Thermal Simulation of Pipeline Flow

Philip Keenan

CRPC-TR91187
September 1991

Center for Research on Parallel Computation
Rice University
6100 South Main Street
CRPC - MS 41
Houston, TX 77005

Revised: August, 1993. Presented at SIAM Annual Meeting (July, 1990) in Chicago.

Thermal Simulation of Pipeline Flow

Philip T. Keenan*

August 2, 1993

Abstract

A new numerical method for studying one dimensional fluid flow through pipelines is presented and analyzed. This work extends in a certain direction the collocation method described by Luskin["An Approximation Procedure for Nonsymmetric, Nonlinear Hyperbolic Systems with Integral Boundary Conditions", SIAM J. Numer. Anal. 1976.]. The pressure and velocity of an isothermal fluid in a pipeline can be described by a coupled pair of nonlinear first order hyperbolic partial differential equations. When thermal effects are important a third equation for temperature is added. While Luskin's method works well for the isothermal situation he discussed, it does not apply in certain common cases when thermal effects are modeled. The analysis of this new method shows how the difficulties that come from the application of standard collocation can be overcome. Experiments indicate that this method is a substantial improvement over standard collocation. It also describes an approach to analyzing nonlinear evolution equations with smooth solutions which produces convergence theorems about the nonlinear system from the corresponding linear theorems with relatively little extra work. This technique also yields an H^1 estimate in the isothermal case.

Key Words: first order hyperbolic, collocation method, upwinding

AMS(MOS) subject classification: 65N35

*Department of Mathematics, University of Chicago, supported in part by a Graduate Fellowship from the National Science Foundation.

Contents

1 Introduction **3**

2 The Partial Differential Equations **3**

3 The Numerical Method **5**

3.1 Rewriting the Partial Differential Equations 5

3.2 Boundary Conditions 7

3.3 Discrete Notation 7

3.4 Defining the Numerical Method 10

4 The General Framework **11**

4.1 The degenerate case 12

4.2 The non-degenerate case 15

5 Theoretical Results **16**

6 Computational Results **17**

7 Remarks and Extensions **19**

8 Error Analysis **26**

8.1 The Error Equation 26

8.2 Diagonalization 32

8.3 The 27 Product Terms 36

8.4 The Induction Hypothesis 43

8.5 The Gronwall Induction 45

1 Introduction

Thermal modeling of fluid flow through pipelines is increasingly necessary as practical engineering applications demand greater fidelity between model and reality. When pipeline simulators are used for leak detection, for instance, real time thermal modeling may be required to distinguish leaks from pressure changes due to heat exchange with the environment. Thermal modeling may also be important in attempting to optimize pipeline usage for maximum economic utility.

Isothermal fluid flow in pipelines is described by a pair of coupled first order nonlinear hyperbolic partial differential equations for pressure and velocity. In 1978 Mitchell Luskin[2] described and analyzed a numerical method applicable to a large class of systems, including the isothermal pipeline case. The method relies upon the fact that the eigenvalues of the matrix which determines the characteristic directions are both bounded far away from zero. In its simplest version, the method uses piecewise linear approximations in space and time and can be described as box centered collocation.

Thermal effects in pipeline flow introduce a third coupled equation for temperature and a resulting third eigenvalue which is small or possibly zero. It can be shown that straightforward application of collocation methods to problems with small eigenvalues produces numerical instability. This work presents a generalization to straightforward collocation which allows the inclusion of thermal effects. We approximate the temperature with velocity upwinded piecewise constants in space, which rectifies the stability problems but substantially increases the complexity of the resulting analysis.

We begin with formal statements of the problem, the method, the theorems and some numerical results. The proofs are postponed to Section 8.

2 The Partial Differential Equations

We remark that the thermal pipeline equations are but one instance of a more general set of coupled first order hyperbolic partial differential equations which may be solved using the methods indicated below. In particular, Luskin's method applies to systems with large eigenvalues, and my method applies to systems with several large and one small eigenvalue.

The primary subject of the investigation here is the numerical approximation of the solution. There does not yet exist a complete theory for the system of nonlinear partial differential equations describing the thermal pipeline model. Therefore, we make certain assumptions about the nature of the solutions which appear reasonable for pipelines. In particular, we assume that the flow speed is much smaller than the speed of sound and that no shocks are present.

Luskin and Blake[3] have demonstrated the existence of smooth solutions for systems such as the isothermal equations.

We state the equations for the *Plug Flow* model of pipeline fluid flow; we are outside the region of laminar flow but in a regime where shocks are not generally observed. These equations have been known for a very long time; Bernoulli described versions in the 1800's.

The three dependent variables are the pressure $p = p(x,t)$, the velocity $v = v(x,t)$, and the temperature $T = T(x,t)$. Here $x \in [0,E]$ and $t \geq 0$, where E is the length of the pipeline. We will restrict attention to a rigid horizontal pipe, though extensions incorporating gravitational effects are straightforward. To be precise, we measure the pressure and temperature at the center of a cross section of the pipe, while the velocity is averaged over the cross section.

From these three variables we also compute the density $\rho = \rho(p,T)$, and the specific internal energy $\mathcal{E} = \mathcal{E}(p,T)$, from tables of thermodynamic constants or an equation of state for the fluid of interest.

Three partial differential equations connect the three independent variables. First is the Conservation of Mass equation,

$$\rho_t + (\rho v)_x = 0.$$

The subscripts denote partial differentiation. This states that no fluid is created or destroyed.

Next is the Conservation of Momentum equation, which is Newton's law of force balance:

$$(\rho v)_t + (\rho v^2)_x + p_x + \frac{\rho f |v| v}{2D} = 0.$$

Here f is a dimensionless positive constant called the coefficient of friction; other forms could also be used for the frictional dissipation term. The parameter D is the internal diameter of the pipe.

Finally we have the Conservation of Energy equation, which refers to the total kinetic and internal energy of the fluid:

$$(\rho\mathcal{E} + \rho v^2/2)_t + ((\rho\mathcal{E} + \rho v^2/2)v)_x + (pv)_x - q = 0.$$

Here q is a function of x and t describing the flow of heat into or out of the pipeline through its walls.

The equations are to be interpreted in some consistent set of units.

The plug flow model has been validated in engineering practice for pipelines carrying such diverse fluids as natural gas, crude oil and water; see [5] for references.

These equations can be thought of as forming a differential-algebraic system. Although some such systems have been analyzed, a general theory is still lacking.

The equations above contain the one-dimensional equations of gas dynamics as the special case $q = f = 0$; hence under certain conditions the solution can involve shocks. The method presented here is appropriate for the solutions which are important in pipeline applications, but this method would probably not be a good choice if shocks were present.

We choose to compute with pressure, velocity and temperature, rather than density, mass flow and internal energy, because the latter variables may be almost discontinuous across the boundaries between different batches of fluid in a pipeline. Pipelines frequently carry several different fluids in successive batches. Some pipelines also contain multiple phases, gaseous as well as liquid components, but we ignore such complications here. Since we are looking for smooth solutions and are using continuous functions to approximate the computing variables, pressure and velocity are appropriate choices. In addition, the discontinuous piecewise constants we use for temperature are also appropriate since contact discontinuities in temperature are possible with appropriate boundary conditions.

The asymptotic accuracy of this method will be limited to first order by the presence of the piecewise constant approximation for temperature. However, we still expect the method to perform well in practice. Temperature effects are generally only a small correction to the pressure and velocity system, as indicated by the wide utility of Luskin's pressure and velocity method. Moreover large temperature changes tend to take place on a much slower time scale than that needed to resolve sonic effects. Thus we expect that in practice the first order errors in temperature will have a small coefficient compared to the second order pressure and velocity errors, leading to better than first order convergence rates at practical mesh sizes.

3 The Numerical Method

3.1 Rewriting the Partial Differential Equations

We begin by formulating the differential system. After expanding the equations in terms of p, v, and T, we make some simple substitutions to get

$$\rho_p(p_t + vp_x) + \rho_T(T_t + vT_x) + \rho v_x = 0, \tag{1}$$

$$v_t + vv_x + \frac{1}{\rho}p_x + \frac{f|v|v}{2D} = 0, \tag{2}$$

$$\mathcal{E}_p(p_t + vp_x) + \mathcal{E}_T(T_t + vT_x) + \frac{1}{\rho}pv_x - \frac{f|v|v^2}{2D} - \frac{q}{\rho} = 0. \tag{3}$$

We next diagonalize the time derivative term, putting the equations into standard form. This has the effect of removing as much temperature dependence as

possible from the pressure and velocity equations, which is helpful since temperature terms will be the main stumbling block in the analysis.

In particular, we analytically invert the 2×2 matrix

$$\begin{pmatrix} \rho_p & \rho_T \\ \mathcal{E}_p & \mathcal{E}_T \end{pmatrix}.$$

The inverse is

$$\frac{1}{k} \begin{pmatrix} \mathcal{E}_T & -\rho_T \\ -\mathcal{E}_p & \rho_p \end{pmatrix},$$

where

$$k = \rho_p \mathcal{E}_T - \mathcal{E}_p \rho_T.$$

The quantity k is positive for fluids such as natural gas.

We write the resulting system of PDE's as

$$u_t + \bar{A} u_x = F, \tag{4}$$

where

$$u = (p, v, T)^{tr},$$

and $\bar{A}$ and F are known functions. Note that the system is non-linear: $\bar{A}$ and F depend on $u(x,t)$; in general they could also depend on x and t explicitly.

In more detail, for some functions a, b, and c, we can write

$$\bar{A}(u) = \begin{pmatrix} v & a & 0 \\ b & v & 0 \\ 0 & c & v \end{pmatrix}. \tag{5}$$

The characteristic polynomial of A is

$$\mathrm{ch}(\lambda) = (v - \lambda)((v - \lambda)^2 - ab).$$

Let $s = \sqrt{ab}$. Then the eigenvalues of $\bar{A}$ are

$$\begin{aligned} \lambda_1 &= s + v, \\ \lambda_2 &= -s + v, \\ \lambda_3 &= v. \end{aligned}$$

These represent three modes, namely, two high speed components called sound waves moving in opposite directions, and one mode moving with the underlying fluid velocity. For this reason s is called the *sonic velocity*. Note that we have assumed $s \gg |v|$ over the operational range of the pipe. This keeps the eigenvalues distinct. For most pipelines, $|v| \leq 0.01s$. It is easy to see that because the third eigenvalue can vanish, straightforward application of collocation to equations of this form can give bizarre behavior; in the case of $v = 0$ a change in temperature at one end of the pipe would be instantly propagated as a saw tooth wave down the entire pipe.

3.2 Boundary Conditions

We need to specify initial conditions for pressure, velocity and temperature. Thus we assume

$$u(x, 0) = u_0(x),$$

for all $x \in [0, E]$ where u_0 is a given, smooth, vector function.

In addition, we need to specify one boundary condition for each in-flowing component of the solution, at each end of the pipe. As a simple example, we may specify pressure at each end of the pipe, and temperature at each inlet end. That is, we set

$$p(0, t) = p_l(t),$$

$$p(E, t) = p_r(t),$$

where p_l and p_r are given smooth functions. Depending on the sign of the velocity we make analogous specifications for temperature at both ends, one end, or not at all. In particular, we specify

$$T(0, t) = T_l(t) \text{ whenever } v(0, t) > 0, \text{ and}$$

$$T(E, t) = T_r(t) \text{ whenever } v(E, t) < 0.$$

These conditions make the problem well-posed. They are representative of the conditions for which the theorems and analysis hold. In Section 4 we will describe the general class of boundary conditions which we consider.

3.3 Discrete Notation

We now introduce some notation for describing the discrete problem. Given a positive integer N, let

$$\mathcal{P}_x = \{x_0, x_1, \ldots, x_N\},$$

where $x_k = kh$, $h = E/N$ and E is the length of the pipe. This gives a uniform partition of $[0, E]$ into N intervals.

We write the spatial midpoints as

$$x_{j+1/2} = (x_j + x_{j+1})/2.$$

We collect these with

$$\hat{\mathcal{P}}_x = \{x_{j+1/2} : j = 0, 1, \ldots, N-1\}.$$

Next we define two shift operators on $[0, E]$:

$$x_+ = \begin{cases} x_{j+1/2} & \text{for } x = x_j,\ j < N, \\ x_N & \text{for } x = x_N, \\ x_{j+1} & \text{for } x \in (x_j, x_{j+1}), \end{cases}$$

and

$$x_- = \begin{cases} x_{j-1/2} & \text{for } x = x_j,\ j > 0, \\ x_0 & \text{for } x = x_0, \\ x_j & \text{for } x \in (x_j, x_{j+1}). \end{cases}$$

If u is any function of x, we let

$$u_+(x) = u(x_+),$$

$$u_-(x) = u(x_-).$$

Define

$$\triangle x = x_+ - x_-,$$

and the centered divided difference operator by

$$\partial_x u = (u_+ - u_-)/\triangle x.$$

Note that for our uniform space partition, $\triangle x = h$ everywhere except at $x = 0$ and $x = E$. We also use a centered approximation to u given by

$$u_{,c} = (u_+ + u_-)/2.$$

We adopt analogous time operators using superscripts, based on a set of time levels

$$\mathcal{P}_t = \{t^0, t^1, \ldots, t^M\},$$

where $t^0 = 0$, $t^M = \tau$ is some fixed final time, and $t^n < t^{n+1}$ for each n. This gives a possibly non-uniform partition of $[0, \tau]$ into M time steps.

Given $\theta \in [0, 1]$ we set

$$t^{n+\theta} = \theta t^{n+1} + (1 - \theta) t^n,$$

and write the set of theta-weighted time levels as

$$\hat{\mathcal{P}}_t = \{t^{n+\theta} : n = 0, 1, \ldots, M - 1\}.$$

We also define

$$u'^{\theta} = \theta u^+ + (1 - \theta) u^-.$$

Finally,

$$\hat{\mathcal{Q}} = \hat{\mathcal{P}}_x \times \hat{\mathcal{P}}_t,$$

is the set of space time points at which we will apply collocation.

Except in Section 7 we assume the time partition is uniform, with

$$t^{n+1} - t^n = \triangle t = \tau / M \text{ for all } n.$$

For example, using this notation we can write

$$\partial_t u(x,t) + \partial_x u(x,t) = 0 \quad \text{for all } (x,t) \in \hat{\mathcal{Q}},$$

by which we mean

$$\frac{u(x_{j+1/2}, t^{n+1}) - u(x_{j+1/2}, t^n)}{t^{n+1} - t^n} + \frac{u(x_{j+1}, t^{n+\theta}) - u(x_j, t^{n+\theta})}{x_{j+1} - x_j} = 0,$$

$$\text{for all } j = 0, 1, \ldots, N - 1, \text{ and } n = 0, 1, \ldots, M - 1.$$

We will henceforth suppress the (x, t) arguments to all functions.

We will need to define several bilinear forms, including the usual $L^2(0, E)$ inner product

$$(f, g)_{L^2} = \int_0^E f\, g\, dx,$$

and two discrete approximations to it,

$$< f, g >_{m^2} = \sum_{\hat{P}_x} f\, g \triangle x,$$

$$< f, g >_{l^2} = \sum_{P_x} f\, g \triangle x.$$

With these are associated the norms and semi-norms

$$\|f\|_{L^2} = \sqrt{(f,f)_{L^2}},$$
$$|f|_{m^2} = \sqrt{<f,f>_{m^2}},$$
$$|f|_{l^2} = \sqrt{<f,f>_{l^2}}.$$

Finally, when dealing with functions of several variables we will use tensor indexing notation and the summation convention. Thus if $A = (A_{ij})$ is a matrix function of a vector u, then we write

$$A_{ij,kl} = \frac{\partial}{\partial u_k}\frac{\partial}{\partial u_l}A_{ij},$$

where A_{ij} is the element of A in the i^{th} row and j^{th} column. If B is another matrix of compatible dimensions we write

$$A_{ik}B_{kj} = \sum_k A_{ik}B_{kj}.$$

3.4 Defining the Numerical Method

Returning to the problem at hand, we recall $u = (p, v, T)^{tr}$. The numerical method involves approximating u by U, where pressure and velocity are piecewise linear in space and in time and temperature is discontinuous piecewise constant in space and piecewise linear in time. Temperature is also velocity advected or *upwinded*. Upwinding allows us to make temperature well defined at the knots, which further allows us to use all the above indexing and discrete derivative notations. U_3^n is upwinded based on U_2^{n-1}, by the rule that

$$U_3^n(x_j) = \begin{cases} U_3^n(x_{j+1/2}) & \text{if} \quad U_2^{n-1}(x_j) < 0, \\ U_3^n(x_{j-1/2}) & \text{if} \quad U_2^{n-1}(x_j) \geq 0. \end{cases}$$

We choose U^0 based on the initial condition u^0; it can be the interpolant of u^0 into the spaces defining U.

To compute $U(t^{n+1})$ from $U(t^n)$, we discretize equation (4). Before doing so, we first make the substitution

$$vT_x = (vT)_x - v_xT.$$

This is useful in the analysis in make the upwinding of temperature work out. In writing this term we will make use of two important matrices,

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{6}$$

and $Q = I - P$ where I is the 3×3 identity matrix. We also let $A = \bar{A}P$, so that $\bar{A} = AP + \lambda_3 Q$.

Next, writing $u = u_0 + (u - u_0)$ where $u - u_0$ is small, we Taylor expand around u_0 and ignore higher order terms beginning with $(u - u_0)^2$. Thinking of u_0 as $u(x, t^n)$ and u as $u(x, t^{n+\theta})$, we have an equation we can discretize. We do so with collocation at each point of $\hat{\mathcal{Q}}$. We write the resulting equations in tensor notation, with $k = 1, 2, 3$. All the coefficient functions A and F in this equation are evaluated at U^-. We thus obtain

$$\begin{aligned} \partial_t U_k + A_{kj}\partial_x U_j + \delta_{k3}\partial_x(U_2^- U_3) - \delta_{k3} U_3 \partial_x U_2^- + A_{kj,i}\partial_x U_j^-(U_i - U_i^-) \\ + \delta_{k3}\partial_x((U_2 - U_2^-)U_3^-) - \delta_{k3} U_3^- \partial_x (U_2 - U_2^-) \\ = f_k + f_{k,i}(U_i - U_i^-) \quad \text{at every point in } \hat{\mathcal{Q}}. \end{aligned} \tag{7}$$

One can view this as a discretization of the integral average of the PDE over $[x_j, x_{j+1}] \times [t_n, t_{n+1}]$.

We also discretize the boundary conditions of Section 3.2. We set

$$U_{1,0}^{n+1} = p_l(t^{n+1}),$$

$$U_{1,N}^{n+1} = p_r(t^{n+1}).$$

To handle the temperature boundary condition we set

$$U_{3,-1/2}^{n+1} = T_l(t^{n+1}),$$

$$U_{3,N+1/2}^{n+1} = T_r(t^{n+1}),$$

and let the usual rule for upwinding determine when this determines $U_{3,0}$ and $U_{3,N}$.

We should mention that the system (7) *is* soluble numerically; it produces a nonsingular square matrix system which has a small bandwidth. To facilitate upwinding of temperature it is convenient to let the midpoint temperature values be variables as well as the pressure, velocity and temperature knot values. Under plausible assumptions the variable ordering $p_j, T_j, v_j, T_{j+1/2}$ produces a matrix which can be solved without pivoting.

4 The General Framework

The thermal pipeline equations formulated in Section 3 are but one instance of a more general set of equations to which my method applies. In this section we state the general problem we propose to solve, which we will call the degenerate case, because of the possibility of one eigenvalue becoming arbitrarily small or even changing

sign. For comparison we also state a version of Luskin's 1978 equations which we call the non-degenerate case. This is the case where all the eigenvalues are bounded away from zero, and in the pipeline context corresponds to assuming some functional form for temperature and then solving only the pressure and velocity equations. The reader interested more in pipelines than in mathematics may skip this section; such a reader should substitute "thermal pipeline equations" for "degenerate case" and "pressure-velocity equations" for "non-degenerate case" hereafter.

We consider the following system of μ first order hyperbolic nonlinear partial differential equations:

$$u_t + \bar{A}(u)u_x = F(u), \tag{8}$$

where $u(x,t) \in \mathcal{R}^\mu$, $x \in [0,E]$, $t \in [0,\tau]$, $\bar{A}$ is a $\mu \times \mu$ matrix, and F is a μ-vector, and where $\bar{A}$ and F are smooth functions of x, t, and $u(x,t)$.

We are given the initial condition

$$u(x,0) = u_0(x) \text{ for all } x \in [0,E], \tag{9}$$

as well as suitable boundary conditions. We need one boundary condition at $(0,t)$ for each eigenvalue of $\bar{A}$ which is positive there, and one boundary condition at (E,t) for each eigenvalue of $\bar{A}$ which is negative there. We describe these more fully in the following two special cases, based on the form of the matrix $\bar{A}$.

In both the following cases we assume that all the eigenvalues of $\bar{A}$ are real and that there exists a smooth and bounded transformation matrix $S(u(x,t),x,t)$ with smooth and bounded inverse such that $S^{-1}\bar{A}S$ is diagonal.

4.1 The degenerate case

Suppose that all the eigenvalues but one of $\bar{A}$ are uniformly bounded away from zero. Suppose also that $\bar{A}$ has the form

$$\bar{A}(u) = \left(\begin{array}{c|c} \mathcal{A}_{11} & 0 \\ \hline \mathcal{A}_{21} & \lambda \end{array}\right), \tag{10}$$

where λ is that one eigenvalue which is not bounded away from zero. Here $\mathcal{A}_{11}$ represents a $\mu - 1 \times \mu - 1$ submatrix, and $\mathcal{A}_{21}$ a $1 \times \mu - 1$ submatrix. Suppose also that

$$\frac{\partial \lambda(u)}{\partial u_\mu} = 0, \tag{11}$$

so that λ does not depend on u_μ. Finally, if the eigenvalues of $\bar{A}$ are $\{\lambda_1, \ldots, \lambda_{\mu-1}, \lambda\}$, then suppose

$$|\lambda_i| \gg |\lambda| \text{ for } 1 \le i \le \mu - 1. \tag{12}$$

The boundary conditions for u_μ are handled exactly as was done for temperature in the previous section, namely we specify

$$u_\mu(0,t) = T_l(t), \text{ whenever } \lambda(0,t) > 0, \text{ and}$$

$$u_\mu(E,t) = T_r(t), \text{ whenever } \lambda(E,t) < 0.$$

Now let S_{11} be a matrix such that $S_{11}^{-1}\mathcal{A}_{11}S_{11}$ is the diagonal matrix $\text{diag}(\lambda_1, \ldots, \lambda_{\mu-1})$ where $\lambda_i > 0$ for $1 \le i \le k$ and $\lambda_i < 0$ for $k+1 \le i \le \mu - 1$. Then we must specify k boundary conditions at $x = 0$ and $\mu - 1 - k$ at $x = E$. At each boundary the ingoing components must be specified as an affine linear function of the outgoing ones, in order for the boundary conditions to be realizable. Let $v = (u_1, \ldots, u_{\mu-1})$. Then we allow the following boundary conditions:

$$R_1 v(0,t) = v_a(t),$$

$$R_2 v(E,t) = v_b(t),$$

where v_a and v_b are given smooth vector functions, and R_1 is a constant $k \times \mu - 1$ matrix and R_2 is a constant $\mu - 1 - k \times \mu - 1$ matrix, and R_1 and R_2 satisfy

$$R_1 S_{11}^{-1} = (E_{11}|E_{12}),$$

$$R_2 S_{11}^{-1} = (E_{21}|E_{22}),$$

where E_{11} and E_{22} are *non-singular* $k \times k$ and $\mu - 1 - k \times \mu - 1 - k$ matrices, respectively.

Assumption 1 *We assume that the system given by (8), with initial data given by (9) and boundary data as described above, does have a smooth solution u for all $t \in [0, \tau]$.*

For example in the pipeline case the degenerate eigenvalue is the flow velocity, which is much smaller than the sonic velocity. It arises in the temperature equation but is independent of temperature. In addition T_x does not appear in the p and v equations. This was achieved by an analytic transformation prior to discretizing the equations. This special form seems important to the analysis. The matrix S_{11}^{-1} has all entries non zero, so choosing $R_1 = R_2 = (1, 0)$ yields non zero scalars for E_{11} and E_{22}. This illustrates the acceptability of the pressure boundary conditions described in Section 3.

Definition 1 *When all the above suppositions hold, we say that we are considering the* **degenerate case** *of (8). In this case we also define a $\mu \times \mu$ matrix P as follows:*

$$P = \text{diag}(1, 1, \ldots, 1, 0). \tag{13}$$

We now define $Q = I - P$, where I is the $\mu \times \mu$ identity matrix, and $A = \bar{A}P$.

With this notation, we can write the degenerate case of (8) as

$$u_t + A(u)u_x + \lambda(u)Qu_x = F(u). \tag{14}$$

Equation (14) has the form

$$u_t + G(u, u_x) = 0,$$

where G is a smooth function of 2μ arguments. Applying Taylor's theorem and using the summation convention that repeated subscripts of i, j, k, or l indicate summation from 1 to μ, we have

$$G_k(v,w) = G_k(v_0,w_0) + G_{k,i}(v_0,w_0)(v_i - v_{0i}) + G_{k,i+\mu}(v_0,w_0)(w_i - w_{0i}) + \text{Quad},$$

where Quad contains only terms quadratic in $v - v_0$ or $w - w_0$. In particular, using $v = u(x,t)$, $w = u_x(x,t)$, $v_0 = u_0(x,t_0)$, and $w_0 = u_{0x}(x,t_0)$, (14) becomes

$$\begin{aligned} &u_{kt} + A_{kj}(u_0)u_{xj} + A_{kj,i}(u_0)u_{0xj}(u_i - u_{0i}) \\ &\quad + \lambda(u_0)\delta_{k\mu}u_{\mu x} + \lambda_{,i}(u_0)\delta_{k\mu}u_{0\mu x}(u_i - u_{0i}) \\ &\quad = F_k(u_0) + F_{k,i}(u_0)(u_i - u_{0i}) + \text{Quad}. \end{aligned} \tag{15}$$

We now choose $u_0(x,t_0) = u(x,t^-)$, and evaluate at all points in $\hat{\mathcal{Q}}$. We also use the identity $ab_x = (ab)_x - a_x b$ on the λ-terms, obtaining

$$\begin{aligned} &u_{kt} + A_{kj}(u^-)u_{xj} + A_{kj,i}(u^-)u^-_{xj}(u_i - u_i^-) \\ &\quad + \delta_{k\mu}\Big((\lambda(u^-)u_\mu)_x - \lambda_x(u^-)u_\mu + (\lambda_{,i}(u^-)u^-_\mu(u_i - u_i^-))_x \\ &\quad - (\lambda_{,i}(u^-)(u_i - u_i^-))_x u^-_\mu\Big) \\ &\quad = F_k(u^-) + F_{k,i}(u^-)(u_i - u_i^-) + \text{Quad}. \end{aligned} \tag{16}$$

We will approximate u by U, where each component U_k is piecewise linear in time between the time levels $\mathcal{P}_t$. Each U_k is piecewise linear in space between the knots $\mathcal{P}_x$, except for the last component U_μ. U^n_μ is piecewise constant in space and is upwinded by $\lambda(U^{n-1})$, meaning that

$$U^n_\mu(x_j) = \begin{cases} U^n_\mu(x_{j+1/2}) & \text{if} \quad \lambda^{n-1}(x_j) < 0, \\ U^n_\mu(x_{j-1/2}) & \text{if} \quad \lambda^{n-1}(x_j) \geq 0. \end{cases}$$

We now discretize (16) by collocation at the points of $\hat{\mathcal{Q}}$, except that we ignore the quadratic terms and use essentially an integral average on the $(ab)_x$ terms. We thus obtain as our proposed numerical method

$$\begin{aligned}\partial_t U_k + A_{kj}(U^-)\partial_x U_j + A_{kj,i}(U^-)\partial_x U_j^-(U_i - U_i^-) \\ + \delta_{k\mu}\Big(\partial_x(\lambda(U^-)U_\mu) - U_\mu\partial_x\lambda(U^-) + \partial_x(\lambda_{,i}(U^-)U_\mu^-(U_i - U_i^-)) \\ - U_\mu^-\partial_x(\lambda_{,i}(U^-)(U_i - U_i^-))\Big) = F_k(U^-) + F_{k,i}(U^-)(U_i - U_i^-). \qquad (17)\end{aligned}$$

We discretize the boundary conditions exactly as in the previous section. We discretize the initial conditions by setting $U^0 = W^0$, where W^0 is the interpolant of the initial data u_0 into the discrete space in which U is defined.

4.2 The non-degenerate case

Suppose that all the eigenvalues of $\bar{A}$ are uniformly bounded away from zero.

In this case we use the same affine linear boundary conditions as in the degenerate case, except that we no longer treat the u_μ component differently from the rest.

Assumption 2 *We assume that the system given by (8), with initial data given by (9) and boundary data as described above, does have a smooth solution u for all $t \in [0, \tau]$.*

For example, in the pipeline case, if one assumes a given functional form $T = T(p)$, and ignores equation (3), the remaining two equations for pressure and velocity have $v + s$ and $v - s$ for eigenvalues. Since $s \gg |v|$, both eigenvalues are bounded well away from zero. Specifying pressure at each end remains a realizable set of boundary conditions.

Definition 2 *When the above supposition holds, we say that we are considering the* **non-degenerate case** *of (8). In this case we also define a $\mu \times \mu$ matrix P as follows:*

$$P = I. \qquad (18)$$

We again define $Q = I - P$, and $A = \bar{A}P$. Note that now $Q = 0$ and $A = \bar{A}$.

With this notation, we can write the non-degenerate case of (8) as

$$u_t + A(u)u_x + \lambda(u)Q u_x = F(u), \qquad (19)$$

which looks the same as equation (14) but which uses a different definition for P, Q, and A. This allows us to carry out the analysis of the degenerate case once and then get the corresponding results for the non-degenerate case simply by redefining the

matrix P. In particular, we use the same numerical method as for the degenerate case, namely that of equation (17). Note, however, that since $P = I$ and $Q = 0$, that equation simplifies considerably, to

$$\begin{aligned} &\partial_t U_k + A_{kj}(U^-)\partial_x U_j + A_{kj,i}(U^-)\partial_x U_j^-(U_i - U_i^-) \\ &\quad = F_k(U^-) + F_{k,i}(U^-)(U_i - U_i^-). \end{aligned} \tag{20}$$

In particular, there is no upwinding to worry about, and all components of U are piecewise linear in space.

5 Theoretical Results

In this section we state some asymptotic convergence results. We need to make the following assumption.

Assumption 3 *Assume $\theta \in (\frac{1}{2}, 1]$ is a given constant. Assume there is a constant K_0, independent of h and $\triangle t$, such that*

$$\frac{1}{K_0} \leq \frac{h}{\triangle t} \leq K_0,$$

as both h and $\triangle t$ go to zero.

For the nonlinear thermal pipeline equations we obtain an L^2 convergence result.

Theorem 1 (The Degenerate Case) *Consider the degenerate case of equation (8). Let assumptions 1 and 3 hold. Then there is a final time $0 < \bar{\tau} \leq \tau$ and a constant C which depends on K_0 and on Sobolev norms for u but remains bounded even when $\lambda = 0$, and which is otherwise independent of h and $\triangle t$, such that for h and $\triangle t$ sufficiently small,*

$$||U - u||_{l^\infty(L^2)} \leq Ch,$$

and

$$||P(U - u)||_{l^\infty(L^\infty)} \leq Ch^{3/4},$$

where the l^∞ norm in time is taken over the range $0 \leq t \leq \bar{\tau}$.

Note that although we assumed $h \sim \triangle t$, there is no CFL type constraint, since this is an implicit method.

In general the constant $\bar{\tau}$ can be order one. However, strengthening of the hypotheses allows us to make $\bar{\tau}$ arbitrarily large by choosing h sufficiently small, in which case the method converges for as long as you wish to run it. Computational examples suggest that this is possible even without extra hypotheses.

Theorem 2 *Consider the degenerate case of equation (8). Let assumptions 1 and 3 hold. Suppose that the following additional assumption holds: The matrix $A(u)$ is independent of u_x. Then the parameter $\bar{\tau}$ in Theorem 1 depends upon h, and may be made as large as desired by choosing h sufficiently small.*

If we assume that the matrix $\bar{A}$ in (4) is independent of u, and that F depends only linearly on u, then the system is linear and we can in fact prove H^1 convergence on pressure and velocity, as described in Keenan[1]. In this case there is no need for an induction, the standard Gronwall inequality applies and the convergence is for all time.

In the nonlinear non-degenerate case, Luskin[2] proved an L^2 estimate for the case $\theta = 1/2$. The same proof I use for the degenerate case yields the following H^1 result in the non-degenerate case when $\theta > 1/2$.

Theorem 3 (Non-degenerate Case) *Consider the non-degenerate case of equation (8). Let assumptions 2 and 3 hold. Then given any final time $0 < \bar{\tau} \leq \tau$, there exists a constant C which depends on K_0 and on Sobolev norms for u but remains bounded even when $\lambda = 0$, and which is otherwise independent of h and $\triangle t$, such that for h and $\triangle t$ sufficiently small,*

$$||U - u||_{l^\infty(H^1)} \leq Ch.$$

6 Computational Results

I have implemented my method for the nonlinear thermal pipeline equations and informally compared results against state of the art commercial codes currently in widespread use. Such codes use ad-hoc methods to incorporate temperature effects which generally require very small time steps to maintain stability. Such time step limitations are poorly understood since convergence analyses do not exist for these methods. Due to the proprietary nature of commercial pipeline codes we cannot present a detailed comparison. However, my method does seem to be able to use much longer time steps than the comparison methods, and the analysis does not require any limitation on the time step. For instance, there is no CFL constraint as would occur in an explicit method. This is important in networks of pipelines of different lengths where the time step for the system would be limited by the smallest natural time step in the network.

I emphasize that my numerical experiments are for the fully nonlinear thermal pipeline equations, in which the matrix A does depend on temperature and the velocity indeed changes sign; yet there has been no evidence of the error blowing up in finite time. Computationally there does not seem to be a restriction on the $\bar{\tau}$ of Theorem 1.

We begin with two illustrations of actual computed solutions.

The first example deals with a pipeline for methane gas. We use a 150 km insulated pipe with a 75 cm internal diameter, carrying gaseous methane. Initially everything is at rest, with a pressure of 8000 kpa and a temperature of 20°C throughout the pipe. We then open the ends of the pipe and drop the outlet pressure to 5500 kpa over 1 minute. Over the next 8 to 16 hours the flow evolves to a steady state. Using 10 km space intervals and 10 minute time steps, we compute the solution after 12 hours. Figure 1 illustrates the resulting pressure, velocity and temperature along the pipe. To fit all three variables on one graph, temperature is shown in degrees C, velocity in meters per second, and pressure is in mega-pascals.

The second example deals with a pipeline for liquid n-octane. We use a 100 km insulated pipe with a 60 cm internal diameter, carrying liquid n-octane. Initially everything is at rest, with a pressure of 1400 kpa and a temperature of 20°C throughout the pipe. We then apply a ten second pulse of extra pressure at the inlet end. This creates a smooth traveling wave in pressure which propagates down the pipe at the sonic velocity of 1.6 km/sec. The pressure wave has amplitude equal to ten percent of the initial pressure, or 140 kpa. As it travels it excites identical looking pulses in velocity and temperature. Using 1 km space intervals and 5/8 sec. time steps, we compute the solution during the first 45 seconds. Figure 2 illustrates the resulting pressure at 15, 30 and 45 seconds. We notice a decay in the wave amplitude, due both to friction and to numerical dissipation in the upwinding process. In both examples the friction factor $f = 0.014$.

We now present a table of empirical convergence rates based on these two example scenarios.

Table 1 indicates the convergence rates obtained for pressure, velocity and temperature in each of the two scenarios described above. In each case we measured both the L^2 and L^∞ norms of the error in each component, as compared to a reference solution on a much finer mesh. As h was decreased, $\triangle t$ was decreased proportionately. For the steady state simulation we examined the norm of the error 12 hours after opening the valves. For the small amplitude wave case, we used the norm of the error after 15 seconds.

Figure 3 is a log-log plot of the errors as h decreases. It shows sample points for both the L^2 and L^∞ norms of the errors in pressure, velocity and temperature in each of the two scenarios described above. In each case the base ten logarithm of the error is plotted against the base ten logarithm of the number of spatial intervals. These are the sample points used in constructing Table 1. Note that pressure is in pascals, velocity in meters per second, and temperature in degrees Celsius. These units separate the error curves into three bands, with pressure on top, then temperature, and finally velocity. The three bands on the left are for the steady state case; the three on the right are for the small amplitude waves. Each band shows the L^2 and

L^∞ norms almost overlapping, because the length of the pipe has been scaled out of the L^2 norm.

The table and graph illustrate our empirical observation that even though the first order nature of piecewise constants appears in the asymptotic convergence rates of the above theorems, in practice we may obtain close to second order convergence rates. This is not too surprising since the pressure and velocity approximations are second order (for $\theta = 1/2$), and since isothermal pressure-velocity simulations give good results in many situations. Temperature effects generally occur on a much slower time scale than sonic effects, so we expect the constant on the first order error terms to be small relative to typical practical values of h. In fact we see that in the near steady state case, where the temperature varies only slowly, the convergence is indeed approximately second order, at least for pressure, over the parameter range shown. We point out that this range is more than sufficient for practical computations, since the errors shown are well below the error of measurement in the real pipeline. We also see first order effects dominating in the small wave case, since here the temperature changes as sharply and rapidly as pressure and velocity.

It is interesting to note the effect of temperature on the sonic speed. If $\mathcal{E}$ is made extremely large, the effect is to force temperature to be essentially constant. This produces a substantial change in the sonic velocity. In the methane pipeline the sonic speed decreases from 415 m/sec in the adiabatic case to 350 m/sec in the isothermal case. In the octane pipeline it decreases from 1630 m/sec to 1312 m/sec. This means that the pulses in Figure 2 would travel about 20% slower if temperature effects were omitted, despite the fact that the overall temperature in that example is virtually constant.

7 Remarks and Extensions

Remark 1 (Higher Order Methods) We begin by remarking on the choice of piecewise constants for temperature rather than some higher order representation. We chose piecewise constants for three main reasons. First, higher order methods are harder to analyze. These methods typically involve slope-limiting procedures for which rigorous convergence results do not yet exist even in the scalar equation case. Second, our numerical experiments indicate very good accuracy with piecewise constants in examples where the parameters were chosen to be of engineering interest. Finally higher order methods are harder to code, making them less useful in engineering practice.

Remark 2 (More General Boundary Conditions) Our analysis holds without change for other simple configurations of boundary conditions such as specifying velocity at each end of the pipe rather than pressure at each end. Luskin[2] treats

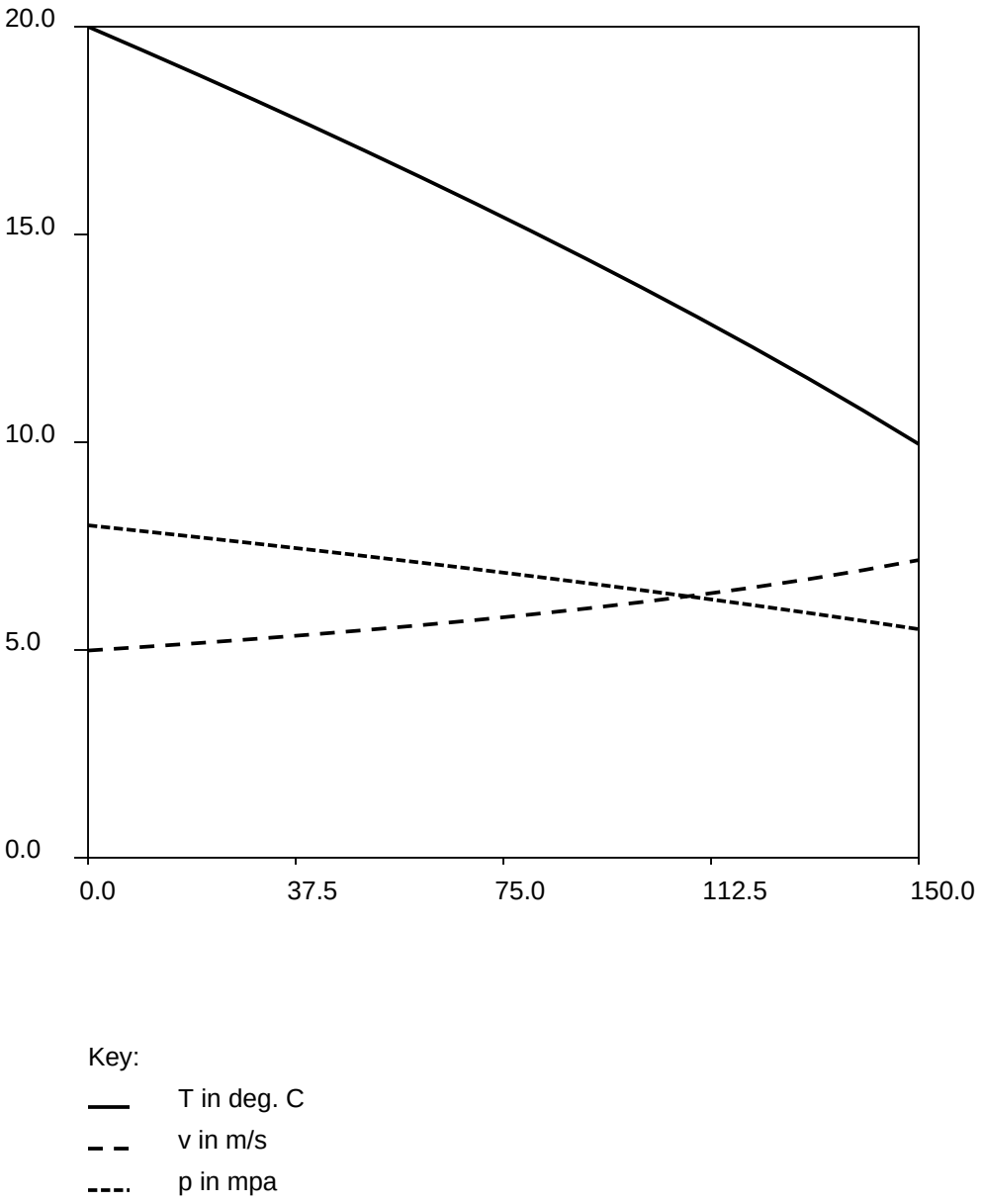

Figure 1: Steady State

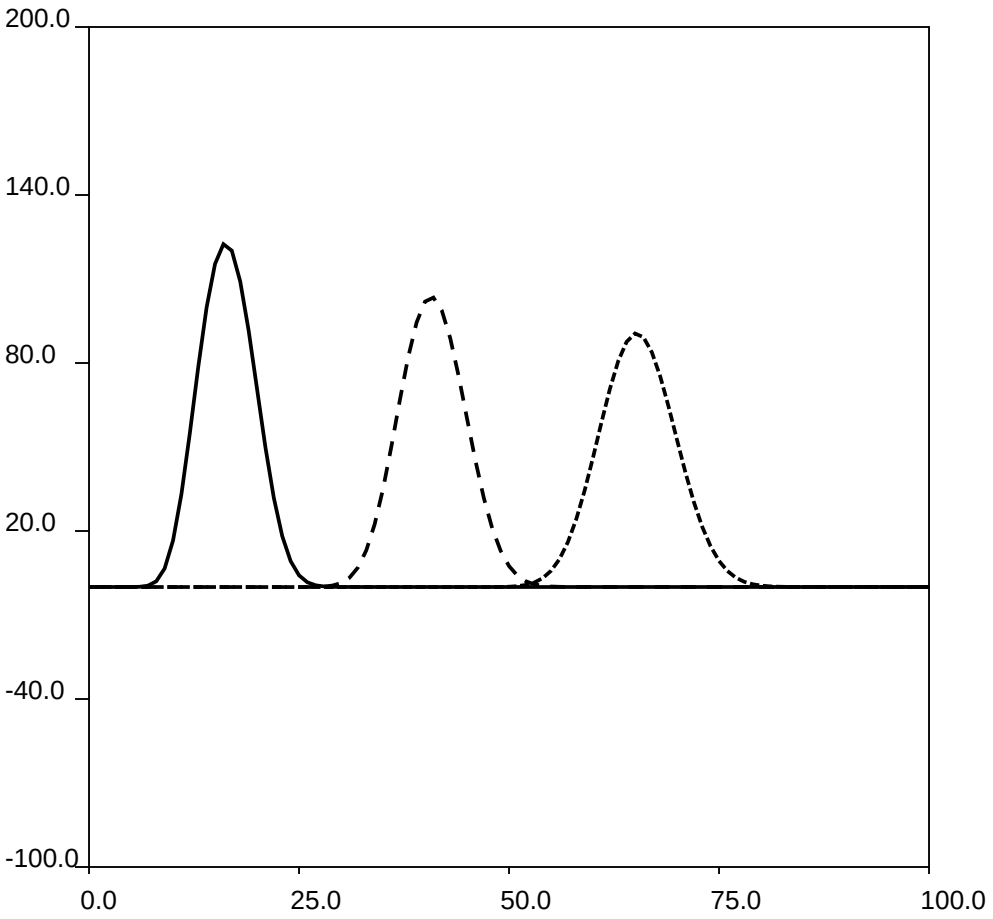

Figure 2: Small Amplitude Wave

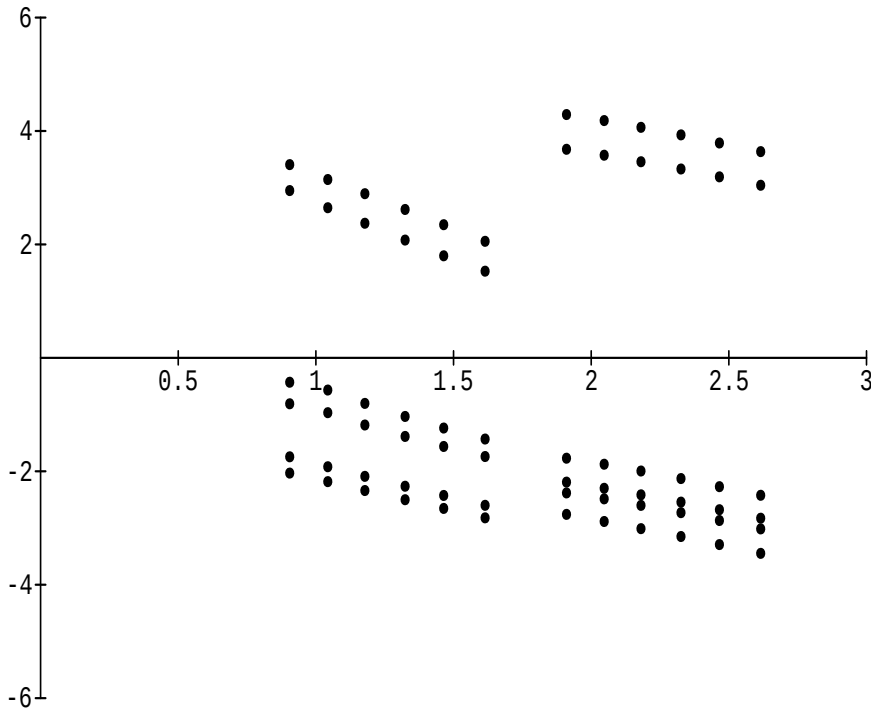

Figure 3: Convergence of the Error

very general nonlinear boundary conditions in the non-degenerate case, including ones involving time integrals. We analyze the boundary conditions using a simplified version of Luskin's analysis, but we do not treat the complexities he considers. Although it is reasonable to expect the analysis to extend to many of the more complicated boundary conditions he describes, we have not undertaken the task of demonstrating this.

Extension 1 (Non-uniform Time Steps) We remark that the same numerical method and proofs hold when $\triangle t$ is allowed to vary in a smooth manner. That is, the three theorems apply when we assume that

$$|\triangle t^{n+1+\theta} - \triangle t^{n+\theta}| \leq K_0 \triangle t^2 \text{ for all } n,$$

and

$$\frac{\max_n \triangle t^{n+1/2}}{\min_m \triangle t^{m+1/2}} \leq K_0,$$

as both h and $\triangle t$ go to zero, for the constant K_0 of Assumption 3, which is independent of h and $\triangle t$.

Extension 2 (Linearizations) We also note that various other linearizations are possible. We completely linearized the nonlinear terms, but in practice one may drop various small lower order terms. In particular, one need only keep the terms identified in the analysis as helping terms in order to achieve numerical stability.

Extension 3 (The Periodic Case) In this section we describe a method for improving a case related to Theorem 1 to yield convergence for any $\bar{\tau}$ given h sufficiently small. In Theorem 2 we accomplished this by weakening the dependence of the problem on temperature. This is not always necessary, as we will show in the following simplified case. Suppose the pipeline problem is periodic, so that the pipe forms a closed ring. We use periodic boundary conditions, which eliminates the complexity of the boundary terms in the error analysis. We will introduce a function $\tilde{u}(x,t)$ which is order h close to the true solution u, but to which the numerically computed solution is h^2 close. The periodicity assumption removes technical problems about the smoothness of $\tilde{u}(x,t)$. This higher order convergence to a comparison function gives the little bit extra needed to extend Theorem 1. In particular, the term $C_0 h^2$ in (24) becomes $C_0 h^4$, which keeps the solutions bounded for any finite time.

In the rest of this section we make this more precise by sketching the necessary modifications to the analysis of Section 8. We assume the reader has already browsed Section 8. In that section the truncation error is defined in (35) and bounded in (38). Using Taylor's theorem we rewrite it here as

$$TE = te_1 h + te_2 h^2, \tag{21}$$

where te_1 is a smooth function consisting of derivatives of u evaluated where TE is, and te_2 consists of integral averages of similar terms. We now define a function Φ by the equation

$$\mathcal{L}(\Phi, \Phi - \Phi^-; \Phi^-) = te_1, \tag{22}$$

with periodic boundary conditions and zero initial data. Then Φ is smooth by the same assumption we make for u. We now let

$$\tilde{u} = u - h\Phi.$$

Now we go through the whole analysis of Section 8, changing only the definition of the discrete interpolant W. Rather than W being the interpolant of u, we let it be the interpolant of $\tilde{u}$. Note that discrete derivatives of W are still bounded as they depend on the smooth functions u and Φ. The matrix L of (66) is now the identity matrix. The entire proof goes through without change, except in (38), where the bound on TE is now second order because the first order terms have been canceled out by the construction of Φ. Note that we do not change the induction hypothesis — Ψ is still only first order small since it involves comparing U with u, not $\tilde{u}$. However, in (121) the h^2 term is now h^4. This carries through the equations of Section 8.5. In addition the h in the H terms is now an h^2. Thus the only real change is in (125) where the right hand side multiplier of h^2 becomes h^4. Thus (126) becomes

$$\begin{aligned} C_{n+1} \leq & \\ & h^2(2h^r + 4(C_J + (C_n + C_n^2)C_I h^q)t^{n+1})^{1/2} \\ & \cdot \exp(4(C_J + 2(C_n + C_n^2)C_I h^q)t^n). \end{aligned} \tag{23}$$

Thus $C(t)$ grows like h^2 times a fixed function of time, and hence can be kept below C_* for arbitrarily long times by making h sufficiently small.

Table 1: Approximate Convergence Rates

	steady state		small waves	
	L^2	L^∞	L^2	L^∞
pressure	2.0	1.9	0.9	0.9
velocity	1.1	1.2	0.9	0.9
temperature	1.3	1.4	1.0	0.9

8 Error Analysis

Note: for an easier to read proof of an important special case, the reader is referred to [1], which treats the linear case of the degenerate problem.

Before launching into the details of the proof we give a brief overview. The *a-priori* error bound will be based on using the discrete scheme on $U - W$ where W is a discrete interpolant of u. The "error equation" for $U - W$ is a version of the discrete scheme that is linearized about the true solution u. We introduce some notation to get this equation in manageable form; the actual error equation is (47). We then diagonalize the discrete scheme by changing variables; this follows the earlier work of Thomeé and of Luskin. Next we develop an evolution inequality (122) for certain norms of the error, using a discrete l^2 inner product of the diagonalized error equation with a test function which is the sum of three terms weighted with carefully chosen powers of h. In developing this evolution inequality there are many terms to estimate; these are summarized in a tableau and estimated one by one. Finally the evolution inequality is used to derive the error bounds.

The fact that $\bar{\tau}$ may be finite in Theorem 1 comes from the unusual form of the evolution inequality. In contrast to the usual Gronwall-type evolution inequality, which in the continuous time case looks like

$$g' = Cg + C_0 h^2,$$

where C and C_0 are constants independent of h, we have an evolution equation of the form

$$g' = Cg(1 + \frac{\sqrt{g}}{h} + \frac{g}{h^2}) + C_0 h^2, \tag{24}$$

for which solutions can blow up in finite time. In addition, the squared norm of the error, g, corresponds in our case to a non-symmetric energy ($K - H$ in the notation of section 8.5); we show that the symmetric part dominates the non-symmetric part.

8.1 The Error Equation

We consider the degenerate case of equation (8), with the numerical method given by equation (17). Recall that the non-degenerate case is a simplification of these equations.

Convention 1 *In what follows we let C be a generic constant whose value in any particular equation depends upon various Sobolev norms of A, λ, F, u, and the constant K_0 of Assumption 3, but which is otherwise* independent *of the discretization parameters h, $\triangle t$ and θ.*

Throughout this section assumptions 1 and 3 hold. In what follows we take $\mu = 3$ as in the thermal pipeline equations, but the proof works in general. The symbol d will be used to distinguish discrete operators from continuous ones.

Let us define two operators $\mathcal{L}$ and $\mathcal{L}^d$ as follows:

$$\begin{aligned}\mathcal{L}_k(a,b;c) &= a_{kt} + A_{kj}(c)a_{xj} + (A_{kj,i}(c)c_{xj} - F_{k,i}(c))b_i \\ &\quad + \delta_{k3}\left((\lambda(c)a_3)_x - \lambda_x(c)a_3 + (\lambda_{,i}(c)c_3b_i)_x - (\lambda_{,i}(c)b_i)_x c_3\right),\end{aligned} \tag{25}$$

and

$$\begin{aligned}\mathcal{L}^d_k(a,b;c) &= \partial_t a_k + A_{kj}(c)\partial_x a_j + (A_{kj,i}(c)\partial_x c_j - F_{k,i}(c))b_i \\ &\quad + \delta_{k3}\left(\partial_x(\lambda(c)a_3) - a_3\partial_x\lambda(c) + \partial_x(\lambda_{,i}(c)c_3b_i) - c_3\partial_x(\lambda_{,i}(c)b_i)\right).\end{aligned} \tag{26}$$

Then we may re-write (14) and (17) in a more compact form: the true solution u satisfies the equation

$$\mathcal{L}(u, u-u^-; u^-) + \mathrm{Quad}(u-u^-; u^-) = F(u^-), \tag{27}$$

while at $\hat{\mathcal{Q}}$ our discrete solution U satisfies

$$\mathcal{L}^d(U, U-U^-; U^-) = F(U^-). \tag{28}$$

Recall that U is piecewise linear in time, and each U^n_k is piecewise linear in space, except U^n_3, which is discontinuous piecewise constant, upwinded by $\lambda(U^{n-1})$. It will now be useful to introduce a discrete interpolant W of u. Such a function is defined in the same discrete space as U. Therefore $U - W$ is also in the discrete space, and thus is easier to analyze than $U - u$. We define W by $W^n_k(x_j) = u^n_k(x_j)$ for $k < 3$, and $W^n_3(x_{j+1/2}) = u^n_3(x_{j+1/2})$, with W^n_3 at the knots upwinded by $\lambda(U^{n-1})$.

We now define the total error

$$\Psi = u - U,$$

the discrete error

$$\zeta = W - U,$$

and the approximation error

$$e = u - W.$$

Under reasonable conditions we know that e is small; it thus suffices to show that ζ is small.

Consider the quantity

$$\eta = \mathcal{L}^d(\zeta, \zeta - \zeta^-; u^-) \quad \text{over } \hat{\mathcal{Q}}. \tag{29}$$

From (27) and (28) we see that

$$\begin{aligned}\eta = &\mathcal{L}^d(W, W - W^-; u^-) - \mathcal{L}(u, u - u^-; u^-) \\ &+ \hat{\mathcal{L}}^d(U, U - U^-; U^-, u^-) - \text{Quad}(u - u^-; u^-) \\ &+ F(u^-) - F(U^-),\end{aligned} \tag{30}$$

where

$$\hat{\mathcal{L}}^d(a, b; c, d) = \mathcal{L}^d(a, b; c) - \mathcal{L}^d(a, b; d), \tag{31}$$

and we used the fact that

$$\mathcal{L}^d(a + a', b + b'; c) = \mathcal{L}^d(a, b; c) + \mathcal{L}^d(a', b'; c). \tag{32}$$

It is also true that

$$\hat{\mathcal{L}}^d(U, U - U^-; U^-, u^-) = \hat{\mathcal{L}}^d(W, W - W^-; U^-, u^-) - \hat{\mathcal{L}}^d(\zeta, \zeta - \zeta^-; U^-, u^-). \tag{33}$$

Now if $U - u$ turns out to be small, the $\hat{\mathcal{L}}^d(\zeta)$ terms will just be small perturbations to the $\mathcal{L}^d(\zeta)$ terms; hence we finally write the discrete error equation

$$\mathcal{L}^d(\zeta, \zeta - \zeta^-; u^-) + \hat{\mathcal{L}}^d(\zeta, \zeta - \zeta^-; U^-, u^-) = TE + NL, \tag{34}$$

where the truncation error is

$$TE = \mathcal{L}^d(W, W - W^-; u) - \mathcal{L}(u, u - u^-; u^-) - \text{Quad}(u - u^-; u^-) \tag{35}$$

and the remaining nonlinear terms are

$$NL = \hat{\mathcal{L}}^d(W, W - W^-; U, u^-) + F(u^-) - F(U^-). \tag{36}$$

Thus the discrete error ζ satisfies a perturbed version of the equation satisfied by U. The perturbations vanish in the linear case; the ζ equation also has truncation error as its main inhomogeneous term.

Recall that P is defined by (13) or (18). To make the expansion of (34) manageable, we introduce the following generic objects, which may be functions, vectors, matrices, three tensors or four tensors, and are all smooth functions of their various arguments.

$$\begin{aligned}M &= M(u, u_x, \text{ and any other derivatives as needed}). \\ \bar{M} &= \bar{M}(\Psi, u, u_x, \ldots). \\ \bar{M}_p &= \bar{M}_p(P\Psi, u, u_x, \ldots).\end{aligned}$$

For instance, we can write $u + \partial_x u = M$, and $U = M - \Psi$, and $\partial_x U = M - \partial_x \Psi$.

We can write the identity

$$f(U) - f(u) = (U_i - u_i) \int_0^1 f_{,i}(u + s(U-u))ds$$

as

$$f(U) - f(u) = \bar{M}\Psi.$$

Other useful identities are

$$\begin{aligned}
f(PU) - f(Pu) &= \bar{M}_p P\Psi, \\
f(U)U - f(u)u &= \bar{M}\Psi, \\
f(U)\partial_x U - f(u)\partial_x u &= \bar{M}\Psi + M\partial_x\Psi + M(\partial_x\Psi)\Psi, \\
f(U)P\partial_x U - f(u)P\partial_x u &= \bar{M}\Psi + MP\partial_x\Psi + \bar{M}(P\partial_x\Psi)\Psi, \\
f(PU)P\partial_x U - f(Pu)P\partial_x u &= \bar{M}_p P\Psi + MP\partial_x\Psi + \bar{M}_p(P\partial_x\Psi)P\Psi.
\end{aligned}$$

We also note the discrete product rule

$$\partial_x(fg) = f_c\partial_x g + g_c\partial_x f,$$

and chain rule

$$\partial_x(f(g)) = \mathcal{I}(f'(g))\partial_x g,$$

where

$$\mathcal{I}(f'(g)) = \int_0^1 f'(g(x_-) + s(g(x_+) - g(x_-)))ds.$$

Finally, we define

$$\bar{\bar{M}} = \bar{M}\Psi,$$

and

$$\bar{\bar{M}}_p = \bar{M}_p P\Psi.$$

Convention 2 *The point of using the M notation is to simplify dealing with the nonlinear terms arising from the differences $U^- - u^-$. To avoid putting minus sign superscripts on everything, we declare that henceforth $\bar{M}$, $\bar{M}_p$, $\bar{\bar{M}}$ and $\bar{\bar{M}}_p$ should be interpreted as involving Ψ^- rather than Ψ.*

Let us write Q_3 for the third column of the matrix $Q = I - P$. Using this notation, we find that the truncation error is

$$\begin{aligned}
TE = &-\text{Quad}(u-u;u^-) + (\partial_t W - u_t) + MP(\partial_x W - u_x) \\
&+ M(W-u) + M(W - W^- - (u-u^-)) + M(u-u^-)(\partial_x u^- - u_x^-) \\
&+ Q_3\left(\partial_x(\lambda(u^-)W_3) - (\lambda(u^-)u_3)_x - W_3\partial_x\lambda(u^-) + u_3\lambda_x(u^-)\right. \\
&+ \partial_x(\lambda_{,i}(u^-)u_3^-(W_i - W_i^-)) - (\lambda_{,i}(u^-)u_3^-(u_i - u^-))_x \\
&\left. - u_3^-(\partial_x(\lambda_{,i}(u^-)(W_i - W_i^-)) - (\lambda_{,i}(u^-)(u_i - u_i^-))_x)\right). \qquad (37)
\end{aligned}$$

These are mainly the usual truncation terms for collocation and can be shown to be small by standard approximation theory methods or simple Taylor expansions; we find that for some C independent of h, $\triangle t$ and θ,

$$|TE|_{m^2} \leq C(h + \triangle t^2 + (\theta - \frac{1}{2})\triangle t), \quad \forall t \in \hat{\mathcal{P}}_t. \tag{38}$$

In the non-degenerate case, the h becomes an h^2, since there are no piecewise constants.

In both cases,

$$|P(TE^{n+1+\theta} - TE^{n+\theta})|_{m^2} \leq C(h\triangle t + \triangle t^2). \tag{39}$$

We pause to remark that

$$\begin{aligned} \partial_x M &= M, \\ \partial_x \bar{M} &= \bar{M} + \bar{M}\partial_x \Psi, \\ \text{and} \quad \partial_x \bar{M}_p &= \bar{M}_p + \bar{M}_p P \partial_x \Psi. \end{aligned}$$

We now turn to the terms in (34) containing $\hat{\mathcal{L}^d}$.

$$\begin{aligned} \hat{\mathcal{L}^d}(a,b;U^-,u^-) &= \bar{\bar{M}} P\partial_x a + \bar{\bar{M}} a + \bar{\bar{M}} b \\ &+ (\bar{\bar{M}} + MP\partial_x \Psi^- + \bar{M}(P\partial_x \Psi^-)\Psi^- b) \\ &+ Q_3 \left(\partial_x((\lambda(U^-) - \lambda(u^-))a_3) + a_3 \partial_x(\lambda(U^-) - \lambda(u^-)) + \partial_x(\bar{\bar{M}} Pb) \right. \\ &+ \left. M\partial_x(\bar{\bar{M}} Pb) + \Psi_3^- \partial_x((M + \bar{\bar{M}})Pb) \right), \end{aligned} \tag{40}$$

where we used the expansion

$$\begin{aligned} &U_3 \partial_x(\lambda_{,i}(U)b_i) - u_3 \partial_x(\lambda_{,i}(u)b_i) \\ &\quad = U_3 \partial_x((\lambda_{,i}(U) - \lambda_{,i}(u))b_i) + (U_3 - u_3)\partial_x(\lambda_{,i}(u)b_i) \\ &\quad = M\partial_x(\bar{\bar{M}} Pb) + \Psi_3 \partial_x((M + \bar{\bar{M}})Pb). \end{aligned}$$

This simplifies to

$$\begin{aligned} \hat{\mathcal{L}^d}(a,b;U^-,u) &= \bar{\bar{M}} P\partial_x a + \bar{\bar{M}} a + \bar{\bar{M}} b + (M + \bar{\bar{M}})(P\partial_x \Psi^-)b \\ &+ Q_3 \left(\partial_x((\lambda(U^-) - \lambda(u^-))a_3) - a_3 \partial_x(\lambda(U^-) - \lambda(u^-)) \right. \\ &+ \left. M\partial_x(\bar{\bar{M}} Pb) + \Psi_3^- \partial_x((M + \bar{\bar{M}})Pb) \right). \end{aligned} \tag{41}$$

Before deriving the analogous expression for $\mathcal{L}^d$, we note that u^- is smooth, so we can safely undo the $ab_x = (ab)_x - a_x b$ transformation used in (26). In particular,

$$\begin{aligned} &\partial_x(\lambda_{,i}(u^-)u_3^- b_i) - u_3^- \partial_x(\lambda_{,i}(u^-)b_i) \\ &\quad = (\partial_x u_3^-)(\lambda_{,i}(u^-)b_i)_{,c} + (u_{3,c}^- - u_3^-)\partial_x(\lambda_{,i}(u^-)b_i) \\ &\quad = MPb + h^2 M\partial_x(MPb). \end{aligned} \tag{42}$$

Thus

$$\mathcal{L}^d(a,b;u^-) = \partial_t a + A^- P\partial_x a + Ma + Mb \\ + Q_3\left(\partial_x(\lambda^- a_3) + h^2 M\partial_x(MPb)\right), \tag{43}$$

where we have written A^- and λ^- for $A(u^-)$ and $\lambda(u^-)$, rather than M, for future convenience. We can now expand (34) in a reasonably simple manner:

$$\begin{aligned}
&\partial_t\zeta + A^- P\partial_x\zeta + M\zeta + M(\zeta-\zeta^-) \\
&\quad + Q_3\left(\partial_x(\lambda^-\zeta_3) + M\partial_x(MP(\zeta-\zeta^-))\right) \\
&\quad + \bar{\bar{M}}P\partial_x\zeta + \bar{\bar{M}}\zeta + \bar{\bar{M}}(\zeta-\zeta^-) \\
&\quad + (M+\bar{\bar{M}})(P\partial_x\Psi^-)(\zeta-\zeta^-) \\
&\quad + Q_3\Big(\partial_x((\lambda(U^-)-\lambda^-)\zeta_3) + \zeta_3\partial_x\bar{\bar{M}}_p + M\partial_x(\bar{\bar{M}}P(\zeta-\zeta^-)) \\
&\quad + \Psi_3^-\partial_x((M+\bar{\bar{M}})P(\zeta-\zeta^-))\Big) = TE + NL.
\end{aligned} \tag{44}$$

We note that in (36), there are still ζ terms:

$$F(u^-) - F(U^-) = \bar{\bar{M}}\triangle t + \bar{M}\triangle t^2 + \bar{M}\triangle t\partial_t\zeta, \tag{45}$$

where we used $u - U = e + \zeta$ and

$$\begin{aligned}
&F(u)(u-u^-) - F(U)(U-U^-) = \\
&\quad (F(u)-F(U))(u-u^-) + F(U)[(u-u^-)-(U-U^-)].
\end{aligned}$$

We will put this $\partial_t\zeta$ term on the left hand side, below.

Since W depends only on u, we have

$$\begin{aligned}
\hat{\mathcal{L}}^d(W, W-W^-; U^-, u^-) &= \bar{\bar{M}} + (M+\bar{\bar{M}})(P\partial_x\Psi^-)\triangle t \\
&\quad + Q_3\left(\partial_x((\lambda(U^-)-\lambda^-)W_3) - W_3\partial_x(\lambda(U^-)-\lambda^-)\right. \\
&\quad \left. + M\partial_x(\bar{\bar{M}}\triangle t) + \Psi_3^-\partial_x((M+\bar{\bar{M}})\triangle t)\right).
\end{aligned} \tag{46}$$

Thus (44) becomes, via (31, 32, 27, 29, 22):

$$LHS = TE + IND, \tag{47}$$

where the new terms are

$$\begin{aligned}
IND &= \bar{\bar{M}} + \bar{M}\triangle t^2 + (M+\bar{\bar{M}})(P\partial_x\Psi^-)\triangle t \\
&\quad + Q_3\left(\bar{\bar{M}}_p + hM\partial_x(P\Psi^-) + M\partial_x(\bar{\bar{M}}\triangle t) + \Psi_3^-\partial_x((M+\bar{\bar{M}})\triangle t)\right),
\end{aligned} \tag{48}$$

and

$$\begin{aligned} LHS = {} & \partial_t\zeta + A^- P\partial_x\zeta + M\zeta + (M+\bar{M})\triangle t\partial_t\zeta \\ & + \bar{\bar{M}}P\partial_x\zeta + \bar{\bar{M}}\zeta + \bar{\bar{M}}\triangle t\partial_t\zeta + (M+\bar{\bar{M}})(P\partial_x\Psi^-)\triangle t\partial_t\zeta \\ & + Q_3\left(\partial_x(\lambda^-\zeta_3) + h^2M\partial_x(MP\triangle t\partial_t\zeta) + \partial_x((\lambda(U^-)-\lambda^-)\zeta_3)\right. \\ & \left. + \zeta_3\partial_x\bar{\bar{M}}_p + M\partial_x(\bar{\bar{M}}P\triangle t\partial_t\zeta) + \Psi_3^-\partial_x((M+\bar{\bar{M}})P\triangle t\partial_t\zeta)\right). \end{aligned} \tag{49}$$

This simplifies to

$$\begin{aligned} LHS = {} & \partial_t\zeta + (M+\bar{M}+\bar{\bar{M}}+(M+\bar{\bar{M}})(P\partial_x\Psi^-))\triangle t\partial_t\zeta \\ & + A^-P\partial_x\zeta + \bar{\bar{M}}P\partial_x\zeta + (M+\bar{\bar{M}})\zeta \\ & + Q_3\left(\partial_x(\lambda^-\zeta_3) + \partial_x((\lambda(U^-)-\lambda^-)\zeta_3) + \zeta_3\partial_x\bar{\bar{M}}_p\right. \\ & \left. + \bar{\bar{M}}\partial_x((M+\bar{\bar{M}})P\partial_t\zeta\triangle t) + M\partial_x((h^2M+\bar{\bar{M}})P\partial_t\zeta\triangle t)\right). \end{aligned} \tag{50}$$

In this form one can see the linear terms and how they are perturbed by *small* nonlinear terms, subject to some induction hypotheses which we formalize later, to the effect that Ψ^- and $P\partial_x\Psi^-$ are small, whence $\bar{\bar{M}}$ and $\partial_x\bar{\bar{M}}_p$ are also small.

To facilitate the diagonalization process described below, we make one more adjustment to this equation. We recall that $A = AP$ and $\bar{A} = AP + \lambda Q$, and that A^- and λ^- depend only on u^- and hence are smooth in space. Thus we may write

$$\begin{aligned} & A^-P\partial_x\zeta + Q_3\partial_x(\lambda^-\zeta_3) = \\ & \qquad \partial_x(\bar{A}^-\zeta) + MP\zeta_{,c} + h^2MP\partial_x\zeta. \end{aligned} \tag{51}$$

8.2 Diagonalization

Before proceeding further with the error analysis of (47), we must first change variables in order to diagonalize the matrix $\bar{A}$ in (51). This is a standard step, also used by Luskin; we remark that it is done only in the proof, not in the numerical computation.

From here on in we need to distinguish two parallel threads in the proof, one for the degenerate case and one for the non-degenerate case. In the non-degenerate case $P = I, Q = 0$, and there are no piecewise constants to keep track of, which simplifies matters to the point where we can prove an H^1 estimate. In the degenerate case we only get L^2 convergence. As much as possible, we will do the two cases together.

The diagonalization is a bit messy, but using our M notation it is not too hard to write it all out.

We define two new matrices S and R, which will be smooth functions of x and t. We define S to be a bounded matrix of column eigenvectors of $\bar{A}$, and $R = S^{-1}$. By assumption in Section 4 the matrices R and S are smooth and bounded functions

of u. We find that in the degenerate case, both S and R can be selected to have the form

$$\begin{pmatrix} * & * & 0 \\ * & * & 0 \\ * & * & 1 \end{pmatrix}.$$

Thus $S = SP + Q$, and $R = RP + Q$.

We define $\Lambda = R\bar{A}S = \mathrm{diag}(\lambda_1, \lambda_2, \lambda)$. By hypothesis, Λ is a bounded function of u. Also for any *scalar* function f, we note that $RQf = Qf$.

We now define, for all x and t,

$$\omega(x,t) = R(x,t)\zeta(x,t). \tag{52}$$

We note that $P\omega = MP\zeta$ and $P\zeta = MP\omega$.

Although ζ is piecewise linear in time, ω is not. However,

$$\omega^{n+\theta} = R^{n+\theta}\zeta^{n+\theta} = R^{n+\theta}(\theta\zeta^{n+1} + (1-\theta)\zeta^n),$$

since ζ *is* piecewise linear time; hence using $\omega = M\zeta$ and $\zeta = M\omega$ we see $\omega^{n+\theta} = M\omega^{n+1} + M\omega^n$, or more compactly

$$\omega = M\omega^+ + M\omega^- \quad \text{in } \hat{\mathcal{P}}_t \quad \text{for all } x. \tag{53}$$

We compute the following transformations over $\hat{\mathcal{Q}}$, writing $S^{+1/2} = S(t^{n+1/2})$:

$$\partial_t\zeta = \partial_t(S\omega) = S_{,c}\partial_t\omega + (\partial_t S)\omega^{,c} = S^{+1/2}\partial_t\omega + \triangle t^2 MP\partial_t\omega + MP\omega^{,c}, \tag{54}$$

in which we note that any derivative of R or S introduces a factor of P, and similarly,

$$\partial_x(\bar{A}^-\zeta) = S^{+1/2}\partial_x(\Lambda^-\omega) + \triangle tMP\partial_x\omega + MP\omega_{,c}, \tag{55}$$

$$P\partial_x\zeta = MP\partial_x\omega + MP\omega_{,c}, \tag{56}$$

and

$$\zeta = S\omega = S^{+1/2}\omega + \triangle tMP\omega. \tag{57}$$

We left multiply equation (47) by $R^{+1/2}$ to create a diagonalized error equation, obtaining

$$\begin{aligned} R^{+1/2}\cdot LHS = {} & \partial_t\omega + \triangle t^2 MP\partial_t\omega + MP\omega^{,c} \\ & + (M + \bar{M} + \bar{\bar{M}} + (M + \bar{\bar{M}})(P\partial_x\Psi^-))\triangle t(M\partial_t\omega + MP\omega^{,c}) \\ & + \partial_x(\Lambda^-\omega) + (\triangle tM + h^2M + \bar{\bar{M}})P\partial_x\omega \end{aligned}$$

$$
\begin{aligned}
&+ (M + \bar{\bar{M}})P\omega_{,c} + (M + \bar{\bar{M}})\omega \\
&+ Q_3 \left(\partial_x((\lambda(U^-) - \lambda^-)\zeta_3) + M\omega\partial_x\bar{\bar{M}}_p \right. \\
&\left. + \bar{\bar{M}}\partial_x((M + \bar{\bar{M}})P\partial_t\zeta\triangle t) + M\partial_x((h^2M + \bar{\bar{M}})P\partial_t\zeta\triangle t) \right) \\
&= M \cdot TE + M \cdot IND. \qquad (58)
\end{aligned}
$$

Notice that there is essentially no change in the right hand side terms. We write $DLHS$ for the diagonalized left hand side of this equation. We simplify it, and also combine the term $Q_3(\partial_x((\lambda(U^-) - \lambda^-)P\zeta_3))$ with the Λ^- matrix, which simply makes the (3,3) entry of Λ^- depend on PU^- rather than Pu^-. This is exactly what we need to handle the upwinding terms, since the upwinding is based on PU^-, rather than Pu^-. Henceforth, Λ^- will mean this modified version. This leaves us with

$$
\begin{aligned}
DLHS = &\partial_t\omega + (M + \bar{M} + \bar{\bar{M}} + (M + \bar{\bar{M}})(P\partial_x\Psi^-))\triangle t\partial_t\omega + \partial_x(\Lambda^-\omega) \\
&+ (\triangle tM + h^2M + \bar{\bar{M}})P\partial_x\omega + (M + \bar{\bar{M}})\omega \\
&+ (M + (\bar{M} + \bar{\bar{M}} + (M + \bar{\bar{M}})(P\partial_x\Psi^-))\triangle t)MP\omega^{,c} + (M + \bar{\bar{M}})P\omega_{,c} \\
&+ Q_3 \left(\partial_x(\bar{\bar{M}}_p P\omega) + M\omega\partial_x\bar{\bar{M}}_p \right. \\
&\left. + \bar{\bar{M}}\partial_x((M + \bar{\bar{M}})P\partial_t\zeta\triangle t) + M\partial_x((h^2M + \bar{\bar{M}})P\partial_t\zeta\triangle t) \right). \qquad (59)
\end{aligned}
$$

We define

$$\epsilon_t = \triangle t(M + \bar{M} + \bar{\bar{M}} + (M + \bar{\bar{M}})P\partial_x\Psi^-), \qquad (60)$$

and

$$\epsilon_x = \triangle tM + h^2M + \bar{\bar{M}}. \qquad (61)$$

Since ζ is piecewise linear in time, and $P\zeta$ is in space, we can write

$$P\omega^{,c} = P\omega + M\triangle tP\partial_t\omega, \qquad (62)$$

and

$$P\omega_{,c} = P\omega + MhP\partial_x\omega. \qquad (63)$$

Thus we also define

$$\epsilon_z = \bar{\bar{M}} + \triangle t(\bar{M} + (M + \bar{\bar{M}})P\partial_x\Psi^-)P, \qquad (64)$$

whence (59) simplifies to

$$DLHS = \boxed{A}\partial_t\omega + \boxed{A'}\epsilon_t\partial_t\omega + \boxed{B}\partial_x(\Lambda^-\omega) + \boxed{B'}\epsilon_x P\partial_x\omega + \boxed{D}(M + \epsilon_z)\omega$$

$$+ Q_3 \left(\boxed{F} \partial_x(\bar{\bar{M}}_p P\omega) + \boxed{G} M\omega\partial_x \bar{\bar{M}}_p \right.$$

$$\left. + \bar{\bar{M}}\partial_x((M+\bar{\bar{M}})P\partial_t\zeta\triangle t) + \boxed{H} M\partial_x((h^2M+\bar{\bar{M}})P\partial_t\zeta\triangle t) \right)$$

$$= M\cdot TE + M\cdot IND = \boxed{C}\ \rho\ , \tag{65}$$

where we label each of the nine terms with a letter for future convenience and lump the two right hand side terms together as ρ.

We have written the general case here; recall that in the non-degenerate case, the only change is that Q becomes zero.

Equation (65) looks very much like the corresponding error equation in the linear case, except for the addition of terms which we hope to show are small, by induction. Note that (65) holds at every point of $\hat{\mathcal{Q}}$.

Let

$$L = \mathrm{diag}(l_1, l_2, l_3), \tag{66}$$

be a diagonal matrix where

$$l_k(x,t) = g(\lambda_k(0,t))(1-\frac{x}{E}) + g(-\lambda_k(E,t))\frac{x}{E}, \text{ for } k = 1,2, \tag{67}$$

and

$$l_3(x,t) = \sigma_0, \tag{68}$$

where

$$g(\lambda) = \begin{cases} 1 & \text{if } \lambda < 0, \\ \sigma_0 & \text{if } \lambda \geq 0. \end{cases} \tag{69}$$

Here σ_0 is a small positive constant to be chosen later, *independent* of h and $\triangle t$.

Note that each l_k is constant in time for each fixed x since the signs of λ_1 and λ_2 never change. Each l_k is also linear in space.

Let us now define r by

$$r = \begin{cases} 1 & \text{in the context of Theorem 1,} \\ 1/2 & \text{in the context of Theorem 2, or} \\ 0 & \text{in the context of Theorem 3.} \end{cases}$$

We now define, at each point in $\hat{\mathcal{Q}}$, the test function we will use for the energy analysis:

$$\varphi = \boxed{1} L\omega + \alpha h^r \boxed{2} P\partial_t\omega + \beta h^r \boxed{3} PL\partial_t\partial_x(\Lambda\omega), \tag{70}$$

where we leave unspecified two non-negative parameters α and β, which will be determined later and will be independent of h and $\triangle t$.

8.3 The 27 Product Terms

We form the vector inner product of both sides of the equation (65) with φ, producing another equation involving 27 product terms which must be considered individually. The following chart summarizes the situation.

		A	A'	B	B'	C	D	F	G	H
$-$	$+$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$
1		L	R	R	R	R	R	R	R	R
2		L	R	R	R	R	R	0	0	0
3		R	R	L	R	R	R	0	0	0

Since $P \cdot Q = 0$, 6 terms vanish automatically. Three other terms marked with an "L" are "helping" or "left hand side terms"; the other 18 are right hand side terms. The analysis of each term proceeds just as in the linear case. The product equation $DLHS \cdot \varphi = \rho \cdot \varphi$ holds at every point of $\hat{\mathcal{Q}}$; for each time level $t^{n+\theta}$ in $\hat{\mathcal{P}}_t$ we multiply by h and sum over all x's in $\hat{\mathcal{P}}_x$, thus forming the discrete spatial midpoint-based m^2 norm. Some right hand side terms will be hidden by direct subtraction; later we will use a time-induction form of Gronwall's inequality to handle the rest.

For each right hand side term we give an upper bound for the sum over $\hat{\mathcal{P}}_x$. For the three terms marked L, we give a positive lower bound instead. The bounds may not be obvious at first, but they follow in straightforward ways from the properties of the objects involved, in particular from knowing that ζ is piecewise linear in time and either piecewise linear or piecewise constant in space.

We begin with the three "helping terms" $A-1$, $A-2$, and $B-3$.

8.3.1 A-1

We write the steps out in some detail for this first left-hand side term:

$$\partial_t\omega \cdot L\omega \geq \frac{1}{2}\partial_t(L\omega \cdot \omega) + (\theta - \frac{1}{2})\triangle t\sigma_0(\partial_t\omega)^2 - C((\omega^-)^2 + (\omega^+)^2), \tag{71}$$

since L is constant in time, linear in x and bounded below by the positive constant σ_0. The above equation holds at each point of $\hat{\mathcal{Q}}$, hence for each point in $\hat{\mathcal{P}}_t$,

$$\sum_{\hat{\mathcal{P}}_x} \partial_t \omega \cdot L\omega h \geq \frac{\sigma_0}{2}\partial_t |\omega|^2_{m^2} + (\theta - \frac{1}{2})\sigma_0 \triangle t |\partial_t \omega|^2_{m^2} - C(|\omega^-|^2_{m^2} + |\omega^+|^2_{m^2}).$$

Here C is a generic constant independent of h, $\triangle t$; it can depend on σ_0 and as always, on norms of u.

8.3.2 A-2

$$\sum_{\hat{\mathcal{P}}_x} \partial_t \omega \cdot \alpha h^r P \partial_t \omega h \geq \alpha h^r |P\partial_t \omega|^2_{m^2}. \tag{72}$$

8.3.3 B-3

$$\begin{aligned}
&\sum_{\hat{\mathcal{P}}_x} \partial_x(\Lambda^- \omega) \cdot \beta h^r PL \partial_t \partial_x (\Lambda \omega) h \\
&\quad \geq \beta h^r \left(\frac{\sigma_0}{2}\partial_t |P\partial_x(\Lambda\omega)|^2_{m^2} + (\theta - \frac{1}{2})\sigma_0 \triangle t |P\partial_t \partial_x(\Lambda\omega)|^2_{m^2} \right. \\
&\qquad \left. - C(|P\partial_x(\Lambda^-\omega^-)|^2_{m^2} + |P\partial_x(\Lambda^+\omega^+)|^2_{m^2}) \right),
\end{aligned} \tag{73}$$

plus a right hand side term of the same form as $B' - 3$ which comes from writing $\Lambda^- = \Lambda + (\Lambda^- - \Lambda)$. We ignore this term here since it will be treated automatically under $B' - 3$.

We now turn to giving upper bounds for right hand side terms. We do some easy terms first.

8.3.4 D-1

$$\sum_{\hat{\mathcal{P}}_x} (M + \epsilon_z)\omega \cdot L\omega h \leq C(1 + |\epsilon_z|_{m^\infty})|\omega|^2_{m^2}. \tag{74}$$

8.3.5 B-2

$$\sum_{\hat{\mathcal{P}}_x} \partial_x(\Lambda^- \omega) \cdot \alpha h^r P \partial_t \omega h \leq \frac{\alpha h^r}{64} |P\partial_t \omega|^2_{m^2} + \alpha h^r C |P\partial_x(\Lambda^-\omega)|^2_{m^2}. \tag{75}$$

8.3.6 D-2

$$\sum_{\hat{\mathcal{P}}_x} (M + \epsilon_z)\omega \cdot \alpha h^r P \partial_t \omega \, h \leq \frac{\alpha h^r}{64} |P\partial_t \omega|^2_{m^2} + \alpha h^r C (1 + |\epsilon_z|_{m^\infty})^2 |\omega|^2_{m^2}. \tag{76}$$

8.3.7 C-1

We now rewrite

$$\rho = Mh + \bar{M}\Psi^- + \bar{M}h^2 + MhP\partial_x\Psi^-.$$

We now find it useful to expand this as

$$\rho = Mh + \bar{M}h^2 + \bar{M}\omega + \bar{M}\triangle t\partial_t\omega + MhP\partial_x\omega + Mh^2P\partial_t\partial_x\omega. \tag{77}$$

We use this in the C terms.

$$\sum_{\hat{\mathcal{P}}_x} \rho \cdot L\omega h \leq$$
$$C(h^2 + |\omega|^2_{m^2} + h^2|\partial_t\omega|^2_{m^2} + h^2|P\partial_x\omega|^2_{m^2} + h^4|P\partial_t\partial_x\omega|^2_{m^2}). \tag{78}$$

8.3.8 C-2

$$\sum_{\hat{\mathcal{P}}_x} \rho \cdot \alpha h^r P\partial_t\omega h \leq \frac{\alpha h^r}{64}|P\partial_t\omega|^2_{m^2} + Ch^{2+r}$$
$$+ \alpha h^r C(|\omega|^2_{m^2} + h^2|\partial_t\omega|^2_{m^2} + h^2|P\partial_x\omega|^2_{m^2} + h^4|P\partial_t\partial_x\omega|^2_{m^2}). \tag{79}$$

8.3.9 B-1

This term contains helping terms as well as right hand side terms, so we give a lower bound.

$$\sum_{\hat{\mathcal{P}}_x} \partial_x(\Lambda^-\omega) \cdot L\omega h \geq$$
$$\frac{1}{2}\sum_{k=1}^{3}(\lambda_k^- l_k\omega_k^2)|_{x=0}^{x=E} - C|\omega_k|^2_{l^2} - Ch|P\partial_x\omega|^2_{m^2}. \tag{80}$$

Note that $|\omega_k|^2_{l^2} \leq C(|\omega_k|^2_{m^2} + h^2|P\partial_x\omega_k|^2_{m^2})$. In the non-degenerate case the first order term in h becomes second order, so an L^2 estimate with $r = 2$ is possible. In the degenerate case, the presence of the first order term requires $r \leq 1$, thus giving a slightly better than L^2 result.

This analysis of the $B-1$ term arises from detailed consideration of the form of products of combinations of piecewise linear and piecewise constant functions and their derivatives.

8.3.10 A-3

$$\sum_{\hat{\mathcal{P}}_x} \partial_t\omega \cdot \beta h^r PL\partial_t\partial_x(\Lambda\omega)h :$$

We note that ω_3 does not appear in this term, so only piecewise linear functions need be considered. We write

$$\partial_t\partial_x(\Lambda\omega) = \partial_x(\boxed{a}\omega^{'c}) + \partial_x(\boxed{b}\partial_t\omega) .$$

From (a) we have

$$\beta h^r \sum_{\hat{\mathcal{P}}_x} \partial_t\omega \cdot PL\partial_x((\partial_t\Lambda)\omega^{'c})h \leq$$
$$\frac{\alpha h^r}{64}|P\partial_t\omega|^2_{m^2} + C\beta h^r|P\partial_x\omega^{'c}|^2_{m^2}, \tag{81}$$

while from (b), we get

$$\beta h^r \sum_{\hat{\mathcal{P}}_x} \partial_t\omega \cdot PL\partial_x(\Lambda^{'c}\partial_t\omega)h \geq$$
$$\frac{\beta h^r}{2}\sum_{k=1}^{2}(\lambda_k^{'c} l_k(\partial_t\omega_k)^2)|_{x=0}^{x=E} - \beta h^r C_1|P\partial_t\omega|^2_{l^2}. \tag{82}$$

We then bound

$$\beta h^r C_1|P\partial_t\omega|^2_{l^2} \leq \beta h^r C_1|P\partial_t\omega|^2_{m^2} + \beta h^r Ch^2|P\partial_t\partial_x\omega|^2_{m^2}.$$

For β sufficiently small relative to α, the first term hides, and since $h \sim \triangle t$ the second term is bounded by

$$\beta h^r C_1(|P\partial_x\omega^+|^2_{m^2} + |P\partial_x\omega^-|^2_{m^2}).$$

The spatial boundary terms in (80) and (82) turn out to give non-negative helping terms, provided σ_0 is chosen sufficiently small relative to certain $O(1)$ constants depending only on u. This follows from the form of L, which is carefully chosen based on a trick used by Luskin and pioneered by Thomeé. Essentially it works as follows.

Consider the term in (80) at $x = 0$; by hypothesis on the signs of the λ_i, this is

$$-\sigma_0|\lambda_1|\omega_1^2 + |\lambda_2|\omega_2^2 - \sigma_0\lambda\omega_3^2.$$

If $\lambda > 0$, then $\zeta_3(0) = 0$ by choice of boundary conditions. Since $|\lambda| \ll |\lambda_1|$ and $|\lambda| \ll |\lambda_2|$, we can bound this term below by

$$-\sigma_0(|\lambda_1| + |\lambda|)\omega_1^2 + (|\lambda_2| - |\lambda|)\omega_2^2.$$

This is positive for σ_0 sufficiently small, given physically realizable boundary conditions. That is, we don't want $\omega_2(0)$ to vanish unless $\omega_1(0)$ also vanishes. For instance, if the boundary conditions are that we specify the pressure at both ends of the pipe, then $\zeta_1(0) = \zeta_1(E) = 0$, whence $\omega_1(0)$ and $\omega_2(0)$ are *proportional* by a constant depending on $A(u)$ and not on h or $\triangle t$. Similar arguments apply at $x = E$ and to the $\partial_t \omega$ terms from (82).

8.3.11 C-3

We "sum by parts in time" on the $Mh + \bar{M}\omega$ terms of ρ and bound the rest directly.

$$\begin{aligned}\sum_{\hat{\mathcal{P}}_x} \rho \cdot \beta h^r PL\partial_t\partial_x(\Lambda\omega)h \leq \\ h^{r-1}(h^2|\partial_t\omega|^2_{m^2} + h^2|P\partial_x\omega|^2_{m^2} + h^4|P\partial_t\partial_x\omega|^2_{m^2} + Ch^4) \\ + C\beta^2 h^{r+1}|P\partial_t\partial_x\omega|^2_{m^2} \\ + \sum_{\hat{\mathcal{P}}_x}(Mh + \bar{M}\omega)\cdot\beta h^r PL\partial_t\partial_x(\Lambda\omega)h. \end{aligned} \tag{83}$$

We recall a formula for summation by parts in time at $t = t^{n+\theta}$:

$$a^{n+\theta}\partial_t b = \frac{1}{\triangle t}a^{n+\theta}(b^+ - b^-) = \frac{1}{\triangle t}(a^{n+\theta}b^+ - a^{n-1+\theta}b^-) - \frac{1}{\triangle t}(a^{n+\theta} - a^{n-1+\theta})b^+.$$

Thus we write

$$\begin{aligned}\sum_{\hat{\mathcal{P}}_x}(Mh + \bar{M}\omega)\cdot\beta h^r PL\partial_t\partial_x(\Lambda\omega)h = \\ \frac{\beta h^r}{\triangle t}\sum_{\hat{\mathcal{P}}_x}((Mh + \bar{M}\omega)\cdot PL\partial_x(\Lambda^+\omega^+) - (Mh + \bar{M}\omega)^{n-1+\theta}\cdot PL\partial_x(\Lambda^-\omega^-))h \\ - \beta h^{r-1}\sum_{\hat{\mathcal{P}}_x}((Mh + \bar{M}\omega) - (Mh + \bar{M}\omega)^{n-1+\theta})\cdot PL\partial_x(\Lambda^+\omega^+)h. \end{aligned} \tag{84}$$

We can bound the second sum by

$$C\beta h^r|P\partial_x(\Lambda^+\omega^+)|^2_{m^2} + \frac{\alpha h^r}{64}|\partial_t\omega|^2_{m^2} + Ch^{2+r} + C\beta h^r|\partial_t\Psi^-|^2_{m^\infty}|\omega|^2_{m^2}.$$

8.3.12 D-3

As in $C-3$, we sum by parts on the $M\omega$ term, which we need not repeat here. We write $\epsilon_z = \bar{M} + C\epsilon_t$, and again sum by parts on the $\bar{M}$ term, while on the ϵ_t term we use the bound

$$\sum_{\hat{\mathcal{P}}_x} \epsilon_t \omega \cdot \beta h^r PL\partial_t \partial_x (\Lambda\omega) h \leq$$
$$C|\epsilon_t|^2_{m^\infty} h^{-3/2} |\omega|^2_{m^2} + \beta^2 h^{2r+3/2} |P\partial_t \partial_x \omega|^2_{m^2}. \tag{85}$$

The summation by parts on the $\bar{\bar{M}}$ term is

$$\sum_{\hat{\mathcal{P}}_x} \bar{\bar{M}}\omega \cdot \beta h^r PL\partial_t \partial_x (\Lambda\omega) h =$$
$$\frac{\beta h^r}{\triangle t} \sum_{\hat{\mathcal{P}}_x} (\bar{\bar{M}}\omega \cdot PL\partial_x(\Lambda^+\omega^+) - (\bar{\bar{M}}\omega)^{n-1+\theta} \cdot PL\partial_x(\Lambda^-\omega^-))h$$
$$- \beta h^{r-1} \sum_{\hat{\mathcal{P}}_x} (\bar{\bar{M}}\omega - (\bar{\bar{M}}\omega)^{n-1+\theta}) \cdot PL\partial_x(\Lambda^+\omega^+)h. \tag{86}$$

We can bound the second sum by

$$C\beta h^r |P\partial_x(\Lambda^+\omega^+)|^2_{m^2} + \frac{\alpha h^r}{64} |\bar{\bar{M}}|^2_{m^\infty} |\partial_t \omega|^2_{m^2}$$
$$+ \beta h^r (|\partial_t \Psi^-|^2_{m^\infty} + |\Psi^-|^2_{m^\infty}) |\omega|^2_{m^2}.$$

8.3.13 A′-1

$$\sum_{\hat{\mathcal{P}}_x} \epsilon_t \partial_t \omega \cdot L\omega h \leq$$
$$C|\omega|^2_{m^2} + \frac{\alpha}{64} |\epsilon_t|^2_{m^\infty} |\partial_t \omega|^2_{m^2}. \tag{87}$$

8.3.14 A′-2

$$\sum_{\hat{\mathcal{P}}_x} \epsilon_t \partial_t \omega \cdot \alpha h^r P\partial_t \omega h \leq$$
$$\alpha h^r |\epsilon_t|_{m^\infty} |\partial_t \omega|^2_{m^2}. \tag{88}$$

8.3.15 A′-3

$$\sum_{\hat{\mathcal{P}}_x} \epsilon_t \partial_t \omega \cdot \beta h^r PL\partial_t \partial_x (\Lambda\omega) h \leq$$
$$|\epsilon_t|^2_{m^\infty} h^{r-3/2} |\partial_t \omega|^2_{m^2} + Ch^{r+3/2} |P\partial_t \partial_x (\Lambda\omega)|^2_{m^2}. \tag{89}$$

8.3.16 B′-1

$$\sum_{\hat{\mathcal{P}}_x} \epsilon_x P\partial_x\omega \cdot L\omega h \leq |\omega|^2_{m^2} + |\epsilon_x|^2_{m^\infty}|P\partial_x\omega|^2_{m^2}. \tag{90}$$

8.3.17 B′-2

$$\sum_{\hat{\mathcal{P}}_x} \epsilon_x P\partial_x\omega \cdot \alpha h^r P\partial_t\omega h \leq \frac{\alpha h^r}{64}|P\partial_t\omega|^2_{m^2} + C\alpha h^r|\epsilon_x|^2_{m^\infty}|P\partial_x\omega|^2_{m^2}. \tag{91}$$

8.3.18 B′-3

$$\begin{aligned}\sum_{\hat{\mathcal{P}}_x} &\epsilon_x P\partial_x\omega \cdot \beta h^r PL\partial_t\partial_x(\Lambda\omega)h \\ &\leq h^r|P\partial_x\omega|^2_{m^2} + |\epsilon_x|^2_{m^\infty}\beta^2 h^r|P\partial_t\partial_x(\Lambda\omega)|^2_{m^2}.\end{aligned} \tag{92}$$

In the non-degenerate case, we are done. In the degenerate case we still have to consider $F-1$, $G-1$, and $H-1$.

8.3.19 F-1

We note that

$$\partial_x\bar{\bar{M}}_p = (\partial_x\bar{M}_p)P\Psi^-_{,c} + \bar{M}_{p,c}P\partial_x\Psi^-.$$

Hence

$$\begin{aligned}\sum_{\hat{\mathcal{P}}_x} &\partial_x(\bar{\bar{M}}_p P\omega)\sigma_0\omega_3 h = \\ &\sum_{\hat{\mathcal{P}}_x}((\partial_x\bar{M}_p)P\Psi^-_{,c} + \bar{M}_{p,c}P\partial_x\Psi^-)\omega_{,c}\sigma_0\omega_3 h + \sum_{\hat{\mathcal{P}}_x}\bar{\bar{M}}_p P\partial_x\omega\sigma_0\omega_3 h \\ &\leq C(|\omega|^2_{l^2} + |\bar{\bar{M}}_p|^2_{m^\infty}|P\partial_x\omega|^2_{m^2}).\end{aligned} \tag{93}$$

Note that $\bar{\bar{M}}_p$ is $h^{3/4}$ in m^∞, and $|\omega_3|_{l^2} = |\omega_3|_{m^2}$, so this is fine.

8.3.20 G-1

$$\sum_{\hat{\mathcal{P}}_x} M\omega\partial_x\bar{\bar{M}}_p\sigma_0\omega_3 h \leq C(|P\Psi^-|_{m^\infty} + |P\partial_x\Psi^-|_{m^\infty})|\omega|^2_{m^2}. \tag{94}$$

8.3.21 H-1

$$\sum_{\hat{\mathcal{P}}_x} \left(\bar{\bar{M}} \partial_x ((M + \bar{\bar{M}}) P \partial_t \zeta \triangle t) + M \partial_x ((h^2 M + \bar{\bar{M}}) P \partial_t \zeta \triangle t) \right) \sigma_0 \omega_3 h \leq$$
$$C|\omega|^2_{m^2} + |\Psi^-|_{m^\infty} \triangle t^2 |\partial_t \partial_x P \omega|^2_{m^2} + |\Psi^-|_{m^\infty} (1 + |P \partial_x \Psi^-|_{m^\infty}) \triangle t^2 |P \partial_t \zeta|^2_{m^2}$$
$$+ |\partial_x \Psi^-|_{m^\infty} \triangle t^2 |P \partial_t \zeta|^2_{m^2}. \tag{95}$$

We also write $|\partial_x \Psi^-|_{m^\infty} \triangle t \leq |\Psi^-|_{l^\infty}$.

8.4 The Induction Hypothesis

We are finally in a position to collect the 27 terms. Before we write it all out, we specify that σ_0 be taken small enough (based on order one constants) to make the $B-1$ and $A-3$ boundary terms non-negative, as described above, and we then take α to be a small fraction of σ_0, based on some other order one constants, and then take β small relative to α again based on order one constants, and finally require h and $\triangle t$ to be sufficiently small with respect to these other constants.

A number of right hand side terms now can be directly subtracted off from left hand side terms. These are terms with a small multiplier on them, usually written as $\frac{1}{64}$ above.

However, other terms which we would like to hide by subtraction involve factors such as $|\epsilon_t|_{m^\infty}$. We must now specify an induction hypothesis that makes these factors sufficiently small. For example, in the bound for term $D-3$ above there is a term

$$\frac{\alpha h^r}{64} |\bar{\bar{M}}|^2_{m^\infty} |\partial_t \omega|^2_{m^2}.$$

To be able to subtract this from the $A-1$ term

$$(\theta - \frac{1}{2}) \sigma_0 \triangle t |\partial_t \omega|^2_{m^2},$$

we need to assume that there are positive constants $\bar{C}$ and δ independent of h such that

$$|\Psi^-|_{m^\infty} \leq \bar{C} h^{\delta + (1-r)/2}.$$

In the non-degenerate case we can subtract it from the $A-2$ term instead, so we need only that

$$|\Psi^-|_{m^\infty} \leq \bar{C} h^\delta.$$

In either case, the inclusion of the $\delta > 0$ means that for h sufficiently small, the term can indeed be subtracted.

There are other right hand side terms which do not hide by subtraction. We wish to treat these by a Gronwall-like time induction. Once again we will need

induction hypotheses to keep the nonlinear terms sufficiently small. For example, in the bound for term $D-1$ above there is a term

$$C|\epsilon_z|_{m^\infty}|\omega|^2_{m^2}.$$

To make this be only a small correction to terms like

$$C|\omega|^2_{m^2},$$

which arise in $A-1$, we need to assume there are positive constants $\bar{C}$ and q independent of h such that

$$|\epsilon_z|_{m^\infty} \leq \bar{C}h^q.$$

It may happen that the best we can get is $q=0$; in this case the convergence estimate may only hold for a fixed finite time regardless of decreasing h.

In order to state the induction hypothesis we begin with a couple definitions. Let

$$K^n = |\omega^n|^2_{m^2} + \beta h^r |P\partial_x(\Lambda^n \omega^n)|^2_{m^2}, \tag{96}$$

and

$$J^n = h^r |P\partial_t \omega^{n-1+\theta}|^2_{m^2} + h|\partial_t \omega^{n-1+\theta}|^2_{m^2} + h^{r+1}|P\partial_t\partial_x\omega^{n-1+\theta}|^2_{m^2}. \tag{97}$$

Now let C_* be a constant independent of h, to be specified later.

Assumption 4 (The Induction Hypothesis) *We assume that at time $t^{n+\theta}$, there is a non-negative number $C_n \leq C_*$ such that $C_{n-1} \leq C_n$ and*

$$K^n + hJ^n \leq (C_n h)^2. \tag{98}$$

The induction hypothesis implies

$$|\zeta|_{m^2} + h^{r/2}|P\partial_x\zeta|_{m^2} \leq Ch, \tag{99}$$

where the C here is proportional to C_n. Standard approximation theory facts like $|\zeta|_{m^\infty} \leq Ch^{-\frac{1}{2}}|\zeta|_{m^2}$ and $|\zeta|_{L^\infty} \leq C(|\zeta|_{L^2}|\zeta|_{H^1})^{\frac{1}{2}}$ allow us to deduce the following additional bounds:

$$\begin{aligned}
|\zeta|_{m^\infty} &\leq Ch^{1/2}, && (100)\\
|P\zeta|_{m^\infty} &\leq Ch^{1-r/4}, && (101)\\
|\Psi|_{m^2} &\leq Ch, && (102)\\
|\Psi|_{m^\infty} &\leq Ch^{1/2}, && (103)\\
|P\Psi|_{m^\infty} &\leq Ch^{1-r/4}, && (104)
\end{aligned}$$

$$
\begin{aligned}
|P\partial_x\Psi|_{m^2} &\leq Ch^{1-r/2}, && (105)\\
|P\partial_x\Psi|_{m^\infty} &\leq Ch^{(1-r)/2}, && (106)\\
|\epsilon_t|_{m^2} &\leq Ch, && (107)\\
|\epsilon_t|_{m^\infty} &\leq Ch, && (108)\\
|\epsilon_x|_{m^2} &\leq Ch, && (109)\\
|\epsilon_x|_{m^\infty} &\leq Ch^{1/2}, && (110)\\
|\epsilon_z|_{m^2} &\leq Ch, && (111)\\
|\epsilon_z|_{m^\infty} &\leq Ch^{1/2}. && (112)
\end{aligned}
$$

In the non-degenerate case we get the improved bounds

$$|\epsilon_x|_{m^\infty} \leq Ch^{1-r/4},$$

and

$$|\epsilon_z|_{m^\infty} \leq Ch^{1-r/4}.$$

The induction hypothesis also implies

$$h^r|P\partial_t\Psi^-|^2_{m^2} + h|\partial_t\Psi^-|^2_{m^2} + h^{r+1}|P\partial_t\partial_x\Psi^-|^2_{m^2} \leq C^2h, \qquad (113)$$

from which we also get the bounds

$$
\begin{aligned}
|\partial_t\Psi|_{m^2} &\leq Ch^0, && (114)\\
|\partial_t\Psi|_{m^\infty} &\leq Ch^{-1/2}, && (115)\\
|P\partial_t\Psi|_{m^2} &\leq Ch^{(1-r)/2}, && (116)\\
|P\partial_t\Psi|_{m^\infty} &\leq Ch^{1/4-r/2}. && (117)
\end{aligned}
$$

In the non-degenerate case these improve to

$$
\begin{aligned}
|\partial_t\Psi|_{m^2} &\leq Ch^{(1-r)/2}, && (118)\\
|\partial_t\Psi|_{m^\infty} &\leq Ch^{1/4-r/2}. && (119)
\end{aligned}
$$

In all of these induction related bounds, the C is proportional to C_n. We point out that the above relations imply analogous bounds on the various $\bar{\bar{M}}$ constructs.

8.5 The Gronwall Induction

Let us now define q by

$$q = \begin{cases} 0 & \text{in the context of Theorem 1,} \\ 1/4 & \text{in the context of Theorem 2, or} \\ 1/2 & \text{in the context of Theorem 3.} \end{cases}$$

Returning to the error equation, we now multiply by $2\triangle t$, subtract the terms that can now be subtracted based on the induction hypothesis, and are left with the inequality

$$G_l \leq G_r,$$

where

$$\begin{aligned} G_l = \sigma_0(|\omega^+|^2_{m^2} - |\omega^-|^2_{m^2}) + (\theta - \frac{1}{2})\sigma_0 \triangle t^2 |\partial_t \omega|^2_{m^2} \\ + \alpha h^{r+1} |P\partial_t \omega|^2_{m^2} + \beta h^r \sigma_0 (|P\partial_x(\Lambda^+\omega^+)|^2_{m^2} - |P\partial_x(\Lambda^-\omega^-)|^2_{m^2}) \\ + \beta h^r (\theta - \frac{1}{2})\sigma_0 \triangle t^2 |P\partial_t\partial_x(\Lambda\omega)|^2_{m^2}. \end{aligned} \tag{120}$$

The reader can check that the above induction hypothesis allows us to write that for some constants C_I and C_J independent of h,

$$\begin{aligned} G_r = (C_J + C_I(C_n + C_n^2)h^q)\triangle t \cdot \\ \left(|\omega^-|^2_{m^2} + |\omega^+|^2_{m^2} + \beta h^r (|P\partial_x(\Lambda^-\omega^-)|^2_{m^2} + |P\partial_x(\Lambda^+\omega^+)|^2_{m^2}) + h^2\right) \\ + \beta h^r \sum_{\hat{\mathcal{P}}_x} \Big((Mh + \bar{M}\omega + \bar{\bar{M}}\omega)^{n+\theta} \cdot PL\partial_x(\Lambda^+\omega^+) \\ -(Mh + \bar{M}\omega + \bar{\bar{M}}\omega)^{n-1+\theta} \cdot PL\partial_x(\Lambda^-\omega^-)\Big) h. \end{aligned} \tag{121}$$

We must now investigate how C_n grows with n.

We define

$$H^{n+1} = \beta h^r \sum_{\hat{\mathcal{P}}_x} \Big((Mh + \bar{M}\omega + \bar{\bar{M}}\omega)^{n+\theta} \cdot PL\partial_x(\Lambda^+\omega^+)\Big) h.$$

Then we have

$$\begin{aligned} K^{n+1} + hJ^{n+1} - H^{n+1} \leq \\ K^n - H^n + (C_J + (C_n + C_n^2)C_I h^q)\triangle t (K^{n+1} + K^n + h^2). \end{aligned} \tag{122}$$

Summing in time many terms telescope, yielding

$$\begin{aligned} K^{n+1} - H^{n+1} + hJ^{n+1} \leq \\ K^0 - H^0 + (C_J + (C_n + C_n^2)C_I h^q) n h^3 + \sum_{k=0}^{n} 2K^k (C_J + 2(C_k + C_k^2)C_I h^q)\triangle t \\ + h(C_J + 2(C_n + C_n^2)C_I h^q)K^{n+1}. \end{aligned} \tag{123}$$

Now by the correct choice of initial conditions, namely $U^0 = W^0$, the induction hypothesis is satisfied at the initial time level with $C_0 = 0$. Similarly $K^0 = H^0 = 0$.

Also note that even when $q = 0$, for h sufficiently small $h(C_J + 2(C_n + C_n^2)C_I h^q) < 1/2$. Now,

$$|H^{n+1}| \leq \beta^{3/2} h^r |PL\partial_x(\Lambda^+ \omega^+)|_{m^2}^2 + \beta^{1/2} h^r (Ch^2 + C(1 + C_*^2 h)(|\omega^+|_{m^2}^2 + \triangle t^2 |\partial_t \omega|_{m^2}^2)),$$

Thus for β sufficiently small with respect $1/(C_*^2 h)$, $H^{n+1} < K^{n+1}/2 + h^{2+r}$. This is a harmless extra constraint on β since below we will be keeping $C_*^2 h^q$ bounded. Thus,

$$\begin{aligned} K^{n+1} + hJ^{n+1} \leq \\ 4(C_J + (C_n + C_n^2)C_I h^q)t^{n+1} h^2 + 2h^{2+r} \\ + 8 \sum_{k=0}^{n} K^k (C_J + 2(C_k + C_k^2)C_I h^q) \triangle t. \end{aligned} \tag{124}$$

So by the usual discrete Gronwall inequality,

$$\begin{aligned} K^{n+1} + hJ^{n+1} \leq \\ h^2 \cdot (2h^r + 4(C_J + (C_n + C_n^2)C_I h^q)t^{n+1}) \\ \cdot \exp(8(C_J + 2(C_n + C_n^2)C_I h^q)t^n). \end{aligned} \tag{125}$$

Hence we determine that at the next time level the induction hypothesis continues to hold, with

$$\begin{aligned} C_{n+1} \leq (2h^r + 4(C_J + (C_n + C_n^2)C_I h^q)t^{n+1})^{1/2} \\ \cdot \exp(4(C_J + 2(C_n + C_n^2)C_I h^q)t^n). \end{aligned} \tag{126}$$

However, this is only true as long as $C_{n+1} \leq C_*$. Thinking of C_n as defining a function $C(t)$ by n corresponding to t^n, we may ask "for what t does $C(t)$ first exceed the stated upper bound?" This time will be the $\bar{\tau}$ of the theorems.

Let $C_G = 4C_J + 2h^r + 8C_I(C_* + C_*^2)h^q$. To see how $C(t)$ grows, we note that $C_n < C_*$ implies

$$C(t) \leq (C_G t)^{1/2} \exp(C_G t).$$

Thus given any fixed values of C_* and h, $C(t)$ grows roughly exponentially with time, and will eventually exceed C_*. If $q > 0$, then we can increase C_* yet decrease h so as to leave C_G unchanged; this allows us to make $\bar{\tau}$ as large as desired.

Acknowledgment: I would like to thank my thesis advisor, Professor Todd Dupont, for suggesting this problem and for many hours of helpful advice, and my wife Mary Jo and my parents for all their love and support.

This work was supported in part by a National Science Foundation Graduate Fellowship.

References

[1] P. T. Keenan, *An error estimate for a new scheme for the general variable coefficient linearized thermal pipeline equations*, Tech. Report 90–20, Department of Computer Science, University of Chicago, 1990.

[2] M. Luskin, *An approximation procedure for nonsymmetric, nonlinear hyperbolic systems with integral boundary conditions*, SIAM J. Numer. Anal., 16 (1979), pp. 145–164.

[3] M. Luskin and T. Blake, *The existence of a global weak solution to the nonlinear waterhammer problem*, Comm. Pure Appl. Math, 35 (1982), pp. 697–735.

[4] V. Thomée, *A stable difference scheme for the mixed boundary problem for a hyperbolic, first order system in two dimensions*, J. Soc. Indust. Appl. Math., 10 (1962), pp. 229–245.

[5] E. B. Wylie and V. Streeter, *Fluid Transients*, McGraw Hill, New York, 1978.