Analisi Numerica. — *Adapting Meshes and Time-Steps for Phase Change Problems.* Nota di Ricardo H. Nochetto, Alfred Schmidt e Claudio Verdi, presentata dal Socio E. Magenes.

Abstract. — We address the numerical approximation of the two-phase Stefan problem and discuss an adaptive finite element method based on rigorous *a posteriori* error estimation and refinement/coarsening. We also investigate how to restrict coarsening for the resulting method to be stable and convergent. We review implementation issues associated with bisection and conclude with simulations of a persistent corner singularity, for which adaptivity is an essential tool.

Key words: degenerate parabolic equations; Stefan problem; finite elements; parabolic duality; *a posteriori* estimates; adaptivity.

Riassunto. — *Metodi Adattativi per Problemi di Cambiamento di Fase.* Si considera l'approssimazione numerica del problema di Stefan bifase e si discute un metodo adattativo di elementi finiti basato su stime dell'errore *a posteriori* rigorose e su tecniche di raffinamento/deraffinamento della reticolazione. Si dimostra che il metodo è stabile e convergente sotto opportune restrizioni dell'operazione di deraffinamento e si illustra l'implementazione dell'algoritmo adattativo con un metodo di bisezione. Si conclude, infine, con alcune simulazioni di un problema che presenta una singolarità di tipo angolo, per catturare la quale è essenziale l'uso di metodi di raffinamento locale.

1. Introduction

Using fixed domain methods for phase change problems is very attractive numerically in that interfaces disappear as explicit unknowns and their tracking and related difficulties are thus avoided. However, the lack of regularity across interfaces is responsible for global numerical pollution effects that degrade accuracy, most noticeably for fixed quasi-uniform meshes and constant time-steps. The purpose of this article is to review some recent results concerning mesh and time-step modification and show that adaptivity can cope with pollution and make accurate computations feasible, reliable, and robust.

We consider the simplest solid-liquid phase transition, namely the classical one or two-phase Stefan problem in enthalpy form

$$\partial_t u - \Delta\beta(u) = f \quad \text{in } Q = \Omega \times (0,T), \tag{1.1}$$

where $\beta(s) = \min(s,0) + \max(s-1,0)$, $\theta = \beta(u)$ is the temperature, and u is the enthalpy or internal energy. This constitutive relation β corresponds to an ideal material with constant thermal properties and unit latent heat.

The finite element analysis of (1.1) was started by Jerome and Rose [8]. They examined the effect of artificial viscosity, namely the replacement of β by the strictly increasing function $\beta_\varepsilon(s) = \beta(s) + \varepsilon s$, and finite element discretization with a quasi-uniform mesh of size h and constant time-step τ. Recently, Rulla and Walkington [29] showed an essentially linear rate of convergence $h + \tau$ for $\varepsilon = 0$ in two space dimensions, thereby extending the result of Rulla for time discretization [28]; see [9] for further extensions to a diffuse interface model. The resulting numerical schemes of [8], [29] provide the best scenario for error analysis, but are difficult to implement and solve because they do not include quadrature.

Numerical integration has been extensively studied by Nochetto and Verdi [26], who obtained an optimal rate of convergence $h^{1/2}$ for $\varepsilon \approx h \approx \tau$ under minimal regularity of data; related results were obtained by Elliott [6]. The techniques developed in [26] have been instrumental in analyzing linearization methods [15], [27] and related models [10], [31].

The first attempt to use properly refined meshes for Stefan problems is due to Nochetto, Paolini, and Verdi [20], [21]. The basic idea was to equidistribute *a priori* discretization errors in the maximum norm, design a mesh with a refined region of thickness of order $\tau^{1/2}$ so as to contain the discrete interface for about $\tau^{-1/2}$ time-steps, and regenerate the entire triangulation upon failure of some mesh admissibility tests. This methodology has been extended to linear schemes [22] and a phase relaxation model [11]. Its main drawback is the accurate computation of interface velocity, needed only for *a priori* mesh design, which is rather problematic for degenerate situations and diffuse interfaces. See [19], [30] for an overview.

Below we describe our current approach to mesh and time-step modification, which is based on *a posteriori* error estimation and refinement/coarsening [23], [24], [25]. The finite element solution U of (1.1), introduced in §2, satisfies

$$\partial_t U - \Delta\beta(U) = f - \mathcal{R} \quad \text{in } Q. \tag{1.2}$$

The parabolic residual $\mathcal{R}$ is a distribution with singular components and oscillatory behavior. In §3 we investigate how to restrict coarsening for the resulting method to be stable and convergent. In §4 we show how to represent the errors $e_u = u - U$ and $e_{\beta(u)} = \beta(u) - \beta(U)$ in terms of negative norms of $\mathcal{R}$, which entail averaging and thus quantify oscillations better. In §5 we derive *a posteriori* error estimates

$$\|e_u\|_{L^\infty(0,T;H^{-1}(\Omega))} + \|e_{\beta(u)}\|_{L^2(Q)} \le \mathcal{E}(u_0, f, T, \Omega; U, h, \tau)$$

with computable right-hand side $\mathcal{E}$; hereafter, h and τ stand for a variable meshsize and time-step. A further localization step in space and time, briefly discussed in §6, is needed for the estimators to be practical. We thus end up with the adaptive algorithm of §6, which equidistributes space discretization errors for a uniform error distribution in time. This strategy leads to the optimal meshes of §7, which possess fewer degrees of freedom than those of [20], [21], and requires no estimate of interface velocity nor restrictions in the number of mesh changes. In §8 we review implementation issues associated with "bisection", the method of choice for refinement/coarsening operations, but the basic ideas go back to Bänsch [2]. We conclude in §9 with simulations of a persistent corner singularity, for which adaptivity turns out to be an essential tool. The example, still unpublished, is due to Athanasopoulos, Caffarelli, and Salsa, whom we thank for bringing it to our attention and for several illuminating discussions.

2. Continuous and Discrete Problems

Let $\Omega \subset \mathbb{R}^d$ $(d \ge 1)$ be a bounded convex polyhedral domain; set $Q = \Omega \times (0,T)$ for $T > 0$. Let u_0 denote the initial enthalpy, let $\theta_0 = \beta(u_0) \in W_0^{1,\infty}(\Omega)$ be the initial temperature, and let f be sufficiently smooth.

Continuous problem. *Find* $u \in L^\infty(0,T;L^2(\Omega)) \cap W^{1,\infty}(0,T;H^{-1}(\Omega))$ *and* $\theta \in L^\infty(0,T;H_0^1(\Omega)) \cap H^1(0,T;L^2(\Omega))$ *such that* $u_{|t=0} = u_0$,

$$\theta(x,t) = \beta(u(x,t)) \quad \textit{a.e. } (x,t) \in Q,$$

and for a.e. $t \in (0,T)$ *and all* $\eta \in H_0^1(\Omega)$ *the following equation holds*

$$\langle \partial_t u, \eta\rangle + \langle \nabla\theta, \nabla\eta\rangle = \langle f, \eta\rangle. \tag{2.1}$$

Hereafter, $\langle \cdot, \cdot \rangle$ stands for either the inner product in $L^2(\Omega)$ or the duality pairing between $H^{-1}(\Omega)$, $H_0^1(\Omega)$. Existence and uniqueness for this problem are known [32].

We now introduce the fully discrete problem, which combines continuous piecewise linear finite elements in space with backward differences in time. We denote by τ_n the time-step at the n-th step and set $t^n = \sum_{i=1}^n \tau_i$; let $t^N \geq T$. Let $\mathcal{M}^n$ be a uniformly regular partition of Ω into simplices S [3] with meshsize density h_n and let $\mathcal{B}^n$ be the collection of interior interelement boundaries e of $\mathcal{M}^n$ in Ω; h_S (resp. h_e) stands for the diameter of $S \in \mathcal{M}^n$ (resp. $e \in \mathcal{B}^n$). Mesh $\mathcal{M}^n$ is obtained from $\mathcal{M}^{n-1}$ by refining/coarsening and thus $\mathcal{M}^n$ and $\mathcal{M}^{n-1}$ are compatible. Therefore the only loss of information between $\mathcal{M}^{n-1}$ and $\mathcal{M}^n$ is due to coarsening, which must be limited in order to preserve stability and convergence; see §3.

Let $\mathbb{V}^n \subset H_0^1(\Omega)$ (resp. $\mathbb{W}^n \subset L^2(\Omega)$) be the usual space of continuous (resp. discontinuous) piecewise linear finite elements over $\mathcal{M}^n$. Let $I_S^n : C^0(S) \to \mathbb{P}^1(S)$ be the local Lagrange interpolation operator; I^n indicates the global operator. The discrete inner product $\langle \cdot, \cdot \rangle^n$ is defined by the vertex quadrature rule [3],

$$\langle \varphi_1, \varphi_2 \rangle^n = \sum_{S \in \mathcal{M}^n} \int_S I_S^n(\varphi_1 \varphi_2)\, dx \quad \forall\, \varphi_1, \varphi_2 \in \mathbb{W}^n,$$

which leads to mass lumping. Set $\|\varphi\|_n = (\langle \varphi, \varphi \rangle^n)^{1/2}$ and $\|\varphi\|_{n,S} = \left(\int_S I_S^n \varphi^2 \right)^{1/2}$ for all $S \in \mathcal{M}^n$ and $\varphi \in \mathbb{W}^n$. Finally, let $P^n : L^2(\Omega) \to \mathbb{W}^n$ indicate the local L^2-projection operator over $\mathbb{W}^n$ defined by

$$P^n \eta \in \mathbb{W}^n : \quad \langle P^n \eta, \varphi \rangle = \langle \eta, \varphi \rangle \quad \forall \varphi \in \mathbb{W}^n.$$

Either I^n or P^n is used as a transfer operator T^n between consecutive meshes $\mathcal{M}^{n-1}$ and $\mathcal{M}^n$. The computation of T^n is simple in both cases [24].

Discrete problem. *Let $U^0 \in \mathbb{V}^0$ be a suitable approximation of u_0. Given U^{n-1}, $\Theta^{n-1} \in \mathbb{V}^{n-1}$, then $\mathcal{M}^{n-1}$ and τ_{n-1} are modified as described below to get $\mathcal{M}^n$ and τ_n and thereafter $U^n, \Theta^n \in \mathbb{V}^n$ computed according to $\Theta^n = I^n \beta(U^n)$ and*

$$\frac{1}{\tau_n} \langle U^n - T^n U^{n-1}, \varphi \rangle^n + \langle \nabla \Theta^n, \nabla \varphi \rangle = \langle I^n f(\cdot, t^n), \varphi \rangle^n \quad \forall\, \varphi \in \mathbb{V}^n. \tag{2.2}$$

Using mass lumping and enforcing the constitutive relation only at the nodes introduces some consistency errors but amounts to having a monotone problem. Thus (2.2) is easy to implement and solved via an optimized nonlinear SOR [21] or monotone multigrid methods [12].

We now conclude with further notation. The jump J_e^n of $\nabla \Theta^n$ across $e \in \mathcal{B}^n$ is

$$J_e^n = [\![\nabla \Theta^n]\!]_e \cdot \nu_e = (\nabla \Theta^n_{|S_1} - \nabla \Theta^n_{|S_2}) \cdot \nu_e.$$

If the unit normal vector ν_e to e always points from S_2 to S_1, then J_e^n is well defined. For any element $S \in \mathcal{M}^n$, J_S^n stands for the jumps of $\nabla \Theta^n$ across $\partial S \backslash \partial \Omega$.

Let U be the piecewise constant extension of $\{U^n\}$ defined by $U(\cdot, 0) = U^0(\cdot)$ and $U(\cdot, t) = U^n(\cdot)$ for all $t^{n-1} < t \leq t^n$ with $n \geq 1$. The interior residual R^n is

$$R^n(\cdot) = I^n f(\cdot, t^n) - \frac{U^n(\cdot) - T^n U^{n-1}(\cdot)}{\tau_n}. \tag{2.3}$$

3. Mesh and Time-Step Modification

That arbitrary mesh changes may lead to convergence to a wrong solution or divergence is known even for the heat equation [5]. In [24], we show that the discrete scheme (2.2) with changing meshes and time-steps remains stable and convergent if two restrictions are imposed for coarsening, whereas refinement operations are always allowed. Both constraints below can be imposed locally on each element $S \in \mathcal{M}^n$ and thus can be checked in practice.

Constraint 1. For all $2 \le n \le N$

$$\|T^n U^{n-1}\|_n^2 - \|U^{n-1}\|_{n-1}^2 \le \tau_{n-1} \|\nabla\Theta^{n-1}\|_{L^2(\Omega)}^2. \tag{3.1}$$

This constraint accounts for the increase in time of the L_x^2 energy due to mesh coarsening and is required for weak stability; (3.1) would be empty with exact integration and $T^n = P^n$.

Constraint 2. For all $1 \le n \le N$ and an arbitrary constant A

$$\|U^{n-1} - T^n U^{n-1}\|_{H^{-1}(\Omega)} \le \tau_{n-1} A \|h_{n-1} \nabla\Theta^{n-1}\|_{L^2(\Omega)}. \tag{3.2}$$

This constraint is required for convergence and limits coarsening in H_x^{-1}, in which I^n does not possess superconvergence properties. This choice $T^n = I^n$ imposes a restriction on the number of mesh coarsenings reminiscent of the mesh constraints in [20], [21]. On the other hand, since P^n exhibits the superconvergence property

$$\|\eta - P^n\eta\|_{H^{-1}(\Omega)} \le C\|h_n(\eta - P^n\eta)\|_{L^2(\Omega)} \quad \forall\ \eta \in L^2(\Omega),$$

then (3.2) results from the local constraint

$$\|U^{n-1} - P^n U^{n-1}\|_{n-1,S} \le \tau_{n-1} A \|\nabla\Theta^{n-1}\|_{L^2(S)} \quad \forall\ S \in \mathcal{M}^n.$$

Note that it could happen that $T^n U^{n-1} = U^{n-1}$ over a coarsened triangle $S \in \mathcal{M}^n$, but $\|T^n U^{n-1}\|_{n,S}^2 - \|U^{n-1}\|_{n-1,S}^2 > 0$, and viceversa. This shows that (3.2) and (3.1) are independent, and in fact of quite different nature.

Assuming that $\mathcal{M}^n$ is acute (weakly acute in two space dimensions), then (3.1) guarantees the weak stability of the discrete scheme (2.2) [24]

$$\tfrac{1}{2} \max_{1\le n\le N} \|U^n\|_n^2 + \sum_{n=1}^N \|U^n - T^n U^{n-1}\|_n^2 + \sum_{n=1}^N \tau_n \|\nabla\Theta^n\|_{L^2(\Omega)}^2 \le C. \tag{3.3}$$

If in addition (3.2) is imposed, then, for $H_n = \max_{x\in\Omega} h_n$ and C independent of T,

$$E = \|e_{\beta(u)}\|_{L^2(Q)} + \|e_u\|_{L^\infty(0,T;H^{-1}(\Omega))} \le CT^{1/2} \max_{1\le n\le N} \left(\tau_n + H_n + T^{1/2}\frac{H_n^2}{\tau_n}\right). \tag{3.4}$$

We infer that, if $H_n = o(\tau_n^{1/2})$, which allows for highly graded meshes, the discrete scheme with variable meshes and time-steps converges. Moreover, if $H_n = O(\tau_n)$, then the following error estimate is valid

$$E \le C \max(T, T^{1/2}) \max_{1\le n\le N} H_n.$$

This result extends the error analysis of [26], thereby incorporating mesh and time-step changes, but does not take full advantage of the underlying structure of the Stefan problem. Such a structure is hidden into (3.4) and is exploited in [24].

We finally would like to compare our analysis with the *a priori* error analysis of [7] for the linear heat equation. We first notice that our analysis is hybrid in the sense that it uses the discrete regularity (3.3) as well as continuous regularity, makes no assumption on the mesh grading, such as $|\nabla h| \ll 1$, as well as between consecutive meshes except for (3.2) and (3.1). The quasi-optimal analysis of [7] does not take quadrature into account. We do not know whether our results are sharp.

4. Error Representation Formula

We represent the errors $e_u = u - U$ and $e_{\beta(u)} = \beta(u) - \beta(U)$ in terms of the residual $\mathcal{R}$ in (1.2). We subtract (1.2) from (1.1) and integrate by parts over Q against a smooth test function ζ vanishing on $\partial\Omega \times (0,T)$. The error e_u satisfies

$$\langle e_u, \zeta\rangle_{|t=T} - \int_0^T \langle e_u, \partial_t \zeta + b\Delta\zeta\rangle = \langle e_u, \zeta\rangle_{|t=0} + \mathcal{R}(\zeta), \tag{4.1}$$

where $0 \le b(x,t) \le 1$ is the discontinuous function

$$b(x,t) = \frac{\beta(u(x,t)) - \beta(U(x,t))}{u(x,t) - U(x,t)} \text{ if } u(x,t) \ne U(x,t), \quad b(x,t) = 1 \text{ otherwise,}$$

and the parabolic residual $\mathcal{R}(\zeta)$ is the distribution

$$\mathcal{R}(\zeta) = \langle U, \zeta\rangle_{|t=0} - \langle U, \zeta\rangle_{|t=T} + \int_Q \big(f\zeta + U\partial_t\zeta + \beta(U)\Delta\zeta\big). \tag{4.2}$$

Together with the initial error, $\mathcal{R}(\zeta)$ is a measure of the amount by which U misses to be a solution of (2.1) and must be evaluated in negative norms. We can represent the error $\|e_{\beta(u)}\|_{L^2(Q)} + \|e_u\|_{L^\infty(0,T;H^{-1}(\Omega))}$ in terms of $\mathcal{R}(\zeta)$ by making judicious choices of $\zeta(\cdot,T)$ and $\partial_t\zeta + b\Delta\zeta$. Given a regularization parameter $\delta > 0$ to be chosen later, we consider two backward parabolic problems, with operator in nondivergence form and vanishing diffusion coefficient b,

$$\partial_t\psi + (b+\delta)\Delta\psi = -b^{1/2}\chi \quad \text{in } Q, \quad \psi(\cdot,T) = 0 \quad \text{in } \Omega, \tag{4.3}$$

$$\partial_t\phi + (b+\delta)\Delta\phi = 0 \quad \text{in } Q, \quad \phi(\cdot,T) = \rho \quad \text{in } \Omega, \tag{4.4}$$

$\psi, \phi = 0$ on $\partial\Omega \times (0,T)$, and $\chi \in L^2(Q)$, $\rho \in H_0^1(\Omega)$. Evaluation of $\mathcal{R}(\zeta)$ depends on regularity of ζ. The theory of nonlinear strictly parabolic problems [14] yields existence of unique solutions $\psi, \phi \in H^{2,1}(Q)$ which satisfy [18], [23]

$$2 \sup_{0\le t\le T} \|\nabla\psi(\cdot,t)\|^2_{L^2(\Omega)}, \ \tfrac{1}{1+\delta}\|\partial_t\psi\|^2_{L^2(Q)}, \ 4\delta\int_0^T |\psi|^2_{H^2(\Omega)} \le \|\chi\|^2_{L^2(Q)}, \tag{4.5}$$

$$\sup_{0\le t\le T} \|\nabla\phi(\cdot,t)\|^2_{L^2(\Omega)}, \ \tfrac{2}{1+\delta}\|\partial_t\phi\|^2_{L^2(Q)}, \ 2\delta\int_0^T |\phi|^2_{H^2(\Omega)} \le \|\nabla\rho\|^2_{L^2(\Omega)}. \tag{4.6}$$

In contrast to the heat equation, problems (4.3) and (4.4) do not exhibit any regularizing effect (in fact as $\delta \to 0$ the information on second space derivatives is lost) and are not computable in that b is discontinuous and depends on both u and U. Since the regularity of the dual problem dictates the weights (powers of meshsize and time-step) of the *a posteriori* error estimators of §5, this indicates the striking difference between degenerate and strictly parabolic problems.

We define the following negative norms of the residuals $\mathcal{R}(\psi)$ and $\mathcal{R}(\phi)$

$$\Psi_{-1} = \sup_{\chi \in L^2(Q)} \frac{|\mathcal{R}(\psi)|}{\|\chi\|_{L^2(Q)}}, \quad \Phi_{-1} = \sup_{\rho \in H_0^1(\Omega)} \frac{|\mathcal{R}(\phi)|}{\|\nabla \rho\|_{L^2(\Omega)}},$$

which, in view of (4.5) and (4.6), involve first derivatives of ψ and ϕ. On using that

$$\psi(\cdot, T) = 0, \quad -\int_0^T \langle e_u, \partial_t \psi + b\Delta\psi \rangle = \langle e_u, b^{1/2}\chi \rangle + \delta \langle e_u, \Delta\psi \rangle,$$

$$\phi(\cdot, T) = \rho, \quad -\int_0^T \langle e_u, \partial_t \phi + b\Delta\phi \rangle = \delta \langle e_u, \Delta\phi \rangle,$$

$$e_u b^{1/2} = (e_u e_{\beta(u)})^{1/2} \geq |e_{\beta(u)}|, \quad |e_u| \leq 1 + |e_{\beta(u)}|,$$

from (4.1) we easily obtain

$$\big(1 - \tfrac{a}{\sqrt{2}}\delta^{1/2}\big)\|e_{\beta(u)}\|_{L^2(Q)} + \|e_u(\cdot, T)\|_{H^{-1}(\Omega)} \leq a\|e_u^0\|_{H^{-1}(\Omega)} + R_{-1} + \tfrac{a}{\sqrt{2}}|Q|^{1/2}\delta^{1/2},$$

where $R_{-1} = \Psi_{-1} + \Phi_{-1}$ and $a = 1 + 1/\sqrt{2}$. Upon taking $\delta \to 0$, we obtain the following representation formula which is valid for any numerical method

$$E \leq a\|e_u^0\|_{H^{-1}(\Omega)} + R_{-1}. \tag{4.7}$$

Estimate (4.7) leads to Approach I below and is pessimistic in that it uses only first space derivatives of ψ and ϕ. An alternative and fruitful avenue consists of exploiting the additional, but nonuniform, H^2 space regularity of ψ and ϕ, thus keeping $\delta > 0$ and optimizing δ later without sending it to 0; this yields Approach II below and works best. To this end, we define the following negative norms of the residuals $\mathcal{R}(\psi)$ and $\mathcal{R}(\phi)$ with second space derivatives of ψ and ϕ

$$\Psi_{-2} = \sup_{\chi \in L^2(Q)} \frac{|\mathcal{R}(\psi)|}{\|\Delta\psi\|_{L^2(Q)}}, \quad \Phi_{-2} = \sup_{\rho \in H_0^1(\Omega)} \frac{|\mathcal{R}(\phi)|}{\|\Delta\phi\|_{L^2(Q)}},$$

and set $R_{-2} = \Psi_{-2} + \Phi_{-2}/\sqrt{2}$. Since (4.5) and (4.6) yield

$$\frac{|\mathcal{R}(\psi)|}{\|\chi\|_{L^2(Q)}} \leq \frac{1}{2\delta^{1/2}} \frac{|\mathcal{R}(\psi)|}{\|\Delta\psi\|_{L^2(Q)}}, \quad \frac{|\mathcal{R}(\phi)|}{\|\nabla\rho\|_{L^2(\Omega)}} \leq \frac{1}{\sqrt{2}\delta^{1/2}} \frac{|\mathcal{R}(\phi)|}{\|\Delta\phi\|_{L^2(Q)}},$$

from (4.1) we readily get

$$\big(2 - a\delta^{1/2}\big)\|e_{\beta(u)}\|_{L^2(Q)} + \|e_u(\cdot, T)\|_{H^{-1}(\Omega)} \leq \sqrt{2}a\|e_u^0\|_{H^{-1}(\Omega)} + q(\delta),$$

where $q(\delta) = q_-\delta^{-1/2} + q_+\delta^{1/2}$ with $q_- = R_{-2}$ and $q_+ = a|Q|^{1/2}$. Noting that $\delta_0 = R_{-2}/a|Q|^{1/2}$ minimizes $q(\delta)$, we obtain the representation formula

$$E \le \sqrt{2}a\|e_u^0\|_{H^{-1}(\Omega)} + 2\begin{cases} \left(a|Q|^{1/2}R_{-2}\right)^{1/2} & \text{if } a^2\delta_0 \le 1, \\ aR_{-2} & \text{if } a^2\delta_0 > 1. \end{cases} \tag{4.8}$$

In fact, if $a\delta_0^{1/2} \le 1$, the first bound follows from $q(\delta_0) = 2(q_-q_+)^{1/2}$. Otherwise, if $a\delta_0^{1/2} > 1$, then $q_+ < a^2q_-$ and $q(1/a^2) < 2aq_-$. We expect δ_0 to be small because it involves $\mathcal{R}(\psi)$ and $\mathcal{R}(\phi)$; however this cannot be guaranteed *a priori*.

If U is a finite element solution, then the parabolic residuals in (4.7) and (4.8) can be further evaluated via Galerkin orthogonality; this is accomplished in §5.

We stress that U need not be a discrete solution as the following application of (4.8) illustrates. Let U be the vanishing viscosity approximation of u, that is the solution of (1.1) with $\beta_\varepsilon(s) = \beta(s) + \varepsilon s$ instead of $\beta(s)$. The usual rate of convergence $E \le C\varepsilon^{1/2}$ is an easy consequence of (4.8) upon realizing that $R_{-2} \le a\varepsilon\|U\|_{L^2(Q)} \le C\varepsilon$, because $\mathcal{R}(\zeta) = -\varepsilon\int_Q U\Delta\zeta$. This derivation is different from the original one in [18] and is valid under minimal regularity of $u_0 \in H^{-1}(\Omega)$ and $f \in L^1(0,T;H^{-1}(\Omega))$, which precludes compactness [23].

5. A Posteriori Error Estimates

We now state two rigorous *a posteriori* error estimates. Their derivation parallels that in §4, but exploits Galerkin orthogonality to express negative norms of the residuals $\mathcal{R}(\psi)$ and $\mathcal{R}(\phi)$ in terms of computable quantities [23].

Integrating (4.2) by parts and using the definition (2.3), $\mathcal{R}(\zeta)$ becomes

$$\begin{aligned} \mathcal{R}(\zeta) = &\sum_{n=1}^N \int_{t^{n-1}}^{t^n} \left(\langle R^n, \zeta\rangle - \langle \nabla\beta(U^n), \nabla\zeta\rangle\right) \\ &+ \sum_{n=1}^N \langle U^{n-1} - T^nU^{n-1}, \zeta(\cdot, t^{n-1})\rangle \\ &+ \sum_{n=1}^N \int_{t^{n-1}}^{t^n} \tfrac{1}{\tau_n}\langle U^n - T^nU^{n-1}, \zeta - \zeta(\cdot, t^{n-1})\rangle \\ &+ \sum_{n=1}^N \int_{t^{n-1}}^{t^n} \langle f - I^nf(\cdot, t^n), \zeta\rangle. \end{aligned} \tag{5.1}$$

We next use Galerkin orthogonality, that is we rewrite the discrete problem (2.2)

$$\begin{aligned} &\langle R^n, \varphi\rangle - \langle \nabla\beta(U^n), \nabla\varphi\rangle \\ &\quad = (\langle R^n, \varphi\rangle - \langle R^n, \varphi\rangle^n) + \langle \nabla(I^n\beta(U^n) - \beta(U^n)), \nabla\varphi\rangle \quad \forall\ \varphi \in \mathbb{V}^n, \end{aligned} \tag{5.2}$$

and subtract this expression from the right hand side of (5.1) to arrive at

$$
\begin{aligned}
\mathcal{R}(\zeta) = &\sum_{n=1}^{N}\int_{t^{n-1}}^{t^n} \big(\langle R^n, \zeta-\varphi\rangle - \langle \nabla I^n\beta(U^n), \nabla(\zeta-\varphi)\rangle\big) \\
&+\sum_{n=1}^{N}\langle U^{n-1}-T^nU^{n-1}, \zeta(\cdot,t^{n-1})\rangle \\
&+\sum_{n=1}^{N}\int_{t^{n-1}}^{t^n} \tfrac{1}{\tau_n}\langle U^n - T^nU^{n-1}, \zeta-\zeta(\cdot,t^{n-1})\rangle \\
&+\text{consistency terms} = \mathrm{I}+\cdots+\mathrm{IV}+\text{c.t.}
\end{aligned}
$$

Note that the right-hand side of (5.2) would be 0 without variational crimes (mass lumping and lumped constitutive relation).

We now estimate each term I to IV separately. We argue with $\zeta=\psi$, solution of (4.3) with $\chi\in L^2(Q)$. Since we need to approximate ζ under minimal regularity, we resort to the Clément interpolation operator $\Pi^n : L^2(\Omega)\to\mathbb{V}^n$, which satisfies [4], for all $\eta\in H^k(\Omega)$ and $k=1,2$,

$$
\begin{aligned}
&\|\eta-\Pi^n\eta\|_{L^2(S)} + h_S\|\nabla(\eta-\Pi^n\eta)\|_{L^2(S)} \le \tilde{C}h_S^k|\eta|_{H^k(\tilde{S})}, \\
&\qquad\qquad \|\eta-\Pi^n\eta\|_{L^2(e)} \le \tilde{C}h_e^{k-1/2}|\eta|_{H^k(\tilde{S})},
\end{aligned}
\tag{5.3}
$$

where $\tilde{S}$ is the union of all elements surrounding $S\in\mathcal{M}^n$ or $e\in\mathcal{B}^n$. Constants $\tilde{C}$ depend solely on the minimum angle of the mesh $\mathcal{M}^n$. An important by-product of uniform mesh regularity is that the number of adjacent simplices to a given element is bounded by a constant M independent of n, meshsizes, and time-steps. We decompose the integral $\langle\nabla\Theta^n, \nabla(\zeta-\varphi)\rangle$ over all elements $S\in\mathcal{M}^n$ and next integrate by parts to obtain the equivalent expression

$$
-\langle \nabla I^n\beta(U^n), \nabla(\zeta-\varphi)\rangle = \sum_{e\in\mathcal{B}^n}\langle\!\langle J_e^n, \zeta-\varphi\rangle\!\rangle_e \quad \forall\,\varphi\in\mathbb{V}^n,
$$

where $\langle\!\langle\cdot,\cdot\rangle\!\rangle_e$ denotes the L^2-scalar product on $e\in\mathcal{B}^n$. Selecting $\varphi(\cdot,t)=\Pi^n\psi(\cdot,t)$ for $t^{n-1}<t\le t^n$ and using (5.3) and H^1_x regularity of ψ, terms I and II can be bounded as follows (Approach I)

$$
\begin{aligned}
|\mathrm{I}| &\le C_1\sum_{n=1}^{N}\int_{t^{n-1}}^{t^n}\Big(\sum_{S\in\mathcal{M}^n} h_S^2\|R^n\|^2_{L^2(S)}\Big)^{1/2}\|\nabla\psi(\cdot,t)\|_{L^2(\Omega)}, \\
|\mathrm{II}| &\le C_2\sum_{n=1}^{N}\int_{t^{n-1}}^{t^n}\Big(\sum_{e\in\mathcal{B}^n} h_e\|J_e^n\|^2_{L^2(e)}\Big)^{1/2}\|\nabla\psi(\cdot,t)\|_{L^2(\Omega)};
\end{aligned}
$$

constants C_1 and C_2 depend upon $\tilde{C}$ and M. Alternatively, on using H^2_x regularity of ψ, we can also write (Approach II)

$$
\begin{aligned}
|\mathrm{I}| &\le C_1\sum_{n=1}^{N}\int_{t^{n-1}}^{t^n}\Big(\sum_{S\in\mathcal{M}^n} h_S^4\|R^n\|^2_{L^2(S)}\Big)^{1/2}|\psi(\cdot,t)|_{H^2(\Omega)}, \\
|\mathrm{II}| &\le C_2\sum_{n=1}^{N}\int_{t^{n-1}}^{t^n}\Big(\sum_{e\in\mathcal{B}^n} h_e^3\|J_e^n\|^2_{L^2(e)}\Big)^{1/2}|\psi(\cdot,t)|_{H^2(\Omega)}.
\end{aligned}
$$

Term III is simply bounded via $H^{-1}(\Omega) - H_0^1(\Omega)$ duality whereas, on using that $\psi(\cdot,t) - \psi(\cdot,t^{n-1}) = \int_{t^{n-1}}^t \partial_s\psi(\cdot,s)$, we readily obtain

$$|\mathrm{IV}| \le \Big(\sum_{n=1}^N \tau_n \|U^n - T^n U^{n-1}\|^2_{L^2(\Omega)}\Big)^{1/2} \Big(\int_0^T \|\partial_t\psi\|^2_{L^2(\Omega)}\Big)^{1/2}.$$

Similar estimates are valid for $\zeta = \phi$, solution of (4.4) with $\rho \in H_0^1(\Omega)$. Neglecting the consistency terms just for simplicity and making use of the *a priori* extimates (4.5) and (4.6), we get the following bounds for the residuals $\mathcal{R}(\psi)$ and $\mathcal{R}(\phi)$

$$\frac{|\mathcal{R}(\psi)|}{\|\chi\|_{L^2(Q)}} \le \tfrac{1}{\sqrt{2}}\mathcal{E}_3 + (1+\delta)^{1/2}\mathcal{E}_4 + \begin{cases} \frac{1}{\sqrt{2}}(\mathcal{E}_1^{\mathrm{I}} + \mathcal{E}_2^{\mathrm{I}}), \\ \frac{1}{2}\delta^{-1/2}(\mathcal{E}_1^{\mathrm{II}} + \mathcal{E}_2^{\mathrm{II}}), \end{cases}$$

$$\frac{|\mathcal{R}(\phi)|}{\|\nabla\rho\|_{L^2(\Omega)}} \le \mathcal{E}_3 + \tfrac{1}{\sqrt{2}}(1+\delta)^{1/2}\mathcal{E}_4 + \begin{cases} \mathcal{E}_1^{\mathrm{I}} + \mathcal{E}_2^{\mathrm{I}}, \\ \frac{1}{\sqrt{2}}\delta^{-1/2}(\mathcal{E}_1^{\mathrm{II}} + \mathcal{E}_2^{\mathrm{II}}), \end{cases}$$

where the error indicators for Approach I are given by

$$\mathcal{E}_1^{\mathrm{I}} = C_1 \sum_{n=1}^N \tau_n \Big(\sum_{S\in\mathcal{M}^n} h_S^2 \|R^n\|^2_{L^2(S)}\Big)^{1/2} \quad \text{interior residual,}$$

$$\mathcal{E}_2^{\mathrm{I}} = C_2 \sum_{n=1}^N \tau_n \Big(\sum_{e\in\mathcal{B}^n} h_e \|J_e^n\|^2_{L^2(e)}\Big)^{1/2} \quad \text{jump residual,}$$

$$\mathcal{E}_3 = \sum_{n=1}^N \|U^{n-1} - T^n U^{n-1}\|_{H^{-1}(\Omega)} \quad \text{coarsening,}$$

$$\mathcal{E}_4 = \Big(\sum_{n=1}^N \tau_n \|U^n - T^n U^{n-1}\|^2_{L^2(\Omega)}\Big)^{1/2} \quad \text{time residual,}$$

and for Approach II by

$$\mathcal{E}_1^{\mathrm{II}} = C_1 \Big(\sum_{n=1}^N \tau_n \sum_{S\in\mathcal{M}^n} h_S^4 \|R^n\|^2_{L^2(S)}\Big)^{1/2} \quad \text{interior residual,}$$

$$\mathcal{E}_2^{\mathrm{II}} = C_2 \Big(\sum_{n=1}^N \tau_n \sum_{e\in\mathcal{B}^n} h_e^3 \|J_e^n\|^2_{L^2(e)}\Big)^{1/2} \quad \text{jump residual.}$$

Setting $\mathcal{E}_0 = \|u_0 - U^0\|_{H^{-1}(\Omega)}$ (initial error), we can argue as in deriving (4.7) and (4.8) and conclude with the following two *a posteriori* error estimates

$$\|e_{\beta(u)}\|_{L^2(Q)} + \|e_u\|_{L^\infty(0,T;H^{-1}(\Omega))} \le \mathcal{E}^{\mathrm{k}}(u_0, f, T, \Omega; U, h, \tau) \tag{5.4}$$

for Approaches $\mathrm{k} = \mathrm{I}$ and $\mathrm{k} = \mathrm{II}$, where

$$\mathcal{E}^{\mathrm{I}} = a(\mathcal{E}_0 + \mathcal{E}_1^{\mathrm{I}} + \mathcal{E}_2^{\mathrm{I}} + \mathcal{E}_3 + \mathcal{E}_4),$$

$$\mathcal{E}^{\mathrm{II}} = \sqrt{2}a(\mathcal{E}_0 + \mathcal{E}_3) + 3a\mathcal{E}_4 + 2a\begin{cases} |Q|^{1/4}(\mathcal{E}_1^{\mathrm{II}} + \mathcal{E}_2^{\mathrm{II}})^{1/2} & \text{if } a^2(\mathcal{E}_1^{\mathrm{II}} + \mathcal{E}_2^{\mathrm{II}}) \le |Q|^{1/2}, \\ a(\mathcal{E}_1^{\mathrm{II}} + \mathcal{E}_2^{\mathrm{II}}) & \text{if } a^2(\mathcal{E}_1^{\mathrm{II}} + \mathcal{E}_2^{\mathrm{II}}) > |Q|^{1/2}. \end{cases}$$

The indicators $\mathcal{E}^{\mathrm{k}}$ can be evaluated explicitly in terms of the computed solution U, initial datum u_0, and source term f. They are essential and are also present for the heat equation, but with different weights and cumulative effect in time [7]. The error accumulation is measured here in L^1 or L^2, whereas it is in L^∞ for the heat equation. The powers of meshsize in $\mathcal{E}_1^{\mathrm{I}}$ and $\mathcal{E}_2^{\mathrm{I}}$ are smaller than those for the heat equation, namely h_e^3 and h_S^4, respectively, thereby reflecting the degenerate nature of (1.1), or equivalently the lack of H^2 space regularity of ψ and ϕ. Approach II yields the same weights for $\mathcal{E}_1^{\mathrm{II}}$ and $\mathcal{E}_2^{\mathrm{II}}$ as the heat equation, but yet with a worse error accumulation in time and a smaller outermost power. The indicators associated to the consistency terms are not essential and could in principle be removed at the expense of complicating the implementation of (2.2).

It is not obvious that $\mathcal{E}^{\mathrm{k}}(u_0, f, T, \Omega; U, h, \tau) \to 0$ as $h, \tau \to 0$, because $\mathcal{E}^{\mathrm{k}}$ depends on discrete quantities that change with h and τ. On the other hand, the stability and error analysis of §3 demonstrate that this goal is achievable. Convergence of the adaptive algorithm remains a challenging open problem.

6. Adaptive Algorithm

Since the error estimators $\mathcal{E}^{\mathrm{k}}$ of §5 entail an L^1 or L^2 accumulation in time, they are impractical in that the entire evolution history would be needed to control the error. In [23], we overcome this hurdle upon equidistributing the errors in time in the L^∞ norm and optimizing the error distribution in space at each time step. If $E_h^n(S)$ denotes the contribution to the total error due to an element $S \in \mathcal{M}^n$, then the optimality condition reads

$$E_h^n(S) = \frac{\text{scaled tolerance}}{\text{cardinality of } \mathcal{M}^n}. \tag{6.1}$$

We now explain our strategy for Approach I, just for simplicity, and introduce the element error indicators and all tests necessary for mesh and time-step admissibility.

For each $S \in \mathcal{M}^n$, the local spatial error indicators are defined as follows: $E_0(S)$ denotes the initial error and

$$E_h^n(S) = 3a^2T^2\big(C_1^2 h_S^2 \|R^n\|_{L^2(S)}^2 + \tfrac{1}{2} C_2^2 h_S \|J_S^n\|_{L^2(\partial S)}^2 + \tau_n^{-2} \|U^{n-1} - T^n U^{n-1}\|_{L^2(S)}^2\big)$$

involves the internal residual, jump residual, and coarsening term. We also set

$$E_h^n = \Big(\sum_{S\in\mathcal{M}^n} E_h^n(S)\Big)^{1/2}, \quad E_h = \max_{1\le n\le N} E_h^n, \quad E_0 = \Big(\sum_{S\in\mathcal{M}^n} E_0(S)\Big)^{1/2}.$$

We finally define the time residual to be

$$E_\tau^n = aT^{1/2}\|U^n - T^n U^{n-1}\|_{L^2(\Omega)}, \quad E_\tau = \max_{1\le n\le N} E_\tau^n.$$

Then it is easy to check that

$$\mathcal{E}^{\mathrm{I}}(u_0, f, T, \Omega; U, h, \tau) \le E_0 + E_\tau + E_h.$$

Given an error tolerance ε, the objective is to adaptively select time-steps τ and mesh densities h in such a way that E_τ^n have comparable size for all $1 \le n \le N$ and $E_h^n(S)$ satisfies (6.1) for all $S \in \mathcal{M}^n$ and $1 \le n \le N$ (equidistribution of spatial degrees of freedom for a uniform error distribution in time) and

$$E_0 + E_\tau + E_h \le \varepsilon. \tag{6.2}$$

Given refinement parameters $\Gamma > 0$ and coarsening parameters $\gamma > 0$ satisfying

$$\Gamma_0 + \Gamma_\tau + \Gamma_h \le 1, \quad \gamma_\tau < \Gamma_\tau, \quad \gamma_h < \Gamma_h,$$

then (6.2) is achieved provided time-steps and mesh densities are modified until

$$E_0 \le \Gamma_0\, \varepsilon, \quad \gamma_\tau\, \varepsilon \le E_\tau \le \Gamma_\tau\, \varepsilon, \quad \gamma_h\, \varepsilon \le E_h \le \Gamma_h\, \varepsilon.$$

7. Optimal Meshes

We study the relative meshsizes and degrees of freedom for optimal meshes, that is meshes satisfying (6.1). We use the notation h_P and h_H to designate the meshsize away from the discrete interface (parabolic region) and near the interface (hyperbolic region). We also denote with M_P and M_H the corresponding number of elements in these two regions. For simplicity, we only argue with the jump residuals J_S^n and suppress the superscript n. We assume that $M_P = O(h_P^{-d})$, $M_H = O(h_H^{1-d})$ and that $J_S \approx 1$ in the hyperbolic region, $J_S \approx h_P$ in the parabolic counterpart, as expected for a smooth free boundary.

Approach I. Condition (6.1) yields $h_H \|J_S\|^2_{L^2(\partial S)} \approx h_P \|J_S\|^2_{L^2(\partial S)}$, whence

$$h_H \approx h_P^{(d+2)/d}, \quad \frac{M_H}{M_P} = O\big(h_P^{(2-d)/d}\big).$$

We then see that $h_H \approx h_P^2$ for $d = 2$, which is consistent with the relation imposed in [20], [21]. Moreover, M_H and M_P are comparable for $d = 2$, whereas the former dominates for $d = 3$.

Approach II. This time (6.1) implies $h_H^3 \|J_S\|^2_{L^2(\partial S)} \approx h_P^3 \|J_S\|^2_{L^2(\partial S)}$, whence

$$h_H \approx h_P^{(d+4)/(d+2)}, \quad \frac{M_H}{M_P} = O\big(h_P^{(4-d)/(d+2)}\big).$$

We conclude that $M_H \ll M_P$ for $1 \le d \le 3$, and so Approach II requires fewer degrees of freedom than Approach I for a given accuracy. This assertion is confirmed by numerical evidence [23], [25].

8. Implementation

Equidistribution Strategy (ES). Let M^n denote the cardinality of $\mathcal{M}^n$ at any step of ES. Given a coarse initial mesh $\mathcal{M}^0$, ES bisects all $S \in \mathcal{M}^0$ such that

$$E_0(S) > \frac{\Gamma_0^2\, \varepsilon^2}{M^0}.$$

To select the time-step τ_n, starting with $\tau_n = \tau_{n-1}$, the algorithm checks whether

$$E_\tau^n > \Gamma_\tau\, \varepsilon, \quad E_\tau^n < \gamma_\tau\, \varepsilon.$$

In the first case τ_n is reduced, whereas in the second one (corresponding to τ_n being too small) τ_n is accepted but the initial guess for the next time-step size is enlarged. Next, starting from $\mathcal{M}^n = \mathcal{M}^{n-1}$, for any $S \in \mathcal{M}^n$, ES checks whether

$$E_h^n(S) > \frac{\Gamma_h^2\, \varepsilon^2}{M^n}, \quad E_h^n(S) < \frac{\gamma_h^2\, \varepsilon^2}{M^n}.$$

Then refinement and coarsening operations are performed accordingly, with the precaution of choosing $\gamma_h \ll \Gamma_h$ properly to prevent ES from alternating such operations over the same elements. Elements are either refined or coarsened via "bisection" as explained later in this section.

FLOW CHART. A flow diagram of the adaptive algorithm reads as follows.

---

start with $\mathcal{M}^0$, τ^0, U^0
for $n \geq 1$
(1) solve for U^n, Θ^n
  compute error estimators E_τ^n and $E_h^n(S)$ for $S \in \mathcal{M}^n$
  if E_τ^n is large, reduce τ^n and goto (1)
(2) for every $S \in \mathcal{M}^n$
    if $E_h^n(S)$ is large, refine S
    if $E_h^n(S)$ is small, coarsen S if possible
  if the mesh was changed
    solve again for U^n, Θ^n
    compute error estimators
  if E_τ^n is large, reduce τ^n and goto (1)
  if E_h^n is large, goto (2)
  accept U^n, Θ^n and $\mathcal{M}^n$, τ^n
  if E_τ^n is small, enlarge τ^{n+1}

After each iteration of ES, both U^n and Θ^n are recalculated on the new mesh using the new time-step size. To reduce the overhead computational cost, a compromise is reached between the optimization of the degrees of freedom and iterations of ES. ES stops iterating as soon as $E_h^n \leq \Gamma_h\, \varepsilon$ is fulfilled, thereby allowing some elements to violate the local error tolerance. Consequently, discretization errors might not be equidistributed correctly. As implemented in [23], [25] and shown in the flow diagram, at least one mesh modification per time-step is performed in order to permit elements near the moving interface to be refined, even if the global error bound is already fulfilled by the old mesh.

MESH REFINEMENT AND COARSENING. We describe shortly the local mesh refinement and coarsening used to adapt the meshes according to information extracted from local error estimators. Several refinement algorithms are available for local refinement of triangular meshes in two space dimensions (see Mitchell [17]). Popular methods are "regular refinement" and "bisection".

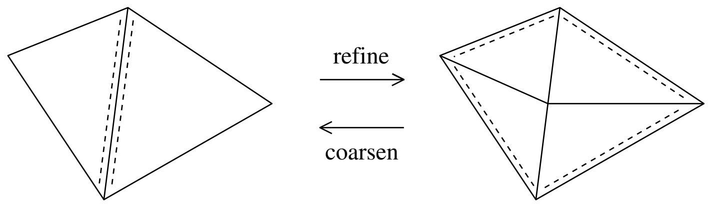

FIGURE 8.1. Atomic refinement and coarsening operations. Refinement edges of triangles are marked with dashed lines.

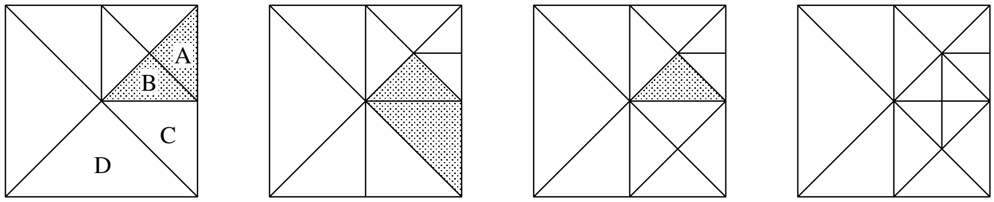

FIGURE 8.2. Recursive refinement. Triangles A and B are initially marked for refinement.

With "regular refinement", triangles are divided into four similar triangles. Unfortunately, nonconforming nodes arise during local regular refinement, which force additional bisection of some triangles. In order to keep triangulations regular, these bisections have to be removed before further refinement. This results in (partly) incompatible meshes and introduces interpolation errors between meshes even during refinement operations. Extensions to three dimensions are highly nontrivial.

In "bisection" methods, instead, triangles are divided into two triangles by inserting the midpoint of one of their edges, called the "refinement edge". We use the "newest vertex" bisection, where the refinement edge for a new triangle is chosen as the edge opposite the newly created vertex. This ensures that shape regularity of the triangulations is preserved. Extensions of such algorithms to three dimensions (meshes of tetrahedra) are described in [2], [13], [16].

The refinement algorithm can be implemented using a recursive bisectioning of triangles. Elements are refined by bisecting their refinement edge. To keep the mesh conforming, bisection of an edge is only allowed when this edge is also the refinement edge for the adjacent triangle which shares this edge. Bisection of an edge and thus of both elements around the edge is the *atomic refinement operation*, see Figure 8.1, and no other refinement operations are allowed. For triangles at the boundary, the corresponding atomic refinement operation involves only one triangle. After each such atomic refinement, the resulting triangulation is conforming.

If a triangle is selected for refinement, but the adjacent triangle across the refinement edge does not share the common edge as its refinement edge, first this neighbour is refined recursively. This generates a compatible refinement edge of the original triangle and its (new) neighbour. This recursion is guaranteed to stop for every triangle of a refined triangulation if the recursive refinement does not create cycles on the (usually very coarse) initial triangulation; this can be easily checked for a triangulation with a given choice of refinement edges. In Figure 8.2 we show a situation where recursion is needed. For all triangles, the longest edge is the refinement edge. Let us assume that triangles A and B are marked for refinement.

Triangle A can be refined at once, as its refinement edge is a boundary edge. For refinement of triangle B, we have to recursively refine triangles C and D. Again, triangle D can be directly refined, so recursion stops there. This is shown in the second part of the figure. Back in triangle C, this can now be refined together with its neighbour. After this, also triangle B can be refined together with its neighbour.

The inverse operation to atomic refinement is the *atomic coarsening operation*. It reduces a coarsening patch of four triangles to two larger triangles. The four smaller trangles must have been created by refinement of the two larger ones. At the boundary, a coarsening patch involves only two small triangles. Again, the triangulation is still conforming after each atomic coarsening operation.

For the usage in the adaptive method described above, triangles are marked for refinement if the local estimator $E_h^n(S)$ is larger than the refinement tolerance $\Gamma_h^2\,\varepsilon^2/M^n$, and marked for coarsening if the local estimator is smaller than $\gamma_h^2\,\varepsilon^2/M^n$. In order for the adaptive method to be able to give guaranteed error bounds, at least all triangles which are marked for refinement have to be refined and at most those triangles which are marked for coarsening are allowed to be coarsened. This does not impose any restriction for refinement, but coarsening operations are allowed only if all involved triangles are marked for coarsening.

9. Simulations

We conclude with an intriguing example with persistent corner singularity for a one-phase Stefan problem in two space dimensions. The key question, posed by Athanasopoulos, Caffarelli, and Salsa with this example, is whether or not $\pi/2$ is the critical angle beyond which the interface immediately regularizes. Our simulations seem to indicate that the critical angle is actually larger than $\pi/2$. To address this issue, the use of adaptive local refinements has been essential.

Let $\Omega = (-0.1, 0.1)^2$, $T = 0.1$, and $g(t) = 2.1 - 10\,t$. Consider polar coordinates (ρ, ω) and the function

$$u^+(\rho,\omega,t) = \begin{cases} \rho^{g(t)}\cos\big(\omega g(t)\big) & \text{if } g(t)|\omega| < \frac{\pi}{2}, \\ 0 & \text{otherwise.} \end{cases} \tag{9.1}$$

An elementary calculation shows that u^+ is a supersolution provided

$$-\frac{\pi\, g'(t)}{2\, g^3(t)} \geq \rho^{g(t)-2}, \tag{9.2}$$

which is not satisfied for all $t \leq T$. We point out that (9.2) can only be enforced for $g(t) > 2$ in a shrinking domain Ω as $g(t) \downarrow 2$, and provides some support to the above conjecture that $\omega_0 = \pi/4$, and thus the opening $\pi/2 = 2\omega_0$, could be critical.

We only use u^+ to set up the Dirichlet condition on the parabolic boundary of Q in our simulations. We also use an angle shift of 0.2, and thus consider $\omega - 0.2$ in (9.1), to avoid grid orientation effects.

The *a posteriori* error estimates (5.4) give no direct control over the accurate resolution of the interface position and motion, which is here the main information of interest. Moreover, the solution u behaves like ρ^2 near the origin, which results in very small local estimators. In order to reduce the meshsize, and thus hope for

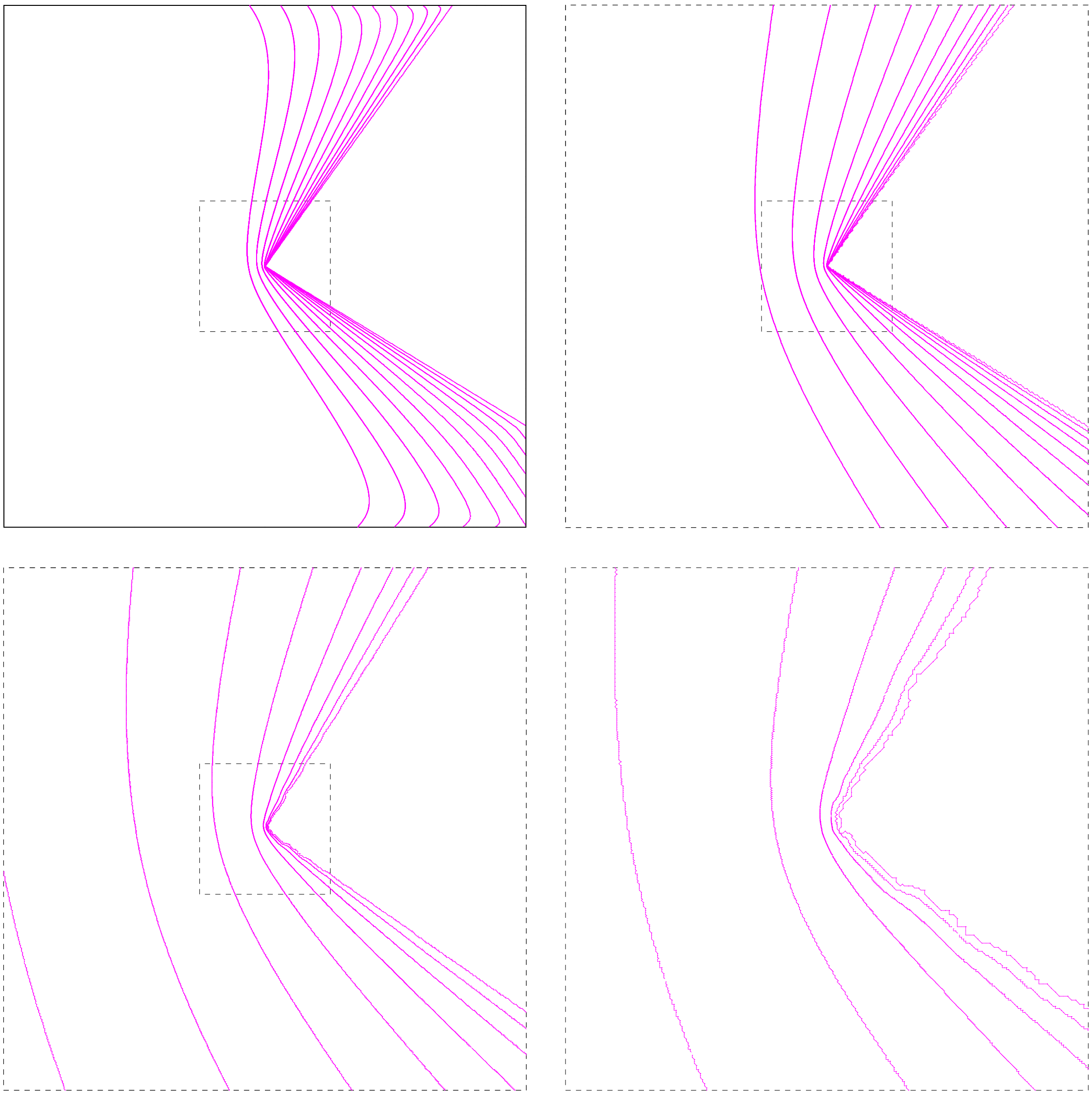

FIGURE 9.1. Zoom of interfaces with scaling factors 1, 4, 16, 64 for smallest tolerance $\varepsilon = 1$, at times $t = 0.01\,k$, $0 \le k \le 10$ (top) resp. $3 \le k \le 10$ (bottom).

a good approximation of the interface near the corner, we introduce an additional weighting factor $1/\max(\rho^2, h_{\min}, U)$, with the parameter $h_{\min}$ given. This weight is essentially proportional to ρ^{-2}, and thus compensates for the behavior of the solution near the origin.

We run our adaptive Approach I for various tolerances $\varepsilon = 1,\ 2,\ 4,\ 8$, fixed time-steps $\tau = 0.5,\ 1,\ 2,\ 5 \times 10^{-3}$, with parameters $h_{\min} = 1,\ 2,\ 4,\ 8 \times 10^{-5}$, and stopping criteria $1,\ 2,\ 4,\ 8 \times 10^{-7}$ for the nonlinear SOR with fixed relaxation parameter 1.5.

Figure 9.1 displays the interfaces and zooms with scaling factors 1, 4, 16, 64 for the most accurate run with $\varepsilon = 1$. The solution shows a corner at the origin which persists for some time, while the angle widens. The corner seems to regularize for $0.04 \le t \le 0.05$, when the angle is already larger than $\pi/2$. Figures 9.2 and 9.3 depict a mesh study for zoom scaling factors 1 and 16, respectively. The relative location of interfaces, as well as time of regularization, are consistent with mesh refinement. Finally, Figure 9.4 contains a representative locally refined mesh and

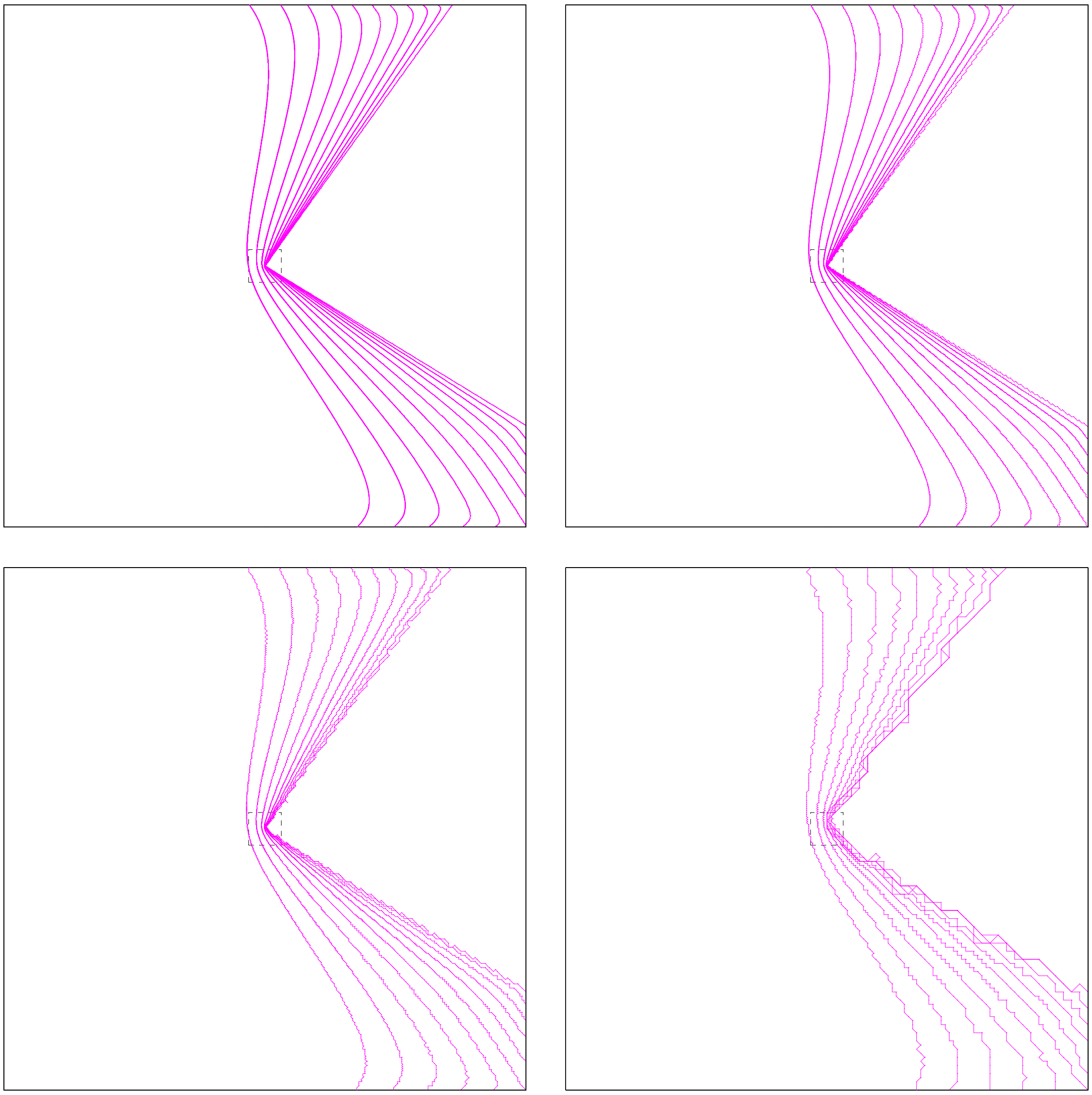

FIGURE 9.2. Interfaces for tolerances $\varepsilon = 1, 2, 4, 8$, at times $t = 0.01\,k$, $0 \le k \le 10$.

its zoom with scaling factor 16. It is clear that the interface is correctly captured by the algorithm even though the solution is very degenerate near the origin.

We stress that the minimum value of the meshsize is 1.2×10^{-5} for $\varepsilon = 1$, which would require 10^8 elements in our domain for a fixed quasi-uniform mesh to capture the singularity: adaptivity and local mesh refinement are thus essential. The number of triangles M vs. time for various tolerances is plotted in Figure 9.5: reducing the tolerance ε by a factor 2 entails an increase of M by a factor 4. The oscillations of M for $\varepsilon = 1$ are due to the upper limit $M_{\max}$ of triangles that the Fortran code is allowed to generate; when $M_{\max}$ is exceeded, the code automatically increases the tolerance ε. For $\varepsilon = 1$, we set $M_{\max} = 160,000$ and observe that oscillations likely occur after the corner smooths out, thereby not affecting our conclusions.

It may be surprising at first glance that a problem governed by a parabolic partial differential equation can exhibit a stationary corner singularity. This is a

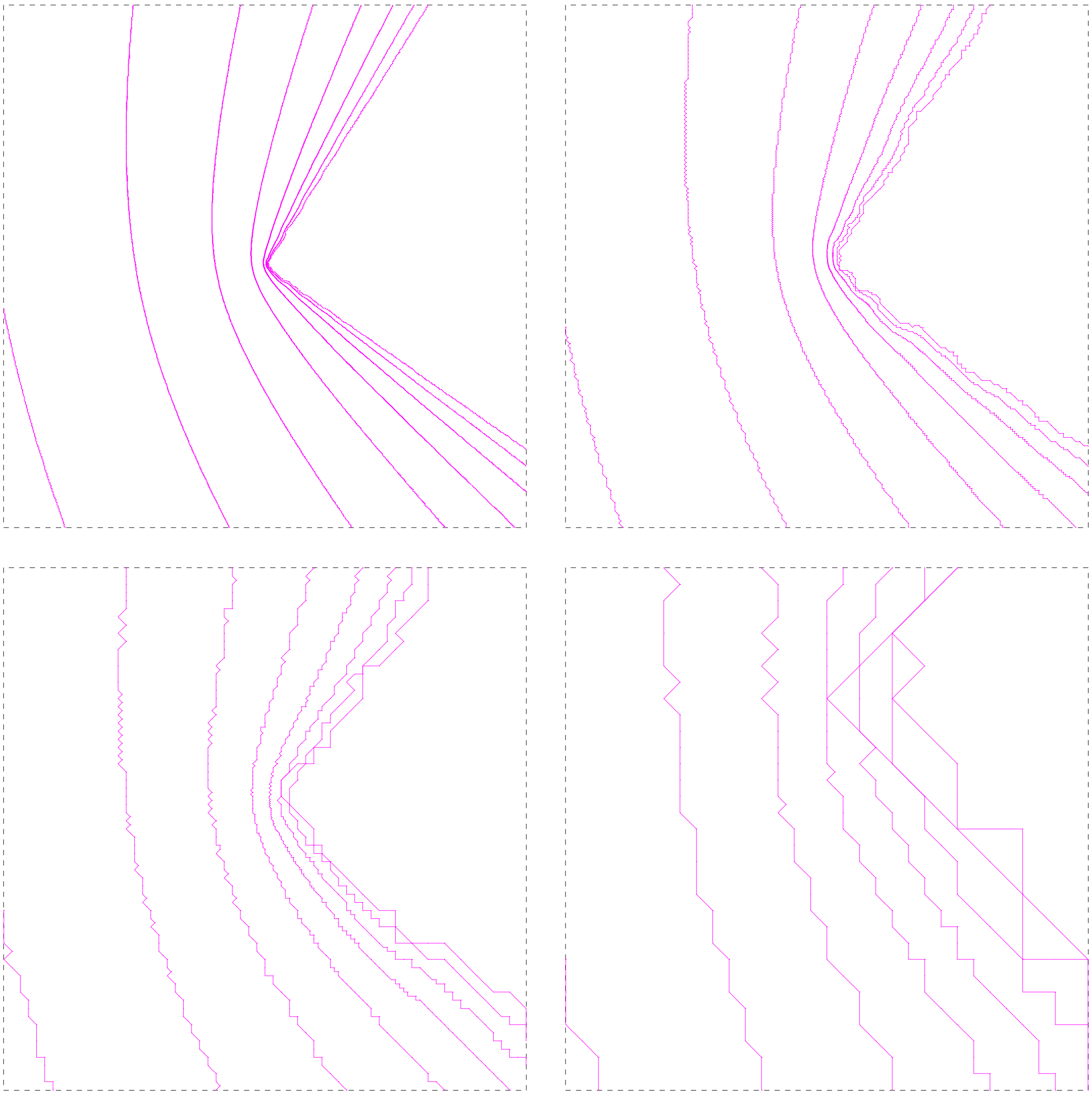

FIGURE 9.3. Zoom of interfaces with scaling factor 16 for tolerances $\varepsilon = 1, 2, 4, 8$, at times $t = 0.01\,k$, $3 \leq k \leq 10$.

hyperbolic effect consistent with the scaling at the interface. We could infer that the Stefan problem possesses a hyperbolic behavior near the interface. This is a structural property already used in [20], [21] for *a priori* design of refined meshes and a consequence of the *a posteriori* mesh design of §7.

This example corroborates the theory of Athanasopoulos, Caffarelli, and Salsa [1], namely that the following two factors may prevent smoothing: (a) the angle is not sufficiently large; (b) the heat fluxes of both phases vanish simultaneously. The theory predicts that failure of either (a) or (b) yields immediate smoothing, but it does not address the behavior for intermediate angles. This delicate question can only be explored numerically as shown above: the critical angle appears to be larger than $\pi/2$.

REFERENCES

1. I. Athanasopoulos, L. Caffarelli, and S. Salsa, *Degenerate phase transition problems of para-*

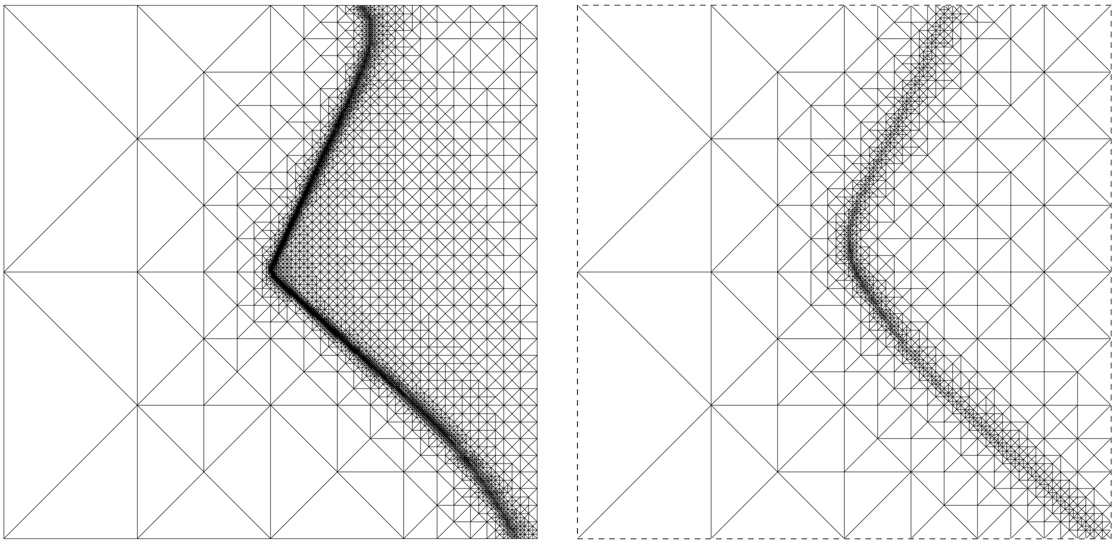

FIGURE 9.4. Mesh and Zoom with scaling factor 16 for tolerance $\varepsilon = 4$, at time $t = 0.06$ (number of triangles $M = 15,273$).

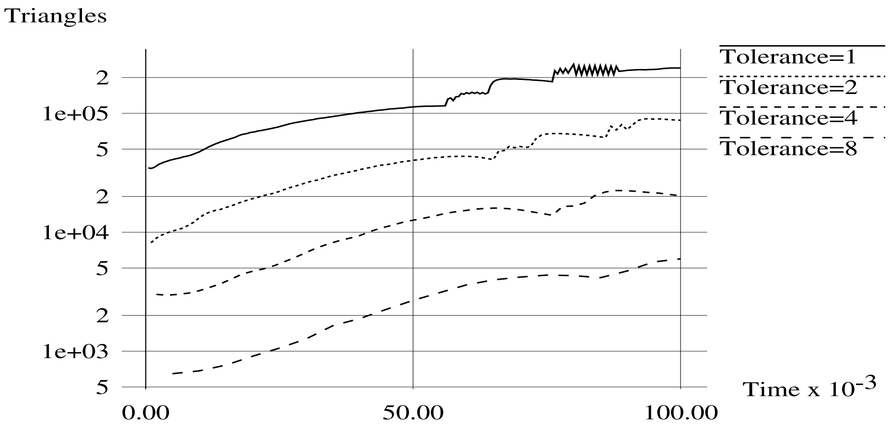

FIGURE 9.5. Number of triangles M vs. time for various tolerances; at $t = 0.05$, $M = 113,136$ for $\varepsilon = 1$, $M = 40,306$ for $\varepsilon = 2$, $M = 12,618$ for $\varepsilon = 4$, $M = 2,677$ for $\varepsilon = 8$.

bolic type. Smoothness of the front (to appear).

2. E. Bänsch, *Local mesh refinement in 2 and 3 dimensions*, IMPACT Comput. Sci. Engrg. **3** (1991), 181–191.
3. P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
4. Ph. Clément, *Approximation by finite element functions using local regularization*, RAIRO Modél. Math. Anal. Numér. **9** (1975), 77–84.
5. T. Dupont, *Mesh modification for evolution equations*, Math. Comp. **29** (1982), 85–107.
6. C.M. Elliott, *Error analysis of the enthalpy method for the Stefan problem*, IMA J. Numer. Anal. **7** (1987), 61–71.
7. K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems I: a linear model problem*, SIAM J. Numer. Anal. **28** (1991), 43–77.
8. J.W. Jerome and M.E. Rose, *Error estimates for the multidimensional two-phase Stefan problem*, Math. Comp. **39** (1982), 377–414.
9. X. Jiang and R.H. Nochetto, *Optimal error estimates for semidiscrete phase relaxation mod-*

els, RAIRO Modél. Math. Anal. Numér. **31** (1997), 91–120.
10. ———, *A finite element method for a phase relaxation model. Part I: quasi-uniform mesh*, SIAM J. Numer. Anal. (to appear).
11. X. Jiang, R.H. Nochetto, and C. Verdi, *A P^1-P^1 finite element method for a phase relaxation model. Part II: adaptively refined meshes*, SIAM J. Numer. Anal. (to appear).
12. R. Kornhuber, *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems*, Teubner, Stuttgart, 1997.
13. I. Kossaczký, *A recursive approach to local mesh refinement in two and three dimensions*, J. Comput. Appl. Math. **55** (1994), 275–288.
14. O.A. Ladyzenskaja, V. Solonnikov, and N. Ural'ceva, *Linear and Quasilinear Equations of Parabolic Type*, vol. TMM 23, AMS, Providence, 1968.
15. E. Magenes, R.H. Nochetto, and C. Verdi, *Energy error estimates for a linear scheme to approximate nonlinear parabolic problems*, RAIRO Modél. Math. Anal. Numér. **21** (1987), 655–678.
16. J.M. Maubach, *Local bisection refinement for n-simplicial grids generated by reflection*, SIAM J. Sci. Statist. Comput. **16** (1995), 210–227.
17. W. Mitchell, *A comparison of adaptive refinement techniques for elliptic problems*, ACM Trans. Math. Softw. **15** (1989), 326–347.
18. R.H. Nochetto, *Error estimates for multidimensional singular parabolic problems*, Japan J. Indust. Appl. Math. **4** (1987), 111–138.
19. ———, *Finite element methods for parabolic free boundary problems*, Advances in Numerical Analysis, (W. Light ed.), vol. I: Nonlinear Partial Differential Equations and Dynamical Systems, Oxford University Press, Oxford, 1991, pp. 34–88.
20. R.H. Nochetto, M. Paolini, and C. Verdi, *An adaptive finite elements method for two-phase Stefan problems in two space dimensions. Part I: stability and error estimates. Supplement*, Math. Comp. **57** (1991), 73–108, S1–S11.
21. ———, *An adaptive finite elements method for two-phase Stefan problems in two space dimensions. Part II: implementation and numerical experiments*, SIAM J. Sci. Statist. Comput. **12** (1991), 1207–1244.
22. ———, *A fully discrete adaptive nonlinear Chernoff formula*, SIAM J. Numer. Anal. **30** (1993), 991–1014.
23. R.H. Nochetto, A. Schmidt, and C. Verdi, *A posteriori error estimation and adaptivity for degenerate parabolic problems*, Math. Comp. (to appear).
24. ———, *Mesh and time step modification for degenerate parabolic problems*, in preparation.
25. ———, *Adaptive algorithm and simulations for Stefan problems in two and three dimensions*, in preparation.
26. R.H. Nochetto and C. Verdi, *Approximation of degenerate parabolic problems using numerical integration*, SIAM J. Numer. Anal. **25** (1988), 784–814.
27. ———, *An efficient linear scheme to approximate parabolic free boundary problems: error estimates and implementation*, Math. Comp. **51** (1988), 27–53.
28. J. Rulla, *Error analysis for implicit approximations to solutions to Cauchy problems*, SIAM J. Numer. Anal. **33** (1996), 68–87.
29. J. Rulla and Walkington, *Optimal rates of convergence for degenerate parabolic problems in two dimensions*, SIAM J. Numer. Anal. **33** (1996), 56–67.
30. C. Verdi, *Numerical aspects of parabolic free boundary and hysteresis problems*, Phase Transition and Hysteresis (A. Visintin ed.), vol. Lectures Notes in Mathematics 1584, Springer-Verlag, Berlin, 1994, pp. 213–284.
31. C. Verdi and A. Visintin, *Error estimates for a semiexplicit numerical scheme for Stefan-type problems*, Numer. Math. **52** (1988), 165–185.
32. A. Visintin, *Models of Phase Transitions*, Birkhäuser, Boston, 1996.

This work was partially supported by NSF GrantDMS-9623394, ESF Grant "Mathematical Treatment of Free Boundary Problems", MURST, and CNR Contract 96.03847.PS01.

Ricardo H. Nochetto: Department of Mathematics
and Institute for Physical Science and Technology
University of Maryland
College Park, MD 20742, USA.

Alfred Schmidt: Institut für Angewandte Mathematik
A.Ludwigs-Universität Freiburg
Hermann-Herder-Str. 10
79104 Freiburg, Germany.

Claudio Verdi: Dipartimento di Matematica
Università di Milano
Via Saldini 50
20133 Milano, Italy