A $P^1 - P^1$ FINITE ELEMENT METHOD FOR A PHASE RELAXATION MODEL. PART II: ADAPTIVELY REFINED MESHES*

XUN JIANG†, RICARDO H. NOCHETTO‡, AND CLAUDIO VERDI♯

Abstract. We examine the effect of adaptively generated refined meshes on the $P^1 - P^1$ finite element method with semi-explicit time stepping of Part I, which applies to a phase relaxation model with small parameter $\varepsilon > 0$. A typical mesh is highly graded in the so-called refined region, which exhibits a local meshsize proportional to the time step τ, and is coarse in the remaining parabolic region where the meshsize is of order $\sqrt{\tau}$. Three admissibility tests guarantee mesh quality and, upon failure, lead to remeshing and so to incompatible consecutive meshes. The most severe test checks whether the transition region, where phase changes take place, belongs to the refined region. The other two tests monitor equidistribution of pointwise interpolation errors. The resulting adaptive scheme is shown to be stable in various Sobolev norms and to converge with a rate of order $\mathcal{O}(\tau/\sqrt{\varepsilon})$ in the natural energy spaces. Several numerical experiments illustrate the scheme's efficiency and enhanced performance as compared with those of Part I.

Key words. phase relaxation, diffuse interface, piecewise linear finite elements, graded meshes, numerical integration, double obstacle

AMS(MOS) subject classifications. 65N15, 65N30, 35R35

1. Introduction. We continue our study of a $P^1 - P^1$ finite element method for the phase relaxation model of Visintin [20]; see Part I [11] and also [18,19,21]. If θ denotes temperature and χ phase variable, the model is written as

$$\partial_t\theta + \partial_t\chi - \Delta\theta = f, \tag{1.1}$$

$$\varepsilon\partial_t\chi + \Lambda(\chi) \ni \theta, \tag{1.2}$$

and is subject to a Neumann boundary condition $\partial_\nu\theta = 0$ and initial data $\theta(\cdot,0) = \theta_0$ and $\chi(\cdot,0) = \chi_0$. Hereafter $\varepsilon > 0$ is a small but fixed parameter and Λ is the inverse of the sign graph. Function χ varies from -1 to 1, the obstacles, within a thin transition region $\mathcal{T}(t) = \{|\chi(\cdot,t)| < 1\}$ (or noncoincidence set) of thickness $\mathcal{O}(\sqrt{\varepsilon})$, where phase changes occur. Both $\nabla\theta$ and χ vary most rapidly within $\mathcal{T}(t)$. Exploiting this localized singular behavior for numerical purposes is the main theme of this paper.

In Part I we examine stability and accuracy of the semi-explicit time discretization

$$\tfrac{1}{\tau}(\Theta^n - \Theta^{n-1}) + \tfrac{1}{\tau}(X^n - X^{n-1}) - \Delta\Theta^n = f^{n-1},$$
$$\tfrac{\varepsilon}{\tau}(X^n - X^{n-1}) + \Lambda(X^n) \ni \Theta^{n-1},$$

* This work was partially supported by NSF Grants DMS-9305935 and DMS-9623394, ESF-FBProgramme "Mathematical Treatment of Free Boundary Problems", EU Grant HCM "Phase Transitions and Surface Tension", MURST, and CNR Contract 95.00735.01. The first author was in residence at the Istituto di Analisi Numerica del CNR di Pavia with a CNR fellowship during part of this work.

† Department of Mathematics, University of California at Davis, Davis, CA 95616, USA – e-mail: xun@math.ucdavis.edu

‡ Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA – e-mail: rhn@math.umd.edu

♯ Dipartimento di Matematica, Università di Milano, 20133 Milano, Italy – e-mail: verdi@paola.mat.unimi.it

combined with continuous piecewise linear finite elements for both Θ^n and X^n over a quasi-uniform mesh of size h. Under a natural compatibility condition between θ_0 and χ_0, we obtain an error estimate of linear order $\mathcal{O}((h+\tau)/\sqrt{\varepsilon})$ provided $\tau \leq \varepsilon$, which is optimal for fixed ε.

In Part II we show how to adaptively produce a sequence of highly graded meshes with information extracted from the fully discrete solutions Θ_h^n and X_h^n; hereafter $h > 0$ stands for the largest element size. Meshes are refined near the *discrete transition region* $\mathcal{T}^n = \{|X_h^n| < 1\}$, where the meshsize is of order $\mathcal{O}(\tau)$, and rather coarse in the remaining parabolic region where the meshsize is of order $\mathcal{O}(\sqrt{\tau})$. The objective is to equidistribute pointwise interpolation errors both in space and time, which corresponds to optimizing the number of degrees of freedom for a desired accuracy. Three admissibility tests monitor mesh quality and, upon failure, lead to remeshing and so to incompatible consecutive meshes. The most severe test checks whether $\mathcal{T}^n$ belongs to the so-called *refined region* where meshsize is of order $\mathcal{O}(\tau)$. However remeshing is expensive if performed very often. This reveals two conflicting conditions for mesh design: the refined region must be sufficiently wide to contain $\mathcal{T}^n$ for as long as possible but also sufficiently thin to minimize the numbers of degrees of freedom. An interpolation error of order $\mathcal{O}(\tau)$ is incurred at every mesh change, which may accumulate in time compromising accuracy. We then limit the number of mesh changes to $\mathcal{O}(1/\sqrt{\varepsilon})$ which, in turn, yields a refined region width of order $\mathcal{O}(\sqrt{\varepsilon})$ and so comparable with that of $\mathcal{T}^n$. These design considerations are fully discussed here and shown to yield stability and a rate of convergence of optimal order $\mathcal{O}((\tau+h^2)/\sqrt{\varepsilon})$. Therefore the present adaptive method improves upon that in Part I in terms of fewer degrees of freedom for the same asymptotic accuracy, provided certain discrete regularity properties are valid. This order is clearly *optimal* as expected from approximation theory, but deteriorates as $\varepsilon \downarrow 0$ as predicted by the semidiscrete traveling waves of [16]; whether $\mathcal{O}(h^2/\sqrt{\varepsilon})$ is the correct asymptotics if both $h, \varepsilon \downarrow 0$ is an intriguing open question.

This study of diffuse interfaces in $\mathbb{R}^d$ extends those in [13,14,15] valid for sharp interfaces and $d = 2$; the case $\tau = \varepsilon$ reduces to the fully discrete Chernoff formula of [15] and improves upon it. We envision applications of similar ideas for the design of adaptive solvers for more complex phase transitions such as *phase field models* [2,3] and models of *polymer crystallization* [12,18]. We believe that our results, both theory and implementation, apply rather directly to the models in [1,6,18].

This paper is organized as follows. In §2 we discuss the semidiscrete and fully discrete problems along with a compatibility condition for the discrete data. We address mesh design in §3 based on interpolation theory for incompatible meshes, which is briefly reviewed in §4. We derive a priori estimates in §5, most notably a new strong stability result. This turns out to be crucial in our error analysis of §6. We conclude in §7 with numerical experiments revealing enhanced performance, efficiency and potentials of our adaptive approach with respect to those of Part I.

2. Problem Statement.

We state semidiscrete and fully discrete problems, along with results from [10]. We then discuss properties of the discrete initial data.

2.1. Semidiscrete Setting.

Let Ω be a bounded convex polyhedron (or smooth domain) in $\mathbb{R}^d$, and set $Q := \Omega \times (0,T)$. Let $z_0 \in \Lambda(\chi_0)$ be the following L^2 function

$$z_0 := \min(\theta_0, 0) \text{ if } \chi_0 = -1, \quad z_0 := 0 \text{ if } |\chi_0| < 1, \quad z_0 := \max(\theta_0, 0) \text{ if } \chi_0 = 1,$$

and assume the compatibility condition

$$\|\theta_0 - z_0\|_{L^\infty(\Omega)} \leq A\sqrt{\varepsilon}, \tag{2.1}$$

along with the regularity properties

$$\|\nabla\theta_0\|_{L^\infty(\Omega)}, \quad \sqrt{\varepsilon}\|D^2\theta_0\|_{L^\infty(\Omega)}, \quad \sqrt{\varepsilon}\|\nabla\chi_0\|_{L^\infty(\Omega)} \le C. \tag{2.2}$$

We note that (2.2) is stronger than the regularity assumption (2.1) in [11] but is still consistent with the regularity of traveling waves [16]. Hereafter C denotes a generic positive constant independent of the relevant parameters ε, h, τ, and n.

Let $\tau > 0$ be the (uniform) time step, $N := T/\tau$ be a positive integer, $t^n := n\tau$, and $I^n := (t^{n-1}, t^n]$ for $1 \le n \le N$. Let $\langle\cdot,\cdot\rangle$ denote the $L^2(\Omega)$ pairing. The semi-explicit discretization in time studied in [10,18,19] reads as follows: *Given $\Theta^0 := \theta_0$, $X^0 := \chi_0$, for any $1 \le n \le N$, find $\Theta^n \in H^1(\Omega)$, $|X^n| \le 1$, and $Z^n \in \Lambda(X^n)$ a.e. in Ω such that*

$$\tfrac{1}{\tau}\langle\Theta^n - \Theta^{n-1}, \phi\rangle + \tfrac{1}{\tau}\langle X^n - X^{n-1}, \phi\rangle + \langle\nabla\Theta^n, \nabla\phi\rangle = 0 \quad \forall\phi \in H^1(\Omega), \tag{2.3}$$

$$\tfrac{\varepsilon}{\tau}\langle X^n - X^{n-1}, \varphi\rangle + \langle Z^n, \varphi\rangle = \langle\Theta^{n-1}, \varphi\rangle \quad \forall\varphi \in L^2(\Omega). \tag{2.4}$$

Note that we have taken $f = 0$ in (1.1) just for simplicity. Results in [10] are proven in the case of a Dirichlet boundary condition, but all proofs apply to the case of a Neumann boundary condition as well. If the stability constraint $\tau \le \varepsilon$ is enforced, we then have the a priori bounds [10]

$$\begin{aligned} &\sum_{n=1}^N \tfrac{1}{\tau}\|\Theta^n - \Theta^{n-1}\|^2_{L^2(\Omega)} + \max_{1\le n\le N}\|\nabla\Theta^n\|^2_{L^2(\Omega)} + \Big\|\sum_{n=1}^N \tau\Theta^n\Big\|_{H^2(\Omega)} \\ &\quad + \varepsilon\sum_{n=1}^N \tau\|\Theta^n\|^2_{H^2(\Omega)} + \tfrac{\varepsilon}{\tau^2}\max_{1\le n\le N}\|X^n - X^{n-1}\|^2_{L^2(\Omega)} \\ &\quad + \sum_{n=1}^N \tfrac{1}{\tau}\langle X^n - X^{n-1}, Z^n - Z^{n-1}\rangle \le C\big(A^2 + \|\nabla\theta_0\|^2_{L^2(\Omega)}\big), \end{aligned} \tag{2.5}$$

with A defined in (2.1). If $H^1(\Omega)^*$ stands for the dual of $H^1(\Omega)$, we also have the optimal error estimates [10]

$$\begin{aligned} &\|u(t^m) - U^m\|_{H^1(\Omega)^*} + \sqrt{\varepsilon}\|\chi(t^m) - X^m\|_{L^2(\Omega)} \\ &\quad + \Big(\sum_{n=1}^m \int_{I^n} \tfrac{\varepsilon}{\tau}\|\tau\partial_t\chi - (X^n - X^{n-1})\|^2_{L^2(\Omega)}\,dt\Big)^{1/2} \\ &\quad + \Big(\sum_{n=1}^m \int_{I^n} \big(\|\theta - \Theta^n\|^2_{L^2(\Omega)} + \tau\|\theta - \Theta^n\|^2_{H^1(\Omega)}\big)\,dt\Big)^{1/2} \le C\tfrac{\tau}{\sqrt{\varepsilon}}. \end{aligned} \tag{2.6}$$

2.2. Fully Discrete Setting. In contrast to [11], where one quasi-uniform mesh is generated and used throughout the computation, we now consider a sequence of graded meshes $\{\mathcal{M}^n\}_{n=1}^N$ which are shape regular uniformly with respect to $1 \le n \le N$ [4,p.132]. The size h_S of each simplex $S \in \mathcal{M}^n$ verifies

$$\lambda_1\tau \le h_S \le \lambda_2\sqrt{\tau}, \tag{2.7}$$

where $\lambda_1, \lambda_2 > 0$ are two fixed parameters. The nth *refined region* $\mathcal{R}^n$ is defined to be

$$\mathcal{R}^n := \cup\big\{S \in \mathcal{M}^n : h_S = \mathcal{O}(\tau)\big\}.$$

Let $h = \mathcal{O}(\sqrt{\tau})$ be the largest element size of all meshes $\mathcal{M}^n$. Let $\mathcal{E}^n$ denote the set of interelement boundaries (or edges). We denote by $\mathbb{V}^n \subset H^1(\Omega)$ the space of continuous piecewise linear finite elements over $\mathcal{M}^n$ and by Π^n the associated Lagrange interpolation operator. We denote by $\{x_j^n\}_{j=1}^{J^n}$ the nodes of $\mathcal{M}^n$ and by $\{\phi_j^n\}_{j=1}^{J^n}$ the canonical basis of $\mathbb{V}^n$. For all $\psi, \varphi \in C^0(\bar{\Omega})$, we denote the vertex quadrature rule by

$$\langle \psi, \varphi \rangle_n := \int_\Omega \Pi^n(\psi\varphi)\, dx, \quad \|\varphi\|_n := \sqrt{\langle \varphi, \varphi \rangle_n},$$

which is easy to evaluate in practice [4]. It is well known that

$$\|\varphi\|_{L^2(\Omega)} \le \|\varphi\|_n \le C^* \|\varphi\|_{L^2(\Omega)} \quad \forall \varphi \in \mathbb{V}^n$$

and that the following interpolation estimate is valid, for all $S \in \mathcal{M}^n$ and $\varphi, \psi \in \mathbb{V}^n$,

$$\int_S \left| \varphi\psi - \Pi^n(\varphi\psi) \right| dx \le C^* h_S^{1+k} \|\nabla^k \varphi\|_{L^2(S)} \|\nabla \psi\|_{L^2(S)}, \quad k = 0, 1. \tag{2.8}$$

For a subset $D \subset \Omega$ which is the union of elements $S \in \mathcal{M}^n$ we use the notation

$$\langle \psi, \varphi \rangle^D := \int_D \psi\varphi\, dx, \quad \langle \psi, \varphi \rangle_n^D := \int_D \Pi^n(\psi\varphi)\, dx, \quad \|\varphi\|_{D,n} := \sqrt{\langle \varphi, \varphi \rangle_n^D}.$$

For each $\mathcal{M}^n$ we introduce the dual mesh $\mathcal{D}^n$, as in [11,§3]. For a continuous function φ we indicate with $\tilde{\varphi} = P^n(\varphi)$ the piecewise constant interpolant over $\mathcal{D}^n$ defined by $\tilde{\varphi}(x_j^n) = \varphi(x_j^n)$ for all $1 \le j \le J^n$. If $\varphi \in \mathbb{V}^n$, then $\|\tilde{\varphi}\|_{L^2(\Omega)} = \|\varphi\|_n$ and

$$\int_S \left| \varphi - \tilde{\varphi} \right| dx \le C h_S \|\nabla \varphi\|_{L^2(S)} \quad \forall S \in \mathcal{M}^n. \tag{2.9}$$

We suppose that $\{\mathcal{M}^n\}_{n=1}^N$ are *weakly acute*, so that the Discrete Maximum Principle holds. This, in turn, compensates for the lack of regularity [5,13,15].

If we have two consecutive meshes $\mathcal{M}^{n-1}$, $\mathcal{M}^n$ and corresponding finite element spaces $\mathbb{V}^n$, $\mathbb{V}^{n-1}$, we set $\hat{\varphi} = \Pi^n \varphi \in \mathbb{V}^n$ for all $\varphi \in \mathbb{V}^{n-1}$. The critical issue of determining $\mathcal{M}^n$ from $\mathcal{M}^{n-1}$ will be discussed later in §3. For the moment we assume that $\mathcal{M}^n$ is known and explain how to compute Θ_h^n and X_h^n. The fully discrete problem then reads: *Given $\mathcal{M}^n$ and $\hat{\Theta}_h^{n-1}$, $\hat{X}_h^{n-1} \in \mathbb{V}^n$, find Θ_h^n, X_h^n, $Z_h^n \in \mathbb{V}^n$ such that*

$$\tfrac{\varepsilon}{\tau}(X_h^n - \hat{X}_h^{n-1}) + Z_h^n = \hat{\Theta}_h^{n-1}, \quad Z_h^n(x_j^n) \in \Lambda(X_h^n(x_j^n)) \quad \forall 1 \le j \le J^n, \tag{2.10}$$

$$\tfrac{1}{\tau}\langle \Theta_h^n - \hat{\Theta}_h^{n-1}, \varphi \rangle + \tfrac{1}{\tau}\langle X_h^n - \hat{X}_h^{n-1}, \varphi \rangle_n + \langle \nabla \Theta_h^n, \nabla \varphi \rangle = 0 \quad \forall \varphi \in \mathbb{V}^n, \tag{2.11}$$

and set $U_h^n := \Theta_h^n + X_h^n$. We notice that (2.10) is equivalent to the nodewise correction

$$X_h^n(x_j^n) = (I + \Lambda)^{-1}\big(\tfrac{\tau}{\varepsilon}\hat{\Theta}_h^{n-1}(x_j^n) + \hat{X}_h^{n-1}(x_j^n)\big) \tag{2.12}$$

and that (2.11) can be rewritten as

$$\langle \Theta_h^n, \varphi \rangle + \tau \langle \nabla \Theta_h^n, \nabla \varphi \rangle = \langle \hat{\Theta}_h^{n-1}, \varphi \rangle - \langle X_h^n - \hat{X}_h^{n-1}, \varphi \rangle_n \quad \forall \varphi \in \mathbb{V}^n, \tag{2.13}$$

which is a positive definite symmetric linear system for Θ_h^n. Thus the discrete problem is solved first by updating X_h^n using the nonlinear inexpensive nodewise relation (2.12), and then by solving the linear system (2.13) (see §7.1).

In contrast to the discrete problem in [11], we see that *exact integration* is used for Θ_h^n in (2.13) but the resulting scheme is still *practical.* In fact a quadrature rule exact for quadratics, like the midpoint rule [4,p.183], would suffice to compute $\langle \Theta_h^n, \varphi \rangle$ while symmetry and positive definiteness are preserved. On the other hand, exact integration yields the asserted rate of convergence $h^2/\sqrt{\varepsilon}$ that does not seem feasible otherwise (see §6). However, the use of quadrature for X_h^n in (2.13) is essential because the constitutive relation $Z_h^n \in \Lambda(X_h^n)$ is only enforced at nodes.

The discrete phase variable X_h^n is used to define the liquid $\mathcal{L}^n$, crystal $\mathcal{C}^n$, and transition region $\mathcal{T}^n$ as follows

$$\mathcal{L}^n\,(\text{resp. } \mathcal{C}^n) := \{x \in \Omega : X_h^n(x) = 1\,(\text{resp. } -1)\}, \quad \mathcal{T}^n := \{x \in \Omega : |X_h^n(x)| < 1\};$$

phase change takes place within the diffuse interface $\mathcal{T}^n$. Since X_h^n is piecewise linear and $|X_h^n| \le 1$, $|X_h^n|$ restricted to the interior of an element is either 1 or < 1. Consequently $\mathcal{L}^n$, $\mathcal{C}^n$, and $\mathcal{T}^n$ are unions of elements, are mutually disjoint and constitute a partition of Ω: $\Omega = \mathcal{L}^n \cup \mathcal{C}^n \cup \mathcal{T}^n$. Another important set is the so-called *enlarged transition region*, defined by

$$\mathcal{T}_+^n := \mathcal{T}^n \cup \mathcal{T}^{n+1} \cup \mathcal{A}^n, \quad \text{where} \quad \mathcal{A}^n := (\mathcal{L}^n \cap \mathcal{C}^{n+1}) \cup (\mathcal{L}^{n+1} \cap \mathcal{C}^n), \tag{2.14}$$

which must be kept fixed when changing meshes (see §3.2). The set $\mathcal{A}^n$ accounts for phase changes which occur in one single time step, and will be shown to be small or even empty for $\tau \ll \varepsilon$ (see Lemma 5.2).

2.3. Discrete Initial Data. The initial graded mesh $\mathcal{M}^1$ is designed with information extracted from θ_0 and χ_0. $\mathcal{M}^1$ is refined near the enlarged initial transition region $\mathcal{T}_+^0$, where the local meshsize is $\mathcal{O}(\tau)$, and coarse elsewhere so as to guarantee equidistribution of spatial interpolation errors in the maximum norm. We define the discrete initial functions Θ_h^0, X_h^0, Z_h^0, $U_h^0 \in \mathbb{V}^1$ as

$$\Theta_h^0 := \Pi^1\theta_0, \quad X_h^0 := \Pi^1\chi_0, \quad Z_h^0 := \Pi^1 z_0, \quad U_h^0 := \Theta_h^0 + X_h^0.$$

In view of (2.2) we readily have

$$\|\nabla\Theta_h^0\|_{L^\infty(\Omega)} \le C, \quad \sqrt{\varepsilon}\|\nabla X_h^0\|_{L^\infty(\Omega)} \le C,$$
$$\|\Theta_h^0 - \theta_0\|_{L^\infty(\Omega)} \le Ch^2\|D^2\theta_0\|_{L^\infty(\Omega)} \le C\frac{\tau}{\sqrt{\varepsilon}}.$$

Since $X_h^0 = \chi_0 = 1$ or -1 outside $\mathcal{T}_+^0$ and $h_S = \mathcal{O}(\tau)$ within $\mathcal{T}_+^0$, we get

$$\|X_h^0 - \chi_0\|_{L^\infty(\Omega)} \le C\tau\|\nabla\chi_0\|_{L^\infty(\Omega)} \le C\frac{\tau}{\sqrt{\varepsilon}},$$
$$\|U_h^0 - u_0\|_{H^1(\Omega)^*} \le \|\Theta_h^0 - \theta_0\|_{L^\infty(\Omega)} + \|X_h^0 - \chi_0\|_{L^\infty(\Omega)} \le C\frac{\tau}{\sqrt{\varepsilon}}.$$

As in [11], in order to carry out stability and error analysis, we select $\Theta_h^{-1} := \Theta_h^0 \in \mathbb{V}^1$, and $X_h^{-1} \in \mathbb{V}^1$ such that $\frac{\varepsilon}{\tau}(X_h^0 - X_h^{-1}) + Z_h^0 = \Theta_h^{-1}$; thus (2.10) holds for all $0 \le n \le N$. In light of (2.1) we deduce

$$\|X_h^0 - X_h^{-1}\|_{L^\infty(\Omega)} \le \frac{\tau}{\varepsilon}\|\Theta_h^0 - Z_h^0\|_{L^\infty(\Omega)} \le \frac{\tau}{\varepsilon}\|\theta_0 - z_0\|_{L^\infty(\Omega)} \le A\frac{\tau}{\sqrt{\varepsilon}}. \tag{2.15}$$

3. Mesh Adaptation. In this section, we discuss mesh design and mesh quality tests. We begin with our motivation for mesh adaptation.

3.1. Motivation. Consider the heat equation, a prototype of parabolic PDE,

$$\partial_t\theta - \Delta\theta = f,$$

discretized with piecewise linear finite elements in space over graded meshes $\{\mathcal{M}^n\}_{n=1}^N$ and backward differences in time. The rate of convergence in $L_t^\infty L_x^2$ is known to be $\mathcal{O}(h^2 + \tau)$ provided $\theta \in L^\infty(0,T;H^2(\Omega)) \cap W^{1,\infty}(0,T;L^2(\Omega))$ [7]. We have seen in [10] that we can only expect a rate $\mathcal{O}(\tau/\sqrt{\varepsilon})$ for our time discretization. We have also shown in Theorem 5.1 of [11] that space discretization yields an additional rate

$\mathcal{O}(h/\sqrt{\varepsilon}\,)$, which is caused by the singular behavior of Θ_h^n and X_h^n in the transition region $\mathcal{T}^n$. Since $X_h^n = X_h^{n-1} = 1$ or -1, and thus Θ_h^n solves a heat equation, outside $\mathcal{T}_+^{n-1} \subset \mathcal{R}^n$, we would expect the usual *parabolic* relation $h_S = \mathcal{O}(\sqrt{\tau}\,)$ away from $\mathcal{R}^n$. The *hyperbolic* relation $h_S = \mathcal{O}(\tau)$ should still be enforced in $\mathcal{R}^n$, as suggested by Theorem 5.1 of [11]. It is our objective to show that this heuristics is indeed correct to achieve a global rate of convergence $\mathcal{O}((h^2+\tau)/\sqrt{\varepsilon}\,)$ in the natural energy norms.

However, we have to be more specific about the relation between local meshsize h_S and time step τ. To elaborate upon this relation, we set $y := x + Vt$ and consider the traveling waves of [16] with velocity V moving toward left and obstacles 0 and 1,

$$\bar\theta(y) = \gamma\sqrt{\varepsilon}\sinh\left(\tfrac{y}{\sqrt{\varepsilon}}\right) + \mathcal{O}(\varepsilon), \quad \bar\chi(y) = \tfrac{\gamma}{V}\left(\cosh\left(\tfrac{y}{\sqrt{\varepsilon}}\right) - 1\right) + \mathcal{O}(\sqrt{\varepsilon}), \quad \forall y \in (0, y_\varepsilon).$$

Here $y_\varepsilon = \mathcal{O}(\sqrt{\varepsilon}\,)$ and $\mathcal{T}(t) = (-Vt, y_\varepsilon - Vt)$ is the transition region. These functions satisfy (1.1), (1.2) with $f = 0$. Since $\partial_t\chi = \mathcal{O}(1/\sqrt{\varepsilon})$ and $\partial_x\chi = \mathcal{O}(1/V\sqrt{\varepsilon})$, the discretization errors in time and space for χ obey the estimates

$$\chi(x,t^n) - \chi(x,t^{n-1}) \approx \tfrac{\tau}{\sqrt{\varepsilon}}, \quad \Pi_h\chi(x,t) - \chi(x,t) \approx \tfrac{h}{V\sqrt{\varepsilon}}, \quad \forall x \in \mathcal{T}(t).$$

Since $\partial_{xx}\theta = \mathcal{O}(1/\sqrt{\varepsilon})$, the interpolation error for θ is of higher order: $\Pi_h\theta - \theta \approx h^2/\sqrt{\varepsilon}$. So the discretization errors for χ dominate. Balancing those errors leads to

$$h \approx V\tau. \tag{3.1}$$

In addition to (3.1), the 1D traveling waves in $\mathcal{T}(t)$ suggest how to estimate velocity V, a critical quantity for a diffuse interface problem. Since $\bar\theta$ and $\bar u := \bar\theta + \bar\chi$ satisfy (1.1), namely $V\bar u' - \bar\theta'' = 0$, integration from y_1 to y_2 yields the *modified Stefan condition*

$$V\big(u(x_1,t) - u(x_2,t)\big) = \partial_x\theta(x_1,t) - \partial_x\theta(x_2,t) \quad \forall x_1, x_2 \in \mathcal{T}(t). \tag{3.2}$$

Using (1.2) instead, we obtain the alternative expression

$$\varepsilon V\partial_x\chi(x,t) = \theta(x,t) \quad \forall x \in \mathcal{T}(t). \tag{3.3}$$

Away from $\mathcal{T}(t)$ we resort to θ to extract relevant information. The pointwise error due to time discretization is $\mathcal{O}(\tau/\sqrt{\varepsilon}\,)$ for 1D traveling waves [16]. On the other hand, the error incurred by interpolation between incompatible meshes is formally

$$\|\zeta(\Theta_h^n) - \Pi^{n+1}\zeta(\Theta_h^n)\|_{L^\infty(S)} \le Ch_S^2\big(\|D^2\Theta_h^n\|_{L^\infty(S)} + \|D\Theta_h^n\|_{L^\infty(S)}^2\big) \tag{3.4}$$

for a smooth nonlinear ζ. This formula is rigorously justified by the interpolation theory for incompatible meshes of [13], of which the L^∞ case is reviewed in §4. For (3.4) to be $\mathcal{O}(\tau)$, h_S needs to satisfy

$$h_S \le \sqrt{\tau}\min\big(\mu_1\|D^2\Theta_h^n\|_{L^\infty(S)}^{-1/2}, \mu_2\|D\Theta_h^n\|_{L^\infty(S)}^{-1}\big). \tag{3.5}$$

Moreover, we impose

$$N_c := \text{number of mesh changes} \le M/\sqrt{\varepsilon}, \tag{3.6}$$

for the cumulative effect of (3.4) not to exceed the global error $\mathcal{O}(\tau/\sqrt{\varepsilon}\,)$ of [11].

Finally to avoid large interpolation errors on X_h^n, X_h^{n+1} due to their rapid variation in $\mathcal{T}_+^n$, we assume that $\mathcal{T}_+^n$ is fixed when changing mesh $\mathcal{M}^n \to \mathcal{M}^{n+1}$. This is feasible because $\mathcal{T}_+^n \subset \mathcal{R}^n$ where elements possess size $\mathcal{O}(\tau)$.

3.2. Mesh Refinement Principles. We are now in the position to describe the principle of mesh adaptation at a generic nth step, assuming $\mathcal{M}^{n-1}, \Theta_h^{n-1}, X_h^{n-1}$ have already been accepted. Since the definition (2.14) of $\mathcal{T}_+^n$ needs both X_h^n and X_h^{n+1}, the required information for mesh design is extracted from the current solution Θ_h^n, X_h^n, and X_h^{n+1}. We will drop the indexes h and n, because no confusion is possible in the present context, and denote

$$X = X_h^n, \quad \Theta = \Theta_h^n, \quad \mathcal{M} := \mathcal{M}^n, \quad \mathcal{E} := \mathcal{E}^n, \quad \mathcal{R} := \mathcal{R}^n, \quad \mathcal{T}_+ = \mathcal{T}_+^n, \quad \hat{\mathcal{M}} = \mathcal{M}^{n+1}.$$

For any given set $D \subset \bar{\Omega}$, we set

$$\mathcal{M}_D := \{S \in \mathcal{M} : (S \backslash \partial S) \cap D \neq \emptyset\}, \quad \mathcal{E}_D := \{e \in \mathcal{E} : e \subset \partial S, S \in \mathcal{M}_D\}.$$

Finally, we define the discrete first and second derivatives of Θ as

$$d_S := \left|\nabla\Theta|_S\right| \quad \forall S \in \mathcal{M}, \quad D_e := |[\![\nabla\Theta]\!]_e|/h_e \quad \forall e \in \mathcal{E},$$

where $[\![\cdot]\!]_e$ indicates the jump operator across an edge e. If e is on the boundary, we set $D_e := |\nabla\Theta|/h_e$. Note that these quantities are easy to evaluate in practice.

3.2.1. Local Mesh Parameters. The desired element and edge sizes in the parabolic region of $\mathcal{M}$, as motivated by (3.5), are given by the local parameters

$$\hat{h}_e := \mu_1 \frac{\tau^{1/2}}{D_e^{1/2}} \quad \forall e \in \mathcal{E} \backslash \mathcal{E}_{\mathcal{T}}, \quad \hat{h}_S := \mu_2 \frac{\tau^{1/2}}{d_S} \quad \forall S \in \mathcal{M} \backslash \mathcal{M}_{\mathcal{T}}. \tag{3.7}$$

Constants μ_1, μ_2 are arbitrary and, in practice, will be computed to balance the number of degrees of freedom of both local parameters $\hat{h}_e$ and $\hat{h}_S$ with that dictated by $\hat{h}_{\mathcal{C}_S}$ in (3.9) below. We also introduce the "bad" set

$$\mathcal{B} := \cup\Big\{S \in \mathcal{M} \backslash \mathcal{M}_{\mathcal{T}_+} : \min_{e \in \mathcal{E} \backslash \mathcal{E}_{\mathcal{T}_+} : e \subset \partial S} (\hat{h}_S, \hat{h}_e) < \lambda_1 \tau\Big\},$$

where derivatives are so large that would lead to a local meshsize smaller than the minimum $\lambda_1\tau$. In our experiments $\mathcal{B}$ is empty, but it might not be in general [9]. Set

$$\mathcal{W} := \mathcal{T}_+ \cup \mathcal{B}. \tag{3.8}$$

In order to determine a local parameter within $\mathcal{T}_+$ we must impose (3.1). First, for each $S \in \mathcal{M}_{\mathcal{T}_+}$ we let S_+ (resp. S_-) indicate an element displayed in the direction of $\nu := \nabla\Theta|_S/d_S$ (resp. $-\nu$) and satisfy $\operatorname{dist}(S_+, S) \geq h_S$ (resp. $\operatorname{dist}(S_-, S) \geq h_S$); see [13,§3.4]. In light of (3.2), we define the velocity v_S along $-\nu$ as

$$v_S := \frac{(\nabla\Theta|_{S_+} - \nabla\Theta|_{S_-}) \cdot \nu}{U(x_+) - U(x_-)},$$

where x_+ (resp. x_-) denotes the barycenter of S_+ (resp. S_-). We then define $\mathcal{C}_S$ to be the cone of axis $-v_S\nu$, vertex at the barycenter x_S of S, opening $\pi/2$ and height $\mu_3|v_S|\sqrt{\varepsilon}$. Since ν and v_S are reliable only when $\mathcal{T}_+$ exhibits a truely one dimensional behavior along ν, we also resort to a discrete version of (3.3), namely

$$\hat{v}_S := \frac{\Theta(x_S)}{\varepsilon \nabla X \cdot \nu}.$$

If v_S is too small or too different from $\hat{v}_S$, the computation is uncertain. We then replace the cone by a ball of radius $\mu_3\lambda_1\sqrt{\varepsilon}$, which plays the role of an uncertainty region. The union of all these cones and balls is the region most likely to contain the evolution of $\mathcal{T}_+$ for at least $\mathcal{O}(\sqrt{\varepsilon}/\tau)$ time steps. This is consistent with the expected number of time steps, namely $\mathcal{O}(\sqrt{\varepsilon}/\tau)$, between consecutive mesh changes as arising from (3.6). We define

$$\hat{h}_{\mathcal{C}_S} := \tau \min\big(\max(\lambda_1, |v_S|), H\big), \tag{3.9}$$

for the meshsize within the cone $\mathcal{C}_S$, or $\tau\lambda_1$ in case $\mathcal{C}_S$ is a ball, which is a truncated version of (3.1). The constants μ_3 and H are arbitrary and result from computational considerations; μ_3 may depend on n.

3.2.2. Mesh Admissibility Tests. We enforce the following three quality criteria.

TEST 1. It checks whether

$$\mathcal{T}_+ \subset \mathcal{R}. \tag{3.10}$$

In the event the enlarged transition region $\mathcal{T}_+$ escapes from the refined region $\mathcal{R}$, we say that TEST(n) = FAIL. Then the solution X, Θ is discarded and the previous solution is used to generate the next mesh $\hat{\mathcal{M}}$. Rejection of $\mathcal{M}$ is mostly dictated by failure of Test 1. If $\mathcal{T}_+$ just touches the boundary elements of $\mathcal{R}$ (RED ZONE), then TESTS(n) = ALERT, and X, Θ are accepted and used to generate the next mesh $\hat{\mathcal{M}}$.

TEST 2. This test ascertains that interpolation errors outside $\mathcal{W}$ (defined in (3.8)) are still equidistributed correctly in the maximum norm

$$h_e \le \mu_1^* \hat{h}_e \quad \forall e \in \mathcal{E}\backslash\mathcal{E}_{\mathcal{W}}, \quad h_S \le \mu_2^* \hat{h}_S \quad \forall S \in \mathcal{M}\backslash\mathcal{M}_{\mathcal{W}}; \tag{3.11}$$

here $\mu_1^*, \mu_2^* > 0$ are suitable constants. This rules out the possibility of an excessive refinement induced by large discrete derivative. However, the new local meshsize might be much smaller than the current one, if influenced by the new refined region. If (3.11) is violated, then TESTS(n) = FAIL.

TEST 3. This test finally verifies whether or not (3.1) is valid within $\mathcal{T}_+$. Since the velocity of $\mathcal{T}_+$ may vary substantially during a period of time of order $\mathcal{O}(\tau/\sqrt{\varepsilon})$, thus making (3.9) inadequate as an error indicator, we enforce the constraints

$$\mu_3^- \hat{h}_{\mathcal{C}_S} \le h_S \le \mu_3^+ \hat{h}_{\mathcal{C}_S} \quad \forall S \in \mathcal{M}_{\mathcal{T}_+}; \tag{3.12}$$

$\mu_3^- < 1 < \mu_3^+$ are suitable constants. If (3.12) is violated, then TESTS(n) = FAIL.

3.3. Flow Chart. Our adaptive FEM is summarized schematically as follows:

$n = 1$	
$\mathcal{H}^1$ = H (1)	(constructs mesh density from θ_0, χ_0)
$\mathcal{M}^1$ = MESH (1)	(generates $\mathcal{M}^1$ from $\mathcal{H}^1$)
INTERPOLATION (1)	(computes $\hat{\Theta}_h^0 = \Pi^1\theta_0$, $\hat{X}_h^0 = \Pi^1\chi_0$)
do (while $n < N$)	
{	
MATRIX (n)	(constructs stiffness and mass matrices)
if ($\mathcal{M}^n \ne \mathcal{M}^{n-1}$ or $n = 1$)	
{X_h^n = PHASE (n)}	(solves (2.12) with $\hat{\Theta}_h^{n-1}$, $\hat{X}_h^{n-1}$)
Θ_h^n = TEMPERATURE (n)	(solves (2.13) with $\hat{\Theta}_h^{n-1}$, $\hat{X}_h^{n-1}$, X_h^n)
X_h^{n+1} = PHASE $(n+1)$	(solves (2.12) with Θ_h^n, X_h^n)
TESTS (n)	(constructs mesh parameters from Θ_h^n, X_h^n, X_h^{n+1} and checks the three admissibility tests on $\mathcal{M}^n$)
if (TESTS (n) = PASS)	
{$\mathcal{M}^{n+1} = \mathcal{M}^n$, $\hat{\Theta}_h^n = \Theta_h^n$, $\hat{X}_h^n = X_h^n$, accept X_h^{n+1}}	
else	

```
  {
    if (TESTS (n) = FAIL)
    {
      SOL_UPDATE (n)           (\Theta_h^n \leftarrow \Theta_h^{n-1}, X_h^n \leftarrow X_h^{n-1}, \mathcal{T}_+^n \leftarrow \mathcal{T}_+^{n-1})
      n \leftarrow n - 1
    }
    \mathcal{H}^{n+1} = H (n+1)          (constructs mesh density from \Theta_h^n, X_h^n, \mathcal{T}_+^n)
    \mathcal{M}^{n+1} = MESH (n+1)       (generates \mathcal{M}^{n+1} with \mathcal{H}^{n+1})
    INTERPOLATION (n+1)       (computes \hat{\Theta}_h^n = \Pi^{n+1}\Theta_h^n, \hat{X}_h^n = \Pi^{n+1} X_h^n)
  }
  n \leftarrow n + 1
}
```

3.4. Main Computer Procedures. We now describe the various procedures in the flow chart above. A uniform background grid $\mathcal{G}$ of size $\mathcal{O}(\sqrt{\tau})$ is first constructed and used throughout the computation to store the desired local mesh size. This is required by the advancing front mesh generator TMG of [17], which is called by MESH.

For $n = 1$, H(1) constructs $\hat{h}_e$, $\hat{h}_S$ and $\hat{h}_{\mathcal{C}_S}$ over $\mathcal{G}$ with information extracted from θ_0 and χ_0, averages these quantities out, and creates the mesh density $\mathcal{H}^1$ over $\mathcal{G}$. The first mesh $\mathcal{M}^1 =$ MESH(1) is next generated by TMG, and both θ_0 and χ_0 interpolated over $\mathcal{M}^1$ via INTERPOLATION(1) to start the calculation.

For a generic n, MATRIX(n) computes the stiffness and mass matrices. The procedure PHASE(n) computes X_h^n upon updating $\hat{X}_h^{n-1}$ nodewise according to (2.12). The routine TEMPERATURE(n) solves the linear system (2.13) via a conjugate gradient method preconditioned with an incomplete Cholesky factorization; the output is Θ_h^n. With Θ_h^n, X_h^n, X_h^{n+1}, TESTS(n) computes the local mesh parameters $\hat{h}_e$, $\hat{h}_S$, and $\hat{h}_{\mathcal{C}_S}$, and checks whether $\mathcal{M}^n$ meets the mesh admissibility tests. Whenever TESTS$(n) =$ FAIL, the computed solutions X_h^n, Θ_h^n are discarded, SOL_UPDATE(n) retrieves the previous solutions X_h^{n-1}, Θ_h^{n-1} and updates n to $n-1$. This is an undesirable event that sometimes can be avoided with the option TESTS$(n) =$ ALERT, which accepts the current solution but leads to immediate remeshing.

Once remeshing is invoked by TESTS$(n) =$ FAIL or ALERT, a new *graded* mesh $\mathcal{M}^{n+1}$ is generated as follows. The procedure H$(n+1)$ constructs the mesh density $\mathcal{H}^{n+1}$ over $\mathcal{G}$: it first transfers $\hat{h}_e$, $\hat{h}_S$, $\hat{h}_{\mathcal{C}_S}$ to $\mathcal{G}$ and next smooths these parameters out to eliminate any possible oscillation. The set $\mathcal{W} = \mathcal{W}^n$ of (3.8), where interpolation errors may be inadmissibly large, must not be modified by the mesh change. Therefore

$$S \in \mathcal{M}^{n+1} \quad \forall S \in \mathcal{M}^n_{\mathcal{W}} \tag{3.13}$$

is the first restriction on $\mathcal{M}^{n+1}$. The desired mesh density $\mathcal{H}^{n+1}$ satisfies

$$\lambda_1 \tau \le h_{\hat{S}} \le \min_{\substack{S' \in \mathcal{M}^n_{\mathcal{T}_+} : \mathcal{C}_{S'} \cap \hat{S} \neq \emptyset \\ e \in \mathcal{E}^n \setminus \mathcal{E}^n_{\mathcal{W}} : e \cap \hat{S} \neq \emptyset \\ S \in \mathcal{M}^n \setminus \mathcal{M}^n_{\mathcal{W}} : S \cap \hat{S} \neq \emptyset}} \left(\lambda_2 \tau^{1/2}, \hat{h}_{\mathcal{C}_{S'}}, \hat{h}_e, \hat{h}_S\right) \quad \forall \hat{S} \in \hat{\mathcal{M}}. \tag{3.14}$$

A new mesh $\mathcal{M}^{n+1} =$ MESH$(n+1)$ of properly graded triangles, satisfying (3.13) and (3.14), is generated by TMG.

Finally INTERPOLATION$(n+1)$ constructs the piecewise linear interpolants of Θ_h^n and X_h^n over $\mathcal{M}^{n+1}$. Since $\mathcal{M}^n$ and $\mathcal{M}^{n+1}$ are incompatible, interpolation is a nontrivial task that requires an efficient implementation via quadtrees; see [14].

Note that regardless of whether $\mathcal{M}^n$ is admissible or not, (3.10) and (3.13) yield

$$\Pi^n = \Pi^{n+1} \quad \text{over} \quad \mathcal{W}^n = \mathcal{T}_+^n \cup \mathcal{B}^n, \tag{3.15}$$

along with

$$\mathcal{T}_+^n \cup \mathcal{T}_+^{n+1} \subset \mathcal{R}^{n+1}. \tag{3.16}$$

Since compatibility is enforced over $\mathcal{T}_+^n$, namely $\Pi^n = \Pi^{n+1}$ on $\mathcal{T}_+^n$, and $X_h^n = X_h^{n+1} = 1$ or -1 in $\Omega \backslash \mathcal{T}_+^n$, no error will be incurred on interpolating X_h^n and X_h^{n+1}, that is

$$\hat{X}_h^n = \Pi^{n+1} X_h^n = X_h^n, \quad \Pi^n X_h^{n+1} = X_h^{n+1}. \tag{3.17}$$

Fixing all elements of $\mathcal{M}_{\mathcal{W}}^n$ is nearly optimal in a general setting; the resulting method is labeled AM1. For computational purposes, however, it is always preferable to remove the constraint (3.13), which leads to the alternative method AM0. We refer to [13,14] for more details in the context of the two-phase Stefan problem, and to §7.2 for a comparison between AM1 and AM0. They are practically equivalent.

4. Interpolation Theory for Incompatible Meshes. We now briefly review the interpolation theory developed in [13,§5] for piecewise smooth functions, and employ the notation of §3.2. In view of (3.13), we only have to be concerned with the complement of $\mathcal{W}$ for interpolation purposes. Therefore, from now on elements $S \in \mathcal{M}$ and $\hat{S} \in \hat{\mathcal{M}}$ do not belong to $\mathcal{W}$. Given a set D, let D_+ indicate the union of all simplices $S \in \mathcal{M}$ intersecting D. Regularity of $\mathcal{M}$ yields the existence of $0 < a < 1$ such that, for all $S \in \mathcal{M}$,

$$\rho_{S'} \geq a h_S \quad \forall S' \in \mathcal{M}_S, \quad \operatorname{dist}(x, S) \geq a h_S \quad \forall x \in \bar{\Omega} \backslash S_+.$$

We then have the following two mutually exclusive situations. We only state L^∞ estimates without proof. We refer to [13] for L^p variants of these estimates for $1 \leq p \leq \infty$, along with complete proofs in 2D which easily extend to $d \geq 2$.

LEMMA 4.1. (Derefinement Case) *Let $\hat{S} \in \hat{\mathcal{M}}$ be such that $a h_S < h_{\hat{S}}$ for all $S \in \mathcal{M}_{\hat{S}}$. Then*

$$\|\Theta - \hat{\Theta}\|_{L^\infty(\hat{S})} + h_{\hat{S}} \|\nabla(\Theta - \hat{\Theta})\|_{L^\infty(\hat{S})} \leq C h_{\hat{S}}^2 \max_{e \in \mathcal{E}_{\hat{S}_+}} D_e.$$

LEMMA 4.2. (Refinement Case) *Let $\hat{S} \in \hat{\mathcal{M}}$ be such that there exists $S \in \mathcal{M}_{\hat{S}}$ satisfying $a h_S \geq h_{\hat{S}}$. Then*

$$\|\Theta - \hat{\Theta}\|_{L^\infty(S)} + h_{\hat{S}} \|\nabla(\Theta - \hat{\Theta})\|_{L^\infty(\hat{S})} \leq C h_{\hat{S}} \max_{e \in \mathcal{E}_{\hat{S}}} h_e D_e.$$

The following result resembles Lemma 5.3 of [15] which is also a consequence of the above abstract interpolation theory, but our result does not rely on a 2D argument and thus holds for $\mathbb{R}^d$ with $d \geq 1$.

COROLLARY 4.1. *There holds* $\|\Theta - \hat{\Theta}\|_{L^\infty(\Omega)} + \|U - \hat{U}\|_{L^\infty(\Omega)} \leq C\tau$.

Proof. Since $\hat{X} = X$ and $U = \Theta + X$, then $U - \hat{U} = \Theta - \hat{\Theta}$. It thus suffices to estimate $\Theta - \hat{\Theta}$. First we consider the derefinement case. On using Lemma 4.1, (3.7), and (3.14) we get

$$\|\Theta - \hat{\Theta}\|_{L^\infty(\hat{S})} \leq C \max_{e \in \mathcal{E}_{\hat{S}_+}} \left(\hat{h}_e^2 D_e\right) \leq C\tau.$$

Likewise, for the refinement case we make use of Lemma 4.2 together with (3.7), (3.11), and (3.14), to arrive at

$$\|\Theta - \hat{\Theta}\|_{L^\infty(\hat{S})} \le C \max_{e\in\mathcal{E}_{\hat{S}}} \left(h_{\hat{S}} h_e D_e\right) \le C \max_{e\in\mathcal{E}_{\hat{S}}} \left(\hat{h}_e^2 D_e\right) \le C\tau. \qquad \square$$

5. Stability. In what follows, we will assume the *structural* property

$$\sum_{e\in\mathcal{E}^n} h_e^d \left(D_e^n\right)^2 \le B\varepsilon^{-1}, \tag{5.1}$$

which is a little stronger than $\varepsilon \sum_{n=1}^N \tau \|\Theta^n\|_{H^2(\Omega)}^2 \le C$ of (2.5). It is worth comparing our structural assumption with the one in [13,15] for the two-phase Stefan problem with $d = 2$, namely $\sum_{e\in\mathcal{E}^n} h_e^2 D_e^n \le C$. This mimics the a priori bound $\|\Delta\theta\|_{L^\infty(0,T;M(\Omega))} \le C$, where $M(\Omega)$ is the space of finite regular Baire measure [8]. Since the latter is not known for the phase relaxation model (1.1), (1.2), we can view (5.1) as a proper replacement that even removes the restriction of [13,14,15] to $d = 2$.

We now derive a priori estimates in energy norms. Comparing with [13,15], we now require a limitation on the number of mesh changes of order $\mathcal{O}(1/\sqrt{\varepsilon})$ instead of $\mathcal{O}(1/\sqrt{\tau})$. Summation over n counting only mesh changes is denoted by $\sum_n^*$.

Lemma 5.1. *Let $\tau \le \varepsilon$ be small and the number of mesh changes $N_c \le M/\sqrt{\varepsilon}$. There exists a constant C, depending only on A in* (2.1)*, B in* (5.1)*, and M, such that*

$$\frac{1}{\tau}\sum_{n=1}^N \|\Theta_h^n - \hat{\Theta}_h^{n-1}\|_{L^2(\Omega)}^2 + \max_{1\le n\le N} \|\nabla\Theta_h^n\|_{L^2(\Omega)}^2 + \sum_{n=1}^N \|\nabla(\Theta_h^n - \hat{\Theta}_h^{n-1})\|_{L^2(\Omega)}^2 \le C,$$

$$\sum_{n=1}^N \|(X_h^n - X_h^{n-1}) - (X_h^{n-1} - X_h^{n-2})\|_{L^2(\Omega)}^2 + \max_{1\le n\le N} \|X_h^n - X_h^{n-1}\|_{L^2(\Omega)}^2 \le C\frac{\tau^2}{\varepsilon}.$$

The first estimate yields $\|\nabla\hat{\Theta}_h^{n-1}\|_{L^2(\Omega)} \le C$, which will be used in (6.6) below.

Proof. Since $\hat{X}_h^{n-1} = \Pi^n X_h^{n-1} = X_h^{n-1}$ in view of (3.17), we will avoid writing the hat. We take $\varphi := \Theta_h^n - \hat{\Theta}_h^{n-1} \in \mathbb{V}^n$ as a test function in (2.11). Upon adding the resulting expression over $1 \le n \le m$, we obtain

$$\begin{aligned} 0 = \frac{1}{\tau}\sum_{n=1}^m \|\Theta_h^n - \hat{\Theta}_h^{n-1}\|_{L^2(\Omega)}^2 + \frac{1}{\tau}\sum_{n=1}^m \langle X_h^n - X_h^{n-1}, \Theta_h^n - \hat{\Theta}_h^{n-1}\rangle_n \\ + \sum_{n=1}^m \langle\nabla\Theta_h^n, \nabla(\Theta_h^n - \hat{\Theta}_h^{n-1})\rangle =: \mathrm{I} + \mathrm{II} + \mathrm{III}. \end{aligned}$$

The presence of $\hat{\Theta}_h^{n-1}$ instead of Θ_h^{n-1} in terms II and III is the main difference with respect to Lemma 4.1 of [11]. We examine this effect separately in terms II and III.

Since $X_h^n - X_h^{n-1} = 0$ in $\Omega\backslash\mathcal{T}_+^{n-1}$, as a consequence of (2.14), the integral in term II can be restricted to $\mathcal{T}_+^{n-1}$. We further invoke (3.15), which allows us to replace Π^n by Π^{n-1} over $\mathcal{T}_+^{n-1}$, and thereby exchange $\langle\cdot,\cdot\rangle_n$ with $\langle\cdot,\cdot\rangle_{n-1}$, to write

$$\mathrm{II} = \frac{1}{\tau}\sum_{n=1}^m \langle X_h^n - X_h^{n-1}, \Theta_h^n - \hat{\Theta}_h^{n-1}\rangle_{n-1}.$$

Upon subtracting (2.10) for n and $n-1$ we readily arrive at

$$\Theta_h^n - \hat\Theta_h^{n-1} = \tfrac{\varepsilon}{\tau}\big((X_h^n - X_h^{n-1}) - (X_h^{n-1} - X_h^{n-2})\big) + (Z_h^n - Z_h^{n-1}) \\ + \big((\Theta_h^n - \hat\Theta_h^{n-1}) - (\Theta_h^{n-1} - \hat\Theta_h^{n-2})\big) + (\Theta_h^{n-1} - \hat\Theta_h^{n-1}).$$

We multiply this expression by $X_h^n - X_h^{n-1} =: \tau\partial X_h^n$, and observe that $\Theta_h^{n-1} - \hat\Theta_h^{n-1} = 0$ in $\mathcal{T}_+^{n-1}$, to infer that

$$\mathrm{II} = \varepsilon\sum_{n=1}^m \langle\partial X_h^n, \partial X_h^n - \partial X_h^{n-1}\rangle_{n-1} + \sum_{n=1}^m\langle\partial X_h^n, Z_h^n - Z_h^{n-1}\rangle_{n-1} \\ + \sum_{n=1}^m\big(\langle\partial X_h^n, \Theta_h^n - \hat\Theta_h^{n-1}\rangle_n - \langle\partial X_h^n, \Theta_h^{n-1} - \hat\Theta_h^{n-2}\rangle_{n-1}\big) =: \mathrm{II}_1 + \mathrm{II}_2 + \mathrm{II}_3.$$

We now recall the elementary identity

$$2a(a-b) = (a-b)^2 + a^2 - b^2 \quad \forall a, b\in\mathbb{R}, \tag{5.2}$$

to rewrite term II_1 as

$$2\mathrm{II}_1 = \varepsilon\sum_{n=1}^m\big(\|\partial X_h^n - \partial X_h^{n-1}\|_{n-1}^2 + \|\partial X_h^n\|_{n-1}^2 - \|\partial X_h^{n-1}\|_{n-1}^2\big).$$

In view of (3.17) we have $\Pi^{n-2}(X_h^{n-1} - X_h^{n-2}) = \Pi^{n-1}(X_h^{n-1} - X_h^{n-2}) = X_h^{n-1} - X_h^{n-2}$. Consequently $\|\partial X_h^{n-1}\|_{n-1} = \|\partial X_h^{n-1}\|_{n-2}$, whence the above summation telescopes. By virtue of (2.15), we conclude

$$2\mathrm{II}_1 = \varepsilon\|\partial X_h^m\|_m^2 + \varepsilon\sum_{n=1}^m\|\partial X_h^n - \partial X_h^{n-1}\|_{n-1}^2 - C.$$

To tackle term II_2 we note that there is no contribution from outside of $\mathcal{T}_+^{n-1}$. Therefore (3.15), in conjunction with $Z_h^n(x_j^n)\in\Lambda(X_h^n(x_j^n))$ for all nodes x_j^n and monotonicity of Λ, implies $\mathrm{II}_2\ge 0$. It is not difficult to check the validity of the following identity (summation by parts), where $\Theta_h^0 = \hat\Theta_h^{-1}$ yields

$$\mathrm{II}_3 = \langle\partial X_h^m, \Theta_h^m - \hat\Theta_h^{m-1}\rangle_m - \sum_{n=1}^m\langle\partial X_h^n - \partial X_h^{n-1}, \Theta_h^{n-1} - \hat\Theta_h^{n-2}\rangle_{n-1}.$$

Since each summand vanishes outside $\mathcal{T}_+^{n-1}\cup\mathcal{T}_+^{n-2}$, (3.16) leads to

$$|\mathrm{II}_3| \le \tfrac13\tau\|\partial X_h^m\|_m^2 + \tfrac13\tau\sum_{n=1}^m\|\partial X_h^n - \partial X_h^{n-1}\|_{n-1}^2 + \tfrac34\tfrac1\tau\sum_{n=1}^m\|\Theta_h^n - \hat\Theta_h^{n-1}\|_{\mathcal{R}^n,n}^2.$$

We resort now to (2.8) for $k=1$ and $\varphi = \psi = \Theta_h^n - \hat\Theta_h^{n-1}$, in conjunction with $h_S\le C\tau$ within $\mathcal{R}^n$, to arrive at

$$|\mathrm{II}_3| \le \tfrac13\tau\|\partial X_h^m\|_m^2 + \tfrac13\tau\sum_{n=1}^m\|\partial X_h^n - \partial X_h^{n-1}\|_{n-1}^2 \\ + \tfrac34\tfrac1\tau\sum_{n=1}^m\|\Theta_h^n - \hat\Theta_h^{n-1}\|_{L^2(\Omega)}^2 + C^*\tau\sum_{n=1}^m\|\nabla(\Theta_h^n - \hat\Theta_h^{n-1})\|_{L^2(\Omega)}^2,$$

where $C^*>0$ is an interpolation constant.

It only remains to estimate term III, which is first decomposed as

$$\mathrm{III} = \sum_{n=1}^{m} \langle \nabla\Theta_h^n, \nabla(\Theta_h^n - \Theta_h^{n-1})\rangle + \sum_{n=1}^{m} \langle \nabla\Theta_h^n, \nabla(\Theta_h^{n-1} - \hat\Theta_h^{n-1})\rangle =: \mathrm{III}_1 + \mathrm{III}_2.$$

It is clear from (5.2) that

$$2\mathrm{III}_1 = \|\nabla\Theta_h^m\|_{L^2(\Omega)} - \|\nabla\Theta_h^0\|^2_{L^2(\Omega)} + \sum_{n=1}^{m} \|\nabla(\Theta_h^n - \Theta_h^{n-1})\|^2_{L^2(\Omega)}.$$

To bound term III_2, we resort to our structural assumption (5.1) and the restriction on the number of mesh changes. On integrating elementwise by parts, we get

$$\begin{aligned}|\mathrm{III}_2| &\le {\sum_{n=1}^{m}}^{*} \sum_{e\in\mathcal{E}^n} \big|\langle\!\langle [\![\nabla\Theta_h^n]\!]_e \cdot \nu_e, \Theta_h^{n-1} - \hat\Theta_h^{n-1}\rangle\!\rangle_e\big| \\ &\le {\sum_{n=1}^{m}}^{*} \|\Theta_h^{n-1} - \hat\Theta_h^{n-1}\|_{L^\infty(\Omega)} \sum_{e\in\mathcal{E}^n} h_e^d D_e^n \\ &\le C\, M\, \tfrac{\tau}{\sqrt{\varepsilon}} \max_{1\le n\le m} \Big(\sum_{e\in\mathcal{E}^n} h_e^d\Big)^{1/2} \Big(\sum_{e\in\mathcal{E}^n} h_e^d (D_e^n)^2\Big)^{1/2} \le C\tfrac{\tau}{\varepsilon}.\end{aligned}$$

Collecting the above estimates, and assuming $1 - 2C^*\tau \ge 1/2$, we readily have

$$\begin{aligned}&\tfrac12\tfrac1\tau \sum_{n=1}^{m} \|\Theta_h^n - \hat\Theta_h^{n-1}\|^2_{L^2(\Omega)} + \|\nabla\Theta_h^m\|^2_{L^2(\Omega)} + \tfrac12 \sum_{n=1}^{m} \|\nabla(\Theta_h^n - \Theta_h^{n-1})\|^2_{L^2(\Omega)} \\ &\quad + \varepsilon\big(1 - \tfrac23\tfrac\tau\varepsilon\big)\|\partial X_h^m\|_m^2 + \varepsilon\big(1 - \tfrac23\tfrac\tau\varepsilon\big) \sum_{n=1}^{m} \|\partial(X_h^n - X_h^{n-1})\|^2_{n-1} \le C\big(1 + \tfrac\tau\varepsilon\big).\end{aligned}$$

Imposing the stability constraint $\tau \le \varepsilon$ we obtain the asserted estimates. ∎

COROLLARY 5.1. *There holds* $\max_{1\le n\le N} \|U_h^n\|_{L^2(\Omega)} \le C$.

Proof. Since $|X_h^n| \le 1$, it is enough to prove the assertion for Θ_h^n. To this end we simply use the telescopic expansion $\Theta_h^m = \Theta_h^0 + \sum_{n=1}^m \big((\Theta_h^n - \hat\Theta_h^{n-1}) + (\hat\Theta_h^{n-1} - \Theta_h^{n-1})\big)$, and apply Lemma 5.1 and Corollary 4.1. ∎

The last bound of Lemma 5.1 yields the following weak statement about the distance between consecutive transition regions $\mathcal{T}^n$ and $\mathcal{T}^{n+1}$, via $\mathcal{A}^n$ of (2.14).

COROLLARY 5.2. *There holds* $\max_{1\le n\le N} |\mathcal{A}^n| \le C\tau^2/\varepsilon$.

Proof. Simply use the fact that $X_h^{n+1} - X_h^n = 2$ or -2 in the set $\mathcal{A}^n$ and $\|X_h^{n+1} - X_h^n\|^2_{L^2(\Omega)} \le C\tau^2/\varepsilon$. ∎

Set $\mathcal{A}^n$ is in fact empty provided $\tau \ll \varepsilon$ and $d = 2$ as the following result reveals.

LEMMA 5.2. *Let* $d = 2$. *There exists* $C_0 > 0$ *such that if* $\tau \le C_0\varepsilon|\log\varepsilon|^{-1/2}$, *then* $\mathcal{A}^n$ *is empty.*

Proof. Since $d = 2$ and Θ_h^n is bounded in $H^1(\Omega)$, an inverse inequality implies [4]

$$\|\Theta_h^n\|_{L^\infty(\Omega)} \le C|\log\tau|^{1/2}\|\Theta_h^n\|_{H^1(\Omega)}.$$

Here we have used that the smallest meshsize is of order $\mathcal{O}(\tau)$. We then consider a node $x_j^n \in \mathcal{A}^n$ such that $X_h^{n+1}(x_j^n) - X_h^n(x_j^n) = 2$. In view of (2.10) and the monotonicity of Λ, we deduce

$$2\tfrac{\varepsilon}{\tau} = \tfrac{\varepsilon}{\tau}\big(X_h^{n+1}(x_j^n) - X_h^n(x_j^n)\big) \le \Theta_h^n(x_j^n) \le C|\log\tau|^{1/2},$$

which is a contradiction if τ is chosen as asserted. ∎

LEMMA 5.3. $\max_{1\le n\le N}\|\nabla X_h^n\|^2_{L^2(\Omega)} + \sum_{n=1}^N \|\nabla(X_h^n - X_h^{n-1})\|^2_{L^2(\Omega)} \le C/\varepsilon$.

Proof. In light of (3.17), proceed exactly as in Lemma 4.2 of [11]. □

6. Error Analysis. We are now in a position to carry out the error analysis. Rather than using the *integration method* of [11], which would need test functions to be admissible for all time steps, we argue as in [10,13,15] with the *inverse Laplacian*. Let $H^1(\Omega)^*$ be the dual of $H^1(\Omega)$, and $G: H^1(\Omega)^* \to H^1(\Omega)$ be the Green's operator

$$\langle G\psi,\phi\rangle + \langle\nabla G\psi,\nabla\phi\rangle = \langle\psi,\phi\rangle \quad \forall\phi\in H^1(\Omega). \tag{6.1}$$

Since Ω is smooth (or convex), elliptic regularity theory implies that G is *regular* [4,p.138], that is

$$\|G\psi\|_{H^2(\Omega)} \le C\|\psi\|_{L^2(\Omega)} \quad \forall\psi\in L^2(\Omega). \tag{6.2}$$

In addition, the norm in $H^1(\Omega)^*$ can be represented in terms of G as

$$\|\psi\|^2_{H^1(\Omega)^*} := \|G\psi\|^2_{L^2(\Omega)} + \|\nabla G\psi\|^2_{L^2(\Omega)} = \langle\psi, G\psi\rangle \quad \forall\psi\in H^1(\Omega)^*. \tag{6.3}$$

We now consider the critical issue of extending the nodewise relation $Z_h^n(x_j^n)\in\Lambda(X_h^n(x_j^n))$ to the entire Ω. With the notion of dual mesh $\mathcal{D}^n$ of $\mathcal{M}^n$ and interpolation operator P^n introduced in [11], and recalled before (2.9), the discrete phase equation (2.10) can be rewritten equivalently as

$$\tfrac{\varepsilon}{\tau}(\tilde X_h^n - \tilde X_h^{n-1}) + \tilde Z_h^n = \tilde{\hat\Theta}_h^{n-1} = P^n(\Pi^n\Theta_h^{n-1}), \quad \tilde Z_h^n\in\Lambda(\tilde X_h^n). \tag{6.4}$$

Since $h_S = \mathcal{O}(\tau)$ within $\mathcal{T}^n\subset\mathcal{R}^n$, (2.9) combined with Lemmas 5.3 and 5.1 yields

$$\|X_h^n - \tilde X_h^n\|_{L^2(\Omega)} \le C\tau\|\nabla X_h^n\|_{L^2(\mathcal{R}^n)} \le C\tfrac{\tau}{\sqrt{\varepsilon}}, \tag{6.5}$$

$$\|\hat\Theta_h^{n-1} - \tilde{\hat\Theta}_h^{n-1}\|_{L^2(\mathcal{R}^n)} \le C\tau\|\nabla\hat\Theta_h^{n-1}\|_{L^2(\mathcal{R}^n)} \le C\tau. \tag{6.6}$$

In contrast to [11], where it is used globally, (6.4) is no longer adequate away from $\mathcal{R}^n$ in the present context. The reason, which hinges on accuracy considerations, will become apparent in the proof of Theorem 6.1. We exploit the fact that the nodewise relation (2.10) is valid pointwise away from $\mathcal{T}^n$, and so in $\Omega\backslash\mathcal{R}^n$, that is

$$\tfrac{\varepsilon}{\tau}(X_h^n - X_h^{n-1}) + Z_h^n = \hat\Theta_h^{n-1}, \quad Z_h^n\in\Lambda(X_h^n) \quad \text{in } \Omega\backslash\mathcal{T}^n. \tag{6.7}$$

In fact, if $X^n = 1$ (resp. -1) in an element S, then $0\le$ (resp. $\ge$) $Z_h^n\in\Lambda(X_h^n)$ in S.

Now we are ready to prove our main result of this paper. The abbreviations

$$e_v^n := V^n - V_h^n, \quad \tilde e_v^n := V^n - \tilde V_h^n$$

are used for $v=\theta,\chi,u$, where V^n represents the semi-discrete solution of (2.3), (2.4) and V_h^n the finite element solution of (2.10), (2.11). Note that (6.8) below does not have a factor $|\log\tau|$ as in [13,15].

THEOREM 6.1. *Let $\{\mathcal{M}^n\}_{n=1}^N$ be a sequence of meshes satisfying the admissibility criteria of §3. Let $\tau\le\varepsilon$ be small and the number of mesh changes satisfy $N_c\le M/\sqrt{\varepsilon}$. Then there exists a constant C, solely depending on A in (2.1), B in (5.1), and M, such that the following asymptotic rates of convergence are valid*

$$\varepsilon\max_{1\le n\le N}\|e_\chi^n\|^2_{L^2(\Omega)} + \sum_{n=1}^N\tau\|e_\theta^n\|^2_{L^2(\Omega)} + \max_{1\le n\le N}\|e_u^n\|^2_{H^1(\Omega)^*} \le C\,\tfrac{\tau^2}{\varepsilon}. \tag{6.8}$$

Proof. We subtract (2.11) from (2.3) to obtain the first error equation

$$\begin{aligned}\langle e_u^n - e_u^{n-1}, \phi\rangle + \tau\langle\nabla e_\theta^n, \nabla\phi\rangle &= \big(\langle X_h^n - X_h^{n-1}, \varphi\rangle_n - \langle X_h^n - X_h^{n-1}, \varphi\rangle\big)\\ &+ \langle \hat{U}_h^{n-1} - U_h^n, \phi - \varphi\rangle + \langle \Theta_h^{n-1} - \hat{\Theta}_h^{n-1}, \phi\rangle + \tau\langle\nabla\Theta_h^n, \nabla(\varphi-\phi)\rangle,\end{aligned} \tag{6.9}$$

valid for all $\phi \in H^1(\Omega)$, $\varphi \in \mathbb{V}^n$. We take the test functions to be $\phi = \phi^n := Ge_u^n$ and $\varphi = \varphi^n$ the Clément interpolant of ϕ^n, which is local, requires no continuity of ϕ^n, and satisfies $\|\nabla\varphi^n\|_{L^2(\Omega)} \le C\|\nabla\phi^n\|_{L^2(\Omega)} \le C\|e_u^n\|_{H^1(\Omega)^*}$ [4,p.145]. Let

$$\|e_u^m\|_{H^1(\Omega)^*} = \max_{1\le n\le N} \|e_u^n\|_{H^1(\Omega)^*} \tag{6.10}$$

be satisfied for $1 \le m \le N$. We then add over n from 1 to m, and use (6.1) to derive

$$\begin{aligned}\mathrm{I} + \mathrm{II} : &= \sum_{n=1}^m \langle e_u^n - e_u^{n-1}, Ge_u^n\rangle + \sum_{n=1}^m \tau\langle e_\theta^n, e_u^n\rangle\\ &= \sum_{n=1}^m \big(\langle X_h^n - X_h^{n-1}, \varphi^n\rangle_n - \langle X_h^n - X_h^{n-1}, \varphi^n\rangle\big)\\ &\quad + \sum_{n=1}^m \langle \hat{U}_h^{n-1} - U_h^n, \phi^n - \varphi^n\rangle + \sum_{n=1}^m \langle \Theta_h^{n-1} - \hat{\Theta}_h^{n-1}, \phi^n\rangle\\ &\quad + \sum_{n=1}^m \tau\langle\nabla\Theta_h^n, \nabla(\varphi^n - \phi^n)\rangle + \sum_{n=1}^m \tau\langle e_\theta^n, \phi^n\rangle\\ &=: \mathrm{III} + \mathrm{IV} + \mathrm{V} + \mathrm{VI} + \mathrm{VII}.\end{aligned} \tag{6.11}$$

Now we subtract (6.4) from (2.4), multiply by $\tau\tilde{e}_\chi^n$ and integrate over $\mathcal{R}^n$. We then subtract (6.7) from (2.4), multiply by τe_χ^n and integrate it over $\Omega\backslash\mathcal{R}^n$. Adding the two resulting inequalities over n from 1 to m, the monotonicity of Λ gives rise to

$$\begin{aligned}\mathrm{VIII} := \varepsilon\sum_{n=1}^m \langle \tilde{e}_\chi^n - \tilde{e}_\chi^{n-1}, \tilde{e}_\chi^n\rangle \le &\sum_{n=1}^m \tau\langle \Theta^{n-1} - \tilde{\hat{\Theta}}_h^{n-1}, X^n - \tilde{X}_h^n\rangle^{\mathcal{R}^n}\\ &+ \sum_{n=1}^m \tau\langle \Theta^{n-1} - \hat{\Theta}_h^{n-1}, X^n - X_h^n\rangle^{\Omega\backslash\mathcal{R}^n} =: \mathrm{IX}.\end{aligned} \tag{6.12}$$

Here we have used that $X_h^n = \tilde{X}_h^n = 1$ or -1 in $\Omega\backslash\mathcal{R}^n$, which implies $e_\chi^n = \tilde{e}_\chi^n$ in $\Omega\backslash\mathcal{R}^n$. To derive our main result we add (6.11) and (6.12) and examine each term separately. First, it is clear from (5.2) and (6.3) that

$$2\mathrm{I} = \|e_u^m\|^2_{H^1(\Omega)^*} - \|e_u^0\|^2_{H^1(\Omega)^*} + \sum_{n=1}^m \|e_u^n - e_u^{n-1}\|^2_{H^1(\Omega)^*}.$$

Term II is simply decomposed as

$$\mathrm{II} = \sum_{n=1}^m \tau\|e_\theta^n\|^2_{L^2(\Omega)} + \sum_{n=1}^m \tau\langle e_\theta^n, e_\chi^n\rangle.$$

For term III, we notice that there is no contribution coming from $\Omega\backslash\mathcal{R}^n$, because $X_h^n - X_h^{n-1} = 0$ in $\Omega\backslash\mathcal{T}_+^{n-1}$, as a consequence of (2.14). Since $h_S = \mathcal{O}(\tau)$ within $\mathcal{R}^n$, we can use (2.8) with $k = 1$ and Lemma 5.3 to get

$$\begin{aligned}\mathrm{III} &\le C\tau^2 \sum_{n=1}^m \|\nabla(X_h^n - X_h^{n-1})\|_{L^2(\Omega)} \|e_u^n\|_{H^1(\Omega)^*} \\ &\le C\tau^3 \sum_{n=1}^m \|\nabla(X_h^n - X_h^{n-1})\|_{L^2(\Omega)}^2 + \tau \sum_{n=1}^m \|e_u^n\|_{H^1(\Omega)^*}^2 \le C\tau^2 + \tau \sum_{n=1}^m \|e_u^n\|_{H^1(\Omega)^*}^2.\end{aligned}$$

Now we notice that $\|\phi^n - \varphi^n\|_{L^2(S)} \le C h_S^2 \|D^2\phi^n\|_{L^2(S_+)}$ for all $S \in \mathcal{M}^n$, where S_+ stands for S and all surrounding elements. Then we split term IV into two parts

$$\mathrm{IV} = \sum_{n=1}^m \langle X_h^{n-1} - X_h^n, \phi^n - \varphi^n \rangle^{\mathcal{R}^n} + \sum_{n=1}^m \langle \hat\Theta_h^{n-1} - \Theta_h^n, \phi^n - \varphi^n \rangle := \mathrm{IV}_1 + \mathrm{IV}_2.$$

Since $h_S = \mathcal{O}(\tau)$ within $\mathcal{R}^n$, on using (6.2), Lemma 5.1, and Corollary 5.1 we see that

$$\mathrm{IV}_1 \le C\tau^2 \sum_{n=1}^m \|X_h^n - X_h^{n-1}\|_{L^2(\Omega)} \|e_u^n\|_{L^2(\Omega)} \le C\frac{\tau^2}{\sqrt{\varepsilon}}.$$

On the other hand, on using $h = \mathcal{O}(\sqrt{\tau})$, (6.2) and Lemma 5.1 once again we obtain

$$\begin{aligned}\mathrm{IV}_2 &\le C \sum_{n=1}^m h^2 \|\Theta_h^n - \hat\Theta_h^{n-1}\|_{L^2(\Omega)} \|e_u^n\|_{L^2(\Omega)} \\ &\le C\frac{\tau^2}{\varepsilon} + \varepsilon \sum_{n=1}^m \tau \|e_\chi^n\|_{L^2(\Omega)}^2 + \frac{1}{6} \sum_{n=1}^m \tau \|e_\theta^n\|_{L^2(\Omega)}^2.\end{aligned}$$

Since the number of mesh changes is $N_c \le M/\sqrt{\varepsilon}$, by virtue of Corollary 4.1 we can estimate term V as

$$\begin{aligned}\mathrm{V} &\le \max_{1\le n\le m} \|G e_u^n\|_{L^2(\Omega)} \sum_{n=1}^m{}^* \|\Theta_h^{n-1} - \hat\Theta_h^{n-1}\|_{L^2(\Omega)} \le C\tau N_c \|e_u^m\|_{H^1(\Omega)^*} \\ &\le \frac14 \|e_u^m\|_{H^1(\Omega)^*}^2 + C\frac{\tau^2}{\varepsilon}.\end{aligned}$$

We next integrate term VI by parts and split the resulting sum over edges $\mathcal{E}_{\mathcal{R}}^n$, namely $e \in \mathcal{E}^n$ inside $\mathcal{R}^n$, and the complement $\mathcal{E}_{\Omega\setminus\mathcal{R}}^n$,

$$\mathrm{VI}_1 + \mathrm{VI}_2 := \sum_{n=1}^m \tau \sum_{e\in\mathcal{E}_{\mathcal{R}}^n} \langle\!\langle [\![\nabla\Theta_h^n]\!]_e \cdot \nu_e, \phi^n - \varphi^n \rangle\!\rangle_e + \sum_{n=1}^m \tau \sum_{e\in\mathcal{E}_{\Omega\setminus\mathcal{R}}^n} \langle\!\langle [\![\nabla\Theta_h^n]\!]_e \cdot \nu_e, \phi^n - \varphi^n \rangle\!\rangle_e.$$

For any $e \in \mathcal{E}^n$, let $S \in \mathcal{M}^n$ be an element containing such an edge e. We make use of the trace inequality

$$\|\zeta\|_{L^2(e)} \le C h_e^{-1/2} \|\zeta\|_{L^2(S)} + C h_e^{1/2} \|\nabla\zeta\|_{L^2(S)},$$

in conjunction with $h_e \ge a h_S$ for some $0 < a < 1$ independent of n, which is a consequence of meshes $\mathcal{M}^n$ being regular uniformly in n, to deduce

$$\|\phi^n - \varphi^n\|_{L^2(e)} \le C h_e^{3/2} \|D^2\phi^n\|_{L^2(S_+)}.$$

Invoking the inverse inequality $\|\nabla\Theta_h^n\|_{L^2(e)} \le C h_e^{-1/2}\|\nabla\Theta_h^n\|_{L^2(S)}$, together with $h_e = \mathcal{O}(\tau)$ in $\mathcal{R}^n$ and (6.2), we get

$$\mathrm{VI}_1 \le C\tau^2 \sum_{n=1}^m \|\nabla\Theta_h^n\|_{L^2(\Omega)}\|e_u^n\|_{L^2(\Omega)} \le C\tfrac{\tau^2}{\varepsilon} + \varepsilon\sum_{n=1}^m \tau\|e_\chi^n\|_{L^2(\Omega)}^2 + \tfrac{1}{6}\sum_{n=1}^m \tau\|e_\theta^n\|_{L^2(\Omega)}^2.$$

For term VI_2, we recall (3.7) and (3.11), namely $h_e|[\![\nabla\Theta_h^n]\!]_e| = h_e^2 D_e \le C\tau$, and (6.2) once again, to obtain

$$\begin{aligned}
\mathrm{VI}_2 &\le C\sum_{n=1}^m \tau \sum_{e\in\mathcal{E}^n_{\Omega\setminus\mathcal{R}}} h_e^{(d+2)/2}|[\![\nabla\Theta_h^n]\!]_e|\ \|D^2\phi^n\|_{L^2(S_+)} \\
&\le C\sum_{n=1}^m \tau^2\Big(\sum_{e\in\mathcal{E}^n} h_e^d\Big)^{1/2}\|e_u^n\|_{L^2(\Omega)} \\
&\le C\tfrac{\tau^2}{\varepsilon} + \varepsilon\sum_{n=1}^m \tau\|e_\chi^n\|_{L^2(\Omega)}^2 + \tfrac{1}{6}\sum_{n=1}^m \tau\|e_\theta^n\|_{L^2(\Omega)}^2.
\end{aligned}$$

For term VII, we easily get

$$\mathrm{VII} \le \sum_{n=1}^m \tau\|e_\theta^n\|_{L^2(\Omega)}\|Ge_u^n\|_{L^2(\Omega)} \le \tfrac{1}{6}\sum_{n=1}^m \tau\|e_\theta^n\|_{L^2(\Omega)}^2 + C\sum_{n=1}^m \tau\|e_u^n\|_{H^1(\Omega)^*}^2.$$

For term VIII, we employ (5.2) in conjunction with

$$\|\tilde e_\chi^0\|_{L^2(\Omega)} \le \|\chi_0 - X_h^0\|_{L^2(\Omega)} + \|X_h^0 - \tilde X_h^0\|_{L^2(\Omega)} \le C\tfrac{\tau}{\sqrt{\varepsilon}},$$

as results from the choice of initial data, $\mathcal{T}^0 \subset \mathcal{R}^1$, and (6.5), to arrive at

$$2\mathrm{VIII} = \varepsilon\|\tilde e_\chi^m\|_{L^2(\Omega)}^2 - \varepsilon\|\tilde e_\chi^0\|_{L^2(\Omega)}^2 + \varepsilon\sum_{n=1}^m \|\tilde e_\chi^n - \tilde e_\chi^{n-1}\|_{L^2(\Omega)}^2 \ge \tfrac{1}{2}\varepsilon\|e_\chi^m\|_{L^2(\Omega)}^2 - C\tau^2.$$

Here we have used (6.5) to replace $\tilde X_h^m$ by X_h^m. We finally decompose term IX as

$$\begin{aligned}
\mathrm{IX} &= \sum_{n=1}^m \tau\langle\Theta^n - \Theta_h^n, X^n - X_h^n\rangle + \sum_{n=1}^m \tau\langle\Theta^{n-1} - \Theta^n, X^n - X_h^n\rangle \\
&\quad + \sum_{n=1}^m \tau\langle\Theta_h^n - \hat\Theta_h^{n-1}, X^n - X_h^n\rangle + \sum_{n=1}^m \tau\langle\hat\Theta_h^{n-1} - \tilde{\hat\Theta}_h^{n-1}, X^n - X_h^n\rangle^{\mathcal{R}^n} \\
&\quad + \sum_{n=1}^m \tau\langle\Theta^{n-1} - \Theta_h^{n-1}, X_h^n - \tilde X_h^n\rangle^{\mathcal{R}^n} + \sum_{n=1}^m \tau\langle\Theta_h^{n-1} - \tilde{\hat\Theta}_h^{n-1}, X_h^n - \tilde X_h^n\rangle^{\mathcal{R}^n} \\
&=: \mathrm{IX}_1 + \cdots + \mathrm{IX}_6.
\end{aligned}$$

We first note that term IX_1 equals a term in II. We next use (2.5), Lemmas 5.1, 5.3, Corollary 4.1, together with (6.5) and (6.6), to deduce

$$\mathrm{IX}_2 \le C\tfrac{\tau^2}{\varepsilon}\sum_{n=1}^m \tfrac{1}{\tau}\|\Theta^n - \Theta^{n-1}\|_{L^2(\Omega)}^2 + \varepsilon\sum_{n=1}^m \tau\|e_\chi^n\|_{L^2(\Omega)}^2 \le C\tfrac{\tau^2}{\varepsilon} + \varepsilon\sum_{n=1}^m \tau\|e_\chi^n\|_{L^2(\Omega)}^2,$$

$$\mathrm{IX}_3 \le C\tfrac{\tau^2}{\varepsilon}\sum_{n=1}^m \tfrac{1}{\tau}\|\Theta_h^n - \hat\Theta_h^{n-1}\|_{L^2(\Omega)}^2 + \varepsilon\sum_{n=1}^m \tau\|e_\chi^n\|_{L^2(\Omega)}^2 \le C\tfrac{\tau^2}{\varepsilon} + \varepsilon\sum_{n=1}^m \tau\|e_\chi^n\|_{L^2(\Omega)}^2,$$

$$\mathrm{IX}_4 \le C\sum_{n=1}^m \tau\|\hat\Theta_h^{n-1} - \tilde{\hat\Theta}_h^{n-1}\|_{L^2(\mathcal{R}^n)}\|e_\chi^n\|_{L^2(\Omega)} \le C\tfrac{\tau^2}{\varepsilon} + \varepsilon\sum_{n=1}^m \tau\|e_\chi^n\|^2_{L^2(\Omega)},$$

$$\mathrm{IX}_5 \le \sum_{n=1}^m \tau\|e_\theta^{n-1}\|_{L^2(\Omega)}\|X_h^n - \tilde X_h^n\|_{L^2(\Omega)} \le C\tfrac{\tau^2}{\varepsilon} + \tfrac16\sum_{n=1}^m \tau\|e_\theta^{n-1}\|^2_{L^2(\Omega)},$$

$$\mathrm{IX}_6 \le \sum_{n=1}^m \tau\|\Theta_h^{n-1} - \tilde{\hat\Theta}_h^{n-1}\|_{L^2(\mathcal{R}^n)}\|X_h^n - \tilde X_h^n\|_{L^2(\Omega)} \le C\tfrac{\tau^2}{\sqrt{\varepsilon}}.$$

In estimating term IX_6 we have added and subtracted $\hat\Theta_h^{n-1}$. In summary, we conclude

$$\mathrm{IX}_2 + \cdots + \mathrm{IX}_6 \le \tfrac16\sum_{n=1}^m \tau\|e_\theta^n\|^2_{L^2(\Omega)} + 3\varepsilon\sum_{n=1}^m \tau\|e_\chi^n\|^2_{L^2(\Omega)} + C\tfrac{\tau^2}{\varepsilon}.$$

Inserting all the estimates above into (6.11) and (6.12) we end up with

$$\|e_u^m\|^2_{H^1(\Omega)^*} + \sum_{n=1}^m \|e_u^n - e_u^{n-1}\|^2_{H^1(\Omega)^*} + \varepsilon\|e_\chi^m\|^2_{L^2(\Omega)} + \sum_{n=1}^m \tau\|e_\theta^n\|^2_{L^2(\Omega)}$$
$$\le C\tfrac{\tau^2}{\varepsilon} + C\sum_{n=1}^m \tau\|e_u^n\|^2_{H^1(\Omega)^*} + C\varepsilon\sum_{n=1}^m \tau\|e_\chi^n\|^2_{L^2(\Omega)}.$$

The discrete Gronwall inequality finally yields the asserted estimate (6.8) with N replaced by m defined in (6.10). The general estimate follows at once upon realizing that $\mathrm{V} \le C\tau^2/\varepsilon$. ∎

The effect of using (6.7) away from $\mathcal{R}^n$ is apparent in term IX_3. We would have otherwise obtained, on using (2.9),

$$\mathrm{IX}_3 = \sum_{n=1}^m \tau\langle\Theta_h^n - \tilde{\hat\Theta}_h^{n-1}, X^n - X_h^n\rangle \le C\tfrac{\tau}{\varepsilon} + \varepsilon\sum_{n=1}^m \tau\|e_\chi^n\|^2_{L^2(\Omega)},$$

which does not exhibit the correct order. We finish this section with an error estimate for the flux $\nabla\theta$. The conditions are the same as in Theorem 6.1.

THEOREM 6.2. *There holds* $\sum_{n=1}^N \tau\|\nabla e_\theta^n\|^2_{L^2(\Omega)} \le C\tau/\varepsilon$.

Proof. We take $\phi = e_\theta^n \in H^1(\Omega)$ in (6.9) and $\varphi \in \mathbb{V}^n$ the elliptic projection of ϕ

$$\langle\nabla(\phi-\varphi), \nabla\psi\rangle + \langle\phi-\varphi, \psi\rangle = 0 \quad \forall\psi \in \mathbb{V}^n.$$

We then replace $\langle\nabla\Theta_h^n, \nabla(\phi-\varphi)\rangle$ by $\langle\Theta_h^n, \varphi-\phi\rangle$ in (6.9), and add the resulting expression over n. To tackle $\sum_{n=1}^N\langle e_u^{n-1} - e_u^n, e_\theta^n\rangle$ we notice that the proof of Theorem 6.1 yields $\sum_{n=1}^N \|e_u^n - e_u^{n-1}\|^2_{H^1(\Omega)^*} \le C\tau^2/\varepsilon$. Invoking this, we see that

$$\sum_{n=1}^N \langle e_u^{n-1} - e_u^n, e_\theta^n\rangle \le C\tfrac{\tau}{\varepsilon} + \tfrac12\sum_{n=1}^N \tau\|\nabla e_\theta^n\|^2_{L^2(\Omega)},$$

which dictates the rate of convergence. On using the approximation property $\|\phi - \varphi\|_{L^2(\Omega)} \le Ch\|e_\theta^n\|_{H^1(\Omega)}$, it is not difficult to conclude the proof following that of Theorem 6.1. Details are left to the reader. ∎

Combining Theorems 6.1 and 6.2 with (2.6), for time discretization, we end up with error estimates for u, θ, $\nabla\theta$, and χ with the optimal orders of (2.6).

7. Numerical Experiments. In this section, we discuss some computational issues along with the adaptive solution of the 2D examples of [11] via AM1 and AM0.

7.1. Implementation. We implemented the adaptive methods AM1 and AM0 of §§3.3 and 3.4 using mass lumping also for Θ_h^n, that is writing $\langle \Theta_h^n - \hat{\Theta}_h^{n-1}, \varphi \rangle_n$ instead of $\langle \Theta^n - \hat{\Theta}_h^{n-1}, \varphi \rangle$ in (2.11). Let $\mathbf{M}^n$ and $\mathbf{K}^n$ be the mass and stiffness matrices

$$\mathbf{M}^n := \left(\langle \phi_i^n, \phi_j^n \rangle_n \right)_{i,j=1}^{J^n}, \quad \mathbf{K}^n := \left(\langle \nabla \phi_i^n, \nabla \phi_j^n \rangle \right)_{i,j=1}^{J^n}.$$

Identifying Θ_h^n and X_h^n with the vector of their nodal values, the *linear* algebraic system (2.13) for temperature Θ_h^n becomes

$$\mathbf{A}^n \Theta_h^n := \mathbf{M}^n \Theta_h^n + \tau \mathbf{K}^n \Theta_h^n = \mathbf{M}^n (\hat{\Theta}_h^{n-1} + \hat{X}_h^{n-1} - X_h^n),$$

where $\mathbf{A}^n$ is symmetric and positive definite. This system is actually solved by the Conjugate Gradient method (CG) preconditioned with an incomplete Cholesky factorization. The phase variable X_h^n is updated via the simple and inexpensive nodewise algebraic correction (2.12). The refined meshes are generated by the mesh generator TMG of [17], after the mesh density $\mathcal{H}^{n+1}$ is constructed and smoothed out by procedure H, as described in §3.4; no mesh has a nonempty set of bad triangles $\mathcal{B}$. The finite element code is written in C in order to avoid very large static memory space allocations and to use special data structures such as quadtrees; see [14].

We stress that parameters μ_1 and μ_2 in (3.7) are computed automatically so as to balance the number of elements produced by each expression in (3.7). The constant μ_3 is crucial in that it determines the thickness of the Refined Region, and is also computed so that the resulting element cardinality matches that of (3.7) [14].

Figures 7.1 and 7.4 contain all consecutive graded meshes for two relevant examples, discussed next. Figures 7.2, 7.3, and 7.5 depict zooms of several *Enlarged Transition Regions* (ETRs) together with the so-called *Red Zones* (RZs), namely blackened triangles which belong to the boundary of *Refined Regions* (RRs). Enforcing (3.10) is equivalent to requiring ETR to lie within RR. The layer of triangles outside RR adjacent to RZ is called *Enlarged Red Zone* (ERZ), and is used by AM0 but not by AM1. For AM0, ETR is allowed to touch ERZ without lossing a step, which violates (3.10). This yields fewer mesh changes and calls to SOL_UPDATE for AM0, because the computation can continue a few more steps.

7.2. Example: Modified Traveling Wave. This is Example 6.1 of [11] with

$$\Omega = (0,4)^2, \quad V = 1, \quad \gamma = 0.5, \quad T = 1, \quad \varepsilon = 0.003, \quad y_\varepsilon = 0.09526;$$

also

$$\|\theta\|_{L^2(Q)} \approx 12, \ \|u\|_{L^\infty(0,T;L^2(\Omega))} \approx 18, \ \sqrt{\varepsilon}\|\chi\|_{L^2(Q)} \approx 0.2, \ \|\theta\|_{L^\infty(Q)} \approx 7.$$

We have chosen the local parameters of (2.7), (3.7), (3.9), (3.11), and (3.12) as

$$\lambda_1 = 4.6, \ \lambda_2 = 5.0, \ \mu_1 \approx 0.9, \ \mu_2 \approx 3.9, \ H = 2.0, \ \mu_3^- = 0.\bar{3}, \ \mu_3^+ = 3.0, \ \mu_1^* = \mu_2^* = 3.5.$$

The solver CG stops when consecutive iterates differ less than a given error tolerance $\sigma(N) = CN^{-2}$ in the ∞-norm, with $\sigma(100) = 128\,10^{-6}$.

Figure 7.1 shows the set of 5 consecutive graded meshes for AM1 and $N = 100$. Those for AM0 are similar, and thus not reported.

Figures 7.2 and 7.3 illustrate failure of TEST 1, which is the main cause of remeshing for all examples. Figures 7.2(a) and (b) depict ETR for AM1, first inside RR at $n = 86$ and next exiting RR at $n = 87$, in violation of (3.10). The step $n = 87$ is thus discarded and repeated after SOL_UPDATE retrieves the previous solution, and a new mesh is generated by MESH. The new RZ, together with the recomputed ETR

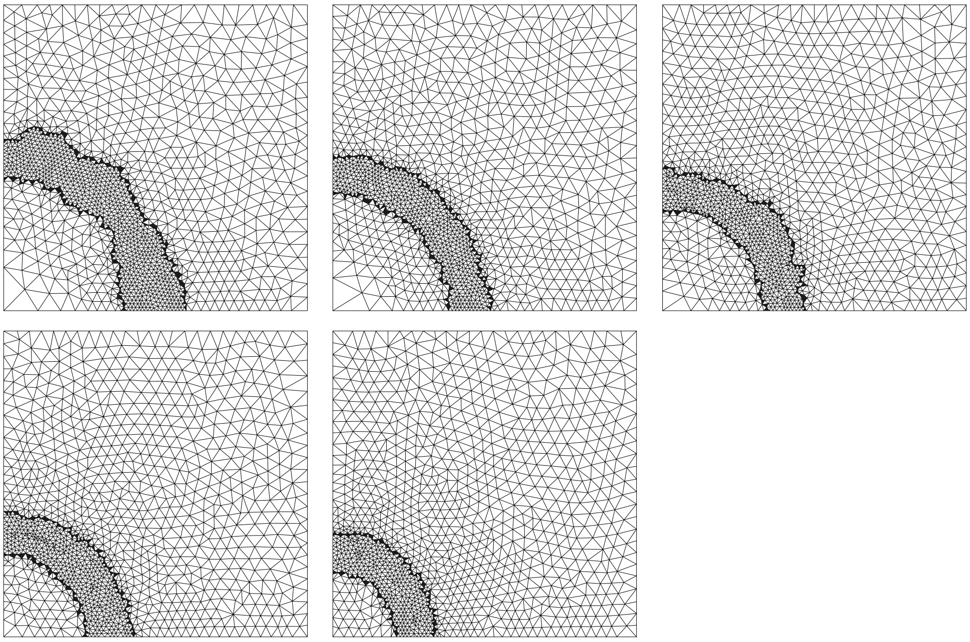

FIGURE 7.1. **Modified Traveling Wave**. Adaptive meshes for $N = 100$ using AM1.

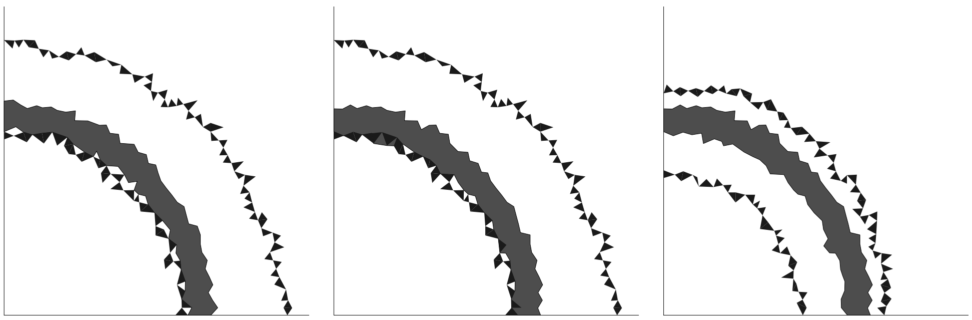

FIGURE 7.2. **Modified Traveling Wave**.
Enlarged Transition Regions and Red Zones at $n = 86, 87, 87$ for $N = 100$ using AM1.

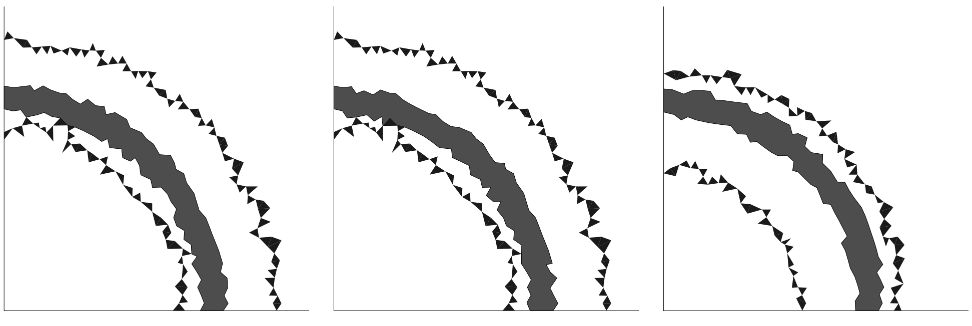

FIGURE 7.3. **Modified Traveling Wave**.
Enlarged Transition Regions and Red Zones at $n = 78, 79, 80$ for $N = 100$ using AM0.

TABLE 7.1: Comparison of methods (errors scaled by 10^2).

	N	h	**J**	**S**	**CG**	$E_\theta^{2,2}$	$E_u^{\infty,2}$	$E_{\chi,\varepsilon}^{\infty,2}$	$E_\theta^{\infty,\infty}$	N_C	**t**
	100	0.060	3570	7402	5	0.6405	3.3975	0.1583	3.2961	1	48
	200	0.030	14372	29274	6.5	0.2587	1.1691	0.0560	1.3676	1	444
FMM	300	0.020	32123	65044	8.1	0.1583	0.6821	0.0337	0.8919	1	1705
	400	0.015	57882	116826	9.2	0.1134	0.4814	0.0242	0.5630	1	4404
	500	0.012	90013	181356	10.6	0.0895	0.3700	0.0187	0.4304	1	9244
		$h_{\min}$ $h_{\max}$		**SR** **SP**							
	100	.0450 .3665	1208	1365 1170	6.2	0.5643	2.8447	0.1438	3.3458	5	31
	200	.0193 .3139	3000	3689 2486	8.3	0.2818	1.098	0.0534	1.1516	6	143
AM0	300	.0118 .2809	5316	7204 3644	10.1	0.1830	0.6717	0.0315	0.6656	6	387
	400	.0073 .2791	8186	11673 4955	12.6	0.1412	0.4751	0.0224	0.4533	6	869
	500	.0053 .2544	11618	17307 6217	13.6	0.1097	0.3680	0.0171	0.3333	6 (2)	1608
	100	.0320 .3729	1206	1356 1175	6.1	0.5334	2.7203	0.1317	2.2445	5 (1)	31
	200	.0133 .3136	3029	3735 2497	8.7	0.2870	0.9208	0.0459	0.9043	7 (1)	153
AM1	300	.0093 .2969	5345	7278 3628	10.5	0.2108	0.5536	0.0280	0.5461	8 (6)	427
	400	.0057 .2759	8434	12190 4933	12.7	0.1606	0.3852	0.0196	0.3818	8 (6)	943
	500	.0042 .2494	11458	16925 6281	13.7	0.1243	0.3086	0.0152	0.2836	8 (6)	1715

TABLE 7.2: % cpu times for AM0 / AM1.

N	$\mathbf{t}_{\text{TEMPERATURE}}$	$\mathbf{t}_{\text{TESTS}}$	$\mathbf{t}_{\text{H}}$	$\mathbf{t}_{\text{MESH}}$	$\mathbf{t}_{\text{INTERPOLATION}}$	$\mathbf{t}_{\text{MATRIX}}$
100	33 / 33	20 / 18	6.2 / 7	17 / 18	4.5 / 4.5	2.2 / 2.2
200	46 / 46	17 / 15	4.6 / 5.4	13.5 / 13.9	3.5 / 3.9	1.5 / 1.7
300	56 / 55	14 / 13.5	3.3 / 3.7	9.7 / 10.1	2.4 / 2.5	0.9 / 1
400	62 / 61	12 / 12	2.8 / 2.9	7.5 / 8.3	1.8 / 2	0.7 / 0.9
500	65 / 64	11 / 11	2.4 / 2.5	7 / 7.4	1.4 / 1.7	0.6 / 0.7

at $n = 87$ lying somewhat in the middle, is depicted in Figure 7.2(c). Such a situation would have entailed immediate remeshing for AM0, but not a call to SOL_UPDATE and loss of a step. As a matter of fact, Figure 7.3(b) illustrates failure of TEST 1 for AM0. This causes immediate remeshing, as shown in Figure 7.3(c), but step $n = 79$ is accepted.

7.3. Performance. The performance of both AM0 and AM1, along with that of the Fixed Mesh Method FMM of Part I, is documented in Tables 7.1 and 7.2 for Example 7.2. We employ the notation

N := number of (uniform) time steps,
h := quasi-uniform mesh size, $h_{\min}$ ($h_{\max}$) := smallest (largest) element size,
J := (average) degrees of freedom (dimension of linear system),
S := (average) total number of elements,
SR := (average) number of elements in the Refined Region,
SP := (average) number of elements in the Parabolic Region,
CG := (average) preconditioned Conjugate Gradient iterations,
$E_v^{p,q} := \|e_v\|_{L^p(0,T;L^q(\Omega))}$, $E_{\chi,\varepsilon}^{\infty,2} := \sqrt{\varepsilon}\|e_\chi\|_{L^\infty(0,T;L^2(\Omega))}$,
N_C := number of meshes (number of SOL_UPDATE calls),
t := total CPU time in seconds (without graphics) on a Sun Sparc 20,
$\mathbf{t}_{\text{PROCEDURE}}$:= CPU time for a procedure in the flow chart of §3.3.

Table 7.1 corroborates that performance of AM0 and AM1 is comparable and much superior than that of FMM. The rates of convergence seem to be linear in the time step τ, as predicted in §6. The number of mesh changes N_C, is somewhat insensitive to τ, especially for AM0, as expected from (3.6); this is an essential difference with respect to [13,14,15]. Note that N_C and the number of SOL_UPDATE calls, may in general be larger for AM1 than for AM0, which is consistent with the explanation of §7.1.

Table 7.2 contains a comparison of relative times required by the main procedures of §3.3 and §3.4, and shows that AM0 and AM1 are comparable. The solver TEMPERATURE dominates the calculation, specially for large N; the time for PHASE is insignificant and is included in that of TEMPERATURE. Even though remeshing is an extremal approach to mesh adaptation, the relative small time required by H and MESH demonstrates that the mesh design of §3 is adequate. Remeshing leads to incompatible consecutive meshes, thereby complicating the interpolation process between them. The implementation of INTERPOLATION via quadtrees is thus clearly efficient. So is the computation of local parameters of §3.2, as well as the testing of mesh quality, performed every time step by TESTS, and the construction of mesh densities by H.

7.4. Example: Cusp Formation. This is Example 6.2 of [11] with data

$$\Omega = (-2,4)\times(0,5), \quad T=1, \quad \varepsilon = 0.0021, \quad N = 300,$$

and solved with AM0. The local parameters of AM0 are

$$\lambda_1=\lambda_2=10.0,\ \mu_1\approx 6.35,\ \mu_2\approx 26.0,\ H=2.0, \mu_3^-=0.25,\ \mu_3^+=4.0,\ \mu_1^*=\mu_2^*=4.0.$$

The error tolerance for CG is $\sigma = 15\,10^{-6}$, and $\mathbf{CG}\approx 7$. There is a total of 11 meshes, which occur at $n = 1,9,30,54,78,104,125,148,178,216,259$. At the beginning $\mathbf{SR} = 4874$ and $\mathbf{SP} = 1372$, but the relation changes as the cusp forms and eventually the solid phase disappears at $n = 259$. In fact RR shrinks until it becomes empty and so both $\mathbf{SR}$ and $\mathbf{SP}$ diminish. For mesh 10 we have $\mathbf{SR} = 1393$ out of a total number of elements $\mathbf{S} = 2852$, whereas the last mesh has no RR and $\mathbf{S} = 1379$. Overall we have the average values $\mathbf{J} = 2311$, $\mathbf{SP} = 1649$, $\mathbf{SR} = 3012$. The topological changes of ETRs are accompanied by similar ones of RRs, as depicted in Figure 7.4. These changes are handled by AM0 without user's intervention.

FIGURE 7.4 illustrates the dependence of local refinements on both location of singularities and discrete regularity. The absence of RR in the last mesh parallels a corresponding absence of solid phase. The concentration of degrees of freedom near the right and left boundaries as time evolves is due to large discrete first and second derivatives of temperature.

Mesh changes are caused mainly by failure of TEST 1, except for the first one which is due to TEST 2. There are three SOL_UPDATE calls, of which two are due to TEST 1 and one to TEST 2. Figure 7.5 shows a sequence of ETRs and RZs at the mesh changes, and specially the mesh changes $n = 78, 79$ and $n = 125, 126$. At $n = 78$, ETR touches ERZ and forces remeshing, but the solution is accepted by AM0. At $n = 125$, ETR exits ERR and the current solution is discarded. At $n = 259$, ETR disappears.

REFERENCES

[1] T. ARBOGAST, *The existence of weak solutions to single porosity and simple dual-porosity models of two-phase incompressible flow*, Nonlinear Analysis, Theory, Methods and Applications, 19 (1992), pp. 1009–1031.

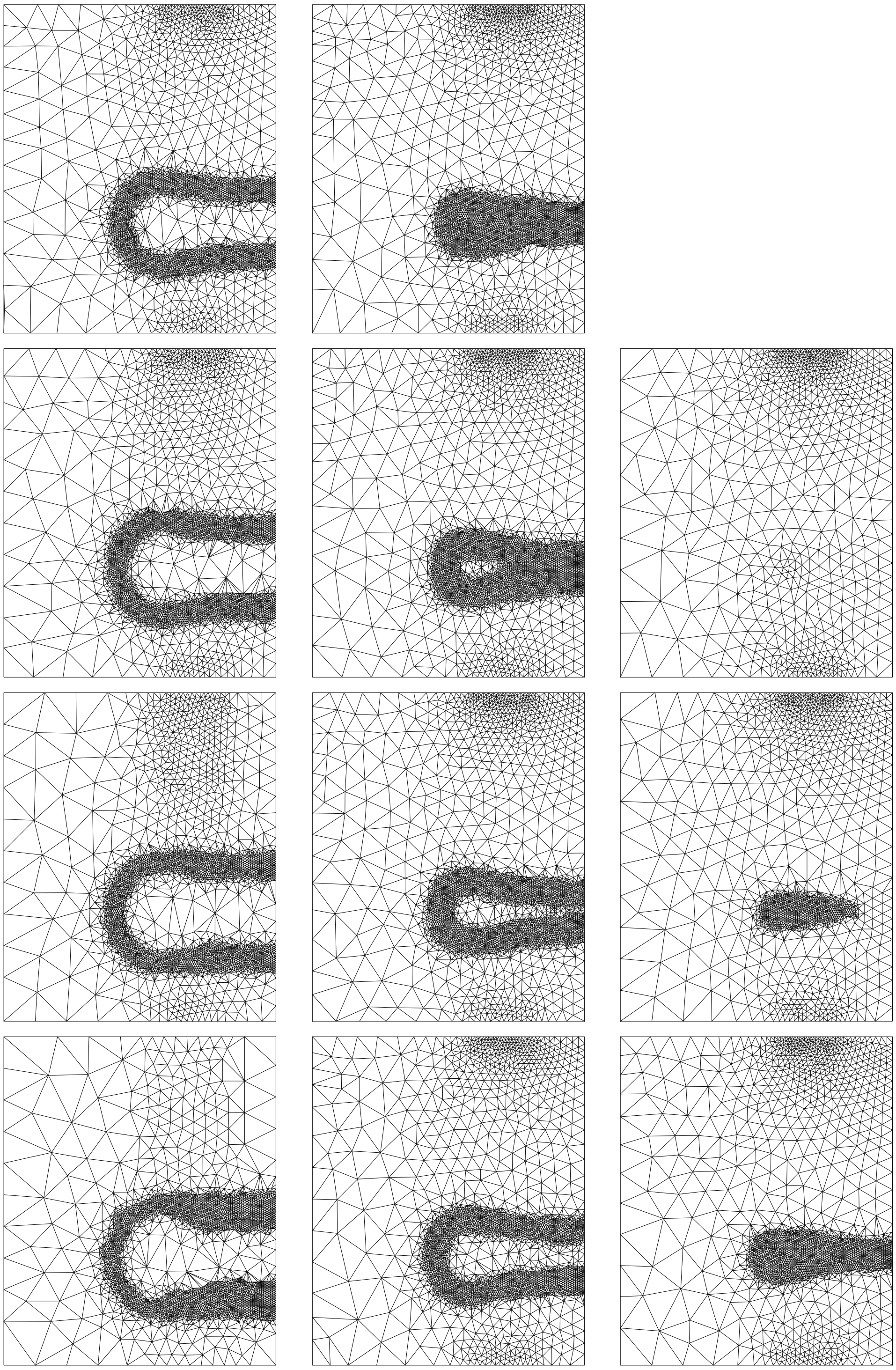

FIGURE 7.4. **Cusp Formation**. Adaptive meshes using AM0.

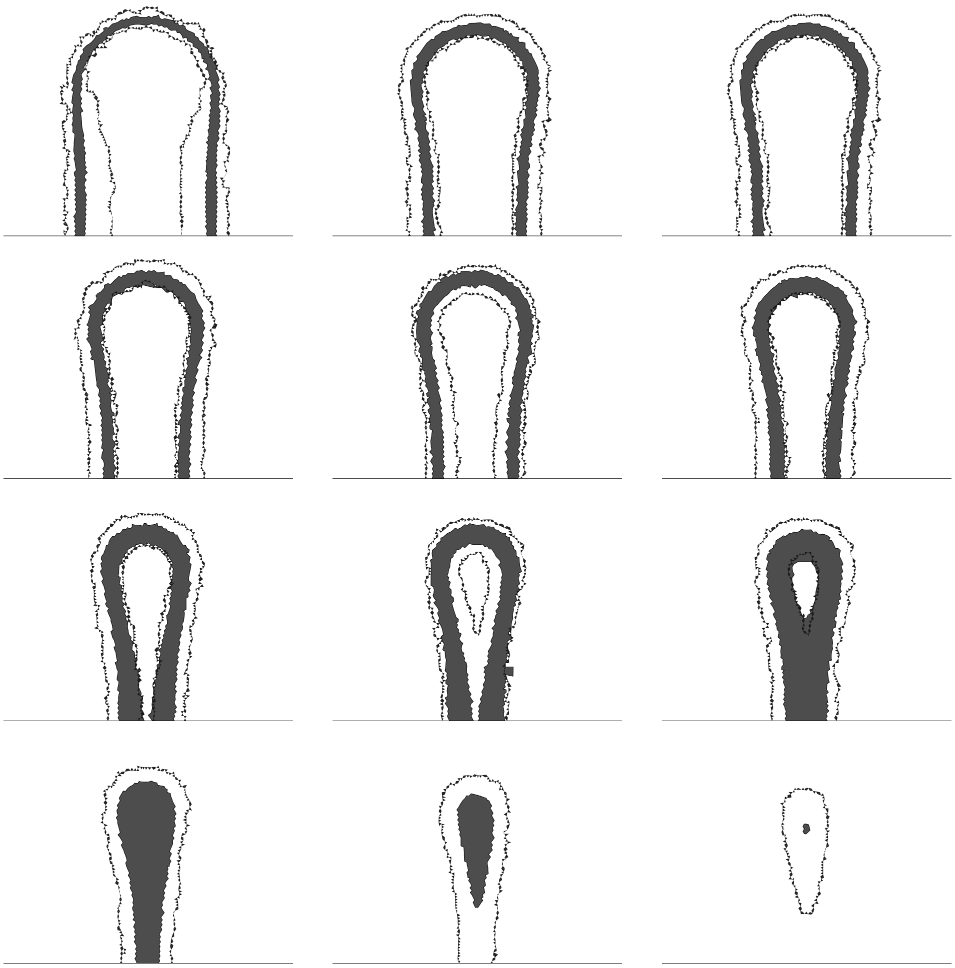

FIGURE 7.5. **Cusp Formation**. Enlarged Transition Regions and Red Zones at $n = 9, 30, 54, 78, 79, 104, 125, 126, 148, 178, 216, 258$.

[2] J.F. BLOWEY AND C.M. ELLIOTT, *Curvature dependent phase boundary motion and parabolic double obstacle problems*, IMA Series, Springer (1993).

[3] G. CAGINALP, *The dynamics of a conserved phase filed system. Stefan like, Hele-Shaw, and Cahn -Hilliard models as asymptotic limit*, IMA J. Appl. Math, 44 (1990), pp. 77–94.

[4] P. G. CIARLET, *The Finite Element Method For Elliptic Problems*, North-Holland, Amsterdam, 1978.

[5] P. C. CIARLET AND P. -A. RAVIART, *Maximum principle and uniform convergence for the finite element method*, Comput. Methods Appl. Mech. Engrg, 2 (1973), pp. 17–31.

[6] C. N. VAN DUIJN AND P. KNABNER, *Travelling waves in the transport of reactive solutes through porous media: Adsorption and binary ion exchanges. Part I*, Transport in Porous Media, 8 (1992), pp. 167–184.

[7] K. ERIKSSON AND C. JOHNSON, *Adaptive finite element methods for parabolic problems II: optimal error extimates in $L_\infty L_2$ and $L_\infty L_\infty$*, SIAM J. Numer. Anal., 32 (1995), pp. 706–740.

[8] J. W. JEROME AND M. E. ROSE, *Error estimates for the multidimentional two-phase Stefan problem*, Math. Comp., 39 (1982), pp. 377–414.

[9] X. JIANG, *Numerical approximations of a phase relaxation model*, Ph.D. dissertation, University of Maryland at College Park, 1994.

[10] X. Jiang and R. H. Nochetto, *Optimal error estimates for semidiscrete phase relaxation models*, RAIRO Model. Math. Anal. Numer., to appear.

[11] X. Jiang and R. H. Nochetto, *A $P^1 - P^1$ finite element method for a phase relaxation model. Part I: Quasi-uniform mesh*, SIAM J. Numer. Anal., submitted.

[12] X. Jiang, R.H. Nochetto and C. Verdi, *A $P^1 - P^0$ finite element method for a model of polymer crystallization*, Comput. Methods Appl. Mech. Engrg., 125 (1995), pp. 303-317.

[13] R. H. Nochetto, M. Paolini and C. Verdi, *An adaptive finite element method for two-phase stefan problems in two space demensions. Part I: Stability and error estimates*, Math. Comp., 57 (1991), pp. 73–108; Supplement, Math. Comp., 57 (1991), pp. S1–S11.

[14] R. H. Nochetto, M. Paolini and C. Verdi, *An adaptive finite element method for two-phase stefan problems in two space demensions. Part II: inplementation and numerical experiments*, SIAM J. Sci. Statist. Comput., 12 (1991), pp. 1207–1244.

[15] R. H. Nochetto, M. Paolini and C. Verdi, *A fully discrete adaptive nonlinear Chernoff formula*, SIAM J. Numer. Anal., 30 (1993), pp. 991–1014.

[16] R. H. Nochetto, M. Paolini and C. Verdi, *Continuous and semidiscrete travelling waves for a phase relaxation model*, European J. Appl. Math., 5 (1994), pp. 177–199.

[17] M. Paolini and C. Verdi, *An automatic mesh generator for planar domain*, Riv. Inform., 20 (1990), pp. 251–267.

[18] C. Verdi and A. Visintin, *Numerical analysis of the multidimensional Stefan problem with supercooling and superheating*, Boll. Unione Mat. Ital. I-B, 7 (1987), pp. 795–814.

[19] C. Verdi and A. Visintin, *Error estimates for a semi-explicit numerical scheme for Stefan-type problems*, Numer. Math., 52 (1988), pp. 165–185.

[20] A. Visintin, *Stefan problem with phase relaxation*, IMA. J. Appl. Math., 34 (1985), pp. 225–245.

[21] A. Visintin, *Supercooling and superheating effects in phase transitions*, IMA J. Appl. Math, 35 (1985), pp. 233–256.