Uniformly Convergent Finite Element Methods for Singularly Perturbed Parabolic Partial Differential Equations

A thesis submitted for the degree of Doctor of Philosophy

at the National University of Ireland

Wen Guo

Supervisor: Dr Martin Stynes

Department of Mathematics

University College

Cork

April 1993

Summary

Uniformly Convergent Finite Element Methods for Singularly Perturbed Parabolic Partial Differential Equations

Wen Guo

Department Of Mathematics

University College, Cork, Ireland

A thesis submitted for the degree of Doctor of Philosophy

April 1993

This thesis is concerned with uniformly convergent finite element methods for numerically solving singularly perturbed parabolic partial differential equations in one space variable.

First, we use Petrov-Galerkin finite element methods to generate three schemes for such problems, each of these schemes uses exponentially fitted elements in space. Two of them are lumped and the other is non-lumped. On meshes which are either arbitrary or slightly restricted, we derive global energy norm and L^2 norm error bounds, uniformly in the diffusion parameter. Under some reasonable global assumptions together with realistic local assumptions on the solution and its derivatives, we prove that these exponentially fitted schemes are locally uniformly convergent,

with order one, in a discrete L^{∞} norm both outside and inside the boundary layer.

We next analyse a streamline diffusion scheme on a Shishkin mesh for a model singularly perturbed parabolic partial differential equation. The method with piecewise linear space-time elements is shown, under reasonable assumptions on the solution, to be convergent, independently of the diffusion parameter, with a pointwise accuracy of almost order 5/4 outside layers and almost order 3/4 inside the boundary layer.

Numerical results for the above schemes are presented.

Finally, we examine a cell vertex finite volume method which is applied to a model time-dependent convection-diffusion problem. Local errors away from all layers are obtained in the l_2 seminorm by using techniques from finite element analysis.

Acknowledgements

I would like to take this opportunity to thank the people who have helped me write this thesis.

First of all I am sincerely grateful to my supervisor Dr Martin Stynes. He has introduced me to this interesting and challenging topic and helped me enormously with numerous discussions and practical advice. He always consoles me when I encounter difficulties and encourages me to overcome them. He also helps a lot in improving my English. Any words of thanks could only be inadequate. I also wish to express my appreciation to Mrs Stynes for her concern and help.

I must also thank Dr Eugene O'Riordan, Dr Grigorii Shishkin and Professor Hans-Görg Roos for short but invaluable discussions and Dr Alan Hegarty for his advice on some of experimental aspects of the thesis.

Thanks must also go to my colleague Guangfu Sun for helpful discussions on different aspects of the work.

In addition I would like to thank Professor Patrick Barry and Professor Finbarr Holland for their concern and support and Dr Thomas Carroll for his help in solving problems in Latexing the thesis.

I must thank my husband who urged me to pursue a Ph.D. in Ireland and has taken care of our little daughter since I was away. I owe a great deal to my parents and my sisters. They have given me great support to carry out my Ph.D. study and helped tremendously in looking after my daughter. Without their encouragement and help I would not be writing this today.

Finally, I gratefully acknowledge the Department of Mathematics at University

College Cork for the financial support and the facilities made available to me while I was writing the thesis.

Contents

Chapter 1

Introduction

1.1 Statement of the Problem

We consider the singularly perturbed parabolic problem

$$Lu(x,t) \equiv -\varepsilon u_{xx} + a(x,t)u_x + b(x,t)u + r(x,t)u_t = f(x,t) \quad \forall (x,t) \in \Omega, \quad (1.1.1)$$

where $\Omega = (0,1) \times (0,T]$, T is a positive constant,

$$u(0,t) = q_0(t) \text{ and } u(1,t) = q_1(t) \quad \text{for} \quad 0 < t \leq T, \quad (1.1.2)$$

$$u(x,0) = u^0(x) \quad \text{for} \quad 0 \leq x \leq 1, \quad (1.1.3)$$

ε is a small positive parameter, and the functions a, b, r and f are sufficiently smooth with

$$0 < \alpha \leq a(x,t) \leq \alpha^*, \quad 0 < \nu \leq r(x,t) \leq \nu^*, \quad \forall (x,t) \in \bar{\Omega}. \quad (1.1.4)$$

We also assume that q_0, q_1 and u^0 are piecewise smooth.

Under these hypotheses, we may assume without loss of generality (by making a change of dependent variable if necessary) that

$$b(x,t) - \frac{1}{2}a_x(x,t) \geq 2C_1 \quad \text{on } \bar{\Omega}, \quad (1.1.5)$$

where $C_1 \in (0,1]$ is a positive constant independent of ε and any mesh used.

The conditions (1.1.1) – (1.1.5) define a time-dependent convection-diffusion problem. Problems of this type arise, for example, in the modelling of steady and unsteady viscous flow problems with large Reynolds numbers (see Peaceman and Rachford [37] and Van Dyke [50]), convective heat transport problems with large Peclet numbers (see Jacob [18]), oil reservoir simulation (see Ewing [11]), radioactive corrosion in the water cycles of an atomic reactor, adsorption processes in gas piplines, spread of medicaments with the blood circulation or of plumes of poisonous industrial wastes in river systems (see Baumert *et al.* [3]), petroleum reservoir mechanics (see Price and Varga [38]) and electromagnetic field problems in moving media (see Hahn [42]). In (1.1.1) ε is a diffusion coefficient and the function a is a flow rate.

The differential operator in (1.1.1) is of mixed parabolic-hyperbolic type and has mainly hyperbolic nature when ε is small when compared to a, b, r, and f. In the limit case $\varepsilon = 0$, (1.1.1) degenerates into a purely hyperbolic equation where the initial-boundary condition is restricted only to the inflow sides $t = 0$ and $x = 0$. The solution of this hyperbolic problem, which is called the reduced solution, can be obtained by integrating along the characteristics starting on the inflow boundary (see Bobisud [4]). It is known that when $\varepsilon \to 0$, the solution $u(x,t)$ of the full problem (1.1.1) – (1.1.5) converges weakly in $L^2(\Omega)$ towards the reduced solution (see Vishik and Lyusternik [51]). However, the reduced solution is in general not identical with $u(x,t)$ at the outflow boundary $x = 1$. Hence the solution $u(x,t)$ will generally vary rapidly in a layer region of width $O(\varepsilon \ln(1/\varepsilon))$ at boundary $x = 1$, even for smooth initial-boundary data. This layer region is called a boundary layer. The boundary layer phenomenon has been discussed by many authors since Prandtl's original work in 1905; see, e.g., Vishik and Lyusternik [51], Eckhaus and de Jager [10], Nayfeh [30] and O'Malley [33].

In addition to having a boundary layer along $x = 1$, the solution $u(x,t)$ also shows internal layers of width $O(\sqrt{\varepsilon})$ (see Nävert [29] and Eckhaus and de Jager [10]) along those characteristics at which the reduced solution is discontinuous. Such discontinuities typically occur if the inflow boundary data have a jump discontinuity on the inflow boundary $x = 0$ or $t = 0$, or if f has a jump discontinuity across a characteristic.

Due to the presence of these layers, the solution $u(x,t)$ will in general not be globally smooth; it will vary rapidly in layer regions. This causes serious difficulties when solving (1.1.1) – (1.1.5) numerically. In the next section, we will review some numerical methods proposed for this singularly perturbed parabolic problem.

1.2 Previous Numerical Analyses

For parabolic partial differential equations, conventional numerical schemes such as the finite element Galerkin method or finite difference methods typically yield centered difference approximations for the convective term. For small values of ε, such methods will produce severely oscillating solutions, which do not accurately approximate the exact solution of (1.1.1) – (1.1.5) unless an unacceptable large number of mesh points is used or the exact solution happens to be globally smooth. Indeed, the inadequacy of the conventional methods for the singularly perturbed parabolic problem is a well-documented fact (cf. for example, Hindmarsh *et al.* [15]). In the finite element approach, this deficiency is usually remedied by replacing the

Galerkin method by a so-called Petrov-Galerkin method, in which the test functions may be selected differently from the trial functions. The key problem then is how to choose test and trial functions. We give here a brief survey of the extensive literature on this topic.

For a singularly perturbed ordinary differential equation, Barrett and Morton [2] constructed a set of special test functions from a set of trial functions by approximately symmetrizing the bilinear form and obtained an almost optimal approximation in a special norm. This symmetrization method based on a symmetrized bilinear form has been applied to singularly perturbed parabolic equations with mixed boundary conditions and periodic boundary conditions by Wu [52]. Optimal estimates in the H^1 norm are derived. However, the analysis does not apply to parabolic problems with Dirichlet boundary conditions due to a lack of coercivity.

The streamline diffusion method of Hughes and Brooks [16] was initially introduced in the case of steady convection-diffusion problems. In this method, the test functions w are constructed from the trial functions v by taking $w = v + \delta v_\beta$, where δ is a small positive parameter of order h, h being the mesh diameter and v_β the derivative in the streamline direction (i.e., the direction of propagation of the reduced problem). Johnson *et al.* [20] extended the method to time dependent convection-diffusion problems. Nävert [29] proved that when piecewise polynomial finite elements of degree k are used, the method is of order $k + 1/2$ in the L^2 norm in smooth regions (i.e., away from any layers).

The Eulerian-Lagrangian localized adjoint method of Celia, Ewing, Herrera and Russell [5, 40] is a finite element method specially proposed for time dependent convection-diffusion problems. It combines finite element techniques with the method of characteristics by using test functions satisfying a local adjoint condition that introduces a Lagrangrian frame of reference.

The previous pair of methods were constructed to reflect the almost hyperbolic nature of the problem (1.1.1) – (1.1.5). Hence they perform well provided that one is far from the boundary layer near $x = 1$. However, they are not accurate inside layers. All global error bounds obtained for these methods involve Sobolev norms of u. Since such norms generally involve negative powers of the parameter ε, the bounds are in general of large magnitude and do not provide evidence of convergence of the methods.

It is desirable to have numerical methods whose accuracy inside the boundary layer is retained irrespective of the value of ε. In the singular perturbation literature, numerical methods with this property are said to be uniformly convergent.

Most uniformly convergent methods have been obtained for singularly perturbed ordinary differential equations (see, e.g., Doolan, Miller and Schilders [8], O'Riordan and Stynes [34], Gartland [13] and Liseikin and Petrenko [24]).

Uniformly convergent methods for problem (1.1.1) – (1.1.5) have been examined by Duffy [9], Han and Kellogg [14] and Ng-Stynes *et al.* [31] on uniform meshes, by Stynes and O'Riordan [45] on an arbitrary mesh and by Shishkin [41] on a special nonuniform mesh.

Duffy [9] gave an algorithm which consists of Allen and Southwell [1]/Il'in [17] differencing in the x-direction and forward differencing in time. For this scheme he claimed a uniform convergence result of order one in the discrete L^∞ norm. However his argument relies on an unjustified differentiation of an asymptotic expansion of the solution u.

Ng-Stynes *et al.* [31] and Stynes and O'Riordan [45] presented a family of finite difference schemes which are generated from Petrov-Galerkin finite element methods with exponential test functions. Under certain compatibility conditions on the data of (1.1.1) – (1.1.5), they showed that the scheme is uniformly convergent with order one in the discrete L^∞ norm. However, their compatibility assumptions are very strong and in practice unlikely to be satisfied.

Han and Kellogg [14] considered a semi-discrete finite element method for (1.1.1) – (1.1.5) while assuming that all functions in (1.1.1) depend on x only. They used an enriched finite element space consisting of piecewise linear functions plus an extra function which is chosen to model the behaviour of the solution in the boundary layer. A uniform error bound of order 3/4 in an energy norm was obtained. However their argument is not applicable when the functions in (1.1.1) depend also on the variable t.

Shishkin [41] constructed a special piecewise uniform mesh, which is fine in part of the boundary layer. When it is used with a upwinded finite difference scheme, an error estimate of order one in the discrete L^∞ norm is obtained.

In most published research, analyses of theoretical uniform convergence are essentially carried out in a consistency/stability framework associated with finite difference methods, which require the scheme considered to satisfy a discrete maximum principle and also need strong global assumptions on the solution and its derivatives.

In contrast to the finite difference situation, there are only a few results which use finite element arguments to yield error bounds which are uniform in ε for approximate solutions to (1.1.1) – (1.1.5). Recently, Stynes and O'Riordan [47] presented a framework for the finite element analysis of a singularly perturbed two-point boundary value problem which yields uniform global L^2 and energy norm bounds. They also successfully applied the method to singularly perturbed elliptic problems in two dimensions [36]. However, they do not obtain pointwise error bounds.

Johnson *et al.* [22] and Niijima [32] used purely finite element analyses to derive localized pointwise error bounds outside any layers. This type of analysis requires only reasonable local and global assumptions on the behaviour of the solution u and its derivatives. When we carried out the research for this thesis, as far as we knew, no counterpart of the results of [22] and [32] existed for singularly perturbed parabolic problems. (We later became aware of the work of Zhou [53], which overlaps slightly with our Chapter 5. See Remark 5.6.2.) Furthermore, no finite element analysis of pointwise convergence inside the boundary layer exists in the literature.

1.3 Outline of the Thesis

The main aim of this work is to propose suitable finite element methods for solving (1.1.1) – (1.1.5) numerically and to analyse uniform convergence of these methods in global L^2 and energy norms and the local L^∞ norm using purely finite element techniques. We will analyse several uniformly convergent methods: exponentially fitted schemes on an arbitrary mesh and the streamline diffusion scheme on a Shishkin mesh. Results are obtained not only outside but also inside the boundary layer. The analyses are carried out under reasonable assumptions on the solution and its derivatives. We will also analyse a cell vertex finite volume method applied to (1.1.1) – (1.1.5), using finite element techniques; in this way we obtain local and global convergence results.

The outline of the thesis is as follows. In Chapter 2 we give a brief description of a combination of Petrov-Galerkin and finite difference methods which will be used in Chapters 3 and 4. We also derive necessary conditions for uniform convergence (in the discrete L^∞ norm) of a scheme on a uniform mesh, which will motivate the choices of trial and test functions in Chapters 3 and 4.

In Chapter 3, two exponentially fitted lumped schemes are presented, using various choices of trial and test functions. The inner product involving the time derivative term in a weak form of (1.1.1) – (1.1.5) is approximated by two suitable discrete inner products. Global error bounds in the discrete L^2 and energy norms are derived uniformly in ε. The error analyses also show that the two schemes are both first order accurate, uniformly in ε, at all nodes.

Chapter 4 examines a non-lumped exponentially fitted scheme. Unlike the schemes in Chapter 3, this scheme integrates the inner product involving the time derivative term exactly. This small modification leads to a rather different scheme which needs a certain stability condition to guarantee uniform convergence. Uniform convergence results similar to those of Chapter 3 are obtained.

We note that the schemes considered in Chapters 3 and 4 are similar to the ones studied in Ng-Stynes *et al.* [31] and Stynes and O'Riordan [45], but the finite element analysis presented here is valid under weaker hypotheses than those required for the finite difference analysis of [45]. In these two Chapters, in order to derive pointwise error bounds we assume only that the solution of (1.1.1) – (1.1.5) and its first order derivatives are uniformly bounded in a variant of the global L^1 norm (cf. (3.2.68)) and that locally the solution is either smooth or exhibits typical boundary layer behaviour (cf. (3.2.74)).

In Chapter 5 we combine the streamline diffusion method with a Shishkin mesh.

Our method uses space-time finite elements. The global assumptions we make in this Chapter are that the right hand side f of (1.1.1) and the initial data u^0 are bounded in the L^2 norm, uniformly in ε, and the streamline derivative of the solution u and εu_{ss} are both bounded in the L^1 norm, uniformly in ε. Under these reasonable global assumptions, we prove that the pointwise error bound is of order almost 5/4 in smooth regions and almost 3/4 inside a typical boundary layer, uniformly in ε. This improves the results of Nävert [29], who did not obtain convergence inside the boundary layer and who analysed L^2 rather that L^∞ local convergence.

In Chapter 6 we introduce a cell vertex finite volume method for (1.1.1) – (1.1.5) with constant coefficients. The method has been widely used in the arospace industry. However, analysis of this method has lagged far behind the application of the method. Up to now, no fully satisfactory analysis of the cell vertex finite volume method has been published. The best estimates available are in Morton and Stynes [27], where a sharp convergence result for a two-point boundary value singularly perturbed problem is obtained in a weighted discrete Sobolev H^1 norm. There is no previous convergence result for this method applied to a singularly perturbed parabolic problem. Here we derive a local l_2 error estimate for the model parabolic problem by using finite element techniques. Under the assumption that the right hand side f of (1.1.1) and the initial data u^0 are bounded in the L^2 norm, uniformly in ε, we show that the method is locally first order accurate on a general tensor product mesh. This result can be sharpened to second order accurate, if either ε is very small compared to the mesh diameter or the mesh is locally almost uniform.

Numerical results are given in the last section of Chapters 1 – 5.

Chapter 2

Discretizing the Problem

2.1 Petrov-Galerkin Finite Elements in Space and Finite Differences in Time

An equivalent formulation of (1.1.1) – (1.1.5) is got by replacing (1.1.1) for each t by

$$B(u, v) + (ru_t, v) = (f, v) \quad \forall v \in H_0^1(0,1). \tag{2.1.1}$$

Here $(\cdot,\cdot)$ is the usual $L^2(0,1)$ inner product, $H_0^1(0,1)$ is the usual Hilbert space given by

$$H_0^1(0,1) = \{v = v(x) : \ \|v\| + \|v_x\| < \infty, \ v(0) = v(1) = 0\},$$

where $\|\cdot\|$ is the norm in $L^2(0,1)$, and we define

$$B(w, z) \equiv \varepsilon(w_x, z_x) + (aw_x, z) + (bw, z). \tag{2.1.2}$$

The weak form (2.1.1) can be discretized by means of a Petrov-Galerkin finite element method with space elements. This yields a semidiscrete problem, which corresponds a system of first order differential equations with an initial condition. Then differencing in time gives a fully discrete problem.

To discretize the problem, we introduce an arbitrary tensor product grid on $\bar{\Omega}$. Let M, N be positive integers. In the x-direction, let

$$0 = x_0 < x_1 < \ldots < x_N = 1 \tag{2.1.3}$$

with $h_i = x_i - x_{i-1}$ for $i = 1, \ldots, N$ and set $H = \max_i h_i$. In the t-direction, let

$$0 = t_0 < t_1 < \ldots < t_M = T \tag{2.1.4}$$

with $k_j = t_j - t_{j-1}$ for $j = 1, \ldots, M$. Set $k = \min_j k_j$ and $K = \max_j k_j$.

Then for each $j \in \{1, \ldots, M\}$, we wish to define trial and test functions on $[0,1] \times \{t_j\}$. The question now is how to choose these trial and test functions so that the resulting scheme is convergent uniformly in ε. To show that standard piecewise polynomial elements are inadequate for this purpose, we shall in Section 2.2 derive necessary conditions for uniform convergence (in the discrete L^∞ norm) of a scheme on a uniform mesh. These conditions imply that the coefficients of the scheme must possess a certain exponential nature.

2.2 Necessary Conditions for Uniform Convergence of a Scheme on a Uniform Mesh

We consider (1.1.1) – (1.1.5) with constant coefficients and zero initial-boundary data, viz.,

$$\left.\begin{array}{ll} Lu(x,t) = -\varepsilon u_{xx} + a u_x + bu + r u_t = f(x,t) & \text{in } \Omega, \\ u(0,t) = u(1,t) = 0, \quad 0 < t \leq T, & \\ u(x,0) = 0, \quad 0 \leq x \leq 1. & \end{array}\right\} \tag{2.2.1}$$

Assume that we have a uniform square mesh of diameter H. Applying a Petrov-Galerkin method in the space variable and finite differencing in the time variable typically leads to a difference scheme of the form

$$\sum_{m=0}^{1} \sum_{n=-1}^{1} \alpha_{n,m} U(x_{i+n}, t_{j+m}) = H \bar{f}_{i,j}, \tag{2.2.2}$$

for $i = 1, \ldots, N-1$ and $j = 1, \ldots, M-1$, where $U(x_i, t_j)$ is our computed solution at the point (x_i, t_j) and $\bar{f}_{i,j}$ is an approximation to $f(x_i, t_j)$. Schemes which involve more than three points in the x-direction and/or more than two points in the t-direction can be treated similarly.

Assume that for all i and j,

$$|U(x_i, t_j) - u(x_i, t_j)| \leq CH^s \tag{2.2.3}$$

for some $s > 0$, where C and s are constants independent of ε and of the mesh parameter H. We will derive necessary conditions on the coefficients $\{\alpha_{n,m}\}$ of the scheme (2.2.2).

Necessary conditions for various singularly perturbed problems have been previously examined by Doolan *et al.* [8] and Roos [39]. A result similar to that presented below was claimed by O'Riordan and Stynes [35], but their proof omits any mention of compatibility assumptions on the data, which are needed for their argument. Our proof uses a uniformly valid asymptotic expansion of Bobisud [4], then finishes using the same argument as in O'Riordan and Stynes [35].

First, from the proof of Theorem 2 in Bobisud [4], we have

Lemma 2.2.1 *Assume that $f \in C^2(\bar{\Omega})$. Let $u_0(x,t)$ be the solution of the reduced problem*

$$a(u_0)_x + bu_0 + r(u_0)_t = f(x,t), \quad \forall (x,t) \in \Omega,$$

$$u_0(0,t) = 0, \quad 0 < t \leq T,$$

$$u_0(x,0) = 0, \quad 0 \le x \le 1.$$

Then for the solution $u(x,t)$ of (2.2.1), we have

$$u(x,t) = u_0(x,t) - u_0(1,t)\exp(-a(1-x)/\varepsilon) + z(x,t), \quad \forall (x,t) \in \Omega, \tag{2.2.4}$$

where

$$|z(x,t)| \le C\sqrt{\varepsilon}, \quad \forall (x,t) \in \Omega. \tag{2.2.5}$$

Next, set $\rho = H/\varepsilon$. We will assume that the coefficients $\{\alpha_{n,m}\}$ in (2.2.2) depend only on n, m and ρ. (This assumption holds for all schemes of which we are aware.) Rewrite (2.2.2) in the form

$$\sum_{m=0}^{1}\sum_{n=-1}^{1} \alpha_{n,m} U(1 - x_{i+n}, T - t_{j+m}) = H \bar{f}_{N-i,M-j}.$$

Fix ρ, i and j. Letting $H \to 0$ yields

$$\begin{aligned} 0 &= \sum_{m=0}^{1}\sum_{n=-1}^{1} \alpha_{n,m} \lim_{H\to 0} U(1 - x_{i+n}, T - t_{j+m}) \\ &= \sum_{m=0}^{1}\sum_{n=-1}^{1} \alpha_{n,m} \lim_{H\to 0} u(1 - x_{i+n}, T - t_{j+m}), \quad \text{using (2.2.3)}, \\ &= \sum_{m=0}^{1}\sum_{n=-1}^{1} \alpha_{n,m} \{u_0(1,T) - u_0(1,T)\exp(-a(i+n)\rho)\}, \end{aligned}$$

using (2.2.4) and the fact that for fixed ρ, $H \to 0$ implies that $\varepsilon \to 0$, so that by (2.2.5), $z \to 0$.

In general $u_0(1,T) \ne 0$, so

$$\sum_{m=0}^{1}\sum_{n=-1}^{1} \alpha_{n,m}\{1 - \exp(-a(i+n)\rho)\} = 0. \tag{2.2.6}$$

This holds for any fixed $i \in \{1, \ldots, N-1\}$. Taking $i = 1, 2$,

$$\sum_{m=0}^{1}\sum_{n=-1}^{1} \alpha_{n,m}\{1 - \exp(-a(1+n)\rho)\} = 0, \tag{2.2.7}$$

$$\sum_{m=0}^{1}\sum_{n=-1}^{1}\alpha_{n,m}\{1-\exp(-a(2+n)\rho)\}=0. \tag{2.2.8}$$

Multiply (2.2.7) by $\exp(-a\rho)$ and then subtract the resulting equation from (2.2.8) to get

$$\sum_{m,n}\alpha_{n,m}=0. \tag{$C1$}$$

Put this into (2.2.6) to obtain

$$\sum_{m,n}\alpha_{n,m}\exp(-a(i+n)\rho)=0,$$

whence

$$\sum_{m,n}\alpha_{n,m}\exp(-an\rho)=0. \tag{$C2$}$$

Conditions (C1) and (C2) are the necessary conditions for uniform in ε convergence. We note that polynomial based schemes in general satisfy (C1), but they cannot satisfy (C2).

In Chapters 3 and 4 we will present some schemes generated by Petrov-Galerkin methods with suitable spacial trial and test functions and backward differencing in the time variable. These schemes have coefficients based on exponentials and satisfy (C1) and (C2).

2.3 Notation

We now introduce some notation which will be used in Chapters 3 and 4.

For any function $v(x,t)$ and $m \in \{1,\ldots,M\}$,

$$v^m(\cdot)=v(\cdot,t_m),$$

$$\partial v^m(\cdot)=(v^m(\cdot)-v^{m-1}(\cdot))/k_m.$$

For any finite element space $D \subset C([0,1] \times \{t_m\})$, where m is fixed, $(v)_D$ denotes the interpolant from D to v at $\{(x_i, t_m)\}_{i=0}^N$. Set

$$[0,1]^\wedge \equiv [0,1] \backslash \{x_0, \ldots, x_N\}.$$

We use $(\cdot,\cdot)$ to denote the usual $L^2(0,1)$ inner product, and $(\cdot,\cdot)^\wedge$ denotes that the integration is only over $[0,1]^\wedge$.

Define

$$\Omega_K = \{(x, t_m) \in \Omega : 0 < x < 1, 1 \leq m \leq M\},$$

$$\Omega_H = \{(x_i, t) \in \Omega : 1 \leq i \leq N-1, 0 < t \leq T\}.$$

Note that Ω_K is a distance k_1 from the boundary $t = 0$ of Ω. We will use the following mesh-dependent norms for $v \in C(\Omega)$:

$$\|v\|_{L^1(L^1(\Omega_K))} = \sum_{m=1}^{M} k_m \|v^m\|_{L^1(0,1)}, \tag{2.3.1}$$

$$\|v\|_{L^1(L^\infty(\Omega_K))} = \sum_{m=1}^{M} k_m \|v^m\|_{L^\infty(0,1)}, \tag{2.3.2}$$

$$\|v\|_{L^2(L^1(\Omega_K))} = \left\{ \sum_{m=1}^{M} k_m \|v^m\|^2_{L^1(0,1)} \right\}^{1/2}, \tag{2.3.3}$$

$$\|v\|_{L^2(L^\infty(\Omega_K))} = \left\{ \sum_{m=1}^{M} k_m \|v^m\|^2_{L^\infty(0,1)} \right\}^{1/2}, \tag{2.3.4}$$

$$\|v\|^*_{L^2(L^1(\Omega_H))} = \left\{ \sum_{i=1}^{N-1} (h_i + h_{i+1}) \left(\int_0^T |v(x_i, t)|\, dt \right)^2 \right\}^{1/2}, \tag{2.3.5}$$

where $\|\cdot\|_{L^1(0,1)}$, $\|\cdot\|_{L^2(0,1)}$ and $\|\cdot\|_{L^\infty(0,1)}$ are the usual L^1, L^2 and L^∞ norms.

When the norms are over all of Ω independently of the mesh, we omit Ω from the notation; define

$$\|v\|_{L^2(L^1)} = \left\{ \int_0^T \left(\int_0^1 |v(x,t)|\,dx \right)^2 dt \right\}^{1/2}, \tag{2.3.6}$$

$$\|v\|^*_{L^2(L^1)} = \left\{ \int_0^1 \left(\int_0^T |v(x,t)|\,dt \right)^2 dx \right\}^{1/2}. \tag{2.3.7}$$

$$\|v\|_{L^2(L^1),m} = \left\{ \int_{t_{m-1}}^{t_m} \left(\int_0^1 |v(x,t)|\,dx \right)^2 dt \right\}^{1/2}, \tag{2.3.8}$$

Throughout the paper, C is a generic positive constant which is independent of ε and of any mesh used. When we say a quantity y is $O(H)$, it means that $|y| \leq CH$.

The analysis will frequently use the arithmetic-geometric mean inequality

$$yz \leq \beta y^2 + z^2/(4\beta) \quad \forall \beta > 0,\ \forall y, z \in \mathcal{R}. \tag{2.3.9}$$

Chapter 3

Exponentially Fitted Lumped Schemes

3.1 Introduction

In this chapter we will consider two schemes generated by Petrov-Galerkin finite element methods with various choices of trial and test functions. To approximate the time derivative term (ru_t, v) in the weak form (2.1.1), we first use a discrete L^2 inner product $(ru_t, v)_d$ which will be defined in terms of the test functions used. We then apply backward differencing in the time variable. Schemes generated by using an approximation of this type to the time derivative are called lumped schemes.

We will define two combinations of trial functions and test functions. They are

(i) any trial functions and $\bar{L}^*$-spline test functions (Section 3.2);

(ii) $\tilde{L}$-Spline trial functions and piecewise linear test functions (Section 3.3).

Each trial or test function is defined on $[0,1] \times \{t_m\}$ for some m.

The motivation for studying the two schemes is to explore the relationship between the choice of trial and test spaces and the norm in which one can prove a global convergence result which is uniform in ε. When $\bar{L}^*$-spline test functions are combined with any reasonable choice of trial functions, we obtain convergence results in the discrete L^2 norm. Here $\bar{L}^*$ is an approximation of L_s^*, the adjoint of the spatial part L_s of the differential operator L. (A related idea is used by Celia *et al.* [5], whose test functions are approximate solutions of $L^*z = 0$.) If we then choose the trial functions to be $\tilde{L}$-splines, where $\tilde{L}$ is an approximation of L_s, we can also prove convergence in an energy norm. These general results can be sharpened if one makes compatibility assumptions on the data of the problem. If, on the other hand, $\tilde{L}$-spline trial functions and piecewise linear test functions are used, we obtain corresponding results in this situation. Our analysis shows that the scheme with $\bar{L}^*$-spline test functions is better for discrete L^∞ and L^2 convergence, and the other scheme, which uses $\tilde{L}$-spline trial functions, is more suitable for energy norm convergence.

Both schemes have coefficients based on exponentials; it was shown in Chapter 2 that a scheme on a uniform mesh must possess this property if it is to be globally L^∞ convergent, uniformly in ε. The results mentioned above are all global in nature.

When internal layers are also present in the solution, it is important to also have local convergence results for a numerical method. We provide such a local convergence analysis for both of our schemes. We prove that, under reasonable assumptions on the behaviour of u and its derivatives, our schemes are locally pointwise convergent *inside* the boundary layer. To the best of our knowledge, no result of this type exists in the literature on (1.1.1) – (1.1.5). When the cell Reynolds number is large, the schemes are similar to upwinding, which is often described as being "formally first order accurate", but has not previously been analysed in the literature as regards its local behaviour. We note that the schemes considered in this paper are similar to the ones studied in Ng-Stynes *et al.* [31] and Stynes and O'Riordan [45] but the finite element analysis presented here is valid under weaker hypotheses than those required for the finite difference analysis of [31, 45], and furthermore no local convergence results are proven in these two papers.

The structure of this Chapter is as follows. In Sections 3.2 and 3.3 we examine the two respective schemes. We will derive global estimates in energy and L^2 norms using finite element techniques similar to those of Stynes and O'Riordan [47]. We also use a discrete Green's function, based on Niijima's analysis of an elliptic problem (see Niijima [32]), to derive local pointwise error estimates inside and outside the boundary layer for the schemes. (This entails a novel definition of a discrete Green's function for our parabolic problem.) In particular we emphasize that many of our error bounds are uniform in ε. Numerical results are presented in Section 3.4.

3.2 $\bar{L}^*$-Spline Test Functions

3.2.1 Description of Scheme and Norms

In this section we shall consider a Petrov-Galerkin finite element method with $\bar{L}^*$-spline test functions. A semidiscrete approximation with a similar idea was considered in Stynes and Guo [44] where the trial functions were chosen as $\tilde{L}$-splines. We do not need at this stage to precisely specify the trial functions $\{\phi_{i,m}(x) : i = 0, \ldots, N \text{ and } m = 0, \ldots, M\}$. We only assume that for $m \in \{0, \ldots, M\}$, each $\phi_{i,m}(x)$ is defined and satisfies the following properties:

$$\text{(i) } \phi_{i,m} \text{ is continuous on } [0,1] \times \{t_m\} \text{ and differentiable on } [0,1]^{\wedge} \times \{t_m\}; \tag{3.2.1}$$

$$(\text{ii})\ \phi_{i,m}(x_j) = \delta_{i,j}, \text{ for } j = 0, 1, \ldots, N, \tag{3.2.2}$$

where

$$\delta_{i,j} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases}$$

Then it turns out that the scheme generated is completely determined, due to the choice of the test functions.

We assume that the mesh is arbitrarily graded in the x-direction, i.e., $h_i \leq h_{i-1}$ for each i. This is not a practical restriction, since the boundary layer is at $x = 1$. To define the $\bar{L}^*$-spline functions, we first introduce two approximations of the function $a(x, t)$ on $t = t_m$ for each m:

$$\begin{aligned} \tilde{a}_i^m &= \tilde{a}^m(x)\,|_{(x_{i-1}, x_i]} \\ &= (a(x_{i-1}, t_m) + a(x_i, t_m))/2, \end{aligned} \tag{3.2.3}$$

and

$$\begin{aligned} \bar{a}_i^m &= \bar{a}^m(x)\,|_{(x_{i-1}, x_i]} \\ &= g(\tilde{\rho}_i^m)a(x_{i-1}, t_m) + (1 - g(\tilde{\rho}_i^m))a(x_i, t_m), \end{aligned} \tag{3.2.4}$$

for $i = 1, \ldots, N$, where

$$g(z) = (1 - \exp(-z))^{-1} - z^{-1} \quad \text{for all } z > 0,$$

and $\tilde{\rho}_i^m = \tilde{a}_i^m h_i/\varepsilon$. Now the $\bar{L}^*$-spline functions $\{\bar{\psi}_{i,m} : i = 1, \ldots, N-1; m = 0, \ldots, M\}$ are defined as:

$$\bar{L}^*\bar{\psi}_{i,m} \equiv -\varepsilon\bar{\psi}_{i,m}''(x) - \bar{a}_i^m\bar{\psi}_{i,m}'(x) = 0 \quad \text{for } x \in [0,1]^{\hat{}}, \tag{3.2.5}$$

$$\bar{\psi}_{i,m}(x_j) = \delta_{i,j}, \qquad \text{for } j = 0, \ldots, N. \tag{3.2.6}$$

A maximum principle argument yields $\bar{\psi}_{i,m}(x,t_m) \geq (x - x_{i-1})/h_i$ for $x_{i-1} \leq x \leq x_i$. Hence

$$(1, \bar{\psi}_{i,m}) \geq h_i/2 \quad \text{for all } i \text{ and } m. \tag{3.2.7}$$

To discretize (2.1.1) we define a discrete inner product for each t_m:

$$(v,w)_{d^*} = \sum_{i=1}^{N-1} (1, \bar{\psi}_{i,m}) v(x_i, t_m) w(x_i, t_m) \quad \forall v, w \in C([0,1] \times \{t_m\}). \tag{3.2.8}$$

For each m, let $\bar{V}_m$ and S_m be the linear spans of $\{\bar{\psi}_{i,m}(x) : i = 1, \ldots, N-1\}$ and $\{\phi_{i,m}(x) : i = 0, \ldots, N\}$ respectively. Then our first lumped Petrov-Galerkin approximation can be formulated as follows: for each $m \in \{1, \ldots, M\}$, find $U^m \in S_m$ such that

$$\bar{B}(U^m, v^m) + (r^m \partial U^m, v^m)_{d^*} = (f^m, v^m)_{d^*} \quad \forall v^m \in \bar{V}_m, \tag{3.2.9}$$

$$U^m(0) = q_0(t_m) \text{ and } U^m(1) = q_1(t_m), \tag{3.2.10}$$

$$U^0 = (u^0)_{S_0}, \tag{3.2.11}$$

where

$$\bar{B}(w,z) = \varepsilon(w_x, z_x) + (\bar{a} w_x, z) + (bw, z)_{d^*}. \tag{3.2.12}$$

The existence and uniqueness of U^m are implied by Lemma 3.2.1 below. The scheme (3.2.9) – (3.2.11) is similar to one given in Stynes and O'Riordan [45] where $\tilde{a}$ was used instead of $\bar{a}$. We have used $\bar{a}$ in order to obtain the coercivity result of Lemma 3.2.1. The approximation $\bar{a}$ is essentially that of Stynes [43]. In implementing the scheme, at each time step we have to solve a linear system whose coefficient matrix is tridiagonal and diagonally dominant; this can be done efficiently in $O(N)$ operations.

To analyse the scheme we introduce the following discrete norms, which are defined for all $w \in H^1([0,1] \times \{t_m\})$:

$$\text{(i) discrete } L^2 \text{ norm:} \quad \|w\|_{d^*} = (w, w)_{d^*}^{1/2}, \tag{3.2.13}$$

$$\text{(ii) discrete energy norm:} \quad |||w|||_{d^*} = \{\varepsilon\|w_x\|^2 + \|w\|_{d^*}^2\}^{1/2}, \tag{3.2.14}$$

where $\|\cdot\|$ is the usual $L^2(0,1)$ norm. For notational simplicity we do not refer to m when we write these norms; its value will always be clear from the context in which the norms are used.

We begin our analysis by showing the following coercivity result.

Lemma 3.2.1 *For each $v^m \in S_m$, $m \in \{0,\ldots,M\}$, and H sufficiently small (independently of ε),*

$$\begin{aligned}
&\bar{B}(v^m, (v^m)_{V_m}) + (r^m \partial v^m, (v^m)_{V_m})_{d^*} \\
&\quad \geq C_1 |||(v^m)_{V_m}|||_{d^*}^2 + \frac{1}{2k_m}\{(r^m v^m, v^m)_{d^*} - (r^m v^{m-1}, v^{m-1})_{d^*}\},
\end{aligned}$$

where C_1 is as in (1.1.5).

Proof. Recalling the definition (3.2.12), we have

$$\begin{aligned}
&\bar{B}(v^m, (v^m)_{V_m}) \\
&\quad = \bar{B}((v^m)_{V_m}, (v^m)_{V_m}) + \bar{B}(v^m - (v^m)_{V_m}, (v^m)_{V_m}) \\
&\quad = \varepsilon \| ((v^m)_{V_m})_x \|^2 + \sum_{i=1}^{N-1} \left\{ b_i^m (1, \bar{\psi}_{i,m}) - \frac{1}{2}(\bar{a}_{i+1}^m - \bar{a}_i^m) \right\} (v_i^m)^2,
\end{aligned} \tag{3.2.15}$$

on integrating by parts and using (3.2.5).

A calculation shows that

$$\bar{a}_{i+1}^m - \bar{a}_i^m = (1, \bar{\psi}_{i,m})(a_x(x_i, t_m) + O(H)).$$

Combining the above with (3.2.15) and using (1.1.5), we obtain, for H sufficiently small (independently of ε),

$$\bar{B}(v^m, (v^m)_{\bar{V}_m}) \geq \varepsilon \| ((v^m)_{\bar{V}_m})_x \|^2 + C_1 \|v^m\|^2_{d^*}. \tag{3.2.16}$$

For the other term, we have

$$\begin{aligned}
&(r^m \partial v^m, (v^m)_{\bar{V}_m})_{d^*} \\
&= \frac{1}{2k_m}\{(r^m v^m, v^m)_{d^*} - (r^m v^{m-1}, v^{m-1})_{d^*} + (r^m, (v^m - v^{m-1})^2)_{d^*}\} \\
&\geq \frac{1}{2k_m}\{(r^m v^m, v^m)_{d^*} - (r^m v^{m-1}, v^{m-1})_{d^*}\}.
\end{aligned}$$

Combining this with (3.2.16) yields the desired result. □

In the following subsections we will derive error estimates for the scheme in various norms.

3.2.2 Global L^2 and Energy Norm Error Estimates

In this subsection we will derive a global discrete norm error estimate for (3.2.9) – (3.2.11). To this end, for each m, let $u^I(x, t_m)$ be the interpolant from S_m to the exact solution $u(x, t_m)$, and let U^m be the solution of (3.2.9) – (3.2.11). Set

$$Z^m = u^I(x, t_m) - U^m \quad \text{and} \quad \eta^m = u^I(x, t_m) - u(x, t_m).$$

Then using (2.1.1), (1.1.2), (1.1.3), (3.2.9) – (3.2.11) and $\eta^m(x_i) = 0$ for all i, we have

$$\bar{B}(Z^m, v^m) + (r^m \partial Z^m, v^m)_{d^*} = \bar{R}(u^m, v^m) \quad \forall v^m \in \bar{V}_m, \tag{3.2.17}$$

$$Z^m(0) = Z^m(1) = 0, \quad Z^0 = 0, \tag{3.2.18}$$

where

$$\bar{R}(u^m, v^m) = ((\bar{a}^m - a^m)u_x^m, v^m) + \{(\theta^m, v^m) - (\theta^m, v^m)_{d^*}\}$$
$$+ (r^m(\partial u^m - u_t^m), v^m)_{d^*} \tag{3.2.19}$$

and

$$\theta = f - bu - ru_t.$$

Lemma 3.2.2 *For each $m \in \{1, \dots, M\}$ and any $v^m \in \bar{V}_m$,*

$$|\bar{R}(u^m, v^m)| \le CH \left\{ \|u_x^m\|_{L^1(0,1)}^2 + \|\theta_x^m\|_{L^1(0,1)}^2 \right\}$$
$$+ C\|\partial u^m - u_t^m\|_{d^*}^2 + (C_1/2)\|v^m\|_{d^*}^2,$$

where C_1 is as in (1.1.5).

Proof. We bound the three terms in (3.2.19) separately. First,

$$\begin{aligned}
&|((\bar{a}^m - a^m)u_x^m, v^m)| \\
&\le C \sum_{i=1}^{N} h_i \int_{x_{i-1}}^{x_i} |u_x^m||v^m|\, dx \\
&\le C \sum_{i=1}^{N-1} (h_i + h_{i+1})|v_i^m| \int_{x_{i-1}}^{x_{i+1}} |u_x^m|\, dx \\
&\le (C_1/12) \sum_{i=1}^{N-1} h_i |v_i^m|^2 + C \sum_{i=1}^{N-1} h_i \left(\int_{x_{i-1}}^{x_{i+1}} |u_x^m|\, dx \right)^2, \\
&\qquad \text{by the mesh grading and (2.3.9),} \\
&\le (C_1/6) \sum_i (1, \bar{\psi}_{i,m})\, |v_i^m|^2 + CH \left\{ \sum_i \int_{x_{i-1}}^{x_{i+1}} |u_x^m|\, dx \right\}^2, \quad \text{using (3.2.7),} \\
&\le (C_1/6)\|v^m\|_{d^*}^2 + CH\|u_x^m\|_{L^1(0,1)}^2.
\end{aligned} \tag{3.2.20}$$

Next, similarly to the derivation of (3.2.20), we have

$$|(\theta^m, v^m) - (\theta^m, v^m)_{d^*}|$$

$$= \left| \sum_{i=1}^{N-1} \left(\theta^m - \theta^m(x_i), \bar{\psi}_{i,m} \right) v_i^m \right|$$
$$\le C \sum_i (1, \bar{\psi}_{i,m}) |v_i^m| \int_{x_{i-1}}^{x_{i+1}} |\theta_x^m(x)|\, dx$$
$$\le (C_1/6) \|v^m\|_{d^*}^2 + CH \|\theta_x^m\|_{L^1(0,1)}^2. \tag{3.2.21}$$

Finally,

$$|(r^m(\partial u^m - u_t^m), v^m)_{d^*}| \le (C_1/6) \|v^m\|_{d^*}^2 + C \|\partial u^m - u_t^m\|_{d^*}^2. \tag{3.2.22}$$

Combining (3.2.20) – (3.2.22), we are done. □

We can now bound the error of our computed solution in a discrete L^2 norm.

Theorem 3.2.1 *Let U^m be defined by (3.2.9) – (3.2.11). For H sufficiently small (independently of ε), and any $n \in \{1, \ldots, M\}$,*

$$\|U^n - u^n\|_{d^*} \le CH^{1/2} \left\{ \|u_x\|_{L^2(L^1(\Omega_K))} + \|\theta_x\|_{L^2(L^1(\Omega_K))} \right\}$$
$$+ CK^{1/2} \|u_{tt}\|_{L^2(L^1(\Omega_H))}^*,$$

where $\|\cdot\|_{L^2(L^1(\Omega_K))}$ and $\|\cdot\|_{L^2(L^1(\Omega_H))}^*$ are defined in (2.3.3) and (2.3.5) respectively.

Proof. Take $v^m = Z^m$ in Lemma 3.2.1 and $v^m = (Z^m)_{\bar{V}_m}$ in Lemma 3.2.2 to get for each m,

$$C_1 |||(Z^m)_{\bar{V}_m}|||^2 + \frac{1}{k_m} \left\{ (r^m Z^m, Z^m)_{d^*} - (r^m Z^{m-1}, Z^{m-1})_{d^*} \right\}$$
$$\le CH \left\{ \|u_x^m\|_{L^1(0,1)}^2 + \|\theta_x^m\|_{L^1(0,1)}^2 \right\} + C \|\partial u^m - u_t^m\|_{d^*}^2.$$

Multiplying by k_m, and summing from $m = 1$ to $m = n \le M$,

$$C_1 \sum_{m=1}^{n} k_m |||(Z^m)_{\bar{V}_m}|||^2 + \nu \|Z^n\|_{d^*}^2$$

$$\leq CH\left\{\|u_x\|^2_{L^2(L^1(\Omega_K))}+\|\theta_x\|^2_{L^2(L^1(\Omega_K))}\right\}$$
$$+CK\left(\|u_{tt}\|^*_{L^2(L^1(\Omega_H))}\right)^2+C\sum_{m=1}^{n-1}k_m\|Z^m\|^2_{d^*},$$

since

$$\begin{aligned}
&\sum_{m=1}^{M}k_m\|\partial u^m-u_t^m\|^2_{d^*}\\
&=\sum_{m=1}^{M}k_m\sum_{i=1}^{N-1}(1,\bar{\psi}_{i,m})|\partial u^m(x_i)-u_t(x_i,t_m)|^2\\
&\leq CK\sum_{i=1}^{N-1}(h_i+h_{i+1})\sum_{m=1}^{M}\left(\int_{t_{m-1}}^{t_m}|u_{tt}(x_i,t)|\,dt\right)^2\\
&\leq CK\sum_{i=1}^{N-1}(h_i+h_{i+1})\left(\int_0^T|u_{tt}(x_i,t)|\,dt\right)^2\\
&=CK\left(\|u_{tt}\|^*_{L^2(L^1(\Omega_H))}\right)^2.
\end{aligned}$$

Use a discrete Gronwall's inequality to complete the argument. □

Remark 3.2.1 Theorem 3.2.1 is valid for any trial functions $\{\phi_{i,m}(x)\}$ which satisfy (3.2.1) and (3.2.2). If the trial functions are specified as $\tilde{L}$-spline functions, then one can prove the same bounds in the discrete energy norm.

That is, we have

Theorem 3.2.2 *Suppose that the trial functions* $\{\tilde{\phi}_{i,m} : i = 0,\ldots,N$ *and* $m = 0,\ldots,M\}$ *are given by*

$$\tilde{L}\tilde{\phi}_{i,m}\equiv-\varepsilon\tilde{\phi}''_{i,m}(x)+\tilde{a}_i^m\tilde{\phi}'_{i,m}(x)=0\quad \textit{for } x\in[0,1]^{\hat{}},\tag{3.2.23}$$

$$\tilde{\phi}_{i,m}(x_j)=\delta_{i,j}\quad \textit{for } j=0,\ldots,N.\tag{3.2.24}$$

If $U^m = \sum_{i=0}^{N} U_i^m \tilde{\phi}_{i,m}(x)$ satisfies (3.2.9) - (3.2.11), then for H sufficiently small (independently of ε), and any $n \in \{1,\dots,M\}$,

$$\|U^n - u^n\|_{d^*} + \left\{\sum_{m=1}^{n} k_m |||U^m - u^m|||_{d^*}^2\right\}^{1/2}$$
$$\leq CH^{1/2}\left\{\|u_x\|_{L^2(L^1(\Omega_K))} + \|\theta_x\|_{L^2(L^1(\Omega_K))} + \|\theta\|_{L^2(L^\infty(\Omega_K))}\right\}$$
$$+ CK^{1/2}\|u_{tt}\|_{L^2(L^1(\Omega_H))}^*,$$

where the norms on the right hand side are defined in (2.3.3) - (2.3.5).

Proof. Noting that $|||Z^m|||_{d^*} \leq C|||(Z^m)_{V_m}|||_{d^*}$, it is straightforward to get, from the proof of Theorem 3.2.1,

$$\|Z^n\|_{d^*} + \left\{\sum_{m=1}^{n} k_m |||Z^m|||_{d^*}^2\right\}^{1/2}$$
$$\leq CH^{1/2}\left\{\|u_x\|_{L^2(L^1(\Omega_K))} + \|\theta_x\|_{L^2(L^1(\Omega_K))}\right\}$$
$$+ CK^{1/2}\|u_{tt}\|_{L^2(L^1(\Omega_H))}^*. \tag{3.2.25}$$

We now need to estimate the interpolation error η^m for all m. We have, for $t = t_m$,

$$B(\eta^m, \eta^m) = \left((a^m - \tilde{a}^m)u_x^I(\cdot, t_m), \eta^m\right) + (b^m\eta^m, \eta^m) - (\theta^m, \eta^m).$$

Since for $x \in [x_{i-1}, x_i]$,

$$|\eta^m(x)| \leq \int_{x_{i-1}}^{x_i} |u_x^m(x)|\,dx,$$

we also have

$$\int_0^1 |\eta^m(x)|\,dx = \sum_{i=1}^{N}\int_{x_{i-1}}^{x_i} |\eta^m(x)|\,dx \leq H\|u_x^m\|_{L^1(0,1)}.$$

Hence

$$B(\eta^m,\eta^m) \le CH\|\eta^m\|_{L^\infty(0,1)}\|u_x^m\|_{L^1(0,1)} + CH\|\theta^m\|_{L^\infty(0,1)}\|u_x^m\|_{L^1(0,1)}$$
$$\le CH\left\{\|u_x^m\|^2_{L^1(0,1)} + \|\theta^m\|^2_{L^\infty(0,1)}\right\}.$$

On the other hand,

$$B(\eta^m,\eta^m) \ge \varepsilon\|\eta_x^m\|^2 + C_1\|\eta^m\|^2 \ge C_1|||\eta^m|||^2_{d^*}.$$

Thus

$$|||\eta^m|||^2_{d^*} \le CH\left\{\|u_x^m\|^2_{L^1(0,1)} + \|\theta^m\|^2_{L^\infty(0,1)}\right\}. \tag{3.2.26}$$

Combining this with (3.2.25) completes the proof. □

3.2.3 Improved Accuracy Under Compatibility Assumptions

In this subsection we assume that we have certain bounds on the derivatives of the solution $u(x,t)$. Such bounds follow, for example, if we assume that the data satisfies certain compatibility conditions at the corners $(0,0)$ and $(1,0)$ of Ω. These conditions are given in Stynes and O'Riordan [45], and it is shown how to use them to obtain the bounds

$$\left|D_x^i D_t^j u(x,t)\right| \le C\{1+\varepsilon^{-i}\exp(-\alpha(1-x)/\varepsilon)\} \quad \forall (x,t)\in\Omega, \tag{3.2.27}$$

for $0 \le i \le 1$ and $0 \le i+j \le 2$.

Using (3.2.27), it is possible to significantly improve the results stated in the previous subsection. Before doing this, we give the following technical result, which can be proved in the same way as Lemma 5.6 of Stynes and O'Riordan [46].

Lemma 3.2.3 *For each $m \in \{1,\ldots,M\}$ and any $v^m \in \bar{V}_m$*

$$\sum_{i=0}^{N-1}\left(\int_{x_i}^1 |v_x^m|\,dx\right)\left(\int_{x_i}^{x_{i+1}} \varepsilon^{-1}\exp(-\alpha(1-x)/\varepsilon)\,dx\right) \le C\varepsilon^{1/2}\|v_x^m\|.$$

Using this lemma and (3.2.27) we now sharpen Lemma 3.2.2.

Lemma 3.2.4 *Assume that (3.2.27) holds. For each $m \in \{1, \ldots, M\}$ and any $v^m \in \bar{V}_m$,*

$$|\bar{R}(u^m, v^m)| \le C(H^2 + K^2) + (C_1/2)|||v^m|||_{d^*}^2,$$

where C_1 is as in (1.1.5).

Proof. Recall (3.2.19). First, by (3.2.27),

$$
\begin{aligned}
&|((\bar{a}^m - a^m)u_x^m, v^m)| \\
&\le C \sum_{i=1}^{N} h_i \int_{x_{i-1}}^{x_i} |u_x^m||v^m|\, dx \\
&\le C \sum_{i=1}^{N} h_i \int_{x_{i-1}}^{x_i} \{1 + \varepsilon^{-1} \exp(-\alpha(1-x)/\varepsilon)\}|v^m|\, dx \\
&\le CH \sum_{i=1}^{N-1} (1, \bar{\psi}_{i,m})|v_i^m| \\
&\quad + CH \sum_{i=0}^{N-1} \left(\int_{x_i}^{1} |v_x^m|\, dx\right) \left(\int_{x_i}^{x_{i+1}} \varepsilon^{-1} \exp(-\alpha(1-x)/\varepsilon)\, dx\right) \\
&\le CH\|v^m\|_{d^*} + CH\varepsilon^{1/2}\|v_x^m\|, \quad \text{by Lemma 3.2.3,} \\
&\le (C_1/6)|||v^m|||_{d^*}^2 + CH^2. \qquad (3.2.28)
\end{aligned}
$$

Next, we can show by an argument similar to the above that

$$|(\theta^m, v^m) - (\theta^m, v^m)_{d^*}| \le (C_1/6)|||v^m|||_{d^*}^2 + CH^2. \qquad (3.2.29)$$

Finally, since by (3.2.27)

$$\|\partial u^m - u_t^m\|_{d^*} \le CK,$$

we get

$$|(r^m(\partial u^m - u_t^m), v^m)_{d^*}| \le (C_1/6)\|v^m\|_{d^*}^2 + CK^2. \qquad (3.2.30)$$

Combining (3.2.28) – (3.2.30) yields the desired result. □

Using Lemmas 3.2.1 and 3.2.4, it is easy to get the following results.

Theorem 3.2.3 *Assume that (3.2.27) and the hypotheses of Theorem 3.2.1 hold. Then*

$$\|U^n - u^n\|_{d^*} \leq C(H + K), \quad \text{for } n = 1, \ldots, M.$$

Theorem 3.2.4 *Assume that (3.2.27) and the hypotheses of Theorem 3.2.2 hold. Then*

$$\|U^n - u^n\|_{d^*} \leq C(H + K),$$

and

$$\left\{ \sum_{m=1}^{n} k_m |||U^m - u^m|||_{d^*}^2 \right\}^{1/2} \leq C(H^{1/2} + K),$$

for $n = 1, \ldots, M$.

Remark 3.2.2 It can be shown that the factor $H^{1/2}$ in Theorem 3.2.4 is the optimal order attainable.

Remark 3.2.3 It can be shown that results analogous to those of Sections 3.2.2 and 3.2.3 hold for the semidiscrete version of (3.2.9) – (3.2.11), except that no u_{tt} term is present.

3.2.4 Localized Pointwise Error Estimates

In this subsection we use a variant of Niijima's approach [32] to derive pointwise error estimates under local smoothness assumptions. For simplicity, we consider a constant coefficient problem with $a(x,t) = b(x,t) = r(x,t) = 1$, and we employ a uniform mesh with space mesh size H and time mesh size K. Our trial space S_m

is as in Subsection 3.2.1, i.e., its basis functions $\phi_{i,m}$ need only satisfy (3.2.1) and (3.2.2).

Let (x_{i_0}, t_{m_0}) be an interior mesh point. We associate a discrete Green's function $G(x,t)$ with this point. That is, $G^m(x) \equiv G(x,t_m) \in \bar{V}_m$ is defined for $m = 0, \ldots, M$ by

$$\bar{B}(\chi^m, G^m) - (\chi^m, \partial G^{m+1})_{d^*} = K^{-1}\delta_{m,m_0}\chi^m(x_{i_0}) \quad \forall \chi^m \in S_m^0, \tag{3.2.31}$$

$$G^m(0) = G^m(1) = 0, \tag{3.2.32}$$

where $S_m^0 = \{\chi^m \in S_m : \chi^m(0) = \chi^m(1) = 0\}$, and we formally set

$$G^{M+1}(x) \equiv 0. \tag{3.2.33}$$

Note that the definition of the discrete Green's function here is not immediate from the definition of such functions for elliptic problems (as used in [32]).

The following lemma gives the existence and uniqueness of G^m.

Lemma 3.2.5 *The equations (3.2.31) - (3.2.33) have a unique solution G^m. Furthermore,*

$$G_i^m \geq 0, \text{ for } i = 0, \ldots, N \text{ and } m = 0, 1, \ldots, M. \tag{3.2.34}$$

Proof. Equations (3.2.31) – (3.2.33) may be written as

$$\begin{aligned} G_{i-1}^m \bar{B}(\phi_{i,m}, \bar{\psi}_{i-1,m}) &+ G_i^m \{\bar{B}(\phi_{i,m}, \bar{\psi}_{i,m}) + K^{-1}(1, \bar{\psi}_{i,m})\} \\ &+ G_{i+1}^m \bar{B}(\phi_{i,m}, \bar{\psi}_{i+1,m}) \\ &= K^{-1}\{\delta_{m,m_0}\delta_{i,i_0} + (1, \bar{\psi}_{i,m})G_i^{m+1}\}, \end{aligned} \tag{3.2.35}$$

for $i = 1, \ldots, N-1$ and $m = 0, \ldots, M$,

$$G_0^m = G_N^m = 0 \text{ for } m = 1, \ldots, M, \tag{3.2.36}$$

$$G_i^{M+1} = 0 \text{ for } i = 0, \ldots, N. \tag{3.2.37}$$

Thus it suffices to show that the system (3.2.35) – (3.2.37) has a unique solution. Inspecting the coefficients of G_j^m on the left hand side, we see that the coefficient matrix of the system is a strictly diagonally dominant tridiagonal matrix with positive diagonal terms and nonpositive off-diagonal terms. Consequently it is an M-matrix and so is invertible. That is, the system (3.2.35) – (3.2.37) can be solved iteratively for G^m in terms of G^{m+1} for $m = M, M-1, \ldots, 0$, since G^{M+1} is known. This completes the proof of the existence and uniqueness.

Since for each m the inverse of the coefficient matrix is nonnegative, it is straightforward to show (3.2.34) by using induction on m. □

Next, we derive an L^1 estimate on G along mesh-lines parallel to the x-axis.

Lemma 3.2.6 *For each $n \in \{0, \ldots, M\}$,*

$$(1, G^n)_{d^*} \leq 1.$$

Proof. Fix $m \in \{0, \ldots, M\}$. Taking $\chi^m = \sum_{i=1}^{N-1} \phi_{i,m}(x)$ in (3.2.31), we have

$$\begin{aligned} \int_0^{x_1} (\varepsilon G_x^m + G^m)\phi_{1,m}'(x)\,dx + \int_{x_{N-1}}^1 (\varepsilon G_x^m + G^m)\phi_{N-1,m}'(x)\,dx \\ + (1, G^m)_{d^*} + K^{-1}(1, G^m - G^{m+1})_{d^*} = K^{-1}\delta_{m,m_0}. \qquad (3.2.38) \end{aligned}$$

Integrating by parts and using the definition of the test functions, we get

$$\begin{aligned} &\int_0^{x_1} (\varepsilon G_x^m + G^m)\phi_{1,m}'(x)\,dx + \int_{x_{N-1}}^1 (\varepsilon G_x^m + G^m)\phi_{N-1,m}'(x)\,dx \\ &= G_1^m\left(\varepsilon\bar{\psi}_{1,m}'(x_1 - 0) + 1\right) - G_{N-1}^m\left(\varepsilon\bar{\psi}_{N-1,m}'(x_{N-1} + 0) + 1\right) \\ &= G_1^m\big(1 - \exp(-H/\varepsilon)\big)^{-1} + G_{N-1}^m\big(\exp(H/\varepsilon) - 1\big)^{-1}. \qquad (3.2.39) \end{aligned}$$

Thus, we get from (3.2.34), (3.2.38) and (3.2.39) that

$$(1, G^m - G^{m+1})_{d^*} \leq \delta_{m,m_0}.$$

Summing from $m = n$ to $m = M$ for any $n \leq M$, and noting that $G^{M+1} = 0$, we obtain

$$(1, G^n)_{d^*} \leq 1$$

as desired. □

Using the L^1 estimate on G, one now is able to estimate the error at (x_{i_0}, t_{m_0}), provided (3.2.27) holds on the entire region Ω. However this assumption is stronger than needed. We shall need such an assumption (in Theorem 3.2.6 below) only on a narrow region extending upstream from (x_{i_0}, t_{m_0}). This is because the discrete Green's function dies off outside region

$$\Omega_0 = \left\{(x,t) \in \Omega : 0 < x \leq x_{i_0} + K_0 \varepsilon^* \ln\left(\frac{1}{HK}\right), \right.$$
$$\left. |x - t - (x_{i_0} - t_{m_0})| \leq K_0 \sqrt{\varepsilon^*} \ln\left(\frac{1}{HK}\right)\right\},$$

with $\varepsilon^* = \max\{\varepsilon, H, K\}$, where $K_0 > 0$ is a constant independent of ε, H and K, which we choose later. We shall prove this fact in Lemma 3.2.7. Without loss of generality, we assume that Ω_0 is a mesh domain.

Lemma 3.2.7 *Given a nonnegative integer s, there exists a positive constant $C = C(s)$ such that*

$$\max_{(x,t_m) \in \Omega \setminus \Omega_0} G(x, t_m) \leq C(s)(HK)^s$$

for each $m \in \{0, \ldots, M\}$.

Proof. We introduce a cut-off function $\Phi(\lambda)$, defined on $(-\infty, \infty)$ by

$$\Phi(\lambda) = \int_\lambda^\infty \exp(-\xi(\tau))\, d\tau \tag{3.2.40}$$

with $\xi(\tau) \in C^2(-\infty, +\infty)$ and $\xi(\tau) = |\tau|$ for $|\tau| \geq 1$. Then define a cut-off function $\omega(x,t)$ on Ω by

$$\omega(x,t) = \Phi\left(\frac{x-A}{\sigma_s}\right)\Phi\left(\frac{x-t-P}{\sigma_\eta}\right)\Phi\left(\frac{P-x+t}{\sigma_\eta}\right) \tag{3.2.41}$$

where $A = x_{i_0}$, $P = x_{i_0} - t_{m_0}$, $\sigma_s = \gamma\varepsilon^*$ and $\sigma_\eta = \gamma\sqrt{\varepsilon^*}$. Here $\gamma > 1$ is some constant (to be specified later) independent of ε, H and K.

For all differentiable $v(x,t)$, set

$$v_\beta = v_s + v_t. \tag{3.2.42}$$

Then it is easy to show that

$$-\omega_\beta > 0 \qquad \text{on } \Omega, \tag{3.2.43}$$

$$\max_{\Delta_i^m} \omega / \min_{\Delta_i^m} \omega \leq C, \quad \max_{\Delta_i^m} |\omega_\beta| / \min_{\Delta_i^m} |\omega_\beta| \leq C, \tag{3.2.44}$$

for $i = 1, \ldots, N$ and $m = 1, \ldots, M$, with

$$\Delta_i^m = \{(x,t) : \ x_{i-1} \leq x \leq x_i, \quad t_{m-1} \leq t \leq t_m\}.$$

Furthermore,

$$\left|D_\beta^j D_t^l \omega\right| \leq C\sigma_s^{-j}\sigma_\eta^{-l}\omega \quad \text{ on } \Omega \text{ for } j + l \leq 2, \tag{3.2.45}$$

$$\left|D_\beta^j D_t^l \omega\right| \leq C\sigma_s^{-j+1}\sigma_\eta^{-l}|\omega_\beta| \quad \text{ on } \Omega \text{ for } j + l \leq 2 \text{ and } j \geq 1, \tag{3.2.46}$$

$$\omega(x_{i_0}, t_{m_0}) \geq C, \tag{3.2.47}$$

$$\omega(x,t) \leq C(HK)^{K_0/\gamma} \qquad \text{on } \Omega \backslash \Omega_0. \tag{3.2.48}$$

Now we take $\chi^m = (\frac{G^m}{\omega^m})_{S_m}$ in (3.2.31) to get

$$\bar{B}\left(\left(\frac{G^m}{\omega^m}\right)_{S_m}, G^m\right) - \left(\frac{G^m}{\omega^m}, \partial G^{m+1}\right)_{d^*} = K^{-1}\delta_{m,m_0}\left(\frac{G^m}{\omega^m}\right)(x_{i_0}). \tag{3.2.49}$$

We derive a lower bound for the left hand side of (3.2.49).

First, by integrating by parts,

$$\bar{B}\left(\left(\frac{G^m}{\omega^m}\right)_{S_m} - \frac{G^m}{\omega^m}, G^m\right) = \left(\left(\frac{G^m}{\omega^m}\right)_{S_m} - \frac{G^m}{\omega^m}, -\varepsilon G^m_{xx} - G^m_x\right)^{\hat{}} = 0,$$

since $G^m \in \bar{V}_m$. Thus,

$$\begin{aligned}
&\bar{B}\left(\left(\frac{G^m}{\omega^m}\right)_{S_m}, G^m\right)\\
&= \bar{B}\left(\frac{G^m}{\omega^m}, G^m\right)\\
&= \varepsilon\|(\omega^m)^{-1/2}G^m_x\|^2 + \|(\omega^m)^{-1/2}G^m\|^2_{d^*}\\
&\quad + \frac{1}{2}\left(\left(\frac{1}{\omega^m}\right)_x, (G^m)^2\right) + \varepsilon\left(\left(\frac{1}{\omega^m}\right)_x G^m, G^m_x\right). \qquad (3.2.50)
\end{aligned}$$

For the other term on the left hand side of (3.2.49), we have

$$\begin{aligned}
&-\left(\frac{G^m}{\omega^m}, \partial G^{m+1}\right)_{d^*}\\
&\geq \frac{1}{2K}\left\{\|(\omega^m)^{-1/2}G^m\|^2_{d^*} - \|(\omega^{m+1})^{-1/2}G^{m+1}\|^2_{d^*}\right\}\\
&\quad + \frac{1}{2K}\left(\frac{\omega^m - \omega^{m+1}}{(\omega^m)^2}, (G^m)^2\right)_{d^*}. \qquad (3.2.51)
\end{aligned}$$

Therefore from (3.2.50) and (3.2.51) we have

$$\begin{aligned}
&\bar{B}\left(\left(\frac{G^m}{\omega^m}\right)_{S_m}, G^m\right) - \left(\frac{G^m}{\omega^m}, \partial G^{m+1}\right)_{d^*}\\
&\geq \frac{1}{2K}\left\{\|(\omega^m)^{-1/2}G^m\|^2_{d^*} - \|(\omega^{m+1})^{-1/2}G^{m+1}\|^2_{d^*}\right\} + I_m + Q_m,
\end{aligned}$$

(3.2.52)

where

$$I_m = \varepsilon\|(\omega^m)^{-1/2}G^m_x\|^2 + \|(\omega^m)^{-1/2}G^m\|^2_{d^*} + \frac{1}{2}\left\|\left(\frac{1}{\omega^m}\right)^{1/2}_{\beta} G^m\right\|^2, \qquad (3.2.53)$$

$$Q_m = \varepsilon\left(\left(\frac{1}{\omega^m}\right)_x G^m, G^m_x\right) + \frac{1}{2}\left(\frac{\omega^m - \omega^{m+1}}{K} + \omega^m_t, \left(\frac{G^m}{\omega^m}\right)^2\right)$$
$$+ \frac{1}{2K}\left\{\left(\frac{\omega^m - \omega^{m+1}}{(\omega^m)^2}, (G^m)^2\right)_{d^*} - \left(\frac{\omega^m - \omega^{m+1}}{(\omega^m)^2}, (G^m)^2\right)\right\}. \quad (3.2.54)$$

We bound the three terms in (3.2.54) separately. For the first term, we use Cauchy-Schwarz' inequality and (2.3.9), to obtain

$$\left|\varepsilon\left(\left(\frac{1}{\omega^m}\right)_x G^m, G^m_x\right)\right|$$
$$\leq \frac{\varepsilon}{4}\|(\omega^m)^{-1/2}G^m_x\|^2 + C\varepsilon\left\|\left(\frac{1}{\omega^m}\right)_x (\omega^m)^{1/2}G^m\right\|^2$$
$$\leq \frac{\varepsilon}{4}\|(\omega^m)^{-1/2}G^m_x\|^2 + C\varepsilon\left\|\left(\frac{1}{\omega^m}\right)_\beta (\omega^m)^{1/2}G^m\right\|^2$$
$$+ C\varepsilon\left\|\left(\frac{1}{\omega^m}\right)_t (\omega^m)^{1/2}G^m\right\|^2$$
$$\leq \frac{\varepsilon}{4}\|(\omega^m)^{-1/2}G^m_x\|^2 + C\varepsilon \max\frac{|\omega_\beta|}{\omega}\left\|\left(\frac{1}{\omega^m}\right)^{1/2}_\beta G^m\right\|^2$$
$$+ C\varepsilon \max\frac{|\omega_t|^2}{\omega^2}\|(\omega^m)^{-1/2}G^m\|^2$$
$$\leq \frac{\varepsilon}{4}\|(\omega^m)^{-1/2}G^m_x\|^2 + C\varepsilon\sigma_x^{-1}\left\|\left(\frac{1}{\omega^m}\right)^{1/2}_\beta G^m\right\|^2$$
$$+ C\varepsilon\sigma_\eta^{-2}\|(\omega^m)^{-1/2}G^m\|^2, \qquad \text{by (3.2.45)},$$
$$\leq \frac{\varepsilon}{4}\|(\omega^m)^{-1/2}G^m_x\|^2 + C\varepsilon(\sigma_x^{-1} + \sigma_\eta^{-2})I_m. \quad (3.2.55)$$

We next proceed to estimate the second term in (3.2.54).

$$\frac{1}{2}\left|\left(\frac{\omega^m - \omega^{m+1}}{K} + \omega^m_t, \left(\frac{G^m}{\omega^m}\right)^2\right)\right|$$
$$\leq CK\sum_{i=1}^{N} \max_{\Delta_i^{m+1}} |\omega_{tt}| \int_{x_{i-1}}^{x_i} \left(\frac{G^m}{\omega^m}\right)^2 dx, \quad \text{by a Taylor expansion},$$
$$\leq CK\sigma_\eta^{-2}\sum_{i=1}^{N} \frac{\max_{\Delta_i^{m+1}}\omega}{\min_{\Delta_i^{m+1}}\omega}\int_{x_{i-1}}^{x_i} (\omega^m)^{-1}(G^m)^2\, dx, \quad \text{using (3.2.45)},$$
$$\leq CK\sigma_\eta^{-2}\|(\omega^m)^{-1/2}G^m\|^2_{d^*}, \quad (3.2.56)$$

since

$$\int_{x_{i-1}}^{x_i} (\omega^m)^{-1}(G^m)^2\,dx$$
$$\le C\left((1,\bar\psi_{i-1,m})(\omega_{i-1}^m)^{-1}|G_{i-1}^m|^2 + (1,\bar\psi_{i,m})(\omega_i^m)^{-1}|G_i^m|^2\right) \quad (3.2.57)$$

follows from (3.2.44) and (3.2.7).

To bound the last term in (3.2.54), we set

$$\Pi^m = \frac{\omega^m - \omega^{m+1}}{(\omega^m)^2}.$$

Then writing G^m as $\sum_{i=1}^{N-1} G_i^m \bar\psi_{i,m}$ we get

$$\begin{aligned}
&\frac{1}{2K}\left|(\Pi^m,(G^m)^2)_{d^*} - (\Pi^m,(G^m)^2)\right| \\
&\quad= \frac{1}{2K}\left|\sum_{i=1}^{N-1} G_i^m\left\{(\Pi_i^m G_i^m,\bar\psi_{i,m}) - (\Pi^m G^m,\bar\psi_{i,m})\right\}\right| \\
&\quad\le \frac{1}{2K}\sum_{i=1}^{N-1}(1,\bar\psi_{i,m})|G_i^m|\int_{x_{i-1}}^{x_{i+1}} |(\Pi^m G^m)_x|\,dx \\
&\quad\le \frac{1}{2K}\sum_{i=1}^{N-1}(1,\bar\psi_{i,m})|G_i^m|\left\{\max_{\Delta_i}|\Pi_x^m|\int_{x_{i-1}}^{x_{i+1}}|G^m|\,dx\right. \\
&\qquad\left.+\max_{\Delta_i}|\Pi^m|\int_{x_{i-1}}^{x_{i+1}}|G_x^m|\,dx\right\}. \qquad (3.2.58)
\end{aligned}$$

Set $\Delta_i = [x_{i-1}, x_i]$. Using (3.2.44) – (3.2.46), we have for $i = 1,\ldots,N-1$,

$$\begin{aligned}
\max_{\Delta_i}|\Pi^m(x)| &\le K\max_{\Delta_i^{m+1}}|\omega_t|/\min_{\Delta_i}(\omega^m)^2 \\
&\le CK\sigma_\eta^{-1}/\min_{\Delta_i}\omega^m, \qquad (3.2.59)
\end{aligned}$$

and

$$\begin{aligned}
&\max_{\Delta_i}|\Pi_x^m(x)| \\
&\quad\le K\max_{\Delta_i^{m+1}}|\omega_{xt}|/\min_{\Delta_i}(\omega^m)^2
\end{aligned}$$

$$+2K\max_{\Delta_i^{m+1}}|\omega_t|\max_{\Delta_i}|\omega_s^m(x)|/\min_{\Delta_i}(\omega^m)^3$$

$$\leq K\max_{\Delta_i^{m+1}}(|\omega_{tt}|+|\omega_{\beta t}|)/\min_{\Delta_i}(\omega^m)^2$$

$$+2K\max_{\Delta_i^{m+1}}|\omega_t|\max_{\Delta_i}(|\omega_t^m(x)|+|\omega_\beta^m(x)|)/\min_{\Delta_i}(\omega^m)^3$$

$$\leq CK\left(\sigma_\eta^{-2}/\min_{\Delta_i}\omega^m+\sigma_\eta^{-1}\sigma_s^{-1/2}\max_{\Delta_i}|\omega_\beta^m|^{1/2}/\min_{\Delta_i}(\omega^m)^{3/2}\right). \quad (3.2.60)$$

Cauchy-Schwarz' inequality gives

$$\int_{s_{i-1}}^{s_i}|G^m|\,dx\leq H^{1/2}\left(\int_{s_{i-1}}^{s_i}(G^m)^2\,dx\right)^{1/2}. \quad (3.2.61)$$

Since $G^m\in\bar{V}_m$, from the definition of the test space $\bar{V}_m$ one can prove

$$\int_{s_{i-1}}^{s_i}|G_s^m|\,dx\leq\varepsilon^{1/2}\left(\int_{s_{i-1}}^{s_i}|G_s^m|^2\,dx\right)^{1/2}, \quad (3.2.62)$$

(cf. Lemma 4.2 of Stynes and O'Riordan [47]). Thus from (3.2.58) – (3.2.62) and (3.2.44) we get

$$\frac{1}{2K}\left|(\Pi^m,(G^m)^2)_{d^*}-(\Pi^m,(G^m)^2)\right|$$

$$\leq C\sum_{i=1}^{N-1}(1,\bar{\psi}_{i,m})(\omega_i^m)^{-1/2}|G_i^m|\left\{\sigma_\eta^{-2}H^{1/2}\left(\int_{s_{i-1}}^{s_{i+1}}(\omega^m)^{-1}(G^m)^2\,dx\right)^{1/2}\right.$$

$$+\sigma_\eta^{-1}\sigma_s^{-1/2}H^{1/2}\left(\int_{s_{i-1}}^{s_{i+1}}\left(\frac{1}{\omega^m}\right)_\beta(G^m)^2\,dx\right)^{1/2}$$

$$\left.+\sigma_\eta^{-1}\varepsilon^{1/2}\left(\int_{s_{i-1}}^{s_{i+1}}(\omega^m)^{-1}|G_s^m|^2\,dx\right)^{1/2}\right\}$$

$$\leq C\sigma_\eta^{-2}H\|(\omega^m)^{-1/2}G^m\|_{d^*}\|(\omega^m)^{-1/2}G^m\|_{d^*}$$

$$+C\sigma_\eta^{-1}\sigma_s^{-1/2}H\|(\omega^m)^{-1/2}G^m\|_{d^*}\left\|\left(\frac{1}{\omega^m}\right)_\beta^{1/2}G^m\right\|$$

$$+C\sigma_\eta^{-1}H^{1/2}\|(\omega^m)^{-1/2}G^m\|_{d^*}\varepsilon^{1/2}\|(\omega^m)^{-1/2}G_s^m\|,$$

by (3.2.57) and Cauchy-Schwarz' inequality,

$$\leq C\left\{\sigma_\eta^{-2}H+\sigma_\eta^{-1}\sigma_s^{-1/2}H+\sigma_\eta^{-1}H^{1/2}\right\}I_m, \quad (3.2.63)$$

using (2.3.9).

Collecting (3.2.55), (3.2.56) and (3.2.63) into (3.2.54) we get

$$\begin{aligned}
|Q_m| &\leq \frac{1}{4}I_m + C\{(\varepsilon^*)^2(\sigma_s^{-2}+\sigma_\eta^{-4}) + \varepsilon^*(\sigma_s^{-1}+\sigma_\eta^{-2})\}I_m \\
&\leq \frac{1}{4}I_m + C\gamma^{-1}I_m, \quad \text{recalling the definition of } \sigma_s \text{ and } \sigma_\eta, \\
&\leq \frac{1}{2}I_m, \qquad (3.2.64)
\end{aligned}$$

by choosing γ sufficiently large, independently of ε, H and K. Consequently, from (3.2.49), (3.2.52) and (3.2.64) we get

$$\begin{aligned}
&\frac{1}{2}I_m + \frac{1}{2K}\{\|(\omega^m)^{-1/2}G^m\|_{d^*}^2 - \|(\omega^{m+1})^{-1/2}G^{m+1}\|_{d^*}^2\} \\
&\leq K^{-1}\delta_{m,m_0}\left(\frac{G^m}{\omega^m}\right)(x_{i_0}). \qquad (3.2.65)
\end{aligned}$$

Multiplying this by K , then summing it from $m = n$ to $m = M$, and noting that $G^{M+1} = 0$, we obtain

$$\sum_{m=n}^{M} KI_m + \|(\omega^n)^{-1/2}G^n\|_{d^*}^2 \leq \begin{cases} 0 & \text{if } n > m_0, \\ 2G_{i_0}^{m_0}/\omega_{i_0}^{m_0} & \text{if } n \leq m_0. \end{cases} \qquad (3.2.66)$$

Using (3.2.47) and Lemma 3.2.6,

$$\frac{G_{i_0}^{m_0}}{\omega_{i_0}^{m_0}} \leq CH^{-1}.$$

Hence from (3.2.66), we get

$$\sum_{m=n}^{M} KI_m + \|(\omega^n)^{-1/2}G^n\|_{d^*}^2 \leq CH^{-1}, \quad \text{for } n = 0,\ldots,M. \qquad (3.2.67)$$

Choose $K_0 = \gamma(2s+2)$. Then by (3.2.48),

$$\omega(x,t) \leq C(HK)^{2s+2} \quad \text{on } \Omega\backslash\Omega_0.$$

For each $(x,t_m) \in \Omega\setminus\Omega_0$, there exists $i' \in \{1,\dots,N\}$ such that $x \in [x_{i'-1}, x_{i'}]$. Thus

$$\begin{aligned}
G^m(x) &\le G^m_{i'-1} + G^m_{i'} \\
&\le C(HK)^{s+1}(\omega^m(x))^{-1/2}(G^m_{i'-1} + G^m_{i'}), \\
&\le C(HK)^{s+1}\left((\omega^m_{i'-1})^{-1/2}G^m_{i'-1} + (\omega^m_{i'})^{-1/2}G^m_{i'}\right), \quad \text{by (3.2.44)}, \\
&\le C(HK)^{s+1/2}K^{1/2}\|(\omega^m)^{-1/2}G^m\|_{d^s}.
\end{aligned}$$

Using this and (3.2.67) we get the desired result. □

Now we can prove our first main result in this subsection. It is a pointwise convergence result, which is obtained under assumptions of local smoothness and reasonable global behaviour of the solution.

Theorem 3.2.5 *Assume that the solution $u(x,t)$ of (1.1.1) - (1.1.5) satisfies*

$$\|u\|_{L^1(L^\infty(\Omega_K))} + \|u^0\|_{L^\infty(0,1)} + \|f\|_{L^1(L^\infty(\Omega_K))} + K\|u_t\|_{L^1(L^1(\Omega_K))} \le C \tag{3.2.68}$$

and

$$\|u\|_{C^2(\Omega_0^+)} + \|f\|_{C^1(\Omega_0^+)} \le C, \tag{3.2.69}$$

where $\|\cdot\|_{L^1(L^\infty(\Omega_K))}$ and $\|\cdot\|_{L^1(L^1(\Omega_K))}$ are defined in (2.3.2) and (2.3.1), and

$$\Omega_0^+ = \{(x,t) \in \Omega : \ \operatorname{dist}((x,t), \Omega_0) \le H + K\}.$$

Then

$$|U(x_{i_0}, t_{m_0}) - u(x_{i_0}, t_{m_0})| \le C(H+K).$$

Proof. With the discrete Green's function G, the pointwise error can be expressed as

$$|U(x_{i_0}, t_{m_0}) - u(x_{i_0}, t_{m_0})|$$

$$
\begin{aligned}
&= |Z(x_{i_0}, t_{m_0})| \\
&= \left| \sum_{m=1}^{M} K\bar{R}(u^m, G^m) \right|, \quad \text{by (3.2.31) -- (3.2.33), (3.2.17) and (3.2.18),} \\
&\le \sum_{m=1}^{M} \sum_{i=1}^{N-1} KG_i^m \left| \bar{R}(u^m, \bar{\psi}_{i,m}) \right|, \qquad (3.2.70)
\end{aligned}
$$

writing G^m as $\sum_{i=1}^{N-1} G_i^m \bar{\psi}_{i,m}$ and using (3.2.34).

Split the sum into two parts:

$$
\sum_{m=1}^{M} \sum_{i=1}^{N-1} = \sum_{(x_i,t_m)\in\Omega_0} + \sum_{(x_i,t_m)\in\Omega\setminus\Omega_0} .
$$

Recalling (3.2.19),

$$
\begin{aligned}
&\sum_{(x_i,t_m)\in\Omega_0} KG_i^m \left| \bar{R}(u^m, \bar{\psi}_{i,m}) \right| \\
&\quad \le \sum_{(x_i,t_m)\in\Omega_0} KG_i^m \{ |(\theta^m - \theta_i^m, \bar{\psi}_{i,m})| + |(\partial u_i^m - u_{t,i}^m)(1, \bar{\psi}_{i,m})| \} \\
&\quad \le \sum_{(x_i,t_m)\in\Omega_0} KG_i^m(1, \bar{\psi}_{i,m}) \left\{ \int_{x_{i-1}}^{x_{i+1}} |\theta_x^m| \, dx + \int_{t_{m-1}}^{t_m} |u_{tt}(x_i, t)| \, dt \right\} \qquad (3.2.71) \\
&\quad \le C(H+K), \qquad (3.2.72)
\end{aligned}
$$

by (3.2.69) and Lemma 3.2.6.

To bound the other sum, we rewrite (3.2.19) as

$$
\begin{aligned}
\bar{R}(u^m, v^m) = {} & ((\bar{a}^m - a^m)u_x^m, v^m) + (f^m - b^m u^m, v^m) \\
& - (f^m - b^m u^m, v^m)_{d^*} + (r^m \partial u^m, v^m)_{d^*} - (r^m u_t^m, v^m),
\end{aligned}
$$

so that, using $\bar{a}^m \equiv a^m \equiv 1$,

$$
\begin{aligned}
\left| \bar{R}(u^m, \bar{\psi}_{i,m}) \right| \le C \{ & H \|f^m\|_{L^\infty(\Delta_i)} + HK^{-1}(\|u^m\|_{L^\infty(\Delta_i)} \\
& + \|u^{m-1}\|_{L^\infty(\Delta_i)}) + \|u_t^m\|_{L^1(\Delta_i)} \},
\end{aligned}
$$

where $\Delta_i = (x_{i-1}, x_{i+1})$. Thus

$$\begin{aligned}
&\sum_{(x_i,t_m)\in\Omega\backslash\Omega_0} KG_i^m \left|\bar{R}(u^m, \bar{\psi}_{i,m})\right| \\
&\quad \le C\left(\max_{(x_i,t_m)\in\Omega\backslash\Omega_0} G_i^m\right)\{\|f\|_{L^1(L^\infty(\Omega_K))} + K^{-1}(\|u\|_{L^1(L^\infty(\Omega_K))} \\
&\qquad\qquad + \|u^0\|_{L^\infty(0,1)}) + \|u_t\|_{L^1(L^1(\Omega_K))}\} \\
&\quad \le CH, \qquad (3.2.73)
\end{aligned}$$

using Lemma 3.2.7 (with $s = 1$) and the assumption (3.2.68).

Combining (3.2.70) with (3.2.72) and (3.2.73) completes the proof. □

Remark 3.2.4 The assumption that $K\|u_t\|_{L^1(L^1(\Omega_K))} \le C$ is reasonable in many cases (see, e.g., Johnson [21] p.147 – 149). In fact, an inspection of the proof of Theorem 3.2.5 shows that in (3.2.68) one can replace C by $CK^{-\delta}$ for any fixed positive constant δ without affecting the conclusion of this theorem.

We note that the assumption (3.2.69) implies that (x_{i_0}, t_{m_0}) is outside the boundary layer. Theorem 3.2.5 gives a pointwise estimate of $O(H+K)$, in regions where the solution is smooth. To get a local pointwise error estimate inside the layer, we need the following technical result.

Lemma 3.2.8 *Set* $W(x) = \exp(\rho/2)\exp(-(1-x)/(2\varepsilon))$ *with* $\rho = H/\varepsilon$. *Then*

$$(-\varepsilon W_{xx} + W_x, \bar{\psi}_{i,m}) \ge (1/16)\int_{x_{i-1}}^{x_{i+1}} \varepsilon^{-1}\exp(-(1-x)/\varepsilon)\,dx,$$

for $i = 1, \ldots, N-1$ *and* $m = 0, \ldots, M$.

Proof. We have

$$(-\varepsilon W_{xx} + W_x, \bar{\psi}_{i,m})$$

$$= (1/4)\exp(\rho/2)\,(\varepsilon^{-1}\exp(-(1-x)/(2\varepsilon)), \bar{\psi}_{i,m})$$
$$\geq (1/4)\exp(\rho/2)\int_{x_{i-1}}^{x_i} \varepsilon^{-1}\exp(-(1-x)/(2\varepsilon))\,\{1-\exp(-(x-x_{i-1})/\varepsilon)\}$$
$$\times\{1-\exp(-\rho)\}^{-1}\,dx, \quad \text{from (3.2.5) and (3.2.6),}$$
$$= (1/2)\exp(-(1-x_{i+1})/(2\varepsilon))(1-\exp(-\rho))/(1+\exp(-\rho/2))^2$$
$$\geq (1/16)\int_{x_{i-1}}^{x_{i+1}} \varepsilon^{-1}\exp(-(1-x)/(2\varepsilon))\,dx$$
$$\geq (1/16)\int_{x_{i-1}}^{x_{i+1}} \varepsilon^{-1}\exp(-(1-x)/\varepsilon)\,dx,$$

as desired. □

We can now give a pointwise error estimate inside the layer, under a realistic local assumption on the behaviour of derivatives of the solution in the layer.

Theorem 3.2.6 *Assume that the exact solution $u(x,t)$ of (1.1.1) - (1.1.5) satisfies (3.2.68) and*

$$\left|D_x^i D_t^j u(x,t)\right| + \left|D_x^i f(x,t)\right|$$
$$\leq C\{1+\varepsilon^{-i}\exp(-(1-x)/\varepsilon)\} \quad \forall (x,t)\in\Omega_0^+, \qquad (3.2.74)$$

for $0\leq i\leq 1$ and $i+j\leq 2$, where Ω_0^+ is as in Theorem 3.2.5. Then

$$|(U-u)(x_{i_0}, t_{m_0})| \leq C(H+K).$$

Proof. Note that (3.2.73) is valid under the assumption (3.2.68). Hence from (3.2.70), (3.2.71), (3.2.73) and (3.2.74),

$$|(U-u)(x_{i_0}, t_{m_0})|$$
$$\leq C\sum_{(x_i,t_m)\in\Omega_0} K(1,\bar{\psi}_{i,m})G_i^m\int_{x_{i-1}}^{x_{i+1}} \varepsilon^{-1}\exp(-(1-x)/\varepsilon)\,dx + C(H+K)$$

$$\leq C\sum_{m=1}^{M}K\sum_{i=1}^{N-1}(1,\bar{\psi}_{i,m})G_i^m(-\varepsilon W_{ss}+W_s,\bar{\psi}_{i,m})+C(H+K),$$

by Lemma 3.2.8,

$$\leq CH\sum_{m=1}^{M}K(-\varepsilon W_{ss}+W_s,G^m)+C(H+K),$$

since $(1,\bar{\psi}_{i,m})\leq CH$,

$$=CH\sum_{m=1}^{M}K\{\varepsilon(W_s,G_s^m)+(W_s,G^m)\}+C(H+K)$$

$$\leq CH\sum_{m=1}^{M}K\{\varepsilon((W_{S_m^0})_s,G_s^m)+((W_{S_m^0})_s,G^m)\}+C(H+K),$$

since $(W-W_{S_m^0},-\varepsilon G_{ss}^m-G_s^m)\hat{}=0$,

$G_s(x_0+0)\geq 0$ and $G_s(x_N-0)\leq 0$,

$$=CH\sum_{m=1}^{M}K\{K^{-1}\delta_{m,m_0}W(x_{i_0})-K^{-1}(W_{S_m^0},G^m-G^{m+1})_{d^*}$$

$$-(W_{S_m^0},G^m)_{d^*}\}+C(H+K),\quad \text{by (3.2.31)},$$

$$=CH\left\{W(x_{i_0})-(W,G^1)_{d^*}+\sum_{m=1}^{M}K(W,G^m)_{d^*}\right\}+C(H+K),$$

by telescoping,

$$\leq CH\max_{1\leq i\leq N-1}W(x_i)+C(H+K),\ \text{using Lemma 3.2.6 and (3.2.34)},$$

$$\leq C(H+K),$$

since $\max_{1\leq i\leq N-1}W(x_i)\leq 1$. This completes the proof. □

Remark 3.2.5 The analysis in this subsection is carried out for a constant coefficient problem, but the conclusions are valid for variable coefficient problems, provided that one also assumes that $\|u_s\|_{L^1(L^1(\Omega_K))}\leq C$.

3.3 $\tilde{L}$-Spline Trial and Hat Test Functions

In this section we choose the $\tilde{L}$-spline functions $\{\tilde{\phi}_{i,m}(x)\}$ defined in (3.2.23) and (3.2.24) to be the trial functions and piecewise linear "hat" functions $\{\psi_{i,m}(x)\}$ to be the test functions. A scheme of this type was considered in Stynes and O'Riordan [47] for two-point boundary value problems. Let $\tilde{S}_m$ and V_m be the linear spans of $\{\tilde{\phi}_{i,m}(x)\}$ and $\{\psi_{i,m}(x)\}$ respectively. We work with an arbitrary mesh (no grading requirement). Because of the new choice of the test functions, we use a discrete L^2 inner product which is different from $(v, w)_{d^*}$: set

$$(v, w)_d = \sum_{i=1}^{N-1} (1, \psi_{i,m}) v(x_i, t_m) w(x_i, t_m) \quad \forall v, w \in C([0,1] \times \{t_m\}). \tag{3.3.1}$$

The discrete L^2 and energy norms are correspondingly changed into

$$\|w\|_d = (w, w)_d^{1/2} \quad \text{and} \quad |||w|||_d = (\varepsilon \|w_x\|^2 + \|w\|_d^2)^{1/2}. \tag{3.3.2}$$

The scheme considered here has a form similar to the one studied in the previous section, with the approximation $\bar{a}$ replaced by $\tilde{a}$, defined in (3.2.3). It is formulated as follows: find $U^m \in \tilde{S}_m$ such that

$$\tilde{B}(U^m, v^m) + (r^m \partial U^m, v^m)_d = (f^m, v^m)_d \quad \forall v^m \in V_m, \tag{3.3.3}$$

$$U^m(0) = q_0(t_m) \text{ and } U^m(1) = q_1(t_m), \text{ for } m = 1, \ldots, M, \tag{3.3.4}$$

$$U^0(x) = (u^0)_{\tilde{S}_0}, \tag{3.3.5}$$

where

$$\tilde{B}(w, z) = \varepsilon(w_x, z_x) + (\tilde{a} w_x, z) + (bw, z)_d. \tag{3.3.6}$$

By an argument similar to that for Lemma 3.2.1 , we have

Lemma 3.3.1 *For each $v^m \in \tilde{S}_m$, $m \in \{1, \ldots, M\}$ and H sufficiently small (independently of ε),*

$$\tilde{B}(v^m, (v^m)_{V_m}) + (r^m \partial v^m, (v^m)_{V_m})_d$$
$$\geq C_1 |||v^m|||_d^2 + \frac{1}{2k_m}\{(r^m v^m, v^m)_d - (r^m v^{m-1}, v^{m-1})_d\},$$

where C_1 is as in (1.1.5).

This lemma guarantees that (3.3.3) – (3.3.5) have a unique solution U^m in $\tilde{S}_m$.

To analyze the scheme (3.3.3) – (3.3.5), let us introduce a set of $\tilde{L}^*$-spline functions $\{\tilde{\psi}_{i,m}(x)\}$, which are defined similarly to (3.2.5) and (3.2.6) except that $\bar{a}_i^m$ is replaced by $\tilde{a}_i^m$. For any $v^m \in V_m$, set $\tilde{v}^m = \sum_{i=1}^{N-1} v_i^m \tilde{\psi}_{i,m}(x)$. Then integrating by parts and using the definitions of $\tilde{\phi}_{i,m}(x)$ and $\tilde{\psi}_{i,m}(x)$, one can prove that

$$\tilde{B}(u^I(\cdot, t_m), v^m) = \tilde{B}(u^m, \tilde{v}^m), \tag{3.3.7}$$

where $u^I(\cdot, t_m)$ is the interpolant from $\tilde{S}_m$ to $u(x, t_m)$.

Set $Z^m = u^I(\cdot, t_m) - U^m$. Then (2.1.1), (1.1.2), (1.1.3), (3.3.3) – (3.3.5) and (3.3.7) yield

$$\tilde{B}(Z^m, v^m) + (r^m \partial Z^m, v^m)_d = \tilde{R}(u^m, v^m) \quad \forall v^m \in V_m, \tag{3.3.8}$$

$$Z^m(0) = Z^m(1) = 0 \text{ and } Z^0 = 0, \tag{3.3.9}$$

where

$$\tilde{R}(u^m, v^m) = ((\tilde{a}^m - a^m)u_x^m, \tilde{v}^m) + \{(\theta^m, \tilde{v}^m) - (\theta^m, v^m)_d\}$$
$$+ (r^m(\partial u^m - u_t^m), v^m)_d. \tag{3.3.10}$$

Here $\theta = f - bu - ru_t$ as in the last section.

To bound $\tilde{R}(u^m, v^m)$, we need a technical result given in Stynes and O'Riordan [47].

Lemma 3.3.2 *For each $m \in \{1, \ldots, M\}$ and any $w^m \in \tilde{S}_m$,*

$$\int_{x_{i-1}}^{x_i} |w_x^m(x)|\, dx \le C\varepsilon^{1/2} \left\{ \int_{x_{i-1}}^{x_i} |w_x^m(x)|^2\, dx \right\}^{1/2}, \quad \textit{for } i = 1, \ldots, N.$$

With this we can prove

Lemma 3.3.3 *For each $m \in \{1, \ldots, M\}$ and any $v^m \in V_m$,*

$$|\tilde{R}(u^m, v^m)| \le CH \left\{ \|u_x^m\|_{L^1(0,1)}^2 + \|\theta_x^m\|_{L^1(0,1)}^2 + \|\theta^m\|_{L^2(0,1)}^2 \right\}$$
$$+ C\|\partial u^m - u_t^m\|_d^2 + (C_1/2)|||v_{\tilde{S}_m}^m|||_d^2.$$

where C_1 is as in (1.1.5).

Proof. From (3.3.10), we have

$$|\tilde{R}(u^m, v^m)| \le |((\tilde{a}^m - a^m)u_x^m, \tilde{v}^m)| + |(\theta^m, v^m) - (\theta^m, v^m)_d|$$
$$+ |(r^m(\partial u^m - u_t^m), v^m)_d| + |(\theta^m, \tilde{v}^m - v^m)|. \qquad (3.3.11)$$

The first three terms can be bounded similarly to (3.2.20) - (3.2.22). For the last term, if $x \in [x_{i-1}, x_i]$ for some $i \in \{1, \ldots, N\}$, then

$$\begin{aligned} &|\tilde{v}^m(x) - v^m(x)| \\ &= |(v_{i-1}^m - v_i^m)(\tilde{\psi}_{i,m}(x) - \psi_{i,m}(x))| \\ &\le 2\int_{x_{i-1}}^{x_i} \left|\left(v_{\tilde{S}_m}^m\right)_x\right| dx, \quad \text{since } |\tilde{\psi}_{i,m}(x)| + |\psi_{i,m}(x)| \le 2, \\ &\le C\varepsilon^{1/2} \left\{ \int_{x_{i-1}}^{x_i} \left|\left(v_{\tilde{S}_m}^m\right)_x\right|^2 dx \right\}^{1/2}, \end{aligned}$$

by Lemma 3.3.2. Thus

$$\|\tilde{v}^m - v^m\| \le CH^{1/2}\varepsilon^{1/2} \left\| \left(v^m_{\hat{S}_m} \right)_s \right\|.$$

Consequently

$$|(\theta^m, \tilde{v}^m - v^m)| \le (C_1/6)\varepsilon \left\| \left(v^m_{\hat{S}_m} \right)_s \right\|^2 + CH\|\theta^m\|^2_{L^2(0,1)}.$$

Combining this and the bounds for the first three terms (cf. (3.2.20) – (3.2.22)) completes the proof. □

Using Lemmas 3.3.1 and 3.3.3, we now have an analogue of Theorem 3.2.2.

Theorem 3.3.1 *Let U^n be defined by (3.3.3) – (3.3.5). For H sufficiently small (independently of ε) and each $n \in \{1, \ldots, M\}$,*

$$\begin{aligned}
&\|U^n - u^n\|_d + \left\{ \sum_{m=1}^{n} k_m |||U^m - u^m|||_d^2 \right\}^{1/2} \\
&\quad \le CH^{1/2} \left\{ \|u_s\|_{L^2(L^1(\Omega_K))} + \|\theta_s\|_{L^2(L^1(\Omega_K))} + \|\theta\|_{L^2(L^\infty(\Omega_K))} \right\} \\
&\qquad + CK^{1/2} \|u_{tt}\|^*_{L^2(L^1(\Omega_H))},
\end{aligned}$$

where the norms on the right hand side are defined in (2.3.3) – (2.3.5).

Proof. Similar to the proof of Theorem 3.2.2. □

Remark 3.3.1 It can be shown that results analogous to Theorem 3.3.1 (except that no u_{tt} term is present) hold for the semi-discrete version of (3.3.3) – (3.3.5).

In what follows, we shall give a pointwise error analysis for the scheme on a uniform mesh with space mesh size H and time mesh size K. For simplicity, we assume, as in Subsection 3.2.4, that $a(x,t) = b(x,t) = r(x,t) = 1$.

For any mesh point$(x_{i_0}, t_{m_0}) \in \Omega$, we define a discrete Green's function $G(x,t)$ associated with (3.3.3) – (3.3.5), similarly to (3.2.31) – (3.2.33): for $m = 0, \ldots, M$,

$$\tilde{B}(\chi^m, G^m) - (\chi^m, \partial G^{m+1})_d = K^{-1}\delta_{m,m_0}\chi^m(x_{i_0}), \quad \forall \chi^m \in \tilde{S}_m^0, \tag{3.3.12}$$

$$G^m(0) = G^m(1) = 0, \quad G^{M+1} = 0, \tag{3.3.13}$$

where $\tilde{S}_m^0 = \left\{\chi^m \in \tilde{S}_m : \chi^m(0) = \chi^m(1) = 0\right\}$. Then with arguments similar to before we can show that Lemmas 3.2.5 and 3.2.6 are valid for this $G(x,t)$. Lemma 3.2.7 still holds; the main difference in the proof is that instead of I_m given by (3.2.53), here we work with

$$I'_m = \varepsilon\|\omega^{-1/2}(G^m_{\bar{S}_m})_s\|^2 + \|(\omega^m)^{-1/2}G^m\|_d^2 + \frac{1}{2}\left\|\left(\frac{1}{\omega^m}\right)_\beta^{1/2} G^m\right\|^2.$$

We now can derive the following local pointwise error estimate.

Theorem 3.3.2 *Assume that the hypotheses of Theorem 3.2.5 or Theorem 3.2.6 hold. If U and u are the solutions of (3.3.3) – (3.3.5) and (1.1.1) – (1.1.5) respectively, then*

$$|U(x_{i_0}, t_{m_0}) - u(x_{i_0}, t_{m_0})| \le C(H + K).$$

Proof. The error at (x_{i_0}, t_{m_0}) can (cf. proof of Theorem 3.2.5) be bounded by

$$|U(x_{i_0}, t_{m_0}) - u(x_{i_0}, t_{m_0})| \le \sum_{m=1}^{M}\sum_{i=1}^{N-1} KG_i^m \left|\tilde{R}(u^m, \tilde{\psi}_{i,m})\right|. \tag{3.3.14}$$

Analogously to (3.2.73), we have

$$\sum_{(x_i, t_m)\in\Omega\setminus\Omega_0} KG_i^m \left|\tilde{R}(u^m, \tilde{\psi}_{i,m})\right| \le CH. \tag{3.3.15}$$

By (3.3.10),

$$
\begin{aligned}
&\left|\tilde{R}(u^m, \tilde{\psi}_{i,m})\right| \\
&\quad \le \{|(\theta^m, \tilde{\psi}_{i,m}) - \theta_i^m(1, \psi_{i,m})| + |(\partial u_i^m - u_{t,i}^m)(1, \psi_{i,m})|\} \\
&\quad \le \{|(\theta^m - \theta_i^m, \tilde{\psi}_{i,m})| + CH^2|\theta_i^m| + |(\partial u_i^m - u_{t,i}^m)(1, \psi_{i,m})|\},
\end{aligned}
$$

since

$$
|(1, \tilde{\psi}_{i,m} - \psi_{i,m})| \le CH^2
$$

follows from the uniformity of the mesh in the x-direction. Hence

$$
\begin{aligned}
&\sum_{(x_i,t_m)\in\Omega_0} KG_i^m \left|\tilde{R}(u^m, \tilde{\psi}_{i,m})\right| \\
&\le \sum_{(x_i,t_m)\in\Omega_0} KG_i^m H \left\{ \int_{x_{i-1}}^{x_{i+1}} |\theta_x^m|\, dx + \int_{t_{m-1}}^{t_m} |u_{tt}(x_i, t)|\, dt + CH|\theta_i^m| \right\} \\
&\le C(H + K), \qquad (3.3.16)
\end{aligned}
$$

similarly to the proof of Theorem 3.2.5 (or Theorem 3.2.6).

The desired result then follows from (3.3.14) – (3.3.16). □

Remark 3.3.2 The previous uniform convergence result also holds for variable coefficient problems, provided that one also assumes that $\|u_x\|_{L^1(L^1(\Omega_K))} \le C$. Theorem 3.3.2 indicates that on a uniform mesh in the x-direction, the scheme (3.3.3) – (3.3.5) satisfies the same error bounds as the scheme (3.2.9) – (3.2.11). However, it is not clear how to get the same convergence results for the scheme (3.3.3) – (3.3.5) on arbitrary meshes.

3.4 Numerical Results

In this final section, we shall present the results of some numerical experiments and compare the actual performance with the theoretical predictions for the schemes discussed above. We solved two problems for various values of ε, H and K on uniform meshes. In each experiment we held the ratio K/H equal to 1. Similar rates of convergence are observed when K/H is some other constant.

All computation was performed in C double precision on an IBM PC.

Example 3.4.1 (global convergence) We examine how both schemes perform when applied to the variable coefficient problem

$$-\varepsilon u_{xx} + (4+2x)u_x + (2+\exp(xt))u + (1+x^2+t^2)u_t = f(x,t) \quad \text{on } \Omega, \quad (3.4.1)$$

with analytical solution

$$u(x,t) = t\exp((x^2+4x-5)/\varepsilon) + x^2 + t^2, \quad (3.4.2)$$

where $\Omega = (0,1)\times(0,1)$. The function $f(x,t)$ and the initial-boundary conditions on $\bar{\Omega}$ were chosen to fit this data. Here $u(x,t)$ exhibits typical boundary-layer behaviour near $x = 1$.

The global discrete L^∞ errors $E_\varepsilon^{(\varepsilon,H)}$ and corresponding convergence rates $p_\varepsilon^{(\varepsilon,H)}$ of the scheme (3.3.3) – (3.3.5) are listed in Tables 3.5.1 and 3.5.2 respectively. These are computed from

$$E_\varepsilon^{(\varepsilon,H)} = \max_{i,m} \left| u(x_i,t_m) - U^{(\varepsilon,H)}(x_i,t_m) \right| \quad (3.4.3)$$

and

$$p_\varepsilon^{(\varepsilon,H)} = \left\{ \ln E_\varepsilon^{(\varepsilon,2H)} - \ln E_\varepsilon^{(\varepsilon,H)} \right\} / \ln 2, \quad (3.4.4)$$

where u is the exact solution and $U^{(\varepsilon,H)}$ is our computed solution with space mesh size H. The rate of uniform convergence is estimated by the so-called "p_ε^+-method" proposed by Farrell and Hegarty [12]. That is,

$$p_\varepsilon^+ = \text{average}_H \; p_\varepsilon^H, \quad (3.4.5)$$

where

$$p_\varepsilon^H = \left\{ \ln E_\varepsilon^{2H} - \ln E_\varepsilon^H \right\} / \ln 2 \tag{3.4.6}$$

and

$$E_\varepsilon^H = \max_\varepsilon E_\varepsilon^{(\varepsilon,H)}. \tag{3.4.7}$$

Table 3.5.1 Global Maximum Errors

ε	N=8	16	32	64	128
1.00000e+00	2.601e-02	1.168e-02	5.463e-03	2.637e-03	1.295e-03
2.50000e-01	6.803e-02	2.536e-02	1.015e-02	4.424e-03	2.046e-03
6.25000e-02	1.106e-01	5.070e-02	1.944e-02	7.272e-03	2.892e-03
1.56250e-02	1.225e-01	6.323e-02	3.042e-02	1.328e-02	5.099e-03
3.90625e-03	1.255e-01	6.637e-02	3.365e-02	1.652e-02	7.774e-03
9.76562e-04	1.263e-01	6.716e-02	3.445e-02	1.734e-02	8.594e-03
2.44141e-04	1.264e-01	6.736e-02	3.465e-02	1.754e-02	8.799e-03
6.10352e-05	1.265e-01	6.741e-02	3.470e-02	1.759e-02	8.850e-03
1.52588e-05	1.265e-01	6.742e-02	3.472e-02	1.761e-02	8.863e-03
3.81470e-06	1.265e-01	6.742e-02	3.472e-02	1.761e-02	8.866e-03
9.53674e-07	1.265e-01	6.742e-02	3.472e-02	1.761e-02	8.867e-03
E_ε^H	1.265e-01	6.742e-02	3.472e-02	1.761e-02	8.867e-03

Table 3.5.2 Global Convergence Rates

ε	N=8	16	32	64	Average
1.00000e+00	1.15	1.10	1.05	1.03	1.08
2.50000e-01	1.42	1.32	1.20	1.11	1.26
6.25000e-02	1.13	1.38	1.42	1.33	1.31
1.56250e-02	0.95	1.06	1.20	1.38	1.15
3.90625e-03	0.92	0.98	1.03	1.09	1.00
9.76562e-04	0.91	0.96	0.99	1.01	0.97
2.44141e-04	0.91	0.96	0.98	1.00	0.96
6.10352e-05	0.91	0.96	0.98	0.99	0.96
1.52588e-05	0.91	0.96	0.98	0.99	0.96
3.81470e-06	0.91	0.96	0.98	0.99	0.96
9.53674e-07	0.91	0.96	0.98	0.99	0.96
p_ε^H	0.91	0.96	0.98	0.99	0.96

From Table 3.5.2 we see that the rates obtained numerically tend to 1 as the number N of nodes increases, and the uniform rate of convergence is $p_\varepsilon^+ = 0.96$. This agrees with the predictions of Theorem 3.3.2.

Similar results were obtained for the scheme (3.2.9) – (3.2.11). As mentioned in Remark 3.3.2, the two schemes have similar error bounds when a uniform mesh is employed.

Example 3.4.2 (local convergence) We now test the local performance of our schemes when applied to the following problem, which has discontinuous initial data:

$$-\varepsilon u_{xx} + (1+x^2)(1+t)u_x + 4u + (1+x)u_t = f(x,t) \text{ on } \Omega, \tag{3.4.8}$$

where

$$u(0,t) = t^3, \quad u(1,t) = (1+t)^3, \quad \text{for } 0 < t \le 1, \tag{3.4.9}$$

$$u(x,0) = \begin{cases} x^3, & \text{when } x \in [0, 0.376), \\ x^3 + \exp(-x/3), & \text{when } x \in [0.376, 1]. \end{cases} \tag{3.4.10}$$

The function $f(x,t)$ is chosen such that the reduced solution $u_0(x,t)$ of (3.4.8) is

$$u_0(x,t) = \begin{cases} (x+t)^3 & \text{on } \Omega_1, \\ (x+t)^3 + \exp(-x/3)\exp(-t/4) & \text{on } \Omega_2, \end{cases}$$

where Ω_1 is defined by

$$\Omega_1 = \left\{ (x,t) \in \Omega : \arctan x + \frac{1}{2}\ln(1+x^2) - \frac{1}{2}(1+t)^2 < C_0 \right\},$$

with

$$C_0 = \arctan 0.376 + \frac{1}{2}\ln(1+(0.376)^2) - \frac{1}{2},$$

and $\Omega_2 = \Omega \setminus \Omega_1$. The solution $u(x,t)$ will have an internal layer along $\partial\Omega_1 \cap \partial\Omega_2$.

We solve this problem using the scheme (3.3.3) – (3.3.5).

Table 3.5.3 Local Maximum Errors

ε	N=8	16	32	64	128
1.00000e+00	2.772e-02	1.495e-02	7.736e-03	3.929e-03	1.983e-03
2.50000e-01	6.448e-02	2.732e-02	1.238e-02	5.857e-03	2.848e-03
6.25000e-02	1.205e-01	4.799e-02	1.812e-02	7.307e-03	3.196e-03
1.56250e-02	1.302e-01	6.736e-02	3.175e-02	1.251e-02	4.683e-03
3.90625e-03	1.302e-01	6.740e-02	3.436e-02	1.732e-02	8.058e-03
9.76562e-04	1.302e-01	6.740e-02	3.435e-02	1.735e-02	8.717e-03
2.44141e-04	1.302e-01	6.740e-02	3.435e-02	1.735e-02	8.715e-03
6.10352e-05	1.302e-01	6.740e-02	3.435e-02	1.735e-02	8.715e-03
1.52588e-05	1.302e-01	6.740e-02	3.435e-02	1.735e-02	8.715e-03
3.81470e-06	1.302e-01	6.740e-02	3.435e-02	1.735e-02	8.715e-03
9.53674e-07	1.302e-01	6.740e-02	3.435e-02	1.735e-02	8.715e-03
E_d^H	1.302e-01	6.740e-02	3.435e-02	1.735e-02	8.715e-03

Table 3.5.4 Local Convergence Rates

ε	N=8	16	32	64	Average
1.00000e+00	0.89	0.95	0.98	0.99	0.95
2.50000e-01	1.24	1.14	1.08	1.04	1.13
6.25000e-02	1.33	1.41	1.31	1.19	1.31
1.56250e-02	0.95	1.09	1.34	1.42	1.20
3.90625e-03	0.95	0.97	0.99	1.10	1.00
9.76562e-04	0.95	0.97	0.99	0.99	0.98
2.44141e-04	0.95	0.97	0.99	0.99	0.98
6.10352e-05	0.95	0.97	0.99	0.99	0.98
1.52588e-05	0.95	0.97	0.99	0.99	0.98
3.81470e-06	0.95	0.97	0.99	0.99	0.98
9.53674e-07	0.95	0.97	0.99	0.99	0.98
p_d^H	0.95	0.97	0.99	0.99	0.98

In Tables 3.5.3 and 3.5.4 we display the local discrete $L^\infty(\Omega')$ errors $E_d^{(\varepsilon,H)}$ and the corresponding rates $p_d^{(\varepsilon,H)}$ of convergence based on the double mesh method, where

$$\Omega' = \{(x,t) : 0 < x \leq 0.5, \quad 0.5 \leq t \leq 1\}.$$

Here

$$E_d^{(\epsilon,H)} = \max_{i,m} |U^{(\epsilon,H)}(x_i, t_m) - U^{(\epsilon,2H)}(x_i, t_m)|$$

and the rate $p_d^{(\epsilon,H)}$ is defined analogously to (3.4.4). We use the "p_d^+-method" (see Farrell and Hegarty [12]) to determine the rate of uniform convergence; the quantities p_d^+ and p_d^H are defined analogously to (3.4.4) – (3.4.7) based on $E_d^{(\epsilon,H)}$.

Note that the solution $u(x,t)$ is smooth in Ω'. The results indicate that the scheme (3.3.3) – (3.3.5) is first order accurate in Ω', as predicted by Theorem 3.3.2.

Similar results were obtained for the scheme (3.2.9) – (3.2.11).

Chapter 4

An Exponentially Fitted Non-lumped Scheme

4.1 Introduction

In this chapter, a scheme generated by a Petrov-Galerkin method is examined using finite element techniques. The scheme is exponentially fitted in the x-direction. Recall from Chapter 2 that any scheme on a uniform mesh must have coefficients based on exponentials if it is to be globally L^∞ convergent, uniformly in ε. In Chapter 3, two similar exponentially fitted schemes have been studied, where the integral involving the time derivative term was approximated by some quadrature rule. This obviously introduces an error. An alternative approach is to integrate this term exactly. This so-called "non-lumped method" has been analysed by Ng-Stynes *et al.* [31] for the case $a(x,t) = a(x)$, using finite difference techniques.

In this chapter, we shall prove that, under assumptions similar to those of Chapter 3 on the behaviour of u and its derivatives, the non-lumped scheme is globally uniformly convergent in an energy norm and L^2 norm. If instead we make weaker global assumptions together with some local assumptions on the behaviour of u and its derivatives, we can then prove that our scheme is locally uniformly convergent in a discrete L^∞ norm both outside *and* inside the boundary layer. Our analysis does not need assumptions as strong as those in [31] on the behaviour of the solution u. Our results are similar to those of Chapter 3, except that a less restrictive condition on the ratio of time mesh size to space mesh size is required here. Furthermore, the analysis here is more complicated.

The chapter is structured as follows: in Section 4.2 we describe our non-lumped scheme and norms used in later analysis. In Section 4.3 global estimates in energy and L^2 norms are derived by finite element techniques. Section 4.4 discusses how the results of Section 4.3 are improved when one has more information on the behaviour of derivatives of u. In Section 4.5 we provide a local convergence analysis, using a discrete Green's function and a cut-off function, based on Niijima's analysis of an elliptic problem [32]. Numerical results are presented in Section 4.6.

4.2 Description of Scheme

Consider a tensor product mesh which is arbitrarily graded in the x-direction and arbitrarily spaced in the t-direction as in Subsection 3.2.1.

On each time level t_m with $m \in \{0, \ldots, M\}$, we define a basis $\{\bar{\phi}_{i,m}(x) : i = 0, \ldots, N\}$ for the trial space $\bar{S}_m$ by

$$-\varepsilon \bar{\phi}''_{i,m}(x) + \bar{a}_i^m \bar{\phi}'_{i,m}(x) = 0, \quad \text{for } x \in [0,1]^\wedge, \tag{4.2.1}$$

$$\bar{\phi}_{i,m}(x_j) = \delta_{i,j}, \quad \text{for } j = 0, \ldots, N, \tag{4.2.2}$$

where $[0,1]^\wedge$ is defined in Section 2.3. Here $\bar{a}_i^m$ is an approximation to $a(x, t_m)$ on $(x_{i-1}, x_i]$ given by (3.2.4). A basis $\{\bar{\psi}_{i,m} : i = 1, \ldots, N-1\}$ for the test space $\bar{V}_m$ is given by (3.2.5) and (3.2.6).

Now we formulate our non-lumped scheme as follows: for each $m \in \{1, \ldots, M\}$, find $U^m \in \bar{S}_m$ such that

$$\bar{B}(U^m, v^m) + (r^m \partial U^m, v^m) = (f^m, v^m)_{d^*} \quad \forall v^m \in \bar{V}_m, \tag{4.2.3}$$

$$U^m(0) = q_0(t_m) \text{ and } U^m(1) = q_1(t_m), \tag{4.2.4}$$

$$U^0 = (u^0)_{\bar{S}_0}, \tag{4.2.5}$$

where $\bar{B}(\cdot,\cdot)$ and $(\cdot,\cdot)_{d^*}$ are as in (3.2.12) and (3.2.8) respectively.

The existence and uniqueness of U^m will follow from Lemma 4.2.2 below.

In addition to the discrete L^2 and energy norms introduced in (3.2.13) and (3.2.14), we will also use the following continuous energy norm:

$$|||w||| = \{\varepsilon \|w_x\|^2 + \|w\|^2\}^{1/2},$$

for each $w \in H^1([0,1] \times \{t_m\})$, where $\|\cdot\|$ is the usual $L^2(0,1)$ norm. A useful relationship between the discrete and usual L^2 norms is given by

Lemma 4.2.1 *Fix $m \in \{1, \ldots, M\}$. Suppose that either $v^m \in \bar{S}_m$ with $v^m(0) = v^m(1) = 0$, or $v^m \in \bar{V}_m$. Then*

$$\|v^m\| \leq C\|v^m\|_{d^*}.$$

Proof. Let $v^m \in \bar{S}_m$, with $v^m(0) = v^m(1) = 0$. Then

$$\|v^m\|^2 = \sum_{i=1}^{N} \int_{x_{i-1}}^{x_i} \left(v_{i-1}^m \bar{\phi}_{i-1,m} + v_i^m \bar{\phi}_{i,m}\right)^2 dx.$$

Note that $0 \leq \bar{\phi}_{i,m} \leq 1$ for all i and $(a+b)^2 \leq 2(a^2 + b^2)$. Consequently

$$\|v^m\|^2 \leq 2 \sum_{i=1}^{N-1} (h_i + h_{i+1})(v_i^m)^2$$

$$\leq 4 \sum_{i=1}^{N-1} h_i (v_i^m)^2, \quad \text{by the mesh grading,}$$
$$\leq 8 \sum_{i=1}^{N-1} (1, \bar{\phi}_{i,m})(v_i^m)^2,$$

since it is easy to show that $(1, \bar{\phi}_{i,m}) \geq h_i/2$.

Similarly one can prove the result for $v^m \in \bar{V}_m$. □

We now demonstrate the following coercivity result.

Lemma 4.2.2 *Let $\rho = \alpha H/\varepsilon$. Assume that H is sufficiently small (independently of ε) and that*

$$\frac{\alpha k}{\nu^* H} > 2\Gamma(\rho), \tag{4.2.6}$$

where

$$\Gamma(\rho) = \big(\rho \coth(\rho/2) - 2\big)/\rho. \tag{4.2.7}$$

Then for each $\chi^m \in \bar{S}_m$, $m \in \{0, \ldots, M\}$, we have

$$\bar{B}(\chi^m, (\chi^m)_{\bar{V}_m}) + (r^m \partial \chi^m, (\chi^m)_{\bar{V}_m})$$
$$\geq (C_1/2) |||\chi^m|||_{d^*}^2 + \frac{1}{2k_m} \{ (r^m \chi^m, \chi^m) - (r^m \chi^{m-1}, \chi^{m-1}) \},$$

where C_1 is as in (1.1.5).

Proof. From (3.2.16), for each $\chi^m \in \bar{S}^m$ and H sufficiently small (independently of ε), we have

$$\bar{B}(\chi^m, (\chi^m)_{\bar{V}_m}) \geq \varepsilon \| ((\chi^m)_{\bar{V}^m})_x \|^2 + C_1 \|\chi^m\|_{d^*}^2.$$

Since for $j = i - 1$ or i,

$$\int_{x_j}^{x_{j+1}} (\bar{\psi}'_{i,m})^2 \, dx = \int_{x_j}^{x_{j+1}} (\bar{\phi}'_{i,m})^2 \, dx, \tag{4.2.8}$$

we have

$$\|((\chi^m)_{\mathcal{V}_m})_x\| = \|\chi_x^m\|. \tag{4.2.9}$$

Hence

$$\bar{B}(\chi^m, (\chi^m)_{\mathcal{V}_m}) \ge \varepsilon\|\chi_x^m\|^2 + C_1\|\chi^m\|_{d^*}^2. \tag{4.2.10}$$

For the other term, we have

$$\begin{aligned}
&(r^m\partial\chi^m, (\chi^m)_{\mathcal{V}_m}) \\
&\quad = k_m^{-1}\{(r^m\chi^m, (\chi^m)_{\mathcal{V}_m}) - (r^m\chi^{m-1}, (\chi^m)_{\mathcal{V}_m})\} \\
&\quad \ge k_m^{-1}\Big\{(r^m\chi^m, (\chi^m)_{\mathcal{V}_m}) - \frac{1}{2}(r^m\chi^{m-1}, \chi^{m-1}) \\
&\qquad\quad -\frac{1}{2}(r^m(\chi^m)_{\mathcal{V}_m}, (\chi^m)_{\mathcal{V}_m})\Big\} \\
&\quad = \frac{1}{2k_m}\{(r^m\chi^m, \chi^m) - (r^m\chi^{m-1}, \chi^{m-1})\} \\
&\qquad\quad -\frac{1}{2k_m}\left(r^m, (\chi^m - (\chi^m)_{\mathcal{V}_m})^2\right) \\
&\quad \ge \frac{1}{2k_m}\{(r^m\chi^m, \chi^m) - (r^m\chi^{m-1}, \chi^{m-1})\} \\
&\qquad\quad -\frac{\nu^*}{2k_m}\|\chi^m - (\chi^m)_{\mathcal{V}_m}\|^2.
\end{aligned} \tag{4.2.11}$$

For $i = 1, \ldots, N-1$, let $\bar{\rho}_i^m = \bar{a}_i^m h_i/\varepsilon$. Then elementary computations show that

$$\begin{aligned}
&\int_{x_{i-1}}^{x_i} (\bar{\phi}_{i,m}(x) - \bar{\psi}_{i,m}(x))^2\,dx \\
&\quad = \frac{2\varepsilon h_i}{\bar{a}_i^m}\left\{\Gamma(\bar{\rho}_i^m) + \frac{\bar{\rho}_i^m}{\sinh\bar{\rho}_i^m} - 1\right\}\int_{x_{i-1}}^{x_i} (\bar{\phi}_{i,m}'(x))^2\,dx \\
&\quad \le \frac{2\varepsilon h_i}{\bar{a}_i^m}\Gamma(\bar{\rho}_i^m)\int_{x_{i-1}}^{x_i} (\bar{\phi}_{i,m}'(x))^2\,dx,
\end{aligned} \tag{4.2.12}$$

where $\Gamma(\cdot)$ is given in (4.2.7). For $z > 0$, write $\Gamma(z)$ as

$$\Gamma(z) = z\Theta(z),$$

with

$$\Theta(z) = \left(z\coth(z/2) - 2\right)/z^2,$$

and set $\rho_i = \alpha h_i/\varepsilon$. Then

$$\Theta(\bar\rho_i^m) \le \Theta(\rho_i),$$

since $\Theta'(z) > 0$ for $z > 0$. Hence

$$\Gamma(\bar\rho_i^m) \le \bar\rho_i^m\Theta(\rho_i) = \frac{\bar a_i^m}{\alpha}\Gamma(\rho_i) \le \frac{\bar a_i^m}{\alpha}\Gamma(\rho),$$

since $\Gamma'(z) > 0$ for $z > 0$. Inserting this into (4.2.12), we have

$$\int_{x_{i-1}}^{x_i} (\bar\phi_{i,m}(x) - \bar\psi_{i,m}(x))^2\,dx \le \frac{2H}{\alpha}\Gamma(\rho)\varepsilon\int_{x_{i-1}}^{x_i} \left(\bar\phi_{i,m}'(x)\right)^2\,dx. \tag{4.2.13}$$

Hence, using $\bar\phi_{i-1,m} + \bar\phi_{i,m} \equiv 1 \equiv \bar\psi_{i-1,m} + \bar\psi_{i,m}$ on each $[x_{i-1}, x_i]$, we get

$$\|\chi^m - (\chi^m)_{\bar V_m}\|^2 \le \frac{2H}{\alpha}\Gamma(\rho)\varepsilon\|\chi_x^m\|^2. \tag{4.2.14}$$

Therefore, by (4.2.11) and (4.2.14),

$$\begin{aligned}&(r^m\partial\chi^m, (\chi^m)_{\bar V_m})\\ &\ge \frac{1}{2k_m}\left\{(r^m\chi^m, \chi^m) - (r^m\chi^{m-1}, \chi^{m-1})\right\} - \frac{H\nu^*}{k\alpha}\Gamma(\rho)\varepsilon\|\chi_x^m\|^2.\end{aligned}$$

Combining this with (4.2.10) and using the assumption (4.2.6) yields the desired result, since $0 < C_1 \le 1$. □

Remark 4.2.1 Condition (4.2.6) is not a serious restriction on the ratio k/H, as $\lim_{\rho\to 0}\Gamma(\rho) = 0$, $\Gamma'(\rho) > 0$ for $\rho > 0$, and $\lim_{\rho\to+\infty}\Gamma(\rho) = 1$. A discussion of a condition similar to (4.2.6) may be found in Ng-Stynes *et al.* [31].

4.3 Global L^2 and Energy Norm Error Estimates

In this section we will derive global L^2 and energy norm error estimates for our non-lumped scheme using finite element techniques.

For each $m \in \{1, \dots, M\}$, let $u^I(x, t_m)$ be the interpolant from $\bar{S}_m$ to the exact solution $u(x, t_m)$ at $\{(x_i, t_m)\}_{i=0}^N$. Recall that U^m is the solution of (4.2.3) – (4.2.5). Set

$$Z^m = u^I(x, t_m) - U^m \quad \text{and} \quad \eta^m = u^I(x, t_m) - u(x, t_m).$$

Then using (2.1.1), (4.2.3) – (4.2.5) and $\eta^m(x_i) = 0$ for all i, we obtain, for any $v^m \in \bar{V}_m$,

$$\bar{B}(Z^m, v^m) + (r^m \partial Z^m, v^m) = \bar{R}(u^m, v^m) + (r^m \partial \eta^m, v^m), \tag{4.3.1}$$

$$Z^m(0) = Z^m(1) = 0, \quad Z^0 = 0, \tag{4.3.2}$$

where

$$\begin{aligned} \bar{R}(u^m, v^m) = {} & ((\bar{a}^m - a^m)u_x^m, v^m) + \{(f^m - b^m u^m, v^m) \\ & - (f^m - b^m u^m, v^m)_{d^*}\} + (r^m(\partial u^m - u_t^m), v^m). \end{aligned} \tag{4.3.3}$$

This has a form similar to equation (3.2.19), except that the last term in (4.3.3) is an ordinary $L^2(0,1)$ inner product instead of a discrete one. Hence one can bound differences of similar terms by arguments similar to those in Chapter 3 and by using Lemma 4.2.1, to get

Lemma 4.3.1 *For each $m \in \{1, \dots, M\}$ and any $v^m \in \bar{V}_m$,*

$$\begin{aligned} |\bar{R}(u^m, v^m)| \le {} & CH \left\{ \|u_x^m\|_{L^1(0,1)}^2 + \|(f^m - b^m u^m)_x\|_{L^1(0,1)}^2 \right\} \\ & + C\|\partial u^m - u_t^m\|^2 + (C_1/8)\|v^m\|_{d^*}^2, \end{aligned}$$

where C_1 is as in (1.1.5).

Next, we bound the other term on the right hand side of (4.3.1).

Lemma 4.3.2 *For each $m \in \{1,\dots,M\}$ and any $v^m \in \bar{V}_m$,*

$$|(r^m \partial\eta^m, v^m)| \le CH\left\{k_m^{-1}\|u_{st}\|^2_{L^2(L^1),m} + \|u_s^m\|^2_{L^1(0,1)}\right\} + (C_1/8)\|v^m\|^2_{d^*},$$

where C_1 is as in (1.1.5), $\|\cdot\|_{L^2(L^1),m}$ is defined in (2.3.8).

Proof. Fix $m \in \{1,\dots,M\}$. For $i = 1,\dots,N$ and $x \in [x_{i-1}, x_i]$,

$$u^I(x,t_m) = u_i^m \bar{\phi}_{i,m}(x) + u_{i-1}^m \bar{\phi}_{i-1,m}(x),$$

and

$$u(x,t_j) = u(x,t_j)\bar{\phi}_{i,m}(x) + u(x,t_j)\bar{\phi}_{i-1,m}(x), \qquad \text{for } j = m-1, m.$$

Hence, for $x_{i-1} \le x \le x_i$,

$$\begin{aligned}\partial\eta^m(x) = k_m^{-1}&\left\{\left(\int_{t_{m-1}}^{t_m}\int_x^{x_i} u_{st}(s,t)\,ds\,dt\right)\bar{\phi}_{i,m-1}(x)\right.\\ &\left.+\left(\int_{t_{m-1}}^{t_m}\int_x^{x_{i-1}} u_{st}(s,t)\,ds\,dt\right)\bar{\phi}_{i-1,m-1}(x)\right\}\\ &+\left(\int_{x_{i-1}}^{x_i} u_s(s,t_m)\,ds\right)\partial\bar{\phi}_{i,m}(x).\end{aligned} \tag{4.3.4}$$

Consequently,

$$\begin{aligned}&|(r^m\partial\eta^m, v^m)|\\ &\le \nu^*\sum_{i=1}^{N-1}|v_i^m|\int_{x_{i-1}}^{x_{i+1}}|\partial\eta^m(x)|\,\bar{\psi}_{i,m}(x)\,dx\\ &\le \nu^*\sum_{i=1}^{N-1}|v_i^m|\left\{\left(k_m^{-1}\int_{t_{m-1}}^{t_m}\int_{x_{i-1}}^{x_{i+1}}|u_{st}(x,t)|\,dx\,dt\right)(1,\bar{\psi}_{i,m})\right.\end{aligned}$$

$$+\left(\int_{x_{i-1}}^{x_{i+1}} |u_x(x,t_m)|\,dx\right)(1,|\partial\bar\phi_{i,m}|)\Bigg\}$$

$$\leq Ck_m^{-1}\int_{t_{m-1}}^{t_m}\left\{\sum_{i=1}^{N-1}|v_i^m|\left(\int_{x_{i-1}}^{x_{i+1}}|u_{xt}(x,t)|\,dx\right)(1,\bar\psi_{i,m})\right\}dt$$

$$+\sum_{i=1}^{N-1}|v_i^m|\left(\int_{x_{i-1}}^{x_{i+1}}|u_x(x,t_m)|\,dx\right)(1,\bar\psi_{i,m}), \tag{4.3.5}$$

since an elementary calculation shows that

$$(1,|\partial\bar\phi_{i,m}|)\leq C(h_i+h_{i+1})\leq C(1,\bar\psi_{i,m}).$$

Using (2.3.9) then completes the proof. □

We can now obtain global L^2 and energy norm estimates for the error between our computed solution and the exact solution.

Theorem 4.3.1 *Assume that the hypotheses of Lemma 4.2.2 hold. Then for each $n\in\{1,\ldots,M\}$, we have*

$$\|U^n-u^n\|+\left\{\sum_{m=1}^{n}k_m|||U^m-u^m|||^2\right\}^{1/2}$$

$$\leq CH^{1/2}\left\{\|u_x\|_{L^2(L^1(\Omega_K))}+\|(f-bu)_x\|_{L^2(L^1(\Omega_K))}+\|u_{xt}\|_{L^2(L^1)}\right.$$

$$\left.+\|f-bu-ru_t\|_{L^2(L^\infty(\Omega_K))}\right\}+CK^{1/2}\|u_{tt}\|^*_{L^2(L^1)}, \tag{4.3.6}$$

where the norms on the right hand side are defined in (2.3.3), (2.3.6) (2.3.4) and (2.3.7).

Proof. Take $v^m=(Z^m)_{\bar V_m}$ in (4.3.1) and use Lemma 4.2.2 with $\chi^m=Z^m$ and Lemmas 4.2.1, 4.3.1 and 4.3.2 with $v^m=(Z^m)_{\bar V_m}$ to get for each m,

$$(C_1/2)|||Z^m|||^2_{d^*}+\frac{1}{k_m}\left\{(r^mZ^m,Z^m)-(r^mZ^{m-1},Z^{m-1})\right\}$$

$$\leq CH\left\{\|u_x^m\|^2_{L^1(0,1)}+\|(f^m-b^mu^m)_x\|^2_{L^1(0,1)}\right.$$

$$\left.+k_m^{-1}\|u_{xt}\|^2_{L^2(L^1),m}\right\}+C\|\partial u^m-u_t^m\|^2.$$

Multiplying this by k_m, and summing from $m = 1$ to $m = n \le M$, we get

$$
\begin{aligned}
&(C_1/2) \sum_{m=1}^{n} k_m |||Z^m|||_{d^*}^2 + \nu \|Z^n\|^2 \\
&\quad \le CH \left\{ \|u_x\|_{L^2(L^1(\Omega_K))}^2 + \|(f - bu)_x\|_{L^2(L^1(\Omega_K))}^2 \right. \\
&\quad\quad \left. + \|u_{xt}\|_{L^2(L^1)}^2 \right\} + CK \left(\|u_{tt}\|_{L^2(L^1)}^* \right)^2 + C \sum_{m=1}^{n-1} k_m \|Z^m\|^2,
\end{aligned}
$$

since

$$
\begin{aligned}
&\sum_{m=1}^{M} k_m \|\partial u^m - u_t^m\|^2 \\
&\quad \le \sum_{m=1}^{M} k_m \int_0^1 \left(\int_{t_{m-1}}^{t_m} |u_{tt}(x,t)| \, dt \right)^2 dx \\
&\quad \le CK \int_0^1 \left(\int_0^T |u_{tt}(x,t)| \, dt \right)^2 dx \\
&\quad \le CK \left(\|u_{tt}\|_{L^2(L^1)}^* \right)^2 .
\end{aligned}
$$

Using a discrete Gronwall's inequality and Lemma 4.2.1 we obtain

$$
\begin{aligned}
&\sum_{m=1}^{n} k_m |||Z^m|||^2 + \|Z^n\|^2 \\
&\quad \le CH \left\{ \|u_x\|_{L^2(L^1(\Omega_K))}^2 + \|(f - bu)_x\|_{L^2(L^1(\Omega_K))}^2 \right. \\
&\quad\quad \left. + \|u_{xt}\|_{L^2(L^1)}^2 \right\} + CK \left(\|u_{tt}\|_{L^2(L^1)}^* \right)^2 .
\end{aligned}
\tag{4.3.7}
$$

Combining this with

$$
|||\eta^m|||^2 \le CH \left\{ \|u_x^m\|_{L^1(0,1)}^2 + \|f^m - b^m u^m - r^m u_t^m\|_{L^\infty(0,1)}^2 \right\}, \tag{4.3.8}
$$

which is similar to inequality (3.2.26), completes the argument. □

Corollary 4.3.1 *If*

$$
\begin{aligned}
&\|u_x\|_{L^2(L^1(\Omega_K))} + \|(f - bu)_x\|_{L^2(L^1(\Omega_K))} + \|u_{xt}\|_{L^2(L^1)} \\
&\quad + \|f - bu - ru_t\|_{L^2(L^\infty(\Omega_K))} + \|u_{tt}\|_{L^2(L^1)}^* \le C,
\end{aligned}
\tag{4.3.9}
$$

then

$$\|U^n - u^n\| + \left\{\sum_{m=1}^{n} k_m |||U^m - u^m|||^2\right\}^{1/2} \le C(H^{1/2} + K^{1/2}).$$

Remark 4.3.1 It can be shown that the factor $H^{1/2}$ in the energy norm bounds of Theorem 4.3.1 and Corollary 4.3.1 is sharp.

4.4 Improved Accuracy Under Extra Conditions

In this section we will show that under certain compatibility assumptions on the data at the corners $(0,0)$ and $(1,0)$ of Ω, the error estimates in the last section can be significantly sharpened. These compatibility conditions are given in Stynes and O'Riordan [45]; using them, one can derive pointwise bounds (3.2.27) on the solution and its derivatives.

Using (3.2.27), we can bound the interpolation error in the L^∞ norm.

Lemma 4.4.1 *(Interpolation error in the L^∞ norm) Assume that (3.2.27) holds. Let $u^I(x,t_m)$ be the interpolant from $\bar{S}^m$ to $u(x,t_m)$ at $\{(x_i,t_m)\}_{i=0}^N$. Then for $x \in [x_{i-1},x_i]$ and each $m \in \{1,\dots,M\}$,*

$$\left|u(x,t_m) - u^I(x,t_m)\right| \le Ch_i.$$

Proof. Fix $i \in \{1,\dots,N\}$ and $m \in \{1,\dots,M\}$. Set $\bar{L}_s Z \equiv -\varepsilon Z_{ss} + \bar{a}_i^m Z_s$. Then the operator $\bar{L}_s$ satisfies a maximum principle on $[x_{i-1},x_i]$, since $\bar{a}_i^m > 0$. On (x_{i-1},x_i) we have

$$
\begin{aligned}
&|\bar{L}_s(u - u^I)(x,t_m)| \\
&\quad = |\bar{L}_s u(x,t_m)|, \quad \text{by (4.2.1)}, \\
&\quad = |f^m(x) + (\bar{a}_i^m - a^m(x))u_s^m - b^m(x)u^m - r^m(x)u_t^m|
\end{aligned}
$$

$$\leq C\left(1 + h_i\varepsilon^{-1}\exp(-\alpha(1-x)/\varepsilon)\right),$$

using (3.2.27).

Consider a barrier function

$$w(x) = (x - x_i) + h_i \exp(-\alpha(1-x)/(2\varepsilon)).$$

Then

$$\bar{L}_s w = \bar{a}_i^m + h_i\varepsilon^{-1}(-\alpha^2/4 + \alpha\bar{a}_i^m/2)\exp(-\alpha(1-x)/(2\varepsilon)).$$

Choose a positive constant C sufficiently large, independent of ε and mesh size , to get

$$\left|\bar{L}_s(u - u^I)(x, t_m)\right| \leq C\bar{L}_s w.$$

Note that $w(x_j) \geq 0 = \left|(u - u^I)(x_j, t_m)\right|$, for $j = i-1$ and i. Applying the maximum principle we obtain

$$\left|(u - u^I)(x, t_m)\right| \leq Cw(x) \leq Ch_i, \quad \text{on } [x_{i-1}, x_i],$$

as desired. □

We now improve the estimates of Theorem 4.3.1 as follows.

Theorem 4.4.1 *Assume that the hypotheses of Theorem 4.3.1 and (3.2.27) hold. Then for each* $n \in \{1, \ldots, M\}$*,*

$$\|U^n - u^n\| \leq C(H + K),$$

$$\left\{\sum_{m=1}^{n} k_m |||U^m - u^m|||^2\right\}^{1/2} \leq C(H^{1/2} + K).$$

Proof. Similarly to the proof of Lemma 3.2.4, we have, for any $m \in \{1, \ldots, M\}$ and $v^m \in \bar{V}_m$,

$$|\bar{R}(u^m, v^m)| \le C(H^2 + K^2) + (C_1/8)|||v^m|||^2_{d^*}. \tag{4.4.1}$$

The term $(r^m \partial\eta^m, v^m)$ can be bounded in the same way. In fact, using (3.2.27) in (4.3.5) we have

$$\begin{aligned} &|(r^m \partial\eta^m, v^m)| \\ &\le C \sum_{i=1}^{N-1} |v_i^m|(1, \psi_{i,m}) \int_{x_{i-1}}^{x_{i+1}} \left\{1 + \varepsilon^{-1} \exp(-\alpha(1-x)/\varepsilon)\right\} dx \\ &\le CH^2 + (C_1/8)|||v^m|||^2_{d^*}, \end{aligned} \tag{4.4.2}$$

(cf. the proof of Lemma 3.2.4). The norm $|||v^m|||_{d^*}$ in (4.4.1) and (4.4.2) can be replaced by $|||v^m_{S_m}|||_{d^*}$, using (4.2.9). Thus, in the same manner as the derivation of (4.3.7) we get from (4.3.1), Lemma 4.2.2, (4.4.1) and (4.4.2),

$$\sum_{m=1}^{n} k_m |||Z^m|||^2 + \|Z^n\|^2 \le C(H^2 + K^2). \tag{4.4.3}$$

By Lemma 4.4.1,

$$\|\eta^m\| \le CH. \tag{4.4.4}$$

Combining this with (4.4.3) and (4.3.8) gives the desired results. □

4.5 Localized Pointwise Error Estimates

In this section we use a variant of Niijima's approach [32] to derive pointwise error estimates under reasonable assumptions on the solution u and its derivatives. For simplicity, we consider a constant coefficient problem with $a(x,t) = a > 0$, $b(x,t) = b > 0$ and $r(x,t) = r > 0$, and we assume that the mesh is uniform with the space and time mesh sizes H and K respectively.

Let (x_{i_0}, t_{m_0}) be an interior mesh point. To estimate the error at this point, we define a discrete Green's function $G(x,t)$ associated with the point (x_{i_0}, t_{m_0}) as follows: for each $m \in \{0, \ldots, M\}$, find $G(x, t_m) \equiv G^m \in \bar{V}_m$ such that

$$\bar{B}(\chi^m, G^m) - (r\chi^m, \partial G^{m+1}) = K^{-1}\delta_{m,m_0}\chi^m(x_{i_0}) \quad \forall \chi^m \in \bar{S}^0_m, \tag{4.5.1}$$

$$G^m(0) = G^m(1) = 0, \tag{4.5.2}$$

where $\bar{S}^0_m = \{\chi^m \in \bar{S}_m : \chi^m(0) = \chi^m(1) = 0\}$, and we formally set

$$G^{M+1}(x) \equiv 0. \tag{4.5.3}$$

The following lemma shows the existence and uniqueness of G^m.

Lemma 4.5.1 *Assume that (4.2.6) holds. Then there exists a unique solution G^m for the equations (4.5.1) – (4.5.3). Furthermore,*

$$G^m_i \geq 0, \text{ for } i = 0, \ldots, N \text{ and } m = 0, 1, \ldots, M. \tag{4.5.4}$$

Proof. Equations (4.5.1) – (4.5.3) may be written as

$$\left(A^m + K^{-1}D^m_1 + J\right)\mathrm{G}_m = K^{-1}\left(\delta^m + D^m_2\mathrm{G}_{m+1}\right) \tag{4.5.5}$$

where A^m, D^m_1 and D^m_2 are $N+1$ by $N+1$ tridiagonal matrices with rows 0 and N identically zero, and for rows $i = 1, \ldots, N-1$ having non-zero entries in columns $i-1, i$ and $i+1$ given by

$$\bar{B}(\bar{\phi}_{i,m}, \bar{\psi}_{i-1,m}), \quad \bar{B}(\bar{\phi}_{i,m}, \bar{\psi}_{i,m}), \quad \bar{B}(\bar{\phi}_{i,m}, \bar{\psi}_{i+1,m});$$

$$r(\bar{\phi}_{i,m}, \bar{\psi}_{i-1,m}), \quad r(\bar{\phi}_{i,m}, \bar{\psi}_{i,m}), \quad r(\bar{\phi}_{i,m}, \bar{\psi}_{i+1,m});$$

$$r(\bar{\phi}_{i,m}, \bar{\psi}_{i-1,m+1}), \quad r(\bar{\phi}_{i,m}, \bar{\psi}_{i,m+1}), \quad r(\bar{\phi}_{i,m}, \bar{\psi}_{i+1,m+1}),$$

respectively; J (which is used solely to incorporate the boundary conditions) is the $N+1$ by $N+1$ matrix with $(0,0)$ and (N,N) entries equal to 1 and all other entries 0; δ^m and $\mathbf{G}_m$ are $N+1$ by 1 matrices with the ith entries

$$\delta_{m,m_0}\delta_{i,i_0}, \quad G_i^m$$

respectively.

Thus it suffices to show that the system (4.5.5) has a unique solution. Inspecting the coefficient matrix of $\mathbf{G}_m$, we see that when (4.2.6) is fulfilled $(A^m + K^{-1}D_1^m + J)$ is an irreducibly diagonally dominant matrix with positive diagonal terms and non-positive off-diagonal terms. Hence it is an M-matrix and so is invertible. That is, the system (4.5.5) can be solved iteratively for $\mathbf{G}_m$ in terms of $\mathbf{G}_{m+1}$ for $m = M, M-1, \ldots, 0$, since $\mathbf{G}_{M+1}$ is known. This completes the proof of the existence and uniqueness.

Since for each m the matrices $(A^m + K^{-1}D_1^m + J)^{-1}$, δ^m and D_2^m are nonnegative, it is straightforward to show (4.5.4) by using induction on m. □

Next, we derive an L^1 estimate on G along mesh-lines parallel to the x-axis.

Lemma 4.5.2 *Assume that (4.2.6) holds. Then for each $n \in \{0, \ldots, M\}$,*

$$(1, G^n)_{d^*} = (1, G^n) \le C.$$

Proof. Taking $\chi^m = \sum_{i=1}^{N-1} \bar{\phi}_{i,m}(x)$ in (4.5.1), multiplying by K, then summing from $m = n \le M$ to $m = M$, we get

$$\sum_{m=n}^{M} K\bar{B}(\chi^m, G^m) + (r\chi^n, G^n) \le 1, \tag{4.5.6}$$

using (4.5.3). Integrating by parts and using (3.2.5), we have

$$\bar{B}(\chi^m, G^m) = \frac{aG_1^m}{1-e^{-\rho}} - \frac{aG_{N-1}^m}{1-e^{\rho}} + (b, G^m)_{d^*}. \tag{4.5.7}$$

From (4.5.4), (4.5.6) and (4.5.7) we deduce that

$$K\left\{\frac{aG_1^n}{1-e^{-\rho}} - \frac{aG_{N-1}^n}{1-e^{\rho}}\right\} + (r\chi^n, G^n) \leq 1. \tag{4.5.8}$$

Now

$$\begin{aligned}
&(\chi^n, G^n) - (1, G^n)\\
&= G_1^n \int_0^{x_1} (\bar{\phi}_{1,n}(x)\bar{\psi}_{1,n}(x) - \bar{\psi}_{1,n}(x))\,dx\\
&\quad + G_{N-1}^n \int_{x_{N-1}}^1 (\bar{\phi}_{N-1,n}(x)\bar{\psi}_{N-1,n}(x) - \bar{\psi}_{N-1,n}(x))\,dx\\
&= G_1^n \left\{\frac{2\rho - e^{\rho} + e^{-\rho}}{\rho(1-e^{\rho})(1-e^{-\rho})} - \frac{\rho - 1 + e^{-\rho}}{\rho(1-e^{-\rho})}\right\} H\\
&\quad + G_{N-1}^n \left\{\frac{2\rho - e^{\rho} + e^{-\rho}}{\rho(1-e^{\rho})(1-e^{-\rho})} - \frac{\rho + 1 - e^{\rho}}{\rho(1-e^{\rho})}\right\} H.
\end{aligned} \tag{4.5.9}$$

Thus (4.5.8) may be written as

$$(r, G^n) + W^n \leq 1 \tag{4.5.10}$$

with

$$W^n = G_1^n W_1^n + G_{N-1}^n W_{N-1}^n, \tag{4.5.11}$$

where

$$W_1^n = \frac{aK}{1-e^{-\rho}} + rH\left\{\frac{2\rho - e^{\rho} + e^{-\rho}}{\rho(1-e^{\rho})(1-e^{-\rho})} - \frac{\rho - 1 + e^{-\rho}}{\rho(1-e^{-\rho})}\right\}$$

and

$$W_{N-1}^n = \frac{aK}{e^{\rho}-1} + rH\left\{\frac{2\rho - e^{\rho} + e^{-\rho}}{\rho(1-e^{\rho})(1-e^{-\rho})} - \frac{\rho + 1 - e^{\rho}}{\rho(1-e^{\rho})}\right\}.$$

By (4.2.6),

$$W_1^n \geq 2rH\frac{\rho\coth(\rho/2)-2}{\rho(1-e^{-\rho})} + rH\left\{\frac{2\rho-e^{\rho}+e^{-\rho}}{\rho(1-e^{\rho})(1-e^{-\rho})} - \frac{\rho-1+e^{-\rho}}{\rho(1-e^{-\rho})}\right\}$$
$$= rH\frac{2+\rho-e^{\rho}+\rho e^{\rho}}{\rho(e^{\rho}-1)(1-e^{-\rho})}. \tag{4.5.12}$$

Set

$$y(\rho) = 2+\rho-2e^{\rho}+\rho e^{\rho}.$$

Then

$$y(0) = 0,$$
$$y'(\rho) = 1-e^{\rho}+\rho e^{\rho}, \quad y'(0)=0,$$
$$y''(\rho) = \rho e^{\rho} > 0, \quad \text{for } \rho > 0.$$

Hence $y(\rho) > 0$ for $\rho > 0$. The denominator in (4.5.12) is obviously positive for $\rho > 0$. Consequently, $W_1^n \geq 0$. Similarly, one can prove that $W_{N-1}^n \geq 0$. Then it follows from (4.5.11) and (4.5.4) that

$$W^n \geq 0. \tag{4.5.13}$$

Putting (4.5.13) into (4.5.10) completes the proof. □

Now define a subdomain Ω_0 associated with (x_{i_0}, t_{m_0}) by

$$\Omega_0^+ = \Big\{(x,t)\in\Omega : 0 < x \leq x_{i_0} + 2K_0\varepsilon^*\ln\left(\frac{1}{HK}\right),$$
$$|rx - at - (rx_{i_0} - at_{m_0})| \leq 2K_0\sqrt{\varepsilon^*}\ln\left(\frac{1}{HK}\right)\Big\}, \tag{4.5.14}$$

with $\varepsilon^* = \max\{\varepsilon, H, K\}$, where $K_0 > 0$ is a constant independent of ε, H and K, which we choose near the end of the proof of Lemma 4.5.3. In Lemma 4.5.3 we will demonstrate that the discrete Green's function G dies off outside a subdomain

Ω_0 of Ω_0^+, where Ω_0 is defined similarly to Ω_0^+, except that $2K_0$ is replaced by K_0. Without loss of generality, we assume that Ω_0 is a mesh domain.

Lemma 4.5.3 *Assume that (4.2.6) holds. Then for any nonnegative integer s, there exists a positive constant $C = C(s)$ such that*

$$\max_{(x,t_m)\in\Omega\setminus\Omega_0} G(x,t_m) \le C(s)(HK)^s$$

for each $m \in \{0,\dots,M\}$.

Proof. We define a cut-off function $\omega(x,t)$ on Ω by

$$\omega(x,t) = \Phi\left(\frac{x-A}{\sigma_x}\right)\Phi\left(\frac{rx-at-P}{\sigma_\eta}\right)\Phi\left(\frac{P-rx+at}{\sigma_\eta}\right) \tag{4.5.15}$$

where $\Phi(\lambda)$ is defined in (3.2.40), and

$$A = x_{i_0},\quad P = rx_{i_0} - at_{m_0},\quad \sigma_x = \gamma\varepsilon^*,\quad \sigma_\eta = \gamma\sqrt{\varepsilon^*}.$$

Here $\gamma > 1$ is some constant (to be specified later) independent of ε, H and K.

Clearly, ω defined in (4.5.15) satisfies (3.2.43) – (3.2.48).

Now we take $\chi^m = (\frac{G^m}{\omega^m})_{S_m}$ in (4.5.1) to get

$$\bar{B}\left(\left(\frac{G^m}{\omega^m}\right)_{S_m}, G^m\right) - \left(\left(\frac{G^m}{\omega^m}\right)_{S_m}, r\partial G^{m+1}\right) = K^{-1}\delta_{m,m_0}\left(\frac{G^m}{\omega^m}\right)(x_{i_0}). \tag{4.5.16}$$

Similarly to (3.2.50), we have

$$\begin{aligned}\bar{B}&\left(\left(\frac{G^m}{\omega^m}\right)_{S_m}, G^m\right)\\ &= \varepsilon\|(\omega^m)^{-1/2}G_x^m\|^2 + b\|(\omega^m)^{-1/2}G^m\|_{d^*}^2\\ &\quad + \frac{a}{2}\left(\left(\frac{1}{\omega^m}\right)_x, (G^m)^2\right) + \varepsilon\left(\left(\frac{1}{\omega^m}\right)_x G^m, G_x^m\right).\end{aligned} \tag{4.5.17}$$

By Cauchy-Schwarz' inequality and (2.3.9),

$$
\begin{aligned}
&-\left(\left(\frac{G^m}{\omega^m}\right)_{\bar{s}_m}, r\partial G^{m+1}\right)\\
&\geq \frac{r}{K}\left\{-\frac{1}{2}\left\|(\omega^{m+1})^{1/2}\left(\frac{G^m}{\omega^m}\right)_{\bar{s}_m}\right\|^2 - \frac{1}{2}\left\|(\omega^{m+1})^{-1/2}G^{m+1}\right\|^2\right.\\
&\qquad\left. + \left(\left(\frac{G^m}{\omega^m}\right)_{\bar{s}_m}, G^m\right)\right\}\\
&= \frac{r}{2K}\left\{\|(\omega^m)^{-1/2}G^m\|^2 - \|(\omega^{m+1})^{-1/2}G^{m+1}\|^2\right\}\\
&\qquad + \frac{r}{2K}\left(\frac{\omega^m-\omega^{m+1}}{K}, \left(\frac{G^m}{\omega^m}\right)^2_{\bar{s}_m}\right)\\
&\qquad - \frac{r}{2K}\left\|(\omega^m)^{1/2}\left(\left(\frac{G^m}{\omega^m}\right)_{\bar{s}_m} - \frac{G^m}{\omega^m}\right)\right\|^2.
\end{aligned}
\tag{4.5.18}
$$

It follows from (4.5.17) and (4.5.18) that

$$
\begin{aligned}
&\bar{B}\left(\left(\frac{G^m}{\omega^m}\right)_{\bar{s}_m}, G^m\right) - \left(\left(\frac{G^m}{\omega^m}\right)_{\bar{s}_m}, r\partial G^{m+1}\right)\\
&\geq \frac{r}{2K}\left\{\|(\omega^m)^{-1/2}G^m\|^2 - \|(\omega^{m+1})^{-1/2}G^{m+1}\|^2\right\} + I_m + Q_m,
\end{aligned}
\tag{4.5.19}
$$

where

$$
I_m = \varepsilon\|(\omega^m)^{-1/2}G^m_x\|^2 + b\|(\omega^m)^{-1/2}G^m\|^2_{d^*} + \frac{1}{2}\left\|\left(\frac{1}{\omega^m}\right)^{1/2}_{\beta} G^m\right\|^2 \tag{4.5.20}
$$

and

$$
\begin{aligned}
Q_m &= \varepsilon\left(\left(\frac{1}{\omega^m}\right)_x G^m, G^m_x\right) + \frac{r}{2}\left(\frac{\omega^m-\omega^{m+1}}{K} + \omega^m_t, \left(\frac{G^m}{\omega^m}\right)^2\right)\\
&\quad + \frac{r}{2}\left(\frac{\omega^m-\omega^{m+1}}{K}, \left(\frac{G^m}{\omega^m}\right)^2_{\bar{s}_m} - \left(\frac{G^m}{\omega^m}\right)^2\right)\\
&\quad - \frac{r}{2K}\left\|(\omega^m)^{1/2}\left(\left(\frac{G^m}{\omega^m}\right)_{\bar{s}_m} - \frac{G^m}{\omega^m}\right)\right\|^2.
\end{aligned}
\tag{4.5.21}
$$

We notice that I_m given in (4.5.20) has the same form as in equation (3.2.53), and the first two terms in Q_m have the same form as the first two terms of equation (3.2.54). Hence, by analogous arguments, one can get

$$\left|\varepsilon\left(\left(\frac{1}{\omega^m}\right)_x G^m, G^m_x\right)\right| + \left|\frac{r}{2}\left(\frac{\omega^m - \omega^{m+1}}{K} + \omega^m_t, \left(\frac{G^m}{\omega^m}\right)^2\right)\right| \le \frac{1}{16} I_m, \quad (4.5.22)$$

on choosing γ sufficiently large, independently of ε, H and K.

Next, we bound the third and fourth terms in (4.5.21). For the fourth term, we have, using $(y-z)^2 \le (y-w)^2 + 2(y-w)(z-w) + (z-w)^2$ and (2.3.9),

$$\frac{r}{2K}\left\|(\omega^m)^{1/2}\left(\left(\frac{G^m}{\omega^m}\right)_{S_m} - \frac{G^m}{\omega^m}\right)\right\|^2$$
$$\le \frac{r}{2K}\left\{\frac{3}{2}\left\|\frac{(G^m)_{S_m} - G^m}{(\omega^m)^{1/2}}\right\|^2 + C\left\|(\omega^m)^{1/2}\left(\frac{(G^m)_{S_m}}{\omega^m} - \left(\frac{G^m}{\omega^m}\right)_{S_m}\right)\right\|^2\right\}. \quad (4.5.23)$$

Set $\Delta_i = [x_{i-1}, x_i]$. Then

$$\left\|\frac{(G^m)_{S_m} - G^m}{(\omega^m)^{1/2}}\right\|^2$$
$$\le \sum_{i=1}^{N}\left(\min_{\Delta_i}\omega^m\right)^{-1}(G^m_i - G^m_{i-1})^2\int_{x_{i-1}}^{x_i}(\bar{\phi}_{i,m}(x) - \bar{\psi}_{i,m}(x))^2\,dx$$
$$\le \sum_{i=1}^{N}\left(\min_{\Delta_i}\omega^m\right)^{-1}(G^m_i - G^m_{i-1})^2\,\frac{2H}{\alpha}\Gamma(\rho)\varepsilon\int_{x_{i-1}}^{x_i}(\bar{\psi}'_{i,m}(x))^2\,dx,$$

by (4.2.13) and (4.2.8),

$$\le \sum_{i=1}^{N}\left(\min_{\Delta_i}\omega^m\right)^{-1}\frac{K}{\nu^*}\varepsilon\int_{x_{i-1}}^{x_i}(G^m_x)^2\,dx, \quad \text{by (4.2.6)},$$
$$= \frac{K}{\nu^*}\varepsilon\|(\omega^m)^{-1/2}G^m_x\|^2$$
$$+ \frac{K}{\nu^*}\sum_{i=1}^{N}\varepsilon\int_{x_{i-1}}^{x_i}\left(\frac{1}{\min_{\Delta_i}\omega^m} - \frac{1}{\omega^m(x)}\right)(G^m_x)^2\,dx$$
$$\le \frac{K}{\nu^*}\varepsilon\|(\omega^m)^{-1/2}G^m_x\|^2$$

$$+\frac{K}{\nu^*}\sum_{i=1}^{N}\frac{\max_{\Delta_i}|\omega_x^m|}{\min_{\Delta_i}\omega^m}H\varepsilon\int_{x_{i-1}}^{x_i}(\omega^m)^{-1}(G_x^m)^2\,dx$$

$$\leq\frac{K}{\nu^*}\left(1+CH(\sigma_x^{-1}+\sigma_\eta^{-1})\right)\varepsilon\|(\omega^m)^{-1/2}G_x^m\|^2, \tag{4.5.24}$$

by (3.2.45) and (3.2.44).

For the other term in (4.5.23),

$$\left\|(\omega^m)^{1/2}\left(\frac{(G^m)_{\mathcal{S}_m}}{\omega^m}-\left(\frac{G^m}{\omega^m}\right)_{\mathcal{S}_m}\right)\right\|^2$$

$$=\sum_{i=1}^{N}\int_{x_{i-1}}^{x_i}\omega^m\left\{\left(\frac{\omega_{i-1}^m-\omega^m(x)}{\omega^m(x)\,\omega_{i-1}^m}\right)G_{i-1}^m\bar{\phi}_{i-1,m}\right.$$

$$\left.+\left(\frac{\omega_i^m-\omega^m(x)}{\omega^m(x)\,\omega_i^m}\right)G_i^m\bar{\phi}_{i,m}\right\}^2\,dx$$

$$\leq\sum_{i=1}^{N}\frac{\max_{\Delta_i}|\omega_x^m|^2}{\min_{\Delta_i}(\omega^m)^2}H^2\int_{x_{i-1}}^{x_i}(\omega^m)^{-1}(G_{i-1}^m\bar{\phi}_{i-1,m}+G_i^m\bar{\phi}_{i,m})^2\,dx,$$

by a Taylor expansion and (4.5.4),

$$\leq C\sum_{i=1}^{N}\frac{\max_{\Delta_i}|\omega_\beta^m|^2+\max_{\Delta_i}|\omega_t|^2}{\min_{\Delta_i}(\omega^m)^2}H^2\int_{x_{i-1}}^{x_i}(\omega^m)^{-1}(G^m)_{\mathcal{S}_m}^2\,dx$$

$$\leq C\sum_{i=1}^{N}\frac{\max_{\Delta_i}|\omega_\beta^m|^2+\max_{\Delta_i}|\omega_t|^2}{\min_{\Delta_i}(\omega^m)^2}H^2$$

$$\times\left\{\int_{x_{i-1}}^{x_i}(\omega^m)^{-1}(G^m)^2\,dx+\int_{x_{i-1}}^{x_i}(\omega^m)^{-1}\left((G^m)_{\mathcal{S}_m}-G^m\right)^2\,dx\right\},$$

using $y^2\leq 2\left(z^2+(y-z)^2\right)$,

$$\leq CH^2\sum_{i=1}^{N}\left\{\left(\frac{\max_{\Delta_i}|\omega_\beta^m|^2}{\min_{\Delta_i}(\omega^m)^2}\right)\left(\frac{\max_{\Delta_i}\omega^m}{\min_{\Delta_i}|\omega_\beta^m|}\right)\int_{x_{i-1}}^{x_i}\left(\frac{1}{\omega^m}\right)_\beta(G^m)^2\,dx\right.$$

$$\left.+2\left(\frac{\max_{\Delta_i}|\omega_t^m|^2}{\min_{\Delta_i}(\omega^m)^3}\right)\int_{x_{i-1}}^{x_i}\left((G_{i-1}^m)^2\bar{\psi}_{i-1,m}+(G_i^m)^2\bar{\psi}_{i,m}\right)\,dx\right\}$$

$$+CH^2(\sigma_x^{-2}+\sigma_\eta^{-2})\left\|(\omega^m)^{-1/2}\left((G^m)_{\mathcal{S}_m}-G^m\right)\right\|^2,$$

using $(G^m(x))^2\leq 2\left((G_{i-1}^m)^2\bar{\psi}_{i-1,m}(x)+(G_i^m)^2\bar{\psi}_{i,m}(x)\right)$,

$$\leq CH^2\sigma_x^{-1}\left\|\left(\frac{1}{\omega^m}\right)_\beta^{1/2}G^m\right\|^2+CH^2\sigma_\eta^{-2}\|(\omega^m)^{-1/2}G^m\|_{d^*}^2.$$

$$+CH^2(\sigma_s^{-2}+\sigma_\eta^{-2})K\left\{1+CH(\sigma_s^{-1}+\sigma_\eta^{-1})\right\}\varepsilon\|(\omega^m)^{-1/2}G_s^m\|^2,$$

by the properties of ω, (4.5.24) and Lemma 4.2.1,

$$\le C\left\{H^2(\gamma\varepsilon^*)^{-1}+K\gamma^{-1}\right\}I_m, \tag{4.5.25}$$

from the expressions for σ_s and σ_η.

Inserting (4.5.24) and (4.5.25) into (4.5.23), we get the following bound for the fourth term in (4.5.21):

$$\frac{r}{2K}\left\|(\omega^m)^{1/2}\left(\left(\frac{G^m}{\omega^m}\right)_{S_m}-\frac{G^m}{\omega^m}\right)\right\|^2 \le \left\{\frac{3}{4}+C\gamma^{-1}\left(\frac{H^2}{K}(\varepsilon^*)^{-1}+1\right)\right\}I_m.$$

From (4.2.6),

$$\frac{H}{K}\le C\frac{\rho}{\rho\coth(\rho/2)-2}\le\begin{cases} C\rho^{-1}, & \text{for } \rho<1,\\ C, & \text{for } \rho\ge 1,\end{cases}$$

so

$$\frac{H^2}{K}\le C\varepsilon^* \quad \text{for all } \rho.$$

Hence

$$\frac{r}{2K}\left\|(\omega^m)^{1/2}\left(\left(\frac{G^m}{\omega^m}\right)_{S_m}-\frac{G^m}{\omega^m}\right)\right\|^2\le\left(\frac{3}{4}+C\gamma^{-1}\right)I_m. \tag{4.5.26}$$

We now turn to bound the third term in (4.5.21). We have

$$\left|\frac{r}{2}\left(\frac{\omega^m-\omega^{m+1}}{K},\left(\frac{G^m}{\omega^m}\right)_{S_m}^2-\left(\frac{G^m}{\omega^m}\right)^2\right)\right|$$

$$\le C\sum_{i=1}^{N}\max_{\Delta_i^m}|\omega_t|\int_{s_{i-1}}^{s_i}\left(\left(\frac{G^m}{\omega^m}\right)_{S_m}-\frac{G^m}{\omega^m}\right)\left(\left(\frac{G^m}{\omega^m}\right)_{S_m}+\frac{G^m}{\omega^m}\right)dx$$

$$\le C\sum_{i=1}^{N}\frac{\max_{\Delta_i^m}|\omega_t|}{\min_{\Delta_i}\omega^m}\left(\int_{s_{i-1}}^{s_i}\omega^m\left(\left(\frac{G^m}{\omega^m}\right)_{S_m}-\frac{G^m}{\omega^m}\right)^2dx\right)^{1/2}$$

$$\times\left(\int_{x_{i-1}}^{x_i}\omega^m\left(\left(\frac{G^m}{\omega^m}\right)_{\bar{x}_m}+\frac{G^m}{\omega^m}\right)^2 dx\right)^{1/2}$$

$$\le C\sigma_\eta^{-1}\left\|(\omega^m)^{1/2}\left(\left(\frac{G^m}{\omega^m}\right)_{\bar{x}_m}-\frac{G^m}{\omega^m}\right)\right\|\left\|(\omega^m)^{1/2}\left(\left(\frac{G^m}{\omega^m}\right)_{\bar{x}_m}+\frac{G^m}{\omega^m}\right)\right\|,$$

by (3.2.45),

$$\le CK^{1/2}\sigma_\eta^{-1}I_m^{1/2}\left\|(\omega^m)^{1/2}\left(\left(\frac{G^m}{\omega^m}\right)_{\bar{x}_m}+\frac{G^m}{\omega^m}\right)\right\|, \tag{4.5.27}$$

by (4.5.26). Next,

$$\left\|(\omega^m)^{1/2}\left(\left(\frac{G^m}{\omega^m}\right)_{\bar{x}_m}+\frac{G^m}{\omega^m}\right)\right\|^2$$

$$\le C\sum_{i=1}^{N}\frac{\max_{\Delta_i}\omega^m}{(\min_{\Delta_i}\omega^m)^2}H\left((G_{i-1}^m)^2+(G_i^m)^2\right)$$

$$\le C\sum_{i=1}^{N}\frac{(\max_{\Delta_i}\omega^m)^2}{(\min_{\Delta_i}\omega^m)^2}H\left((\omega_{i-1}^m)^{-1}(G_{i-1}^m)^2+(\omega_i^m)^{-1}(G_i^m)^2\right)$$

$$\le C\sum_{i=1}^{N-1}H(\omega_i^m)^{-1}(G_i^m)^2, \quad \text{by (3.2.44)},$$

$$\le C\|(\omega^m)^{-1/2}G^m\|_{d^*}^2, \quad \text{since } H\le 2(1,\bar{\psi}_{i,m}),$$

$$\le CI_m. \tag{4.5.28}$$

Substituting (4.5.28) into (4.5.27) and using the definition of σ_η, we obtain

$$\left|\frac{\tau}{2}\left(\frac{\omega^m-\omega^{m+1}}{K},\left(\frac{G^m}{\omega^m}\right)_{\bar{x}_m}^2-\left(\frac{G^m}{\omega^m}\right)^2\right)\right|\le C\gamma^{-1}I_m. \tag{4.5.29}$$

Collecting (4.5.22), (4.5.26) and (4.5.29) into (4.5.21) gives

$$|Q_m|\le\left(\frac{13}{16}+C\gamma^{-1}\right)I_m\le\frac{7}{8}I_m, \tag{4.5.30}$$

by choosing γ sufficiently large, independently of ε, H and K. Consequently, from (4.5.16), (4.5.19) and (4.5.30) we get

$$\frac{1}{8}I_m+\frac{\tau}{2K}\{\|(\omega^m)^{-1/2}G^m\|^2-\|(\omega^{m+1})^{-1/2}G^{m+1}\|^2\}$$

$$\leq K^{-1}\delta_{m,m_0}\left(\frac{G^m}{\omega^m}\right)(x_{i_0}). \tag{4.5.31}$$

It follows from (3.2.47), Lemma 4.5.2 and (4.5.31) that

$$\sum_{m=n}^{M} KI_m + \|(\omega^n)^{-1/2}G^n\|^2 \leq CH^{-1}, \quad \text{for } n = 0,\ldots,M. \tag{4.5.32}$$

Choose $K_0 = \gamma(2s+2)$. Then by (3.2.48),

$$\omega(x,t) \leq C(HK)^{2s+2}, \quad \text{on } \Omega \setminus \Omega_0. \tag{4.5.33}$$

For each $(x,t_m) \in \Omega\setminus\Omega_0$, there exists $i' \in \{1,\ldots,N\}$ such that $x \in [x_{i'-1}, x_{i'}]$. Thus

$$\begin{aligned} G^m(x) &\leq G^m_{i'-1} + G^m_{i'} \\ &\leq C(HK)^{s+1}(\omega^m(x))^{-1/2}(G^m_{i'-1} + G^m_{i'}), \quad \text{using (4.5.33)}, \\ &\leq C(HK)^{s+1}\left((\omega^m_{i'-1})^{-1/2}G^m_{i'-1} + (\omega^m_{i'})^{-1/2}G^m_{i'}\right), \quad \text{by (3.2.44)}, \\ &\leq C(HK)^{s+1/2}K^{1/2}\|(\omega^m)^{-1/2}G^m\|_{d^s}. \end{aligned}$$

Using this and (4.5.32) we get the desired result. □

We are now ready to derive an error estimate at (x_{i_0}, t_{m_0}) under reasonable assumptions on the global and local behaviour of the solution and its derivatives.

Theorem 4.5.1 *Assume that (4.2.6) holds and that the solution $u(x,t)$ of (1.1.1) - (1.1.5) satisfies (3.2.68) and either (3.2.69) or (3.2.74). Then*

$$|U(x_{i_0}, t_{m_0}) - u(x_{i_0}, t_{m_0})| \leq C(H+K).$$

Proof. With the discrete Green's function G, the pointwise error can be expressed as

$$|U(x_{i_0}, t_{m_0}) - u(x_{i_0}, t_{m_0})|$$

$$
\begin{aligned}
&= |Z(x_{i_0}, t_{m_0})| \\
&= \left| \sum_{m=1}^{M} K \left\{ \bar{R}(u^m, G^m) + (r\partial\eta^m, G^m) \right\} \right|, \quad \text{by (4.5.1) and (4.3.1),} \\
&\leq \sum_{m=1}^{M} \sum_{i=1}^{N-1} K G_i^m \left| \bar{R}(u^m, \bar{\psi}_{i,m}) + (r\partial\eta^m, \bar{\psi}_{i,m}) \right|, \qquad (4.5.34)
\end{aligned}
$$

writing G^m as $\sum_{i=1}^{N-1} G_i^m \bar{\psi}_{i,m}$ and using (4.5.4).

Split the sum into two parts:

$$
\sum_{m=1}^{M} \sum_{i=1}^{N-1} = \sum_{(x_i, t_m) \in \Omega_0} + \sum_{(x_i, t_m) \in \Omega \setminus \Omega_0} .
$$

Recall (4.3.3) and use (4.3.5) to get

$$
\begin{aligned}
&\sum_{(x_i, t_m) \in \Omega_0} K G_i^m \left| \bar{R}(u^m, \bar{\psi}_{i,m}) + (r\partial\eta^m, \bar{\psi}_{i,m}) \right| \\
&\quad \leq \sum_{(x_i, t_m) \in \Omega_0} K G_i^m \left\{ \int_{x_{i-1}}^{x_{i+1}} |(f^m - bu^m)_x| \, dx + C \max_{x \in \Delta_i} \int_{t_{m-1}}^{t_m} |u_{tt}(x,t)| \, dt \right. \\
&\qquad \left. + \int_{x_{i-1}}^{x_{i+1}} K^{-1} \int_{t_{m-1}}^{t_m} |u_{xt}(x,t)| \, dt \, dx + \int_{x_{i-1}}^{x_{i+1}} |u_x^m(x)| \, dx \right\}, \qquad (4.5.35)
\end{aligned}
$$

where $\Delta_i = (x_{i-1}, x_{i+1})$.

When (3.2.69) is satisfied, the terms in the brackets can be bounded by $C(H+K)$. Hence by Lemma 4.5.2, (4.5.35) is bounded by $C(H+K)$. If instead (3.2.74) is used, we have

$$
\begin{aligned}
&\sum_{(x_i, t_m) \in \Omega_0} K G_i^m \left| \bar{R}(u^m, \bar{\psi}_{i,m}) + (r\partial\eta^m, \bar{\psi}_{i,m}) \right| \\
&\quad \leq C(H+K) + C \sum_{(x_i, t_m) \in \Omega_0} K G_i^m (1, \bar{\psi}_{i,m}) \int_{x_{i-1}}^{x_{i+1}} \varepsilon^{-1} \exp\left(-\alpha(1-x)/\varepsilon \right) dx \\
&\quad \leq C(H+K), \qquad (4.5.36)
\end{aligned}
$$

(cf. proof of Theorem 3.2.6.)

In order to estimate the sum over $\Omega \setminus \Omega_0$, we rewrite (4.3.3) to get

$$
\begin{aligned}
&\bar{R}(u^m, v^m) + (r\partial\eta^m, v^m) \\
&\quad = ((\bar{a}^m - a^m)u_x^m, v^m) + (f^m - b^m u^m, v^m) - (f^m - b^m u^m, v^m)_{d^*} \\
&\qquad + (r^m \partial u^I(\cdot, t_m), v^m) - (r^m u_t^m, v^m).
\end{aligned}
$$

Thus, using $\bar{a}^m \equiv a^m \equiv a$,

$$
\begin{aligned}
&\left|\bar{R}(u^m, \bar{\psi}_{i,m}) + (r\partial\eta^m, \bar{\psi}_{i,m})\right| \\
&\quad \le C\left\{H\|f^m\|_{L^\infty(\Delta_i)} + HK^{-1}\left(\|u^m\|_{L^\infty(\Delta_i)}\right.\right. \\
&\qquad \left.\left. + \|u^{m-1}\|_{L^\infty(\Delta_i)}\right) + \|u_t^m\|_{L^1(\Delta_i)}\right\}.
\end{aligned}
$$

Consequently, using Lemma 4.5.3 (with $s = 1$) we can bound the sum over $\Omega \setminus \Omega_0$ as follows:

$$
\begin{aligned}
&\sum_{(x_i, t_m) \in \Omega \setminus \Omega_0} KG_i^m \left|\bar{R}(u^m, \bar{\psi}_{i,m}) + (r\partial\eta^m, \bar{\psi}_{i,m})\right| \\
&\quad \le C(HK)\left\{\|f\|_{L^1(L^\infty(\Omega_K))} + K^{-1}\left(\|u\|_{L^1(L^\infty(\Omega_K))}\right.\right. \\
&\qquad \left.\left. + \|u^0\|_{L^\infty(0,1)}\right) + \|u_t\|_{L^1(L^1(\Omega_K))}\right\} \\
&\quad \le CH,
\end{aligned}
\tag{4.5.37}
$$

using (3.2.68).

Collecting (4.5.36) and (4.5.37) into (4.5.35) completes the proof. □

Remark 4.5.1 The global assumption of Theorem 4.5.1 is reasonable in many cases (see Remark 3.2.4).

We note that (3.2.69) implies that Ω_0 is outside the boundary layer, while (3.2.74) permits (x_{i_0}, t_{m_0}) to lie inside the layer. Thus Theorem 4.5.1 gives a pointwise error bound both outside and inside the boundary layer.

Remark 4.5.2 The analysis in this section is carried out for a constant coefficient problem, but the conclusions are valid for variable coefficient problems, provided that one also assumes that $\|u_x\|_{L^1(L^1(\Omega_K))} \leq C$.

4.6 Numerical Results

In this section, we shall present some numerical results for the non-lumped scheme (4.2.3) – (4.2.5). The numerical experiments were conducted on two problems to examine the global and local performance of the scheme. The two problems were solved for various values of ε, H and K on uniform meshes. In each experiment we held the ratio K/H equal to 1. Similar rates of convergence are observed when K/H is some other constant for which the stability condition (4.2.6) is satisfied. The experimental results will be compared with the theoretical predictions of the previous sections.

All computation was performed in C double precision on an IBM PC.

Example 4.6.1 (global convergence) We examine how the scheme performs when applied to the variable coefficient problem

$$-\varepsilon u_{xx} + (1+\sin x)u_x + (\cos x + \exp(t))u + u_t = f(x,t) \quad \text{on } \Omega, \tag{4.6.1}$$

with analytical solution

$$u(x,t) = \exp(t - (1 - x - \cos 1 + \cos x)/\varepsilon) + x^2, \tag{4.6.2}$$

where $\Omega = (0,1)\times(0,1)$. The function $f(x,t)$ and the initial-boundary conditions on $\bar{\Omega}$ were chosen to fit this data. Here $u(x,t)$ exhibits typical boundary-layer behaviour near $x = 1$.

The global discrete L^∞ errors $E_\varepsilon^{(\varepsilon,H)}$ and corresponding convergence rates $p_\varepsilon^{(\varepsilon,H)}$ of the scheme (4.2.3) are listed in Tables 4.6.1 and 4.6.2 respectively. These are computed from (3.4.3) and (3.4.4). The rate of uniform convergence, which is estimated by (3.4.5) – (3.4.7), is given in the last line of Table 4.6.2.

Table 4.6.1 Global Maximum Errors

ε	N=8	16	32	64	128
1.00000e+00	7.367e-03	3.556e-03	1.747e-03	8.660e-04	4.309e-04
2.50000e-01	1.173e-02	3.905e-03	1.437e-03	5.852e-04	2.592e-04
6.25000e-02	3.879e-02	1.152e-02	3.075e-03	8.206e-04	2.349e-04
1.56250e-02	7.264e-02	3.182e-02	1.111e-02	3.117e-03	8.135e-04
3.90625e-03	8.407e-02	4.316e-02	2.040e-02	8.475e-03	2.875e-03
9.76562e-04	8.700e-02	4.613e-02	2.338e-02	1.140e-02	5.249e-03
2.44141e-04	8.774e-02	4.688e-02	2.413e-02	1.215e-02	6.003e-03
6.10352e-05	8.792e-02	4.706e-02	2.432e-02	1.234e-02	6.191e-03
1.52588e-05	8.797e-02	4.711e-02	2.437e-02	1.239e-02	6.239e-03
3.81470e-06	8.798e-02	4.712e-02	2.438e-02	1.240e-02	6.250e-03
9.53674e-07	8.798e-02	4.712e-02	2.438e-02	1.240e-02	6.253e-03
E_ε^H	8.798e-02	4.712e-02	2.438e-02	1.240e-02	6.253e-03

Table 4.6.2 Global Convergence Rates

ε	N=8	16	32	64	Average
1.00000e+00	1.05	1.03	1.01	1.01	1.02
2.50000e-01	1.59	1.44	1.30	1.17	1.38
6.25000e-02	1.75	1.91	1.91	1.80	1.84
1.56250e-02	1.19	1.52	1.83	1.94	1.62
3.90625e-03	0.96	1.08	1.27	1.56	1.22
9.76562e-04	0.92	0.98	1.04	1.12	1.01
2.44141e-04	0.90	0.96	0.99	1.02	0.97
6.10352e-05	0.90	0.95	0.98	1.00	0.96
1.52588e-05	0.90	0.95	0.98	0.99	0.95
3.81470e-06	0.90	0.95	0.98	0.99	0.95
9.53674e-07	0.90	0.95	0.98	0.99	0.95
p_ε^H	0.90	0.95	0.98	0.99	0.95

From Table 4.6.2 we see that the rates obtained numerically tend to 1 as N

increases, and the uniform rate of convergence is $p_\varepsilon^+ = 0.95$. This agrees with the prediction of Theorem 4.5.1.

Example 4.6.2 (local convergence) We now test the local performance of our scheme when applied to Example 3.4.2, which has discontinuous initial data (see (3.4.8) – (3.4.10)).

In Tables 4.6.3 and 4.6.4 we display the local discrete $L^\infty(\Omega')$ errors $E_d^{(\varepsilon,H)}$ and the corresponding rates $p_d^{(\varepsilon,H)}$ of convergence based on the double mesh method, where

$$\Omega' = \{(x,t) : 0 < x \leq 0.5, \quad 0.5 \leq t \leq 1\}.$$

Here

$$E_d^{(\varepsilon,H)} = \max_{i,m} |U^{(\varepsilon,H)}(x_i,t_m) - U^{(\varepsilon,2H)}(x_i,t_m)|$$

and the rate $p_d^{(\varepsilon,H)}$ is defined analogously to (3.4.4). We use the "p_d^+-method" (see Farrell and Hegarty [12]) to determine the rate of uniform convergence; the quantities p_d^+ and p_d^H are defined analogously to (3.4.5) – (3.4.7) based on $E_d^{(\varepsilon,H)}$.

Table 4.6.3 Local Maximum Errors

ε	N=8	16	32	64	128
1.00000e+00	1.812e-02	9.937e-03	5.300e-03	2.381e-03	1.261e-03
2.50000e-01	7.264e-02	2.469e-02	9.258e-03	3.878e-03	1.899e-03
6.25000e-02	1.940e-01	6.501e-02	2.021e-02	6.612e-03	2.404e-03
1.56250e-02	2.480e-01	1.216e-01	5.047e-02	1.683e-02	5.205e-03
3.90625e-03	2.496e-01	1.258e-01	6.307e-02	3.074e-02	1.274e-02
9.76562e-04	2.499e-01	1.260e-01	6.321e-02	3.164e-02	1.583e-02
2.44141e-04	2.500e-01	1.261e-01	6.323e-02	3.166e-02	1.585e-02
6.10352e-05	2.500e-01	1.261e-01	6.324e-02	3.166e-02	1.585e-02
1.52588e-05	2.501e-01	1.261e-01	6.324e-02	3.166e-02	1.585e-02
3.81470e-06	2.501e-01	1.261e-01	6.324e-02	3.167e-02	1.584e-02
9.53674e-07	2.500e-01	1.261e-01	6.325e-02	3.166e-02	1.584e-02
E_d^H	2.501e-01	1.261e-01	6.325e-02	3.167e-02	1.585e-02

Table 4.6.4 Local Convergence Rates

ε	N=8	16	32	64	Average
1.00000e+00	0.87	0.91	1.15	0.92	0.96
2.50000e-01	1.56	1.41	1.26	1.03	1.31
6.25000e-02	1.58	1.69	1.61	1.46	1.58
1.56250e-02	1.03	1.27	1.58	1.69	1.39
3.90625e-03	0.99	1.00	1.04	1.27	1.07
9.76562e-04	0.99	1.00	1.00	1.00	1.00
2.44141e-04	0.99	1.00	1.00	1.00	0.99
6.10352e-05	0.99	1.00	1.00	1.00	0.99
1.52588e-05	0.99	1.00	1.00	1.00	0.99
3.81470e-06	0.99	1.00	1.00	1.00	1.00
9.53674e-07	0.99	1.00	1.00	1.00	1.00
p_d^H	0.99	1.00	1.00	1.00	1.00

Note that the solution $u(x,t)$ is smooth in Ω'. The results indicate that the scheme (4.2.3) – (4.2.5) is first order accurate in Ω', as predicted by Theorem 4.5.1.

Chapter 5

A Streamline Diffusion Scheme on a Shishkin Mesh

5.1 Introduction

The streamline diffusion method is a finite element method introduced in the case of stationary convection-diffusion problems by Hughes and Brooks [16]. Mathematical analyses of the method have been performed by Johnson *et al.* [20, 22] and Niijima [32] for stationary problems. Nävert [29] extended the method to time-dependent convection-diffusion problems and obtained local L^2 error estimates of order $k+1/2$, with piecewise polynomial finite elements of degree k, in smooth regions (i.e., regions away from any layers). However, in the literature there is no previous pointwise convergence result, which is uniform in the diffusion parameter, for the method inside the boundary layer.

In the present paper, we will improve the results just mentioned for the problem:

$$-\varepsilon u_{xx} + au_x + u + u_t = f(x,t) \quad \forall (x,t) \in \Omega, \tag{5.1.1}$$

$$u(0,t) = u(1,t) = 0 \quad \text{for} \quad 0 < t \leq T, \tag{5.1.2}$$

$$u(x,0) = u^0(x) \quad \text{for} \quad 0 \leq x \leq 1, \tag{5.1.3}$$

where $\Omega = (0,1) \times (0,T]$, ε is a small positive parameter, $a > 0$ is a constant, and $u^0 \in L^2[0,1]$, $f \in L^2(\Omega)$. Here for simplicity we have taken the coefficients of the differential equation to be constant.

We shall give pointwise error analyses for the streamline diffusion method both outside and inside the boundary layer at $x = 1$. We obtain convergence, uniformly in ε, at nodes inside the layer by introducing a special piecewise uniform mesh which resolves part of the boundary layer. Our analysis shows that when the streamline diffusion method is combined with this special mesh, it retains its usual accuracy in smooth regions. In the case of piecewise linear finite elements, the pointwise error bound is almost order $5/4$ away from layers and almost order $3/4$ near and inside the boundary layer. The analysis uses techniques of Niijima [32] and of Johnson *et al.* [22], who considered an elliptic problem on a quasiuniform mesh. In contrast we deal here with a parabolic problem on a highly nonuniform mesh, which leads to many differences and complications in our analysis. Indeed, our approach leads to a slight sharpening of Niijima's results; see Remark 5.4.1 below.

The idea of using a piecewise uniform mesh to guarantee accurate numerical results inside the boundary layer is due to Shishkin [41]. His analysis is set in a finite difference framework and in particular seems applicable only to difference schemes which satisfy a discrete maximum principle. It is therefore inappropriate for the streamline diffusion method; an alternative approach, such as that presented here, is needed.

The Shishkin mesh is remarkable in two ways: firstly, it resolves part but not all of the boundary layer, yet still yields convergence which is uniform in ε; secondly, despite the fact that there is an abrupt change in mesh size, this does not destabilize the difference scheme.

An outline of the chapter is as follows: in Section 5.2 we introduce a special piecewise uniform mesh and construct a streamline diffusion scheme on this mesh. Section 5.3 discusses the properties of our finite element space and analyzes the interpolation errors. In Section 5.4 we define a discrete Green's function G associated with the scheme and estimate it. Our main uniform convergence results are given in Section 5.5. Finally, Section 5.6 presents some numerical results.

5.2 Mesh and Scheme

Let N and M be two positive integers, satisfying

$$\max\{N/M, M/N\} \leq C. \tag{5.2.1}$$

We assume that N is even and

$$N > 4. \tag{5.2.2}$$

Let $\lambda \in (0, 1/2)$ denote a mesh transition parameter, which may depend on N and ε, and will be specified in Section 5.3. We write $\Omega = \Omega_1 \cup \Omega_2$ with $\Omega_1 = (0, 1-\lambda) \times (0, T]$ and $\Omega_2 = \Omega \setminus \Omega_1$. Introduce a set of mesh points $\{(x_i, t_j) \in \Omega : i = 0, \ldots, N \text{ and } j = 0, \ldots, M\}$ with

$$x_i = \begin{cases} 2(1-\lambda)N^{-1}i, & \text{for } i = 0, \ldots, N/2, \\ 1 - \lambda + 2\lambda N^{-1}(i - N/2), & \text{for } i = N/2+1, \ldots, N, \end{cases} \tag{5.2.3}$$

and

$$t_j = TM^{-1}j, \quad \text{for } j = 0, \ldots, M. \tag{5.2.4}$$

By drawing lines through these mesh points parallel to the x- and t-axes, Ω_1 and Ω_2 are each partitioned into $MN/2$ rectangles. Divide each rectangle into two triangles by drawing the diagonal of the rectangle which runs from northwest to southeast (here, as is customary, we have taken the x-axis running west to east and the t-axis south to north). This yields a triangulation of Ω_l denoted by Ω_l^N, for $l = 1, 2$. Each of these Ω_l^N is a uniform triangulation by means of right angled triangles τ, with base

$$H_\tau = \begin{cases} 2(1-\lambda)N^{-1} & \text{for } \tau \in \Omega_1^N, \\ 2\lambda N^{-1} & \text{for } \tau \in \Omega_2^N, \end{cases} \tag{5.2.5}$$

and altitude

$$K = TM^{-1} \quad \text{for all } \tau \in \Omega^N, \tag{5.2.6}$$

where $\Omega^N = \Omega_1^N \cup \Omega_2^N$.

Since we are interested in the singularly perturbed case, we assume throughout that

$$\varepsilon \leq N^{-1}. \tag{5.2.7}$$

Our next aim is to formulate a time-stepping procedure for (5.1.1) – (5.1.3) so that the discrete solution can be computed successively on a sequence of time levels.

On each time slab

$$S_j = [0,1] \times (t_{j-1}, t_j] \quad \text{for } j = 1, \ldots, M, \tag{5.2.8}$$

we define a finite element space V_j by

$$\begin{aligned} V_j = \{ v \in C(S_j) :\ & v(0,t) = v(1,t) = 0 \ \ \forall t \in (t_{j-1}, t_j], \\ & v\big|_\tau \text{ is linear } \forall \tau \in \Omega^N \text{ such that } \tau^0 \subseteq S_j \}, \end{aligned} \tag{5.2.9}$$

where $C(S_j)$ denotes the space of continuous functions on S_j and τ^0 is the interior of τ.

We also introduce the streamline derivative w_β for all differentiable functions w by defining

$$w_\beta = aw_x + w_t.$$

We shall apply the streamline diffusion method [20, 29] to the problem (5.1.1) – (5.1.3) successively on each slab S_j, imposing the initial value at $t = t_{j-1}$ weakly and the boundary condition strongly. To this end, we introduce the finite element space on Ω:

$$V \equiv \left\{ v \in L^2(\Omega) : \ v\big|_{S_j} \in V_j \ \text{ for } j = 1, \ldots, M \right\}, \tag{5.2.10}$$

and define, for each $v \in V_j$, for $j = 1, \ldots, M$ and $0 \le x \le 1$,

$$v^+(x, t_{j-1}) = \lim_{s \to +0} v(x, t_{j-1} + s), \tag{5.2.11}$$

$$v^-(x, t_j) = \lim_{s \to -0} v(x, t_j + s). \tag{5.2.12}$$

We also use the notation of (5.2.11) and (5.2.12) for those functions in $C(S_j)$ for which the indicated limits exist.

Notation. For all measurable $D \subseteq \Omega$, set

$$(v, w)_D = \iint_D vw \, dxdt \quad \forall v, w \in L^2(\Omega), \tag{5.2.13}$$

$$\|v\|_D = (v, v)_D^{1/2} \quad \forall v \in L^2(\Omega). \tag{5.2.14}$$

For $j = 1, \ldots, M$, set

$$\Lambda_j(D) = \{(x, t) \in D : \ t = t_j\}, \tag{5.2.15}$$

and define

$$< v, w >_{j,D} = \int_{\Lambda_j(D)} v(x, t_j) w(x, t_j) \, dx \quad \forall v, w \in L^2(\Lambda_j), \tag{5.2.16}$$

$$|v|_{j,D} = < v, v >_{j,D}^{1/2} \quad \forall v \in L^2(\Lambda_j). \tag{5.2.17}$$

When $D = \Omega$, we omit D from the notation.

We now formulate our streamline diffusion method as follows: for $j = 1, \ldots, M$, find $U \in V$ such that

$$\begin{aligned} &\varepsilon(U_s, v_s)_{S_j} + (U_\beta + U, v + \rho v_\beta)_{S_j} + < U^+, v^+ >_{j-1} \\ &= (f, v + \rho v_\beta)_{S_j} + < U^-, v^+ >_{j-1} \quad \forall v \in V_j, \end{aligned} \tag{5.2.18}$$

where we set

$$< U^-, v^+ >_0 = < u^0, v^+ >_0 \quad \forall v \in V_1, \tag{5.2.19}$$

and

$$\rho = \rho(x) = \begin{cases} 2(1-\lambda)N^{-1} & \text{for } x \in (0, 1-\lambda), \\ 0 & \text{otherwise.} \end{cases} \tag{5.2.20}$$

Remark 5.2.1 Scheme (5.2.18) – (5.2.20) is essentially the same as that given in [20, 29]. The only difference is that we take $\rho = 0$ on Ω_2. This is because later λ is chosen quite small, which implies that the mesh in Ω_2 is very fine in the x-direction. Consequently our scheme is not upwinded on Ω_2.

Remark 5.2.2 For each j, (5.2.18) – (5.2.20) is equivalent to a linear system of equations. Since the space V is defined independently on each slab with no continuity requirements from one slab to the next, the solution U will in general have jumps across each time level t_j.

Define

$$\begin{aligned} H^1_\Lambda(\Omega) = \{w \in H^1(\Omega) :\ & w^+(\cdot, t_j) \text{ and } w^-(\cdot, t_j) \text{ exist} \\ & \text{and lie in } L^2(\Lambda_j) \text{ for } j = 0, \ldots, M\}, \end{aligned}$$

where we set $w^-(x,0) = 0$ and $w^+(x,T) = 0$ for $0 \le x \le 1$.

In order to write (5.2.18) – (5.2.20) in a compact form suitable for analysis we introduce the jump $[v]$ of $v \in H^1_\Lambda(\Omega)$ across each time level by defining, for $0 \le x \le 1$,

$$[v](x,t_j) = (v^+ - v^-)(x,t_j), \quad \text{for } j = 0, \ldots, M.$$

By summation of (5.2.18) over all S_j, we get the following discrete analogue of (5.1.1) – (5.1.3): find $U \in V$ such that

$$B(U,v) = (f, v + \rho v_\beta) + < u^0, v^+ >_0 \quad \forall v \in V, \tag{5.2.21}$$

where for all $w, v \in H^1_\Lambda(\Omega)$ we set

$$B(w,v) = \varepsilon(w_s, v_s) + (w_\beta + w, v + \rho v_\beta) + \sum_{j=0}^{M-1} < [w], v^+ >_j . \tag{5.2.22}$$

In our later analysis, we will also use the following expression for $B(\cdot,\cdot)$ which is equivalent to (5.2.22):

$$\begin{aligned} B(w,v) = {} & \varepsilon(w_s, v_s) + (w_\beta, \rho v_\beta) + (w, v) \\ & + (w, (\rho - 1)v_\beta) - \sum_{j=1}^{M} < w^-, [v] >_j . \end{aligned} \tag{5.2.23}$$

This can be obtained by integrating by parts the term (w_β, v) in (5.2.22).

Notation. For all measurable $D \subseteq \Omega$ and for all $v \in H^1_\Lambda(\Omega)$, set

$$|||v|||^2_D \equiv \varepsilon \|v_s\|^2_D + \|\rho^{1/2} v_\beta\|^2_D + \|v\|^2_D + \sum_{j=0}^{M} \|[v]\|^2_{j,D} \,.$$

When $D = \Omega$ we omit D from the above norms.

The following theorem states a stability inequality for (5.2.21) which also guarantees the existence and uniqueness of the discrete solution U.

Theorem 5.2.1 *If U is the solution of (5.2.21), then*

$$|||U|||^2 \le 8\|f\|^2 + 4\|u^0\|^2_{L^2(0,1)}.$$

Proof. From (5.2.22), we have

$$B(U,U) = \varepsilon\|U_x\|^2 + (U_\beta, U) + \|\rho^{1/2}U_\beta\|^2 + \|U\|^2 \\ + (U, \rho U_\beta) + \sum_{j=0}^{M-1} <[U], U^+>_j . \tag{5.2.24}$$

An integration yields

$$\begin{aligned}
&(U_\beta, U) + \sum_{j=0}^{M-1} <[U], U^+>_j \\
&= \frac{1}{2}\sum_{j=1}^{M} \{<U^-, U^->_j - <U^+, U^+>_{j-1}\} \\
&\quad + \sum_{j=0}^{M-1} \{<U^+, U^+>_j - <U^-, U^+>_j\} \\
&= \frac{1}{2}\sum_{j=1}^{M} <U^-, U^->_j + \frac{1}{2}\sum_{j=0}^{M-1} \{<U^+, U^+>_j - 2<U^-, U^+>_j\} \\
&= \frac{1}{2}\sum_{j=0}^{M} |[U]|_j^2 .
\end{aligned} \tag{5.2.25}$$

Cauchy-Schwarz' inequality and the arithmetic-geometric mean inequality give

$$|(U, \rho U_\beta)| \le \frac{1}{2}\left\{\|U\|^2 + \|\rho U_\beta\|^2\right\}, \tag{5.2.26}$$

since $\rho \le 2N^{-1} \le 1$. Thus from (5.2.24) – (5.2.26),

$$B(U,U) \ge \frac{1}{2}|||U|||^2. \tag{5.2.27}$$

On the other hand,

$$|(f, U + \rho U_\beta) + <u^0, U^+>_0|$$

$$\leq 2\|f\|^2 + \frac{1}{4}\left\{\|U\|^2 + \|\rho^{1/2}U_\beta\|^2 + |U^+|_0^2\right\} + \|u^0\|^2_{L^2(0,1)}.$$

(5.2.28)

Taking $v = U$ in (5.2.21), the desired result follows from (5.2.27) and (5.2.28). □

5.3 Properties of V and Interpolation Error

In this section, we shall discuss inverse and interpolation properties of our finite element space V which will be used in the sequel. We also specify the transition parameter λ of the mesh.

First of all, we consider some properties of V.

Lemma 5.3.1 *For any $v \in V$, we have*

(i) for $\tau \in \Omega_1^N$, $1 \leq q \leq p \leq \infty$ and $l = 0, 1$,

$$\|v\|_{W_p^l(\tau)} \leq CN^{l+2(1/q-1/p)}\|v\|_{L_q(\tau)}, \tag{5.3.1}$$

where W_p^l denotes the usual Sobolev space;

(ii) for $\tau \in \Omega_2^N$,

$$\|v_t\|_\tau \leq CN\|v\|_\tau, \tag{5.3.2}$$

$$\|v\|_{L^\infty(\tau)} \leq CN^{1/2}\lambda^{1/2}\|v_s\|_{S\cap\Omega_2}, \tag{5.3.3}$$

where S is the unique slab (see (5.2.8)) containing τ^0;

(iii) for $j = 1, \ldots, M$,

$$\|v_\beta\|_{S_j\cap\Omega_2} \leq (1 + CN\lambda)\|v_s\|_{S_j\cap\Omega_2}. \tag{5.3.4}$$

Proof. The first conclusion of the lemma is a standard inverse inequality, because our assumption (5.2.1) implies that Ω_1^N is a regular triangulation.

Next, (5.3.2) follows by the standard argument of transforming to a reference triangle of unit diameter.

We now turn to prove (5.3.3). For $(x,t) \in \tau \in \Omega_2^N$, let S be the slab containing τ^0. Then since $v(1,t) = 0$ for all t and v_s is constant on each τ in S,

$$\begin{aligned} |v(x,t)| &\leq \int_s^1 |v_s(\xi,t)|\, d\xi \\ &\leq CK^{-1} \iint_{S\cap\Omega_2} |v_s(\xi,t)|\, d\xi dt \\ &\leq CN[\text{meas}(S\cap\Omega_2)]^{1/2} \|v_s\|_{S\cap\Omega_2}, \end{aligned} \tag{5.3.5}$$

by Cauchy-Schwarz' inequality. Noting that

$$\text{meas}(S\cap\Omega_2) \leq CN^{-1}\lambda,$$

we deduce (5.3.3).

It remains to show (5.3.4). From the definition of v_β, it is sufficient to prove that

$$\|v_t\|_{S_j\cap\Omega_2} \leq CN\lambda\|v_s\|_{S_j\cap\Omega_2} \quad \text{for } j = 1,\ldots,M. \tag{5.3.6}$$

In fact, for each τ in $S_j \cap \Omega_2$, by (5.3.2) and (5.3.3),

$$\begin{aligned} \|v_t\|_\tau &\leq CN\|v\|_{L^\infty(\tau)}[\text{meas}(\tau)]^{1/2} \\ &\leq CN[\lambda N^{-1}M^{-1}]^{1/2}N^{1/2}\lambda^{1/2}\|v_s\|_{S_j\cap\Omega_2} \\ &\leq CN^{1/2}\lambda\|v_s\|_{S_j\cap\Omega_2}. \end{aligned}$$

Thus

$$\|v_t\|^2_{S_j\cap\Omega_2} = \sum_{\tau\subseteq S_j\cap\Omega_2} \|v_t\|^2_\tau \leq CN^2\lambda^2\|v_s\|^2_{S_j\cap\Omega_2},$$

since $S_j \cap \Omega_2^N$ contains N triangles. This implies (5.3.6). □

Lemma 5.3.2 *Let $p \in (1,\infty]$. Assume that $w \in W_p^2(\Omega)$. Let $\tau \in \Omega^N$. Let w^I denote the linear function which interpolates to w at the nodes of τ. Then*

$$\|w - w^I\|_{L^p(\tau)} \le C\left\{H_\tau^2\|w_{ss}\|_{L^p(\tau)} + H_\tau K\|w_{st}\|_{L^p(\tau)} + K^2\|w_{tt}\|_{L^p(\tau)}\right\}, \tag{5.3.7}$$

$$\left\|(w - w^I)_s\right\|_{L^p(\tau)} \le C\left\{H_\tau\|w_{ss}\|_{L^p(\tau)} + K\|w_{st}\|_{L^p(\tau)}\right\}, \tag{5.3.8}$$

$$\left\|(w - w^I)_t\right\|_{L^p(\tau)} \le C\left\{H_\tau\|w_{st}\|_{L^p(\tau)} + K\|w_{tt}\|_{L^p(\tau)}\right\}, \tag{5.3.9}$$

$$\left|w - w^I\right|_{j,\tau} \le CK^{-1/2}\left\{H_\tau^2\|w_{ss}\|_\tau + H_\tau K\|w_{st}\|_\tau + K^2\|w_{tt}\|_\tau\right\} \tag{5.3.10}$$

for $j = 0, \ldots, M$.

Proof. Let $\hat\tau$ denote the reference triangle with vertices at $(0,0), (1,0)$ and $(0,1)$. Let F be a one-to-one linear function which maps $\hat\tau$ onto τ. Set

$$\hat w(z) = w(F(z)) \quad \forall z \in \hat\tau.$$

Then it is well known that

$$\left\|\hat w - (\hat w)^I\right\|_{L^p(\hat\tau)} \le C\left\{\|\hat w_{\hat s\hat s}\|_{L^p(\hat\tau)} + \|\hat w_{\hat s\hat t}\|_{L^p(\hat\tau)} + \|\hat w_{\hat t\hat t}\|_{L^p(\hat\tau)}\right\}, \tag{5.3.11}$$

where $(\hat w)^I$ is the linear function which interpolates to $\hat w$ at the vertices of $\hat\tau$ and $\hat s$, $\hat t$ are the variables used in $\hat\tau$. On observing that $(\hat w)^I = w^I \circ F$, $\hat w_{\hat s\hat s} = H_\tau^2 w_{ss} \circ F$, $\hat w_{\hat s\hat t} = H_\tau K w_{st} \circ F$ and $\hat w_{\hat t\hat t} = K^2 w_{tt} \circ F$, transforming (5.3.11) to integrals over τ yields (5.3.7).

Next, using Lemma 2.3 from Křížek [23], we get

$$\left\|\left(\hat w - (\hat w)^I\right)_{\hat s}\right\|_{L^p(\hat\tau)} \le C\left\{\|\hat w_{\hat s\hat s}\|_{L^p(\hat\tau)} + \|\hat w_{\hat s\hat t}\|_{L^p(\hat\tau)}\right\}$$

and

$$\left\|\left(\hat w - (\hat w)^I\right)_{\hat t}\right\|_{L^p(\hat\tau)} \le C\left\{\|\hat w_{\hat s\hat t}\|_{L^p(\hat\tau)} + \|\hat w_{\hat t\hat t}\|_{L^p(\hat\tau)}\right\}.$$

The desired estimates (5.3.8) and (5.3.9) now follow by transforming the above to integrals over τ.

Finally, using the Bramble-Hilbert lemma [6],

$$\left\|\hat{w} - (\hat{w})^I\right\|_{L^2(F^{-1}(\Lambda_j(\tau)))} \leq C\left\{\|\hat{w}_{\hat{s}\hat{s}}\|_{\hat{\tau}} + \|\hat{w}_{\hat{s}\hat{t}}\|_{\hat{\tau}} + \|\hat{w}_{\hat{t}\hat{t}}\|_{\hat{\tau}}\right\}.$$

Again transforming to τ, one obtains (5.3.10). This completes the proof. □

An immediate inference from Lemma 5.3.2 is the following interpolation error results on those τ where the solution is smooth.

Theorem 5.3.1 *For any $\tau \in \Omega^N$, if $\|u\|_{C^2(\tau)} \leq C$, then*

$$\left\|u - u^I\right\|_{L^\infty(\tau)} \leq CN^{-2} \tag{5.3.12}$$

and

$$\left\|\left(u - u^I\right)_s\right\|_{L^\infty(\tau)} + \left\|\left(u - u^I\right)_\beta\right\|_{L^\infty(\tau)} \leq CN^{-1}. \tag{5.3.13}$$

In order to obtain a satisfactory pointwise error bound in our later estimates, we shall require the local L^∞ interpolation error for a solution with typical boundary layer behaviour to be at least first order in Ω_1 and second order in Ω_2. A calculation, based on (5.3.7) and (5.3.23) below, then motivates the choice

$$\lambda = 2\alpha^{-1}\varepsilon \ln N \tag{5.3.14}$$

with the constant α chosen to satisfy $0 < \alpha \leq a$. We shall assume from now on that (5.3.14) holds.

Note that this choice of λ implies (cf. (5.2.5)) that the mesh in the x-direction is very fine in Ω_2 and coarse in Ω_1. Note also that the boundary layer at $x = 1$ is typically of width $O(\varepsilon \ln(1/\varepsilon))$ and in practice one usually has $\varepsilon^{-1} > N$; consequently (cf. (5.3.14)) the mesh resolves only part of the layer.

The following theorem gives interpolation error bounds on each τ where the solution exhibits boundary layer behaviour.

Theorem 5.3.2 *For any $\tau \in \Omega^N$, assume that*

$$\left|\frac{\partial u}{\partial x}(x,t)\right| \le C\left\{1+\varepsilon^{-1}\exp(-a(1-x)/\varepsilon)\right\} \quad \forall (x,t)\in\tau \tag{5.3.15}$$

and that for $i+j\le 2$ we have

$$\left|\frac{\partial^{i+j}u}{\partial x^i \partial t^j}(x,t)\right| \le C\varepsilon^{-i} \quad \forall (x,t)\in\tau. \tag{5.3.16}$$

Then

(i) if $\tau \in \Omega_1^N$, we have

$$\left\|u-u^I\right\|_{L^\infty(\tau)} \le CN^{-1}; \tag{5.3.17}$$

(ii) if $\tau \in \Omega_2^N$, we have

$$\left\|u-u^I\right\|_{L^\infty(\tau)} \le CN^{-2}\ln^2 N \tag{5.3.18}$$

and

$$\varepsilon\left\|\left(u-u^I\right)_x\right\|_{L^\infty(\tau)} \le CN^{-1}\ln N. \tag{5.3.19}$$

Proof. Without loss of generality, we assume that the vertices of τ are (x_{i-1},t_m), (x_i,t_{m-1}) and (x_{i-1},t_{m-1}). Thus on τ, we have

$$\begin{aligned} u^I(x,t) = {} & u(x_{i-1},t_m)\phi_1(x,t) + u(x_i,t_{m-1})\phi_2(x,t) \\ & + u(x_{i-1},t_{m-1})\phi_3(x,t), \end{aligned} \tag{5.3.20}$$

where

$$\begin{cases} \phi_1(x,t) = (t-t_{m-1})/K, \\ \phi_2(x,t) = (x-x_{i-1})/H_\tau, \\ \phi_3(x,t) = (x_i-x)/H_\tau - (t-t_{m-1})/K. \end{cases} \tag{5.3.21}$$

Clearly

$$\sum_{i=1}^{3} \phi_i(x,t) = 1 \quad \text{and} \quad 0 \le \phi_i(x,t) \le 1 \quad \text{on } \tau. \tag{5.3.22}$$

By (5.3.20) and (5.3.22), for $(x,t) \in \tau$ we get

$$\begin{aligned}
&\left|\left(u - u^I\right)(x,t)\right| \\
&\quad = \left|u(x,t)(\phi_1 + \phi_2 + \phi_3)(x,t) - u^I(x,t)\right| \\
&\quad \le |u(x,t) - u(x_{i-1},t_m)| + |u(x,t) - u(x_i,t_{m-1})| + |u(x,t) - u(x_{i-1},t_{m-1})| \\
&\quad \le |u(x,t) - u(x_{i-1},t)| + |u(x_{i-1},t) - u(x_{i-1},t_m)| \\
&\qquad + |u(x,t) - u(x_i,t)| + |u(x_i,t) - u(x_i,t_{m-1})| \\
&\qquad + |u(x,t) - u(x_{i-1},t)| + |u(x_{i-1},t) - u(x_{i-1},t_{m-1})| \\
&\quad \le 3\int_{x_{i-1}}^{x_i} |u_x(\xi,t)|\, d\xi + \int_{t_{m-1}}^{t_m} \left(2|u_t(x_{i-1},s)| + |u_t(x_i,s)|\right) ds \\
&\quad \le C\left\{\int_{x_{i-1}}^{x_i} \left[1 + \varepsilon^{-1}\exp(-a(1-\xi)/\varepsilon)\right]\, d\xi + \int_{t_{m-1}}^{t_m} ds\right\}, \\
&\qquad\qquad \text{by (5.3.15) and (5.3.16),} \\
&\quad \le C\left\{N^{-1} + \exp(-a(1-x_i)/\varepsilon)\right\} \\
&\quad \le C\left\{N^{-1} + \exp(-a\lambda/\varepsilon)\right\},
\end{aligned} \tag{5.3.23}$$

since $\tau \in \Omega_1^N$ means that $x_i \le 1 - \lambda$. This, together with the choice (5.3.14) of λ, implies the result of part (i).

To prove part (ii), we note that for $\tau \in \Omega_2^N, H_\tau = 2\lambda N^{-1} = 4a^{-1}\varepsilon \ln N$. Using this and (5.3.16) in (5.3.7) and (5.3.8), we obtain (5.3.18) and (5.3.19). □

5.4 Discrete Green's Function

Let (x^*, t^*) be a mesh node in Ω. The discrete Green's function $G \in V$ associated with (x^*, t^*) is defined by

$$B(\chi, G) = \chi(x^*, t^*) \quad \forall \chi \in V, \tag{5.4.1}$$

where we recall that $G(x,t) \equiv 0$ for $t > T$. From (5.2.27), G is well defined.

In this section we will derive a global estimate for G in the energy norm $|||\cdot|||$ and prove that G is negligible outside a narrow region extending upstream from (x^*, t^*). This region is defined by

$$\begin{aligned}\Omega_0 = \{(x,t) \in \Omega : 0 < x \le x^* + K_0\sigma_\beta \ln N,\\ |x - at - (x^* - at^*)| \le K_0\sigma_\eta \ln N\},\end{aligned} \tag{5.4.2}$$

where K_0 is a positive constant independent of ε, N and M. We choose K_0 in the proof of Theorem 5.4.2 below; σ_β and σ_η will be given in (5.4.13) and (5.4.14) respectively.

We start by introducing a cut-off function with exponential decay. Set

$$\omega(x,t) = g\left(\frac{x - x^*}{\sigma_\beta}\right) g\left(\frac{x - at - (x^* - at^*)}{\sigma_\eta}\right) g\left(\frac{x^* - at^* - (x - at)}{\sigma_\eta}\right) \tag{5.4.3}$$

where

$$g(r) = \frac{2}{1 + \exp(r)} \quad \text{for } r \in (-\infty, +\infty). \tag{5.4.4}$$

By some elementary calculations, one can easily show

Lemma 5.4.1 *For $\omega(x,t)$ defined in (5.4.3), there hold*

(i) $0 < \omega(x,t) \le 8$ *on* Ω;

(ii) $-\omega_\beta(x,t) > 0$ *on* Ω;

(iii) for each $\tau \in \Omega^N$, *if* $\sigma_\beta \geq H_\tau$ *and* $\sigma_\eta \geq K$, *then*

$$\max_\tau \omega / \min_\tau \omega \leq C, \quad \max_\tau |\omega_\beta| / \min_\tau |\omega_\beta| \leq C;$$

(iv) for all l *and* m,

$$\left| \frac{\partial^{l+m}\omega(x,t)}{\partial\beta^l \partial t^m} \right| \leq C\sigma_\beta^{-l}\sigma_\eta^{-m}\omega(x,t) \quad \text{on } \Omega;$$

(v) for all $l \geq 1$ *and all* m,

$$\left| \frac{\partial^{l+m}\omega(x,t)}{\partial\beta^l \partial t^m} \right| \leq C\sigma_\beta^{1-l}\sigma_\eta^{-m}|\omega_\beta(x,t)| \quad \text{on } \Omega;$$

(vi) on any triangle τ^* *which contains* (x^*, t^*),

$$\omega(x,t) \geq C;$$

(vii) $\omega(x,t) \leq CN^{-K_0}$ *on* $\Omega \backslash \Omega_0$.

We shall first derive a global estimate on G in a weighted energy norm, defined by

$$|||G|||_\omega^2 = \varepsilon\|\omega^{-1/2}G_s\|^2 + \|\omega^{-1/2}\rho^{1/2}G_\beta\|^2 + \left\| \left(\frac{1}{\omega}\right)_\beta^{1/2} G \right\|^2 + \|\omega^{-1/2}G\|^2 + \sum_{j=0}^{M} |[\omega^{-1/2}G]|_j^2. \tag{5.4.5}$$

This estimate will be obtained by demonstrating the following three lemmas.

Lemma 5.4.2 *If* $\sigma_\beta \geq \gamma N^{-1}$ *and* $\sigma_\eta^2 \geq \gamma\varepsilon$, *then for* $\gamma \geq 1$ *sufficiently large and independent of* N, M *and* ε,

$$B\left(\frac{G}{\omega}, G\right) \geq \frac{1}{4}|||G|||_\omega^2. \tag{5.4.6}$$

Proof.

$$
\begin{aligned}
B\left(\frac{G}{\omega},G\right) &= \varepsilon\left(\left(\frac{G}{\omega}\right)_s,G_s\right)+\left(\left(\frac{G}{\omega}\right)_\beta,\rho G_\beta\right)+\left(\frac{G}{\omega},G\right)\\
&\quad+\left(\frac{G}{\omega},(\rho-1)G_\beta\right)-\sum_{j=1}^{M}\left\langle\frac{G^-}{\omega},[G]\right\rangle_j\\
&=\varepsilon\|\omega^{-1/2}G_s\|^2+\varepsilon\left(\left(\frac{1}{\omega}\right)_s G,G_s\right)+\|\omega^{-1/2}\rho^{1/2}G_\beta\|^2\\
&\quad+\left(\left(\frac{1}{\omega}\right)_\beta G,\rho G_\beta\right)+\|\omega^{-1/2}G\|^2+\left(\frac{G}{\omega},\rho G_\beta\right)\\
&\quad-\left(\frac{G}{\omega},G_\beta\right)-\sum_{j=1}^{M}\left\langle\frac{G^-}{\omega},[G]\right\rangle_j.
\end{aligned}
\tag{5.4.7}
$$

Integrating by parts gives

$$
\left(\frac{G}{\omega},G_\beta\right)=\frac{1}{2}\sum_{j=1}^{M}\left\{\left\langle\frac{G^-}{\omega},G^-\right\rangle_j-\left\langle\frac{G^+}{\omega},G^+\right\rangle_{j-1}\right\}-\frac{1}{2}\left\|\left(\frac{1}{\omega}\right)_\beta^{1/2}G\right\|^2.
$$

Substituting this into (5.4.7), we get

$$
\begin{aligned}
B\left(\frac{G}{\omega},G\right) &= \varepsilon\|\omega^{-1/2}G_s\|^2+\|\omega^{-1/2}\rho^{1/2}G_\beta\|^2+\frac{1}{2}\left\|\left(\frac{1}{\omega}\right)_\beta^{1/2}G\right\|^2\\
&\quad+\|\omega^{-1/2}G\|^2+\frac{1}{2}\sum_{j=0}^{M}\left|[\omega^{-1/2}G]\right|_j^2+\varepsilon\left(\left(\frac{1}{\omega}\right)_s G,G_s\right)\\
&\quad+\left(\left(\frac{1}{\omega}\right)_\beta G,\rho G_\beta\right)+\left(\frac{G}{\omega},\rho G_\beta\right).
\end{aligned}
\tag{5.4.8}
$$

We bound the last three terms separately. First, by Cauchy-Schwarz' inequality and the arithmetic-geometric mean inequality, we have

$$
\begin{aligned}
&\left|\varepsilon\left(\left(\frac{1}{\omega}\right)_s G,G_s\right)\right|\\
&\le\frac{\varepsilon}{2}\|\omega^{-1/2}G_s\|^2+\frac{\varepsilon}{2}\left\|\omega^{1/2}\left(\frac{1}{\omega}\right)_s G\right\|^2\\
&\le\frac{\varepsilon}{2}\|\omega^{-1/2}G_s\|^2+C\varepsilon\left\{\left\|\omega^{1/2}\left(\frac{1}{\omega}\right)_\beta G\right\|^2+\left\|\omega^{1/2}\left(\frac{1}{\omega}\right)_t G\right\|^2\right\}
\end{aligned}
$$

$$\leq \frac{\varepsilon}{2}\|\omega^{-1/2}G_s\|^2 + C\varepsilon\left\{\max_{\Omega}\left(\omega\left(\frac{1}{\omega}\right)_\beta\right)\left\|\left(\frac{1}{\omega}\right)_\beta^{1/2}G\right\|^2\right.$$

$$\left. + \max_{\Omega}\left(\omega^2\left(\frac{1}{\omega}\right)_t^2\right)\|\omega^{-1/2}G\|^2\right\}$$

$$\leq \frac{\varepsilon}{2}\|\omega^{-1/2}G_s\|^2 + C\varepsilon\left\{\sigma_\beta^{-1}\left\|\left(\frac{1}{\omega}\right)_\beta^{1/2}G\right\|^2 + \sigma_\eta^{-2}\|\omega^{-1/2}G\|^2\right\}, \quad (5.4.9)$$

using Lemma 5.4.1 (iv). Similarly, using (5.2.20),

$$\left|\left(\left(\frac{1}{\omega}\right)_\beta G, \rho G_\beta\right)\right| \leq \frac{1}{4}\|\omega^{-1/2}\rho^{1/2}G_\beta\|^2 + CN^{-1}\sigma_\beta^{-1}\left\|\left(\frac{1}{\omega}\right)_\beta^{1/2}G\right\|^2. \quad (5.4.10)$$

Finally,

$$\left|\left(\frac{G}{\omega}, \rho G_\beta\right)\right| \leq \frac{1}{4}\|\omega^{-1/2}\rho^{1/2}G_\beta\|^2 + \|\rho^{1/2}\omega^{-1/2}G\|^2$$

$$\leq \frac{1}{4}\|\omega^{-1/2}\rho^{1/2}G_\beta\|^2 + \frac{1}{2}\|\omega^{-1/2}G\|^2, \quad (5.4.11)$$

since (5.2.20) and (5.2.2) imply $\rho < 1/2$.

Collecting (5.4.9) – (5.4.11) into (5.4.8), we get

$$B\left(\frac{G}{\omega}, G\right) \geq \frac{1}{2}|||G|||_\omega^2 - C\theta|||G|||_\omega^2$$

with $\theta = \max\left\{\varepsilon\sigma_\beta^{-1},\ \varepsilon\sigma_\eta^{-2},\ N^{-1}\sigma_\beta^{-1}\right\}$. Using $\varepsilon \leq N^{-1}$, we have

$$\theta = \max\left\{N^{-1}\sigma_\beta^{-1},\ \varepsilon\sigma_\eta^{-2}\right\} \leq \gamma^{-1}$$

by hypothesis. Choosing γ sufficiently large, independently of N, M and ε, completes the proof. □

Lemma 5.4.3 *Assume that $\sigma_\beta \geq N^{-1}$. Then*

$$\left|\left(\frac{G}{\omega}\right)(x^*, t^*)\right| \leq \frac{1}{16}|||G|||_\omega^2 + CN\ln^\delta N,$$

where $\delta = \begin{cases} 0, & \textit{when } (x^*, t^*) \in \Omega_1, \\ 1, & \textit{otherwise.} \end{cases}$

Proof. Let τ^* be a mesh triangle containing (x^*, t^*) and satisfying $t \leq t^*$ for all $(x, t) \in \tau^*$.

Suppose that $(x^*, t^*) \in \Omega_1$. We must have $\tau^* \in \Omega_1^N$. By Lemma 5.4.1 (vi) and (5.3.1), we have

$$\left|\left(\frac{G}{\omega}\right)(x^*, t^*)\right| \leq CN\|G\|_{\tau^*}$$
$$\leq CN\left(\max_{\tau^*}\left(\frac{1}{\omega}\right)_\beta^{-1/2}\right)\left\|\left(\frac{1}{\omega}\right)_\beta^{1/2} G\right\|_{\tau^*}. \tag{5.4.12}$$

A direct calculation shows that for $(x, t) \in \tau^*$

$$\left(\frac{1}{\omega}\right)_\beta^{-1}(x, t)$$
$$= 2\frac{\sigma_\beta}{a}\exp\left(-\frac{x - x^*}{\sigma_\beta}\right) g\left(\frac{x - at - (x^* - at^*)}{\sigma_\eta}\right) g\left(\frac{x^* - at^* - (x - at)}{\sigma_\eta}\right)$$
$$\leq 8\frac{\sigma_\beta}{a}\exp\left(-\frac{H_{\tau^*}}{\sigma_\beta}\right)$$
$$\leq CN^{-1},$$

by hypothesis. Inserting this into (5.4.12) and using an arithmetic-geometric inequality yields the desired result for $(x^*, t^*) \in \Omega_1$.

Now suppose instead that $(x^*, t^*) \in \Omega_2$. Let S^* be the slab which contains τ^*. Using (5.3.3),

$$\left|\left(\frac{G}{\omega}\right)(x^*, t^*)\right| \leq CN^{1/2}\lambda^{1/2}\|G_s\|_{S^*\cap\Omega_2}$$
$$\leq CN^{1/2}\sqrt{\varepsilon}\ln^{1/2} N\|\omega^{-1/2}G_s\|_{S^*\cap\Omega_2},$$

by Lemma 5.4.1 (vi) and (i). This proves the lemma for $(x^*, t^*) \in \Omega_2$. □

Lemma 5.4.4 *Assume that*

$$\sigma_\beta = \gamma N^{-1} \tag{5.4.13}$$

and

$$\sigma_\eta = \begin{cases} \gamma\varepsilon^{1/2}, & \text{if } N^{-3/2} \le \varepsilon \le N^{-1}, \\ \gamma N^{-3/2}\varepsilon^{-1/2}, & \text{if } N^{-2} \le \varepsilon \le N^{-3/2}, \\ \gamma N^{-1/2}, & \text{if } \varepsilon \le N^{-2}, \end{cases} \tag{5.4.14}$$

with $\gamma > 1$ sufficiently large, independent of N, M and ε. Then

$$\left| B\left(\left(\frac{G}{\omega}\right)^I - \frac{G}{\omega}, G\right)\right| \le \frac{1}{16}|||G|||_\omega^2,$$

where $\left(\frac{G}{\omega}\right)^I$ is the interpolant from V to $\frac{G}{\omega}$.

Proof. For convenience we set

$$E(x,t) = \left(\frac{G}{\omega}\right)^I (x,t) - \left(\frac{G}{\omega}\right)(x,t).$$

Then Cauchy-Schwarz' inequality gives

$$\begin{aligned} &|B(E,G)| \\ &\quad \le \varepsilon\|\omega^{1/2}E_x\|\,\|\omega^{-1/2}G_x\| + \|\rho^{1/2}\omega^{1/2}E_\beta\|\,\|\omega^{-1/2}\rho^{1/2}G_\beta\| \\ &\qquad + \|\omega^{1/2}E\|\,\|\omega^{-1/2}G\| + \|\rho^{-1/2}\omega^{1/2}E\|_{\Omega_1}\|\omega^{-1/2}\rho^{1/2}G_\beta\|_{\Omega_1} \\ &\qquad + \|\omega^{1/2}E\|_{\Omega_2}\|\omega^{-1/2}G_\beta\|_{\Omega_2} + \sum_{j=1}^{M}\left|\omega^{1/2}E^-\right|_j\left|[\omega^{-1/2}G]\right|_j. \end{aligned} \tag{5.4.15}$$

Note that there is no term $\|\omega^{-1/2}G_\beta\|_{\Omega_2}$ in $|||G|||_\omega^2$, so we will deal with this first. Obviously

$$\|\omega^{-1/2}G_\beta\|_{\Omega_2} \le \|\omega^{-1/2}G_x\|_{\Omega_2} + \|\omega^{-1/2}G_t\|_{\Omega_2}. \tag{5.4.16}$$

For each $\tau \in \Omega_2^N$,

$$\begin{aligned} \|\omega^{-1/2}G_t\|_\tau &\le \max_\tau \omega^{-1/2}\|G_t\|_\tau \\ &\le CN\|G\|_\tau / \min_\tau \omega^{1/2}, \quad \text{by (5.3.2)}, \\ &\le CN\|\omega^{-1/2}G\|_\tau, \end{aligned}$$

using Lemma 5.4.1 (iii). Hence

$$\|\omega^{-1/2}G_t\|_{\Omega_2} \le CN\|\omega^{-1/2}G\|_{\Omega_2}. \tag{5.4.17}$$

Since $G(1,t) = 0$ for all $t \in (0,T]$, we have for $(x,t) \in \Omega_2$,

$$\begin{aligned} |(\omega^{-1/2}G)(x,t)| &\le \int_x^1 \left|(\omega^{-1/2}G)_x(\xi,t)\right| d\xi \\ &\le \lambda^{1/2}\left\{\int_{1-\lambda}^1 \left|(\omega^{-1/2}G)_x(\xi,t)\right|^2 d\xi\right\}^{1/2}, \end{aligned} \tag{5.4.18}$$

thus

$$\begin{aligned} \|\omega^{-1/2}G\|_{\Omega_2} &\le \left\{\int_0^T \int_{1-\lambda}^1 \left[\lambda \int_{1-\lambda}^1 \left|(\omega^{-1/2}G)_x(\xi,t)\right|^2 d\xi\right] dxdt\right\}^{1/2} \\ &\le \lambda\|(\omega^{-1/2}G)_x\|_{\Omega_2} \\ &\le \lambda\left\{\|\omega^{-1/2}G_x\|_{\Omega_2} + \|(\omega^{-1/2})_\beta G\|_{\Omega_2} + \|(\omega^{-1/2})_t G\|_{\Omega_2}\right\} \\ &\le C\lambda\left\{\|\omega^{-1/2}G_x\|_{\Omega_2} + \sigma_\beta^{-1/2}\left\|\left(\frac{1}{\omega}\right)_\beta^{1/2} G\right\|_{\Omega_2} + \sigma_\eta^{-1}\|\omega^{-1/2}G\|_{\Omega_2}\right\}, \\ &\qquad\qquad \text{by Lemma 5.4.1 (iv),} \\ &\le C\lambda\left(\varepsilon^{-1/2} + \sigma_\beta^{-1/2} + \sigma_\eta^{-1}\right)|||G|||_\omega \\ &\le C\varepsilon^{-1/2}\lambda|||G|||_\omega, \end{aligned} \tag{5.4.19}$$

using $\sigma_\beta \ge \varepsilon$ and $\sigma_\eta \ge \varepsilon^{1/2}$.

From (5.4.16), (5.4.17) and (5.4.19), using $\varepsilon \le N^{-1}$, we get

$$\begin{aligned} \|\omega^{-1/2}G_\beta\|_{\Omega_2} &\le \left(\varepsilon^{-1/2} + C\varepsilon^{-1/2}N\lambda\right)|||G|||_\omega \\ &\le C\varepsilon^{-1/2}\ln N|||G|||_\omega. \end{aligned} \tag{5.4.20}$$

Now using (5.4.20) in (5.4.15) and applying the arithmetic-geometric mean inequality, we obtain

$$|B(E,G)| \le \frac{1}{32}|||G|||_\omega^2 + C\Big\{\varepsilon\|\omega^{1/2}E_x\|^2 + \|\rho^{1/2}\omega^{1/2}E_\beta\|^2 + N\|\omega^{1/2}E\|_{\Omega_1}^2$$

$$+\varepsilon^{-1}\ln^2 N\|\omega^{1/2}E\|^2_{\Omega_2}+\sum_{j=1}^{M}\left|\omega^{1/2}E^-\right|_j^2\Big\}. \tag{5.4.21}$$

From Lemma 5.3.2, we have, for any $\tau\in\Omega^N$,

$$\begin{aligned}&\|E\|_\tau+N^{-1/2}|E^-|_{j,\tau}\\ &\quad\le C\left\{H_\tau^2\left\|\left(\frac{G}{\omega}\right)_{ss}\right\|_\tau+H_\tau K\left\|\left(\frac{G}{\omega}\right)_{st}\right\|_\tau+K^2\left\|\left(\frac{G}{\omega}\right)_{tt}\right\|_\tau\right\},\end{aligned} \tag{5.4.22}$$

$$\|E_s\|_\tau\le C\left\{H_\tau\left\|\left(\frac{G}{\omega}\right)_{ss}\right\|_\tau+K\left\|\left(\frac{G}{\omega}\right)_{st}\right\|_\tau\right\} \tag{5.4.23}$$

and

$$\|E_\beta\|_\tau\le C(H_\tau+K)\left\{\left\|\left(\frac{G}{\omega}\right)_{ss}\right\|_\tau+\left\|\left(\frac{G}{\omega}\right)_{st}\right\|_\tau+\left\|\left(\frac{G}{\omega}\right)_{tt}\right\|_\tau\right\}. \tag{5.4.24}$$

On each $\tau\in\Omega^N$, $G_{ss}=G_{st}=G_{tt}=0$. Consequently we have, for each τ,

$$\left\|\left(\frac{G}{\omega}\right)_{ss}\right\|_\tau\le\left\|\left(\frac{1}{\omega}\right)_{ss}G\right\|_\tau+2\left\|\left(\frac{1}{\omega}\right)_{s}G_s\right\|_\tau, \tag{5.4.25}$$

$$\left\|\left(\frac{G}{\omega}\right)_{st}\right\|_\tau\le\left\|\left(\frac{1}{\omega}\right)_{t}G_s\right\|_\tau+\left\|\left(\frac{1}{\omega}\right)_{st}G\right\|_\tau+\left\|\left(\frac{1}{\omega}\right)_{s}G_t\right\|_\tau, \tag{5.4.26}$$

$$\left\|\left(\frac{G}{\omega}\right)_{tt}\right\|_\tau\le\left\|\left(\frac{1}{\omega}\right)_{tt}G\right\|_\tau+2\left\|\left(\frac{1}{\omega}\right)_{t}G_t\right\|_\tau. \tag{5.4.27}$$

Set

$$\tilde{\omega}_\tau=\min_\tau\omega^{1/2}.$$

Using Lemmas 5.3.1 and 5.4.1, and $\sigma_\beta\le\sigma_\eta$, we have the following upper bounds (in some cases two or more bounds are given for the same term in order to handle different values of ε and N later):

$$\left\|\left(\frac{1}{\omega}\right)_{tt}G\right\|_\tau\le C\sigma_\eta^{-2}\|\omega^{-1/2}G\|_\tau/\tilde{\omega}_\tau \tag{5.4.28}$$

for each $\tau \in \Omega^N$;

$$\left\|\left(\frac{1}{\omega}\right)_{ss} G\right\|_\tau + \left\|\left(\frac{1}{\omega}\right)_{st} G\right\|_\tau$$
$$\leq C\left\{\left\|\left(\frac{1}{\omega}\right)_{\beta\beta} G\right\|_\tau + \left\|\left(\frac{1}{\omega}\right)_{\beta t} G\right\|_\tau + \left\|\left(\frac{1}{\omega}\right)_{tt} G\right\|_\tau\right\}$$
$$\leq C\left\{\sigma_\beta^{-3/2}\left\|\left(\frac{1}{\omega}\right)_\beta^{1/2} G\right\|_\tau + \sigma_\eta^{-2}\|\omega^{-1/2}G\|_\tau\right\}\Big/\tilde{\omega}_\tau \tag{5.4.29}$$

for each $\tau \in \Omega^N$;

$$\left\|\left(\frac{1}{\omega}\right)_t G_t\right\|_\tau \leq \begin{cases} C\sigma_\eta^{-1}\left(\|\omega^{-1/2}G_\beta\|_\tau + \|\omega^{-1/2}G_s\|_\tau\right)/\tilde{\omega}_\tau \\ CN\sigma_\eta^{-1}\|\omega^{-1/2}G\|_\tau/\tilde{\omega}_\tau \end{cases} \tag{5.4.30}$$

for each $\tau \in \Omega^N$;

$$\left\|\left(\frac{1}{\omega}\right)_s G_t\right\|_\tau \leq \begin{cases} C\Big(\sigma_\beta^{-1/2}N\left\|\left(\frac{1}{\omega}\right)_\beta^{1/2} G\right\|_\tau + \sigma_\eta^{-1}\|\omega^{-1/2}G_\beta\|_\tau \\ \quad + \sigma_\eta^{-1}\|\omega^{-1/2}G_s\|_\tau\Big)\Big/\tilde{\omega}_\tau, \\ CN\Big(\sigma_\beta^{-1/2}\left\|\left(\frac{1}{\omega}\right)_\beta^{1/2} G\right\|_\tau + \sigma_\eta^{-1}\|\omega^{-1/2}G\|_\tau\Big)\Big/\tilde{\omega}_\tau \end{cases} \tag{5.4.31}$$

for each $\tau \in \Omega^N$;

$$\left\|\left(\frac{1}{\omega}\right)_t G_s\right\|_\tau \leq \begin{cases} C\sigma_\eta^{-1}\|\omega^{-1/2}G_s\|_\tau/\tilde{\omega}_\tau & \forall\tau \in \Omega^N, \\ CN\sigma_\eta^{-1}\|\omega^{-1/2}G\|_\tau/\tilde{\omega}_\tau & \forall\tau \in \Omega_1^N, \end{cases} \tag{5.4.32}$$

$$\left\|\left(\frac{1}{\omega}\right)_s G_s\right\|_\tau \leq \begin{cases} CN\left(\sigma_\beta^{-1/2}\left\|\left(\frac{1}{\omega}\right)_\beta^{1/2} G\right\|_\tau + \sigma_\eta^{-1}\|\omega^{-1/2}G\|_\tau\right)\Big/\tilde{\omega}_\tau \\ \qquad\qquad \forall\tau \in \Omega_1^N, \\ C\left(\sigma_\beta^{-1/2}N\left\|\left(\frac{1}{\omega}\right)_\beta^{1/2} G\right\|_\tau + \sigma_\eta^{-1}\|\omega^{-1/2}G_s\|_\tau\right)\Big/\tilde{\omega}_\tau \\ \qquad\qquad \forall\tau \in \Omega_1^N, \\ C\sigma_\beta^{-1}\|\omega^{-1/2}G_s\|_\tau/\tilde{\omega}_\tau \quad \forall\tau \in \Omega^N. \end{cases} \tag{5.4.33}$$

Let

$$\begin{cases} L_{11} = \sigma_\eta^{-1}\varepsilon^{-1/2}, \qquad L_{12} = \sigma_\beta^{-1}\varepsilon^{-1/2}, \\ L_2 = \sigma_\eta^{-1}N^{1/2}, \\ L_3 = \max\{\sigma_\beta^{-3/2},\ \sigma_\beta^{-1/2}N\}, \\ L_{41} = \sigma_\eta^{-2}, \quad L_{42} = \sigma_\eta^{-1}N, \quad L_4 = \max\{L_{41},\ L_{42}\}. \end{cases} \tag{5.4.34}$$

Let $|||G|||_{\omega,D}$ denote that the integrations and summations in $|||G|||_\omega$ (see (5.4.5)) are extended only over D for any measurable $D \subseteq \Omega$.

Then collecting (5.4.28) – (5.4.33) into (5.4.25) – (5.4.27) yields

$$\left\|\left(\frac{G}{\omega}\right)_{ss}\right\|_\tau \le \begin{cases} C(L_3 + L_4)|||G|||_{\omega,\tau}/\tilde{\omega}_\tau \quad \forall\tau \in \Omega_1^N, \\ C(L_{11} + L_3 + L_{41})|||G|||_{\omega,\tau}/\tilde{\omega}_\tau \quad \forall\tau \in \Omega_1^N, \\ C(L_{12} + L_3 + L_{41})|||G|||_{\omega,\tau}/\tilde{\omega}_\tau \quad \forall\tau \in \Omega_2^N; \end{cases} \tag{5.4.35}$$

$$\left\|\left(\frac{G}{\omega}\right)_{st}\right\|_\tau \le \begin{cases} C(L_3 + L_4)|||G|||_{\omega,\tau}/\tilde{\omega}_\tau \quad \forall\tau \in \Omega_1^N, \\ C(L_{11} + L_2 + L_3 + L_{41})|||G|||_{\omega,\tau}/\tilde{\omega}_\tau \quad \forall\tau \in \Omega_1^N, \\ C(L_{11} + L_3 + L_4)|||G|||_{\omega,\tau}/\tilde{\omega}_\tau \quad \forall\tau \in \Omega_2^N; \end{cases} \tag{5.4.36}$$

$$\left\|\left(\frac{G}{\omega}\right)_{tt}\right\|_\tau \le \begin{cases} C(L_{11} + L_2 + L_{41})|||G|||_{\omega,\tau}/\tilde{\omega}_\tau \quad \forall\tau \in \Omega_1^N, \\ CL_4\|\omega^{-1/2}G\|_\tau/\tilde{\omega}_\tau \quad \forall\tau \in \Omega^N. \end{cases} \tag{5.4.37}$$

We may now bound the terms in (5.4.21). From (5.4.23), (5.4.35) and (5.4.36), using Lemma 5.4.1 (iii), we have

$$\begin{aligned} \varepsilon\|\omega^{1/2}E_s\|^2 &\le \varepsilon \sum_{\tau\in\Omega^N} \left(\max_\tau \omega\right) \|E_s\|_\tau^2 \\ &\le C\varepsilon N^{-2}\left\{\left(L_{11}^2 + L_2^2 + L_3^2 + L_{41}^2\right)|||G|||_{\omega,\Omega_1}^2\right. \\ &\quad \left. + \left[\lambda^2\left(L_{12}^2 + L_3^2 + L_{41}^2\right) + L_{11}^2 + L_3^2 + L_4^2\right]|||G|||_{\omega,\Omega_2}^2\right\} \\ &\le C\varepsilon N^{-2}\left\{L_{11}^2 + L_{12}^2 + L_2^2 + L_3^2 + L_4^2\right\}|||G|||_\omega^2 \end{aligned}$$

$$\leq C\left\{N^{-2}\sigma_\eta^{-2}+N^{-2}\sigma_\beta^{-2}+\varepsilon N^{-1}\sigma_\eta^{-2}+\varepsilon N^{-2}\sigma_\beta^{-3}\right.$$
$$\left.+\varepsilon\sigma_\beta^{-1}+\varepsilon N^{-2}\sigma_\eta^{-4}+\varepsilon\sigma_\eta^{-2}\right\}|||G|||_\omega^2$$
$$\leq C\gamma^{-1}|||G|||_\omega^2, \tag{5.4.38}$$

since $\sigma_\beta=\gamma N^{-1}$, $\sigma_\eta\geq\gamma\varepsilon^{1/2}$ and $\sigma_\beta\leq\sigma_\eta$.

Similarly, using (5.4.35) – (5.4.37), we get two different upper bounds:

$$N\|\omega^{1/2}E\|_{\Omega_1}^2+\|\omega^{1/2}\rho^{1/2}E_\beta\|^2+\sum_{j=1}^{M}\left|\omega^{1/2}E^-\right|_{j,\Omega_1}^2$$
$$\leq\begin{cases} CN^{-3}(L_3^2+L_4^2)|||G|||_\omega^2 \\ CN^{-3}(L_{11}^2+L_2^2+L_3^2+L_{41}^2)|||G|||_\omega^2\end{cases}$$
$$\leq\begin{cases} C\left(N^{-3}\sigma_\beta^{-3}+N^{-1}\sigma_\beta^{-1}+N^{-3}\sigma_\eta^{-4}+N^{-1}\sigma_\eta^{-2}\right)|||G|||_\omega^2 \\ C\left(N^{-3}\varepsilon^{-1}\sigma_\eta^{-2}+N^{-2}\sigma_\eta^{-2}+N^{-3}\sigma_\beta^{-3}+N^{-1}\sigma_\beta^{-1}+N^{-3}\sigma_\eta^{-4}\right)|||G|||_\omega^2\end{cases}$$
$$\leq C\gamma^{-1}|||G|||_\omega^2, \tag{5.4.39}$$

from (5.4.13) and (5.4.14), where the first bound above is used when $\varepsilon\leq N^{-2}$ and the second when $N^{-2}<\varepsilon\leq N^{-1}$.

Finally,

$$\varepsilon^{-1}\ln^2 N\|\omega^{1/2}E\|_{\Omega_2}^2+\sum_{j=1}^{M}\left|\omega^{1/2}E^-\right|_{j,\Omega_2}^2$$
$$\leq C\left(\varepsilon^{-1}N^{-1}\ln^2 N+1\right)N^{-3}\left\{\lambda^4\left(L_{12}^2+L_3^2+L_{41}^2\right)|||G|||_{\omega,\Omega_2}^2\right.$$
$$\left.+\lambda^2\left(L_{11}^2+L_3^2+L_4^2\right)|||G|||_{\omega,\Omega_2}^2+L_4^2\|\omega^{-1/2}G\|_{\Omega_2}^2\right\}$$
$$\leq C\varepsilon^{-1}N^{-4}\lambda^2\ln^2 N\left\{L_{11}^2+L_{12}^2+L_3^2+L_{41}^2+L_4^2\varepsilon^{-1}\right\}|||G|||_\omega^2,$$

using $1\leq\varepsilon^{-1}N^{-1}$, $\lambda\leq C$ and (5.4.19),

$$\leq CN^{-4}\ln^4 N\left\{\sigma_\eta^{-2}+\sigma_\beta^{-2}+\varepsilon\sigma_\beta^{-3}+\varepsilon N^2\sigma_\beta^{-1}+\sigma_\eta^{-4}+N^2\sigma_\eta^{-2}\right\}|||G|||_\omega^2,$$

recalling (5.3.14) and (5.4.34),

$$\leq C\gamma^{-1}|||G|||_{\omega}^2, \tag{5.4.40}$$

since $\sigma_\beta = \gamma N^{-1}$ and $\sigma_\eta \geq \gamma N^{-3/4}$ from (5.4.14).

Substituting (5.4.38) – (5.4.40) into (5.4.21), we get

$$|B(E,G)| \leq \left(\frac{1}{32} + C\gamma^{-1}\right) |||G|||_{\omega}^2.$$

The desired result follows on choosing γ sufficiently large, independently of N, M and ε. □

We are now in a position to present our main results of this section.

Theorem 5.4.1 *Assume that σ_β and σ_η are chosen as in (5.4.13) and (5.4.14) respectively. Then*

$$|||G|||^2 \leq 8|||G|||_{\omega}^2 \leq CN\ln^{\delta} N,$$

where $\delta = \begin{cases} 0, & \textit{when } (x^*, t^*) \in \Omega_1, \\ 1, & \textit{otherwise.} \end{cases}$

Proof. The first inequality can be easily obtained by using Lemma 5.4.1 (i). To show the second inequality, we take $\chi = \left(\frac{G}{\omega}\right)^I$ in (5.4.1) to get

$$B\left(\left(\frac{G}{\omega}\right)^I, G\right) = \left(\frac{G}{\omega}\right)(x^*, t^*).$$

But by Lemmas 5.4.2 and 5.4.4,

$$B\left(\left(\frac{G}{\omega}\right)^I, G\right) = B\left(\frac{G}{\omega}, G\right) + B\left(\left(\frac{G}{\omega}\right)^I - \frac{G}{\omega}, G\right) \geq \frac{3}{16}|||G|||_{\omega}^2.$$

Now Lemma 5.4.3 yields the desired result. □

With Theorem 5.4.1 we may derive our second estimate on G.

Theorem 5.4.2 *Assume that the hypotheses of Theorem 5.4.1 hold. Then for each nonnegative integer s, there exists a positive constant $C = C(s)$ such that*

$$\|G\|_{W^1_\infty(\Omega_1\setminus\Omega_0)} \le CN^{-s}, \tag{5.4.41}$$

and

$$\|G\|_{L^\infty(\Omega_2\setminus\Omega_0)} + \varepsilon|G|_{W^1_\infty(\Omega_2\setminus\Omega_0)} \le CN^{-s}, \tag{5.4.42}$$

where we have used the usual notation for the Sobolev space W^1_∞ and its associated seminorm and norm.

Proof. Define $\Omega_0' \subseteq \Omega$ by (5.4.2) with K_0 replaced by $K_0/2$. Assume without loss of generality that Ω_0' is a mesh domain, by enlarging it slightly where necessary. Given s, choose

$$K_0 = 4(s+3). \tag{5.4.43}$$

Then by Lemma 5.4.1 (vii), $\omega \le CN^{-K_0/2} \le CN^{-2(s+3)}$ on $\Omega \setminus \Omega_0'$. Hence

$$|||G|||^2_{\Omega\setminus\Omega_0'} \le CN^{-2(s+3)}|||G|||^2_\omega \le CN^{-2(s+2)}, \tag{5.4.44}$$

by Theorem 5.4.1. Then (5.4.41) follows using the inverse estimate (5.3.1).

Next, let $(x', t') \in \Omega_2 \setminus \Omega_0$ be arbitrary. Suppose first that

$$x' - at' \le x^* - at^* - K_0\sigma_\eta \ln N.$$

Then we have

$$\begin{aligned} |G(x',t')| &= \left|\int_0^{1-\lambda} G_s(s,t')\,ds + \int_{1-\lambda}^{s} G_s(s,t')\,ds\right| \\ &\le CN\left\{\|G_s\|_{T_1\cap\Omega_1} + \sqrt{\lambda}\|G_s\|_{T_1\cap\Omega_2}\right\}, \end{aligned}$$

since G_s is piecewise constant, where T_1 is the union of those mesh triangles which contain the line segment $\{(x, t') \in \Omega \backslash \Omega_0 : 0 \leq x \leq x'\}$. Note that by our supposition $T_1 \subseteq \Omega \setminus \Omega_0'$.

Now suppose instead that $x' - at' \geq x^* - at^* + K_0\sigma_\eta \ln N$. Then

$$|G(x', t')| = \left| \int_s^1 G_s(s, t')\, ds \right| \leq CN\sqrt{\lambda} \|G_s\|_{T_2 \cap \Omega_2},$$

where $T_2 \subseteq \Omega \backslash \Omega_0'$ is the union of those mesh triangles which contain the line segment $\{(x, t') \in \Omega \setminus \Omega_0 : x' \leq x \leq 1\}$.

Hence for $(x', t') \in \Omega_2 \setminus \Omega_0$ we always have

$$\begin{aligned} |G(x', t')| &\leq CN \left\{ \|G_s\|_{\Omega_1 \backslash \Omega_0'} + \sqrt{\varepsilon} \ln^{1/2} N \|G_s\|_{\Omega_2 \backslash \Omega_0'} \right\} \\ &\leq CN \left\{ N\|G\|_{\Omega_1 \backslash \Omega_0'} + \sqrt{\varepsilon} \ln^{1/2} N \|G_s\|_{\Omega_2 \backslash \Omega_0'} \right\}, \quad \text{using (5.3.1)}, \\ &\leq CN^{-s}, \end{aligned}$$

by (5.4.44). This proves the $L^\infty(\Omega_2 \setminus \Omega_0)$ estimate in (5.4.42).

It remains to show the $\varepsilon|G|_{W^1_\infty(\Omega_2 \backslash \Omega_0)}$ estimates. We may assume that (x', t') does not lie on the boundary of any triangle. Suppose that (x', t') lies in the triangle τ_0. Now G_s and G_t are constant on τ_0. Hence

$$\begin{aligned} |G_s(x', t')| &+ |G_t(x', t')| \\ &= [\text{meas}(\tau_0)]^{-1/2} (\|G_s\|_{\tau_0} + \|G_t\|_{\tau_0}) \\ &\leq C\varepsilon^{-1} N^{-s}, \end{aligned}$$

using (5.3.2), the fact that $\tau_0 \subseteq \Omega \setminus \Omega_0'$ and (5.4.44). □

Remark 5.4.1 In obtaining the global energy norm estimate of G in Theorem 5.4.1, we used a sharpening of Niijima's approach [32] for elliptic problems. Using the ideas above, one can improve the global L^2- and H^1-estimates of the Green's function given in Lemma 2.2 of [32], by removing all ln factors there.

Remark 5.4.2 Theorem 5.4.2 shows that the discrete Green's function G essentially vanishes outside Ω_0 with σ_β, σ_η and K_0 determined by (5.4.13), (5.4.14) and (5.4.43). Since the dimensions of Ω_0 are much greater than the maximum diameter of the mesh triangles, we may assume that Ω_0 is a mesh domain.

5.5 Localized Pointwise Error Estimates

In this section we will estimate the nodal error between the exact solution u and our computed solution U. In order to derive a nodal error formula suitable for an analysis under weak assumptions on u, we need the following lemma.

Lemma 5.5.1 *For any $w \in V$ and any mesh subdomain $D \subseteq \Omega$, there exist $Pw \in V$ and a positive constant C, independent of N, M and ε, such that*

$$Pw = w \qquad \text{on } D, \tag{5.5.1}$$

$$\|Pw\| \le C\|w\|_{L^\infty(D)}, \tag{5.5.2}$$

$$|Pw|_{W_1^1(\Omega)} \le CN\|w\|_{L^\infty(D)}, \tag{5.5.3}$$

$$\langle 1, |(Pw)^-|\rangle_j \le C\|w\|_{L^\infty(D)}, \quad \text{for } j = 1,\dots,M, \tag{5.5.4}$$

where $|\cdot|_{W_1^1(\Omega)}$ denotes the usual seminorm in $W_1^1(\Omega)$.

Proof. For each $j \in \{1,\dots,M\}$, let $\{\phi_{i,j}\}_{i=1}^{2(N-1)}$ be a set of basis functions in V_j. Then any $w \in V$ can be expressed as

$$w(x,t) = \sum_{i=1}^{2(N-1)} w_{i,j}\,\phi_{i,j}(x,t) \quad \text{for } (x,t) \in S_j \text{ and } j = 1,\dots,M.$$

We define Pw by

$$Pw(x,t) = \sum_{i=1}^{2(N-1)} v_{i,j}\,\phi_{i,j}(x,t) \quad \text{for } (x,t) \in S_j \text{ and } j = 1,\ldots,M, \tag{5.5.5}$$

where

$$v_{i,j} = \begin{cases} w_{i,j}, & \text{when } \operatorname{supp}(\phi_{i,j}) \cap D^0 \neq \emptyset, \\ 0, & \text{otherwise,} \end{cases} \tag{5.5.6}$$

and D^0 denotes the interior of D. Then $Pw \in V$, (5.5.1) is satisfied and

$$\operatorname{dist}\big(\operatorname{supp}(Pw), D\big) \leq 2N^{-1}.$$

By (5.5.6), the nonzero terms in (5.5.5) are associated with i's such that

$$\operatorname{supp}(\phi_{i,j}) \cap \tau^0 \neq \emptyset$$

for some τ with $\tau^0 \subseteq D \cap S_j$. Thus the nonzero $v_{i,j}$'s can be bounded by

$$|v_{i,j}| = |w_{i,j}| \leq \|w\|_{L^\infty(\tau_{i,j})} \quad \text{for some } \tau_{i,j}^0 \subseteq D \cap S_j. \tag{5.5.7}$$

From (5.5.5) – (5.5.7) we get

$$\begin{aligned} \|Pw\|_{S_j \setminus D} &\leq \sum_i |v_{i,j}|\,\|\phi_{i,j}\|_{S_j \setminus D} \\ &\leq CN^{-1}\|w\|_{L^\infty(S_j \cap D)}, \end{aligned} \tag{5.5.8}$$

since $\|\phi_{i,j}\|_{S_j \setminus D} \leq CN^{-1}$ for all i and j, and there are at most four terms in the sum.

Summing (5.5.8) over all j we get

$$\|Pw\|_{\Omega \setminus D} \leq CN^{-1/2}\|w\|_{L^\infty(D)}. \tag{5.5.9}$$

Now (5.5.2) follows easily.

Note that for all j and i,

$$|\phi_{i,j}|_{W_1^1(\Omega)} + \|\phi_{i,j}(\cdot, t_j)\|_{L^1(0,1)} \leq CN^{-1}. \tag{5.5.10}$$

Using (5.5.10), one can prove analogously to the above that

$$|Pw|_{W_1^1(\Omega)} \leq |w|_{W_1^1(\Omega)} + C\|w\|_{L^\infty(D)} \tag{5.5.11}$$

and

$$\langle 1, |(Pw)^-|\rangle_j \leq \langle 1, |w^-|\rangle_{j,D} + CN^{-1}\|w\|_{L^\infty(D)}. \tag{5.5.12}$$

Now (5.5.4) is immediate from (5.5.12). To show (5.5.3), we use (5.5.7) and (5.5.10) to get

$$\begin{aligned} |w|_{W_1^1(D)} &\leq C\sum_{j=1}^{M} \sum_{\tau_{i,j}^0 \subseteq D\cap S_j} \|w\|_{L^\infty(\tau_{i,j})}N^{-1} \\ &\leq CN\|w\|_{L^\infty(D)}. \end{aligned} \tag{5.5.13}$$

This, together with (5.5.11), proves (5.5.3). □

Let u^I be the interpolant from V to u. Using Lemma 5.5.1, we define $Pu^I \in V$ such that

$$Pu^I = u^I \quad \text{on } \Omega_0 \tag{5.5.14}$$

and

$$\begin{aligned} &\|Pu^I\| + N^{-1}|Pu^I|_{W_1^1(\Omega)} + \left\langle 1, \left|(Pu^I)^-\right|\right\rangle_j \\ &\leq C\left\|u^I\right\|_{L^\infty(\Omega_0)} \leq C\|u\|_{L^\infty(\Omega_0)}. \end{aligned} \tag{5.5.15}$$

Take $\chi = U - Pu^I$ in (5.4.1). Note that $\left(Pu^I\right)(x^*, t^*) = u(x^*, t^*)$. We get

$$(U - u)(x^*, t^*) = B\left(U - Pu^I, G\right). \tag{5.5.16}$$

Set $\eta(x,t) = \left(u - Pu^I\right)(x,t)$. Using (5.2.21) and (5.1.1), we have

$$\begin{aligned} B\left(U - Pu^I, G\right) &= (f, G + \rho G_\beta) + < u^0, G^+ >_0 - B\left(Pu^I, G\right) \\ &= B(\eta, G) - \varepsilon(u_{ss}, \rho G_\beta). \end{aligned} \tag{5.5.17}$$

Combining (5.5.16) and (5.5.17) and using (5.2.23) we obtain the following nodal error formula:

$$\begin{aligned} (U - u)(x^*, t^*) &= \varepsilon(\eta_s, G_s) + (\eta_\beta, \rho G_\beta) + (\eta, G) + (\eta, (\rho - 1)G_\beta) \\ &\quad - \varepsilon(u_{ss}, \rho G_\beta) - \sum_{j=1}^{M} < \eta^-, [G] >_j \\ &\equiv R(\eta, G). \end{aligned} \tag{5.5.18}$$

Let $R_D(\eta, G)$ denote that the integrations in (5.5.18) are extended only over D, for any domain D. Then

$$(U - u)(x^*, t^*) = R_{\Omega_0}(\eta, G) + R_{\Omega\backslash\Omega_0}(\eta, G), \tag{5.5.19}$$

where $< \eta^-, [G] >_j$ is split by

$$< \eta^-, [G] >_j = < \eta^-, [G] >_{j,\Omega_0} + < \eta^-, [G] >_{j,\Omega\backslash\Omega_0} . \tag{5.5.20}$$

Lemma 5.5.2

$$\begin{aligned} |R_{\Omega\backslash\Omega_0}(\eta, G)| \le C N^{-2} \ln N \, \{ &\|f\| + \|u^0\|_{L^2(0,1)} \\ &+ \|u\|_{L^\infty(\Omega_0)} + \|u_\beta\|_{L^1(\Omega)} + \|\varepsilon u_{ss}\|_{L^1(\Omega)} \} . \end{aligned}$$

Proof. From (5.5.18)

$$\begin{aligned} &|R_{\Omega\backslash\Omega_0}(\eta, G)| \\ &\quad \le \|G\|_{W^1_\infty(\Omega_1\backslash\Omega_0)} \left\{ \varepsilon\|\eta_s\|_{L^1(\Omega_1)} + \|\rho\eta_\beta\|_{L^1(\Omega_1)} + \|\eta\|_{L^1(\Omega_1)} + \|\rho\varepsilon u_{ss}\|_{L^1(\Omega_1)} \right\} \end{aligned}$$

$$+\varepsilon|G|_{W^1_\infty(\Omega_2\setminus\Omega_0)}\left\{\|\eta_s\|_{L^1(\Omega_2)}+\varepsilon^{-1}\|\eta\|_{L^1(\Omega_2)}\right\}$$
$$+\|G\|_{L^\infty(\Omega\setminus\Omega_0)}\left\{\|\eta\|_{L^1(\Omega)}+\sum_{j=1}^{M}\langle 1,|\eta^-|\rangle_j\right\}. \tag{5.5.21}$$

We have

$$\begin{aligned}
&\varepsilon\|\eta_s\|_{L^1(\Omega_1)}+\|\rho\eta_\beta\|_{L^1(\Omega_1)}+\|\eta\|_{L^1(\Omega_1)}\\
&\quad\le C\Big\{\varepsilon\|u_s\|_{L^1(\Omega_1)}+N^{-1}\|u_\beta\|_{L^1(\Omega_1)}+\|u\|_{L^1(\Omega_1)}\\
&\qquad+\varepsilon\left\|(Pu^I)_s\right\|_{L^1(\Omega_1)}+N^{-1}\left\|(Pu^I)_\beta\right\|_{L^1(\Omega_1)}+\|Pu^I\|_{L^1(\Omega_1)}\Big\}\\
&\quad\le C\left\{\varepsilon\|u_s\|_{\Omega_1}+N^{-1}\|u_\beta\|_{L^1(\Omega_1)}+\|u\|_{\Omega_1}+\|u\|_{L^\infty(\Omega_0)}\right\},
\end{aligned} \tag{5.5.22}$$

using $\varepsilon\le N^{-1}$ and (5.5.15).

Since $\eta(1,t)=0$ for all $t\in(0,T]$,

$$\|\eta\|_{L^1(\Omega_2)}\le C\varepsilon\ln N\|\eta_s\|_{L^1(\Omega_2)}.$$

Hence

$$\begin{aligned}
&\|\eta_s\|_{L^1(\Omega_2)}+\varepsilon^{-1}\|\eta\|_{L^1(\Omega_2)}\\
&\quad\le C\ln N\|\eta_s\|_{L^1(\Omega_2)}\\
&\quad\le C\ln N\left\{\|u_s\|_{L^1(\Omega_2)}+\left\|(Pu^I)_s\right\|_{L^1(\Omega_2)}\right\}\\
&\quad\le C\ln N\left\{\sqrt{\varepsilon}\ln^{1/2}N\|u_s\|_{\Omega_2}+N\|u\|_{L^\infty(\Omega_0)}\right\},
\end{aligned} \tag{5.5.23}$$

using (5.5.15) again.

Finally, from (5.5.15)

$$\begin{aligned}
\sum_{j=1}^{M}\langle 1,|\eta^-|\rangle_j&\le\sum_{j=1}^{M}\left\{<1,|u|>_j+\left\langle 1,\left|(Pu^I)^-\right|\right\rangle_j\right\}\\
&\le\sum_{j=1}^{M}|u|_j+CN\|u\|_{L^\infty(\Omega_0)}.
\end{aligned} \tag{5.5.24}$$

From (5.1.1) it is easy to show that for each $j \in \{1, \dots, M\}$,

$$\varepsilon\|u_x\|^2_{\hat{S}_j} + \|u\|^2_{\hat{S}_j} + |u|^2_j \le \|f\|^2_{\hat{S}_j} + \|u^0\|^2_{L^2(0,1)} \qquad (5.5.25)$$

where $\hat{S}_j = \cup_{l=1}^{j} S_j$.

Collecting (5.5.22) – (5.5.24) into (5.5.21), using (5.5.25) and Theorem 5.4.1 with $s = 3$, we obtain

$$\begin{aligned}
&\left|R_{\Omega\setminus\Omega_0}(\eta, G)\right| \\
&\quad \le CN^{-3}\Big\{\varepsilon\|u_x\|_{\Omega_1} + N^{-1}\|u_\beta\|_{L^1(\Omega_1)} + \|u\|_{\Omega_1} + N^{-1}\|\varepsilon u_{xx}\|_{L^1(\Omega_1)} \\
&\qquad + \sqrt{\varepsilon}\ln^{3/2} N\|u_x\|_{\Omega_2} + N\ln N\|u\|_{L^\infty(\Omega_0)} + \sum_{j=1}^{M}|u|_j\Big\} \\
&\quad \le CN^{-3}\Big\{\left(\ln^{3/2} N + N + 1\right)\left(\|f\| + \|u^0\|_{L^2(0,1)}\right) \\
&\qquad + N\ln N\|u\|_{L^\infty(\Omega_0)} + N^{-1}\|u_\beta\|_{L^1(\Omega)} + N^{-1}\|\varepsilon u_{xx}\|_{L^1(\Omega)}\Big\}.
\end{aligned}$$

The desired result follows immediately. □

We are now ready for our main theorems.

Theorem 5.5.1 *(Pointwise error estimate away from layers) Assume that*

$$\|f\| + \|u^0\|_{L^2(0,1)} + \|u_\beta\|_{L^1(\Omega)} + \|\varepsilon u_{xx}\|_{L^1(\Omega)} \le C, \qquad (5.5.26)$$

and

$$\|u\|_{C^2(\Omega_0)} \le C. \qquad (5.5.27)$$

Then

$$|(U - u)(x^*, t^*)| \le C\sigma_\eta^{1/2} N^{-1} \ln^{\xi} N,$$

where σ_η is given in (5.4.14) and $\xi = \begin{cases} 1/2, & \text{when } (x^*, t^*) \in \Omega_1, \\ 1, & \text{otherwise.} \end{cases}$

Proof. Recalling (5.5.18), we have

$$
\begin{aligned}
&|R_{\Omega_0}(\eta, G)| \\
&\quad \le \sqrt{\varepsilon}\|\eta_s\|_{L^\infty(\Omega_0)}\sqrt{\varepsilon}\|G_s\|_{L^1(\Omega_0)} + N^{-1/2}\|\eta_\beta\|_{L^\infty(\Omega_0)}\|\rho^{1/2}G_\beta\|_{L^1(\Omega_0)} \\
&\quad + \|\eta\|_{L^\infty(\Omega_0)}\|G\|_{L^1(\Omega_0)} + \|\eta\|_{L^\infty(\Omega_0)}\|G_\beta\|_{L^1(\Omega_0)} \\
&\quad + \varepsilon N^{-1/2}\|u_{ss}\|_{L^\infty(\Omega_0)}\|\rho^{1/2}G_\beta\|_{L^1(\Omega_0)} + \|\eta\|_{L^\infty(\Omega_0)}\sum_{j=1}^{M} < 1, |[G]| >_{j,\Omega_0}.
\end{aligned}
$$

Using (5.5.27), Theorem 5.3.1 and

$$
\mathrm{meas}(\Omega_0) + \mathrm{meas}(\Lambda_j(\Omega_0)) \le C\sigma_\eta \ln N, \tag{5.5.28}
$$

where $\Lambda_j(\Omega_0)$ is as in (5.2.15), we get

$$
\begin{aligned}
&|R_{\Omega_0}(\eta, G)| \\
&\quad \le C\sigma_\eta^{1/2}\ln^{1/2} N\Big\{\varepsilon N^{-1}\|G_s\|_{\Omega_0} + N^{-3/2}\|\rho^{1/2}G_\beta\|_{\Omega_0} \\
&\quad + N^{-2}\|G\|_{\Omega_0} + N^{-2}\left(\|G_\beta\|_{\Omega_1\cap\Omega_0} + \sqrt{\lambda}\|G_\beta\|_{\Omega_2\cap\Omega_0}\right) \\
&\quad + \varepsilon N^{-1/2}\|\rho^{1/2}G_\beta\|_{\Omega_0} + N^{-2}\sum_{j=1}^{M}\||[G]|\big|_{j,\Omega_0}\Big\}.
\end{aligned}
$$

Hence, using $\varepsilon \le N^{-1}$ and (5.3.4),

$$
\begin{aligned}
|R_{\Omega_0}(\eta, G)| &\le C\sigma_\eta^{1/2}N^{-3/2}\ln^{1/2} N|||G||| \\
&\le C\sigma_\eta^{1/2}N^{-1}\ln^{(1+\delta)/2} N,
\end{aligned} \tag{5.5.29}
$$

by Theorem 5.4.1.

Applying (5.5.29) and Lemma 5.5.2 to (5.5.19) concludes the argument. □

We next give a pointwise convergence result for the case when (x^*, t^*) lies in the boundary layer, under the assumption that the solution u exhibits typical boundary layer behaviour in the neighbourhood Ω_0 of (x^*, t^*).

Theorem 5.5.2 *(Pointwise error estimate inside the boundary layer) Assume that (5.5.26) holds and that*

$$\left|\frac{\partial^{i+j}u(x,t)}{\partial x^i \partial t^j}\right| \leq C\left\{1+\varepsilon^{-i}\exp(a(1-x)/\varepsilon)\right\} \quad \text{on } \Omega_0 \tag{5.5.30}$$

for $i+j \leq 2$. Then

$$|(U-u)(x^*,t^*)| \leq C\sigma_\eta^{1/2}N^{-1/2}\ln^\zeta N,$$

where σ_η is given in (5.4.14) and $\zeta = \begin{cases} 2, & \text{when } (x^*,t^*) \in \Omega_1, \\ 5/2, & \text{otherwise.} \end{cases}$

Proof. Set

$$\lambda_\varepsilon = -2\varepsilon \ln \varepsilon + \mu N^{-1}, \tag{5.5.31}$$

where $0 \leq \mu < 1/2$ is a constant which is chosen so that $1-\lambda_\varepsilon \in \{x_i : i = 1,\dots,N/2\}$.

Divide Ω_0 into three parts:

$$\Omega_0 = I_1 \cup I_2 \cup I_3, \tag{5.5.32}$$

where

$$I_1 = \{(x,t) \in \Omega_0 : \ 0 < x \leq 1-\lambda_\varepsilon\},$$

$$I_2 = \{(x,t) \in \Omega_0 : \ 1-\lambda_\varepsilon < x \leq 1-\lambda\},$$

$$I_3 = \{(x,t) \in \Omega_0 : \ 1-\lambda \leq x\}.$$

Clearly I_i $(i = 1,2,3)$ are mesh domains with $I_1 \cup I_2 \subseteq \Omega_1$ and $I_3 \subseteq \Omega_2$.

In order to estimate $R_{\Omega_0}(\eta, G)$, we bound $R_{I_i}(\eta, G)$ $(i = 1,2,3)$ successively. First, from (5.5.30) we have $\|u\|_{C^2(I_1)} \leq C$. Thus by arguments similar to the proof of Theorem 5.5.1, we have

$$|R_{I_1}(\eta, G)| \leq C\sigma_\eta^{1/2}N^{-1}\ln N. \tag{5.5.33}$$

Next, we bound the terms in $R_{I_2}(\eta, G)$. Integrating $\int_\tau \eta_x$ with respect to x and using Theorem 5.3.2 (i),

$$
\begin{aligned}
|(\varepsilon\eta_x, G_x)_{I_2}| &= \left|\sum_{\tau\subseteq I_2} G_x\big|_\tau \int_\tau \varepsilon\eta_x\right| \\
&\leq C \sum_{\tau\subseteq I_2} \left|G_x\big|_\tau\right| \varepsilon N^{-2} \\
&\leq C\varepsilon \|G_x\|_{L^1(I_2)}.
\end{aligned}
$$

Similarly

$$\left|(\eta_\beta, \rho G_\beta)_{I_2}\right| \leq CN^{-1}\|G_\beta\|_{L^1(I_2)},$$

$$|(\eta, G)_{I_2}| \leq CN^{-1}\|G\|_{L^1(I_2)},$$

$$\left|(\eta, (\rho-1)G_\beta)_{I_2}\right| \leq CN^{-1}\|G_\beta\|_{L^1(I_2)},$$

$$
\begin{aligned}
&\left|\varepsilon\left(u_{xx}, \rho G_\beta\right)_{I_2}\right| \\
&\quad\leq C \sum_{\tau\subseteq I_2} N^{-1}\varepsilon \left|G_\beta\big|_\tau\right| \int_\tau \left[1 + \varepsilon^{-2}\exp(-a(1-x)/\varepsilon)\right] d\tau \\
&\quad\leq C \sum_{\tau\subseteq I_2} N^{-1}\varepsilon \left|G_\beta\big|_\tau\right| \left\{N^{-2} + N^{-1}\varepsilon^{-1}\exp(-a\lambda/\varepsilon)\right\} \\
&\quad\leq CN^{-2}\|G_\beta\|_{L^1(I_2)}.
\end{aligned}
$$

Finally

$$\left|\sum_{j=1}^{M} \langle \eta^-, [G]\rangle_{j,I_2}\right| \leq CN^{-1} \sum_{j=1}^{M} \langle 1, |[G]|\rangle_{j,I_2}.$$

From the above estimates and Cauchy-Schwarz' inequality, we get

$$
\begin{aligned}
|R_{I_2}(\eta, G)| \leq\; & CN^{-1/2}|||G||| \left\{[\text{meas}(I_2)]^{1/2}\right. \\
& \left. + N^{-1/2} \max_j [\text{meas}\,(\Lambda_j(I_2))]^{1/2}\right\}. \qquad (5.5.34)
\end{aligned}
$$

Now

$$\begin{aligned}\text{meas}(I_2) &\le C(\lambda_\varepsilon - \lambda)\sigma_\eta \ln N \\ &\le C\left(\varepsilon \ln(\varepsilon N)^{-1} + N^{-1}\right)\sigma_\eta \ln N \\ &\le C\sigma_\eta N^{-1} \ln N, \end{aligned} \tag{5.5.35}$$

using $\ln(\varepsilon N)^{-1} \le (\varepsilon N)^{-1}$. Also

$$\text{meas}\,(\Lambda_j(I_2)) \le \lambda_\varepsilon - \lambda \le CN^{-1}. \tag{5.5.36}$$

Thus from (5.5.34) – (5.5.36) and Theorem 5.4.1, we obtain

$$|R_{I_2}(\eta, G)| \le C\sigma_\eta^{1/2} N^{-1/2} \ln N. \tag{5.5.37}$$

Finally, for $R_{I_3}(\eta, G)$, recall that $\rho = 0$ on I_3, so that

$$R_{I_3}(\eta, G) = \varepsilon(\eta_s, G_s)_{I_3} + (\eta, G)_{I_3} - (\eta, G_\beta)_{I_3} - \sum_{j=1}^{M} \langle \eta^-, [G] \rangle_{j,I_3}.$$

By Theorem 5.3.2 (ii) and Cauchy-Schwarz' inequality,

$$\begin{aligned}|R_{I_3}(\eta, G)| \le\ & C\left\{N^{-1} \ln N \|G_s\|_{I_3} + N^{-2} \ln^2 N \|G\|_{I_3} \right. \\ & \left. + N^{-2} \ln^2 N \|G_\beta\|_{I_3}\right\} [\text{meas}(I_3)]^{1/2} \\ & + CN^{-2} \ln^2 N \sum_{j=1}^{M} |[G]|_{j,I_3} [\text{meas}\,(\Lambda_j(I_3))]^{1/2}. \end{aligned}$$

Since

$$\text{meas}(I_3) \le C\lambda\sigma_\eta \ln N \le C\sigma_\eta \varepsilon \ln^2 N$$

and

$$\text{meas}\,(\Lambda_j(I_3)) \le C\sigma_\eta \ln N,$$

thus from (5.3.4) and Theorem 5.4.1,

$$\begin{aligned}|R_{I_3}(\eta, G)| &\le C\sigma_\eta^{1/2} N^{-1} \ln^2 N |||G||| \\ &\le C\sigma_\eta^{1/2} N^{-1/2} \ln^{2+\delta/2} N. \end{aligned} \tag{5.5.38}$$

Combining (5.5.33), (5.5.37) and (5.5.38) gives

$$|R_{\Omega_0}(\eta, G)| \le C\sigma_\eta^{1/2} N^{-1/2} \ln^{\zeta} N, \tag{5.5.39}$$

which together with Lemma 5.5.2 and (5.5.19) proves the desired result. □

Recall (5.4.14). From Theorems 5.5.1 and 5.5.2, we reach the following conclusion for our streamline diffusion scheme (5.2.21).

Corollary 5.5.1 *Assume that (5.5.26) holds. Then in smooth regions, the scheme (5.2.21) is pointwise accurate of order almost* $O(\varepsilon^{1/4} N^{-1})$ *when* $N^{-3/2} \le \varepsilon \le N^{-1}$, *order almost* $O(\varepsilon^{-1/4} N^{-7/4})$ *when* $N^{-2} \le \varepsilon \le N^{-3/2}$ *and order almost* $O(N^{-5/4})$ *when* $0 < \varepsilon \le N^{-2}$. *In the regions where the solution exhibits typical boundary layer behaviour, on the other hand, the scheme is almost order* $O(\varepsilon^{1/4} N^{-1/2})$, $O(\varepsilon^{-1/4} N^{-5/4})$ *and* $O(N^{-3/4})$ *in the above three cases respectively. These results are uniform in* ε.

Remark 5.5.1 The assumption (5.5.26) in Theorems 5.5.1 and 5.5.2 is reasonable in many cases. In fact, an inspection of the proof of Lemma 5.5.2 shows that in (5.5.26) one can replace C by $CK^{-\mu}$ for any fixed positive constant μ without affecting the conclusions of our error analysis.

Remark 5.5.2 Applying the initial conditions in a strong form on each time level leads to a three-level scheme. Then the above analysis starting from (5.2.21) still applies, except that the terms involving integrals of the form $< \cdot, \cdot >_j$ now disappear. That is, this three-level scheme is theoretically as accurate as our two-level scheme above.

5.6 Numerical Results

In this section, we verify experimentally the theoretical results obtained in Section 5.5. Nodal errors and convergence rates for our scheme (5.2.18) – (5.2.20) with (5.3.14) are presented for two test problems.

In each computation we take $N = M$ and solve the problems for various ε and N. We note that the characteristics of the reduced solution of (5.1.1) run from southwest to northeast, while our division of rectangles into triangles in Section 5.2 used gridlines running from northwest to southeast. Thus our mesh is not tailored to the reduced problem. In fact, similar rates of convergence are observed when the gridlines coincide with the characteristics.

The scheme (5.2.18) – (5.2.20) is used successively on a sequence of time levels. On each level, the scheme is equivalent to a system of $2(N-1)$ linear equations. The coefficient matrix of the system can be easily permuted to yield a pentadiagonal matrix. Hence it is possible to solve the system by triangular decomposition with $O(N)$ operations.

All calculations were carried out in C double precision on an IBM PC.

Example 5.6.1 We first test the performance of our scheme when applied to a problem with typical boundary layer behaviour:

$$-\varepsilon u_{xx} + u_x + u + u_t = f(x,t) \quad \text{on } \Omega \tag{5.6.1}$$

with analytical solution

$$u(x,t) = t\exp(-(1-x)/\varepsilon) + 1 - x^2 + t^2, \tag{5.6.2}$$

where $\Omega = (0,1) \times (0,1]$. The function $f(x,t)$ and the initial-boundary values on $\bar{\Omega}$ are chosen to fit this data.

The problem is solved with $\alpha = 0.61$ in (5.3.14).

Table 5.6.1 Global Maximum Nodal Errors

ε	N=8	16	32	64	128	256
1.56250e-2	1.645e-1	8.277e-2	3.331e-2	1.336e-2	4.880e-3	1.677e-3
3.90625e-3	1.999e-1	1.055e-1	4.763e-2	2.097e-2	8.765e-3	3.385e-3
9.76562e-4	2.147e-1	1.195e-1	6.186e-2	3.035e-2	1.403e-2	6.097e-3
2.44141e-4	2.191e-1	1.249e-1	6.951e-2	3.687e-2	1.897e-2	9.232e-3
6.10352e-5	2.202e-1	1.265e-1	7.217e-2	4.006e-2	2.201e-2	1.162e-2
1.52588e-5	2.205e-1	1.269e-1	7.291e-2	4.113e-2	2.313e-2	1.279e-2
3.81470e-6	2.206e-1	1.270e-1	7.310e-2	4.142e-2	2.346e-2	1.316e-2
9.53674e-7	2.206e-1	1.270e-1	7.315e-2	4.149e-2	2.354e-2	1.326e-2
2.38419e-7	2.206e-1	1.270e-1	7.316e-2	4.151e-2	2.356e-2	1.329e-2
5.96046e-8	2.206e-1	1.270e-1	7.317e-2	4.151e-2	2.357e-2	1.329e-2
1.49012e-8	2.206e-1	1.270e-1	7.317e-2	4.151e-2	2.357e-2	1.330e-2
3.72529e-9	2.206e-1	1.270e-1	7.317e-2	4.151e-2	2.357e-2	1.330e-2

Table 5.6.2 Global Convergence Rates

ε	N=8	16	32	64	128
1.56250e-2	0.99	1.31	1.32	1.45	1.54
3.90625e-3	0.92	1.15	1.18	1.26	1.37
9.76562e-4	0.85	0.95	1.03	1.11	1.20
2.44141e-4	0.81	0.85	0.91	0.96	1.04
6.10352e-5	0.80	0.81	0.85	0.86	0.92
1.52588e-5	0.80	0.80	0.83	0.83	0.86
3.81470e-6	0.80	0.80	0.82	0.82	0.83
9.53674e-7	0.80	0.80	0.82	0.82	0.83
2.38419e-7	0.80	0.80	0.82	0.82	0.83
5.96046e-8	0.80	0.80	0.82	0.82	0.83
1.49012e-8	0.80	0.80	0.82	0.82	0.83
3.72529e-9	0.80	0.80	0.82	0.82	0.83

The global maximum nodal errors $E^{(\varepsilon,N)}$ between the exact solution u and the computed solution $U^{(\varepsilon,N)}$ and the corresponding convergence rates $p^{(\varepsilon,N)}$ are displayed in Tables 5.6.1 and 5.6.2 respectively, where $E^{(\varepsilon,N)}$ and $p^{(\varepsilon,N)}$ are computed from (3.4.3) and (3.4.4).

We remark that the maximum errors in Table 5.6.1 occur at nodes inside the boundary layer. Table 5.6.2 shows that for this test problem, the uniform convergence rate of our scheme is 0.83 as $N \to \infty$, which is close to the value 0.75 proven in Theorem 5.5.2.

Example 5.6.2 We now examine how our scheme performs locally away from all layers. Consider

$$-\varepsilon u_{xx} + 3\,u_x + u + u_t = f(x,t) \quad \text{on } \Omega \tag{5.6.3}$$

with discontinuous initial data at $x = 0.5$, so that the solution $u(x,t)$ has an internal layer lying along the line $x = 3t + 0.5$.

Using $u_R(x,t)$ to denote the sum of the reduced solution to (5.6.3) and the boundary layer component of $u(x,t)$ at $x = 1$, we choose the initial and boundary data and $f(x,t)$ so that

$$u_R(x,t) = (x+t)^3 + A\exp(-x-t) + (1+t^2)\exp(-3(1-x)/\varepsilon), \tag{5.6.4}$$

where

$$A = \begin{cases} 0, & \text{when } x < 3t + 0.5, \\ 1, & \text{otherwise.} \end{cases} \tag{5.6.5}$$

We do not have an explicit expression for the exact solution $u(x,t)$, so we compare our computed solution $U^{(\varepsilon,N)}$ with $u_R(x,t)$ on Ω', where

$$\Omega' = \{(x,t) \in \Omega : \; 0 \le x \le 0.99, \quad 0.5 \le t \le 1\}.$$

It is valid to use u_R instead of u when ε is small, since Ω' is then outside the internal and boundary layers, i.e., the solution of (5.6.3) – (5.6.5) is smooth in Ω'.

We solve this problem with $\alpha = 1.6$ in (5.3.14).

Table 5.6.3 Local Maximum Nodal Errors

ε	N=8	16	32	64	128	256
9.76562e-4	3.651e-2	1.139e-2	3.108e-3	7.886e-4	1.907e-4	4.569e-5
2.44141e-4	3.668e-2	1.152e-2	3.184e-3	8.286e-4	2.070e-4	5.003e-5
6.10352e-5	3.673e-2	1.155e-2	3.204e-3	8.412e-4	2.137e-4	5.305e-5
1.52588e-5	3.674e-2	1.155e-2	3.210e-3	8.445e-4	2.156e-4	5.408e-5
3.81470e-6	3.674e-2	1.156e-2	3.211e-3	8.453e-4	2.161e-4	5.436e-5
9.53674e-7	3.674e-2	1.156e-2	3.211e-3	8.455e-4	2.162e-4	5.443e-5
2.38419e-7	3.674e-2	1.156e-2	3.211e-3	8.456e-4	2.163e-4	5.445e-5
5.96046e-8	3.674e-2	1.156e-2	3.211e-3	8.456e-4	2.163e-4	5.445e-5
1.49012e-8	3.674e-2	1.156e-2	3.211e-3	8.456e-4	2.163e-4	5.446e-5
3.72529e-9	3.674e-2	1.156e-2	3.211e-3	8.456e-4	2.163e-4	5.446e-5

Table 5.6.4 Local Convergence Rates

ε	N=8	16	32	64	128
9.76562e-4	1.68	1.87	1.98	2.05	2.06
2.44141e-4	1.67	1.85	1.94	2.00	2.05
6.10352e-5	1.67	1.85	1.93	1.98	2.01
1.52588e-5	1.67	1.85	1.93	1.97	2.00
3.81470e-6	1.67	1.85	1.93	1.97	1.99
9.53674e-7	1.67	1.85	1.93	1.97	1.99
2.38419e-7	1.67	1.85	1.93	1.97	1.99
5.96046e-8	1.67	1.85	1.93	1.97	1.99
1.49012e-8	1.67	1.85	1.93	1.97	1.99
3.72529e-9	1.67	1.85	1.93	1.97	1.99

Table 5.6.3 lists the local maximum nodal errors $E_L^{(\varepsilon,N)}$ given by

$$E_L^{(\varepsilon,N)} = \max_{(x_i,t_j)\in\Omega'} \left| U^{(\varepsilon,N)}(x_i,t_j) - u_R(x_i,t_j) \right| . \tag{5.6.6}$$

The corresponding convergence rates $p_L^{(\varepsilon,N)}$ are listed in Table 5.6.4, which are computed from $E_L^{(\varepsilon,N)}$ analogously to (3.4.4). Our numerical results indicate that in this smooth region, the scheme is approximately second order as $N \to \infty$, which is better than the order of 5/4 predicted by Theorem 5.5.2. A similar gap between theory and numerical experience is present in all analyses of the streamline diffusion method (see, e.g., [22, 29]).

We also tested the method on (5.6.3) and (5.6.4) with $A = 0$ on Ω. In this case (5.6.4) gives a smooth exact solution to (5.6.3). The numerical results on Ω' for this smooth problem are identical to those displayed in Tables 5.6.3 and 5.6.4, except for a little difference in the errors when $N = 8$ and 16. This means that for (5.6.3) – (5.6.5), the local performance in Ω' of our scheme is not strongly affected by the presence of the internal and boundary layers.

Remark 5.6.1 Comparing Tables 5.6.1 and 5.6.3, we see that our method is much more accurate away from layers, as predicted by Corollary 5.5.1.

Remark 5.6.2 After the work in this chapter was completed, we became aware of the existence of Zhou [53], who has recently performed a similar analysis, obtaining pointwise error estimates only outside all layers on a quasiuniform mesh.

Chapter 6

A Cell Vertex Finite Volume Method

6.1 Introduction

The cell vertex finite volume method is a commonly used discretization scheme for conservation laws. It has been highly successful in modelling flows in aerodynamics. Since the method fits very naturally with convection problems, it has advantageous properties for convection-diffusion problems. However, all analyses for cell vertex methods have been carried out either for pure convection problems (see, e.g., Morton and Süli [28], Süli [48, 49] and Morton and Stynes [27]), or for convection-diffusion two-point boundary value problems (see, e.g., Mackenzie and Morton [25] and Morton and Stynes [27]). So far, there has been no similar analysis for a parabolic convection-diffusion problem in the literature.

In this final chapter, we examine a cell vertex finite volume method when applied to the following model time-dependent convection-diffusion problem:

$$Lu(x,t) \equiv -\varepsilon u_{xx} + au_x + bu + ru_t = f(x,t) \quad \forall (x,t) \in \Omega, \qquad (6.1.1)$$

$$u(0,t) = u(1,t) = 0 \quad \text{for} \quad 0 < t \leq T, \qquad (6.1.2)$$

$$u(x,0) = u^0(x) \quad \text{for} \quad 0 \le x \le 1, \tag{6.1.3}$$

where $0 < \varepsilon \ll 1$ and Ω is as in the last chapter. For simplicity, we assume that a, b and r are constants with

$$a > 0, \quad b > 0 \quad \text{and} \quad r > 0. \tag{6.1.4}$$

We also assume that $f \in L_2(\Omega)$ and $u^0 \in L_2(0,1)$.

The outline of the chapter is as follows. In Section 6.2 we describe the cell vertex method for (6.1.1) – (6.1.4) and reformulate it as a finite element method. Section 6.3 is devoted to the derivation of a discrete Gårding inequality which guarantees the existence and uniqeness of the finite volume solution. Local errors in the l_2 seminorm (defined in Section 6.3) are analyzed in Section 6.4. (We note that, when restricted to certain piecewise bilinear trial spaces, this seminorm becomes a norm.) Our analysis indicates that on a general tensor product mesh, the method is first order accurate away from all layers, in the l_2 seminorm. We can sharpen this result to local second order accuracy in l_2, if either ε is very small compared to the mesh diameter or the mesh is locally almost uniform.

We hope in the future to continue this analysis of the cell vertex finite volume method. In particular we intend to investigate the causes and treatment of chequer-board modes.

6.2 Description of the Cell Vertex Scheme

To discretize (6.1.1) – (6.1.4), we first define a partition of Ω as follows. For any pair of positive integers N and M, we consider the arbitrary tensor product grid

$$\Omega^h = \Big\{(x_i, t_j) \in \Omega: \ 0 = x_0 < x_1 < \ldots < x_N = 1,$$
$$0 = t_0 < t_1 < \ldots < t_M = T\Big\},$$

with $h_i = x_i - x_{i-1}, k_j = t_j - t_{j-1}$ and $h = \max_{i,j}\{h_i, k_j\}$. Define the "finite volume" or "cell" $K_{i,j}$ by

$$K_{i,j} = (x_{i-1}, x_i) \times (t_{j-1}, t_j), \qquad \text{for } i = 1, \ldots, N \text{ and } j = 1, \ldots, M.$$

In the finite volume context, the discretization of (6.1.1) is performed on each cell. The basic idea is to integrate (6.1.1) over a cell so that the convection and diffusion terms are converted into line integrals of normal fluxes along the cell edges. Then use the trapezoidal rule to approximate the integrals. Thus, letting u^h denote the computed solution, for cell $K_{i,j}$ we have

$$\begin{aligned}
&\iint_{K_{i,j}} f(x,t)\,dxdt \\
&\quad = -\varepsilon \int_{t_{j-1}}^{t_j} \Big(u_x^h(x_i,t) - u_x^h(x_{i-1},t)\Big)\,dt \\
&\qquad + a \int_{t_{j-1}}^{t_j} \Big(u^h(x_i,t) - u^h(x_{i-1},t)\Big)\,dt \\
&\qquad + r \int_{x_{i-1}}^{x_i} \Big(u^h(x,t_j) - u^h(x,t_{j-1})\Big)\,dt + \iint_{K_{i,j}} bu^h(x,t)\,dxdt \\
&\quad \approx \frac{\varepsilon k_j}{2}\Big(u_x^h(x_i,t_j) - u_x^h(x_{i-1},t_j) + u_x^h(x_i,t_{j-1}) - u_x^h(x_{i-1},t_{j-1})\Big) \\
&\qquad + \frac{ak_j}{2}\Big(u^h(x_i,t_j) - u^h(x_{i-1},t_j) + u^h(x_i,t_{j-1}) - u^h(x_{i-1},t_{j-1})\Big) \\
&\qquad + \frac{rh_i}{2}\Big(u^h(x_i,t_j) - u^h(x_i,t_{j-1}) + u^h(x_{i-1},t_j) - u^h(x_{i-1},t_{j-1})\Big) \\
&\qquad + \frac{bh_ik_j}{4}\Big(u^h(x_i,t_j) + u^h(x_{i-1},t_j) + u^h(x_i,t_{j-1}) + u^h(x_{i-1},t_{j-1})\Big).
\end{aligned} \tag{6.2.1}$$

With the approximation $u^h(x,t)$ parameterized by its values at the vertices, this still leaves two problems to be solved. Firstly, how do we define u_x^h at the nodes?

There are several ways in which this may be done, but we consider here the so-called Method A in Mackenzie and Morton [25]. That is, we define

$$u_s^h(x_i, t_j) = \frac{1}{h_i + h_{i+1}} \left(u^h(x_{i+1}, t_j) - u^h(x_{i-1}, t_j) \right), \tag{6.2.2}$$

for $i = 1, \ldots, N-1$, and

$$u_s^h(0, t_j) = \frac{2}{h_1} \left(u^h(x_1, t_j) - u^h(0, t_j) \right) - u_s^h(x_1, t_j). \tag{6.2.3}$$

Similarly to (6.2.3) one can define $u_s^h(1, t_j)$. This solves the first problem.

The second difficulty is as follows. If we perform the discretization (6.2.1) on all cells, we will have a system of NM equations in $(N-1)M$ unknowns, since $u^h(x,t)$ will be prescribed on three sides of Ω using (6.1.2) and (6.1.3). That is, we have M equations too many. To obtain an exact match, we choose upwind control volumes, that is, each nodal unknown is associated with the cell upwind of it. We do this by discarding the equations associated with $K_{N,j}$ for $j = 1, \ldots, M$. We then obtain a system of equations (6.2.1) – (6.2.3), for $i = 1, \ldots, N-1$ and $j = 1, \ldots, M$, which has exactly the same number of unknowns as that of equations. The second problem then disappears.

Finite volume methods are often interpreted as finite difference methods. This is often reflected in the finite difference techniques used to analyse such schemes. However, for a scheme which does not satisfy a discrete maximum principle, such as (6.2.1) – (6.2.3), a satisfactory finite difference analysis is difficult to obtain. In fact, the cell vertex formulation of the finite volume method has a natural interpretation as a Petrov-Galerkin finite element method. The finite element framework then affords the possibility of applying some finite element techniques to estimate errors in the finite volume method; see [28, 48, 49, 27].

To reformulate the cell vertex finite volume scheme (6.2.1) - (6.2.3) as a finite element method, we first define our trial and test spaces. Set

$$\mathcal{U}_0^h = \{v \in H^1(\Omega) \cap C(\bar{\Omega}) : \; v(0,t) = v(1,t) = 0 \text{ for } t \in (0,T], \\ v \text{ is bilinear on each cell } K\},$$

$$\mathcal{M}^h = \{p \in L^2(\Omega) : \; p \text{ is constant on each cell } K, \\ p \equiv 0 \text{ on cells } K_{N,j} \text{ for } j = 1,\dots,M\}.$$

In order to simplify the presentation, we introduce the averaging operators μ, μ_x and μ_t, for $i = 1,\dots,N$ and $j = 1,\dots,M$,

$$\mu w_{i,j} = \frac{1}{h_i k_j} \iint_{K_{i,j}} w(x,t)\,dx dt,$$

$$\mu_x w_{i,j} = \frac{1}{h_i} \int_{x_{i-1}}^{x_i} w(x,t_j)\,dx,$$

$$\mu_t w_{i,j} = \frac{1}{k_j} \int_{t_{j-1}}^{t_j} w(x_i,t)\,dt,$$

for all $w(x,t)$ for which the right hand side is defined.

Remark 6.2.1 One can easily verify that for each $v \in \mathcal{U}_0^h$ and for $i = 1,\dots,N$ and $j = 1,\dots,M$,

$$\mu v_{i,j} = \frac{1}{4}(v_{i-1,j} + v_{i,j} + v_{i-1,j-1} + v_{i,j-1}) \tag{6.2.4}$$

$$= \frac{1}{2}(\mu_x v_{i,j} + \mu_x v_{i,j-1}) \tag{6.2.5}$$

$$= \frac{1}{2}(\mu_t v_{i,j} + \mu_t v_{i-1,j}), \tag{6.2.6}$$

$$\mu(v_x)_{i,j} = \frac{1}{h_i}(\mu_t v_{i,j} - \mu_t v_{i-1,j}), \tag{6.2.7}$$

$$\mu(v_t)_{i,j} = \frac{1}{k_j}(\mu_x v_{i,j} - \mu_x v_{i,j-1}), \tag{6.2.8}$$

where $v_{i,j}$ denotes $v(x_i, t_j)$.

Now the cell vertex finite volume approximation is defined as follows: find $u^h \in \mathcal{U}_0^h$ satisfying

$$\hat{B}\left(u^h, p\right) = (f, p) \qquad \forall p \in \mathcal{M}^h, \tag{6.2.9}$$

$$< u^h(\cdot, 0), p^+ > = < u^0, p^+ > \qquad \forall p \in \mathcal{M}^h, \tag{6.2.10}$$

where $(\cdot,\cdot)$ and $< \cdot,\cdot >$ are the usual $L^2(\Omega)$ and $L^2(0,1)$ inner products,

$$p^+(x) = \lim_{t \to 0+} p(x,t),$$

and we set for any $(v,p) \in H^1(\Omega) \times \mathcal{M}^h$,

$$\begin{aligned} \hat{B}(v,p) = & -\varepsilon \sum_{j=1}^{M} \sum_{i=1}^{N-1} k_j \, \mu p_{i,j} \left\{ \hat{\mu}_t(v_s)_{i,j} - \hat{\mu}_t(v_s)_{i-1,j} \right\} \\ & + (av_s + rv_t + bv, p). \end{aligned} \tag{6.2.11}$$

Here we define, for $j = 1, \ldots, M$ and $v \in C(\Omega)$,

$$\hat{\mu}_t(v_s)_{i,j} = \begin{cases} \frac{2}{h_i + h_{i+1}} (\mu v_{i+1,j} - \mu v_{i,j}), & \text{if } i = 1, \ldots, N-1, \\ \\ \frac{2}{h_1} \mu_t v_{1,j} - \hat{\mu}_t(v_s)_{1,j}, & \text{if } i = 0. \end{cases} \tag{6.2.12}$$

Remark 6.2.2 For the discretization of the diffusion term in $\hat{B}(v,p)$, we do not need nodal values of v_s, just its integral along two sides $x = x_{i-1}$ and $x = x_i$ of $K_{i,j}$. For $i = 1, \ldots, N-1$, $\hat{\mu}_t(v_s)_{i,j}$ is defined by associating $\mu v_{i,j}$ with the cell centre then taking the obvious divided difference. For $i = 0$, we define $\hat{\mu}_t(v_s)_{i,j}$ by an extrapolation. It is easy to check that (6.2.1) – (6.2.3) is equivalent to (6.2.9) – (6.2.12).

In the next section, we will show the existence and uniqueness of the finite volume solution.

6.3 Stability and Convergence

We begin our analysis of the cell vertex finite volume scheme (6.2.9) – (6.2.12) by establishing the stability of the method in some appropriate mesh-dependent norms, which in turn implies the existence and uniqueness of the finite volume solution u^h.

We introduce the following mesh-dependent norms:

$$|v|_{l_2(\Omega^h)} = \left\{ \sum_{j=1}^{M} \sum_{i=1}^{N-1} h_i k_j |\mu v_{i,j}|^2 \right\}^{1/2},$$

$$|v|_{l_2(\partial_+\Omega^h)} = \left\{ \sum_{j=1}^{M} k_j |\mu_t v_{N-1,j}|^2 + \sum_{i=1}^{N-1} h_i |\mu_x v_{i,M}|^2 \right\}^{1/2},$$

$$|v|_{l_2(\partial_-\Omega^h)} = \left\{ \sum_{i=1}^{N-1} h_i |\mu_x v_{i,0}|^2 \right\}^{1/2},$$

$$|v_x|_{\hat{l}_2(\Omega^h)} = \left\{ \sum_{j=1}^{M} \sum_{i=1}^{N-1} \frac{h_i + h_{i+1}}{2} k_j |\hat{\mu}_t(v_x)_{i,j}|^2 + \frac{h_1}{4} \sum_{j=1}^{M} k_j |\hat{\mu}_t(v_x)_{0,j}|^2 \right\}^{1/2},$$

for all $v(x,t)$ for which the right hand sides are defined.

Remark 6.3.1 We note that these norms are seminorms on $L^2(\Omega)$. If $|\cdot|_{l_2(\Omega^h)}$ is restricted to the subspace of $\mathcal{U}_0^h$ defined by

$$\left\{ v \in \mathcal{U}_0^h : \; v(x,0) = 0 \;\; \text{for } 0 \le x \le 1 \right\},$$

then it is a norm. The first three of these seminorms are similar to those used in Süli [49]. The last seminorm is introduced here to deal with the diffusion term.

Define a projection $R : \mathcal{U}_0^h \to \mathcal{M}^h$ by

$$Rv = \begin{cases} \mu v_{i,j}, & \text{on } K_{i,j}, \text{ for } i = 1, \ldots, N-1 \text{ and } j = 1, \ldots, M, \\ 0, & \text{otherwise.} \end{cases}$$

The stability of the finite volume method (6.2.9) – (6.2.12) is proved by the following discrete Gårding inequality.

Theorem 6.3.1 *Assume that $\varepsilon \le a(h_{N-1}+h_N)$. Then for each $v \in \mathcal{U}_0^h$,*

$$\hat{B}(v,Rv) \ge \frac{\varepsilon}{2}|v_s|^2_{l_2(\Omega^h)} + b|v|^2_{l_2(\Omega^h)} + \frac{1}{4}\min\{a,r\}|v|^2_{l_2(\theta_+\Omega^h)} - \frac{r}{2}|v|^2_{l_2(\theta_-\Omega^h)}.$$

Proof. Recall the definition (6.2.11) of $\hat{B}(\cdot,\cdot)$. For each $v \in \mathcal{U}_0^h$,

$$\begin{aligned}\hat{B}(v,Rv) &= -\varepsilon\sum_{j=1}^{M}\sum_{i=1}^{N-1} k_j\,\mu v_{i,j}\,\{\hat{\mu}_t(v_s)_{i,j} - \hat{\mu}_t(v_s)_{i-1,j}\} \\ &\quad + (av_s + rv_t + bv, Rv) \\ &\equiv I_1 + I_2. \end{aligned} \tag{6.3.1}$$

Firstly, by summation by parts,

$$\begin{aligned} I_1 &= \varepsilon\sum_{j=1}^{M} k_j\Big\{ -\mu v_{N-1,j}\,\hat{\mu}_t(v_s)_{N-1,j} + \mu v_{1,j}\,\hat{\mu}_t(v_s)_{0,j} \\ &\qquad + \sum_{i=1}^{N-2}(\mu v_{i+1,j} - \mu v_{i,j})\hat{\mu}_t(v_s)_{i,j}\Big\}. \end{aligned} \tag{6.3.2}$$

Now

$$\begin{aligned} &-\mu v_{N-1,j}\,\hat{\mu}_t(v_s)_{N-1,j} \\ &\quad = (\mu v_{N,j} - \mu v_{N-1,j})\hat{\mu}_t(v_s)_{N-1,j} - \mu v_{N,j}\,\hat{\mu}_t(v_s)_{N-1,j} \\ &\quad = \frac{h_{N-1}+h_N}{2}|\hat{\mu}_t(v_s)_{N-1,j}|^2 - \frac{1}{2}\mu_t v_{N-1,j}\,\hat{\mu}_t(v_s)_{N-1,j}, \\ &\qquad\quad \text{using (6.2.12), (6.2.6) and } \mu_t v_{N,j} = 0, \\ &\quad \ge \frac{h_{N-1}+h_N}{4}|\hat{\mu}_t(v_s)_{N-1,j}|^2 - \frac{1}{4(h_{N-1}+h_N)}|\mu_t v_{N-1,j}|^2. \end{aligned} \tag{6.3.3}$$

Similarly, using (6.2.6) and $\mu_t v_{0,j} = 0$, we have

$$\begin{aligned} &\mu v_{1,j}\,\hat{\mu}_t(v_s)_{0,j} \\ &\quad = \frac{1}{2}\mu_t v_{1,j}\,\hat{\mu}_t(v_s)_{0,j} \end{aligned}$$

$$= \frac{h_1}{4}(\hat{\mu}_t(v_s)_{0,j} + \hat{\mu}_t(v_s)_{1,j})\hat{\mu}_t(v_s)_{0,j}, \quad \text{by (6.2.12)},$$
$$\geq \frac{h_1}{8}\left(|\hat{\mu}_t(v_s)_{0,j}|^2 - |\hat{\mu}_t(v_s)_{1,j}|^2\right). \tag{6.3.4}$$

Also from (6.2.12), for $i = 1, \ldots, N-2$, we obtain

$$(\mu v_{i+1,j} - \mu v_{i,j})\hat{\mu}_t(v_s)_{i,j} = \frac{h_i + h_{i+1}}{2}|\hat{\mu}_t(v_s)_{i,j}|^2. \tag{6.3.5}$$

Substituting (6.3.3) – (6.3.5) into (6.3.2) we get

$$\begin{aligned} I_1 \geq \varepsilon \sum_{j=1}^{M} k_j \Bigg\{ & \frac{h_{N-1} + h_N}{4} |\hat{\mu}_t(v_s)_{N-1,j}|^2 - \frac{1}{4(h_{N-1} + h_N)} |\mu_t v_{N-1,j}|^2 \\ & + \frac{h_1}{8}\left(|\hat{\mu}_t(v_s)_{0,j}|^2 - |\hat{\mu}_t(v_s)_{1,j}|^2\right) \\ & + \sum_{i=1}^{N-2} \frac{h_i + h_{i+1}}{2} |\hat{\mu}_t(v_s)_{i,j}|^2 \Bigg\}. \end{aligned} \tag{6.3.6}$$

Next,

$$\begin{aligned} I_2 &= \sum_{j=1}^{M}\sum_{i=1}^{N-1} h_i k_j \, \mu(av_s + rv_t + bv)_{i,j} \, \mu v_{i,j} \\ &= \sum_{j=1}^{M}\sum_{i=1}^{N-1} \{ak_j(\mu_t v_{i,j} - \mu_t v_{i-1,j}) + rh_i(\mu_s v_{i,j} - \mu_s v_{i,j-1}) + b\mu v_{i,j}\} \, \mu v_{i,j}, \\ &\qquad\qquad \text{by (6.2.7) and (6.2.8)}, \\ &= \frac{a}{2}\sum_{j=1}^{M} k_j \sum_{i=1}^{N-1}\left(|\mu_t v_{i,j}|^2 - |\mu_t v_{i-1,j}|^2\right) + \frac{r}{2}\sum_{i=1}^{N-1} h_i \sum_{j=1}^{M}\left(|\mu_s v_{i,j}|^2 - |\mu_s v_{i,j-1}|^2\right) \\ &\quad + b\sum_{j=1}^{M}\sum_{i=1}^{N-1} h_i k_j |\mu v_{i,j}|^2, \quad \text{using (6.2.5) and (6.2.6)}, \\ &= \frac{a}{2}\sum_{j=1}^{M} k_j |\mu_t v_{N-1,j}|^2 + \frac{r}{2}\sum_{i=1}^{N-1} h_i\left(|\mu_s v_{i,M}|^2 - |\mu_s v_{i,0}|^2\right) \\ &\quad + b\sum_{j=1}^{M}\sum_{i=1}^{N-1} h_i k_j |\mu v_{i,j}|^2, \end{aligned} \tag{6.3.7}$$

by telescoping and using $\mu_t v_{0,j} = 0$, for $j = 1, \ldots, M$.

Hence

$$\begin{aligned}\hat{B}(v, Rv) &= I_1 + I_2 \\ &\geq \frac{\varepsilon}{2}|v_x|^2_{l_2(\Omega^h)} + \left(\frac{a}{2} - \frac{\varepsilon}{4(h_{N-1}+h_N)}\right)\sum_{j=1}^{M} k_j|\mu_t v_{N-1,j}|^2 \\ &\quad + \frac{r}{2}\sum_{i=1}^{N-1} h_i|\mu_x v_{i,M}|^2 - \frac{r}{2}|v|^2_{l_2(\partial_-\Omega^h)} + b|v|^2_{l_2(\Omega^h)}.\end{aligned}$$

The desired result then follows from the assumption of the theorem. □

As a corollary we obtain the following stability result.

Theorem 6.3.2 *Assume that $\varepsilon \leq a(h_{N-1} + h_N)$. Then (6.2.9) - (6.2.12) has a unique solution $u^h \in \mathcal{U}_0^h$ and*

$$\begin{aligned}&\varepsilon \left|u^h_x\right|^2_{l_2(\Omega^h)} + \left|u^h\right|^2_{l_2(\Omega^h)} + \left|u^h\right|^2_{l_2(\partial_+\Omega^h)} \\ &\quad \leq C\left\{|f|^2_{l_2(\Omega^h)} + \left|u^0\right|^2_{l_2(\partial_-\Omega^h)}\right\}. \qquad (6.3.8)\end{aligned}$$

Proof. As the existence of a unique solution follows from (6.3.8) because we are dealing with a norm in this situation (cf. Remark 6.3.1), we only have to establish (6.3.8).

Taking $p = Ru^h$ in (6.2.9) and using the arithmetic-geometric inequality, we obtain

$$\begin{aligned}\hat{B}(u^h, Ru^h) &= (f, Ru^h) \\ &\leq \frac{b}{2}\left|u^h\right|^2_{l_2(\Omega^h)} + \frac{1}{2b}|f|^2_{l_2(\Omega^h)}.\end{aligned}$$

Now an appeal to Theorem 6.3.1 completes the proof. □

As another corollary of Theorem 6.3.1 we have the following global error bound.

Theorem 6.3.3 *Assume that $\varepsilon \le a(h_{N-1}+h_N)$. Let u^I be the interpolant from $\mathcal{U}_0^h$ to u. Then*

$$\begin{aligned}
&\varepsilon \left|\mu_t(u_x) - \hat{\mu}_t(u_x^h)\right|^2_{l_2^\circ(\Omega^h)} + \left|u - u^h\right|^2_{l_2(\Omega^h)} + \left|u - u^h\right|^2_{l_2(\partial_+\Omega^h)} \\
&\quad \le C\left\{\varepsilon \left|\mu_t(u_x) - \hat{\mu}_t(u_x^I)\right|^2_{l_2^\circ(\Omega^h)} + \left|a(u-u^I)_x + r(u-u^I)_t\right|^2_{l_2(\Omega^h)} \right.\\
&\qquad \left. + \left|u - u^I\right|^2_{l_2(\Omega^h)} + \left|u - u^I\right|^2_{l_2(\partial_-\Omega^h)}\right\},
\end{aligned}$$

where

$$\begin{aligned}
\left|\mu_t(u_x) - \hat{\mu}_t(u_x^h)\right|^2_{l_2^\circ(\Omega^h)} &= \sum_{j=1}^{M} k_j \sum_{i=1}^{N-1} \frac{h_i + h_{i+1}}{2} \left|\mu_t(u_x)_{i,j} - \hat{\mu}_t(u_x^h)_{i,j}\right|^2 \\
&\quad + \frac{h_1}{4}\sum_{j=1}^{M} k_j \left|\mu_t(u_x)_{0,j} - \hat{\mu}_t(u_x^h)_{0,j}\right|^2, \qquad (6.3.9)
\end{aligned}$$

and $\left|\mu_t(u_x) - \hat{\mu}_t(u_x^I)\right|^2_{l_2^\circ(\Omega^h)}$ is similarly defined.

Proof. Set

$$\xi = u^h - u^I, \quad \eta = u - u^I.$$

Then

$$u - u^h = \eta - \xi.$$

We begin by estimating ξ.

Applying Theorem 6.3.1 we obtan

$$\begin{aligned}
&\frac{\varepsilon}{2}|\xi_x|^2_{\hat{l}_2(\Omega^h)} + b|\xi|^2_{l_2(\Omega^h)} + \frac{1}{4}\min\{a, r\}|\xi|^2_{l_2(\partial_+\Omega^h)} \\
&\quad \le \hat{B}(\xi, R\xi) + \frac{r}{2}|\xi|^2_{l_2(\partial_-\Omega^h)}. \qquad (6.3.10)
\end{aligned}$$

Set

$$e_{i,j} = \mu_t(u_x)_{i,j} - \hat{\mu}_t(u_x^I)_{i,j}. \qquad (6.3.11)$$

From (6.1.1) – (6.1.3) and (6.2.9) – (6.2.12), we have

$$\hat{B}(\xi, R\xi) = -\varepsilon \sum_{j=1}^{M} \sum_{i=1}^{N-1} k_j \, \mu\xi_{i,j} \{e_{i,j} - e_{i-1,j}\} + (a\eta_s + r\eta_t + b\eta, R\xi) \tag{6.3.12}$$

and

$$|\xi|^2_{l_2(\partial_-\Omega^h)} = |\eta|^2_{l_2(\partial_-\Omega^h)}. \tag{6.3.13}$$

We estimate the term involving ε first.

$$\begin{aligned}
&-\varepsilon \sum_{j=1}^{M} \sum_{i=1}^{N-1} k_j \, \mu\xi_{i,j} \{e_{i,j} - e_{i-1,j}\} \\
&= \varepsilon \sum_{j=1}^{M} k_j \left\{ -\mu\xi_{N,j} \, e_{N-1,j} + \mu\xi_{1,j} \, e_{0,j} + \sum_{i=1}^{N-1} (\mu\xi_{i+1,j} - \mu\xi_{i,j}) \, e_{i,j} \right\}, \\
&\qquad \text{by summation by parts,} \\
&= \varepsilon \sum_{j=1}^{M} k_j \left\{ -\frac{1}{2}\mu_t\xi_{N-1,j} \, e_{N-1,j} + \frac{1}{2}\mu_t\xi_{1,j} \, e_{0,j} \right. \\
&\qquad \left. + \sum_{i=1}^{N-1} \frac{h_i + h_{i+1}}{2} \hat{\mu}_t(\xi_s)_{i,j} \, e_{i,j} \right\}, \\
&\qquad \text{by virtue of (6.2.12), (6.2.6) and } \mu_t\xi_{0,j} = \mu_t\xi_{N,j} = 0, \\
&= \varepsilon \sum_{j=1}^{M} k_j \left\{ -\frac{1}{2}\mu_t\xi_{N-1,j} \, e_{N-1,j} + \frac{h_1}{4} \left(\hat{\mu}_t(\xi_s)_{0,j} + \hat{\mu}_t(\xi_s)_{1,j}\right) e_{0,j} \right. \\
&\qquad \left. + \sum_{i=1}^{N-1} \frac{h_i + h_{i+1}}{2} \hat{\mu}_t(\xi_s)_{i,j} \, e_{i,j} \right\}, \quad \text{by (6.2.12),} \\
&\le \varepsilon \sum_{j=1}^{M} k_j \left\{ \frac{\min\{a, r\}}{8\varepsilon} |\mu_t\xi_{N-1,j}|^2 + C\varepsilon |e_{N-1,j}|^2 \right. \\
&\qquad + \frac{h_1}{16} |\hat{\mu}_t(\xi_s)_{0,j}|^2 + \frac{h_1}{16} |\hat{\mu}_t(\xi_s)_{1,j}|^2 + C h_1 |e_{0,j}|^2 \\
&\qquad \left. + \frac{1}{8} \sum_{i=1}^{N-1} \frac{h_i + h_{i+1}}{2} |\hat{\mu}_t(\xi_s)_{i,j}|^2 + 2 \sum_{i=1}^{N-1} \frac{h_i + h_{i+1}}{2} |e_{i,j}|^2 \right\} \\
&\le \frac{1}{8} \min\{a, r\} |\xi|^2_{l_2(\partial_+\Omega^h)} + \frac{\varepsilon}{4} |\xi_s|^2_{\hat{l}_2(\Omega^h)} + C\varepsilon \left|\mu_t(u_s) - \hat{\mu}_t(u_s^I)\right|^2_{l_2^*(\Omega^h)},
\end{aligned} \tag{6.3.14}$$

using $\varepsilon \le a(h_{N-1}+h_N)$.

As for the other term in (6.3.12), we have

$$(a\eta_x + r\eta_t + b\eta, R\xi) \le \frac{b}{2}|\xi|^2_{l_2(\Omega^k)} + C|a\eta_x + r\eta_t + b\eta|^2_{l_2(\Omega^k)}. \qquad (6.3.15)$$

Thus from (6.3.10), (6.3.12) – (6.3.15), it follows that

$$\frac{\varepsilon}{4}|\xi_x|^2_{l_2(\Omega^k)} + \frac{b}{2}|\xi|^2_{l_2(\Omega^k)} + \frac{1}{8}\min\{a,r\}|\xi|^2_{l_2(\partial_+\Omega^k)}$$
$$\le C\left\{\varepsilon\left|\mu_t(u_x) - \hat{\mu}_t(u_x^I)\right|^2_{l_2^0(\Omega^k)} + |a\eta_x + r\eta_t + b\eta|^2_{l_2(\Omega^k)} + |\eta|^2_{l_2(\partial_-\Omega^k)}\right\}.$$

This, together with the triangle inequality, yields the desired result. □

6.4 Local Error Analysis

In this section we present local convergence results for the cell vertex finite volume scheme (6.2.9) – (6.2.12).

We first derive various interpolation errors. Set

$$I = \{(i,j) : i = 1,\dots,N-1,\ j = 1,\dots,M\}.$$

We have

Lemma 6.4.1 *For any $(i,j)\in I$, assume that $u \in C^3(K_{i,j}\cup K_{i+1,j})$. Then*

$$|\mu\eta_{i,j}| + |\mu(a\eta_x + r\eta_t)_{i,j}| \le C(h_ik_j)^{-1/2}h^2|u|_{H^3(K_{i,j})}, \qquad (6.4.1)$$

$$|\mu_x\eta_{i,0}| \le Ch_i^{3/2}|u^0|_{H^3(x_{i-1},x_i)}, \qquad (6.4.2)$$

$$\left|\mu_t(u_x)_{i,j} - \hat{\mu}_t(u_x^I)_{i,j}\right| \le C\Big\{|h_i - h_{i+1}|\ \|u\|_{C^2(K_{i,j})}$$
$$+\ (h_{i+1}^2 + h_i^2 + k_j^2)\ \|u\|_{C^3(K_{i,j}\cup K_{i+1,j})}\Big\} \qquad (6.4.3)$$

with

$$\begin{aligned}\left|\mu_t(u_x)_{0,j} - \hat{\mu}_t(u_x^I)_{0,j}\right| \le C\Big\{&|h_1 - h_2|\ \|u\|_{C^2(\bar{K}_{1,j})} \\ &+ \left(h_2^2 + h_1^2 + k_j^2\right)\|u\|_{C^3(K_{1,j}\cup K_{2,j})}\Big\}, \qquad (6.4.4)\end{aligned}$$

where $|\cdot|_{H^3(D)}$ *denotes the usual seminorm on* $H^3(D)$.

Proof. The proof of (6.4.1) can be found in the proof of Theorem 4 in Morton and Süli [28]. In a similar manner one can prove (6.4.2).

We only have to prove (6.4.3) and (6.4.4). For $i \ge 1$, using (6.2.12),

$$\begin{aligned}\hat{\mu}_t(u_x^I)_{i,j} &= \frac{2}{h_i + h_{i+1}}\left(\mu u_{i+1,j}^I - \mu u_{i,j}^I\right) \\ &= \frac{1}{2(h_i + h_{i+1})}(u_{i+1,j} + u_{i+1,j-1} - u_{i-1,j} - u_{i-1,j-1}), \quad \text{by (6.2.4)}, \\ &= \frac{1}{2}\left(u_x(x_i,t_j) + u_x(x_i,t_{j-1})\right) + \frac{h_{i+1} - h_i}{4}\left(u_{xx}(x_i,t_j) + u_{xx}(x_i,t_{j-1})\right) \\ &\quad + \frac{1}{12(h_i + h_{i+1})}\left\{h_{i+1}^3\left(u_{xxx}(\theta_1,t_j) + u_{xxx}(\theta_2,t_{j-1})\right)\right. \\ &\quad \left. + h_i^3\left(u_{xxx}(\theta_3,t_j) + u_{xxx}(\theta_4,t_{j-1})\right)\right\},\end{aligned}$$

by a Taylor expansion, where

$$x_{i-1} < \theta_3, \theta_4 < x_i < \theta_1, \theta_2 < x_{i+1}.$$

Thus

$$\begin{aligned}&|\mu_t(u_x)_{i,j} - \hat{\mu}_t(u_x^I)_{i,j}| \\ &\quad \le \left|\frac{1}{k_j}\int_{t_{j-1}}^{t_j} u_x(x_i,t)\,dt - \frac{1}{2}\left(u_x(x_i,t_j) + u_x(x_i,t_{j-1})\right)\right| \\ &\qquad + C|h_{i+1} - h_i|\ \|u\|_{C^2(\bar{K}_{i,j})} + C\left(h_{i+1}^2 + h_i^2\right)\|u\|_{C^3(K_{i,j}\cup K_{i+1,j})} \\ &\quad \le C|h_{i+1} - h_i|\ \|u\|_{C^2(\bar{K}_{i,j})} + C\left(h_{i+1}^2 + h_i^2 + k_j^2\right)\|u\|_{C^3(K_{i,j}\cup K_{i+1,j})},\end{aligned}$$

by the error estimate for the trapezoidal rule.

Similarly, for $i = 0$, by a Taylor expansion about $x = 0$ and using $u_{0,j} = 0$ for all j,

$$\begin{aligned}
\hat{\mu}_t(u_x^I)_{0,j} &= \frac{2}{h_1}\mu_t u_{1,j}^I - \mu_t(u_x^I)_{1,j} \\
&= \frac{1}{h_1}(u_{1,j} + u_{1,j-1}) - \frac{1}{2(h_1+h_2)}(u_{2,j} + u_{2,j-1}) \\
&= u_x(0,t_j) + u_x(0,t_{j-1}) + \frac{h_1}{2}\left(u_{xx}(0,t_j) + u_{xx}(0,t_{j-1})\right) \\
&\quad + \frac{h_1^2}{6}\left(u_{xxx}(\theta_5,t_j) + u_{xxx}(\theta_6,t_{j-1})\right) \\
&\quad - \frac{1}{2}\left(u_x(0,t_j) + u_x(0,t_{j-1})\right) - \frac{h_1+h_2}{4}\left(u_{xx}(0,t_j) + u_{xx}(0,t_{j-1})\right) \\
&\quad - \frac{(h_1+h_2)^2}{12}\left(u_{xxx}(\theta_7,t_j) + u_{xxx}(\theta_8,t_{j-1})\right), \\
&= \frac{1}{2}\left(u_x(0,t_j) + u_x(0,t_{j-1})\right) + \frac{h_1-h_2}{4}\left(u_{xx}(0,t_j) + u_{xx}(0,t_{j-1})\right) \\
&\quad + \frac{h_1^2}{6}\left(u_{xxx}(\theta_5,t_j) + u_{xxx}(\theta_6,t_{j-1})\right) \\
&\quad - \frac{(h_1+h_2)^2}{12}\left(u_{xxx}(\theta_7,t_j) + u_{xxx}(\theta_8,t_{j-1})\right),
\end{aligned}$$

where

$$0 < \theta_l < x_2 \quad \text{for } l = 5, \ldots, 8.$$

Hence

$$\begin{aligned}
&|\mu_t(u_x)_{0,j} - \hat{\mu}_t(u_x^I)_{0,j}| \\
&\quad \le \left| \frac{1}{k_j}\int_{t_{j-1}}^{t_j} u_x(0,t)\,dt - \frac{1}{2}\left(u_x(0,t_j) + u_x(0,t_{j-1})\right) \right| \\
&\qquad + C|h_1 - h_2|\, \|u\|_{C^2(K_{1,j})} + C\left(h_2^2 + h_1^2\right) \|u\|_{C^3(K_{1,j}\cup K_{2,j})}, \\
&\quad \le C|h_1 - h_2|\, \|u\|_{C^2(K_{1,j})} + C\left(h_2^2 + h_1^2 + k_j^2\right) \|u\|_{C^3(K_{1,j}\cup K_{2,j})},
\end{aligned}$$

which completes the proof of (6.4.4). □

For each $(i,j) \in I$, define

$$B_{i,j}(w) = \frac{\varepsilon}{h_i}(\hat{\mu}_t(w_s)_{i,j} - \hat{\mu}_t(w_s)_{i-1,j}) + \mu(aw_s + rw_t + bw)_{i,j}, \tag{6.4.5}$$

for all w for which the right hand side is defined. Then

$$\hat{B}(w,p) = \sum_{j=1}^{M} \sum_{i=1}^{N-1} h_i k_j \, \mu p_{i,j} \, B_{i,j}(w). \tag{6.4.6}$$

For $\tilde{I}$ any nonempty subset of I, let $\tilde{\Omega} = \bigcup_{(i,j)\in\tilde{I}} K_{i,j}$. Set

$$|w|_{\hat{B}(\tilde{\Omega})} = \left\{ \sum_{(i,j)\in\tilde{I}} h_i k_j |B_{i,j}(w)|^2 \right\}^{1/2}. \tag{6.4.7}$$

We note that $|\cdot|_{\hat{B}(\Omega^h)}$ is a seminorm on $L^2(\Omega)$. It can be regarded as a generalization of the seminorm $|\nabla \cdot I^h(a(\cdot))|_{l_2(\Omega^h)}$ introduced in Morton and Stynes [27].

Using Lemma 6.4.1 we get the following error bound in a local $|\cdot|_{\hat{B}(\Omega^h)}$ seminorm.

Theorem 6.4.1 *Let $\tilde{\Omega} = \bigcup_{(i,j)\in\tilde{I}} K_{i,j}$ be arbitrary. Set*

$$\tilde{I}^+ = \left\{(i,j) \in I : \ j = j' \text{ and } |i - i'| \leq 1 \quad \text{for some } (i',j') \in \tilde{I}\right\}.$$

Let $\tilde{\Omega}^+ = \bigcup_{(i,j)\in\tilde{I}^+} K_{i,j}$. Assume that $u \in C^3(\tilde{\Omega}^+)$. Then

$$\begin{aligned}
&\left|u^h - u^I\right|_{\hat{B}(\tilde{\Omega})} \\
&\quad \leq C\varepsilon \left\{ \max_{(i,j)\in\tilde{I}} \{h_i^{-1}|h_{i+1} - h_i|, \ h_i^{-1}|h_i - h_{i-1}|\} \|u\|_{C^2(\tilde{\Omega}^+)} \right. \\
&\qquad \left. + \max_{(i,j)\in\tilde{I}} \{h_{i+1}^2 h_i^{-1}, \ h_i, \ k_j^2 h_i^{-1}\} \|u\|_{C^3(\tilde{\Omega}^+)} \right\} + Ch^2|u|_{H^3(\tilde{\Omega})}.
\end{aligned} \tag{6.4.8}$$

Proof. From (6.1.1), (6.2.9) and (6.2.11), we get

$$\begin{aligned}
\hat{B}\left(u^h - u^I, p\right) = {} & -\varepsilon \sum_{j=1}^{M} \sum_{i=1}^{N-1} k_j \, \mu p_{i,j}\{e_{i,j} - e_{i-1,j}\} \\
& + (a\eta_s + r\eta_t + b\eta, p), \quad \forall p \in \mathcal{M}^h,
\end{aligned} \tag{6.4.9}$$

where $e_{i,j}$ is as in (6.3.11).

Fix $(i,j) \in \tilde{I}$. Take p in (6.4.9) to be the characteristic function of $K_{i,j}$. From (6.4.6) this yields

$$B_{i,j}\left(u^h - u^I\right) = -\varepsilon h_i^{-1}\{e_{i,j} - e_{i-1,j}\} + \mu(a\eta_x + r\eta_t + b\eta)_{i,j}. \tag{6.4.10}$$

Applying Lemma 6.4.1 gives

$$\begin{aligned}\left|B_{i,j}\left(u^h - u^I\right)\right| \le\ & C\varepsilon h_i^{-1}\Big\{(|h_{i+1} - h_i| + |h_i - h_{i-1}|)\,\|u\|_{C^2(K_{i-1,j}\cup K_{i,j})} \\ &+ \left(h_{i+1}^2 + h_i^2 + k_j^2\right)\|u\|_{C^3(K_{i-1,j}\cup K_{i,j}\cup K_{i+1,j})}\Big\} \\ &+ C(h_ik_j)^{-1/2}h^2|u|_{H^3(K_{i,j})}.\end{aligned}$$

The desired inequality follows immediately from the definition (6.4.7). □

In what follows, we will derive a local error bound in an energy seminorm. To this end, we introduce a cut-off function $w(x,t)$ defined by

$$\omega(x,t) = g\left(\frac{x - x^*}{\gamma h}\right)g(t - t^*),$$

where (x^*,t^*) is a fixed node, $\gamma \ge 1$ is a constant (which we choose later to be independent of ε and the mesh), and

$$g(r) = \frac{2}{1 + \exp(r)} \quad \forall r \in (-\infty,\infty).$$

Set

$$\Omega_0 = \{(x,t) \in \Omega : x \le x^*,\ t \le t^*\}, \tag{6.4.11}$$

$$\Omega_0^+ = \left\{(x,t) \in \Omega : x \le x^* + s\gamma h \ln\frac{1}{h},\ t \le t^* + s\gamma h \ln\frac{1}{h}\right\}, \tag{6.4.12}$$

where $s > 0$ is some integer (which we choose later to be independent of ε and h). Without losing generality, we assume that Ω_0^+ consists of cells, that is,

$$\Omega_0^+ = \bigcup_{j=1}^{j'} \bigcup_{i=1}^{i'} K_{i,j},$$

for some $(i', j') \in I$. Set

$$\Omega_0^{++} = \bigcup_{j=1}^{j'} \bigcup_{i=1}^{i'+1} K_{i,j}. \tag{6.4.13}$$

Similarly to the previous chapters, one can easily show that

$$\omega_x < 0 \quad \text{and} \quad \omega_t < 0 \quad \text{on } \Omega, \tag{6.4.14}$$

$$\max_{D_{i,j}} \omega / \min_{D_{i,j}} \omega \le C, \quad \max_{D_{i,j}} |\omega_x| / \min_{D_{i,j}} |\omega_x| \le C, \tag{6.4.15}$$

with $D_{i,j} = \overline{K_{i,j} \cup K_{i+1,j}}$,

$$|\omega_x| \le C\gamma^{-1} h^{-1} \omega, \tag{6.4.16}$$

$$\omega(x_{i'}, t) \le C h^s, \quad \text{for } t \in [0, T], \tag{6.4.17}$$

$$\omega(x, t) \ge 1 \quad \text{on } \Omega_0. \tag{6.4.18}$$

Notation. We introduce the following weighted norms:

$$|v|_{l_2(\Omega_0^+),\omega} = \left\{ \sum_{j=1}^{j'} \sum_{i=1}^{i'} h_i k_j \, \omega_{i,j} |\mu v_{i,j}|^2 \right\}^{1/2},$$

$$|v|_{l_2(\theta - \Omega_0^+),\omega} = \left\{ \sum_{i=1}^{i'} h_i \, \omega_{i,1} |\mu_x v_{i,0}|^2 \right\}^{1/2},$$

$$|v_x|_{\hat{l}_2(\Omega_0^+),\omega} = \left\{ \sum_{j=1}^{j'} \sum_{i=1}^{i'} \frac{h_i + h_{i+1}}{2} \, k_j \, \omega_{i,j} |\hat{\mu}_t(v_x)_{i,j}|^2 \right.$$
$$\left. + \frac{h_1}{4} \sum_{j=1}^{j'} k_j \, \omega_{1,j} |\hat{\mu}_t(v_x)_{0,j}|^2 \right\}^{1/2},$$

for all $v(x,t)$ for which the right hand sides are defined.

Define $R_\omega : \mathcal{U}_0^h \to \mathcal{M}^h$ by

$$R_\omega v = \begin{cases} \omega_{i,j}\ \mu v_{i,j}, & \text{on } K_{i,j},\ \text{for } i = 1,\ldots,i',\ j = 1,\ldots,j', \\ 0, & \text{otherwise.} \end{cases}$$

Then we can prove a weighted Gårding ineqzality.

Lemma 6.4.2 *Assume that there exists a positive constant c_0, which is independent of ε and of the mesh, such that*

$$\varepsilon \le c_0 h_i \quad \textit{for } i = 1,\ldots,i'+1. \tag{6.4.19}$$

Then

$$\begin{aligned}\hat{B}(v, R_\omega v) \ge\ & \frac{\varepsilon}{4}|v_x|^2_{l_2(\Omega_0^+),\omega} + b|v|^2_{l_2(\Omega_0^+),\omega} \\ & - \frac{r}{2}|v|^2_{l_2(\theta_-\Omega_0^+),\omega} - Ch^s h_{i'+1}^{-1}|v|^2_{l_2(\Omega_0^{++})},\end{aligned}$$

where $|\cdot|_{l_2(\Omega_0^{++})}$ is defined analogously to $|\cdot|_{l_2(\Omega^h)}$.

Proof. Similarly to the derivation of (6.3.7), we have

$$\begin{aligned}
&(av_x + rv_t + bv, R_\omega v) \\
&= \frac{a}{2}\sum_{j=1}^{j'}\sum_{i=1}^{i'} k_j\,\omega_{i,j}\left(|\mu_t v_{i,j}|^2 - |\mu_t v_{i-1,j}|^2\right) \\
&\quad + \frac{r}{2}\sum_{j=1}^{j'}\sum_{i=1}^{i'} h_i\,\omega_{i,j}\left(|\mu_x v_{i,j}|^2 - |\mu_x v_{i,j-1}|^2\right) \\
&\quad + b\sum_{j=1}^{j'}\sum_{i=1}^{i'} h_i k_j\,\omega_{i,j}\,|\mu v_{i,j}|^2 \\
&= \frac{a}{2}\sum_{j=1}^{j'} k_j\left\{\sum_{i=1}^{i'}(\omega_{i,j} - \omega_{i+1,j})|\mu_t v_{i,j}|^2 + \omega_{i'+1,j}|\mu_t v_{i',j}|^2\right\}
\end{aligned}$$

$$+\frac{r}{2}\sum_{i=1}^{i'} h_i \left\{ \sum_{j=1}^{j'} (\omega_{i,j}-\omega_{i,j+1})|\mu_s v_{i,j}|^2 + \omega_{i,j'+1}|\mu_s v_{i,j'}|^2 - \omega_{i,1}|\mu_s v_{i,0}|^2 \right\}$$

$$+ b\sum_{j=1}^{j'}\sum_{i=1}^{i'} h_i k_j \,\omega_{i,j}\,|\mu v_{i,j}|^2, \tag{6.4.20}$$

by summation by parts and using $\mu_t v_{0,j} = 0$.

Next, the contribution to $\hat{B}(v, R_\omega v)$ from the diffusion term is

$$\begin{aligned}
Q_\varepsilon &\equiv -\varepsilon \sum_{j=1}^{j'}\sum_{i=1}^{i'} k_j\,\omega_{i,j}\,\mu v_{i,j}\{\hat{\mu}_t(v_s)_{i,j} - \hat{\mu}_t(v_s)_{i-1,j}\} \\
&= \varepsilon \sum_{j=1}^{j'} k_j \Big\{ -\omega_{i',j}\,\mu v_{i',j}\,\hat{\mu}_t(v_s)_{i',j} + \omega_{1,j}\,\mu v_{1,j}\,\hat{\mu}_t(v_s)_{0,j} \\
&\qquad - \sum_{i=1}^{i'-1}(\omega_{i,j}\,\mu v_{i,j} - \omega_{i+1,j}\,\mu v_{i+1,j})\hat{\mu}_t(v_s)_{i,j}\Big\}.
\end{aligned} \tag{6.4.21}$$

Now, analogously to (6.3.3), we have

$$\begin{aligned}
&-\mu v_{i',j}\,\hat{\mu}_t(v_s)_{i',j} \\
&\qquad \geq \frac{h_{i'}+h_{i'+1}}{4}\,|\hat{\mu}_t(v_s)_{i',j}|^2 - \frac{1}{h_{i'}+h_{i'+1}}\,|\mu v_{i'+1,j}|^2.
\end{aligned} \tag{6.4.22}$$

For $i = 1,\ldots,i'-1$,

$$\begin{aligned}
&-(\omega_{i,j}\,\mu v_{i,j} - \omega_{i+1,j}\,\mu v_{i+1,j})\hat{\mu}_t(v_s)_{i,j} \\
&\quad = \{\omega_{i,j}\,(\mu v_{i+1,j} - \mu v_{i,j}) - (\omega_{i,j} - \omega_{i+1,j})\mu v_{i+1,j}\}\,\hat{\mu}_t(v_s)_{i,j} \\
&\quad = \frac{h_i + h_{i+1}}{2}\,\omega_{i,j}\,|\hat{\mu}_t(v_s)_{i,j}|^2 - (\omega_{i,j} - \omega_{i+1,j})\mu v_{i+1,j}\,\hat{\mu}_t(v_s)_{i,j}, \quad \text{by (6.2.12)}, \\
&\quad \geq \frac{h_i + h_{i+1}}{4}\,\omega_{i,j}\,|\hat{\mu}_t(v_s)_{i,j}|^2 - \frac{1}{h_i + h_{i+1}}(\omega_{i,j} - \omega_{i+1,j})^2\omega_{i,j}^{-1}|\mu v_{i+1,j}|^2.
\end{aligned} \tag{6.4.23}$$

Substituting (6.4.22), (6.3.4) and (6.4.23) into (6.4.21) , we obtain

$$Q_\varepsilon \geq \varepsilon \sum_{j=1}^{j'} k_j \left\{ \frac{h_{i'}+h_{i'+1}}{4}\,\omega_{i',j}\,|\hat{\mu}_t(v_s)_{i',j}|^2 - \frac{1}{h_{i'}+h_{i'+1}}\,\omega_{i',j}\,|\mu v_{i'+1,j}|^2 \right.$$

$$+\frac{h_1}{8}\omega_{1,j}(|\hat{\mu}_t(v_s)_{0,j}|^2 - |\hat{\mu}_t(v_s)_{1,j}|^2) + \sum_{i=1}^{i'-1}\frac{h_i+h_{i+1}}{4}\,\omega_{i,j}\,|\hat{\mu}_t(v_s)_{i,j}|^2$$

$$-\sum_{i=1}^{i'-1}\frac{1}{h_i+h_{i+1}}(\omega_{i,j}-\omega_{i+1,j})^2\,\omega_{i,j}^{-1}\,|\mu v_{i+1,j}|^2\Bigg\}$$

$$\ge \frac{\varepsilon}{4}|v_s|^2_{l_2(\Omega_0^+),\omega} - \frac{\varepsilon}{h_{i'}+h_{i'+1}}\sum_{j=1}^{j'}k_j\,\omega_{i',j}\,|\mu v_{i'+1,j}|^2$$

$$-\varepsilon\sum_{j=1}^{j'}k_j\sum_{i=1}^{i'-1}\frac{1}{h_i+h_{i+1}}(\omega_{i,j}-\omega_{i+1,j})^2\,\omega_{i,j}^{-1}\,|\mu v_{i+1,j}|^2. \tag{6.4.24}$$

In (6.4.24), the second term can be bounded by using (6.4.17), to get

$$\begin{aligned}
&\frac{\varepsilon}{h_{i'}+h_{i'+1}}\sum_{j=1}^{j'}k_j\,\omega_{i',j}\,|\mu v_{i'+1,j}|^2\\
&\qquad\le C\frac{\varepsilon}{h_{i'}+h_{i'+1}}h^s\sum_{j=1}^{j'}k_j\,|\mu v_{i'+1,j}|^2\\
&\qquad\le Ch^s h_{i'+1}^{-1}|v|^2_{l_2(\Omega_0^{++})},
\end{aligned}\tag{6.4.25}$$

using (6.4.19).

As for the last term in (6.4.24), using (6.2.6) and $(a+b)^2 \le 2(a^2+b^2)$, we get

$$\begin{aligned}
&\sum_{i=1}^{i'-1}\frac{1}{h_i+h_{i+1}}(\omega_{i,j}-\omega_{i+1,j})^2\omega_{i,j}^{-1}|\mu v_{i+1,j}|^2\\
&\quad\le \frac{1}{2}\sum_{i=1}^{i'-1}\frac{1}{h_i+h_{i+1}}\,h_{i+1}^2\max_{K_{i+1,j}}|\omega_s|^2\,\omega_{i,j}^{-1}(|\mu_t v_{i,j}|^2 + |\mu_t v_{i+1,j}|^2)\\
&\quad\le C\sum_{i=1}^{i'-1}\frac{h_{i+1}^2}{h_i+h_{i+1}}\gamma^{-1}h^{-1}\max_{K_{i+1,j}}|\omega_s|(|\mu_t v_{i,j}|^2 + |\mu_t v_{i+1,j}|^2),\\
&\qquad\qquad\text{using (6.4.16) and (6.4.15),}\\
&\quad\le C\gamma^{-1}\sum_{i=1}^{i'-1}\max_{K_{i+1,j}}|\omega_s|(|\mu_t v_{i,j}|^2 + |\mu_t v_{i+1,j}|^2)\\
&\quad\le C\gamma^{-1}\sum_{i=1}^{i'}\left(\max_{K_{i+1,j}}|\omega_s| + \max_{K_{i,j}}|\omega_s|\right)|\mu_t v_{i,j}|^2
\end{aligned}$$

$$\le C\gamma^{-1}\sum_{i=1}^{i'}\left(\max_{K_{i,j}\cup K_{i+1,j}}|\omega_x| \Big/ \min_{K_{i+1,j}}|\omega_x|\right)\frac{1}{h_{i+1}}(\omega_{i,j}-\omega_{i+1,j})|\mu_t v_{i,j}|^2,$$

since (6.4.14) implies that $\omega_{i,j}-\omega_{i+1,j}>0$,

$$\le C\gamma^{-1}\sum_{i=1}^{i'}\frac{1}{h_{i+1}}(\omega_{i,j}-\omega_{i+1,j})|\mu_t v_{i,j}|^2, \quad \text{using (6.4.15)},$$

$$\le \frac{a}{2c_0}\sum_{i=1}^{i'}\frac{1}{h_{i+1}}(\omega_{i,j}-\omega_{i+1,j})\,|\mu_t v_{i,j}|^2, \tag{6.4.26}$$

on choosing γ sufficiently large, independently of ε and of the mesh used.

Thus, from (6.4.24) – (6.4.26) and (6.4.19), we obtain

$$Q_\varepsilon \ge \frac{\varepsilon}{4}|v_x|^2_{l_2(\Omega_0^+),\omega}-\frac{a}{2}\sum_{j=1}^{j'}k_j\sum_{i=1}^{i'}(\omega_{i,j}-\omega_{i+1,j})|\mu_t v_{i,j}|^2 - Ch^s h_{i'+1}^{-1}|v|^2_{l_2(\Omega_0^{++})}.$$

Combine this with (6.4.20) and use (6.4.14) to complete the proof. □

We now prove the main result of this section.

Theorem 6.4.2 *Assume that $\varepsilon \le a(h_{N-1}+h_N)$ and that (6.4.19) holds. If $u \in C^3(\Omega_0^{++})$, then*

$$\left|u-u^h\right|_{E(\Omega_0)} \le Ch\left(\|u\|_{C^3(\Omega_0^{++})}+|u^0|_{H^3(0,x_{i'})}\right) + C\left(h^s h_{i'+1}^{-1}\right)^{1/2}\left(\|f\|_{L^2(\Omega)}+\left\|u^0\right\|_{L^2(0,1)}+|u|_{H^3(\Omega_0^{++})}\right),$$

where Ω_0 and Ω_0^{++} are as in (6.4.11) and (6.4.13) respectively, s is as in (6.4.12), and

$$\left|u-u^h\right|_{E(\Omega_0)} = \left\{\varepsilon\left|\mu_t(u_x)-\hat{\mu}_t(u_x^h)\right|^2_{l_2^\circ(\Omega_0)}+\left|u-u^h\right|^2_{l_2(\Omega_0)}\right\}^{1/2}$$

with $|\mu_t(u_x)-\hat{\mu}_t(u_x^h)|_{l_2^\circ(\Omega_0)}$ and $|\cdot|_{l_2(\Omega_0)}$ defined similarly to (6.3.9) and $|\cdot|_{l_2(\Omega^h)}$.

Proof. Applying Lemma 6.4.2, we obtain

$$\frac{\varepsilon}{4}|\xi_x|^2_{l_2(\Omega_0^+),\omega}+b|\xi|^2_{l_2(\Omega_0^+),\omega} \le \hat{B}(\xi,R_\omega\xi)+\frac{r}{2}|\xi|^2_{l_2(\theta_-\Omega_0^+),\omega}+Ch^s h_{i'+1}^{-1}|\xi|^2_{l_2(\Omega_0^{++})}. \tag{6.4.27}$$

Recalling (6.4.6),

$$\hat{B}(\xi, R_\omega \xi) = \sum_{j=1}^{j'} \sum_{i=1}^{i'} h_i k_j \, \omega_{i,j} \, \mu \xi_{i,j} B_{i,j}(\xi)$$
$$\leq \frac{b}{2} |\xi|^2_{l_2(\Omega_0^+),\omega} + \frac{1}{2b} \sum_{j=1}^{j'} \sum_{i=1}^{i'} h_i k_j \, \omega_{i,j} \, |B_{i,j}(\xi)|^2. \qquad (6.4.28)$$

Put this into (6.4.27) to get

$$\frac{\varepsilon}{2} |\xi_x|^2_{\bar{l}_2(\Omega_0^+),\omega} + b|\xi|^2_{l_2(\Omega_0^+),\omega}$$
$$\leq \frac{1}{b} \sum_{j=1}^{j'} \sum_{i=1}^{i'} h_i k_j \, \omega_{i,j} \, |B_{i,j}(\xi)|^2 + r|\xi|^2_{l_2(\partial_- \Omega_0^+),\omega} + Ch^s h_{i'+1}^{-1} |\xi|^2_{l_2(\Omega_0^{++})}$$
$$\leq C \left\{ |\xi|^2_{\hat{B}(\Omega_0^+)} + |\xi|^2_{l_2(\partial_- \Omega_0^+)} \right\} + Ch^s h_{i'+1}^{-1} |\xi|^2_{l_2(\Omega_0^{++})}, \qquad (6.4.29)$$

since $\omega(x,t) \leq 2$ on $\bar{\Omega}$, where $|\cdot|_{l_2(\partial_- \Omega_0^+)}$ is defined similarly to $|\cdot|_{l_2(\partial_- \Omega^h)}$.

Appealing to Theorem 6.4.1, we obtain

$$|\xi|_{\hat{B}(\Omega_0^+)} \leq C\varepsilon \max_{1 \leq i \leq i'} \{h_i^{-1}\} \left\{ \max_{1 \leq i \leq i'} \{|h_{i+1} - h_i|\} \|u\|_{C^2(\Omega_0^+)} + h^2 \|u\|_{C^3(\Omega_0^{++})} \right\}$$
$$+ Ch^2 |u|_{H^3(\Omega_0^+)}$$
$$\leq Ch \|u\|_{C^3(\Omega_0^{++})}, \qquad (6.4.30)$$

using (6.4.19).

By (6.2.10),

$$|\xi|_{l_2(\partial_- \Omega_0^+)} = |\eta|_{l_2(\partial_- \Omega_0^+)} \leq Ch^2 \left| u^0 \right|_{H^3(0, x_{i'})}, \qquad (6.4.31)$$

by virtue of (6.4.2).

For the last term in (6.4.29), we have

$$|\xi|_{l_2(\Omega_0^{++})} \leq \left| u^h \right|_{l_2(\Omega_0^{++})} + |u|_{l_2(\Omega_0^{++})} + |\eta|_{l_2(\Omega_0^{++})}$$
$$\leq C \left(|f|_{l_2(\Omega^h)} + \left| u^0 \right|_{l_2(\partial_- \Omega^h)} + |u|_{l_2(\Omega_0^{++})} \right) + Ch^2 |u|_{H^3(\Omega_0^{++})},$$

according to Theorem 6.3.2 and (6.4.1).

Clearly, for any $w \in L^2(\Omega)$,

$$|w|_{l_2(\Omega_0^{++})} \leq |w|_{l_2(\Omega^h)} \leq \|w\|_{L_2(\Omega)},$$

by Cauchy-Schwarz inequality. Also

$$\|u\|_{L^2(\Omega)} \leq C\left(\|f\|_{L^2(\Omega)} + \|u^0\|_{L^2(0,1)}\right).$$

Thus

$$|\xi|_{l_2(\Omega_0^{++})} \leq C\left(\|f\|_{L^2(\Omega)} + \|u^0\|_{L^2(0,1)}\right) + Ch^2|u|_{H^3(\Omega_0^{++})}. \tag{6.4.32}$$

Collecting (6.4.30) – (6.4.32) into (6.4.29) yields

$$\begin{aligned}
&\left\{\varepsilon|\xi_s|^2_{l_2(\Omega_0^+),\omega} + |\xi|^2_{l_2(\Omega_0^+),\omega}\right\}^{1/2} \\
&\quad \leq Ch\|u\|_{C^3(\Omega_0^{++})} + Ch^2|u^0|_{H^3(0,s_{i'})} \\
&\qquad + C\left(h^s h_{i'+1}^{-1}\right)^{1/2}\left(\|f\|_{L^2(\Omega)} + \|u^0\|_{L^2(0,1)} + |u|_{H^3(\Omega_0^{++})}\right). \qquad (6.4.33)
\end{aligned}$$

Note that by (6.4.18),

$$\varepsilon|\xi_s|^2_{l_2(\Omega_0^+)} + |\xi|^2_{l_2(\Omega_0)} \leq \varepsilon|\xi_s|^2_{l_2(\Omega_0^+),\omega} + |\xi|^2_{l_2(\Omega_0^+),\omega}. \tag{6.4.34}$$

Combining (6.4.33) with (6.4.34), invoking the triangle inequality and using Lemma 6.4.1 we obtain the desired result. □

Remark 6.4.1 The assumption that $u \in C^3(\Omega_0^{++})$ in Theorem 6.4.2 can be guaranteed if the data is sufficiently smooth and satisfies certain compatibility conditions at the corner $(0,0)$ of Ω.

Corollary 6.4.1 *Assume that the hypotheses of Theorem 6.4.2 hold and*

$$h_{i+1} = h_i + O(h^2), \quad \textit{for } i = 1, \ldots, i' \tag{6.4.35}$$

and

$$h_{i'+1} = O(h^\kappa) \quad \textit{for some } \kappa > 0. \tag{6.4.36}$$

Then

$$\left|u - u^h\right|_{E(\Omega_0)} \leq Ch^2 \left(\|u\|_{C^3(\Omega_0^{++})} + |u^0|_{H^3(0,s_{i'})} + \|f\|_{L^2(\Omega)} + \left\|u^0\right\|_{L^2(0,1)}\right).$$

Proof. By inspecting the proof of Theorem 6.4.2, we see (cf. (6.4.30)) that when (6.4.35) holds,

$$|\xi|_{B(\Omega_0^+)} \leq Ch^2\|u\|_{C^3(\Omega_0^{++})}. \tag{6.4.37}$$

Hence from (6.4.29), (6.4.37), (6.4.31) and (6.4.32) we have

$$\begin{aligned}
&\left\{\varepsilon|\xi_s|^2_{l_2(\Omega_0^+),\omega} + |\xi|^2_{l_2(\Omega_0^+),\omega}\right\}^{1/2} \\
&\quad\leq Ch^2 \left(\|u\|_{C^3(\Omega_0^{++})} + |u^0|_{H^3(0,s_{i'})}\right) \\
&\qquad + C\left(h^s h_{i'+1}^{-1}\right)^{1/2} \left(\|f\|_{L^2(\Omega)} + \left\|u^0\right\|_{L^2(0,1)} + |u|_{H^3(\Omega_0^{++})}\right). \quad (6.4.38)
\end{aligned}$$

Since (6.4.36) implies $h_{i'+1}^{-1} = O(h^{-\kappa})$, we take $s = \kappa + 4$ in (6.4.38). Now arguments exactly the same as in the proof of Theorem 6.4.2 lead to the desired result. □

Corollary 6.4.2 *Assume that the hypotheses of Theorem 6.4.2 hold and that there exists a positive constant c_1, which is independent of ε and of the mesh, such that*

$$\varepsilon \leq c_1 h_i^2, \quad \textit{for } i = 1, \ldots, i' + 1. \tag{6.4.39}$$

Then

$$\left|u - u^h\right|_{E(\Omega_0)} \leq Ch^2 \left(\|u\|_{C^3(\Omega_0^{++})} + |u^0|_{H^3(0,s_{i'})} + \|f\|_{L^2(\Omega)} + \left\|u^0\right\|_{L^2(0,1)}\right).$$

Proof. From the proof of Lemma 6.4.2, we see that when (6.4.39) holds, one can get (cf. (6.4.25)), for each $v \in \mathcal{U}_0^h$,

$$\hat{B}(v, R_\omega v) \geq \frac{\varepsilon}{4}|v_x|^2_{l_2(\Omega_0^+),\omega} + b|v|^2_{l_2(\Omega_0^+),\omega} - \frac{r}{2}|v|^2_{l_2(\partial_-\Omega_0^+),\omega} - Ch^s|v|^2_{l_2(\Omega_0^{++})}.$$

Hence, similarly to the derivation of (6.4.29), using (6.4.28),

$$\begin{aligned}
&\frac{\varepsilon}{2}|\xi_x|^2_{l_2(\Omega_0^+),\omega} + b|\xi|^2_{l_2(\Omega_0^+),\omega} \\
&\leq \frac{1}{b}\sum_{j=1}^{j'}\sum_{i=1}^{i'} h_i k_j \omega_{i,j}\,|B_{i,j}(\xi)|^2 + r|\xi|^2_{l_2(\partial_-\Omega_0^+),\omega} + Ch^s|\xi|^2_{l_2(\Omega_0^{++})} \\
&\leq C\left\{|\xi|^2_{\hat{B}(\Omega_0^+)} + |\xi|^2_{l_2(\partial_-\Omega_0^+)}\right\} + Ch^s|\xi|^2_{l_2(\Omega_0^{++})}. \qquad (6.4.40)
\end{aligned}$$

Also, from (6.4.30) we see that (6.4.39) implies (6.4.37). It then follows from (6.4.40), (6.4.37), (6.4.31) and (6.4.32) that

$$\begin{aligned}
&\left\{\varepsilon|\xi_x|^2_{l_2(\Omega_0^+),\omega} + |\xi|^2_{l_2(\Omega_0^+),\omega}\right\}^{1/2} \\
&\quad\leq Ch^2\left(\|u\|_{C^3(\Omega_0^{++})} + |u^0|_{H^3(0,x_{i'})}\right) \\
&\qquad + Ch^{s/2}\left(\|f\|_{L^2(\Omega)} + \|u^0\|_{L^2(0,1)} + |u|_{H^3(\Omega_0^{++})}\right).
\end{aligned}$$

Choose $s = 4$ and follow the same argument as in the proof of Theorem 6.4.2 to complete the proof. □

Remark 6.4.2 The assumption (6.4.19) is reasonable, since we are interested in the singularly perturbed case. Theorem 6.4.2 tells us that under this assumption, away from any layers, the scheme (6.2.9) – (6.2.12) on an arbitrary tensor product mesh is first order accurate in the l_2 seminorm, as one can choose s sufficiently large to make the term $Ch^s h_{i'+1}^{-1}$ negligible. Corollary 6.4.1 indicates that if we work with an almost uniform mesh, then the method becomes second order accurate in smooth regions. Corollary 6.4.2 shows that when the diffusion parameter ε is relatively small, the method is second order accurate on any general tensor product mesh, away from any layers. However, this l_2 seminorm is of course not strong enough to exclude chequerboard oscillations from the computed solution.

Bibliography

[1] D. N. De G. Allen and R. Southwell, *Relaxation methods applied to determine the motion, in two dimensions, of a viscous fluid past a fixed cylinder*, Q. J. Mech. Appl. Math. 8 (1955), pp. 129 - 145.

[2] J. W. Barrett and K. W. Morton, *Optimal finite element solutions to diffusion-convection problems in one dimension*, Int. J. Numer. Meth. Engrg., 15 (1980), pp. 1457 - 1474.

[3] H. Baumert, P. Braun, E. Glos, W. Müller and G. Stoyan, *Modelling and computation of water quality problems in river networks*, Lecture Notes in Control and Information Sciences 23, Berlin, Springer, 1981, pp. 482 - 491.

[4] L. Bobisud, *Second-order linear parabolic equations with a small parameter*, Arch. Rational Mech. Anal. 27 (1967), pp. 385 - 397.

[5] M. A. Celia, T. F. Russell, I. Herrera and R. E. Ewing, *An Eulerian-Lagrangian localized adjoint method for the advection-diffusion equation*, Adv. Water Resour. 13 (1990), pp. 187 - 206.

[6] P. G. Ciarlet, *The finite element method for elliptic problems*, North Holland, Amsterdam, 1980.

[7] C. N. Dawson, T. F. Russell and M. F. Wheeler, *Some improved error estimates for the modified method of characteristics*, SIAM J. Numer. Anal. 26 (1989), pp. 1487 – 1512.

[8] E. P. Doolan, J. J. H. Miller and W. H. A. Schilders, *Uniform numerical methods for problems with initial and boundary layers*, Boole Press, Dublin, 1980.

[9] D. Duffy, *Uniformly convergent difference schemes for the convection-diffusion equation*, in BAIL I, Proc. First Internat. Conf. on Boundary and Interior Layers (J. J. H. Miller ed.), Boole Press, Dublin, 1980, pp. 265 – 269.

[10] W. Eckhaus and E. M. de Jager, *Asymptotic solutions of singular perturbation problems for linear differential equations of elliptic type*, Arch. Rational Mech. Anal. 23 (1966), pp. 26 – 86.

[11] R. E. Ewing, *The mathematics of reservoir simulation*, SIAM, Philadelphia, 1983.

[12] P. A. Farrell and A. Hegarty, *On the determination of the order of uniform convergence*, in Proceedings of 13th IMACS World Congress on Computation and Applied Mathematics (R. Vichnevetsky, J. J. H. Miller, eds.), IMACS, 1991, pp. 501 – 502.

[13] E. C. Gartland Jr., *Uniform high-order difference schemes for a singularly perturbed two-point boundary value problem*, Math. Comp. 48 (1987), pp. 551 – 564.

[14] H. Han and R. B. Kellogg, *A method of enriched subspaces for the numerical solution of a parabolic singular perturbation problem*, in BAIL II, Proc. Second

Internat. Conf. on Boundary and Interior Layers (J. J. H. Miller ed.), Boole Press, Dublin, 1982, pp. 46 – 52.

[15] A. C. Hindmash, P. M. Gresho and D. F. Griffiths, *The stability of explicit Euler time-integration for certain finite difference approximations of the multi-dimensional advection-diffusion equation*, Internat. J. Numer. Meth. Fluids 4 (1984), pp. 853 – 897.

[16] T. J. R. Hughes and A. N. Brooks, *A multidimensional upwind scheme with no crosswind diffusion*, in Finite Element Methods for Convection Dominated Flows (Hughes, T. J. R., ed.), AMD. Vol. 34, ASME, New York, N. Y., 1979, pp. 19 – 35.

[17] A. M. Il'in, *Differencing scheme for a differential equation with a small parameter affecting the highest derivative*, Mat. Zametki 6 (1969), pp. 237 – 248, with English transl. in Math. Notes 6 (1969), pp. 596 – 602.

[18] M. Jacob, *Heat transfer*, Wiley, New York, 1959.

[19] C. Johnson and U. Nävert, *An analysis of some finite element methods for advection-diffusion problems*, in Analytical and Numerical Approaches to Asymptotic Problems in Analysis (Axelsson, O., Frank, L. S., van der Sluis, A., eds.), North-Holland, Amsterdam, 1981, pp. 99 – 116.

[20] C. Johnson, U. Nävert and J. Pitkäranta, *Finite element methods for linear hyperbolic problems*, Comput. Meths. Appl. Mech. Engrg. 45 (1984), pp. 285 – 312.

[21] C. Johnson, *Numerical solution of partial differential equations by the finite element method*, Cambridge University Press, Cambridge, 1987.

[22] C. Johnson, A. H. Schatz and L. B. Wahlbin, *Crosswind smear and pointwise errors in streamline diffusion finite element methods*, Math. Comp. 49 (1987), pp. 25 – 38.

[23] M. Křížek, *On semiregular families of triangulations and linear interpolation*, Appl. Math. 36 (1991), pp. 223 – 232.

[24] V. D. Liseikin and V. E. Petrenko, *On the numerical solution of a nonlinear singular perturbation problem*, Preprint 687, Novosibirsk 1987, Siberian Branch, USSR Academy of Sciences.

[25] J. A. Mackenzie and K. W. Morton, *Finite volume solutions for convection-diffusion test problems*, Math. Comp. 60 (1993), pp. 189 – 220.

[26] K. W. Morton and I. J. Sobey, *Discretisation of a convection-diffusion equation*, Technical Report 91/4, Oxford University Computing Laboratory, Numerical Analysis Group, September 1991.

[27] K. W. Morton and M. Stynes, *An analysis of the cell vertex method.* Technical Report 91/7, Oxford University Computing Laboratory, Numerical Analysis Group, May 1992.

[28] K. W. Morton and E. Süli, *Finite volume methods and their analysis*, IMA Journal of Numerical Analysis, 11 (1991), pp. 241 – 260.

[29] U. Nävert, *A finite element method for convection-diffusion problems*, Thesis, Chalmers University of Technology and University of Gothenburg, 1982.

[30] A. Nayfeh, *Perturbation methods*, Wiley, 1973.

[31] M. J. Ng-Stynes, E. O'Riordan and M. Stynes, *Numerical methods for time-dependent convection-diffusion equations*, J. Comp. Appl. Maths. 21 (1988), pp. 289 – 310.

[32] K. Niijima, *Pointwise error estimates for a streamline diffusion finite element scheme*, Numer. Math. 56 (1990), pp. 707 – 719.

[33] R. E. O'Malley, *Singular perturbation methods for ordinary differential equations*, Springer-Verlag, New York, 1991.

[34] E. O'Riordan and M. Stynes, *Analysis of difference schemes for singularly perturbed differential equations using a discretized Green's function*, in BAIL IV, Proc. Fourth Internat. Conf. on Interior and Boundary Layers (J. J. H. Miller ed.), Boole Press, Dublin, 1986, pp. 157 – 168.

[35] E. O'Riordan and M. Stynes, *Necessary conditions for uniform convergence of difference schemes for a singularly perturbed parabolic problem*, in BAIL V, Proc. Fifth Internat. Conf. on Boundary and Interior Layers (B. Guo, J. J. H. Miller, Z. Shi, eds.), Boole Press, Dublin, 1988, pp. 277 – 281.

[36] E. O'Riordan and M. Stynes, *A globally uniformly convergent finite element method for a singularly perturbed elliptic problem in two dimensions*, Math. Comp. 57 (1991), pp. 47 – 62.

[37] D. W. Peaceman and H. H. Rachford, *numerical calculation of multi-dimensional miscible displacement*, Soc. Petroleum Engrg. J., 24 (1962), pp. 327 – 338.

[38] H. S. Price and R. S. Varga, *Error bounds for semidiscrete Galerkin approximations of parabolic problems with applications to petroleum reservoir mechanics*, in Numerical solution of field problems in continuum physics, SIAM-AMS, 2, 1970.

[39] H.-G. Roos, *Necessary convergence conditions for upwind schemes in the two-dimensional case*, Int. J. Numer. Meth. Eng. 21 (1985), pp. 1459 – 1469.

[40] T. F. Russell, *Eulerian-Lagrangian localized adjoint methods for advection-dominated problems*, in Proc. 13th Biennial Conf. on Numer. Anal., University of Dundee, Pitman, 1990.

[41] G. I. Shishkin, *Grid approximation of singularly perturbed boundary value problems with convective terms*, Sov. J. Numer. Anal. Math. Modelling 5 (1990), pp. 173 – 187.

[42] S.-Y. Hahn, *An 'upwind' finite element method for electromagnetic field problems in moving media*, Int. J. Numer. Meth. Eng. 24 (1987), pp. 2071 – 2086.

[43] M. Stynes, *An adaptive uniformly convergent numerical method for a semilinear singular perturbation problem*, SIAM J. Numer. Anal. 26 (1989), pp. 442 – 455.

[44] M. Stynes and W. Guo, *Finite element analysis of an advection-diffusion equation*, in Numerical methods in singularly perturbed problems, Proc. Intl. Sem-

inar on Appl. Math. '91, Dresden, Germany, (H. -G. Roos, A. Felgenhauer, L. Angermann, eds.), Technische Universität Dresden, 1991, pp. 123 - 128.

[45] M. Stynes and E. O'Riordan, *Uniformly convergent difference schemes for singularly perturbed parabolic diffusion-convection problems without turning points*, Numer. Math. 55 (1989), pp. 521 - 544.

[46] M. Stynes and E. O'Riordan, *An analysis of a two-point boundary value problem with a boundary layer, using only finite element techniques*, Technical Report, Department of Mathematics, University College, Cork, October 1989.

[47] M. Stynes and E. O'Riordan, *An analysis of a two-point boundary value problem with a boundary layer, using only finite element techniques*, Math. Comp. 56 (1991), pp. 663 - 675.

[48] E. Süli, *The accuracy of finite volume methods on distorted partitions*, in The Mathematics of Finite Elements and Applications (J. R. Whiteman ed.), VII MAFELAP 1990, Academic Press, 1992, pp. 253 - 260.

[49] E. Süli, *The accuracy of cell vertex finite volume methods on quadrilateral meshes*, Math. Comp. 59 (1992), pp. 359 - 382.

[50] M. Van Dyke, *Perturbation methods in fluid mechanics*, Parabolic Press, 1975.

[51] M. I. Vishik and L. A. Lyusternik, *Regular degeneration and boundary layer for linear differential equations with a small parameter*, Uspekki Mat. Nauk. 12 (1957), pp. 3 - 122= Amer. Math. Soc. Transl. Ser. 2, 20 (1969), pp. 239 - 364.

[52] W. Wu, *Petrov-Galerkin methods for parabolic convection-diffusion problems*, Ph.D. Thesis, Linacre College, University of Oxford, 1987.

[53] G. Zhou, *An adaptive streamline diffusion finite element method for hyperbolic systems in gas dynamics*, Preprint 688, Institute für angewandte Mathematik, Universität Heidelberg, Germany, September 1992.