

## AN ADAPTIVE FINITE ELEMENT METHOD FOR TWO-PHASE STEFAN PROBLEMS IN TWO SPACE DIMENSIONS. PART I: STABILITY AND ERROR ESTIMATES

R. H. NOCHETTO, M. PAOLINI, AND C. VERDI

**ABSTRACT.** A simple and efficient adaptive local mesh refinement algorithm is devised and analyzed for two-phase Stefan problems in 2D. A typical triangulation is coarse away from the discrete interface, where discretization parameters satisfy a parabolic relation, whereas it is locally refined in the vicinity of the discrete interface so that the relation becomes hyperbolic. Several numerical tests are performed on the computed temperature to extract information about its first and second derivatives as well as to predict discrete free boundary locations. Mesh selection is based upon equidistributing pointwise interpolation errors between consecutive meshes and imposing that discrete interfaces belong to the so-called refined region. Consecutive meshes are not compatible in that they are not produced by enrichment or coarsening procedures but rather regenerated. A general theory for interpolation between noncompatible meshes is set up in  $L^p$ -based norms. The resulting scheme is stable in various Sobolev norms and necessitates fewer spatial degrees of freedom than previous practical methods on quasi-uniform meshes, namely  $O(\tau^{-3/2})$  as opposed to  $O(\tau^{-2})$ , to achieve the same global asymptotic accuracy; here  $\tau > 0$  is the (uniform) time step. A rate of convergence of essentially  $O(\tau^{1/2})$  is derived in the natural energy spaces provided the total number of mesh changes is restricted to  $O(\tau^{-1/2})$ , which in turn is compatible with the mesh selection procedure. An auxiliary quasi-optimal pointwise error estimate for the Laplace operator is proved as well. Numerical results illustrate the scheme's efficiency in approximating both solutions and interfaces.

### 1. INTRODUCTION

A common feature in dealing with *degenerate* parabolic equations is the intrinsic lack of regularity of solutions across the interfaces (or free boundaries) which, in turn, are not known in advance. For the two-phase Stefan problem, for instance, the temperature  $\theta$  cannot be better than Lipschitz continuous and the enthalpy  $u$  (or energy density) typically exhibits a jump discontinuity across the interface. They satisfy the PDE

$$(1.1) \quad u_t - \operatorname{div}(k(\theta)\nabla\theta) = f(\theta) \quad \text{in } \Omega \times (0, T),$$

Received September 21, 1989.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 65N15, 65N30, 35R35.

This work was partially supported by NSF Grants DMS-8805218 and DMS-9008999, and by MPI (Fondi per la Ricerca Scientifica 40%) and CNR (IAN, Contract 880032601 and Progetto Finalizzato "Sistemi Informatici e Calcolo Parallelo", Sottoprogetto "Calcolo Scientifico per Grandi Sistemi") of Italy.

subject to the *strongly nonlinear* constitutive relation  $\theta = \beta(u)$ , where  $\beta$  vanishes on  $[0, 1]$ . The situation is more intricate when cusps and/or mushy regions develop. This lack of smoothness makes finite element approximations, defined on quasi-uniform meshes, perform worse than expected according to the interpolation theory. In other words, the singularity located on the interface pollutes the numerical solution everywhere. Numerical experiments for the Stefan problem indicate that the rate of convergence for temperature is never better than linear [17, 18]. Theoretical results are even more pessimistic [4, 8, 12, 13, 14, 17, 24].

Methods studied so far are not completely satisfactory in that they do not take advantage of the fact that singularities are located in a small region compared with the entire domain  $\Omega$ , at least whenever mushy regions do not occur. Consequently, a possible remedy is to be found in terms of a suitably designed adaptive algorithm. In fact, we would like to use a finer mesh near singularities in order to equidistribute interpolation errors but still preserve the number of degrees of freedom, and thus the computational complexity. We refer to [1, 14] for an account of the state-of-the-art on this topic along with numerous references.

In this light, the aim of this paper is to present and analyze an adaptive mesh refinement method for two-dimensional two-phase Stefan problems. We emphasize that such problems are *strongly nonlinear* in that singularities do not smooth out as time evolves and, more notably, they may even develop. This is a striking contrast between degenerate and purely parabolic equations. Therefore, even though various adaptive algorithms have been recently introduced for standard parabolic equations [1, 5, 14], ours appears to be the first one with a rigorous mathematical foundation for Stefan problems. We refer to [15], where a summary of some preliminary results can be found.

The finite element mesh cannot be modified in an arbitrary manner for the discrete scheme to be stable and convergent. Several tests are carried out on the computed temperature to extract information about its first and second derivatives as well as about the location of the discrete interface. Upon failure, the current mesh is discarded and a new one completely regenerated by an efficient automatic mesh generator [19]. Since the new mesh is not produced by enrichment or coarsening procedures, it happens to be *noncompatible* with the previous one. It is designed to be coarse away from the discrete free boundary, where the typical meshsize is  $O(\tau^{1/2})$ , and locally refined near the interface for triangles to reach a size  $O(\tau)$ ; hereafter,  $\tau > 0$  stands for the (uniform) time step. These relations, which come from elementary interpolation considerations, reflect the physical property that the Stefan problem behaves as parabolic away from the interface but possesses a first-order hyperbolic-like structure in its vicinity. On the other hand, even though the cost of generating a mesh at every single time step is asymptotically negligible compared to that of solving the associated nonlinear algebraic systems, frequent remeshing should be avoided for practical purposes. In addition, the interpolation process used to transfer infor-

mation between consecutive meshes incorporates an error  $O(\tau)$  that eventually accumulates in time. To prevent such an error from compromising accuracy, a restriction on the total number of admissible mesh changes is enforced, namely  $O(\tau^{-1/2})$ . As the current mesh is thus to be kept fixed for a prescribed number of time iterations and discrete interfaces are supposed to lie within the so-called *refined region*, a further refinement is required for our strategy to succeed. This is accomplished by predicting the “small” region to be occupied by the discrete interface between consecutive mesh changes, as well as checking that it actually remains there within safe limits. The resulting scheme is stable, convergent and necessitates fewer degrees of freedom than previous methods on fixed meshes, namely  $O(\tau^{-3/2})$  instead of  $O(\tau^{-2})$  for well-behaved interfaces, to achieve the same global asymptotic accuracy. Moreover, it exhibits a superior performance as expressed in terms of computing time for a desired accuracy. This improvement is even more dramatic when accuracy is measured in the maximum norm.

The paper is organized as follows. In §2 we formulate the continuous and discrete problems along with the corresponding assumptions. In §3 we comment on certain heuristic aspects of our local refinement strategy, which is fully discussed in §4. In §5 and Supplement §S1 we prove several  $L^p$ -based interpolation estimates for noncompatible meshes that account for mesh change effects and play a major role in our analysis. Discrete stability in various norms is then derived in §6 and Supplement §S2. As a result, in §7 we demonstrate a rate of convergence of essentially  $O(\tau^{1/2})$  for both  $\theta$  and  $u$  in the natural energy norms, provided the total number of mesh changes is limited to  $O(\tau^{-1/2})$ . This result agrees with previous ones [4, 17, 24] obtained for a fixed mesh under the *stronger* assumption that the meshsize is  $O(\tau)$ . We also prove, in Supplement §S3, an auxiliary quasi-optimal pointwise error estimate for the Laplace operator, that may have some independent interest in that it extends the techniques in [21, 22, 23] to general meshes; it is based upon a new discrete Caccioppoli estimate. To simplify the presentation, we assume that conductivity  $k = 1$  and that mushy regions do not occur. These interesting situations are, however, treated in some detail in Supplement §S4 along with a modification of the local mesh refinement algorithm. We conclude in §8 with several numerical experiments to illustrate the superior performance of the Adaptive Method in approximating both solutions and interfaces. Various computational issues are discussed in §8 as well.

Further numerical results and comparisons with the Fixed Mesh Method as well as implementation details will appear elsewhere [16]. They indicate a (practical) linear rate of convergence, namely  $O(\tau)$ , which is much better than our theoretical prediction. This topic deserves further investigation.

## 2. PROBLEM STATEMENT

Let  $\Omega \subset \mathbf{R}^2$  be a bounded domain with  $\partial\Omega \in C^{1,1}$  and  $T > 0$  be fixed. The case of polygonal domains will be considered in §S4.4. Let  $\beta: \mathbf{R} \rightarrow \mathbf{R}$  be

a continuous and nondecreasing function which satisfies

$$(2.1) \quad 0 < l_\beta \leq \beta'(s) \leq L_\beta < \infty, \quad |\beta''(s)| \leq L_{\beta'} < \infty, \quad \text{a.e. } s \in \mathbf{R} \setminus [0, 1], \\ \beta(s) = 0, \quad \forall s \in [0, 1];$$

hence  $L_\beta$  and  $L_{\beta'}$  are the Lipschitz constants of  $\beta$  and  $\beta'$ , respectively. A typical example is  $\beta(s) := c_1 s^- + c_2 (s - 1)^+$ , where  $c_1 < 0 < c_2$  are fixed; this corresponds to an ideal material with constant thermal properties. Let  $u_0$  indicate the initial enthalpy. Let  $\theta_0 := \beta(u_0)$  denote the initial temperature and let  $I_0 := \{x \in \Omega : \theta_0(x) = 0\}$  be the initial interface. They satisfy

$$(2.2) \quad \theta_0 \in W_0^{1,\infty}(\Omega) \cap W^{2,\infty}(\Omega \setminus I_0),$$

$$(2.3) \quad I_0 \text{ is a Lipschitz curve.}$$

Therefore,  $u_0$  is of bounded variation,  $u_0 \in W^{2,\infty}(\Omega \setminus I_0)$  and it has a jump discontinuity across  $I_0$ . In §S4.2 we will allow the initial interface  $I_0$  to degenerate into a mushy region. The source term  $f$  is also Lipschitz continuous, namely,

$$(2.4) \quad |f(s_1) - f(s_2)| \leq L_f |s_1 - s_2|, \quad \forall s_1, s_2 \in \mathbf{R}.$$

For the moment, the conductivity  $k$  verifies  $k = 1$ ; see §S4.3 for the general case. The continuous problem then reads as follows: *find  $\theta$  and  $u$  such that*

$$(2.5) \quad \theta \in L^2(0, T; H_0^1(\Omega)), \quad u \in L^\infty(0, T; L^\infty(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)),$$

$$(2.6) \quad \theta(x, t) = \beta(u(x, t)), \quad \text{a.e. } x \in \Omega, \quad t \in (0, T),$$

$$(2.7) \quad u(\cdot, 0) = u_0$$

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

$$(2.8) \quad \langle u_t, \varphi \rangle + \langle \nabla \theta, \nabla \varphi \rangle = \langle f(\theta), \varphi \rangle.$$

Hereafter,  $\langle \cdot, \cdot \rangle$  stands for the inner product on  $L^2(\Omega)$ . It is to be observed that the vanishing Dirichlet boundary condition on  $\theta$  is assumed only for simplicity and, in addition, that the *interface*  $I(t) := \{x \in \Omega : \theta(x, t) = 0\}$  does not include  $\partial\Omega$ . Existence and uniqueness for this problem are known as well as the following further regularity results [6, 8, 10]:

$$(2.9) \quad \theta \in H^1(0, T; L^2(\Omega)) \cap L^\infty(0, T; H_0^1(\Omega)), \quad \Delta \theta \in L^\infty(0, T; M(\Omega)),$$

where  $M(\Omega)$  stands for the space of finite regular Baire measures. In the classical situation, the free boundary motion is governed by the so-called *Stefan condition*

$$(2.10) \quad [\nabla \theta^+(x, t) - \nabla \theta^-(x, t)] \cdot \nu_x = V(x),$$

where  $x \in I(t)$ ,  $\nu_x$  is the unit vector normal to  $I(t)$  and  $V(x)$  is the normal velocity of  $I(t)$ , both at point  $x$ . Consequently, if  $V(x) \neq 0$ , the flux

$\nabla\theta(x, t)$  presents a jump discontinuity at  $x \in I(t)$  which prevents  $\theta$  from being better than globally Lipschitz continuous. Equation (2.10) may fail to hold, though, whenever cusps develop. Let  $[[\cdot]]$  denote the jump operator on  $I$ , that is  $[[\nabla\theta]] := (\nabla\theta^+ - \nabla\theta^-)|_I$ . The interface  $I$  may also degenerate into a *mushy region*, in which case (2.10) is to be replaced by the more general expression

$$(2.11) \quad [[\nabla\theta(x, t)]] \cdot \nu_x = [[u(x, t)]] V(x), \quad x \in \partial I(t) \cap \Omega.$$

As already said, this interesting situation is treated in some detail in §S4.2. For the moment, we suppose that mushy regions do not occur.

We now introduce the finite element approximation. Denote by  $\tau := T/N$  the time step and by  $\mathcal{S}^n$  a partition of  $\Omega$  into triangles;  $\mathcal{S}^n$  is assumed to be weakly acute and regular uniformly in  $1 \leq n \leq N$ . The first condition means that for any pair of adjacent triangles the sum of the opposite angles with respect to the common side does not exceed  $\pi$ . Given a triangle  $S \in \mathcal{S}^n$ ,  $h_S$  stands for its size and verifies  $\lambda\tau \leq h_S \leq \Lambda\tau^{1/2}$  ( $0 < \lambda, \Lambda$  fixed) whereas  $\rho_S$  denotes the diameter of the biggest ball contained in  $S$ . The second condition above is then equivalent to requiring  $\rho_S \geq \sigma h_S$  for all  $S \in \mathcal{S}^n$ , where  $0 < \sigma < 1$  is a fixed constant (independent of  $n$  and  $N$ !) [2, p. 132]. The discrete domain  $\Omega^n := \bigcup_{S \in \mathcal{S}^n} S$  does not coincide with  $\bar{\Omega}$ . However, since the technical arguments to handle their discrepancy were introduced in [17], we omit them here by simply assuming  $\Omega^n = \bar{\Omega}$ . Nonetheless, the influence of the pollution effect due to corners will be examined in §S4.4. Let  $\mathbf{V}^n \subset H_0^1(\Omega)$  indicate the usual piecewise linear finite element space over  $\mathcal{S}^n$  and  $\Pi^n: C^0(\bar{\Omega}) \rightarrow \mathbf{V}^n$  the usual Lagrange interpolation operator [2, p. 94]. Finally, let  $\{x_j^n\}_{j=1}^{J^n}$  denote the nodes of  $\mathcal{S}^n$  and  $\{\chi_j^n\}_{j=1}^{J^n}$  the canonical basis of  $\mathbf{V}^n$ . The discrete initial enthalpy  $U^0 \in \mathbf{V}^1$  is defined at a generic node  $x_j^1$  of  $\mathcal{S}^0 := \mathcal{S}^1$  to be

$$(2.12) \quad U^0(x_j^1) := u_0(x_j^1), \quad \forall x_j^1 \in \Omega \setminus I_0, \quad U^0(x_j^1) := 1, \quad \forall x_j^1 \in I_0.$$

Hence,  $U^0$  is easy to evaluate in practice. Set  $\Theta^0 := \Pi^1\theta_0$  ( $= \Pi^1[\beta(U^0)]$ ). Given  $U^{n-1}$ ,  $\Theta^{n-1} \in \mathbf{V}^{n-1}$ , the discrete scheme then reads as follows: for any  $1 \leq n \leq N$  select  $\mathcal{S}^n$  and find  $U^n$ ,  $\Theta^n \in \mathbf{V}^n$  such that

$$(2.13) \quad \Theta^n = \Pi^n \beta(U^n),$$

$$(2.14) \quad \hat{U}^{n-1} := \Pi^n U^{n-1}, \quad \hat{\Theta}^{n-1} := \Pi^n [\beta(U^{n-1})],$$

$$(2.15) \quad \tau^{-1} \langle U^n - \hat{U}^{n-1}, \chi \rangle^n + \langle \nabla \Theta^n, \nabla \chi \rangle = \langle f(\hat{\Theta}^{n-1}), \chi \rangle^n, \quad \forall \chi \in \mathbf{V}^n,$$

where  $\langle \cdot, \cdot \rangle^n$  is defined by

$$(2.16) \quad \langle \varphi, \chi \rangle^n := \int_{\Omega} \Pi^n(\varphi \chi) dx, \quad \forall \varphi, \chi \in C^0(\bar{\Omega}).$$

Note that the integral in (2.16) can be easily evaluated element-by-element via the vertex quadrature rule, which is exact for piecewise linear functions

[2, p. 182]. The discrete interface and refined region are then defined to be  $F^n := \{x \in \Omega : \Theta^n(x) = 0\}$  and  $\mathcal{R}^n := \bigcup\{S \in \mathcal{S}^n : h_S = O(\tau)\}$ , respectively.

In view of (2.16), the equation (2.15) can also be written as follows:

$$(2.17) \quad \tau^{-1} \langle U^n - U^{n-1}, \chi \rangle^n + \langle \nabla \Theta^n, \nabla \chi \rangle = \langle f(\beta(U^{n-1})), \chi \rangle^n, \quad \forall \chi \in \mathbf{V}^n,$$

thus eliminating the definitions in (2.14). It is to be stressed that the nodes of  $\mathcal{S}^n$  are then used for the numerical integration of piecewise linear functions defined in  $\mathbf{V}^{n-1}$ . The interpolation error so incurred may destroy convergence as well as stability. The mesh selection strategy of §4 will account for such an effect.

Observe that, if we first decompose the integral  $\langle \nabla \varphi, \nabla \chi \rangle$  over all triangles of  $\mathcal{S}^n$  and next integrate by parts, we get

$$(2.18) \quad \langle \nabla \varphi, \nabla \chi \rangle = \sum_{e \in \mathcal{E}^n} \langle \llbracket \nabla \varphi \rrbracket_e \cdot \nu_e, \chi \rangle_e, \quad \forall \varphi, \chi \in \mathbf{V}^n,$$

where  $\mathcal{E}^n := \{e : e \text{ is a side (or edge) of } S \text{ in } \Omega, S \in \mathcal{S}^n\}$ ,  $\langle \cdot, \cdot \rangle_e$  denotes the  $L^2$ -scalar product on  $e$ ,  $\nu_e$  is the unit vector normal to  $e$  and  $\llbracket \cdot \rrbracket_e$  indicates the jump operator on  $e$  for all  $e \in \mathcal{E}^n$ . Let  $S_j^n$  denote the interior of  $\text{supp } \chi_j^n$  for  $1 \leq j \leq J^n$  and set  $h_e := \text{length}(e)$  for  $e \in \mathcal{E}^n$ . Then (2.18) results in

$$(2.19) \quad 2 \langle \nabla \varphi, \nabla \chi_j^n \rangle = \sum_{e \subset S_j^n} h_e \llbracket \nabla \varphi \rrbracket_e \cdot \nu_e, \quad \forall \varphi \in \mathbf{V}^n, \quad \forall 1 \leq j \leq J^n.$$

In view of (2.16) and (2.19), another useful relation equivalent to (2.15) reads

$$(2.20) \quad \sum_{e \subset S_j^n} h_e \llbracket \nabla \Theta^n \rrbracket_e \cdot \nu_e = \frac{2}{3} \text{meas}(S_j^n) \left( f(\hat{\Theta}^{n-1}(x_j^n)) - \tau^{-1} (U^n - \hat{U}^{n-1})(x_j^n) \right).$$

From now on,  $C > 0$  will denote a constant independent of  $\tau$  but not necessarily the same at each occurrence. Moreover,  $C$  may depend on the given data as well as on the various constants to be introduced in §4. The notation  $\cdot = O(\tau^\gamma)$  will be often used instead of  $\cdot \leq C\tau^\gamma$ . As usual,  $|x|$  will stand for any norm of  $x \in \mathbf{R}^2$ .

### 3. HEURISTIC GUIDELINES

We now give a heuristic motivation for the local refinement strategy of §4. We first consider the following 1-D problem discretized only in time:

$$(3.1) \quad U - \tau \beta(U)_{xx} = u_0 \quad \text{in } (-1, 1),$$

where  $\beta(s) := (s-1)^+ - s^-$ ,  $V := \gamma_2 - \gamma_1 > 0$  and  $u_0(x) := \gamma_2(e^{Vx} - 1)/V + 1$  if  $x > 0$ ,  $u_0(x) := \gamma_1(e^{Vx} - 1)/V$  if  $x < 0$ . Since  $u(x, \tau) = u_0(x + V\tau)$ , the interface  $I(t)$ , initially at 0, reaches  $-V\tau$  at time  $\tau$ . Let  $\delta$  denote the

position of the discrete-time interface  $F$ . It is not difficult, but tedious, to see that ( $\Theta := \beta(U)$ )

$$(3.2) \quad \|\Theta - \theta(\tau)\|_{L^\infty((-1, 1))} = O(\tau) \geq V^2\tau,$$

$$(3.3) \quad \delta = -V\tau + O(\tau^{3/2}),$$

$$(3.4) \quad \Theta_x(\delta^-) = \Theta_x(\delta^+) = \gamma_1 + O(\tau^{1/2}) \quad (\text{smearing effect!}),$$

$$(3.5) \quad \Theta_x(0) = \gamma_2 + O(\tau^{1/2}),$$

$$(3.6) \quad \Theta_{xx}(x) = \tau^{-1} + O(1), \quad \forall \delta < x < 0.$$

On the other hand, suppose  $(-1, 1)$  is partitioned into equal intervals of size  $h$ . Then, in view of the shape of  $\theta(\tau) := \beta(u(\tau))$  and (3.6), the pointwise interpolation errors in space satisfy

$$(3.7) \quad \begin{aligned} \|\theta(\tau) - \Pi^1\theta(\tau)\|_{L^\infty((-1, 1))} &\leq Vh/4 + O(h^2), \\ \|\Theta - \Pi^1\Theta\|_{L^\infty((-1, 1))} &\leq h^2\tau^{-1}/8 + O(h^2). \end{aligned}$$

What we learn from this relevant example can be expressed as follows. Since we expect to deal with Lipschitz continuous temperatures, the local meshsize  $h_S$  near  $F$  and interface velocity  $V_S$  should verify  $h_S \approx V_S\tau$  to balance the interpolation errors in space (3.7) with the truncation error in time (3.2). In addition, no condition similar to (2.10) is valid for the semidiscrete problem at  $F$ , even though the free boundary moves correctly. To retrieve the proper jump condition, however, we just have to move a distance  $\delta$  backwards along the normal to  $F$  because, by virtue of (3.4) and (3.5),

$$(3.8) \quad \Theta_x(0) - \Theta_x(\delta^-) = V + O(\tau^{1/2}),$$

or, equivalently,  $V = \int_\delta^0 \Theta_{xx}(s)ds + O(\tau^{1/2})$ . Consequently, an overrefinement near the interface is extremely dangerous in that we may lose information on the interface velocity without gaining accuracy and, as a result, we might be in trouble to predict its future position. We thus realize that enforcing these two observations would require a stepwise control of the relation  $h_S \approx V_S\tau$ , where  $V_S$  could be determined by means of (3.8) with  $\delta$  being replaced by  $h_S$ . On the other hand, there is an interval  $O(\tau)$ -long behind  $F$ , namely  $(\delta, 0)$ , on which second derivatives are  $O(\tau^{-1})$ .

Away from the interface  $I$ , problem (1.1) is strictly parabolic, namely,

$$(3.9) \quad c(\theta)\theta_t - \Delta\theta = f(\theta),$$

which is a mildly nonlinear heat equation;  $c(s) := 1/\beta'(\beta^{-1}(s))$  for all  $s \in \mathbb{R} \setminus [0, 1]$ . Hence the discretization parameters should verify the usual parabolic constraint  $h_S = O(\tau^{1/2})$ .

These two distinct behaviors, rephrased here in terms of local regularity, must be reflected in the local refinement algorithm, for instance, as illustrated in §4.

Let us now explore some heuristic properties of the fully discrete scheme (2.15). Near the discrete interfaces, where the best we can say is  $|(U^n - \widehat{U}^{n-1})(x_j^n)| \leq C$ , (2.20) results in  $|\sum_{e \subset S_j^n} h_e \llbracket \nabla \Theta^n \rrbracket_e \cdot \nu_e| \leq C \tau^{-1} \text{meas}(S_j^n)$ . Hence, except for a very unlikely cancellation in the above summation, we can expect discrete second derivatives  $D_e$  to verify

$$(3.10) \quad D_e := h_e^{-1} |\llbracket \nabla \Theta^n \rrbracket_e| \leq C \tau^{-1} h_e^{-2} \text{meas}(S_j^n) \leq C \tau^{-1}.$$

This is consistent with (3.6). Moreover, away from the discrete free boundaries, we can expect  $|(U^n - \widehat{U}^{n-1})(x_j^n)| \leq C |(\Theta^n - \widehat{\Theta}^{n-1})(x_j^n)| \leq C \tau$ , because of (2.1), (2.13), (2.14) and the strict parabolicity of (3.9). Thus  $|\sum_{e \subset S_j^n} h_e \llbracket \nabla \Theta^n \rrbracket_e \cdot \nu_e| \leq C \text{meas}(S_j^n)$ . Therefore, arguing as before, we conclude that

$$(3.11) \quad D_e \leq C h_e^{-2} \text{meas}(S_j^n) \leq C,$$

for all  $e$  in the parabolic region. The heuristic observations (3.10) and (3.11) regarding  $D_e$ , as well as the smearing effect (3.4), were confirmed by 2-D numerical experimentation. It also revealed the validity of the following  $L^1$ -type a priori estimate:

$$(3.12) \quad \sum_{e \in \mathcal{E}^n} h_e^2 D_e \leq C.$$

This property is a discrete analogue of (2.9), i.e.,  $\Delta \theta \in L^\infty(0, T; M(\Omega))$ . It is still in good agreement with numerical evidence. Indeed, actual computations show the occurrence of a strip  $O(\tau)$ -wide behind the discrete interface  $F^n$  where  $D_e = O(\tau^{-1})$ , which in turn is consistent with (3.6) and (3.10). In this case, since the local meshsize near  $F^n$  should be  $h_e = O(\tau)$ , (3.12) imposes a severe regularity restriction on the interface, namely,

$$(3.13) \quad \text{length}(F^n) \leq \sum_{e \in \mathcal{E}^n : e \cap F^n \neq \emptyset} h_e \leq C \sum_{e \in \mathcal{E}^n} h_e^2 D_e \leq C.$$

Such a condition is quite reasonable for practical purposes but is not known to hold in a general setting. We stress that without some kind of additional regularity it is probably hopeless to improve upon the Fixed Mesh Method [4, 8, 12, 13, 17, 24]. In this light, (3.12) is always assumed at the mesh changes and used in §§6, 7, S4, though it constitutes a limitation of the Adaptive Method. It is however partially justified by Lemma 6.5 which, being implicitly guaranteed by the scheme, combines with (2.20) to yield

$$(3.14) \quad \sum_{j=1}^{J^n} \left| \sum_{e \subset S_j^n} h_e \llbracket \nabla \Theta^n \rrbracket_e \cdot \nu_e \right| \leq 2 \int_{\Omega} \Pi^n (|f(\widehat{\Theta}^{n-1})| + \tau^{-1} |U^n - \widehat{U}^{n-1}|) \leq C,$$

for all time steps  $n$  between consecutive mesh changes. We then see that only a cancellation in the above summation could lead to a bound weaker than (3.12). This seems to be unlikely for locally smooth interfaces, as well as for cusps,

because of their local character. At the same time, Lemma 5.10 shows that (3.12) is preserved for  $\hat{\Theta}^n$ . This somehow explains the fact that (3.12) was never violated in our numerical experiments. Designing an algorithm for which (3.12) is implicitly guaranteed constitutes a challenging open problem, though. Note that mushy regions may occur as long as their boundaries are also well-behaved. On the other hand, regarding first derivatives, the following  $L^2$ -type a priori estimate is implicitly guaranteed (see (6.6)):

$$(3.15) \quad \sum_{S \in \mathcal{S}^n} h_S^2 |\nabla \Theta^n|_S|^2 \leq C.$$

We finally comment upon the effect of interpolation between noncompatible meshes. Let  $\zeta: \mathbf{R} \rightarrow \mathbf{R}$  be sufficiently smooth and  $S \in \mathcal{S}^n$  be a generic element. Proceed then formally, as if  $\Theta^{n-1}$  were smooth, to deduce that

$$(3.16) \quad \begin{aligned} & \|\zeta(\Theta^{n-1}) - \Pi^n[\zeta(\Theta^{n-1})]\|_{L^\infty(S)} \\ & \leq Ch_S^2 \left( \|D^2 \Theta^{n-1}\|_{L^\infty(S)} + \|D\Theta^{n-1}\|_{L^\infty(S)}^2 \right), \end{aligned}$$

where  $D$  and  $D^2$  denote discrete first and second derivatives, respectively. In §5 we give proper justification for (3.16). Since we want this interpolation error to be  $O(\tau)$ , the new local meshsize should satisfy

$$(3.17) \quad h_S \leq \tau^{1/2} \min \left( \mu_1 \|D^2 \Theta^{n-1}\|_{L^\infty(S)}^{-1/2}, \mu_2 \|D\Theta^{n-1}\|_{L^\infty(S)}^{-1} \right).$$

This in turn allows second derivatives to blow up without violating  $h_S \geq \lambda\tau$  as long as  $\|D^2 \Theta^{n-1}\|_{L^\infty(S)} \leq (\mu_1/\lambda)^2 \tau^{-1}$ , which is consistent with (3.10). First derivatives may also degenerate without violating  $h_S \geq \lambda\tau$  provided  $\|D\Theta^{n-1}\|_{L^\infty(S)} \leq (\mu_2/\lambda)\tau^{-1/2}$ . Such a degeneracy is expected only whenever cusps develop, this being a local phenomenon. In addition, having control of quadrature errors introduced by (2.16) leads to restrictions on triangle diameters wherever  $\|D\Theta^{n-1}\|_{L^\infty(S)}$  exceeds a certain tolerance; this is accomplished via (3.17) as well.

On the other hand, for all  $S \in \mathcal{S}^{n-1}$  intersecting the discrete interface  $F^{n-1}$  we have  $\{x \in S : 0 \leq U^{n-1}(x) \leq 1\} \neq \emptyset$ . For sample problems having a nicely behaved continuous free boundary and verifying a nondegeneracy property, numerical experiments indicate that  $U^{n-1}$  may vary from 0 to 1 within one single element. Consequently, even a slight perturbation of triangles  $S$  traversed by  $F^{n-1}$  would produce an error  $\|U^{n-1} - \Pi^n U^{n-1}\|_{L^\infty(S)} = O(1)$  and a subsequent optimal lower bound  $\|U^{n-1} - \Pi^n U^{n-1}\|_{L^1(\Omega)} \geq C\tau$ , which could be attained provided length  $(F^{n-1}) = O(1)$ . This property of  $F^{n-1}$  is not enough, however, to ensure the validity of another crucial interpolation estimate (Lemma 5.6), namely  $\|\nabla(\Theta^{n-1} - \hat{\Theta}^{n-1})\|_{L^2(\Omega)} \leq C\tau^{1/2}$ , unless  $\nabla\Theta^{n-1}$  is bounded on  $F^{n-1}$ ; strong stability would thus break down too (Lemma 6.4). Since such a further

constraint on  $\nabla\Theta^{n-1}$  rules out the formation of cusps, we should not modify triangles crossed by  $F^{n-1}$ . For computational purposes, it is always preferable not to impose this constraint, which is feasible whenever the interface is “smooth”; see §S4.1.

#### 4. LOCAL REFINEMENT STRATEGY

The aim of this section is to describe the relation between two successive and noncompatible meshes, say  $\mathcal{S}^{n-1}$  and  $\mathcal{S}^n$ , along with the necessary numerical tests to be performed on  $\Theta^{n-1}$ . The initial mesh  $\mathcal{S}^0$  ( $=\mathcal{S}^1$ ) is constructed along the same lines with  $\theta_0$  in place of  $\Theta^{n-1}$ . Since no confusion is possible, for simplicity we remove the superscripts and use the following notation:  $\mathcal{S} := \mathcal{S}^{n-1}$ ,  $\widehat{\mathcal{S}} := \mathcal{S}^n$ ,  $\mathcal{E} := \mathcal{E}^{n-1}$ ,  $\widehat{\mathcal{E}} := \mathcal{E}^n$ ,  $\mathcal{R} := \mathcal{R}^{n-1}$ ,  $\widehat{\mathcal{R}} := \mathcal{R}^n$ ,  $\Pi := \Pi^{n-1}$ ,  $\widehat{\Pi} := \Pi^n$ ,  $U := U^{n-1}$ ,  $\widehat{U} := \widehat{U}^{n-1}(= \widehat{\Pi}U)$ ,  $\Theta := \Theta^{n-1}(= \Pi\beta(U))$ ,  $\widehat{\Theta} := \widehat{\Pi}\beta(U)$  and  $F := F^{n-1}$ , for  $2 \leq n \leq N$ . In §4.1 we introduce three local parameters that represent the expected value of local meshsize. We discuss the mesh selection algorithm in §4.2 and conclude with several comments in §4.3.

**4.1. Local mesh parameters.** Note that  $\nabla\Theta|_S, \nabla U|_S \in [P^0(S)]^2$  for all  $S \in \mathcal{S}$ . Set  $d_S := |\nabla\Theta|_S|$ ,  $D_S := |\nabla U|_S|$ ,  $\delta_S := h_S d_S$  for all  $S \in \mathcal{S}$  and

$$(4.1) \quad \delta_e := |\nabla\Theta|_{S_1} - \nabla\Theta|_{S_2}| = \|[\nabla\Theta]_e\|, \quad D_e := \frac{\delta_e}{h_e}, \quad \forall e \in \mathcal{E},$$

where  $S_1, S_2 \in \mathcal{S}$  are so that  $e = S_1 \cap S_2$ ; set  $\mathcal{A}_e := \{S_1, S_2\}$ . Note that these quantities are easy to evaluate in practice. We then introduce the following local parameters:

$$(4.2) \quad \hat{h}_e := \mu_1 \frac{\tau^{1/2}}{D_e^{1/2}}, \quad \forall e \in \mathcal{E} \setminus \mathcal{E}_F,$$

$$(4.3) \quad \hat{h}_S := \mu_2 \frac{\tau^{1/2}}{d_S}, \quad \forall S \in \mathcal{S} \setminus \mathcal{S}_F,$$

where  $\mathcal{S}_F := \{S \in \mathcal{S} : S \cap F \neq \emptyset\}$ ,  $\mathcal{E}_F := \{e \in \mathcal{E} : e \subset \partial S, S \in \mathcal{S}_F\}$  and  $\mathcal{F} := \bigcup_{S \in \mathcal{S}_F} S$ . Here,  $\mu_1, \mu_2 > 0$  are arbitrary constants which, in practice, result from computational considerations as well as specific properties of the problem at hand; the same comment applies to  $\lambda, \Lambda$ . The two local parameters above account for the effect of interpolation between noncompatible meshes, as motivated by (3.17). In case they violate the constraint  $\hat{h}_e, \hat{h}_S \geq \lambda\tau$ , we say that discrete derivatives are badly-behaved. This situation will require special care, even though it was never observed in practice. To this end, we set  $\mathcal{S}_B := \{S \in \mathcal{S} \setminus \mathcal{S}_F : \min_{e \in \mathcal{E} \setminus \mathcal{E}_F : e \subset \partial S} (\hat{h}_S, \hat{h}_e) < \lambda\tau\}$ ,  $\mathcal{E}_B := \{e \in \mathcal{E} \setminus \mathcal{E}_F : e \subset \partial S, S \in \mathcal{S}_B\}$ ,  $\mathcal{B} := \bigcup_{S \in \mathcal{S}_B} S$  and  $\mathcal{S}_0 := \mathcal{S} \setminus (\mathcal{S}_F \cup \mathcal{S}_B)$ ,  $\mathcal{E}_0 := \mathcal{E} \setminus (\mathcal{E}_F \cup \mathcal{E}_B)$ ,  $\Omega_0 := \Omega \setminus (\mathcal{F} \cup \mathcal{B})$ .

We now focus our attention on the local meshsize near  $F$  for problems without mushy regions. Inspired by (3.8) and subsequent heuristics  $h_S \approx V_S \tau$ ,

for each  $S \in \mathcal{S}$  crossed by the discrete interface  $F$ , we compute a discrete interface velocity  $V_S$  using a suitable discretization of condition (2.10), namely,

$$(4.4) \quad V_S := (\nabla \Theta|_{S_1} - \nabla \Theta|_{S_2}) \cdot \nu,$$

where  $S_1, S_2 \in \mathcal{S}$  belong to each phase, are displayed on the direction  $\nu$  (unit vector normal to  $F$ ) and satisfy  $\text{dist}(S_i, S) \geq h_S$  ( $i = 1, 2$ ). We next consider a cone  $\mathcal{E}_S$  of axis  $\nu$ , vertex at  $S$ , opening  $\pi/2$  and height  $\mu_3 V_S \tau^{1/2}$  as being the region most likely to contain the evolution of  $F_S := S \cap F$  for at least  $O(\tau^{-1/2})$  time steps. The local parameter associated with the interface is defined by

$$(4.5) \quad \hat{h}_{F_S} := \tau \min\{\max(\lambda, V_S), M\}.$$

The above two new constants  $\mu_3$  and  $M$  are arbitrary at this stage. The same rules of selection as for the previous four constants apply;  $\mu_3$  may depend on  $n$ .

The union of all cones  $\mathcal{E}_S$  constitutes the predicted refined region  $\hat{\mathcal{R}}$  whose width is  $O(\tau^{1/2})$ . Note that  $\hat{\mathcal{R}} \subset \widehat{\mathcal{R}}$ . The size of  $\hat{\mathcal{R}}$  enables discrete interfaces to remain within  $\widehat{\mathcal{R}}$  for at least  $O(\tau^{-1/2})$  time steps, as desired.

**4.2. Mesh selection algorithm.** As already said, the initial mesh  $\mathcal{S}^1$  is built with the required pointwise information extracted from  $\theta_0$ . Assuming now that we have a mesh  $\mathcal{S}$ , we would like to discuss the three tests to be performed on the computed solution  $\Theta$  to either accept or discard  $\mathcal{S}$ .

The first test consists of checking whether the discrete interface  $F$  is within the refined region  $\mathcal{R}$  or not. In the event  $F$  escapes from  $\mathcal{R}$ , we say that the test has failed.

The second test ascertains that interpolation errors are still equidistributed correctly:

$$(4.6) \quad h_e \leq \mu_1^* \hat{h}_e, \quad \forall e \in \mathcal{E}_0, \quad h_S \leq \mu_2^* \hat{h}_S, \quad \forall S \in \mathcal{S}_0;$$

here  $\mu_1^*, \mu_2^* > 1$  are suitable constants. This rules out the possibility of an excessive refinement induced by large discrete derivatives. However, the new local meshsize might be much smaller than the current one, if influenced by the new refined region  $\widehat{\mathcal{R}}$ . The example in §8.3, for instance, makes (4.6) fail; see also [16, §7].

Sometimes the interface velocity may vary substantially during an  $O(\tau^{1/2})$  period of time so as to make (4.5) inadequate. More specifically, the local truncation error (3.2) would not be properly reflected in the local meshsize and also, in case the current meshsize becomes too small, the computation of  $V_S$  via (4.4) might be inaccurate because of the smearing effect (3.4). In addition, the fact that triangles of  $\mathcal{S}_F$  are fixed and new nearby elements might have a much smaller size would create serious programming difficulties in specifying  $\widehat{\mathcal{R}}$ . To prevent that from happening, a third test is enforced, namely,

$$(4.7) \quad \mu_3^- \hat{h}_{F_S} \leq h_S \leq \mu_3^+ \hat{h}_{F_S}, \quad \forall S \in \mathcal{S}_F,$$

where  $\mu_3^- < 1 < \mu_3^+$  are suitable constants. A relevant example that makes (4.7) fail is that in [16, §7.3].

If any one of the above tests fails, then the current mesh  $\mathcal{S}$  is rejected as well as the solution  $\{\Theta, U\}$ , which is overwritten with the previously computed solution. A new *graded* mesh  $\widehat{\mathcal{S}}$  with the following properties is then generated. To preserve the constraint  $h_{\widehat{\mathcal{S}}} \geq \lambda\tau$ , we must keep  $\mathcal{S}_B$  fixed because discrete derivatives are badly-behaved. In addition, in accord with the last heuristic observation of §3, we must not modify  $\mathcal{S}_F$ . Hence,

$$(4.8) \quad S \in \widehat{\mathcal{S}}, \quad \forall S \in \mathcal{S}_B \cup \mathcal{S}_F,$$

is the first restriction on  $\widehat{\mathcal{S}}$ . The second one reads

$$(4.9) \quad \lambda\tau \leq h_{\widehat{\mathcal{S}}} \leq \min_{\substack{S' \in \mathcal{S}_F : \mathcal{E}_{S'} \cap \widehat{\mathcal{S}} \neq \emptyset \\ S \in \mathcal{S}_0, e \in \mathcal{E}_0 : S \cap \widehat{\mathcal{S}} \neq \emptyset, e \subset \partial S}} (\Lambda\tau^{1/2}, \hat{h}_{F_{S'}}, \hat{h}_e, \hat{h}_S), \quad \forall \widehat{S} \in \widehat{\mathcal{S}}.$$

This accounts for both the equidistribution of pointwise interpolation errors (3.17) and the definition of refined region  $\widehat{\mathcal{R}}$ . The effective implementation of (4.9) will be discussed in §8.1; see also [16, §6].

**4.3. Further properties.** The information about discrete derivatives could be extracted from  $U$  rather than  $\Theta$  because, in view of (2.1) and (2.13), they are equivalent on  $\Omega_0$ . In fact, for all  $e \in \mathcal{E}_0$  and  $S \in \mathcal{S}_0$ , we have

$$(4.10) \quad h_e^{-1} |\nabla U|_{S_1} - \nabla U|_{S_2}| \leq C(l_\beta^{-1} D_e + L_\beta l_\beta^{-3} (d_{S_1}^2 + d_{S_2}^2)), \quad S_i \in \mathcal{A}_e, \\ D_S \leq C l_\beta^{-1} d_S.$$

Since  $\mathcal{S}$  was designed to be adequate for at least  $O(N^{1/2})$  time steps, the number of expected mesh changes is at most  $O(N^{1/2})$ . This goal was always achieved in practice.

To avoid rejecting the computed solution  $\{\Theta, U\}$  owing to failure of the first test, we always check if the discrete interface  $F$  has just reached the boundary of the refined region  $\mathcal{R}$ , called RED ZONE, which in turn alerts that an imminent remeshing must be done; see Figures 9.1 and 9.2. On the other hand, to prevent the program from performing a useless time step owing to failure of either (4.6) or (4.7), these tests can be carried out with more stringent constants. In that case, their failure will only warn that  $\mathcal{S}$  cannot be kept any longer. This trick actually succeeds because discrete derivatives may exhibit large oscillations solely near the discrete interface  $F$ , and thus within  $\mathcal{R}$  where the current (local) meshsize is already  $O(\tau)$ . Hence, remeshing is mostly dictated by the free boundary location, as observed in practice.

In the subsequent analysis of §§6, 7, S4, we will assume the following *structural* property which, in view of §3, is partially justified by Lemma 6.5 and numerical evidence:

$$(4.11) \quad \sum_{e \in \mathcal{E}} h_e^2 D_e \leq C.$$

The definition of the sets  $\mathcal{B}$  and  $\widehat{\mathcal{R}}$  might be contradictory unless  $\mathcal{B} \subset \mathcal{R}$ , that is  $h_S = O(\tau)$  for all  $S \in \mathcal{S}_B$ . Discrete derivatives  $D_e$  and  $d_S$  are typically well-behaved outside  $\mathcal{R}$ , because this is a parabolic region where  $D_e, d_S \leq C$  at the previous mesh change. Consequently, we do not expect any rapid variation of either  $D_e$  or  $d_S$  on  $\Omega \setminus \mathcal{R}$ , which means  $\mathcal{B} \subset \mathcal{R}$  and also  $\mathcal{B} \subset \widehat{\mathcal{R}}$ . The set  $\mathcal{B}$ , though, was always empty in our numerical experiments. Note that  $\mathcal{F} \subset \widehat{\mathcal{R}}$  as well.

For a well-behaved interface, (4.5) coincides asymptotically with the usual *hyperbolic* relation  $h = O(\tau)$ . This was proposed here as a means to balance interpolation errors and attenuate the smearing effect, rather than for stability purposes. Stability is always built into the scheme regardless of the number of mesh changes, as shown in §6. Our algorithm is still a *fixed domain method*, even though we predict the region to be invaded by the discrete interface. Indeed, we do not use predicted interfaces to solve uncoupled (nonlinear) heat equations, as customary for front-tracking methods, but rather as a refinement indicator. The behavior of  $D_e$  and  $d_S$  depend certainly upon regularity of the underlying problem. We may think of these quantities as being bounded uniformly in  $\tau$  away from  $F$ , where  $\theta$  is expected to be smooth; see (3.11). Expressions (4.2) and (4.3), combined with (4.9), then result in  $h_S = O(\tau^{1/2})$ , which is the usual *parabolic* relation. We finally observe that the assumption  $\partial\Omega \in C^{1,1}$  avoids further refinements to alleviate the pollution effect produced by corner singularities; see §S4.4.

## 5. INTERPOLATION ESTIMATES FOR NONCOMPATIBLE MESHES

Our goal now is to show that the above criteria for mesh selection guarantee a satisfactory error control. The results in §5.1 are valid in general for regular and noncompatible meshes  $\mathcal{S}$  and  $\widehat{\mathcal{F}}$  and possess some intrinsic interest. They are next applied to the present setting in §§5.2, 5.3. We will stick with the notation of §4.

**5.1. The basic estimates.** Let us first introduce some further notation and a number of useful geometric properties. Given  $W \subset \bar{\Omega}$ , set

$$\mathcal{S}_W := \{S \in \mathcal{S}_0 : S \cap W \neq \emptyset\}, \quad \mathcal{E}_W := \{e \in \mathcal{E}_0 : e \cap W \neq \emptyset\}, \quad \widetilde{W} := \bigcup_{S \in \mathcal{S}_W} S.$$

Let  $B(x, r)$  indicate the ball of center  $x$  and radius  $r$ . The following facts are simple consequences of the regularity of  $\mathcal{S}$ : there exists  $0 < a < 1$  such that for all  $S \in \mathcal{S}_0$

$$(5.1) \quad \rho_{S'} \geq ah_S, \quad \forall S' \in \mathcal{S}_S,$$

$$(5.2) \quad \text{dist}(x, S) \geq ah_S, \quad \forall x \in \bar{\Omega}_0 \setminus \widetilde{S},$$

$$(5.3) \quad \text{card } \mathcal{S}_S, \text{ card } \mathcal{E}_S = O(1).$$

It is then possible to introduce a smooth function  $h(x)$  which is locally comparable with the meshsize (see Lemma 5.1 in Supplement §S1).

In deriving the interpolation estimates below we shall distinguish between two opposite situations according to the relative size of triangles in both meshes  $\mathcal{S}$  and  $\widehat{\mathcal{S}}$ . It is worth noting that there is no assumption on the relative size or location of new and old triangles at this stage. Set  $\widehat{\mathcal{S}}_0 := \widehat{\mathcal{S}} \setminus (\mathcal{S}_F \cup \mathcal{S}_B)$  and define the *derefinement* case to be

$$(5.4) \quad \text{given } \widehat{S} \in \widehat{\mathcal{S}}_0, \quad ah_S < h_{\widehat{S}} \quad \text{for all } S \in \mathcal{S}_{\widehat{S}}.$$

By contrast, the *refinement* situation reads as follows

$$(5.5) \quad \text{given } \widehat{S} \in \widehat{\mathcal{S}}_0, \quad \text{there exists } S \in \mathcal{S}_{\widehat{S}} \quad \text{such that } ah_S \geq h_{\widehat{S}}.$$

These two cases are obviously mutually exclusive. In addition, for all  $\widehat{S}$ 's satisfying (5.5) we have

$$(5.6) \quad \widehat{S} \subset \widetilde{S} \quad \text{for some } S \in \mathcal{S}_{\widehat{S}}, \quad \text{card } \mathcal{E}_{\widehat{S}}, \quad \text{card } \mathcal{S}_{\widehat{S}} = O(1),$$

as results from (5.1), (5.2) and (5.3). Let  $\widehat{\mathcal{S}}_1$  (resp.  $\widehat{\mathcal{S}}_2$ ) indicate the set of all  $\widehat{S}$ 's satisfying (5.4) (resp. (5.5)). Let  $V$  be a generic piecewise linear function defined on  $\mathcal{S}$ . Let  $\zeta: \mathbf{R} \rightarrow \mathbf{R}$  be so that  $\zeta' \in W^{1,\infty}(\mathbf{R})$ . Let  $L_\zeta$  and  $L_{\zeta'}$  denote the Lipschitz constants of  $\zeta$  and  $\zeta'$ , respectively. Let  $\delta_S, d_S, \delta_e$  and  $D_e$  indicate, for the moment, derivatives extracted from  $V$ . We now state, and then prove in Supplement §S1, two crucial estimates, their difference lying essentially in their derivation and further application. The first one refers to the *derefinement* case and can be viewed as a discrete analogue of that in [2, p. 115]. It roughly asserts that regularity (of  $\mathcal{S}$ ) is the sole property that matters for a “discrete” interpolation estimate to hold. The second and more elementary estimate refers to the *refinement* case (5.5).

**Lemma 5.2.** *Let  $\widehat{S} \in \widehat{\mathcal{S}}_1$ . Then*

$$(5.7) \quad \begin{aligned} & \|\Pi\zeta(V) - \widehat{\Pi}\zeta(V)\|_{L^p(\widehat{S})} + h_{\widehat{S}} \|\nabla[\Pi\zeta(V) - \widehat{\Pi}\zeta(V)]\|_{L^p(\widehat{S})} \\ & \leq Ch_{\widehat{S}}^2 \begin{cases} L_\zeta \left( \sum_{e \in \mathcal{E}_{\widehat{S}}} h_e^2 D_e^p \right)^{1/p} + L_{\zeta'} \left( \sum_{S \in \mathcal{S}_{\widehat{S}}} h_S^2 d_S^{2p} \right)^{1/p}, & 1 \leq p < \infty, \\ L_\zeta \max_{e \in \mathcal{E}_{\widehat{S}}} D_e + L_{\zeta'} \max_{S \in \mathcal{S}_{\widehat{S}}} d_S^2, & p = \infty. \end{cases} \end{aligned}$$

**Lemma 5.3.** *Let  $\widehat{S} \in \widehat{\mathcal{F}}_2$ . Then*

$$(5.8) \quad \begin{aligned} & \|\Pi\zeta(V) - \widehat{\Pi}\zeta(V)\|_{L^p(\widehat{S})} \\ & \leq \begin{cases} CL_\zeta h_{\widehat{S}}^{1+2/p} \left( \sum_{e \in \mathcal{E}_{\widehat{S}}} \delta_e^p \right)^{1/p} + CL_{\zeta'} h_{\widehat{S}}^{2/p} \left( \sum_{S \in \mathcal{S}_{\widehat{S}}} \delta_S^{2p} \right)^{1/p}, & 1 \leq p < \infty, \\ CL_\zeta h_{\widehat{S}} \max_{e \in \mathcal{E}_{\widehat{S}}} \delta_e + CL_{\zeta'} \max_{S \in \mathcal{S}_{\widehat{S}}} \delta_S^2, & p = \infty, \end{cases} \end{aligned}$$

and

$$(5.9) \quad \begin{aligned} & \|\nabla[\Pi\zeta(V) - \widehat{\Pi}\zeta(V)]\|_{L^p(\widehat{S})} \\ & \leq \begin{cases} Ch_{\widehat{S}}^{2/p} \left[ L_\zeta \left( \sum_{e \in \mathcal{E}_{\widehat{S}}} \delta_e^p \right)^{1/p} + L_{\zeta'} \left( \sum_{S \in \mathcal{S}_{\widehat{S}}} h_S^p d_S^{2p} \right)^{1/p} \right], & 1 \leq p < \infty, \\ CL_\zeta \max_{e \in \mathcal{E}_{\widehat{S}}} \delta_e + CL_{\zeta'} \max_{S \in \mathcal{S}_{\widehat{S}}} (h_S d_S^2), & p = \infty. \end{cases} \end{aligned}$$

*Remark 5.1.* Consider the simplest case  $\zeta = \text{Identity}$ ; thus  $L_{\zeta'} = 0$ . We point out that we need control on interpolation errors even for the refinement situation, simply because meshes  $\mathcal{S}$  and  $\widehat{\mathcal{S}}$  are not compatible. If they were, these errors would simply vanish.

**5.2. Error estimates.** The first result, whose proof is given in Supplement §S1, deals with the initial triangulation  $\mathcal{S}^1$  and the choice (2.12) for the discrete initial enthalpy  $U^0$ .

**Lemma 5.4.** *We have*

$$\|u_0 - U^0\|_{H^{-1}(\Omega)} \leq C\tau |\log \tau|^{1/2}.$$

We now come to the subtle issue of changing the mesh. Our first task is to apply the two basic interpolation estimates to  $U$ . To this end, let  $V = U$  and  $\zeta$  be as in Lemmas 5.2 and 5.3. Note that the choice  $\zeta = \beta$  is allowed because, in view of (4.8), it is enough to deal with  $\widehat{S} \in \widehat{\mathcal{F}}_0$ . Recall that derivatives of  $U$  in  $\Omega_0$  can be expressed in terms of quantities extracted from  $\Theta$ , as stated in (4.10). Hence, let  $\delta_S$ ,  $d_S$ ,  $\delta_e$ , and  $D_e$  denote from now on derivatives extracted from  $\Theta$ . The following error estimates illustrate the connection between Lemmas 5.2 and 5.3 and the mesh selection algorithm.

**Lemma 5.5.** *The following sharp pointwise error estimate holds:*

$$(5.10) \quad \|\Pi\zeta(U) - \widehat{\Pi}\zeta(U)\|_{L^\infty(\Omega)} \leq C\tau.$$

*Proof.* Let first  $\widehat{S} \in \widehat{\mathcal{F}}_1$ . On using (5.7) and (4.10), in conjunction with (4.2), (4.3) and (4.9), we easily obtain

$$\|\Pi\zeta(U) - \widehat{\Pi}\zeta(U)\|_{L^\infty(\widehat{S})} \leq C \max_{e \in \mathcal{E}_{\widehat{S}}} (h_S^2 D_e) + C \max_{S \in \mathcal{S}_{\widehat{S}}} (h_S^2 d_S^2) \leq C\tau.$$

Likewise, let us take  $\widehat{S} \in \widehat{\mathcal{F}}_2$  and make use of (5.8) and (4.10), together with (4.2), (4.3), (4.6) and (4.9), to arrive at

$$\|\Pi\zeta(U) - \widehat{\Pi}\zeta(U)\|_{L^\infty(\widehat{S})} \leq C \max_{e \in \mathcal{E}_S} (h_S h_e D_e) + C \max_{S \in \mathcal{S}_S} (h_S^2 d_S^2 + h_S h_S d_S^2) \leq C\tau.$$

This concludes the proof because (5.10) is obviously sharp.  $\square$

*Remark 5.2.* Since meshes  $\mathcal{S}$  and  $\widehat{\mathcal{S}}$  are not compatible, we cannot expect a pointwise error estimate for  $\nabla U$  to hold. To make this claim apparent, we consider, for instance, the refinement case (i.e.,  $\widehat{S} \in \widehat{\mathcal{F}}_2$ ) and suppose that  $\zeta := \text{Identity}$  (i.e.,  $L_{\zeta'} = 0$ ). We can write (5.9) as follows:

$$\|\nabla(U - \widehat{U})\|_{L^\infty(\widehat{S})} \leq C \max_{e \in \mathcal{E}_S} h_e D_e,$$

from which we conclude that  $\|\nabla(U - \widehat{U})\|_{L^\infty(\widehat{S})} = O(1)$  provided  $D_e = O(h_e^{-1})$ , as is expected to happen near the interface  $F$ . In any event, setting  $\check{h}_{\widehat{S}} := \min_{e \in \mathcal{E}_S} \check{h}_e$ , we can rewrite (5.9) in the form

$$\check{h}_{\widehat{S}} \|\nabla(U - \widehat{U})\|_{L^\infty(\widehat{S})} \leq C \check{h}_{\widehat{S}} \max_{e \in \mathcal{E}_S} \delta_e \leq C \max_{e \in \mathcal{E}_S} (\check{h}_e h_e D_e) \leq C\tau,$$

as results from (4.2) and (4.6). When  $\check{h}_{\widehat{S}} = O(\tau^{1/2})$ , as happens away from  $F$ , we see that

$$\|\nabla(U - \widehat{U})\|_{L^\infty(\widehat{S})} \leq C\tau^{1/2}.$$

It is worth noting that the critical parameter is  $\check{h}_{\widehat{S}}$  rather than  $h_{\widehat{S}}$  ( $h_{\widehat{S}} \leq \check{h}_{\widehat{S}}!$ ). Similar conclusions hold also for the derefinement case.

In spite of this negative result, we still have an error bound for  $\nabla(U - \widehat{U})$  in energy norm. Under the assumptions (4.11) and (3.15), the following lemma yields  $\|\nabla[\Pi\zeta(U) - \widehat{\Pi}\zeta(U)]\|_{L^2(\Omega)} \leq C\tau^{1/2}$ .

**Lemma 5.6.** *The following sharp  $L^2$ -error estimate holds:*

$$(5.11) \quad \|\nabla[\Pi\zeta(U) - \widehat{\Pi}\zeta(U)]\|_{L^2(\Omega)}^2 \leq C\tau \left( L_\zeta \sum_{e \in \mathcal{E}_0} h_e^2 D_e + L_{\zeta'} \sum_{S \in \mathcal{S}_0} h_S^2 d_S^2 \right).$$

*Proof.* We proceed as in Lemma 5.5 by first examining the derefinement case, namely  $\widehat{S} \in \widehat{\mathcal{F}}_1$ . By virtue of (4.2), (4.3), (4.9) and (4.10), together with (5.7), we easily obtain

$$\begin{aligned} \sum_{\widehat{S} \in \widehat{\mathcal{F}}_1} \|\nabla[\Pi\zeta(U) - \widehat{\Pi}\zeta(U)]\|_{L^2(\widehat{S})}^2 &\leq C \sum_{\widehat{S} \in \widehat{\mathcal{F}}_1} \left( \sum_{e \in \mathcal{E}_S} (h_S^2 D_e) h_e^2 D_e + \sum_{S \in \mathcal{S}_S} (h_S^2 d_S^2) h_S^2 d_S^2 \right) \\ &\leq C\tau \left( \sum_{e \in \mathcal{E}_0} h_e^2 D_e + \sum_{S \in \mathcal{S}_0} h_S^2 d_S^2 \right). \end{aligned}$$

Note that we could replace double summations by single ones because, as a consequence of (5.4),  $\text{card}\{\widehat{S} \in \widehat{\mathcal{F}}_1 : e \cap \widehat{S} \neq \emptyset\} = O(1)$  for all  $e \in \widehat{\mathcal{E}}_0$ . On the other hand, for all  $\widehat{S} \in \widehat{\mathcal{F}}_2$  (refinement case) the above properties, coupled now with (4.6) and (5.9), lead to

$$\begin{aligned} \sum_{\widehat{S} \in \widehat{\mathcal{F}}_2} \|\nabla[\Pi\zeta(U) - \widehat{\Pi}\zeta(U)]\|_{L^2(\widehat{S})}^2 &\leq C \sum_{\widehat{S} \in \widehat{\mathcal{F}}_2} \left( \sum_{e \in \widehat{\mathcal{E}}_S} (h_{\widehat{S}} \delta_e)^2 + \sum_{S \in \mathcal{F}_S} (h_S^2 d_S^2) (h_{\widehat{S}}^2 d_S^2) \right) \\ &\leq C\tau \sum_{\widehat{S} \in \widehat{\mathcal{F}}_2} \left( \sum_{e \in \widehat{\mathcal{E}}_S} (h_{\widehat{S}} h_e D_e) + \sum_{S \in \mathcal{F}_S} h_S^2 d_S^2 \right). \end{aligned}$$

To proceed further, we need the following elementary inequalities:

$$(5.12) \quad \sum_{\widehat{S} \in \widehat{\mathcal{F}}_e^2} h_{\widehat{S}} \leq Ch_e, \quad \forall e \in \mathcal{E}_0, \quad \sum_{\widehat{S} \in \widehat{\mathcal{F}}_S^2} h_{\widehat{S}}^2 \leq Ch_S^2, \quad \forall S \in \mathcal{S}_0,$$

where  $\widehat{\mathcal{F}}_e^2 := \{\widehat{S} \in \widehat{\mathcal{F}}_2 : e \cap \widehat{S} \neq \emptyset\}$  for all  $e \in \mathcal{E}_0$  and  $\widehat{\mathcal{F}}_S^2 := \{\widehat{S} \in \widehat{\mathcal{F}}_2 : S \cap \widehat{S} \neq \emptyset\}$  for all  $S \in \mathcal{S}_0$ . Indeed, this yields

$$\sum_{e \in \mathcal{E}_0} h_e D_e \sum_{\widehat{S} \in \widehat{\mathcal{F}}_e^2} h_{\widehat{S}} \leq C \sum_{e \in \mathcal{E}_0} h_e^2 D_e, \quad \sum_{S \in \mathcal{S}_0} d_S^2 \sum_{\widehat{S} \in \widehat{\mathcal{F}}_S^2} h_{\widehat{S}}^2 \leq C \sum_{S \in \mathcal{S}_0} h_S^2 d_S^2,$$

whence

$$\sum_{\widehat{S} \in \widehat{\mathcal{F}}_2} \|\nabla[\Pi\zeta(U) - \widehat{\Pi}\zeta(U)]\|_{L^2(\widehat{S})}^2 \leq C\tau \left( \sum_{e \in \mathcal{E}_0} h_e^2 D_e + \sum_{S \in \mathcal{S}_0} h_S^2 d_S^2 \right).$$

It only remains to demonstrate (5.12). The second estimate in (5.12) is obvious in view of (5.5), whereas the first one comes from the following consideration.

Let  $\widehat{\mathcal{F}}_e^2 := \{\widehat{S}_k\}_{k=1}^K$  be ordered on  $e$ , and let  $x_k, x_{k-1}$  be the end points of the segment  $e \cap \widehat{S}_k$ . We would like to replace  $h_k := h_{\widehat{S}_k}$  by  $C|x_k - x_{k-1}|$ , but this may not be true for a triangle which is crossed by  $e$  near a vertex. We can however argue as follows to overcome the difficulty. The regularity of  $\widehat{\mathcal{F}}$  yields  $\text{card } \mathcal{J}_k = O(1)$ , where  $\mathcal{J}_k := \{j : \widehat{S}_j \cap B(x_k, ah_k) \neq \emptyset\}$  for  $1 \leq k \leq K$ . Hence,

$$\sum_{k=1}^K h_k \leq C \sum_{k=1}^K \sum_{j \in \mathcal{J}_k} |x_j - x_{j-1}| \leq C \sum_{k=1}^K |x_k - x_{k-1}| \leq Ch_e.$$

Since (5.11) is sharp according to the discrete regularity dealt with, the lemma is thus proved.  $\square$

*Remark 5.3.* In view of the pointwise estimate (5.10) and the a priori discrete  $L^1$ -bound (4.11), the energy error estimate (5.11) for  $\zeta = \text{Identity}$  may be regarded as a 2-D interpolation result, say between  $L^\infty(\Omega)$  and  $W^{2,1}(\Omega)$ .

In §5.3 we will need two trace estimates for the interpolation error  $\Theta - \widehat{\Theta}$  that are stated in Lemmas 5.7 and 5.8 (see Supplement §S1). They correspond to either the derefinement or the refinement situation.

**5.3. A priori estimates.** In this subsection we state a priori estimates related to the crucial bound (4.11) as well as (3.15). Their proofs are reported in Supplement §S1. We begin with (4.11) for the discrete initial temperature  $\Theta^0 := \Pi^1[\beta(U^0)] = \Pi^1\theta_0$ . The symbol  $D_e$  stands here for second derivatives extracted from  $\Theta^0$ .

**Lemma 5.9.** *We have*

$$\|\nabla\Theta^0\|_{L^\infty(\Omega)} + \sum_{e \in \mathcal{S}^1} h_e^2 D_e \leq C.$$

We finally state, in Lemmas 5.10 and 5.11, that (4.11) and (3.15) are inherited by  $\widehat{\mathcal{F}}$  and  $\widehat{\Theta}$ . Let  $D_{\widehat{e}}$  denote the obvious analogue of (4.1) with  $\Theta$  replaced by  $\widehat{\Theta}$ .

**Lemma 5.10.** *The following discrete a priori  $W^{2,1}$ -bound is valid:*

$$(5.13) \quad \sum_{\widehat{e} \in \widehat{\mathcal{E}}} h_{\widehat{e}}^2 D_{\widehat{e}} \leq C \left( L_\beta \sum_{e \in \mathcal{E}} h_e^2 D_e + L_{\beta'} \sum_{S \in \mathcal{S}} h_S^2 d_S^2 \right).$$

Observe that  $\sigma^2 \sum_{S \in \mathcal{S}} h_S^2 d_S^2 \leq \|\nabla\Theta\|_{L^2(\Omega)}^2 \leq \sum_{S \in \mathcal{S}} h_S^2 d_S^2$  is valid as well as the obvious analogue with  $\Theta$  and  $d_S$  replaced by  $\widehat{\Theta}$  and  $d_{\widehat{S}} := |\nabla\widehat{\Theta}|_{\widehat{S}}$ , respectively.

**Lemma 5.11.** *The following discrete a priori  $H^1$ -bound holds:*

$$(5.14) \quad \|\nabla\widehat{\Theta}\|_{L^2(\Omega)}^2 \leq (1 + CL_{\beta'}\tau) \|\nabla\Theta\|_{L^2(\Omega)}^2 + C\tau \sum_{e \in \mathcal{E}_0} h_e^2 D_e.$$

Note that the coefficient in front of  $\|\nabla\Theta\|_{L^2(\Omega)}^2$  becomes 1 whenever  $\beta$  is piecewise linear, i.e.,  $L_{\beta'} = 0$ .

## 6. STABILITY

Our present purpose is to show that the local refinement method (2.13)–(2.15) is stable in various Sobolev norms as soon as the refinement strategy proposed in §4 is enforced. We start by recalling that  $\mathcal{S}^n$  is weakly acute, i.e., the sum of the opposite angles with respect of the common side of any pair of faced triangles does not exceed  $\pi$ . As a consequence, we readily have the following form of the *discrete maximum principle*: let  $\varphi \in \mathbf{V}^n$  attain its maximum at the internal node  $x_j^n$  and let  $\chi_j^n \in \mathbf{V}^n$  be the corresponding basis function; then

$$(6.1) \quad \langle \nabla\varphi, \nabla\chi_j^n \rangle \geq 0.$$

This will serve to exploit monotonicity properties of the problem at hand which in turn compensate for the lack of regularity. Note once more the difference between our approach and that for purely parabolic problems [1, 5]. The first result, proved in Supplement §S2, reads as follows.

**Lemma 6.1.** *The following a priori maximum norm estimate holds:*

$$(6.2) \quad \max_{1 \leq n \leq N} \|U^n\|_{L^\infty(\Omega)} + \max_{1 \leq n \leq N} \|\Theta^n\|_{L^\infty(\Omega)} \leq C.$$

The following well-known result is also a consequence of (6.1) [3]. We include its proof in Supplement §S2, just for the sake of completeness.

**Lemma 6.2.** *Let  $\alpha \in W^{1,\infty}(\mathbf{R})$  satisfy  $\alpha(0) = 0$  and  $0 \leq \alpha'(s) \leq L_\alpha < \infty$  for a.e.  $s \in \mathbf{R}$ . Then*

$$(6.3) \quad \|\nabla \Pi^n \alpha(\varphi)\|_{L^2(\Omega)}^2 \leq L_\alpha \langle \nabla \varphi, \nabla \Pi^n \alpha(\varphi) \rangle, \quad \forall \varphi \in \mathbf{V}^n.$$

In proving the following lemmas, we shall extensively make use of the equivalence of continuous and discrete  $L^2$ -norms for discrete functions, namely

$$\|\varphi\|_{L^2(\Omega)}^2 \leq \langle \varphi, \varphi \rangle^n \leq C \|\varphi\|_{L^2(\Omega)}^2, \quad \forall \varphi \in \mathbf{V}^n.$$

Our next step is to demonstrate a *weak* a priori estimate in energy norm. The first term in (6.4) may be thought of as a discrete  $H^{1/2}(0, T; L^2(\Omega))$ -norm.

**Lemma 6.3.** *The following a priori estimate holds:*

$$(6.4) \quad \sum_{n=1}^N \|U^n - \hat{U}^{n-1}\|_{L^2(\Omega)}^2 + \sum_{n=1}^N \tau \|\nabla \Theta^n\|_{L^2(\Omega)}^2 \leq C.$$

*Proof.* Take  $\chi = \tau U^n \in \mathbf{V}^n$  as a test function in (2.15), and next add the resulting expression over  $n$  for  $1 \leq n \leq m$  ( $\leq N$ ). We have

$$\begin{aligned} \text{I} + \text{II} &:= \sum_{n=1}^m \langle U^n - \hat{U}^{n-1}, U^n \rangle^n + \sum_{n=1}^m \tau \langle \nabla \Theta^n, \nabla U^n \rangle \\ &= \sum_{n=1}^m \tau \langle f(\hat{\Theta}^{n-1}), U^n \rangle^n =: \text{III}. \end{aligned}$$

Using (2.16) together with the elementary identity

$$(6.5) \quad 2a(a-b) = a^2 - b^2 + (a-b)^2, \quad \forall a, b \in \mathbf{R},$$

the first term can be further split as follows ( $\Pi^0 := \Pi^1$ ):

$$\begin{aligned} 2\text{I} &= \sum_{n=1}^m \int_{\Omega} \Pi^n [(U^n)^2] - \Pi^n [(U^{n-1})^2] + \Pi^n [(U^n - U^{n-1})^2] \\ &= \sum_{n=1}^m \int_{\Omega} \Pi^n [(U^n)^2] - \Pi^{n-1} [(U^{n-1})^2] \\ &\quad + \sum_{n=2}^m \int_{\Omega} \Pi^{n-1} [(U^{n-1})^2] - \Pi^n [(U^{n-1})^2] \\ &\quad + \sum_{n=1}^m \int_{\Omega} \Pi^n [(U^n - U^{n-1})^2] =: \text{I}_1 + \text{I}_2 + \text{I}_3. \end{aligned}$$

We now evaluate these three terms separately. We first obtain

$$I_1 = \langle U^m, U^m \rangle^m - \langle U^0, U^0 \rangle^1 \geq \|U^m\|_{L^2(\Omega)}^2 - C\|U^0\|_{L^2(\Omega)}^2,$$

and

$$I_3 = \sum_{n=1}^m \langle U^n - \hat{U}^{n-1}, U^n - \hat{U}^{n-1} \rangle^n \geq \sum_{n=1}^m \|U^n - \hat{U}^{n-1}\|_{L^2(\Omega)}^2.$$

The remaining term, which occurs only when the mesh is changed, can be handled by Lemma 5.5 with  $\zeta(s) = s^2$ . Indeed, since  $U^n$  satisfies (6.2),  $\zeta$  can be suitably modified outside the range of  $U^n$  in such a way that  $\zeta' \in W^{1,\infty}(\mathbf{R})$ . Hence  $I_2 \geq -C \sum_{n=2}^m \tau = -C$ . On the other hand, (2.13) in conjunction with (2.1) and Lemma 6.2 yields

$$\sum_{n=1}^m \tau \|\nabla \Theta^n\|_{L^2(\Omega)}^2 \leq L_\beta \mathbb{II}.$$

For the remaining term III we make use of (6.2) to arrive at

$$|\text{III}| \leq \sum_{n=1}^m \tau \left( |f(0)| \text{meas}(\Omega)^{1/2} + L_f \|\hat{\Theta}^{n-1}\|_{L^2(\Omega)} \right) \|U^n\|_{L^2(\Omega)} \leq C,$$

because  $\hat{\Theta}^{n-1} = \Pi^n[\beta(U^{n-1})]$ . The proof is thus complete.  $\square$

We now derive a *strong* a priori estimate in energy norms. To this end, we need the *structural* assumption (4.11).

**Lemma 6.4.** *Let (4.11) hold. Then*

$$(6.6) \quad \sum_{n=1}^N \tau^{-1} \|\Theta^n - \hat{\Theta}^{n-1}\|_{L^2(\Omega)}^2 + \max_{1 \leq n \leq N} \|\nabla \Theta^n\|_{L^2(\Omega)} \leq C.$$

*Proof.* We argue by induction. Let  $1 =: n_1 < n_2 < \dots < n_K \leq N$  denote the indexes corresponding to the mesh changes; set  $n_{K+1} := N + 1$ . We want to prove the following inequality for all  $m$  between two successive mesh changes, say  $n_k \leq m < n_{k+1}$ :

$$\begin{aligned} A_m &:= (2L_\beta)^{-1} \sum_{n=1}^m \tau^{-1} \|\Theta^n - \hat{\Theta}^{n-1}\|_{L^2(\Omega)}^2 + \max_{0 \leq n \leq m} \|\nabla \Theta^n\|_{L^2(\Omega)}^2 \\ &\leq C_m := \exp(CL_\beta m \tau) \left( C(B_k + 1)m\tau + \|\nabla \Theta^0\|_{L^2(\Omega)}^2 \right), \end{aligned}$$

where  $B_k := \max_{1 \leq i \leq k} \sum_{e \in \mathcal{E}_0^{n_i-1}} h_e^2 D_e^{n_i-1}$ . Since this estimate is obviously valid for  $m = 0$ , assume, by induction, that it holds for all  $m < n_k$  as well.

Take  $\chi = \Theta^n - \widehat{\Theta}^{n-1} \in \mathbf{V}^n$  as a test function in (2.15) and next add the resulting equality from  $n = n_k$  to  $n = m$  ( $< n_{k+1}$ ). We end up with

$$\begin{aligned} \text{I} + \text{II} &:= \sum_{n=n_k}^m \tau^{-1} \langle U^n - \Pi^n U^{n-1}, \Pi^n [\beta(U^n) - \beta(U^{n-1})] \rangle^n \\ &\quad + \sum_{n=n_k}^m \langle \nabla \Theta^n, \nabla(\Theta^n - \widehat{\Theta}^{n-1}) \rangle \\ &= \sum_{n=n_k}^m \langle f(\widehat{\Theta}^{n-1}), \Theta^n - \widehat{\Theta}^{n-1} \rangle^n =: \text{III}. \end{aligned}$$

The first term can be easily evaluated as follows:

$$\begin{aligned} \text{I} &= \sum_{n=n_k}^m \tau^{-1} \int_{\Omega} \Pi^n [(U^n - U^{n-1})(\beta(U^n) - \beta(U^{n-1}))] \\ &\geq L_{\beta}^{-1} \sum_{n=n_k}^m \tau^{-1} \int_{\Omega} \Pi^n [(\beta(U^n) - \beta(U^{n-1}))^2] \\ &\geq L_{\beta}^{-1} \sum_{n=n_k}^m \tau^{-1} \|\Theta^n - \widehat{\Theta}^{n-1}\|_{L^2(\Omega)}^2. \end{aligned}$$

Since  $\widehat{\Theta}^{n-1} = \Theta^{n-1}$  for all  $n_k < n < n_{k+1}$ , as a by-product of (6.5), the next term becomes

$$2\text{II} = \|\nabla \Theta^m\|_{L^2(\Omega)}^2 - \|\nabla \widehat{\Theta}^{n_k-1}\|_{L^2(\Omega)}^2 + \sum_{n=n_k}^m \|\nabla(\Theta^n - \widehat{\Theta}^{n-1})\|_{L^2(\Omega)}^2.$$

By virtue of the induction assumption we can use (5.14) to arrive at

$$\begin{aligned} \|\nabla \widehat{\Theta}^{n_k-1}\|_{L^2(\Omega)}^2 &\leq (1 + CL_{\beta'}\tau) \|\nabla \Theta^{n_k-1}\|_{L^2(\Omega)}^2 + C\tau \sum_{e \in \mathcal{E}_0^{n_k-1}} h_e^2 D_e^{n_k-1} \\ &\leq (1 + CL_{\beta'}\tau) \exp(CL_{\beta'}(n_k - 1)\tau) \left( C(B_{k-1} + 1)(n_k - 1)\tau + \|\nabla \Theta^0\|_{L^2(\Omega)}^2 \right) \\ &\quad + CB_k\tau - ((2L_{\beta})^{-1} + C\tau) \sum_{n=1}^{n_k-1} \tau^{-1} \|\Theta^n - \widehat{\Theta}^{n-1}\|_{L^2(\Omega)}^2, \end{aligned}$$

whence

$$\begin{aligned} &\|\nabla \widehat{\Theta}^{n_k-1}\|_{L^2(\Omega)}^2 + (2L_{\beta})^{-1} \sum_{n=1}^{n_k-1} \tau^{-1} \|\Theta^n - \widehat{\Theta}^{n-1}\|_{L^2(\Omega)}^2 \\ &\leq \exp(CL_{\beta'}n_k\tau) \left( C(B_{k-1} + 1)(n_k - 1)\tau + \|\nabla \Theta^0\|_{L^2(\Omega)}^2 \right) + CB_k\tau, \end{aligned}$$

because  $1 + CL_{\beta}\tau \leq \exp(CL_{\beta}\tau)$ . It only remains to examine term III, for

which we have

$$\begin{aligned} |\text{III}| &\leq \sum_{n=n_k}^m \left( |f(0)| \text{meas}(\Omega)^{1/2} + L_f \|\widehat{\Theta}^{n-1}\|_{L^2(\Omega)} \right) \|\Theta^n - \widehat{\Theta}^{n-1}\|_{L^2(\Omega)} \\ &\leq C(m - n_k + 1)\tau + (2L_\beta)^{-1} \sum_{n=n_k}^m \tau^{-1} \|\Theta^n - \widehat{\Theta}^{n-1}\|_{L^2(\Omega)}^2, \end{aligned}$$

because (6.2) yields  $\|\widehat{\Theta}^{n-1}\|_{L^2(\Omega)} \leq C\|\Theta^{n-1}\|_{L^\infty(\Omega)} \leq C$ . Collecting all partial results, we get the following estimate, for any  $n_k \leq m < n_{k+1}$ ,

$$\begin{aligned} A_m &\leq \exp(CL_{\beta'} n_k \tau) \left( C(B_{k-1} + 1)(n_k - 1)\tau + \|\nabla \Theta^0\|_{L^2(\Omega)}^2 \right) \\ &\quad + CB_k \tau + C(m - n_k + 1)\tau \\ &\leq \exp(CL_{\beta'} m \tau) \left( C(B_k + 1)m\tau + \|\nabla \Theta^0\|_{L^2(\Omega)}^2 \right) = C_m. \end{aligned}$$

The induction argument is thus complete. Finally, the desired estimate follows from Lemma 5.9 and assumption (4.11).  $\square$

*Remark 6.1.* As a by-product of the above argument, combined with (2.1), (2.13) and the fact that  $F^{n-1}, F^n \subset \mathcal{R}^n$  for all  $1 \leq n \leq N$ , we deduce

$$\begin{aligned} C &\geq \sum_{n=1}^N \tau^{-1} \int_{\Omega} \Pi^n \left[ (U^n - U^{n-1})^2 \int_0^1 \beta'(sU^n + (1-s)U^{n-1}) ds \right] \\ &\geq l_\beta \sum_{n=1}^N \tau^{-1} \int_{\Omega \setminus \mathcal{R}^n} \Pi^n [(U^n - U^{n-1})^2], \end{aligned}$$

whence

$$(6.7) \quad \sum_{n=1}^N \tau^{-1} \|U^n - \widehat{U}^{n-1}\|_{L^2(\Omega \setminus \mathcal{R}^n)}^2 \leq C.$$

Our final estimate is a discrete analogue of (2.9):  $u_t \in L^\infty(0, T; M(\Omega))$ .

**Lemma 6.5.** *Let (4.11) hold. Then*

$$(6.8) \quad \max_{1 \leq n \leq N} \|U^n - \widehat{U}^{n-1}\|_{L^1(\Omega)} \leq C\tau.$$

*Proof.* It is enough to prove (6.8) for all steps between two successive mesh changes, say  $1 \leq n_k \leq m < n_{k+1}$ . Moreover, we can assume without loss of generality that  $\beta$  is strictly increasing because the asymptotic constant in (6.8) is independent of the lower bound of  $\beta'$ . As a result,  $\beta^{-1}$  is well defined and monotone increasing. Subtract now two consecutive discrete equations (2.15) for  $n_k \leq n < n_{k+1}$  to arrive at

$$(6.9) \quad \langle \partial U^n - \partial U^{n-1}, \chi \rangle^{n_k} + \tau \langle \nabla \partial \Theta^n, \nabla \chi \rangle = \tau \langle \partial f(\Theta^{n-1}), \chi \rangle^{n_k},$$

where  $\partial z^n := \tau^{-1}(z^n - z^{n-1})$  and  $\Theta^{n_k-2} := \Theta^{n_k-1}$ . For notational convenience, we have used  $U^{n_k-1}$  to denote  $\hat{U}^{n_k-1} = \Pi^{n_k}(U^{n_k-1}) \in \mathbf{V}^{n_k}$ ,  $\Theta^{n_k-1}$  to designate  $\hat{\Theta}^{n_k-1} = \Pi^{n_k} \beta(U^{n_k-1}) \in \mathbf{V}^{n_k}$  and, finally,  $U^{n_k-2} \in \mathbf{V}^{n_k}$  to indicate the solution of the following auxiliary problem:

$$\tau^{-1} \langle U^{n_k-1} - U^{n_k-2}, \chi \rangle^{n_k} + \langle \nabla \Theta^{n_k-1}, \nabla \chi \rangle = \langle f(\Theta^{n_k-1}), \chi \rangle^{n_k}, \quad \forall \chi \in \mathbf{V}^{n_k}.$$

By virtue of (2.15), (2.18) and (6.2), we readily have for all  $\chi \in \mathbf{V}^{n_k}$ ,  $\|\chi\|_{L^\infty(\Omega)} \leq 1$ ,

$$\begin{aligned} |\langle \partial U^{n_k-1}, \chi \rangle^{n_k}| &\leq |\langle f(\Theta^{n_k-1}), \chi \rangle^{n_k}| + \sum_{e \in \mathcal{E}^{n_k}} |\langle \llbracket \nabla \Theta^{n_k-1} \rrbracket_e \cdot \nu_e, \chi \rangle_e| \\ &\leq C \left( 1 + \sum_{e \in \mathcal{E}^{n_k}} h_e^2 D_e^{n_k-1} \right). \end{aligned}$$

Let  $\psi_\varepsilon \in C^\infty(\mathbf{R})$  be a monotone approximation of the function  $\text{sgn}$  such that  $\psi_\varepsilon(0) = 0$ ,  $|\psi_\varepsilon(s)| \leq 1$  and  $\psi_\varepsilon(s) \rightarrow \text{sgn}(s)$  as  $\varepsilon \downarrow 0$  for all  $s \in \mathbf{R} \setminus \{0\}$ . Set  $\chi := \Pi^{n_k}[\psi_\varepsilon(\partial U^{n_k-1})]$  in the above inequality and let  $\varepsilon \downarrow 0$ . Since  $\Theta^{n_k-1}$  is either the initial discrete temperature or that corresponding to a mesh change, it satisfies either Lemma 5.9 or Lemma 5.10. We infer that

$$(6.10) \quad \int_{\Omega} \Pi^{n_k} |\partial U^{n_k-1}| \leq C,$$

because of (4.11) and (3.15). To proceed further, take  $\chi := \Pi^{n_k}[\psi_\varepsilon(\partial \Theta^n)] \in \mathbf{V}^{n_k}$  as a test function in (6.9) and next add the resulting expressions from  $n_k$  to  $m < n_{k+1}$ . We obtain

$$\begin{aligned} \text{I} + \text{II} &:= \sum_{n=n_k}^m \langle \partial(U^n - U^{n-1}), \Pi^{n_k}[\psi_\varepsilon(\partial \Theta^n)] \rangle^{n_k} \\ &\quad + \sum_{n=n_k}^m \tau \langle \nabla \partial \Theta^n, \nabla \Pi^{n_k}[\psi_\varepsilon(\partial \Theta^n)] \rangle \\ &= \sum_{n=n_k}^m \tau \langle \partial f(\Theta^{n-1}), \Pi^{n_k}[\psi_\varepsilon(\partial \Theta^n)] \rangle^{n_k} =: \text{III}, \end{aligned}$$

which is now examined in detail. By virtue of (6.3) we have  $\text{II} \geq 0$ . For  $\text{III}$  we make use of (6.6) to arrive at

$$|\text{III}| \leq C \sum_{n=n_k}^{m-1} \tau L_f \|\partial \Theta^n\|_{L^2(\Omega)} \leq C.$$

For the remaining term  $\text{I}$  we reason as follows. We first observe that

$$0 \leq \partial U^{n-1}(x_j^{n_k}) \psi_\varepsilon(\partial \Theta^n(x_j^{n_k})) \leq |\partial U^{n-1}(x_j^{n_k})|,$$

for all nodes  $x_j^{n_k}$  of  $\mathcal{S}^{n_k}$ . Secondly, we take the limit as  $\varepsilon \downarrow 0$  to obtain

$$\begin{aligned} \partial U^n(x_j^{n_k}) \psi_\varepsilon(\partial \Theta^n(x_j^{n_k})) &\rightarrow \partial U^n(x_j^{n_k}) \text{sgn}(\partial \Theta^n(x_j^{n_k})) \\ &= \partial U^n(x_j^{n_k}) \text{sgn}(\partial U^n(x_j^{n_k})) = |\partial U^n(x_j^{n_k})|, \end{aligned}$$

for all  $x_j^{n_k}$  such that  $\partial U^n(x_j^{n_k}) \neq 0$  or, equivalently,  $\partial \Theta^n(x_j^{n_k}) \neq 0$ . Note that it is precisely here where we need  $\beta$  to be strictly increasing. Therefore, term I becomes

$$\sum_{n=n_k}^m \int_{\Omega} \Pi^{n_k} |\partial U^n| - \int_{\Omega} \Pi^{n_k} |\partial U^{n-1}| = \int_{\Omega} \Pi^{n_k} |\partial U^m| - \int_{\Omega} \Pi^{n_k} |\partial U^{n_k-1}|.$$

Collecting the previous estimates and making use of (6.10), combined with the fact that  $\|\chi\|_{L^1(\Omega)} \leq \int_{\Omega} \Pi^{n_k} |\chi|$  for all  $\chi \in \mathbf{V}^{n_k}$ , the assertion easily follows with an asymptotic constant independent of the lower bound of  $\beta'$ .  $\square$

## 7. ENERGY ERROR ESTIMATES

In this section we derive error estimates in the natural energy spaces for our Adaptive Method. We also state a quasi-optimal error estimate in the maximum norm for an auxiliary elliptic problem which may have some independent interest. Its proof is reported in Supplement §S3. We stress once again that technical arguments to handle the discrepancy between continuous and discrete domains were introduced in full detail in [17] and are thus omitted here by simply assuming  $\Omega^n = \bar{\Omega}$ .

We start by recalling a well-known interpolation estimate for the quadrature formula (2.16), namely, for all  $S \in \mathcal{S}^n$  we have

$$(7.1) \quad \int_S |\chi\varphi - \Pi^n(\chi\varphi)| dx \leq Ch_S \|\chi\|_{L^2(S)} \|\nabla\varphi\|_{L^2(S)}, \quad \forall \chi, \varphi \in \mathbf{V}^n.$$

Let  $G: H^{-1}(\Omega) \rightarrow H_0^1(\Omega)$  designate the Green's operator, that is,

$$(7.2) \quad \langle \nabla G\psi, \nabla\varphi \rangle = \langle \psi, \varphi \rangle, \quad \forall \varphi \in H_0^1(\Omega).$$

Since  $\Omega$  is smooth, the operator  $G$  is *regular* [2, p. 138]:

$$(7.3) \quad \|G\psi\|_{H^2(\Omega)} \leq C\|\psi\|_{L^2(\Omega)}, \quad \forall \psi \in L^2(\Omega).$$

In addition, the norm in  $H^{-1}(\Omega)$  can be represented in terms of  $G$  as follows:

$$(7.4) \quad \|\psi\|_{H^{-1}(\Omega)} = \|\nabla G\psi\|_{L^2(\Omega)} = \langle \psi, G\psi \rangle^{1/2}, \quad \forall \psi \in H^{-1}(\Omega).$$

The discrete Green's operator  $G^n: H^{-1}(\Omega) \rightarrow \mathbf{V}^n$  is defined by

$$(7.5) \quad \langle \nabla G^n\psi, \nabla\chi \rangle = \langle \psi, \chi \rangle, \quad \forall \chi \in \mathbf{V}^n,$$

and satisfies the following error estimates. Set  $h_n := \max_{S \in \mathcal{S}^n} h_S$ ,  $\rho_n := \min_{S \in \mathcal{S}^n} h_S$  and recall that  $\lambda\tau \leq \rho_n < h_n \leq \Lambda\tau^{1/2}$ , as results from (4.9). We have first

$$(7.6) \quad \|(G - G^n)\psi\|_{L^2(\Omega)} \leq Ch_n^2 \|\psi\|_{L^2(\Omega)}, \quad \forall \psi \in L^2(\Omega),$$

under the sole assumption of regularity of both  $\mathcal{S}^n$  and  $G$  [2]. If, in addition,  $h_n \leq C\rho_n^\gamma$  with  $0 < \gamma \leq 1$ , we will prove in §S3.2 the following quasi-optimal pointwise estimate; in the present case  $\gamma = 1/2$ .

**Lemma 7.1.** *There holds  $\|(G - G^n)\psi\|_{L^\infty(\Omega)} \leq Ch_n^2 |\log h_n|^7 \|\psi\|_{L^\infty(\Omega)}$ , for all  $\psi \in L^\infty(\Omega)$ .*

Its proof is based on having suitable local energy estimates as in [21, 22, 23]. However, the novelty here is that  $\mathcal{S}^n$  is not quasi-uniform nor does it have any a priori structure as in [22]. Such a new difficulty is responsible for the extra power of the logarithm.

We now introduce the *error equations*. Integration of (2.8) on the interval  $I^n := (t^{n-1}, t^n]$  results in

$$(7.7) \quad \begin{aligned} \langle u^n - u^{n-1}, \varphi \rangle + \left\langle \int_{I^n} \nabla \theta(t) dt, \nabla \varphi \right\rangle \\ = \left\langle \int_{I^n} f(\theta(t)) dt, \varphi \right\rangle, \quad \forall \varphi \in H_0^1(\Omega), \quad 1 \leq n \leq N, \end{aligned}$$

where  $t^n := n\tau$  and  $u^n := u(t^n)$ . At the same time, (2.15) can be written equivalently as

$$(7.8) \quad \begin{aligned} \langle U^n - U^{n-1}, \varphi \rangle + \tau \langle \nabla \Theta^n, \nabla \chi \rangle \\ = \langle U^n - \widehat{U}^{n-1}, \chi \rangle - \langle U^n - \widehat{U}^{n-1}, \chi \rangle^n \\ + \langle U^n - \widehat{U}^{n-1}, \varphi - \chi \rangle + \langle \widehat{U}^{n-1} - U^{n-1}, \varphi \rangle \\ + \tau \langle f(\widehat{\Theta}^{n-1}), \chi \rangle^n, \quad \forall \varphi \in H_0^1(\Omega), \quad \chi \in \mathbf{V}^n, \quad 1 \leq n \leq N. \end{aligned}$$

Subtraction of (7.8) from (7.7) yields the error equation

$$(7.9) \quad \begin{aligned} \langle e_u^n - e_u^{n-1}, \varphi \rangle + \left\langle \int_{I^n} \nabla e_\theta(t) dt, \nabla \varphi \right\rangle + \tau \langle \nabla \Theta^n, \nabla(\varphi - \chi) \rangle \\ = \langle U^n - \widehat{U}^{n-1}, \chi \rangle^n - \langle U^n - \widehat{U}^{n-1}, \chi \rangle + \langle \widehat{U}^{n-1} - U^n, \varphi - \chi \rangle \\ + \langle U^{n-1} - \widehat{U}^{n-1}, \varphi \rangle + \left\langle \int_{I^n} f(\theta(t)) dt, \varphi \right\rangle - \tau \langle f(\widehat{\Theta}^{n-1}), \chi \rangle^n, \end{aligned}$$

for all  $\varphi \in H_0^1(\Omega)$ ,  $\chi \in \mathbf{V}^n$ , where we have set, for any  $1 \leq n \leq N$ ,

$$e_u(t) := u(t) - U^n, \quad e_\theta(t) := \theta(t) - \Theta^n, \quad \forall t \in I^n, \quad \text{and} \quad e_u^n := e_u(t^n).$$

**Theorem 7.1.** *Let (4.11) hold and the number of mesh changes be bounded above by  $O(\tau^{-1/2})$ . Then*

$$(7.10) \quad \|e_u\|_{L^\infty(0, T; H^{-1}(\Omega))} + \|e_\theta\|_{L^2(0, T; L^2(\Omega))} \leq C\tau^{1/2} |\log \tau|^{7/2}.$$

For the practical range of time steps  $\tau$ , the above rate is essentially  $O(\tau^{1/2})$ . The restriction on the number of mesh changes accounts for the accumulation of interpolation errors  $U^{n-1} - \widehat{U}^{n-1}$  which, by Lemma 5.5, are  $O(\tau)$ . The mesh selection algorithm of §4.2 is so designed as to make such a restriction and (4.11) acceptable constraints. The key issue in Theorem 7.1 is the underlying set of graded meshes  $\{\mathcal{S}^n\}_{n=1}^N$  for which it holds and constitutes the first rigorous result. In fact, similar rates of convergence have been obtained for

quasi-uniform meshes [4, 8, 17, 24]. The improvement upon those results is thus to be expressed in terms of spatial degrees of freedom, as explained in §8.2.

*Proof of Theorem 7.1.* Take first  $\varphi := Ge_u^n \in H_0^1(\Omega)$  and  $\chi := G^n e_u^n \in \mathbf{V}^n$  in (7.9), next add over  $n$  from 1 to  $m$  ( $\leq N$ ) and use (7.2) and (7.5) to arrive at

$$\begin{aligned}
& \sum_{n=1}^m \langle e_u^n - e_u^{n-1}, Ge_u^n \rangle + \sum_{n=1}^m \left\langle \int_{I^n} e_\theta(t) dt, e_u^n \right\rangle =: \text{I} + \text{II} \\
& = \sum_{n=1}^m [\langle U^n - \widehat{U}^{n-1}, G^n e_u^n \rangle^n - \langle U^n - \widehat{U}^{n-1}, G^n e_u^n \rangle] \\
(7.11) \quad & + \sum_{n=1}^m \langle \widehat{U}^{n-1} - U^n, (G - G^n) e_u^n \rangle + \sum_{n=1}^m \langle U^{n-1} - \widehat{U}^{n-1}, Ge_u^n \rangle \\
& + \sum_{n=1}^m \left( \left\langle \int_{I^n} f(\theta(t)) dt, Ge_u^n \right\rangle - \tau \langle f(\widehat{\Theta}^{n-1}), G^n e_u^n \rangle^n \right) \\
& =: \text{III} + \dots + \text{VI}.
\end{aligned}$$

The rest of the proof consists simply of evaluating these six terms separately. In order to simplify notation, set

$$A_m := \max_{1 \leq n \leq m} \|e_u^n\|_{H^{-1}(\Omega)}^2.$$

By virtue of Lemma 5.4, (6.5) and (7.4), the first term yields

$$2\text{I} = \|e_u^m\|_{H^{-1}(\Omega)}^2 - \|u_0 - U^0\|_{H^{-1}(\Omega)}^2 + \sum_{n=1}^m \|e_u^n - e_u^{n-1}\|_{H^{-1}(\Omega)}^2 \geq \|e_u^m\|_{H^{-1}(\Omega)}^2 - C\tau.$$

In view of the constitutive relations (2.6) and (2.13), term II can be further split as follows:

$$\begin{aligned}
\text{II}_1 + \text{II}_2 + \text{II}_3 & := \sum_{n=1}^m \int_{I^n} \langle e_\theta(t), u(t^n) - u(t) \rangle dt \\
& + \sum_{n=1}^m \int_{I^n} \langle \beta(u(t)) - \beta(U^n), u(t) - U^n \rangle dt \\
& + \sum_{n=1}^m \int_{I^n} \langle \beta(U^n) - \Pi^n \beta(U^n), e_u(t) \rangle dt.
\end{aligned}$$

We first make use of (2.5) and (6.4) to evaluate term  $\text{II}_1$  as follows:

$$|\text{II}_1| \leq \sum_{n=1}^m \int_{I^n} \|\nabla e_\theta(t)\|_{L^2(\Omega)} \left\| \int_t^{t^n} u_t(s) ds \right\|_{H^{-1}(\Omega)} dt \leq C\tau.$$

We next recall an elementary interpolation estimate. Let  $\alpha : \mathbf{R} \rightarrow \mathbf{R}$  be a continuous and nondecreasing function; then

$$\begin{aligned}
(7.12) \quad & \|\alpha(\varphi) - \Pi^n \alpha(\varphi)\|_{L^p(S)} \\
& \leq Ch_S \|\nabla \Pi^n \alpha(\varphi)\|_{L^p(S)}, \quad \forall S \in \mathcal{S}^n, \varphi \in \mathbf{V}^n, 1 \leq p \leq \infty.
\end{aligned}$$

Applying this inequality with  $\alpha = \beta$  and  $p = 2$ , in conjunction with (6.4) and the constraint  $h_S \leq \Lambda\tau^{1/2}$  for all  $S \in \mathcal{S}^n$ , we obtain

$$\begin{aligned} \mathbb{II}_2 &\geq (2L_\beta)^{-1} \sum_{n=1}^m \int_{I^n} \|e_\theta(t)\|_{L^2(\Omega)}^2 dt - L_\beta^{-1} \tau \sum_{n=1}^m \|\beta(U^n) - \Pi^n \beta(U^n)\|_{L^2(\Omega)}^2 \\ &\geq (2L_\beta)^{-1} \|e_\theta\|_{L^2(0, t^m; L^2(\Omega))}^2 - C\tau. \end{aligned}$$

The analysis of term  $\mathbb{II}_3$  will be split according to the local meshsize near the discrete interface  $F^n$  and far from it. Consider first  $S \in \mathcal{S}_F^n = \{S \in \mathcal{S}^n : S \cap F^n \neq \emptyset\}$ , for which  $h_S = O(\tau)$ . The inequality (7.12) just used, now with  $p = 1$ , leads also to

$$\|\beta(U^n) - \Pi^n \beta(U^n)\|_{L^1(S)} \leq C\tau \|\nabla \Theta^n\|_{L^1(S)}.$$

On the other hand, if  $S \in \mathcal{S}^n \setminus \mathcal{S}_F^n$  we can exploit the further regularity of  $\beta$ , together with (4.10) and the constraint  $h_S \leq \Lambda\tau^{1/2}$  for all  $S \in \mathcal{S}^n$ , to arrive at

$$\begin{aligned} \|\beta(U^n) - \Pi^n \beta(U^n)\|_{L^1(S)} &\leq Ch_S^2 \|D^2 \beta(U^n)\|_{L^1(S)} \\ &\leq CL_{\beta'} \tau \|\nabla U^n\|_{L^2(S)}^2 \leq CL_{\beta'} l_\beta^{-1} \tau \|\nabla \Theta^n\|_{L^2(S)}^2, \end{aligned}$$

by virtue of a standard interpolation estimate. Hence, from (2.5), (6.2) and (6.4) we readily have

$$\begin{aligned} |\mathbb{II}_3| &\leq \sum_{n=1}^m \left( \sum_{S \in \mathcal{S}^n} \|\beta(U^n) - \Pi^n \beta(U^n)\|_{L^1(S)} \right) \int_{I^n} \|e_u(t)\|_{L^\infty(\Omega)} dt \\ &\leq C\tau \sum_{n=1}^m \tau (\|\nabla \Theta^n\|_{L^1(\Omega)} + \|\nabla \Theta^n\|_{L^2(\Omega)}^2) \leq C\tau. \end{aligned}$$

In summary, for term II we have obtained the lower bound

$$\mathbb{II} \geq (2L_\beta)^{-1} \|e_\theta\|_{L^2(0, t^m; L^2(\Omega))}^2 - C\tau.$$

The analysis of term III on the right-hand side of (7.11) will be also split according to what happens within the refined region  $\mathcal{R}^n$  (defined in §2) and out of it. More precisely, we decompose the integrals in III over all triangles  $S \in \mathcal{S}_R^n := \{S \in \mathcal{S}^n : S \subset \mathcal{R}^n\}$ , where  $h_S = O(\tau)$ , and  $S \in \mathcal{S}^n \setminus \mathcal{S}_R^n$ , where  $h_S \leq \Lambda\tau^{1/2}$ . Inequality (7.1) then yields

$$\begin{aligned} |\mathbb{III}| &\leq \sum_{n=1}^m \sum_{S \in \mathcal{S}^n} \int_S |(U^n - \hat{U}^{n-1})G^n e_u^n - \Pi^n [(U^n - \hat{U}^{n-1})G^n e_u^n]| dx \\ &\leq C \sum_{n=1}^m (\tau \|U^n - \hat{U}^{n-1}\|_{L^2(\Omega)} + \tau^{1/2} \|U^n - \hat{U}^{n-1}\|_{L^2(\Omega \setminus \mathcal{R}^n)}) \|\nabla G^n e_u^n\|_{L^2(\Omega)}. \end{aligned}$$

Hence, applying (7.4) and the property  $\|\nabla G^n \psi\|_{L^2(\Omega)} \leq \|\nabla G \psi\|_{L^2(\Omega)}$ , together with (6.4) and (6.7), we get

$$\begin{aligned} |\text{III}| &\leq \eta A_m + C \eta^{-1} \tau \left( \sum_{n=1}^m \|U^n - \widehat{U}^{n-1}\|_{L^2(\Omega)}^2 + \sum_{n=1}^m \tau^{-1} \|U^n - \widehat{U}^{n-1}\|_{L^2(\Omega \setminus \mathcal{A}^n)}^2 \right) \\ &\leq \eta A_m + C \eta^{-1} \tau, \end{aligned}$$

where  $\eta > 0$  is to be selected. The contribution due to IV can be handled by means of (2.5), (6.2), (6.8) and Lemma 7.1 as follows:

$$|\text{IV}| \leq \sum_{n=1}^m \|U^n - \widehat{U}^{n-1}\|_{L^1(\Omega)} \|(G - G^n)e_u^n\|_{L^\infty(\Omega)} \leq C \tau |\log \tau|^7.$$

Term V is easy to evaluate in view of (5.10) and the fact that  $U^{n-1} \neq \widehat{U}^{n-1}$  only whenever the mesh  $\mathcal{S}^{n-1}$  changes. In fact, we have

$$\begin{aligned} |\text{V}| &\leq \sum_{n=1}^m \|U^{n-1} - \widehat{U}^{n-1}\|_{L^2(\Omega)} \|e_u^n\|_{H^{-1}(\Omega)} \\ &\leq A_m^{1/2} \sum_{n=1}^m \|U^{n-1} - \widehat{U}^{n-1}\|_{L^2(\Omega)} \leq C \tau^{1/2} A_m^{1/2} \leq \eta A_m + C \eta^{-1} \tau. \end{aligned}$$

It only remains to estimate the contribution due to the source  $f$ . To evaluate VI, we decompose it further as follows:

$$\begin{aligned} \text{VI} &= \sum_{n=1}^m \left\langle \int_{I^n} f(\theta(t)) dt, (G - G^n)e_u^n \right\rangle \\ &\quad + \sum_{n=1}^m \left\langle \int_{I^n} [f(\theta(t)) - f(\Theta^n)] dt, G^n e_u^n \right\rangle + \sum_{n=1}^m \tau (f(\Theta^n) - f(\widehat{\Theta}^{n-1}), G^n e_u^n) \\ &\quad + \sum_{n=1}^m \tau \left( \langle f(\widehat{\Theta}^{n-1}), G^n e_u^n \rangle - \langle f(\widehat{\Theta}^{n-1}), G^n e_u^n \rangle^n \right) =: \text{VI}_1 + \dots + \text{VI}_4. \end{aligned}$$

We have first

$$|\text{VI}_1| \leq C \|f(\theta)\|_{L^\infty(0, T; L^2(\Omega))} \max_{1 \leq n \leq m} \|(G - G^n)e_u^n\|_{L^2(\Omega)} \leq C \tau,$$

as results from (2.4), (2.5), (6.2), and (7.6), as well as

$$\begin{aligned} |\text{VI}_2| &\leq C \tau^{1/2} \sum_{n=1}^m \|e_\theta\|_{L^2(I^n; L^2(\Omega))} \|e_u^n\|_{H^{-1}(\Omega)} \\ &\leq \eta \|e_\theta\|_{L^2(0, T^m; L^2(\Omega))}^2 + C \eta^{-1} \sum_{n=1}^m \tau \|e_u^n\|_{H^{-1}(\Omega)}^2. \end{aligned}$$

We next use (6.6) together with (2.4) to obtain

$$|\text{VI}_3| \leq C \tau \sum_{n=1}^m \|\Theta^n - \widehat{\Theta}^{n-1}\|_{L^2(\Omega)} \|e_u^n\|_{H^{-1}(\Omega)} \leq \eta A_m + C \tau^2.$$

We then decompose  $\text{VI}_4$  further and apply (7.1) in conjunction with (5.14) to arrive at

$$\begin{aligned} |\text{VI}_4| &\leq \tau \left| \sum_{n=1}^m \langle f(\widehat{\Theta}^{n-1}) - \Pi^n f(\widehat{\Theta}^{n-1}), G^n e_u^n \rangle \right| \\ &\quad + \tau \left| \sum_{n=1}^m \langle \Pi^n f(\widehat{\Theta}^{n-1}), G^n e_u^n \rangle - \langle \Pi^n f(\widehat{\Theta}^{n-1}), G^n e_u^n \rangle^n \right| \\ &\leq \tau^{3/2} \sum_{n=1}^m \|\nabla \widehat{\Theta}^{n-1}\|_{L^2(\Omega)} \|e_u^n\|_{H^{-1}(\Omega)} \leq \eta A_m + C\tau. \end{aligned}$$

A proper choice of  $\eta$  finally allows terms  $A_m$  and  $\|e_\theta\|_{L^2(0, t^m; L^2(\Omega))}^2$  to be absorbed into the left-hand side of (7.11). Therefore, for all  $1 \leq m \leq N$ , we have obtained the estimate

$$\max_{1 \leq n \leq m} \|e_u^n\|_{H^{-1}(\Omega)}^2 + \|e_\theta\|_{L^2(0, t^m; L^2(\Omega))}^2 \leq C\tau |\log \tau|^7 + C \sum_{n=1}^m \tau \|e_u^n\|_{H^{-1}(\Omega)}^2.$$

The rightmost term can be removed after applying the discrete Gronwall lemma. The desired estimate then follows from the property  $u \in H^1(0, T; H^{-1}(\Omega)) \subset C^{0,1/2}(0, T; H^{-1}(\Omega))$ , which can be used to replace  $\max_{1 \leq n \leq m} \|e_u^n\|_{H^{-1}(\Omega)}$  by  $\|e_u\|_{L^\infty(0, T; H^{-1}(\Omega))}$ . The theorem is thus proved.  $\square$

## 8. COMPUTATIONAL ISSUES

To conclude this paper, we present several numerical experiments and compare the proposed Adaptive Method (AM) with the Fixed Mesh Method (FMM). We also comment on some crucial computational issues. Full implementation details as well as many other relevant numerical tests will appear in Part II [16].

**8.1. Implementation.** Let  $\mathbf{M}^n := (\langle \chi_i^n, \chi_j^n \rangle)_{i,j=1}^{J^n}$  and  $\mathbf{K}^n := (\langle \nabla \chi_i^n, \nabla \chi_j^n \rangle)_{i,j=1}^{J^n}$  denote the mass matrix and the stiffness matrix, respectively. The equation (2.15) can then be written in matrix form as follows:

$$(8.1) \quad \mathbf{M}^n U^n + \tau \mathbf{K}^n \Theta^n = \mathbf{M}^n (\widehat{U}^{n-1} + \tau f(\widehat{\Theta}^{n-1})),$$

where we have identified piecewise linear functions with the vector of their nodal values. Since  $\Theta^n = \Pi^n[\beta(U^n)]$ , the algebraic system (8.1) is (strongly) nonlinear. However, as  $\mathbf{M}^n$  is diagonal, (8.1) can be easily and efficiently solved by a nonlinear SOR method that is known to converge; see [16, 17, 18]. Based on a linear majorant, an approximate optimal relaxation parameter can be determined in advance so as to accelerate the convergence of this iterative method [16].

In order to implement (4.9) efficiently, an auxiliary (or superposed) uniform square mesh  $\mathcal{Q}$  of size  $\rho := O(\tau^{1/2})$  is used. Such a mesh is created and kept fixed from the beginning in such a way that  $\bar{\Omega} \subset \bigcup_{R \in \mathcal{Q}} R$ . We associate with

each square  $R$  a value of desired meshsize, defined as the minimum of the local parameters of §4.1 and §S4.2. This leads to the following piecewise constant function:

$$(8.2) \quad \check{\mathbf{h}}^n|_R := \min_{\substack{S' \in \mathcal{S}_F^n : \mathcal{E}_{S'} \cap R \neq \emptyset \\ S \in \mathcal{S}_0^n, e \in \mathcal{E}_0^n : S \cap R \neq \emptyset, e \subset \partial S}} (\Lambda\tau^{1/2}, \hat{h}_{F_{S'}}, \hat{h}_e, \hat{h}_S), \quad \forall R \in \mathcal{Q}.$$

Such a function may be very oscillatory and thus not appropriate as a meshsize indicator. An efficient smoothing postprocess then constructs a piecewise constant function  $\mathbf{h}^n$  satisfying the following *compatibility* property for all  $R \in \mathcal{Q}$ :

$$(8.3) \quad \check{\mathbf{h}}^n|_P \geq \mathbf{h}^n|_R, \quad \forall P \in \mathcal{Q} : \text{dist}(P, R) \leq \mathbf{h}^n|_R.$$

This is achieved, for instance, by defining  $\mathbf{h}^n$  to be

$$(8.4) \quad \mathbf{h}^n|_R := \min_{P \in \mathcal{Q}} (\max(\check{\mathbf{h}}^n|_P, \text{dist}(P, R) - \rho)), \quad \forall R \in \mathcal{Q}.$$

The function  $\mathbf{h}^n$  is then used by the automatic mesh generator of [19] to produce an admissible triangulation. In fact, by virtue of (8.3), we have  $\lambda\tau \leq h_S \leq \check{\mathbf{h}}^n(x) \leq \Lambda\tau^{1/2}$  for all  $S \in \mathcal{S}^n$  and  $x \in S$ , which in turn implies (4.9). This topic is further discussed in [16].

Since several remeshing operations are to be performed during the solution process, mesh generator efficiency is a crucial issue. We used the mesh generator ADVFRONT of [19], which produces weakly acute meshes for general planar domains with quasi-optimal computational complexity  $O(S^n \log S^n)$ , where  $S^n := \text{card}(\mathcal{S}^n)$ . This is so because binary search techniques are employed on suitable quadtree structured data to update the advancing front. Similar ideas were applied to determine  $\hat{U}^{n-1}$ , the interpolant of  $U^{n-1}$  in  $\mathcal{S}^n$ . More specifically, this crucial interpolation process requires a computational labor of  $O((S^{n-1} + J^n) \log J^n)$ , which is also nearly optimal; recall that  $J^n$  is the number of nodes of  $\mathcal{S}^n$ . Such techniques are essential for the Adaptive Method to be competitive. See [16] for more details.

**8.2. Degrees of freedom.** Suppose that there is no discrete mushy region and any discrete interface  $F^n$  is a polygonal curve with finite length uniformly in  $\tau$ , which in turn is consistent with (3.13). Then the refined region  $\mathcal{R}^n$  is just a strip  $O(\tau^{1/2})$ -wide around  $F^n$ . Since the local meshsize is  $O(\tau)$ , the number of triangles within  $\mathcal{R}^n$  is  $O(\tau^{-3/2})$ . Except possibly for a small transition region, triangles outside of  $\mathcal{R}^n$  are  $O(\tau^{1/2})$ . Consequently, the number of triangles outside of  $\mathcal{R}^n$  becomes  $O(\tau^{-1})$ ; so the required computational labor in  $\mathcal{R}^n$  dominates! Hence, the number of spatial *degrees of freedom* (DOF) for every mesh  $\mathcal{S}^n$  is

$$(8.5) \quad \text{DOF} = O(\tau^{-3/2}),$$

for a global accuracy  $O(\tau^{1/2})$ . This quantity compares quite favorably with similar ones for practical methods involving a single quasi-uniform mesh and is

reflected in Tables 8.1–8.3. In fact, in that case  $h_S = O(\tau)$  for all  $S \in \mathcal{S}$  and so  $\text{DOF} = O(\tau^{-2})$  for the same global accuracy [4, 17, 24]. We also stressed that even with mushy regions, we need fewer DOF. A relevant example is fully discussed in [16].

Consider now the Fixed Mesh Method with preliminary regularization [8, 12, 13, 17]. Assuming that a global nondegeneracy property is valid, a result similar to (8.5) can be obtained [17], but at the expense of a much worse local approximation quality, as numerical evidence indicates [16, 17, 18]. Such a local property is extremely important in determining interfaces and explains the much better resolution associated with the present method. We also underline that the example below does not satisfy a global nondegeneracy property.

**8.3. Numerical experiments.** To illustrate the superior performance of our Adaptive Method with respect to fixed mesh techniques, we have chosen the severe test below. It is a classical two-phase Stefan problem with an interface that moves up and down. The exact temperature is given by the following expression:

$$(8.6) \quad \theta(x, y, t) := \begin{cases} 0.75(r^2 - 1), & r < 1, \\ (1.5 - \alpha'(t) \sin \varphi)(r - 1), & r \geq 1, \end{cases}$$

where  $r := (x^2 + (y - \alpha(t))^2)^{1/2}$ ,  $\alpha(t) := 0.5 + \sin(1.25t)$ ,  $\sin \varphi := (y - \alpha(t))/r$ ,  $\Omega := (0, 5) \times (0, 5)$  and  $T := \pi/1.25$ . Dirichlet boundary conditions are imposed at  $y = 0$ ,  $y = 5$  and  $x = 5$ , and a homogeneous Neumann condition is prescribed at  $x = 0$ . Since the exact interface  $I(t)$  is a circle with center  $(0, \alpha(t))$  and radius 1, the velocity  $V(x, y, t)$  normal to  $I(t)$  at  $(x, y)$  exhibits a significant variation along the front, which makes this example an extremely difficult test for our numerical method. Moreover, since  $V(x, y, t)$  vanishes at both  $(x, y, T/2)$  and  $(1, \alpha(t), t)$ , and is thus very small nearby, this test constitutes a fair measure of robustness under degenerate situations.

Several numerical experiments were performed with both our Adaptive Method (with and without fixed triangles as in §4.2 and §S4.1) and the Fixed Mesh Method [3, 4, 8, 12, 13, 17, 24]. For the latter, the constant of proportionality between  $\tau$  and the (uniform) meshsize was chosen so as to minimize  $\|e_\theta\|_{L^2(Q)}$  for a desired number of DOF, where  $Q := \Omega \times (0, T)$ . The various constants introduced in §§4, S4, 8 are as follows:  $\lambda = 1.5$ ,  $\Lambda = 5$ ,  $M = 5$ ,  $\rho = 0.75\sqrt{\tau}$ ,  $\mu_1 = 2.22$ ,  $\mu_2 = 5.59$ , average value of  $\mu_3 \approx 1.2$ ,  $\mu_1^* = \mu_2^* = 3.5$ ,  $\mu_3^- = 0.5$ ,  $\mu_3^+ = 2$ ,  $C_1 = \|D^2\theta_0\|_{L^\infty(\Omega)}$ ,  $C_2 = 3\|\nabla\theta_0\|_{L^\infty(\Omega)}$ ,  $C_3 = 2 \text{ length}(I_0)$ ,  $C_4 = C_3\lambda$ .

Results are reported in Tables 8.1–8.3, where we have employed the following notation:  $\mathbf{N} :=$  number of (uniform) time steps,  $\mathbf{J} :=$  average number of nodes,  $\mathbf{S}_R :=$  average number of triangles within the refined region,  $\mathbf{S}_p :=$  average number of triangles in the rest of  $\Omega$ ,  $\mathbf{S} :=$  total number of triangles (uniform mesh),  $\mathbf{C} :=$  number of mesh changes (and number of computed solutions rejected, if any),  $\mathbf{CPU} :=$  CPU time in seconds (on a VAX

TABLE 8.1  
*Adaptive method (with fixed triangles)*

N	J	$S_R$	$S_P$	C	$E_\theta^2$	$E_u^2$	$E_\theta^\infty$	$E_I^\infty$	CPU
40	319	496	148	4(1)	16.1	35.4	11.3	5.37	58
60	593	957	233	4(1)	10.5	29.1	6.65	4.44	132
80	875	1430	328	4	7.54	24.2	5.23	3.45	224
120	1413	2320	510	5	5.34	22.7	3.93	2.60	546
160	2089	3504	679	6(1)	4.07	19.6	3.14	2.08	1111
240	3620	6201	1042	8(1)	2.82	17.7	2.41	1.65	2998

TABLE 8.2  
*Adaptive method (without fixed triangles)*

N	J	$S_R$	$S_P$	C	$E_\theta^2$	$E_u^2$	$E_\theta^\infty$	$E_I^\infty$	CPU
40	339	536	144	3	18.8	37.0	10.7	5.83	50
60	592	957	231	4	11.3	28.8	7.42	5.01	126
80	818	1306	334	5	7.59	26.1	5.17	3.26	235
120	1406	2306	510	6	5.73	23.3	3.92	2.66	583
160	2110	3563	662	6	4.60	20.8	3.44	2.32	1068
240	3631	6245	1023	8	2.99	18.5	2.58	1.72	2916

TABLE 8.3  
*Fixed mesh method*

N	J	S	$E_\theta^2$	$E_u^2$	$E_\theta^\infty$	$E_I^\infty$	CPU
50	448	942	15.9	39.7	20.8	13.8	37
75	1017	2104	10.5	30.4	15.3	9.40	113
100	1812	3718	7.81	26.8	12.4	6.88	292
150	4107	8356	5.57	22.3	7.78	5.65	919
200	7361	14912	4.52	18.6	6.31	3.53	2264

8530, vms 4.6) and  $E_\theta^2 := \|e_\theta\|_{L^2(Q)}$ ,  $E_u^2 := \|e_u\|_{L^2(Q)}$ ,  $E_\theta^\infty := \|e_\theta\|_{L^\infty(Q)}$ ,  $E_I^\infty := \max_{1 \leq n \leq N} \text{dist}(I(n\tau), F^n)$ , where the errors are scaled by  $10^2$ . Moreover,  $\|\theta\|_{L^2(Q)} \approx 33.16$ ,  $\|\theta\|_{L^\infty(Q)} \approx 13.38$ ,  $\|u\|_{L^2(Q)} \approx 39.81$ , with an error of one unit in the fourth digit.

In light of these (partial) results, we can certainly claim a superior performance of the Adaptive Method in that it requires less computational labor, say CPU, for a desired global accuracy. Moreover, the proposed local refinement strategies of §4.2 and §4.1, with and without fixed triangles, perform quite similarly. The  $L^2$ -error for temperature  $E_\theta^2$  behaves linearly in  $\tau$ , thus much better than predicted. We also have a (linear) pointwise error  $E_\theta^\infty$  that

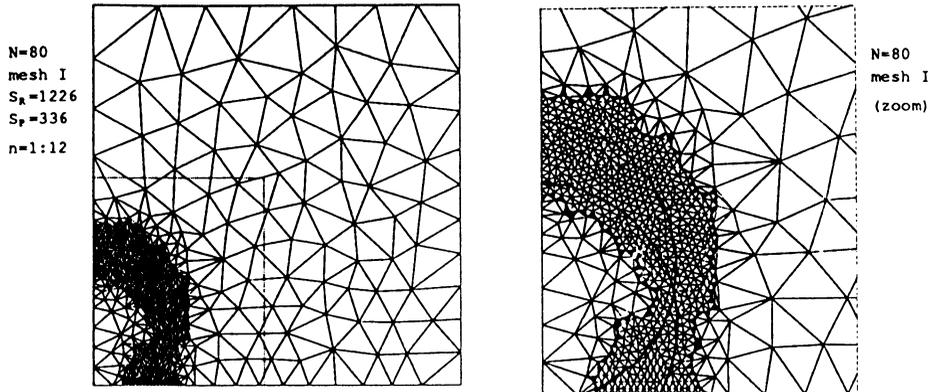


FIGURE 8.1

AM,  $N = 80$ ,  $\mathcal{S}_F \cup \mathcal{S}_B$  free: Mesh I and zoom of refined region

is far from being theoretically explainable. The same happens with  $E_u^2$  and  $E_I^\infty$ . The improvement gained in  $L^\infty$  is clearly more pronounced than that in  $L^2$ . The free boundary is located within one single element, thus confirming the aforementioned local approximation quality. Therefore, we have a practical  $O(\tau)$ -rate of convergence in distance for interfaces, the best one can hope for. Since approximability and nondegeneracy are tied together [14], it is worth noting that the nondegeneracy property is not uniform in the present case.

We finally conclude with several pictures corresponding to the case  $N = 80$  without fixed triangles. The first mesh together with the corresponding Refined Region is depicted in Figure 8.1. The boundary of the Refined Region, called RED ZONE, appears blackened in Figure 8.1b, which also shows the first and last discrete interfaces computed with mesh I. Note that the last interface has escaped from the Refined Region and should thus be discarded as indicated in §4.2. To avoid rejecting a computed solution owing to failure of the free boundary location test, a more flexible strategy for the case without fixed triangles has been designed in [16]. The remaining meshes, generated automatically by ADVFRONT are illustrated in Figure 8.2. Observe the proper grading produced by ADVFRONT. We see how the refined region moves up and down accompanying the interface motion. Note that even in the upmost position, when the interface is motionless, the proposed strategy is successful. Figure 8.3 is a zoom of both the exact and discrete interfaces for a number of time steps. The agreement between these curves is quite remarkable as compared with the local meshsize. Part of the RED ZONE, blackened triangles, can be seen in the second zoom as well. The exact and discrete interfaces obtained with the FMM for  $N = 100$  are depicted in Figure 8.4. It is worth comparing this picture with Figure 8.3: the interface location is drastically improved by the AM.

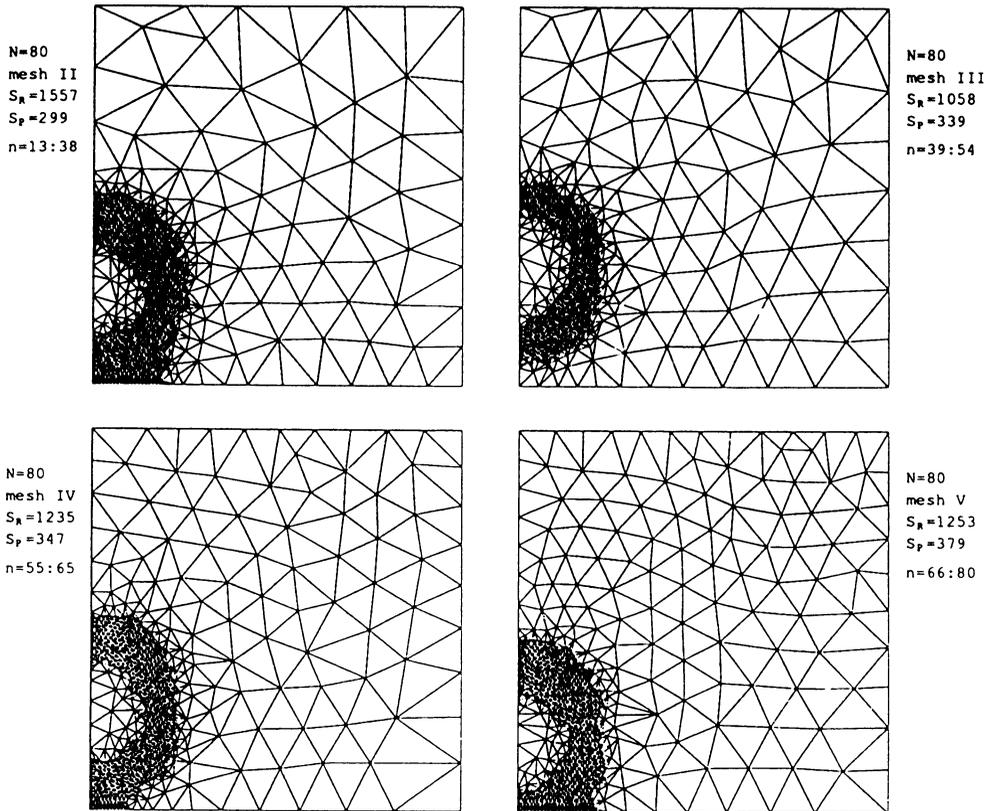


FIGURE 8.2  
AM,  $N = 80$ ,  $\mathcal{S}_F \cup \mathcal{S}_B$  free: Consecutive meshes

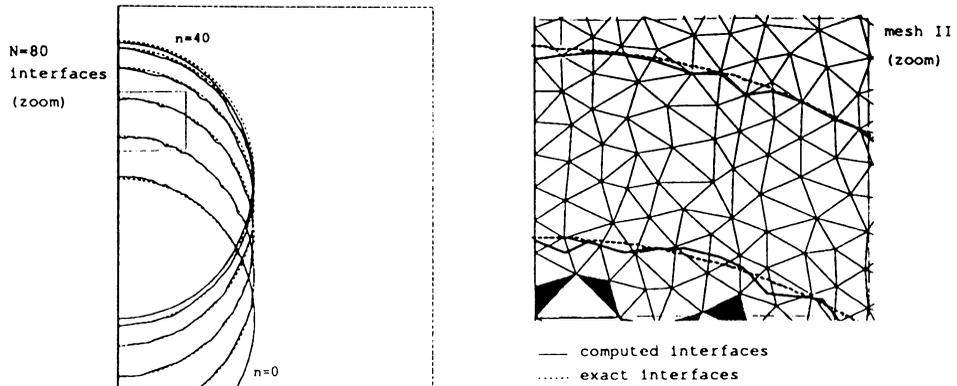


FIGURE 8.3  
AM,  $N = 80$ ,  $\mathcal{S}_F \cup \mathcal{S}_B$  free: Interfaces at  $n = 8k$  ( $0 \leq k \leq 5$ );  
zoom of Mesh II

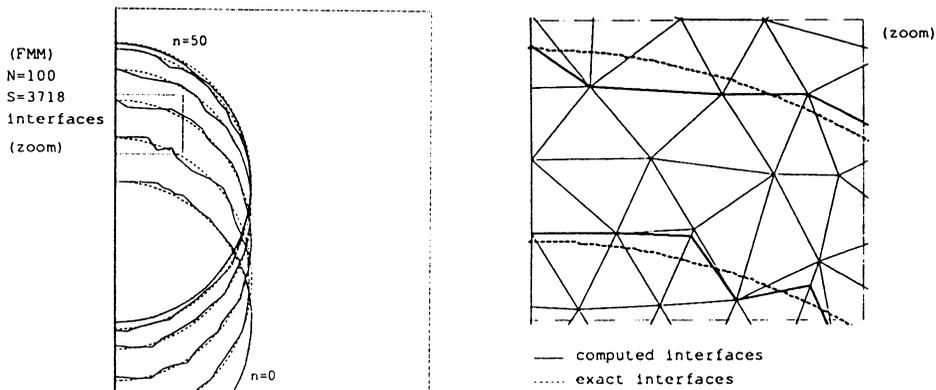


FIGURE 8.4  
FMM,  $N = 100$ : Interfaces at  $n = 10k$  ( $0 \leq k \leq 5$ );  
zoom of the mesh

#### BIBLIOGRAPHY

1. I. Babuška, O. C. Zienkiewicz, J. Gago, and E. R. de A. Oliveira, *Accuracy estimates and adaptive refinements in finite element computations*, Wiley, Chichester, 1986.
2. P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1978.
3. J. F. Ciavaldini, *Analyse numérique d'un problème de Stefan à deux phases par une méthode d'éléments finis*, SIAM J. Numer. Anal. **12** (1975), 464–487.
4. C. M. Elliott, *Error analysis of the enthalpy method for the Stefan problem*, IMA J. Numer. Anal. **7** (1987), 61–71.
5. K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems I: A linear problem*, preprint 31, University of Göteborg, 1988.
6. A. Friedman, *The Stefan problem in several space variables*, Trans. Amer. Math. Soc. **133** (1968), 51–87.
7. P. Grisvard, *Elliptic problems in non-smooth domains*, Pitman, Boston, 1985.
8. J. W. Jerome and M. Rose, *Error estimates for the multidimensional two-phase Stefan problem*, Math. Comp. **39** (1982), 377–414.
9. R. Löhner, *Some useful data structures for the generation of unstructured meshes*, Comm. Appl. Numer. Methods **4** (1988), 123–135.
10. E. Magenes, *Problemi di Stefan bifase in più variabili spaziali*, Matematiche **36** (1981), 65–108.
11. E. Magenes, C. Verdi, and A. Visintin, *Theoretical and numerical results on the two-phase Stefan problem*, SIAM J. Numer. Anal. **26** (1989), 1425–1438.
12. R. H. Nochetto, *Error estimates for two-phase Stefan problems in several space variables, I: Linear boundary conditions*, Calcolo **22** (1985), 457–499.
13. ———, *Error estimates for multidimensional singular parabolic problems*, Japan J. Appl. Math. **4** (1987), 111–138.
14. ———, *Numerical methods for free boundary problems*, Free Boundary Problems: Theory and Applications II (K.-H. Hoffmann and J. Sprekels, eds.), Res. Notes in Math. 186, Pitman, London, 1990, pp. 555–566.

15. R. H. Nochetto, M. Paolini, and C. Verdi, *Local mesh refinements for two-phase Stefan problems in two space variables*, Mathematical Models for Phase Change Problems (J. F. Rodrigues, ed.), Birkhäuser Verlag, Basel, 1989, pp. 261–286.
16. ———, *An adaptive finite element method for two-phase Stefan problems in two space dimensions. Part II: Implementation and numerical experiments*, SIAM J. Sci. Statist. Comput. 1991 (to appear).
17. R.H. Nochetto and C. Verdi, *Approximation of degenerate parabolic problems using numerical integration*, SIAM J. Numer. Anal. **25** (1988), 784–814.
18. M. Paolini, G. Sacchi and C. Verdi, *Finite element approximations of singular parabolic problems*, Internat. J. Numer. Methods. Engrg. **26** (1988), 1989–2007.
19. M. Paolini and C. Verdi, *An automatic triangular mesh generator for planar domains*, Riv. Inform. **20** (1990), 251–267.
20. J. C. W. Rogers and A. E. Berger, *Some properties of the nonlinear semigroup for the problem  $u_t - \Delta f(u) = 0$* , Nonlinear Anal. **8** (1984), 909–939.
21. A. H. Schatz and L. B. Wahlbin, *Maximum norm estimates in the finite element method on plane polygonal domains. Part 1*, Math. Comp. **32** (1978), 73–109.
22. ———, *Maximum norm estimates in the finite element method on plane polygonal domains. Part 2, Refinements*, Math. Comp. **33** (1979), 465–492.
23. ———, *On the quasi-optimality in  $L^\infty$  of the  $H_0^1$ -projection into finite element spaces*, Math. Comp. **38** (1982), 1–22.
24. C. Verdi, *Optimal error estimates for an approximation of degenerate parabolic problems*, Numer. Funct. Anal. Optim. **9** (1987), 657–670.

DEPARTMENT OF MATHEMATICS AND INSTITUTE FOR PHYSICAL SCIENCE AND TECHNOLOGY,  
UNIVERSITY OF MARYLAND, COLLEGE PARK, MD 20742

ISTITUTO DI ANALISI NUMERICA DEL CNR, 27100 PAVIA, ITALY

DIPARTIMENTO DI MECCANICA STRUTTURALE, UNIVERSITÀ DI PAVIA AND ISTITUTO DI ANALISI  
NUMERICA DEL CNR, 27100 PAVIA, ITALY

**Supplement to**  
**AN ADAPTIVE FINITE ELEMENT METHOD FOR TWO-PHASE**  
**STEFAN PROBLEMS IN TWO SPACE DIMENSIONS.**  
**PART I: STABILITY AND ERROR ESTIMATES**

R. H. NOCHETTO† M. PAOLINI‡ AND C. VERDI‡

**S1. Supplement to Section 5.**

LEMMA 5.1. *For all regular meshes  $\mathcal{S}$  there exists a  $C^\infty$ -function  $h: \bar{\Omega} \rightarrow \mathbf{R}^+$  satisfying*

$$(S1.1) \quad D^k h(x) = O(h_S^{1-k}), \quad \forall x \in S \in \mathcal{S}, k \geq 0.$$

*Proof.* For each  $S \in \mathcal{S}$  there exists a (finite) covering  $\{B_i\}$  of  $S$  such that  $\text{card } \{B_i\} = O(1)$ , where  $B_i := B(x_i, r_i)$ ,  $x_i \in S$  and  $r_i := ah_S/2$ ; thus (5.2) yields  $\bar{B}_i := B(x_i, 2r_i) \subset \tilde{S}$ . By virtue of (5.3), this gives rise to a covering  $\{B_i\}_{i=1}^I$  of  $\bar{\Omega}$  satisfying  $b_S := \text{card } \{B_i : S \cap \bar{B}_i \neq \emptyset\} = O(1)$  for all  $S \in \mathcal{S}$  (nonoverlapping property!). Let  $\delta_0 \in C_0^\infty(B(0, 1))$  satisfy  $\int_{\mathbf{R}^2} \delta_0 = 1, 0 \leq \delta_0 \leq 1, \delta_0(x) = 1$  for all  $x \in B(0, 1/2)$  and  $\delta_0$  is radially symmetric (mollifier function). Set  $\mu_i(x) := \delta_0((x - x_i)/(2r_i))$  for all  $x \in \bar{\Omega}$  and note that

$$\text{supp } \mu_i = \bar{B}_i \subset \tilde{S}, \quad 1 \leq \sum_{i=1}^I \mu_i(x) \leq b, \quad \forall x \in \bar{\Omega},$$

where  $b$  depends only on the regularity of  $\mathcal{S}$ ; we certainly have  $b < b_S = O(1)$  for all  $S \in \mathcal{S}$ . Define now  $h \in C^\infty(\bar{\Omega})$  to be

$$h(x) := \frac{a}{b} \sum_{i=1}^I r_i \mu_i(x), \quad \forall x \in \bar{\Omega}.$$

Since  $\rho_{S'} < h_S$  for all  $S' \in \mathcal{S}_S$ , it is easily seen from (5.1) that, if  $x_i \in S$ , then

$$(S1.2) \quad \frac{a^2}{b} r_i = \frac{a^3}{2b} h_S \leq \frac{a^2}{2b} \min_{S' \in \mathcal{S}_S} \rho_{S'} \leq h(x) \leq \frac{a}{2} \max_{S' \in \mathcal{S}_S} \rho_{S'} < \frac{a}{2} h_S = r_i, \quad \forall x \in B_i.$$

Hence  $B(x, h(x)) \subset \bar{B}_i \subset \tilde{S}$  for all  $x \in B_i$ . Since

$$D^k h(x) = \frac{a}{b} \sum_{i=1}^I r_i D^k \mu_i(x) = C \sum_{i=1}^I r_i^{1-k} D^k \delta_0((x - x_i)/(2r_i)), \quad \forall x \in \Omega,$$

the desired result easily follows from (S1.2).  $\square$

†Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742 USA.

‡Istituto di Analisi Numerica del CNR, 27100 Pavia, Italy.

‡Dipartimento di Meccanica Strutturale, Università di Pavia and Istituto di Analisi Numerica del CNR, 27100 Pavia, Italy.

PROOF OF LEMMA 5.2. Via a standard scaling argument, it suffices to derive the following form of (5.7) over the master element, for  $1 \leq p < \infty$ ,

$$(S1.3) \quad \|\Pi\zeta(V) - \tilde{\Pi}\zeta(V)\|_{L^p(S)} + \|\nabla[\Pi\zeta(V) - \tilde{\Pi}\zeta(V)]\|_{L^p(S)} \leq C(\Phi_{\mathcal{E}_S} + \Psi_{\mathcal{S}_S}),$$

where, for convenience, we have set

$$\Phi_{\mathcal{E}_S} := L_{\zeta} \left( \sum_{e \in \mathcal{E}_S} h_e^2 D_e^p \right)^{1/p}, \quad \Psi_{\mathcal{S}_S} := L_{\zeta'} \left( \sum_{S \in \mathcal{S}_S} h_S^2 d_S^{2p} \right)^{1/p}.$$

The case  $p = \infty$  is similar. The proof of (S1.3) will be split into several steps. Let  $\rho \in C_0^\infty$  be defined by

$$\rho(x, y) := h(x)^{-2} \delta_0((x - y)h(x)^{-1}), \quad \forall x \in \hat{S}, \quad y \in \mathbf{R}^2,$$

where both  $h$  and  $\delta_0$  come from Lemma 5.1. Let  $\{B_i\}_{i=1}^I$  be the covering of  $\hat{S}$  used also in Lemma 5.1. By virtue of (5.1) and (S1.2), we have  $B(x, h(x)) \subset \hat{B}_i \subset \hat{S}$  for all  $x \in B_i$ . We now introduce a regularization  $\zeta(V)_\rho$  of  $\zeta(V)$ . Let  $\zeta(V)_\rho$  be the  $C^\infty$ -function defined by

$$\zeta(V)_\rho(x) := \int_{\mathbf{R}^2} \zeta(V(y))\rho(x, y)dy = \int_{B(x, h(x))} \zeta(V(y))\rho(x, y)dy, \quad \forall x \in \hat{S},$$

where  $V$  is extended as a continuous piecewise linear function outside of  $\hat{S}$ , say in an obvious manner, whenever  $\hat{S}$  has one edge on  $\partial\Omega$ . We claim that

$$|D^2\zeta(V)_\rho(x)| \leq CL_\zeta \sum_{e \in \mathcal{E}_{\hat{S}}} D_e + CL_{\zeta'} \sum_{S \in \mathcal{S}_{\hat{S}}} d_S^2 \quad \forall x \in B_i.$$

To prove this, we shall need the following properties of function  $\rho$ : for all  $x \in \hat{S}$ ,

$$\int_{\mathbf{R}^2} \rho(x, y)dy = 1, \quad \int_{\mathbf{R}^2} (x - y)\rho(x, y)dy = 0,$$

from which

$$\int_{\mathbf{R}^2} D_x^k \rho(x, y)dy = 0, \quad k \geq 1, \\ \int_{\mathbf{R}^2} (x - y) D_{x_i} D_{x_j} \rho(x, y)dy = -\delta_{ij}, \quad i, j = 1, 2, \quad \int_{\mathbf{R}^2} (x - y) D_x^2 \rho(x, y)dy = 0, \quad k \geq 2.$$

Here,  $\delta_{ij}$  stands for the Kronecker symbol. Moreover, in view of (S1.1), the following inequalities hold for all  $x \in S \in \mathcal{S}_S$  as well:

$$\|\nabla_x \rho(x, \cdot)\|_{L^1(\mathbf{R}^2)} \leq Ch(x)^{-1} \leq Ch_S^{-1}, \quad \|D_x^2 \rho(x, \cdot)\|_{L^1(\mathbf{R}^2)} \leq Ch(x)^{-2} \leq Ch_S^{-2}.$$

On making use of these properties of  $\rho$ , we easily get for all  $x \in B_i$

$$\begin{aligned} D^2\zeta(V)_\rho(x) &= \int_{\mathbf{R}^2} \zeta(V(y)) D_x^2 \rho(x, y) dy \\ &= \int_{B(x, h(x))} [\zeta(V(y)) - \zeta(V(x)) + \nabla\zeta(V(x)) \cdot (x - y)] D_x^2 \rho(x, y) dy \\ &= \int_{B(x, h(x))} I(y) D_x^2 \rho(x, y) dy, \end{aligned}$$

where

$$I(y) := \left( \int_0^1 \zeta'(V(sy + (1-s)x)) \nabla V(sy + (1-s)x) ds - \zeta'(V(x)) \right) \nabla V(x) \cdot (y - x).$$

But

$$\begin{aligned} I(y) &= (y - x) \cdot \left( \int_0^1 [\zeta'(V(sy + (1-s)x) - \zeta'(V(x))] \nabla V(sy + (1-s)x) ds \right. \\ &\quad \left. + \zeta'(V(x)) \int_0^1 [\nabla V(sy + (1-s)x) - \nabla V(x)] ds \right), \end{aligned}$$

whence

$$|I(y)| \leq CL_\zeta \sum_{e \in \mathcal{E}_{B(x, h(x))}} h_e \delta_e + CL_{\zeta'} \sum_{S \in \mathcal{S}_{B(x, h(x))}} h_S^2 d_S^2.$$

Consequently, for all  $x \in B_i$ , we obtain

$$\begin{aligned} |D^2\zeta(V)_\rho(x)| &\leq C \int_{\mathbf{R}^2} |D_x^2 \rho(x, y)| dy \left( L_\zeta \sum_{e \in \mathcal{E}_{B(x, h(x))}} h_e \delta_e + L_{\zeta'} \sum_{S \in \mathcal{S}_{B(x, h(x))}} h_S^2 d_S^2 \right) \\ &\leq CL_\zeta \sum_{e \in \mathcal{E}_{B(x, h(x))}} D_e + CL_{\zeta'} \sum_{S \in \mathcal{S}_{B(x, h(x))}} d_S^2 \leq CL_\zeta \sum_{e \in \mathcal{E}_{B_i}} D_e + CL_{\zeta'} \sum_{S \in \mathcal{S}_{B_i}} d_S^2. \end{aligned}$$

In view of (5.3) and (S1.2) we have  $\text{card } \mathcal{E}_{B_i} = O(1)$  which, in turn, implies the pointwise estimate  $|D^2\zeta(V)_\rho(x)| \leq CL_\zeta \sum_{e \in \mathcal{E}_{B_i}} D_e^{1/p} + CL_{\zeta'} \sum_{S \in \mathcal{S}_{B_i}} d_S^{2/p}$  for all  $x \in B_i$ , as results from Hölder's inequality. This inequality will serve to derive an  $L^p$ -bound. By virtue of the nonoverlapping property of  $\{B_i\}_{i=1}^I$ , we have

$$\begin{aligned} \|D^2\zeta(V)_\rho\|_{L^p(S)} &\leq \left( \sum_{i=1}^I \int_{B_i} |D^2\zeta(V)_\rho(x)|^p dx \right)^{1/p} \\ &\leq C \left( \sum_{i=1}^I L_\zeta^p \sum_{e \in \mathcal{E}_{B_i}} h_e^2 D_e^p + L_{\zeta'}^p \sum_{S \in \mathcal{S}_{B_i}} h_S^2 d_S^{2p} \right)^{1/p} \leq C(\Phi_{\mathcal{E}_S} + \Psi_{\mathcal{S}_S}), \end{aligned}$$

which is the desired bound. Standard interpolation theory [2, p. 115] then yields

$$(S1.6) \quad \begin{aligned} & \|\zeta(V)_\rho - \tilde{\Pi}\zeta(V)_\rho\|_{L^r(\hat{S})} + \|\nabla(\zeta(V)_\rho - \tilde{\Pi}\zeta(V)_\rho)\|_{L^r(\hat{S})} \\ & \leq C\|D^2\zeta(V)_\rho\|_{L^r(\hat{S})} \leq C(\Phi_{\mathcal{E}_3} + \Psi_{\mathcal{S}_3}). \end{aligned}$$

At the same time, we can compare  $\zeta(V)$  and  $\zeta(V)_\rho$  as follows:

$$\nabla(\zeta(V) - \zeta(V)_\rho)(x) = \int_{B(x, h_\rho(x))} [\zeta(V(x)) - \zeta(V(y)) - \nabla\zeta(V(x)) \cdot (x - y)] \nabla_x \rho(x, y) dy.$$

Hence, for all  $x \in B_1$  we obtain

$$\begin{aligned} \|\nabla(\zeta(V) - \zeta(V)_\rho)(x)\| & \leq C \int_{\mathbf{R}^2} |\nabla_x \rho(x, y)| dy \left( L_\zeta \sum_{e \in \mathcal{E}_{B(x, h_\rho(x))}} h_e \delta_e + L_\zeta \sum_{S \in \mathcal{S}_{B(x, h_\rho(x))}} h_S^2 d_S^2 \right) \\ & \leq CL_\zeta \sum_{e \in \mathcal{E}_{B_1}} h_e D_e + CL_\zeta \sum_{S \in \mathcal{S}_{B_1}} h_S^2 d_S^2. \end{aligned}$$

Arguing as before, we deduce that

$$(S1.5) \quad \|\nabla(\zeta(V) - \zeta(V)_\rho)\|_{L^r(\hat{S})} \leq C(\Phi_{\mathcal{E}_3} + \Psi_{\mathcal{S}_3}).$$

A similar calculation for  $(\zeta(V) - \zeta(V)_\rho)(x)$ , valid for all  $x \in B_1$ , leads to

$$\begin{aligned} |(\zeta(V) - \zeta(V)_\rho)(x)| & = \left| \int_{B(x, h_\rho(x))} [\zeta(V(x)) - \zeta(V(y)) - \nabla\zeta(V(x)) \cdot (x - y)] \rho(x, y) dy \right| \\ & \leq CL_\zeta \sum_{e \in \mathcal{E}_{B_1}} h_e^2 D_e + CL_\zeta \sum_{S \in \mathcal{S}_{B_1}} h_S^2 d_S^2, \end{aligned}$$

and, as a consequence, to

$$(S1.6) \quad \|\zeta(V) - \zeta(V)_\rho\|_{L^r(\hat{S})} \leq C(\Phi_{\mathcal{E}_3} + \Psi_{\mathcal{S}_3}).$$

Moreover, since  $\tilde{\Pi}\zeta(V) - \zeta(V)_\rho$  is linear in  $\hat{S}$ , we see that

$$\|\tilde{\Pi}[\zeta(V) - \zeta(V)_\rho]\|_{W^{1, r}(\hat{S})} \leq C\|\tilde{\Pi}[\zeta(V) - \zeta(V)_\rho]\|_{L^\infty(\hat{S})} \leq C\|\zeta(V) - \zeta(V)_\rho\|_{L^\infty(\hat{S})},$$

whence

$$(S1.7) \quad \|\tilde{\Pi}[\zeta(V) - \zeta(V)_\rho]\|_{W^{1, r}(\hat{S})} \leq C(\Phi_{\mathcal{E}_3} + \Psi_{\mathcal{S}_3}),$$

as results from the previous pointwise bound for  $\zeta(V) - \zeta(V)_\rho$  and Hölder's inequality. To complete the argument, we make use again of standard interpolation estimates [2, p. 115] as follows:

$$(S1.8) \quad \begin{aligned} \|\zeta(V) - \Pi\zeta(V)\|_{W^{1, r}(\hat{S})} & \leq C \left( \sum_{S \in \mathcal{S}_3} h_S^p \|D^2\zeta(V)\|_{L^r(\hat{S})}^p \right)^{1/p} \\ & \leq CL_{\zeta'} \left( \sum_{S \in \mathcal{S}_3} h_S^{2+p} d_S^{2p} \right)^{1/p} \leq C\Psi_{\mathcal{S}_3}. \end{aligned}$$

Combining the inequalities (S1.4)-(S1.8), we finally obtain the desired estimate (S1.3).  $\square$

PROOF OF LEMMA 5.3. In view of (5.6) it is enough to prove the above estimates for  $p = \infty$ . We first examine  $\zeta(V) - \Pi\zeta(V)$ . Standard interpolation estimates yield

$$\|\zeta(V) - \Pi\zeta(V)\|_{L^\infty(\hat{S})} \leq \|\zeta(V) - \Pi\zeta(V)\|_{L^\infty(\tilde{S})} \leq CL_{\zeta'} \max_{S \in \mathcal{S}_3} (h_S d_S)^2,$$

as well as

$$\|\nabla[\zeta(V) - \Pi\zeta(V)]\|_{L^\infty(\hat{S})} \leq \|\nabla[\zeta(V) - \Pi\zeta(V)]\|_{L^\infty(\tilde{S})} \leq CL_{\zeta'} \max_{S \in \mathcal{S}_3} (h_S d_S^2).$$

We next deal with  $\zeta(V) - \tilde{\Pi}\zeta(V)$  and assume, without loss of generality, that  $\hat{S}$  is the master triangle. A standard scaling argument will produce estimates (5.8) and (5.9) as soon as we have

$$\|\nabla[\zeta(V) - \tilde{\Pi}\zeta(V)]\|_{L^\infty(\hat{S})} \leq CL_\zeta \max_{e \in \mathcal{E}_3} \delta_e + CL_{\zeta'} \max_{S \in \mathcal{S}_3} (h_S d_S^2).$$

Indeed, since  $\zeta(V) - \tilde{\Pi}\zeta(V)$  vanishes at the vertices of  $\hat{S}$ , we infer that

$$\|\zeta(V) - \tilde{\Pi}\zeta(V)\|_{L^\infty(\hat{S})} \leq C\|\nabla[\zeta(V) - \tilde{\Pi}\zeta(V)]\|_{L^\infty(\hat{S})},$$

which, coupled with the fact that  $h_S > h_{\tilde{S}} = O(1)$  for all  $S \in \mathcal{S}_3$  as results from (5.1) and (5.5), leads to (5.8). Moreover, since

$$\min_{\tilde{S}} \zeta(V)_x \text{ (resp. } \zeta(V)_y) \leq \tilde{\Pi}\zeta(V)_x \text{ (resp. } \tilde{\Pi}\zeta(V)_y) \leq \max_{\tilde{S}} \zeta(V)_x \text{ (resp. } \zeta(V)_y),$$

we easily see that

$$\|\nabla[\zeta(V) - \tilde{\Pi}\zeta(V)]\|_{L^\infty(\hat{S})} \leq \max_{x, y \in \hat{S}} |\nabla[\zeta(V(x)) - \zeta(V(y))]| =: \max_{x, y \in \hat{S}} I(x, y).$$

But

$$\begin{aligned} I(x, y) &\leq |\zeta'(V(y))\nabla[V(x) - V(y)]| + |\zeta'(V(x)) - \zeta'(V(y))|\nabla V(x)| \\ &\leq L_\zeta \sum_{e \in \mathcal{E}_\zeta} \delta_e + L_{\zeta'} \sum_{S \in \mathcal{S}_\zeta} (h_S d_S^q) \leq C L_\zeta \max_{e \in \mathcal{E}_\zeta} \delta_e + C L_{\zeta'} \max_{S \in \mathcal{S}_\zeta} (h_S d_S^q), \end{aligned}$$

because of (5.6). This provides the desired bound and completes the proof of the lemma.  $\square$

PROOF OF LEMMA 5.4. In view of (2.1), (2.2) and (2.12), we infer that

$$\|u_0 - U^0\|_{L^\infty(S)} \leq C\tau, \quad \forall S \in \mathcal{S}^0 (= \mathcal{S}^1) : S \cap I_0 = \emptyset.$$

Since  $u_0$  is discontinuous across  $I_0$ , the best we can say is  $\|u_0 - U^0\|_{L^\infty(\mathcal{F}^0)} \leq C$ . The fact that  $I_0$  has finite length leads now to

$$\|u_0 - U^0\|_{L^1(\mathcal{F}^0)} \leq C\tau^{1/p}, \quad \forall 1 \leq p < \infty.$$

Let  $\varphi \in H_0^1(\Omega)$  be given. Then, using the Hölder inequality combined with a well-known 2-D Sobolev inequality, we get

$$|(u_0 - U^0, \varphi)| \leq \|u_0 - U^0\|_{L^1(\Omega)} \|\varphi\|_{L^1(\Omega)} \leq Cq^{1/2}\tau^{1-1/q} \|\varphi\|_{H_0^1(\Omega)}.$$

The asserted estimate follows from taking  $q = \lceil \log \tau \rceil$ .  $\square$

We now demonstrate auxiliary trace estimates for the interpolation error  $\Theta - \hat{\Theta}$ . Let  $\hat{\mathcal{E}}_0 := \{\hat{e} \in \mathcal{E} : A_{\hat{e}} \neq \emptyset\}$ , where  $A_{\hat{e}} := (\hat{S}_1 \cup \hat{S}_2) \cap \hat{\Omega}_0$  and  $\{\hat{S}_1, \hat{S}_2\} = \mathcal{A}_{\hat{e}}$ ; note that  $\hat{\mathcal{E}}_0 \supset \hat{\mathcal{E}} \setminus (\mathcal{E}_F \cup \mathcal{E}_B)$ . Let  $\hat{\mathcal{E}}_2 \subset \hat{\mathcal{E}}_0$  be the set of all those  $\hat{e}$ 's for which  $\mathcal{A}_{\hat{e}} \cap \hat{\mathcal{S}}_2 \neq \emptyset$ , set  $\hat{\mathcal{E}}_1 := \hat{\mathcal{E}}_0 \setminus \hat{\mathcal{E}}_2$ .

LEMMA 5.7. Let  $\hat{e} \in \hat{\mathcal{E}}_1$ . Then

$$(S1.9) \quad \|\nabla(\Theta - \hat{\Theta})\|_{L^1(\hat{e})} \leq C \sum_{e \in \mathcal{E}_{A_{\hat{e}}}} h_e^2 D_e + C \sum_{S \in \mathcal{S}_{A_{\hat{e}}}} (h_S d_S)^2.$$

Proof. Let  $\zeta = \beta$ ,  $V = U$  and  $p = 1$  in Lemma 5.2. Let  $\hat{S} \in \mathcal{A}_{\hat{e}} \cap \hat{\mathcal{S}}_0$ . Recall the notation used in Lemma 5.2 and, in addition, the pointwise estimates

$$|D^2 \beta(U)_\rho(x)| \leq C \sum_{e \in \mathcal{E}_B} D_e + C \sum_{S \in \mathcal{S}_B} d_S^q,$$

and

$$\|\nabla|\beta(U) - \beta(U)_\rho|(x)\| \leq C \sum_{e \in \mathcal{E}_B} h_e D_e + C \sum_{S \in \mathcal{S}_B} h_S d_S^q,$$

for all  $x \in B_{\hat{e}}$ . Hence,

$$\begin{aligned} \|\nabla|\beta(U) - \beta(U)_\rho|\|_{L^1(\hat{e})} &\leq C \sum_{i=1}^{\nu_{\hat{e} \cap B_{\hat{e}}}} \int_{e \cap B_{\hat{e}}} |\nabla|\beta(U) - \beta(U)_\rho|(x)| d\sigma(x) \\ &\leq C \sum_{e \in \mathcal{E}_{\hat{e}_2}} h_e^2 D_e + C \sum_{S \in \mathcal{S}_{\hat{e}_2}} h_S^2 d_S^q. \end{aligned}$$

A well-known trace result can now be applied to arrive at

$$\begin{aligned} \|\nabla|\beta(U)_\rho - \hat{\Pi}\beta(U)_\rho|\|_{L^1(\hat{e})} &\leq Ch_{\hat{e}}^{-1} \|\nabla|\beta(U)_\rho - \hat{\Pi}\beta(U)_\rho|\|_{L^1(\hat{e})} + C \|D^2 \beta(U)_\rho\|_{L^1(\hat{e})} \\ &\leq C \sum_{e \in \mathcal{E}_{\hat{e}_2}} h_e^2 D_e + C \sum_{S \in \mathcal{S}_{\hat{e}_2}} h_S^2 d_S^q. \end{aligned}$$

At the same time, a local inverse inequality combined with (S1.7) leads to

$$\begin{aligned} \|\nabla|\hat{\Pi}\beta(U) - \hat{\Pi}\beta(U)_\rho|\|_{L^1(\hat{e})} &\leq Ch_{\hat{e}}^{-1} \|\nabla|\hat{\Pi}\beta(U) - \hat{\Pi}\beta(U)_\rho|\|_{L^1(\hat{e})} \\ &\leq C \sum_{e \in \mathcal{E}_{\hat{e}_2}} h_e^2 D_e + C \sum_{S \in \mathcal{S}_{\hat{e}_2}} h_S^2 d_S^q. \end{aligned}$$

Finally, a standard  $L^\infty$ -interpolation estimate yields

$$\begin{aligned} \|\nabla|\beta(U) - \Pi\beta(U)\|_{L^1(\hat{e})} &= \sum_{S \in \mathcal{S}_{\hat{e}_2}} \int_{e \cap S} |\nabla|\beta(U) - \Pi\beta(U)|(x)| d\sigma(x) \\ &\leq C \sum_{S \in \mathcal{S}_{\hat{e}_2}} h_S^2 \|D^2 \beta(U)\|_{L^\infty(S)} \leq C \sum_{S \in \mathcal{S}_{\hat{e}_2}} h_S^2 d_S^q, \end{aligned}$$

and concludes the proof of the lemma.  $\square$

LEMMA 5.8. Let  $\hat{e} \in \hat{\mathcal{E}}_2$ . Then

$$(S1.10) \quad \|\nabla(\Theta - \hat{\Theta})\|_{L^1(\hat{e})} \leq Ch_{\hat{e}} \left( \sum_{e \in \mathcal{E}_{A_{\hat{e}}}} \delta_e + \sum_{S \in \mathcal{S}_{A_{\hat{e}}}} h_S d_S^q \right).$$

Proof. Let  $\zeta = \beta$ ,  $V = U$  and  $p = \infty$  in Lemma 5.3. Let  $\hat{S} \in \mathcal{A}_{\hat{e}} \cap \hat{\mathcal{S}}_0$ . On using (4.10) in conjunction with (5.6), we readily get

$$\|\nabla(\Theta - \hat{\Theta})\|_{L^1(\hat{e})} \leq Ch_{\hat{e}} \|\nabla(\Theta - \hat{\Theta})\|_{L^\infty(\hat{e})} \leq Ch_{\hat{e}} \left( \sum_{e \in \mathcal{E}_{\hat{e}_2}} \delta_e + \sum_{S \in \mathcal{S}_{\hat{e}_2}} h_S d_S^q \right),$$

which yields the assertion.  $\square$

Estimate (5.10), with  $\zeta = \beta$ , and (2.18), together with the fact that  $\hat{\Theta} = \Theta$  for all  $e \in \mathcal{E}_F \cup \mathcal{E}_B$ , lead to

$$\begin{aligned} \int_{\Omega_0} \nabla(\hat{\Theta} - \Theta) \cdot \nabla \Theta &= \sum_{e \in \mathcal{E}_0} \langle (\hat{\Theta} - \Theta), [\nabla \Theta]_e \cdot \nu_e \rangle \\ &\leq C \|\hat{\Theta} - \Theta\|_{L^\infty(\Omega)} \sum_{e \in \mathcal{E}_0} h_e^2 D_e \leq C\tau \sum_{e \in \mathcal{E}_0} h_e^2 D_e. \end{aligned}$$

In view of (5.11), with  $\zeta = \beta$ , and inequality  $\sigma^2 \sum_{S \in \mathcal{S}} h_S^2 d_S^2 \leq \|\nabla \Theta\|_{L^2(\Omega)}^2$ , the desired estimate follows immediately.  $\square$

## S2. Supplement to Section 6.

**PROOF OF LEMMA 6.1.** Let  $C_0 := \|U^n\|_{L^\infty(\Omega)}$  and  $f_0 := |f(0)|$ . In view of (2.1) and (2.13) it is enough to prove that  $\|U^n\|_{L^\infty(\Omega)} \leq C_n$ , where

$$(S2.1) \quad C_n := C_0 \exp(n\tau L_\beta L_f) + \frac{f_0}{L_\beta L_f} (\exp(n\tau L_\beta L_f) - 1).$$

Since (S2.1) is obviously valid for  $n = 0$ , assume, by induction, that it holds for  $n - 1$  as well. Let  $x \in \Omega$  be a point at which  $U^n$  attains its maximum. Since  $U^n$  is piecewise linear,  $x$  is a node of  $\mathcal{S}^n$  (say  $x = x^n$ ). Function  $\Theta^n$  attains also its maximum at  $x^n$  because of (2.13) and the monotonicity of  $\beta$ . Then, by virtue of (6.1) we can write

$$\begin{aligned} U^n(x^n) \int_{\Omega} \chi^n &= (U^n, \chi^n)^n \leq (\hat{U}^{n-1}, \chi^n)^n + \tau (f(\hat{\Theta}^{n-1}), \chi^n)^n \\ &= (\hat{U}^{n-1}(x^n) + \tau f(\hat{\Theta}^{n-1}(x^n))) \int_{\Omega} \chi^n, \end{aligned}$$

thus

$$U^n(x^n) \leq \hat{U}^{n-1}(x^n) + \tau f(\hat{\Theta}^{n-1}(x^n)).$$

On using the induction assumption, together with the inequalities

$$\|\hat{U}^{n-1}\|_{L^\infty(\Omega)} \leq \|U^{n-1}\|_{L^\infty(\Omega)}, \quad \|\hat{\Theta}^{n-1}\|_{L^\infty(\Omega)} \leq L_\beta \|U^{n-1}\|_{L^\infty(\Omega)},$$

we see that

$$U^n(x^n) \leq C_{n-1}(1 + \tau L_\beta L_f) + \tau f_0.$$

The fact that  $1 + \tau L_\beta L_f \leq \exp(\tau L_\beta L_f)$  results in  $U^n(x^n) \leq C_n$ , where  $C_n$  is defined in (S2.1). This completes the induction argument.  $\square$

**PROOF OF LEMMA 6.2.** Since no confusion is possible, we remove the superscript  $n$  and simply use the following notation:  $\mathcal{S} := \mathcal{S}^n$ ,  $\Pi := \Pi^n$ , and  $k_{ij} := \int_{\Omega} \nabla \chi_i^n \cdot \nabla \chi_j^n$ ,

**PROOF OF LEMMA 5.9.** A bound on the first term follows from  $\|\nabla \Theta^0\|_{L^\infty(\Omega)} \leq \|\nabla \theta_0\|_{L^\infty(\Omega)}$  and (2.2). For the second term, we split  $\mathcal{E}^1$  into the set  $\mathcal{E}_F^1$  of those  $e$ 's such that  $A_e \cap I_0 \neq \emptyset$  and the complement, where  $A_e := S_1 \cup S_2$ ,  $\{S_1, S_2\} = \mathcal{A}_e$ . We use again (2.2) to get  $\delta_e \leq C$  for all  $e \in \mathcal{E}_F^1$ , which, in view of (2.3), leads to

$$\sum_{e \in \mathcal{E}_F^1} h_e^2 D_e = \sum_{e \in \mathcal{E}_F^1} h_e \delta_e \leq C \sum_{e \in \mathcal{E}_F^1} h_e \leq C \text{length}(I_0) \leq C,$$

because  $I_0$  has finite length. On the other hand, for all  $e \notin \mathcal{E}_F^1$ , we have

$$D_e = h_e^{-1} \delta_e \leq h_e^{-1} \|\text{osc } \nabla \theta_0\|_{L^\infty(A_e)} \leq \|D^2 \theta_0\|_{L^\infty(A_e)} \leq C,$$

where "osc" stands for oscillation. This concludes the argument.  $\square$

**PROOF OF LEMMA 5.10.** Since  $\hat{\Theta}|_S = \Theta|_S$  for all  $\hat{S} \in (\mathcal{S}_B \cup \mathcal{S}_F)$ , we only have to consider  $\hat{e} \in \hat{\mathcal{E}}_0$ . Set  $\hat{\mathcal{E}}_1 := \{\hat{e} \in \hat{\mathcal{E}}_2 : A_{\hat{e}} \subset S$ , for some  $S \in \mathcal{S}_0\}$  and  $\hat{\mathcal{E}}_2^c := \hat{\mathcal{E}}_2 \setminus \hat{\mathcal{E}}_1$ . Let first  $\hat{e} \in \hat{\mathcal{E}}_1^c$ . It is easily seen that  $D_{\hat{e}} \leq CL_\beta D_{S_2}^2 \leq C|L_\beta|^2 L_\beta d_{S_2}^2$ , where  $S_2 \supset A_{\hat{e}}$ , and so

$$\sum_{e \in \hat{\mathcal{E}}_1^c} h_e^2 D_e \leq C \sum_{e \in \hat{\mathcal{E}}_1^c} h_e^2 d_{S_2}^2 \leq C \sum_{S \in \mathcal{S}_0} d_S^2 \sum_{e \in \hat{\mathcal{E}}_1^c} h_e^2 \leq C \sum_{S \in \mathcal{S}_0} h_S^2 d_S^2,$$

where  $\hat{\mathcal{E}}_2^c := \{\hat{e} \in \hat{\mathcal{E}}_2^c : A_{\hat{e}} \subset S\}$ . For the two remaining cases, namely  $\hat{\mathcal{E}}_1^c$  and  $\hat{\mathcal{E}}_2$ , we can obviously replace  $D_{\hat{e}}$  by  $h_{\hat{e}} \|\nabla(\hat{\Theta})|_{S_1} - \Theta|_{S_1}\|_{L^1(\hat{e})} + \|\nabla(\hat{\Theta})|_{S_2} - \Theta|_{S_2}\|_{L^1(\hat{e})}$ . Let next  $\hat{e} \in \hat{\mathcal{E}}_2^c$ . Then, Lemma 5.8 implies

$$\begin{aligned} \sum_{e \in \hat{\mathcal{E}}_2^c} h_e^2 D_e &\leq C \sum_{e \in \hat{\mathcal{E}}_2^c} h_e \left( \sum_{e \in \mathcal{E}_{A_e}} h_e D_e + \sum_{S \in \mathcal{S}_{A_e}} h_S d_S^2 \right) \\ &\leq C \sum_{e \in \hat{\mathcal{E}}_2^c} h_e D_e + C \sum_{S \in \mathcal{S}} h_S d_S^2 \sum_{e \in \hat{\mathcal{E}}_2^c} h_e \leq C \sum_{e \in \hat{\mathcal{E}}_2^c} h_e^2 D_e + C \sum_{S \in \mathcal{S}} h_S^2 d_S^2, \end{aligned}$$

where  $\mathcal{E}_{A_e}^c := \{\hat{e} \in \hat{\mathcal{E}}_2^c : e \cap \hat{e} \neq \emptyset\}$  and  $\mathcal{S}_{A_e}^c := \{\hat{e} \in \hat{\mathcal{E}}_2^c : A_{\hat{e}} \cap \partial S \neq \emptyset\}$ . Here we have argued as at the end of Lemma 5.6 to conclude that  $\sum_{e \in \hat{\mathcal{E}}_2^c} h_e \leq Ch_e$  as well as  $\sum_{e \in \hat{\mathcal{E}}_2^c} h_e \leq Ch_S$ .

Suppose now  $\hat{e} \in \hat{\mathcal{E}}_1$ . Lemma 5.7 finally yields

$$\sum_{e \in \hat{\mathcal{E}}_1} h_e^2 D_e \leq C \sum_{e \in \hat{\mathcal{E}}_1} \left( h_e^2 D_e + \sum_{S \in \mathcal{S}_{A_e}} h_S^2 d_S^2 \right) \leq C \sum_{e \in \hat{\mathcal{E}}_1} h_e^2 D_e + C \sum_{S \in \mathcal{S}} h_S^2 d_S^2. \quad \square$$

**PROOF OF LEMMA 5.11.** Since  $\hat{\Theta}|_S = \Theta|_S$  for all  $\hat{S} \in (\mathcal{S}_B \cup \mathcal{S}_F)$ , we only have to prove the estimate in  $\Omega_0$ . We obviously have

$$\int_{\Omega_0} |\nabla \hat{\Theta}|^2 = \int_{\Omega_0} |\nabla \Theta|^2 + \int_{\Omega_0} |\nabla(\hat{\Theta} - \Theta)|^2 + 2 \int_{\Omega_0} \nabla(\hat{\Theta} - \Theta) \cdot \nabla \Theta.$$

where  $\tilde{W} := \cup\{S \in \mathcal{S} : S \cap W \neq \emptyset\}$  and  $\tilde{W} := \text{Interior}(\cup\{S \in \mathcal{S} : S \subset W\})$ .

LEMMA S3.1. Let  $\varphi \in \mathbf{V}$  be discrete harmonic in  $B_r$ , that is,

$$(S3.2) \quad \langle \nabla \varphi, \nabla \chi_j \rangle = 0, \quad \text{for all } 1 \leq j \leq J \text{ such that } x_j \in B_r.$$

Then there exists  $\Gamma > 1$  such that for  $r \geq \Gamma h$  the following energy estimate holds:

$$(S3.3) \quad \|\nabla \varphi\|_{L^2(B_r)} \leq Cr^{-1} \|\varphi\|_{L^2(B_r)}.$$

Proof. It is easily seen, from (S3.1) and (S3.2), that

$$\|\nabla \varphi\|_{L^2(B_{r^*})}^2 \leq 2^{-1}(\varphi^2, \Delta \omega) + \langle \nabla \varphi, \nabla(\omega^2 \varphi - \chi) \rangle =: I + II, \quad \forall \chi \in \mathbf{V}.$$

In view of (S3.1), we have  $I \leq Cr^{-2} \|\varphi\|_{L^2(A_k)}^2$ . On the other hand, taking  $\chi = \Pi(\omega^2 \varphi) \in \mathbf{V}$  and using (S3.1) again yields

$$\begin{aligned} II &\leq C \sum_{S \in \mathcal{S}_{A_k}} h_S \|\nabla \varphi\|_{L^2(S)} \|D^2(\omega \varphi)\|_{L^2(S)} \\ &\leq C \sum_{S \in \mathcal{S}_{A_k}} h_S \|\nabla \varphi\|_{L^2(S)} (r^{-2} \|\varphi\|_{L^2(S)} + r^{-1} \|\nabla \varphi\|_{L^2(S)}) \\ &\leq Chr^{-1} \|\nabla \varphi\|_{L^2(A_k)}^2 + Cr^{-2} \|\varphi\|_{L^2(A_k)}^2. \end{aligned}$$

Hence,

$$\|\nabla \varphi\|_{L^2(B_{r^*})}^2 \leq Chr^{-1} \|\nabla \varphi\|_{L^2(A_k)}^2 + Cr^{-2} \|\varphi\|_{L^2(A_k)}^2, \quad k \geq 1.$$

We now apply this inequality iteratively for  $1 \leq k \leq K$  to arrive at

$$\|\nabla \varphi\|_{L^2(B_r)} \leq (Chr^{-1})^{K/2} \|\nabla \varphi\|_{L^2(B_{Kr})} + Cr^{-1} \|\varphi\|_{L^2(B_{Kr})}.$$

Since there is no a priori structure of  $\mathcal{S}$ , triangles in  $B_{Kr}$  may have very disparate sizes. The best we can say then is

$$\|\nabla \varphi\|_{L^2(B_{Kr})} \leq C\rho^{-1} \|\varphi\|_{L^2(B_{Kr})} \leq Ch^{-1/\gamma} \|\varphi\|_{L^2(B_{Kr})},$$

as results from a standard local inverse inequality [2, p. 140]. Let  $\Gamma$  satisfy  $\log(C\Gamma) \geq 2(1 - \gamma)/\gamma$ , provided  $\gamma < 1$ . Then, for  $K = \lceil \log h \rceil$ , we have

$$(Chr^{-1})^{K/2} h^{-1/\gamma} \leq \Gamma r^{-1},$$

which in turn implies the assertion. In case  $\gamma = 1$ , two iterations suffice, namely  $K = 2$ . This concludes the proof.  $\square$

$\varphi_i := \varphi(x_i^n)$ ,  $\alpha_i := \alpha(\varphi_i)$ , for  $1 \leq i, j \leq J := J^n$ . The fact that  $S$  is weakly acute is reflected in

$$k_{ii} > 0, \quad k_{ij} \leq 0 \quad (j \neq i), \quad \sum_{j=1}^J k_{ij} \geq 0, \quad \forall 1 \leq i \leq J;$$

so the stiffness matrix  $\mathbf{K} = (k_{ij})_{i,j=1}^J$  is an M-matrix. Hence,

$$\begin{aligned} \langle \nabla \varphi, \nabla \Pi \alpha(\varphi) \rangle &= \sum_{i,j=1}^J \alpha_i \varphi_j k_{ij} = \sum_{i=1}^J \left( \sum_{j \neq i} \alpha_i \varphi_j k_{ij} + \alpha_i \varphi_i k_{ii} \right) \\ &= \sum_{i=1}^J \left( \sum_{j \neq i} \alpha_i (\varphi_j - \varphi_i) k_{ij} + \alpha_i \varphi_i (k_{ii} + \sum_{j \neq i} k_{ij}) \right). \end{aligned}$$

Using the symmetry of  $\mathbf{K}$ , i.e.  $k_{ij} = k_{ji}$ , we observe that

$$\sum_{i=1}^J \sum_{j \neq i} \alpha_i (\varphi_j - \varphi_i) k_{ij} = \sum_{j=1}^J \sum_{i \neq j} \alpha_j (\varphi_i - \varphi_j) k_{ij},$$

and, as a consequence, that

$$\begin{aligned} L_\alpha(\nabla \varphi, \nabla \Pi \alpha(\varphi)) &= L_\alpha \sum_{i=1}^J \left( \sum_{j \neq i} \frac{1}{2} (\alpha_i - \alpha_j) (\varphi_j - \varphi_i) k_{ij} + \alpha_i \varphi_i (k_{ii} + \sum_{j \neq i} k_{ij}) \right) \\ &\geq \sum_{i=1}^J \left( \sum_{j \neq i} -\frac{1}{2} (\alpha_i - \alpha_j)^2 k_{ij} + \alpha_i^2 (k_{ii} + \sum_{j \neq i} k_{ij}) \right) = \sum_{i,j=1}^J \alpha_i \alpha_j k_{ij} = \|\nabla \Pi \alpha(\varphi)\|_{L^2(\Omega)}^2, \end{aligned}$$

because  $\alpha(0) = 0$ . The proof is thus complete.  $\square$

**S3. Pointwise Error Estimates.** Our goal now is to prove Lemma 7.1. Our approach is based on the method of local  $H^1$ -error estimates [21,22,23]. The crucial step here is a new version of the energy (or Caccioppoli) estimate of Lemma S3.1 which accounts for the lack of quasi-uniformity and improves upon the argument introduced in [21]. The only assumption on the triangulation, besides regularity, is  $h_n \leq C\rho_h^n$  with  $0 < \gamma \leq 1$ .

Let us start by adding some useful notation. Set  $h := h_n, \rho := \rho_n, v := G\psi, v_h := G^n\psi$  and drop the super(sub)script  $n$  in the various symbols already used. In this spirit,  $\{X_j\}_{j=1}^J$  and  $\{x_j\}_{j=1}^J$  stand for the basis functions and nodes of  $\mathcal{S}$ , respectively, whereas  $\mathbf{V}$  denotes the finite element space. Let  $B_r$  indicate the ball of radius  $r \geq \Gamma h$  ( $\Gamma > 1$  to be selected) and center  $x_0 \in \Omega$ . Set  $\tilde{r} := r \lceil \log h \rceil$ . Let  $\omega$  denote a cutoff function with the following properties:

$$(S3.1) \quad \omega = 1 \text{ in } \tilde{B}_{Kr}, \quad \omega = 0 \text{ in } \Omega_{\tilde{r}(k+1)r}, \quad |D^i \omega| \leq Cr^i \text{ in } A_k := \tilde{B}_{k^*} \setminus \tilde{B}_{(k+1)r},$$

because  $\Pi v - v_h \in \mathbf{V}$ . Moreover, on using Lemma S3.2, we obtain

$$\begin{aligned} \|\Pi v - v_h\|_{H^1(S_0)} &\leq \|v - \Pi v\|_{H^1(B_r)} + \|v - v_h\|_{H^1(B_r)} \\ &\leq C|\log h|\|\nabla(v - \Pi v)\|_{L^2(B_{2r})} + Cr^{-1}\|v - v_h\|_{L^2(B_{2r})}. \end{aligned}$$

We claim that the rightmost term satisfies

$$(S3.5) \quad \|v - v_h\|_{L^2(B_{2r})} \leq Ch^2 |\log h|^6 \|\nabla(v - \Pi v)\|_{L^\infty(\Omega)}.$$

Let us assume, for the moment, that (S3.5) is valid and complete the proof of Lemma 7.1.

We clearly have

$$\|v - v_h\|_{L^\infty(\Omega)} \leq Ch |\log h|^6 \|\nabla(v - \Pi v)\|_{L^\infty(\Omega)} \leq Ch^{2-2/p} |\log h|^6 \|v\|_{W^{2,p}(\Omega)}.$$

The fact that  $\partial\Omega \in C^{1,1}$  results in  $\|v\|_{W^{2,p}(\Omega)} \leq Cp\|\psi\|_{L^\infty(\Omega)}$ ; the precise dependence on  $p$  follows from tracing constants in the singular integrals involved. Then we choose  $p = |\log h|$  in the above inequality to get the desired estimate.

To complete the argument we must prove (S3.5). To this end, we will use the Green's function  $g$  and the local energy estimates. Let  $\varphi \in C_0^\infty(B_{2r})$  and  $\zeta \in H_0^1(\Omega)$  satisfy  $\Delta\zeta = -\varphi$ . Let  $\zeta_h \in \mathbf{V}$  be the  $H^1$ -projection of  $\zeta$  onto  $\mathbf{V}$ , namely,

$$(\nabla(\zeta - \zeta_h), \nabla\chi_j) = 0, \quad \forall 1 \leq j \leq J.$$

Since

$$(v - v_h, \varphi) = (\nabla(v - v_h), \nabla\zeta) = (\nabla(v - \Pi v), \nabla(\zeta - \zeta_h)) \leq \|\nabla(v - \Pi v)\|_{L^\infty(\Omega)} \|\nabla(\zeta - \zeta_h)\|_{L^2(\Omega)},$$

all we have to show is

$$(S3.6) \quad \|\nabla(\zeta - \zeta_h)\|_{L^2(\Omega)} \leq Ch^2 |\log h|^6 \|\varphi\|_{L^2(\Omega)}.$$

For  $1 \leq k \leq K$  ( $\leq C|\log h|$ ), set

$$\begin{aligned} A_k &:= \{x \in \Omega : 4h|\log h|^k \leq |x - x_0| < 4h|\log h|^{k+1}\}, \\ A_k^+ &:= \{x \in \Omega : \text{dist}(x, A_k) < 2h|\log h|^k\}, \end{aligned}$$

and  $A_0 := B_{4r}$ . We claim that

$$\|\nabla\zeta\|_{L^2(\Omega)} \leq C\tilde{r} |\log \tilde{r}|^{1/2} \|\varphi\|_{L^2(\Omega)}.$$

Indeed, using Hölder's inequality combined with a 2-D Sobolev inequality, we deduce that

$$\|\nabla\zeta\|_{L^2(\Omega)}^2 = \langle \zeta, \varphi \rangle \leq \|(\zeta - \zeta_h)\|_{L^2(\Omega)} \|\varphi\|_{L^2(\Omega)} \leq Cr^{2-1} h^q |\log h|^{1/2} \|\nabla\zeta\|_{L^2(\Omega)} \|\varphi\|_{L^2(\Omega)}.$$

**REMARK S3.1.** It is worth noting the contrast between a quasi-uniform mesh (i.e.,  $\gamma = 1$ ) and the general situation for which  $\tilde{r}/r = |\log h|$ . This fact causes the extra powers of the logarithm in Lemma 7.1.

We are now in a position to derive our basic tool, namely, the local  $H^1$ -error estimates.

**LEMMA S3.2.**  $\|v - v_h\|_{H^1(B_r)} \leq C|\log h|\|\nabla(v - \Pi v)\|_{L^2(B_{2r})} + Cr^{-1}\|v - v_h\|_{L^2(B_{2r})}$ .

*Proof.* Let  $\omega$  be the cutoff function defined in (S3.1), this time subordinate to  $B_r$ . Set  $\tilde{v} := \omega v$  and define  $\tilde{v}_h \in \mathbf{V}$  to be the (local)  $H^1$ -projection of  $\tilde{v}$  onto  $\mathbf{V}$ , namely,

$$(\nabla(\tilde{v}_h - \tilde{v}), \nabla\chi_j) = 0, \quad \forall x_j \in \tilde{B}_{2r}^j \quad (1 \leq j \leq J); \quad \tilde{v}_h(x) = 0, \quad \forall x \in \Omega \setminus \tilde{B}_{2r}.$$

Then  $v_h - \tilde{v}_h$  is discrete harmonic in  $\tilde{B}_{2r}^j$  ( $\supset B_r$ ). Hence, (S3.3) yields

$$\|\nabla(v_h - \tilde{v}_h)\|_{L^2(B_r)} \leq Cr^{-1}\|v_h - \tilde{v}_h\|_{L^2(B_r)} \leq Cr^{-1}(\|v_h - v\|_{L^2(B_r)} + \|\tilde{v} - \tilde{v}_h\|_{L^2(B_r)}),$$

because  $\tilde{v} = v$  in  $B_r$ . Since  $\tilde{v} - \tilde{v}_h = 0$  on  $\partial\tilde{B}_{2r}^j$ , the Poincaré inequality implies

$$\begin{aligned} \|\tilde{v} - \tilde{v}_h\|_{L^2(B_r)} &\leq C\tilde{r}\|\nabla(\tilde{v} - \tilde{v}_h)\|_{L^2(B_{2r})} \\ &\leq C\tilde{r}\|\nabla(\omega v)\|_{L^2(B_{2r})} \leq C\tilde{r}(\|\nabla v\|_{L^2(B_{2r})} + \tilde{r}^{-1}\|v\|_{L^2(B_{2r})}). \end{aligned}$$

On the other hand, we obviously have

$$\|\nabla(\tilde{v} - \tilde{v}_h)\|_{L^2(B_r)} \leq \|\nabla\tilde{v}\|_{L^2(B_r)} \leq C(\|\nabla v\|_{L^2(B_{2r})} + \tilde{r}^{-1}\|v\|_{L^2(B_{2r})}).$$

Summarizing, we have obtained

$$\|v - v_h\|_{H^1(B_r)} \leq C|\log h|(\|\nabla v\|_{L^2(B_{2r})} + r^{-1}\|v\|_{L^2(B_{2r})}) + Cr^{-1}\|v - v_h\|_{L^2(B_{2r})}.$$

The assertion finally follows upon writing  $v - v_h = (v - \Pi v) - (v_h - \Pi v)$ .  $\square$

Let  $g(x, y)$  be the Green's function corresponding to the Laplace operator, with vanishing Dirichlet boundary condition, and having the singularity located at  $y \in \Omega$ . Since  $\partial\Omega \in C^{1,1}$ , it is well known that

$$(S3.4) \quad |D_x^2 g(x, y)| \leq C|x - y|^{-2}, \quad \forall x \neq y.$$

**PROOF OF LEMMA 7.1.** Let  $x_0 \in S_0$  be a point at which the maximum norm of  $v - v_h$  is attained. Let  $r = \Gamma h$ . Then

$$\|v - v_h\|_{L^\infty(\Omega)} \leq \|v - \Pi v\|_{L^\infty(S_0)} + \|\Pi v - v_h\|_{L^\infty(S_0)} \leq \|v - \Pi v\|_{L^\infty(S_0)} + C\|\Pi v - v_h\|_{H^1(S_0)},$$

closest corner to  $x$ ,  $0 < \omega_x \leq 2\pi$  be the corresponding internal angle at  $y_x$  and  $\sigma_x := \pi/\omega_x$ . Then, our technique, combined with that in [21], leads to the following *local* result

$$(S3.7) \quad \|v - v_h\|_{L^\infty(B_r)} \leq Ch \log h \|\nabla(v - \Pi v)\|_{L^\infty(B_{2r})} + Cr^{-1} \|v - v_h\|_{L^2(B_{2r})},$$

for  $r \geq Ch^{1-\varepsilon}$  ( $\varepsilon > 0$ ), provided  $h(x) \leq Ch|x - y_x|^{\max(0, 1-\sigma_x)}$ , where  $h(x)$  is the function introduced in Lemma 5.1. Details are omitted.

**S4. Generalizations.** So far we have assumed that triangles intersecting the discrete interface and “bad” triangles are kept fixed. We have also excluded the analysis of mushy regions and of variable conductivity as well as the treatment of polygonal domains. We now wish to investigate these issues in some detail. We will be using the notation of §§2.4.

**S4.1. Eliminating Fixed Triangles.** Recall that  $S_B$  is the set of all triangles on which either first or second derivatives exceed a certain tolerance, and that  $S_F$  are those traversed by the discrete interface  $F$ . All those elements were left fixed, which is nearly the best we can do in a general setting. We want to show that under the following assumptions on the behavior of the numerical method:

$$(S4.1) \quad d_S \leq C_1, \quad \forall S \in S_F \cup S_B,$$

$$(S4.2) \quad D_\varepsilon \leq C_2 \tau^{-1}, \quad \forall \varepsilon \in \mathcal{E}_B \cup \mathcal{E}_F,$$

$$(S4.3) \quad \text{length}(F) \leq C_3, \quad \text{meas}(\mathcal{B}) \leq C_4 \tau,$$

the constraint (4.8) can be removed, namely all elements of  $S$  become free. The entire mesh  $S$  can thus be regenerated, on replacing (4.2) and (4.3) by

$$(S4.4) \quad \hat{h}_\varepsilon := \max \left( \frac{\tau^{1/2}}{\lambda \tau, \mu_1 \frac{\tau^{1/2}}{D_\varepsilon^{1/2}}} \right), \quad \forall \varepsilon \in \mathcal{E}, \quad \hat{h}_S := \max \left( \lambda \tau, \mu_1 \frac{\tau^{1/2}}{d_S} \right), \quad \forall S \in S.$$

Hence, what really matters in defining both  $\hat{h}_\varepsilon$  and  $\hat{h}_S$  is local regularity because  $F$  does not appear explicitly in (S4.4). Condition (S4.1) excludes the formation of cusps but it is still quite reasonable in many circumstances of practical interest. Condition (S4.2) is also acceptable in light of (3.10) whereas (S4.3) is consistent with (4.11). In §8.3 we present a relevant example that fulfils (S4.1)-(S4.3) and compare the performance of this strategy and that of §4.

Let us start by proving the following crucial estimates:

$$(S4.5) \quad \|\Pi(\zeta(U)) - \hat{\Pi}(\zeta(U))\|_{L^p(\Omega)} \leq C \tau^{1/p}, \quad \forall 1 \leq p < \infty, \quad \zeta' \in W^{1,\infty}(\mathbf{R}),$$

$$(S4.6) \quad \|\Theta - \hat{\Theta}\|_{L^\infty(\Omega)} \leq C \tau, \quad \|\nabla(\Theta - \hat{\Theta})\|_{L^2(\Omega)} \leq C \tau^{1/2}.$$

Set

$$\hat{\mathcal{F}} := \cup \{\hat{S} \in \hat{S} : \hat{S} \cap S_F \neq \emptyset\}, \quad \hat{\mathcal{B}} := \cup \{\hat{S} \in \hat{S} : \hat{S} \cap S_B \neq \emptyset\}.$$

The asserted estimate follows from taking  $q = \|\log \bar{h}\|$ . Consequently,

$$\|\nabla(\zeta - \zeta_h)\|_{L^1(A_k)} \leq C \bar{h} \|\nabla(\zeta - \zeta_h)\|_{L^2(\Omega)} \leq C \bar{h}^2 \|\log h\|_{L^2(\Omega)} \leq Ch^2 \|\log h\|_{L^2(\Omega)}^2 \|\varphi\|_{L^2(\Omega)}.$$

We now examine the contribution due to  $A_k$  for  $1 \leq k \leq K$ . Let  $\tilde{B}_k := B(x_k, h \log h^{k-1})$  be a covering of  $A_k$  with the nonoverlapping property

$$1 \leq \sum_{\tilde{B}_k} \chi_{\tilde{B}_k}(x) \leq C \quad (\text{independent of } h \text{ and } k), \quad \forall x \in A_k,$$

where  $\chi_{\tilde{B}_k}$  stands for the characteristic function of  $\tilde{B}_k$ . Set  $B_k^+ := B(x_k, 2h \log h^k)$ . As a consequence of Lemma S3.2, we see that

$$\|\nabla(\zeta - \zeta_h)\|_{L^1(B_k)} \leq C \|\log h\| \|\nabla(\zeta - \Pi \zeta)\|_{L^2(B_k^+)} + C(h \log h^{k-1})^{-1} \|\zeta - \zeta_h\|_{L^2(A_k^+)},$$

Since  $\sum_k \chi_{\tilde{B}_k}(x) \leq C \|\log h\|^2$ , we infer that

$$\begin{aligned} \|\nabla(\zeta - \zeta_h)\|_{L^1(A_k)} &\leq Ch \|\log h\|^{k+1} \|\nabla(\zeta - \zeta_h)\|_{L^2(A_k)} \\ &\leq Ch \|\log h\|^{k+3} \|\nabla(\zeta - \Pi \zeta)\|_{L^2(A_k^+)} + C \|\log h\|^3 \|\zeta - \zeta_h\|_{L^2(A_k^+)}. \end{aligned}$$

Using the representation formula  $\zeta(x) = \int_{B_{2r}} g(x, y) \varphi(y) dy$  for  $x \in \Omega$ , in conjunction with (S3.4), we get

$$\|D^2 \zeta\|_{L^2(A_k^+)} \leq C(h \log h^k)^{-2} \text{meas}(A_k^+) \|\varphi\|_{L^1(B_{2r})} \leq C \|\log h\|^{2-k} \|\varphi\|_{L^2(\Omega)}.$$

Hence,

$$\|\nabla(\zeta - \Pi \zeta)\|_{L^2(A_k^+)} \leq Ch \|D^2 \zeta\|_{L^2(A_k^+)} \leq Ch \|\log h\|^{2-k} \|\varphi\|_{L^2(\Omega)}.$$

Finally, it only remains to add over  $k$  for  $1 \leq k \leq K$  ( $\leq C \|\log h\|$ ), where  $C$  is such that  $\{A_k\}_{k=0}^K$  form a covering of  $\Omega$ , and apply (7.6) to conclude the whole argument.  $\square$

**REMARK S3.2.** We have deliberately omitted the technical details involved in the analysis of the discrepancy between  $\Omega$  and its polygonal approximation, for which we refer to [23]. Lemma 7.1 can be viewed as a sharpening of the arguments in [21], where an interior  $O(h^{2-\varepsilon})$ -rate of convergence is proved for convex polygons. We do not know whether the extra powers of the logarithm occurring in our pointwise estimate are really necessary, but we stress that it is still quasi-optimal.

**REMARK S3.3.** Suppose  $\Omega$  is a rectangle, as in §8.3. It is easily seen, by means of reflection arguments, that the key properties  $\|\varphi\|_{W^{2,p}(\Omega)} \leq C p \|\psi\|_{L^\infty(\Omega)}$  and (S3.4) are still valid and so is Lemma 7.1. For general polygons let  $y_x \in \partial\Omega$ , for a given  $x \in \Omega$ , be the

because the number of mesh changes is bounded by  $O(\tau^{-1/2})$ . This ends the discussion.

**S4.2. MUSHY REGIONS.** It is well known for the continuous problem (2.8) that in the absence of initial mushy regions they can only develop when  $f \neq 0$  [20]. They are a striking manifestation of degeneracy in that the underlying PDE becomes  $u_t = f(0)$ , at least formally. Our objective now is to motivate the proper adjustments needed by the mesh selection algorithm to handle such a situation.

We allow the initial interface  $I_0$  to degenerate into a mushy region, namely  $\text{meas}(\mathcal{F}_0) > 0$ , at the expense of the further assumptions

$$(S4.7) \quad \partial I_0 \cap \Omega \quad \text{is a Lipschitz curve,}$$

$$(S4.8) \quad u_0 \in W^{1,\infty}(I_0).$$

To determine the initial mesh  $S^1$ , we only have to make sure that the initial refined region contains  $I_0$ . Hence, in view of (S4.7) and (S4.8), Lemma 5.4 still holds.

We distinguish between two opposite situations according to the role played by heat conduction. In the first case, heat reaches the interface  $\partial I(t)$  at  $x$  by conduction. Consequently, the flux satisfies  $\nabla \theta^{+,-}(x,t) = \llbracket \nabla \theta(x,t) \rrbracket \neq 0$  for  $x \in \partial I(t) \cap \Omega$  and (2.11) provides a useful expression for the interface velocity  $V(x)$ . Moreover, the mush is *contracting* at  $x$  in the sense that the adjacent phase moves toward the mushy region. In the other case, instead, the internal heat production (or consumption), represented by  $f$ , is solely responsible for the interface motion. Since there is no heat conduction at  $x \in \partial I(t) \cap \Omega$ , namely  $\nabla \theta^{+,-}(x,t) = \llbracket \nabla \theta(x,t) \rrbracket = 0$ , (2.11) no longer provides information on  $V(x)$ . It says, however, that enthalpy  $u$  is continuous at  $x \in \partial I(t) \cap \Omega$ , or equivalently  $\llbracket u(x,t) \rrbracket = 0$ , if  $V(x) \neq 0$ . The mush can either be contracting or expanding at  $x$ ; we say that the mushy region is *degenerate* at  $x$ . Both opposite situations can occur at the same time [16].

A *discrete mushy region*  $\mathcal{M}$  is defined as the union of all triangles  $S \in \mathcal{S}$  such that  $\Theta|_S$  vanishes. We say that  $\mathcal{M}$  is *degenerate* on a side  $e \subset \partial \mathcal{M}$  if

$$(S4.9) \quad d_S \leq \mu_4 h_e, \quad \forall S \in \mathcal{S}_e,$$

where  $\mu_4 > 0$  is a suitable constant. Let  $\partial_e \mathcal{M}$  denote the union of all those  $e$ 's and set  $\partial_e \mathcal{M} := \partial \mathcal{M} \setminus \partial_e \mathcal{M}$ , which stands for the nondegenerate part of  $\partial \mathcal{M}$ .

We can certainly infer that the evolution of a degenerate mushy boundary cannot be predicted by means of the Stefan condition, as we did before. Since that was our basic tool in constructing the new refined region  $\tilde{\mathcal{R}}$ , where the discrete interface is supposed to belong to, we may wonder whether the entire mush should lie in  $\tilde{\mathcal{R}}$ . We want to show that our algorithm is robust in the sense that the *structural* assumptions

$$(S4.10) \quad \|\nabla U\|_{L^1(\Omega)} \leq C \sum_{S \in \mathcal{S}} h_S^2 D_S \leq C, \quad \text{length}(\partial \mathcal{M}) \leq C,$$

Since  $h_S, h_{\mathcal{S}} = O(\tau)$  for all  $S \subset \tilde{\mathcal{F}} \cup \tilde{\mathcal{B}}$  and  $S \in \mathcal{S}_{\tilde{\mathcal{S}}}$ , we have  $\text{card}(\mathcal{S}_{\tilde{\mathcal{S}}}) = O(1)$ . Also, (S4.3) implies  $\text{meas}(\tilde{\mathcal{F}} \cup \tilde{\mathcal{B}}) \leq C\tau$ . Since  $U$  satisfies (6.2), we get

$$\|\Pi(\zeta(U) - \tilde{\Pi}(\zeta(U)))\|_{L^1(\tilde{\mathcal{F}} \cup \tilde{\mathcal{B}})} \leq C \|U\|_{L^\infty(\Omega)} \left( \text{meas}(\tilde{\mathcal{F}} \cup \tilde{\mathcal{B}}) \right)^{1/p} \leq C\tau^{1/p}.$$

This estimate, together with (5.10) in  $\Omega \setminus (\tilde{\mathcal{F}} \cup \tilde{\mathcal{B}})$ , yields (S4.5), which in turn is used with  $p = 1$  to demonstrate Lemma 6.3. The estimates in (S4.6) are the analogues of (5.10) and (5.11) for  $\zeta = \beta$ . We only have to consider  $S \subset \tilde{\mathcal{F}} \cup \tilde{\mathcal{B}}$ , as before. Note that Lemma 5.3, with  $V = U$  and  $\zeta = \beta$ , cannot be applied because of the lack of  $W^{2,\infty}$ -regularity of  $\beta(U)$  in  $\tilde{\mathcal{F}}$ . However we just modify the proof of Lemma 5.3, by splitting  $\Theta - \tilde{\Theta}$  as follows:

$$\Theta - \tilde{\Theta} = \Theta - \tilde{\Pi}\Theta + \tilde{\Pi}(\Pi\beta(U) - \beta(U)).$$

Using now Lemma 5.3, with  $V = \Theta$  and  $\zeta = \text{Identity}$ , combined with (S4.2), we get

$$\|\Theta - \tilde{\Pi}\Theta\|_{L^\infty(S)} + \|\nabla(\Theta - \tilde{\Pi}\Theta)\|_{L^2(S)} \leq Ch_{\mathcal{S}} \max_{e \in \mathcal{E}_S} h_e D_e \leq C\tau.$$

On the other hand, in view of (4.6) we have  $\delta_S = h_S d_S \leq Ch_S d_S$ , and thus (7.12) and standard interpolation estimates yield

$$\begin{aligned} \|\tilde{\Pi}(\Pi\beta(U) - \beta(U))\|_{L^\infty(S)} + \|\nabla \tilde{\Pi}(\Pi\beta(U) - \beta(U))\|_{L^2(S)} &\leq C \|\Pi\beta(U) - \beta(U)\|_{L^\infty(S)} \\ &\leq C \max_{S \in \mathcal{S}_{\tilde{\mathcal{S}}} \cap (\tilde{\mathcal{F}} \cup \tilde{\mathcal{B}})} h_S d_S + CL_{\tilde{\rho}} \rho^{-1} \max_{S \in \mathcal{S}_{\tilde{\mathcal{S}}} \cap (\tilde{\mathcal{F}} \cup \tilde{\mathcal{B}})} h_S^2 d_S^2 \leq C\tau. \end{aligned}$$

as results from (S4.1) and (S4.4). This proves the first estimate in (S4.6). For the second one, the fact that  $h_{\mathcal{S}} = O(\tau)$  in  $\tilde{\mathcal{F}} \cup \tilde{\mathcal{B}}$ , coupled with the two preceding estimates, yields

$$\|\nabla(\Theta - \tilde{\Theta})\|_{L^2(\tilde{\mathcal{F}} \cup \tilde{\mathcal{B}})}^2 \leq C \sum_{S \subset \tilde{\mathcal{F}} \cup \tilde{\mathcal{B}}} \|\nabla(\Theta - \tilde{\Theta})\|_{L^2(S)}^2 \leq C \sum_{S \subset \tilde{\mathcal{F}} \cup \tilde{\mathcal{B}}} h_S^2 \leq C \text{meas}(\tilde{\mathcal{F}} \cup \tilde{\mathcal{B}}) \leq C\tau.$$

The same argument leads to  $\|\nabla(\Theta - \tilde{\Theta})\|_{L^1(\tilde{\mathcal{S}})} \leq C\tau$  for all  $\tilde{e} \in \tilde{\mathcal{E}}$  such that  $S_{\tilde{e}} \cap (\tilde{\mathcal{F}} \cup \tilde{\mathcal{B}}) \neq \emptyset$ . The above estimates for  $\Theta - \tilde{\Theta}$  are the basic ingredients to demonstrate Lemmas 5.10 and 5.11, from which Lemmas 6.4 and 6.5 follow.

It only remains to comment upon the error estimate (7.10). The only novelty arises from treating term  $V$ . In view of (S4.5), we resort to a 2-D Sobolev inequality, with  $q = |\log \tau|$ , to conclude

$$\begin{aligned} |V| &\leq \sum_{n=1}^m \|\tilde{U}^{n-1} - \tilde{U}^{n-1}\|_{L^1(\Omega)} \|G_e^n\|_{L^q(\Omega)} \\ &\leq Cq^{1/2} \max_{1 \leq n \leq m} \|G_e^n\|_{H_0^1(\Omega)} \sum_{n=1}^m \|\tilde{U}^{n-1} - \tilde{U}^{n-1}\|_{L^1(\Omega)} \\ &\leq C\tau^{1/2-1/q} \max_{1 \leq n \leq m} \|e_n^*\|_{H^{-1}(\Omega)} \leq \eta \max_{1 \leq n \leq m} \|e_n^*\|_{H^{-1}(\Omega)} + C\tau^{-1} |\log \tau|, \end{aligned}$$

lead to stability and accuracy provided the new refined region  $\tilde{\mathcal{R}}$  contains the current mushy region  $\mathcal{M}$ . In other words, a degenerate mushy boundary may exit  $\tilde{\mathcal{R}}$  without causing troubles, simply because the enthalpy is regular.

We now would like to motivate the left estimate in (S4.10). To this end, let  $\tilde{U} := \tilde{U}^{n-2}$  and assume that such a bound is valid for  $\tilde{U}$ . Observe first that the assumption  $\text{length}(\partial\mathcal{M}) \leq C$ , together with  $D_S \leq Ch_S^{-1}$  for all  $S \in \mathcal{S}$ , leads to

$$\sum_{S \in \mathcal{S}_{\partial\mathcal{M}}} h_S^2 D_S \leq C \sum_{S \in \mathcal{S}_{\partial\mathcal{M}}} h_S \leq C \text{length}(\partial\mathcal{M}) \leq C.$$

Let now  $x_j \in \mathcal{M}$  satisfy  $\tilde{S}_j \subset \mathcal{M}$ . From (2.20), it is easily seen that  $U(x_j) = \tilde{U}(x_j) + \tau f(\beta(\tilde{U}(x_j)))$ . We next subtract this expression from a similar one for  $x_i$ , generic node adjacent to  $x_j$ , to arrive at

$$|U(x_i) - U(x_j)| \leq (1 + \tau L_f L_\beta) |\tilde{U}(x_i) - \tilde{U}(x_j)|,$$

where we have used (2.1) and (2.4). Hence,

$$\sum_S h_S^2 D_S \leq (1 + \tau L_f L_\beta) \sum_S h_S^2 \tilde{D}_S \leq C,$$

where the summation is carried over all  $S \in \mathcal{S}$  having a vertex  $x_j$  as above. It only remains to consider  $S$ 's belonging to either the solid or liquid phase, where  $U$  and  $\Theta$  turn out to be equivalent (see (4.10)). By virtue of (6.6), which does not entail the absence of mush, we conclude that

$$\sum_{S \in \mathcal{S}_{\Omega_1 \setminus \mathcal{M}}} h_S^2 D_S \leq C \sum_{S \in \mathcal{S}_{\Omega_1 \setminus \mathcal{M}}} h_S^2 d_S \leq C \|\nabla \Theta\|_{L^1(\Omega)} \leq C.$$

Our argument is not completely satisfactory, though, in that we cannot derive the global bound  $\|\nabla U\|_{L^1(\Omega)} \leq C$  depending on the initial enthalpy only, because  $\mathcal{S}$  is not translation invariant (it is not even quasi-uniform!). We refer to [11] for the continuous problem.

Assume (S4.1)-(S4.3) and (S4.10). The *mesh selection algorithm* of §4.2 is then modified as follows. The first test on the interface location is replaced by  $\partial_n \mathcal{M} \subset \mathcal{R}$ , whereas the other two, namely (4.6) and (4.7), remain unchanged; the last refers to  $\partial_n \mathcal{M}$  only. Thus  $\partial_d \mathcal{M}$  may exit  $\tilde{\mathcal{R}}$ . In addition to (4.2), (4.3) and (4.5), the last being enforced on  $\partial_n \mathcal{M}$  only, the following local parameter is used to determine the local meshsize within  $\mathcal{M}$ :

$$(S4.11) \quad \hat{h}_S := \max \left( \lambda \tau, \mu_S \tau / D_S^{1/3} \right), \quad \forall S \in \mathcal{S}_{\mathcal{M} \setminus \partial_n \mathcal{M}},$$

where  $\mu_S > 0$  is a suitable constant; set  $\hat{h}_\varepsilon := \infty$  for all  $\varepsilon \subset \mathcal{M}$ . If  $\mathcal{S}$  is to be discarded because a test of §4.2 fails, then all its elements are free, which means that  $\mathcal{S}$  can be

completely regenerated, but we impose the constraints (4.9) and  $\mathcal{M} \subset \tilde{\mathcal{R}}$ . Note that (S4.11) comes from combining  $\mathcal{M} \subset \tilde{\mathcal{R}}$  with (S4.10) and the fact that the local meshsize within  $\tilde{\mathcal{R}}$  is  $O(\tau)$ .

To prove stability, we have to verify (S4.5) and (S4.6). We first use Lemma 5.3, with  $V = U$ , for all  $\hat{S} \in \hat{\mathcal{S}}_{\mathcal{M}} \setminus \hat{\mathcal{S}}_{\partial_n \mathcal{M}}$ , where  $\hat{\partial}_n \mathcal{M} = \cup \{\hat{S} \in \hat{\mathcal{S}} : \hat{S}_S \cap \mathcal{S}_{\partial_n \mathcal{M}} \neq \emptyset\}$ . Since  $h_{\hat{S}} = O(\tau)$  for all  $\hat{S} \in \hat{\mathcal{S}}_{\mathcal{M}}$  and  $D_{\hat{S}} \leq Ch_{\hat{S}}^{-1}$ ,  $h_{\hat{S}} \leq Ch_S$  for all  $S \in \mathcal{S}_{\hat{S}}$ , (S4.10) yields

$$\begin{aligned} \|\Pi(\zeta(U) - \tilde{\Pi}(\zeta(U)))\|_{L^1(\mathcal{M} \setminus \partial_n \mathcal{M})} &\leq C \left( \sum_{S \in \mathcal{S}_{\mathcal{M}} \setminus \hat{\mathcal{S}}_{\partial_n \mathcal{M}}} h_S^2 \max_{S \in \mathcal{S}_{\hat{S}}} (h_S^2 D_S^p + h_S^{2p} D_S^p) \right)^{1/p} \\ &\leq C \tau^{1/p} \left( \sum_{S \in \mathcal{S}_{\mathcal{M}} \setminus \hat{\mathcal{S}}_{\partial_n \mathcal{M}}} h_S^2 D_S \right)^{1/p} \leq C \tau^{1/p}, \quad \forall 1 \leq p < \infty. \end{aligned}$$

This estimate, together with (5.10), (6.2) and property meas  $(\hat{\partial}_n \mathcal{M}) \leq C\tau$ , yields (S4.5). In deriving (S4.6) we only have to examine what happens on  $\partial_d \mathcal{M}$ , because  $\Theta = 0$  in  $\mathcal{M}$ . In view of (S4.9), (S4.6) follows as in §S4.1, for instance.

Regarding accuracy, novelties arise from treating terms  $II_S$  and  $III$  because properties  $h_S = O(\tau)$  for all  $S \in \mathcal{S}_{\tilde{\mathcal{R}}}$  as well as  $F^{n-1}, F^n \subset \tilde{\mathcal{R}}^n$ , used in Theorem 7.1, are no longer valid. For term  $II_S$ , we observe first that  $\beta(U^n) = 0$  for all  $S \in \mathcal{S}_{\mathcal{M}^c}$  and next use that  $\text{length}(\partial \mathcal{M}^n) \leq C$  and  $d_S \leq C$  to conclude

$$\sum_{S \in \mathcal{S}_{\partial_n \mathcal{M}^c}} \|\beta(U^n) - \Pi^n \beta(U^n)\|_{L^1(S)} \leq \sum_{S \in \mathcal{S}_{\partial_n \mathcal{M}^c}} h_S^3 d_S \leq C\tau \text{length}(\partial \mathcal{M}^n) \leq C\tau.$$

Term  $III$  can be treated by virtue of property  $(U^n - \hat{U}^{n-1})(x) = O(\tau)$  for all  $x \in \mathcal{S}$  such that either  $\hat{S} \subset \mathcal{M}^n$  or  $S \in \mathcal{S}_{\partial_n \mathcal{M}^c}^n$  (use (2.20)). We finally proceed as in §S4.1, first to treat term  $V$ , and next to derive (7.10).

**S4.3. Variable Conductivity.** We now turn our attention to the more realistic situation of temperature-dependent conductivity. Let  $k : \mathbf{R} \rightarrow [1, \infty)$  satisfy

$$(S4.12) \quad k \in W^{1,\infty}(\mathbf{R}).$$

Since the discrete maximum principle was a fundamental tool in our stability analysis, we transform (1.1) so as to preserve such a property. This is accomplished as follows. Let  $\tilde{k} : \mathbf{R} \rightarrow [1, \infty)$  denote a piecewise linear approximation of  $k$  that verifies

$$(S4.13) \quad \|k - \tilde{k}\|_{L^\infty(\mathbf{R})} \leq C\tau^{1/2}, \quad \|\tilde{k}\|_{W^{1,\infty}(\mathbf{R})} \leq C\|k\|_{W^{1,\infty}(\mathbf{R})}.$$

Let  $\tilde{K}, \tilde{K} : \mathbf{R} \rightarrow \mathbf{R}$  be the following Kirchhoff transformations:

$$(S4.14) \quad \tilde{K}(z) := \int^z k(s) ds, \quad \tilde{K}(z) := \int^z \tilde{k}(s) ds,$$

and set  $\alpha := K \circ \beta$ ,  $\tilde{\alpha} := \tilde{K} \circ \beta$ . Note that  $\tilde{K}$ , and so  $\tilde{\alpha}$ , are easy to evaluate in practice. Since (1.1) becomes

$$u_t - \Delta v = f(\theta), \quad \theta = \beta(u), \quad v = \tilde{K}(\theta) = \alpha(u),$$

we propose the following discrete scheme:

$$(S4.15) \quad \tau^{-1}(U^n - \tilde{U}^{n-1}, \chi)^n + (\nabla V^n, \nabla \chi) = (f(\tilde{\Theta}^{n-1}), \chi)^n, \quad \forall \chi \in V^n,$$

where  $V^n := \Pi^n[\tilde{\alpha}(U^n)] = \Pi^n[\tilde{K}(\Theta^n)]$ . It is easily seen that (S4.15) bears the same advantages and stability properties as (2.15), because  $\tilde{\alpha}$  and  $f \circ \tilde{K}^{-1}$  verify (2.1) and (2.4), respectively. The functions  $V^n$  and  $\tilde{V}^{n-1} := \Pi^n[\tilde{\alpha}(\tilde{U}^{n-1})]$  are to be considered in place of  $\Theta^n$  and  $\tilde{\Theta}^{n-1}$  in the stability estimates of §6. It only remains to analyze accuracy. More specifically, we just have to examine term  $II$  in (7.11), namely,  $II = \sum_{n=1}^m \int_{J_n} (v(t) - V^n, e_n^v) dt$ , which is further decomposed as follows:

$$\begin{aligned} II_1 + \dots + II_4 &:= \sum_{n=1}^m \int_{J_n} (v(t) - V^n, u(t^n) - u(t)) dt + \sum_{n=1}^m \int_{J_n} (\tilde{\alpha}(u(t)) - \tilde{\alpha}(U^n), u(t) - U^n) dt \\ &\quad + \sum_{n=1}^m \int_{J_n} (\tilde{\alpha}(U^n) - \Pi^n \tilde{\alpha}(U^n), e_u(t)) dt + \sum_{n=1}^m \int_{J_n} (K(\theta(t)) - \tilde{K}(\theta(t)), e_u(t)) dt. \end{aligned}$$

In view of (S4.13) and (2.5), we get

$$\begin{aligned} |II_4| &= \sum_{n=1}^m \int_{J_n} \left| \int_{J_n} (k(s) - \tilde{k}(s)) ds, e_u(t) \right| dt \\ &\leq \sum_{n=1}^m \int_{J_n} \|k(\theta(t)) - \tilde{k}(\theta(t))\|_{L^2(\Omega)} \|e_u(t)\|_{H^{-1}(\Omega)} dt \\ &\leq C\tau^{1/2} \|\nabla \theta\|_{L^1(0, \tau; L^2(\Omega))} \|e_u\|_{L^\infty(0, \tau; H^{-1}(\Omega))} \leq \eta \|e_u\|_{L^\infty(0, \tau; H^{-1}(\Omega))}^2 + C\eta^{-1}\tau. \end{aligned}$$

We next continue as in Theorem 7.1 to derive the rate of convergence (7.10) for the errors  $\tilde{K}(\theta(t)) - \Pi^n[\tilde{K}(\Theta^n)]$  and  $e_u$ . Finally, (S4.13) and (7.12) lead to the desired estimate for  $\|e_\theta\|_{L^2(0, T; L^2(\Omega))}$ .

It is worth noting that, by virtue of (6.2), the approximation preprocessor (S4.13) must be carried out on a bounded interval only.

**S4.4. Polygonal Domains.** The only limitation of our approach to handle polygonal domains arises in the proof of Theorem 7.1 and, more precisely, in the analysis of term  $IV$  in (7.11). In fact, we can no longer use a global pointwise estimate because of the corner singularities.

For convex domains we split  $IV$  into two terms according to what happens within  $\mathcal{R}$  and outside. The second contribution is easy to tackle in view of (6.7) and (7.6). The first one, instead, requires the use of (S3.7) with  $r$  being a fixed number less than the smallest distance between corners and the center  $z_0$  of  $B_r$ , verifying  $x_0 \in \mathcal{R}$ . Suppose  $0$  is a corner and set  $\sigma := \pi/\omega$ , where  $\omega < \pi$  is the internal angle. The singularity of  $v = G\psi$  is of the form  $v(x) \approx |x|^\sigma$  for  $\sigma \neq$  integer; otherwise a logarithm occurs [7]. Since

$$\|v - v_h\|_{L^\infty(B_r)} \leq C\tau^{1/2} |\log \tau|^6 \|\nabla(v - \Pi v)\|_{L^\infty(B_{2r})} + C\tau,$$

and we need  $\|v - v_h\|_{L^\infty(B_r)} \leq C\tau |\log \tau|^7 \|\psi\|_{L^\infty(\Omega)}$ , we only have to consider  $1 < \sigma < 2$ . Assume

$$(S4.16) \quad \text{dist}(0, B_{2r}) \geq C \max(\tau, \tau^{1/2(2-\sigma)}).$$

Then, the local meshsize  $h(x)$  introduced in Lemma 5.1 must satisfy

$$\lambda\tau \leq h(x) \leq C\tau^{1/2} |x|^{2-\sigma}, \quad \forall x \in B_{2r},$$

which is possible because  $|x| \geq C\tau^{1/2(2-\sigma)}$ . When (S4.16) is violated, or equivalently  $\mathcal{R}$  comes too close to a corner, (S3.7) does not seem to be good enough as compared to what interpolation theory suggests. This is due to the appearance of  $h = O(\tau^{1/2})$  on the right-hand side rather than the local meshsize. A result of such a flavor was derived in [21,22] under the structural constraint that all triangles at about the same distance from a corner have nearly the same size, which may not be true in our setting. Such a case thus constitutes an open problem.

For concave polygonal domains, we need further refinements near the concave corners to eliminate the pollution effect in  $H^1$ , namely  $h(x) \approx \tau^{1/2} |x|^{-\sigma+\epsilon}$  for  $1/2 \leq \sigma < 1$ ,  $\epsilon > 0$  [7], and assume that  $\mathcal{R}$  remains far away from the corners. We can therefore apply (7.6) and (S3.7) to prove Theorem 7.1. The general situation is also an open question.