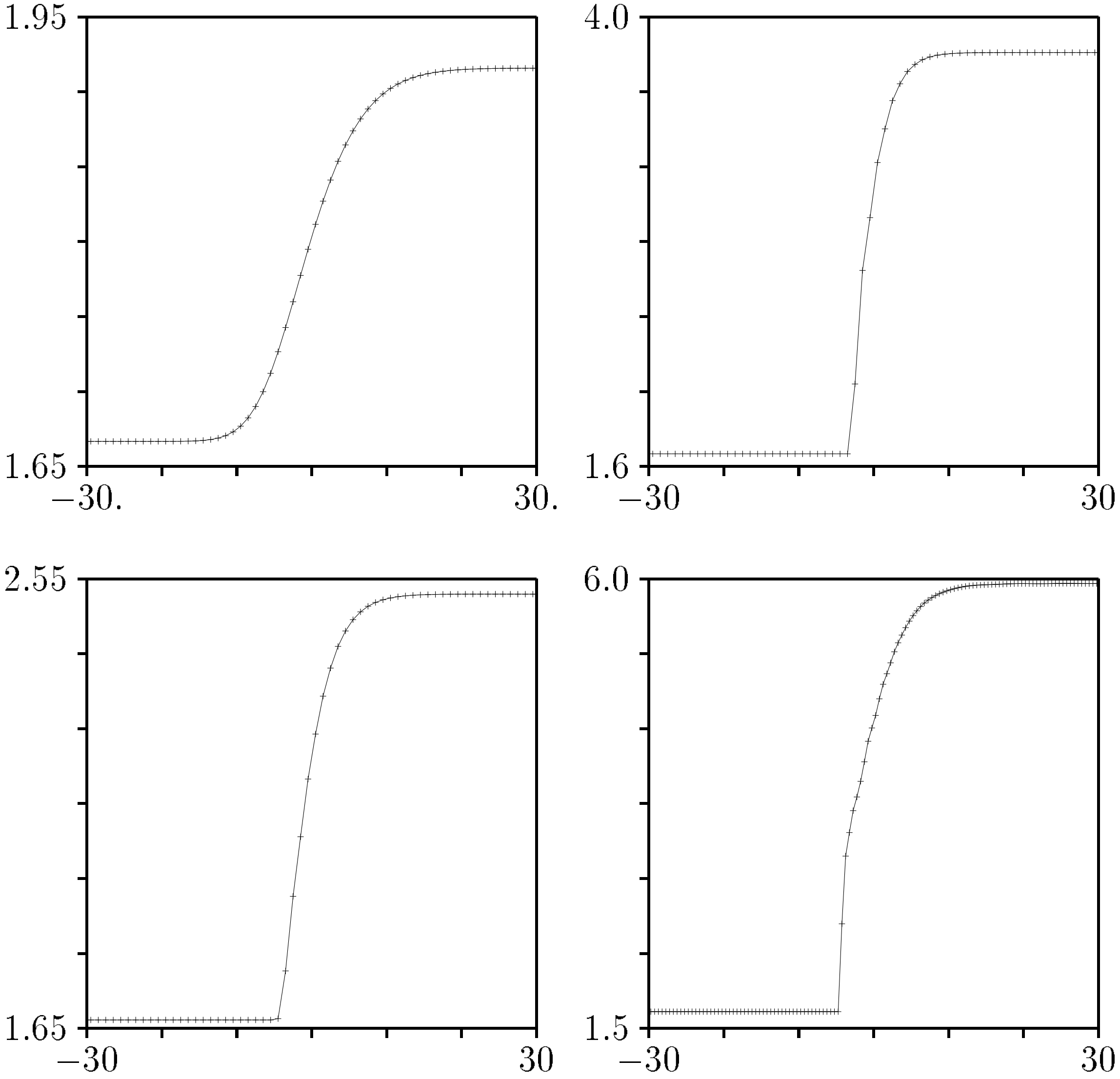


Figure 6.4: Solutions for the density ρ, plotted for the 10–moment equations (Equation 6.15), which may be compared to the results in [13]. Initial data for the left state $\mathbf{w}_L = (\rho, u, p, z)^T = (\frac{5}{3}, M, 1, 1)^T$, where M is the Mach number: $M = 1.1$ (top left), $M = 1.35$ (bottom left), $M = 2$ (top right) and $M = 5$ (bottom right). The relaxation time is $\tau = 1$, and the solutions are computed using a mesh size of $dx = 1$ (except the figure on the bottom right, which uses $dx = 0.5$) and a CFL number of $\nu = 0.5$. All solutions are shown at a final time $t = 100$.

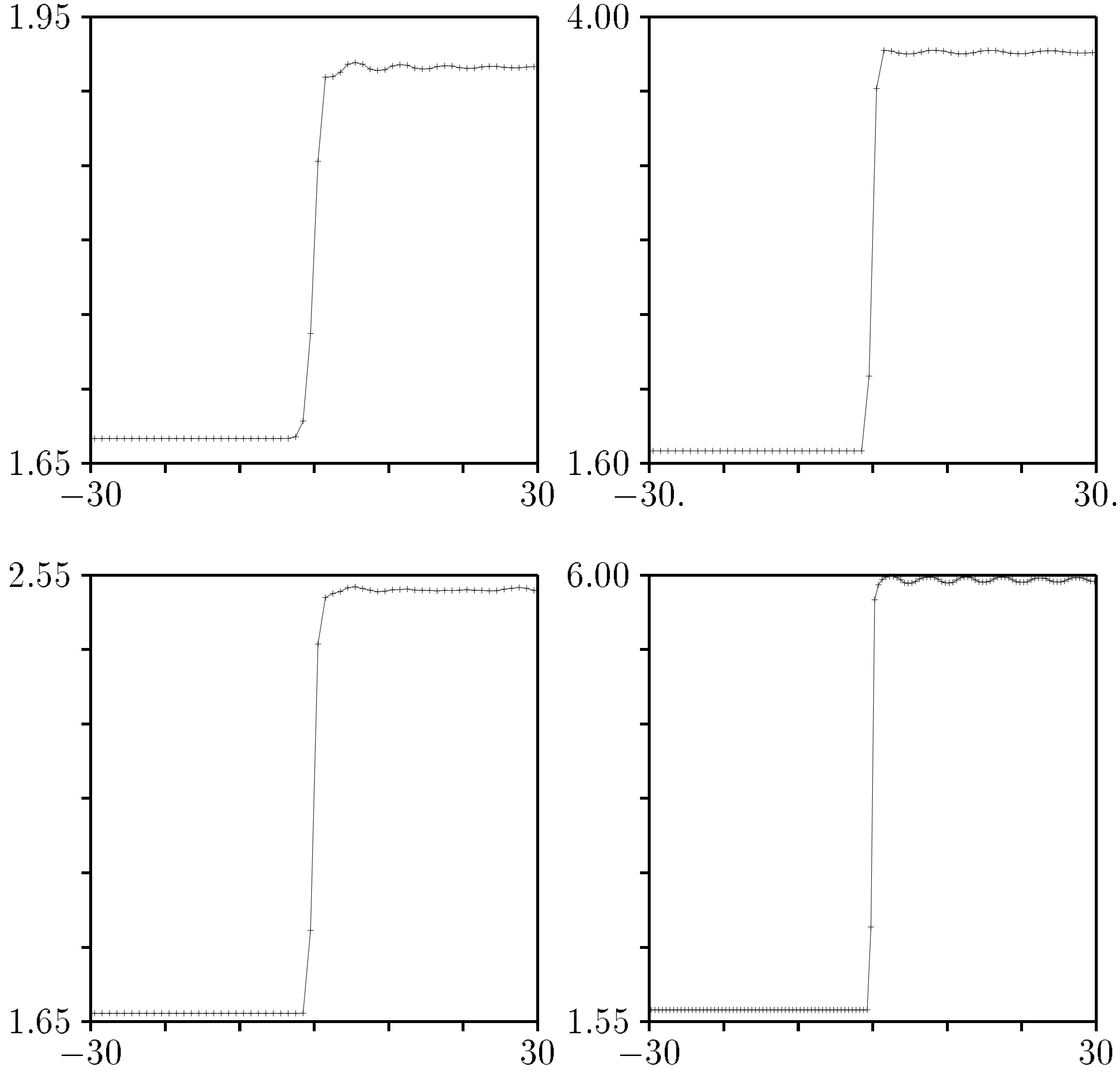


Figure 6.5: Solutions for the density ρ, plotted for the 10–moment equations (Equation 6.15), which may be compared to the results in [13]. Initial data for the left state $\mathbf{w}_L = (\rho, u, p, z)^T = (\frac{5}{3}, M, 1, 1)^T$, where M is the Mach number: $M = 1.1$ (top left), $M = 1.35$ (bottom left), $M = 2$ (top right) and $M = 5$ (bottom right). The relaxation time is $\tau = 10^{-10}$, and the solutions are computed using a mesh size of $dx = 1$ (except the figure on the bottom right, which uses $dx = 0.5$) and a CFL number of $\nu = 0.5$. All solutions are shown at a final time $t = 100$.

shock is stationary. However, it is the *equilibrium* shock that is stationary while a frozen–characteristic method is used; no matter how stiff the test case, the solution will *relax* to equilibrium (by some finite amount), resulting in a numerical shock that is not stationary. Thus, it is quite conceivable that the post–shock phenomenon is back from the crypt [4][5].

---

[5] In [4], it was shown why post–shock oscillations were generated by slowly moving shock waves.

CHAPTER VII

CONCLUSIONS AND FUTURE WORK

The primary goal of this study was to establish a systematic procedure for solving stiff relaxation systems, which has been achieved. The strategy was to start with the simplest possible (model) problem and extend the solution methodology to successively harder, and more realistic, problems.

In Chapter II, a linear 2×2 relaxation model was solved by the method of characteristics. Chapter III described its extension to a finite volume formulation, capable of capturing shocks and satisfying the weak form of the conservation laws (*i.e.*, the integral form). This was further extended in Chapter IV to account for nonlinear effects (the work so far being limited to the 2×2 case). Finally, in Chapter VI, more complex realistic applications were considered. A large number of test–cases were run, most of which have been directly compared to current methods [38, 15, 76]. It was found that the present method is at least competitive with the best solution algorithms currently available. Further, using the analysis and methods developed here, a new approach for solving viscous conservation laws was proposed (Chapter V), which was demonstrated to be a viable alternative to current methods with several added advantages. The most significant gains were the conversion of the higher–order system to a first–order system, which eliminates the problem of accuracy (in computing higher–derivatives) when solving viscous equations on irregular meshes [22][1]. Further, the time–step was dependent on the convection terms alone, and good solutions could be achieved at low cost even for high Peclet numbers.

The method developed is fully discrete, second–order accurate, and performs well even for severely under–resolved calculations (stiffness of one billion). Since it is only point–implicit, it is simple, parallelizable and economical in computational and developmental cost. Moreover, it is very easy to adapt even to relaxation systems for which very little analytical information is available. Furthermore, the captured shocks are typically far narrower than with rival schemes [41, 38, 15], and the CFL (stability) condition is that of the homogeneous system.

The proposed method is a generalization of the point implicit scheme [83], and is stable even for large stiffness (fluid dynamic limit). It does not suffer from the loss of accuracy that is commonly noticed with Strang–splitting in this limit [55, 38]. Based on early lessons, a coupling mechanism is built into the scheme, so that the numerics mimic the physics by letting the wave families interact in the entire spatio–temporal region bounded by the initial line and the domain of dependence. Although it has not been proven, the method very likely projects the data onto an equilibrium solution, since none of the cases attempted so far have led to blow–up or spurious solutions (and it has been shown that lack of this projection property leads to these pitfalls). In the limit of infinite stiffness, the PI$_\mu$ scheme asymptotically approaches the fully point implicit scheme, which automatically satisfies this projection property [38]. Note that while there is no initial–layer fix in the method presented here, the only place where a spurious layer is visible is in the nonlinear Broadwell model (Chapter VI). A successful method for the linear problem, the PI scheme, failed in the nonlinear stiff

---

[1] This still leaves the issue of resolution [22].

cases. It is conjectured that this was due to the method not having the conservation property[2]. For the 10–moment equations, the under–resolved method gave rise to oscillations but these were typically less than 1% of the shock jump for a stiffness of the order of a billion, which is asymptotically the worst case. While the shock in this problem is stationary (based on equilibrium wavespeeds), it is the full relaxation system that is being solved. Therefore, there is a finite relaxation effect even for the most severely under–resolved case, which may lead to a creeping motion of the shock. Slowly–moving shocks also plague Euler solvers (Euler equations being the fluid–dynamic limit of the 10–moment model), and have been studied in detail recently [4]. However, the present results are no worse than Euler results obtained by conventional solvers [4].

The success of the present approach lies in its ability to compute the fluxes more accurately. This is done by using a characteristic tracing step to obtain states at Gauss points along the interface, with the flux integral being approximated by Gaussian quadratures. The characteristic tracing step incorporates the influence of the source terms in a simple manner, leading to a good approximation to the time–varying fluxes. It is seen to be a natural generalization of the Piecewise Steady Approximation (which has been advocated for non–stiff source terms by several researchers, *e.g.*, [71]). Although only a constant mesh size is used, the scheme would readily extend to adaptive–grid methods.

This work also answers the question posed earlier by several researchers (*e.g.*, [55, 41]: which fluxes and wavespeeds, frozen or equilibrium, are the correct ones to use for such computations? It is natural to expect the frozen (equilibrium) fluxes and

---

[2] Note that, intuitively, one would evaluate the source term halfway between the ends of the characteristics for accuracy in the non–stiff cases, while evaluating it at the new time level in the limit of infinite stiffness. These limits are attained quite naturally by the PI_μ scheme, using this conservation property.

wavespeeds to be the correct ones in the respective limits, but the ambiguity in the transition has led to much speculation. The characteristic–based algorithm presented here goes quite naturally from the frozen to the equilibrium limit, but never makes use of equilibrium wavespeeds (*i.e.*, only uses frozen characteristics).

Under–resolved schemes have been shown to be prone to two further shortcomings [38]: first, the generation of spurious initial layers when the initial data are not in local equilibrium, and second, the convergence of the scheme to the wrong (unphysical) solution with (possibly) incorrect shock speeds and jumps. The results presented here adequately demonstrate the robustness of the proposed scheme. It has the correct long–time behavior with both spatial and temporal under–resolution, and the results do not degrade in the intermediate regime. Further, the solutions are always the physical ones, having the correct shock speeds and jumps, without generating noticeable (erroneous) initial layers.

In terms of future work, there are several possibilities. First, an attempt needs to be made to extend this methodology to multi–dimensions. This could be done in a reasonably straightforward manner if dimensional splitting was acceptable. There are two possible ways to use dimensional splitting: first, to use Strang–type splitting for the two directional operators, and second, to solve a one–dimensional problem at each of the four interfaces (which may be better). If, however, the accuracy *does* reduce to first–order, the extension could be much harder than anticipated and a dimensional coupling would need to be sought. It is also not clear how the method could be adapted to genuinely multi–dimensional approaches.

Recall that the 10–moment model was successfully solved for small stiffness, and reasonable estimates were obtained in the fluid–dynamic limit. This suggests a second promising avenue to explore would be to extend this procedure to solve the more complex higher–order moment–closure models. In addition, the two–temperature model of the Euler equations with heat transfer was very similar (at first sight) to models in computational combustion, indicating that numerical combustion may be a third potential area of impact for the current methodology.

A fourth, and extremely promising avenue to explore, would be to attempt to use the new formulation for viscous conservation laws (of Chapter V) to solve the equivalent of the full Navier–Stokes equations, and to compare the algorithm efficiency and accuracy with contemporary methods. Solutions to unsteady problems may be hard, since the correct asymptotics would need to be determined. However, for steady problems, whether or not the asymptotics are correct doesn't matter. Moreover, the advantage of being able to compute higher–Peclet–number flows than conventional methods would speed convergence to steady state, and this is possibly the single most promising area of application.

APPENDICES

APPENDIX A

THE PI$_\mu$ SCHEME

It is easy to show that for any set of linear equations

$$\mathbf{u}_t + (\mathbf{A} \cdot \mathbf{u})_x = \mathbf{P} \cdot \mathbf{u} \quad ,$$

where **A**, **P** are constant matrices, a “conservative” form can be obtained as

$$(e^{-\mathbf{P}t}\mathbf{u})_t + (e^{-\mathbf{P}t}\mathbf{A} \cdot \mathbf{u})_x = 0 \quad ,$$

where $e^{-\mathbf{P}t}$ is a matrix exponential. For the HHE’s, it is simply

$$e^{-\mathbf{P}t} = \mathrm{diag}(1, e^{-t/\tau}) \quad ,$$

resulting in the conservation condition

$$Z(k) = e^{-2k} \quad , \tag{A.1}$$

where $z(k)$ is one of the coefficients that appear in the method of characteristics (see Section 2.3). What is desired here is a closed–form solution for the “best” place at which to evaluate the source term in the context of a point–implicit method of characteristics, the PI$_\mu$ scheme (Equation 2.19). This is determined by satisfying conservation in the above sense for a constant state $\bar{\mathbf{w}} = (\bar{\theta}, \bar{q})^T$, the constraint for which is

$$Z \quad = \quad [1 - 2k(1-\mu)](1+2k\mu)^{-1} \quad = \quad \exp(-2k) \quad .$$

This is easily solved for μ to give

$$\mu = \frac{1}{1 - e^{-2k}} - \frac{1}{2k} \quad , \tag{A.2}$$

which can be expanded in a Taylor series around $k = 0$ to get

$$\mu = \frac{1}{2} + \frac{k}{6} - \frac{k^3}{90} + \frac{k^5}{945} + O(k^6) \quad , \tag{A.3}$$

and asymptotes to

$$\mu(k \to \infty) = \; \uparrow 1 \quad . \tag{A.4}$$

The same analysis holds for the general form of the linear 2 × 2 system, and would hold true for any linear system [check].

From this, two very simple, but nevertheless illuminating, conclusions can be drawn. First, in the limit of $k \to 0$ (*i.e.*, in the limit of vanishing stiffness), $\mu = 0.5$ (the PI scheme, where the source term is evaluated midway along each characteristic). Second, as the stiffness increases, the above analysis suggests evaluating the source term increasingly closer to the new time level until in the limit of $k \to \infty$, $\mu = 1$ (the fully implicit scheme).

APPENDIX B

The GAUSS POINTS AND WEIGHTS

The interface flux that was to be approximated by Gaussian quadratures was (Section 3.4.1)

$$\mathbf{F}_{j+1/2} = \frac{1}{2}\int_{-1}^{1} \mathbf{f}_{j+1/2}(\alpha)d\alpha \quad = \quad \frac{1}{2}\sum_{k=1}^{k=K} w_k \mathbf{f}_{j+1/2}(\alpha_k) \quad ,$$

where

$$\alpha_k = \frac{1}{2}\left(\Delta t \xi_k + \Delta t\right) = \frac{\Delta t}{2}\left(1 + \xi_k\right) \quad .$$

Here, ξ_k and w_k are the location and weight of the k^{th} Gauss point, for a K point quadrature, which are given below for the 3 point quadrature [1].

3 point quadrature: locations and weights

$$\begin{aligned} \xi_1 &= -0.774596669241483 \\ \xi_2 &= 0. \\ \xi_3 &= -\xi_1 \end{aligned}$$

$$\begin{aligned} w_1 &= 0.555555555555556 \\ w_2 &= 0.888888888888889 \\ w_3 &= w_1 \end{aligned}$$

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] M. Abramowitz and I. A. Stegun, eds., *Handbook of Mathematical Functions*, National Bureau of Standards, 10th printing, 1972.

[2] M. Arora and P. L. Roe, *On oscillations due to shock capturing in unsteady flows*, in Proceedings of the 14th International Conference On Numerical Methods in Fluid Dynamics, Springer-Verlag, New York, 1994, to appear.

[3] ——, *Design of algorithms for stiff relaxation systems*, in Proceedings of the 3rd International Congress on Industrial and Applied Mathematics), Akademie–Verlag, Berlin, 1995, to appear.

[4] ——, *On post–shock oscillations due to shock capturing schemes in unsteady flows*, Journal of Computational Physics, (1995, to appear).

[5] E. Becker, *Gas Dynamics*, Academic Press, New York, 1968.

[6] M. J. Berger and J. Oliger, *Adaptive mesh refinement for hyperbolic partial differential equations*, Journal of Computational Physics, 53 (1984), pp. 484–512.

[7] P. L. Bhatnagar, *A model for collisional processes in gases. I. Small amplitude processes in charged and neutral one–component systems*, Physical Review, 94 (1954), pp. 511–525.

[8] P. A. Blythe, *Non-linear wave propagation in a relaxing gas*, Journal of Fluid Mechanics, 37 (1969), pp. 31–50.

[9] A. Bourlioux, A. Majda, and V. Roytburd, *Theoretical and numerical structures for unstable one–dimensional detonations*, SIAM Journal on Applied Mathematics, 51 (1991), pp. 303–343.

[10] M. Brio and C. C. Wu, *An upwind differencing scheme for the equations of ideal magnetohydrodynamics*, Journal of Computational Physics, 75 (1988), pp. 400–422.

[11] J. E. Broadwell, *Shock structure in a simple discrete velocity gas*, Physics of Fluids, 7 (1964), pp. 1243–1247.

[12] L. J. F. Broer, *Characteristics of the equations of motion of a reacting gas*, Journal of Fluid Mechanics, 4 (1958), pp. 276–282.

[13] S. Brown, *The development of Roe-type approximate riemann solvers for the 10–and 35–moment models and the numerical solution of shock structures*, PhD thesis, The University of Michigan, 1996.

[14] R. Caflisch, *Navier–Stokes and Boltzmann shock profiles for a model of gas dynamics*, Communications on Pure and Applied Math, 22 (1979), pp. 521–554.

[15] R. Caflisch, S. Jin, and G. Russo, *Uniformly accurate schemes for hyperbolic systems with relaxation*, SIAM Journal on Numerical Analysis, (1996, to appear).

[16] C. Cattaneo, *A form of heat conduction equations which eliminates the paradox of instantaneous propagation*, Comptes Rendus, 247 (1958), pp. 431–433.

[17] C. Cerecignani, *The Boltzmann Equation and its Applications*, Springer–Verlag, New York, 1988.

[18] G.-Q. Chen, C. D. Levermore, and T.-P. Liu, *Hyperbolic conservation laws with stiff relaxation terms and entropy*, Communications on Pure and Applied Math, 47 (1994), pp. 787–830.

[19] G.-Q. Chen and T.-P. Liu, *Zero relaxation and dissipation limits for hyperbolic conservation laws*, Communications on Pure and Applied Math, 45 (1993), pp. 755–781.

[20] J. F. Clarke, *Gas dynamics with relaxation effects*, Reports on Progress in Physics, 41 (1978), pp. 807–864.

[21] J. F. Clarke and M. McChesney, *Dynamics of Relaxing Gases*, Butterworths, Boston, 1976.

[22] W. L. Coirier, *Simulation of Steady Viscous Flow on an Adaptively Refined Cartesian Grid*, PhD thesis, The University of Michigan, 1994.

[23] P. Colella, A. Majda, and V. Roytburd, *Theoretical and numerical structure for reacting shock waves*, SIAM Journal on Scientific and Statistical Computing, 7 (1986), pp. 1059–1080.

[24] P. Colella and P. R. Woodward, *The piecewise-parabolic method (PPM) for gas-dynamical simulations*, Journal of Computational Physics, 54 (1984), pp. 174–201.

[25] F. Coron and B. Perthame, *Numerical passage from kinetic to fluid equations*, SIAM Journal on Numerical Analysis, 28 (1991), pp. 26–42.

[26] D. De Zeeuw and K. G. Powell, *An adaptively-refined cartesian mesh solver for the Euler equations*, Journal of Computational Physics, 104 (1993), pp. 56–68.

[27] S. Deshpande, *A second order accurate, kinetic–theory based method for inviscid compressible flow*, Tech. Report 2613, NASA Langley, 1986.

[28] D. A. Drew, *Mathematical modelling of two-phase flow*, Annual Review of Fluid Mechanics, 15 (1983), pp. 261–291.

[29] W. Fickett, *Detonation in miniature*, in The Mathematics of Combustion, J. D. Buckmaster, ed., SIAM, Philadelphia, 1985, pp. 133–181.

[30] M. M. Gibson and B. E. Launder, *Ground effects on pressure fluctuations in the atmospheric boundary layer*, Journal of Fluid Mechanics, 86 (1978), pp. 491–511.

[31] H. M. Glaz and T.-P. Liu, *The asymptotic analysis of wave interactions and numerical calculations of transonic nozzle flow*, Advances in Applied Mathematics, 5 (1984), pp. 111–146.

[32] J. Glimm, *The continuous structure of discontinuities*, 344 (1986), pp. 177–186.

[33] T. I. Gombosi, *Gaskinetic Theory*, Cambridge University Press, 1994.

[34] T. I. Gombosi, C. P. T. Groth, P. L. Roe, and S. L. Brown, *35-moment closure for rarefied gases: Derivation, transport equations, and wave structure.* submitted to *Physics of Fluids*, 1994.

[35] T. I. Gombosi, K. G. Powell, and D. L. D. Zeeuw, *Axisymmetric modeling of cometary mass loading on an adaptively refined grid: MHD results.* submitted to *J. Geophys. Research*, 1994.

[36] W. C. Griffith, D. Brickl, and V. Blackman, *Structure of shock waves in polyatomic gases*, Physical Review, 102 (1956), pp. –.

[37] W. C. Griffith and A. Kenny, *On fully-dispersed shock waves in carbon dioxide*, Journal of Fluid Mechanics, pp. 286–289.

[38] S. Jin, *Runge–Kutta methods for hyperbolic conservation laws with stiff relaxation terms*, Journal of Computational Physics, 122 (1995).

[39] S. Jin and C. D. Levermore, *The discrete-ordinate method in diffusive regimes*, Transport Theory and Statistical Physics, 20 (1991), pp. 413–439.

[40] ——, *Fully discrete numerical transfer in diffusive regimes*, Transport Theory and Statistical Physics, 22 (1993), pp. 739–791.

[41] ——, *Numerical schemes for hyperbolic conservation laws with stiff relaxation terms*, Journal of Computational Physics, (1993, submitted).

[42] J. G. Jones, *On the near-equilibrium and near-frozen regions in an expansion wave in a relaxing gas*, Journal of Fluid Mechanics, 19 (1964), pp. 81–102.

[43] E. W. Larsen, *The asymptotic diffusion limit of discretized transport problems*, Nuclear Science and Engineering, 112 (1992), pp. 336–346.

[44] E. W. Larsen, J. E. Morel, and J. W. F. Miller, *Asymptotic solutions of numerical transport problems in optically thick, diffusive regimes*, Journal of Computational Physics, 69 (1987), pp. 283–324.

[45] B. E. Launder and N. Shima, *Second moment closure for the near–wall sublayer: Development and application*, AIAA Journal, 27 (1989), pp. 1319–1325.

[46] P. D. Lax and B. Wendroff, *Systems of conservation laws*, Communications on Pure and Applied Math, 13 (1960), pp. 217–237.

[47] G. Lebon and A. Cloot, *Propagation of ultrasonic sound waves in dissipative dilute gases and extended irreversible thermodynamics*, Wave Motion, 11 (1989), pp. 23–32.

[48] M. J. Lighthill, *Viscosity effects in sound waves of finite amplitude*, in Surveys in Mechanics, G. K. Batchelor and R. M. Davies, eds., Cambridge University Press, London, 1956, pp. 250–351.

[49] T.-P. Liu, *Hyperbolic conservation laws with relaxation*, Communications in Mathematical Physics, 108 (1987), pp. 153–175.

[50] A. Majda, *A qualitative model for dynamic combustion*, SIAM Journal on Applied Mathematics, 41 (1981), pp. 70–93.

[51] F. E. Marble, *Dynamics of dusty gases*, Annual Review of Fluid Mechanics, 2 (1970), pp. 397–446.

[52] F. K. Moore and W. E. Gibson, *Propagation of weak disturbances in a gas subject to relaxation effects*, Journal of Aero/Space Sciences, (1960), pp. 117–127.

[53] H. Ockendon and D. A. Spence, *Non-linear wave propagation in a relaxing gas*, Journal of Fluid Mechanics, 39 (1969), pp. 329–345.

[54] R. B. Pember, *Numerical methods for hyperbolic conservation laws with stiff relaxation, I. Spurious solutions*, SIAM Journal on Applied Mathematics, 53 (1993), pp. 1293–1330.

[55] ——, *Numerical methods for hyperbolic conservation laws with stiff relaxation, II. Higher order Godunov methods*, SIAM Journal on Scientific Computing, 14 (1993), pp. 824–859.

[56] B. Perthame, *Boltzmann type schemes for compressible Euler equations in one and two space dimensions*, SIAM Journal on Numerical Analysis, 27 (1990), pp. 1405–1421.

[57] ——, *Second–order boltzmann schemes for compressible Euler equations in one and two space dimensions*, SIAM Journal on Numerical Analysis, 29 (1992), pp. 1–19.

[58] K. H. Prendergast and K. Xu, *Numerical hydrodynamics from gas-kinetic theory*, Journal of Computational Physics, 109 (1993), pp. 53–66.

[59] J. J. Quirk, *A contribution to the great Riemann solver debate*, International Journal for Numerical Methods in Fluids, 18 (1994), pp. 555–574.

[60] M. Renardy, W. Hrusa, and J. Nehel, *Mathematical Problems in Viscoelasticity*, vol. 35 of Pitman Monographs and Surveys in Pure and Applied Mathematics, Longman Sci. Tech., New York, 1987.

[61] P. L. Roe, *Approximate Riemann solvers, parameter vectors and difference schemes*, Journal of Computational Physics, 43 (1981).

[62] P. L. Roe and M. Arora, *Design of algorithms for a dispersive hyperbolic problem*, in AIAA 10th Computational Fluid Dynamics Conference, AIAA, 1991, pp. 95–105.

[63] ——, *Characteristic-based schemes for dispersive waves I. The method of characteristics for smooth solutions*, Numerical Methods for Partial Differential Equations, 9 (1993), pp. 459–505.

[64] M. Sichel and Y. K. Yin, *Viscous transonic flow in relaxing gases*, Zeitschrift fuer Angewandte Mathematik und Mechanik, 56 (1976), pp. 315–329.

[65] J. Smoller, *Shock Waves and Reaction-Diffusion Equations*, Springer-Verlag, New York, 1983.

[66] J. J. Stoker, *Water Waves*, Wiley, New York, 1958.

[67] G. Strang, *On the construction and comparison of difference schemes*, SIAM Journal on Numerical Analysis, 5 (1968), pp. 506–517.

[68] I. Suliciu, *Modeling phase transitions by rate–type constitutive equations: Shock wave structures*, Int. J. Eng. Sci., 28 (1990), pp. 829–841.

[69] B. van Leer, *Towards the ultimate conservative difference scheme. IV. A new approach to numerical convection*, Journal of Computational Physics, 23 (1977).

[70] ——, *Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method*, Journal of Computational Physics, 32 (1979).

[71] ——, *On the relation between the upwind-differencing schemes of Godunov, Engquist-Osher and Roe*, SIAM Journal on Scientific and Statistical Computing, 5 (1984).

[72] P. VERNOTTE, *Les paradoxes de la théorie continue de l'equation de la chaleur*, Comptes Rendus, 246 (1958).

[73] W. G. VINCENTI AND C. H. KRUGER, *Introduction to Physical Gas Dynamics*, John Wiley, New York, 1965.

[74] R. VON MISES, *Mathematical Theory of Compressible Fluid Flow*, Academic Press, New York, 1958.

[75] G. B. WALLIS, *One-dimensional Two-phase Flow*, McGraw-Hill, New York, 1969.

[76] S. WEEKES, *The Travelling Wave Method for the Navier–Stokes Equations*, PhD thesis, The University of Michigan, Ann Arbor, 1995.

[77] P. P. WEGENER, B.-T. CHU, AND W. A. KLIKOFF, *Weak waves in relaxing flows*, Journal of Fluid Mechanics, 23 (1965), pp. 787–800.

[78] G. B. WHITHAM, *Some comments on wave propagation and shock wave structure with application to magnetohydrodynamics*, Communications on Pure and Applied Math, 12 (1959), pp. 113–158.

[79] ——, *Linear and Nonlinear Waves*, John Wiley, New York, 1974.

[80] D. C. WIGGERT, *Analysis of early-time transient heat conduction by method of characteristics*, Journal of Heat Transfer, 99 (1977), pp. 35–40.

[81] W. W. WOOD, *Characteristic equations for reactive flow*, Journal of Chemical Physics, p. 596.

[82] W. W. WOOD AND J. G. KIRKWOOD, *Hydrodynamics of a reacting and relaxing fluid*, Journal of Applied Physics, 28 (1957), pp. 395–398.

[83] H. C. YEE AND J. L. SHINN, *Semi-implicit and fully implicit shock capturing methods for nonequilibrium flows*, AIAA Journal, 27 (1989), pp. 299–307.

ABSTRACT

EXPLICIT CHARACTERISTIC–BASED HIGH–RESOLUTION ALGORITHMS
FOR HYPERBOLIC CONSERVATION LAWS WITH STIFF SOURCE TERMS

by

Mohit Arora

Chairman: Philip L. Roe

A new algorithm for the numerical solution of stiff hyperbolic relaxation systems is presented. It is an extremely simple characteristic–based, fully–discrete, explicit, second–order scheme that is computationally inexpensive. It is accurate in the non–stiff (rarefied) regime, and maintains its accuracy for moderate stiffness and under–resolution. As the under–resolution becomes more severe, the results gracefully deteriorate, and even in the fluid dynamic limit, it appears to give the physical solution (*i.e.*, correct shock jumps and locations), although under–resolution does introduce errors. Results are presented for a variety of linear and nonlinear test cases, most of which are compared to current literature. In all cases, the present results are at least competitive with the best available solutions, and often surpass rival approaches. Moreover, a fresh approach is proposed for the solution of viscous conservation laws via a transformation to a relaxation system, which becomes stiff

as the Peclet number (P_Δ) increases. This has the potential of eliminating the errors arising from higher–order differencing on irregular meshes, a severe problem in cartesian adaptive–grid methods. In short, since the method is robust, economical and results in physical solutions (even in the fluid dynamic limit), it is a viable method for computing under–resolved flows.