OPTIMAL ERROR ESTIMATES FOR SEMIDISCRETE PHASE RELAXATION MODELS*

XUN JIANG† AND RICARDO H. NOCHETTO‡

Abstract. This paper examines and compares semi-implicit, semi-explicit and extrapolation time-discretizations of a simple phase relaxation model with small parameter ε. The model consists of a diffusion-advection-reaction PDE for temperature coupled with an ODE with double obstacle ± 1 for phase variable. Sharp interfaces are thereby replaced by thin transition layers of thickness $\mathcal{O}(\sqrt{\varepsilon})$. As time-step $\tau \downarrow 0$, the semi-implicit and extrapolation schemes are shown to converge with optimal orders $\mathcal{O}(\tau)$ for temperature and enthalpy, and $\mathcal{O}(\sqrt{\tau})$ for heat flux, irrespective of ε, provided $\tau \leq \varepsilon/2$ for the extrapolation scheme. The second scheme may be viewed as a linearization of the first. For the semi-explicit counterpart, which is also a linearization subject to the stability constraint $\tau \leq \varepsilon$, these orders are further multiplied by an extra factor $1/\sqrt{\varepsilon}$, and are sharp. The results for the semi-implicit scheme are preserved in the singular limit $\varepsilon \downarrow 0$, namely the Stefan problem with temperature-dependent convection and reaction.

Key words. phase relaxation, double obstacle, error estimate, semi-implicit, semi-explicit, extrapolation

AMS subject classifications. 35K65, 35K85, 35R35, 65M15

1. Introduction. In spite of numerical evidence, it was long believed that the rate of convergence of the backward Euler method for *degenerate* parabolic problems could be at most $\mathcal{O}(\sqrt{\tau})$, $\tau > 0$ being the uniform time step. This order was proven by Crandall and Liggett for nonlinear semigroups of contractions in Banach spaces [5]. A striking recent result of Rulla [12] asserts that such a rate is in fact $\mathcal{O}(\tau)$ in Hilbert spaces, provided the underlying nonlinear (possibly multivalued) operator $\mathcal{A}$ is a *subgradient* and the initial datum u_0 belongs to its domain $\mathcal{D}(\mathcal{A})$. A relevant example is the Stefan problem

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

$$\theta = \beta(u), \tag{1.2}$$

with $\beta(s) := \min(s+1, 0) + \max(s-1, 0)$, that models heat transfer in a body undergoing phase change at a prescribed melting temperature. The key property used in [12] in estimating the error incurred by the implicit time discretization (backward Euler method)

$$\tfrac{1}{\tau}(U^n - U^{n-1}) - \Delta\beta(U^n) = 0, \tag{1.3}$$

*This work was partially supported by NSF Grant DMS-9305935

†Department of Mathematics, University of California, Davis, CA 95616, USA.

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

is that $\mathcal{A}u := -\Delta\beta(u)$ is the subgradient in $H^{-1}(\Omega)$ of the convex lower semicontinuous function $\varphi : H^{-1}(\Omega) \to \mathbb{R}$ defined by $\varphi(u) := \int_\Omega \int_0^{u(x)} \beta(s)\,ds\,dx$ if $u \in L^2(\Omega)$ and $+\infty$ otherwise [4,p.123]. Cleverly combined with the underlying Hilbert structure and the fact that $\|\nabla\beta(u(t))\|^2_{L^2(\Omega)}$ is a Liapunov functional for (1.1)-(1.2), this yields *optimal* error estimates of order $\mathcal{O}(\tau)$ for (1.3) without resorting to second time derivatives [12], which in fact do not exist for (1.1). The relation between approximability and regularity is further explored in [13], where u is shown to possess an almost 3/2 derivative in time with values in $H^{-1}(\Omega)$. Since u is discontinuous, such a regularity result turns out to be sharp.

The Stefan problem is the simplest solid-liquid phase transition in that it presumes constant melting temperature θ, say $\theta = 0$, as expressed by the constitutive relation (1.2). The sign of θ thus determines the phase. Upon introducing the phase variable $\chi := u - \theta$, we can rewrite (1.2) as $\chi \in \operatorname{sign}(\theta) = \Lambda^{-1}(\theta)$, or equivalently as

$$\theta \in \Lambda(\chi) := \begin{cases} [-\infty, 0], & \chi = -1 \\ 0, & -1 < \chi < 1 \\ [0, +\infty], & \chi = +1, \end{cases} \tag{1.4}$$

where $\in$ is interpreted in the sense of graphs. The fact that χ can be used as a phase indicator (order parameter) is clearly stated in (1.4) and in fact is more natural than θ. Equation (1.1) now becomes

$$\partial_t\theta + \partial_t\chi - \Delta\theta = f, \tag{1.5}$$

and together with (1.4) constitutes a parabolic system with initial conditions $\theta = \theta_0$ and $\chi = \chi_0$. Experiments, however, often suggest violation of $\theta \in \Lambda(\chi)$ in melting processes of interest. To incorporate these superheating and undercooling effects, without surface tension, Visintin proposed a simple model of phase relaxation [17,18]. The equilibrium condition (1.4) is replaced by the following dynamic relation

$$\varepsilon\partial_t\chi + \Lambda(\chi) \ni \theta, \tag{1.6}$$

where $\varepsilon > 0$ is a small relaxation parameter. Since (1.4) is the formal limit of (1.6), we expect the resulting problem to converge to (1.1) as $\varepsilon \downarrow 0$. This was proven in [16] together with a suboptimal error estimate of order $\mathcal{O}(\tau^{1/4})$. The sharp interface model (1.1)-(1.2) is thus replaced by a diffuse one exhibiting a thin transition layer of size $\mathcal{O}(\sqrt{\varepsilon}\,)$ [10], where the phase change takes place, and independent thermodynamic variables θ and χ. Even though system (1.5)-(1.6) is smoother than (1.1)-(1.2), it is still strongly nonlinear with regularity deteriorating as $\varepsilon \downarrow 0$.

In this paper we examine the combined effect of phase relaxation and time discretization for a heat transfer equation with temperature-dependent convection and reaction

$$\partial_t\theta + \partial_t\chi - \Delta\theta = \operatorname{div}\mathbf{b}(\theta) + f(\theta) =: F(\theta). \tag{1.7}$$

We impose a vanishing Dirichlet boundary condition for θ, along with initial conditions θ_0 and χ_0 for θ and χ. Our discussion applies also to other *degenerate* parabolic problems. Petroleum reservoir and groundwater diffusion modeling typically lead to equations like (1.7) for a modified phase saturation [1]. The modeling of reactive solute transport with an equilibrium adsorption process also gives rise to (1.7) [3]. In both cases the relation between u and $\theta = \beta(u)$ is monotone and degenerate ($\beta'(u)$ vanishes for some u's). Since the Stefan problem is certainly more extreme in that $\beta'(u)$ vanishes in an entire interval, our results will extend to these cases as well. The implicit Euler's scheme reads

$$\frac{1}{\tau}(X^n - X^{n-1}) + \frac{1}{\tau}(\Theta^n - \Theta^{n-1}) - \Delta\Theta^n = F(\Theta^n), \tag{1.8}$$

$$\frac{\varepsilon}{\tau}(X^n - X^{n-1}) + \Lambda(X^n) \ni \Theta^n, \tag{1.9}$$

and falls into the general framework of [12] solely when $F = 0$, as explained in §3. In this restrictive case, [12] gives rise to a linear rate of convergence $\mathcal{O}(\tau)$, independent of ε, provided θ_0 and χ_0 satisfy a compatibility condition; see (2.4). If $F \neq 0$ the question arises as to whether a similar result holds. Its affirmative answer is discussed in §6. Being coupled, the above system leads to a strongly nonlinear algebraic system upon space discretization, which requires an iterative solver and is thus computationally inconvenient.

We intend to study several means of linearization. We first introduce a *semi-implicit* time stepping for treating the convection and reaction terms in (1.7), namely

$$\frac{1}{\tau}(X^n - X^{n-1}) + \frac{1}{\tau}(\Theta^n - \Theta^{n-1}) - \Delta\Theta^n = F(\Theta^{n-1}). \tag{1.10}$$

We consider next the following three time discretizations of (1.6) with $\Theta^{-1} = \Theta^0$:

$$\frac{\varepsilon}{\tau}(X^n - X^{n-1}) + \Lambda(X^n) \ni T^n := \begin{cases} \Theta^n & \text{semi-implicit} \\ \Theta^{n-1} & \text{semi-explicit} \\ 2\Theta^{n-1} - \Theta^{n-2} & \text{extrapolation.} \end{cases} \tag{1.11}$$

Both the semi-explicit and extrapolation methods decouple X^n and Θ^n, thereby yielding linear problems. In fact (1.11) is merely an algebraic correction that produces X^n, and therefore (1.10) becomes a *linear* elliptic PDE for Θ^n. However for these schemes to be stable a constraint must be imposed on the ratio τ/ε. Stability is studied in §4, and rates of convergence derived in the subsequent sections. If

$$\begin{aligned} E(\varepsilon,\tau) &:= \max_{1\le n\le N} \Big(\|u(t^n) - U^n\|_{H^{-1}(\Omega)} + \sqrt{\varepsilon}\,\|\chi(t^n) - X^n\|_{L^2(\Omega)} \Big) \\ &+ \Big(\sum_{n=1}^N \int_{I^n} \tau\varepsilon \|\partial_t \chi - \partial X^n\|^2_{L^2(\Omega)}\, dt \Big)^{\frac{1}{2}} \\ &+ \Big(\sum_{n=1}^N \int_{I^n} \big(\|\theta - \Theta^n\|^2_{L^2(\Omega)} + \tau \|\theta - \Theta^n\|^2_{H^1_0(\Omega)} \big)\, dt \Big)^{\frac{1}{2}} \end{aligned} \tag{1.12}$$

denotes the approximation error, then Table 1.1 summarizes our results. The semidiscrete traveling waves of [10] reveal the resulting order for the semi-explicit method to be the best possible; it improves upon those in [15,16] indeed. A new space discretization of this scheme using piecewise linear finite elements for both Θ^n and X^n is further analyzed in [6]. When $\tau = \varepsilon$ and $F = 0$ the semi-explicit problem coincides with the so-called *nonlinear Chernoff formula* studied in [7]; see also the surveys [9,14]. The somewhat disturbing factor $1/\sqrt{\varepsilon}$ can be eliminated via extrapolation, but at the expense of keeping two consecutive time iterates, namely Θ^{n-1} and Θ^{n-2}, and a more restrictive stability constraint.

Method	PDE	Stability	$E(\tau,\varepsilon)$
semi-implicit	nonlinear		$\mathcal{O}(\tau)$
semi-explicit	linear	$\tau \le \varepsilon$	$\mathcal{O}(\tau/\sqrt{\varepsilon})$
extrapolation	linear	$\tau \le \varepsilon/2$	$\mathcal{O}(\tau)$

Table 1.1: Summary of Results

This paper is organized as follows. In §2 we discuss weak formulations of the various semidiscretizations. We then show how Rulla's general framework applies to the implicit scheme (1.8)–(1.9) for the simplest case $F = 0$. New strong stability estimates for all three semidiscrete schemes are derived in §4. They are crucial for the error analyses of the remaining sections. In §5 we examine the semi-explicit scheme. We discuss the semi-implicit scheme in §6, along with intermediate regularity for $\partial_t \chi$ and $\nabla\theta$. The latter is crucial to study the extrapolation scheme in §7. We conclude in §8 with the semi-implicit discretization of the Stefan problem with temperature-dependent convection and reaction.

2. Basic Setting. Let Ω be a bounded domain of $\mathbb{R}^d$ ($d \ge 1$) with a $C^{0,1}$ boundary, occupied by liquid and solid of a certain material, and set $Q := \Omega \times (0,T)$ where $0 < T < +\infty$ is fixed. The variational formulations are weak forms of the classical equations (1.5)-(1.6) that also include boundary conditions. Hereafter, the symbol $\langle\cdot,\cdot\rangle$ will indicate the L^2 pairing and ocassionally the duality pairing between $H_0^1(\Omega)$ and $H^{-1}(\Omega)$. The variational formulation of (1.5)-(1.6) reads: *Find* $\theta \in L^2(0,T;H_0^1(\Omega)) \cap H^1(0,T;L^2(\Omega))$, $\chi \in H^1(0,T;L^2(\Omega))$ *and* $z \in L^2(0,T;L^2(\Omega))$ *such that* $z \in \Lambda(\chi)$ *a.e. in* Q *and for all* $\phi \in H_0^1(\Omega)$ *and* $\varphi \in L^2(\Omega)$ *the following holds*

$$\langle \partial_t\theta + \partial_t\chi, \phi\rangle + \langle\nabla\theta,\nabla\phi\rangle = -\langle \mathbf{b}(\theta),\nabla\phi\rangle + \langle f(\theta),\phi\rangle \quad a.e.\ in\ (0,T), \tag{2.1}$$

$$\varepsilon\langle\partial_t\chi,\varphi\rangle + \langle z,\varphi\rangle = \langle\theta,\varphi\rangle \qquad a.e.\ in\ (0,T), \tag{2.2}$$

$$\theta(\cdot,0) = \theta_0 \quad \chi(\cdot,0) = \chi_0 \qquad a.e.\ in\ \Omega, \tag{2.3}$$

where $\theta_0 \in H_0^1(\Omega)$ *and* $-1 \le \chi_0 \le 1$. The functions $\mathbf{b} : Q \times \mathbb{R} \to \mathbb{R}^d$ and $f : Q \times \mathbb{R} \to \mathbb{R}$ are assumed to be Lipschitz continuous for all $(x,t) \in Q$:

$$\big|\mathbf{b}(\theta_1) - \mathbf{b}(\theta_2)\big| \le L_{\mathbf{b}}\big|\theta_1 - \theta_2\big|, \qquad \big|f(\theta_1) - f(\theta_2)\big| \le L_f\big|\theta_1 - \theta_2\big|.$$

Dependence on $(x,t) \in Q$ will not be made explicit. Sometimes we will abuse notation and write $\Lambda(\chi)$ instead of z as the symbol $\Lambda(\chi)$ is more suggestive.

The study of traveling waves of [10] reveals that the transition region $\{|\chi(\cdot,t)| < 1\}$ is a layer of thickness $\mathcal{O}(\sqrt{\varepsilon}\,)$ in which $|\theta| \le C\varepsilon^{1/2}$, and that it is precisely there where the condition $\theta \in \Lambda(\chi)$ is violated, thereby leading to superheating effects. It is therefore quite reasonable to assume the *compatibility condition* between θ_0 and χ_0

$$\int_{\{\chi_0=1\}} \big|\min(\theta_0,0)\big|^2 + \int_{\{\chi_0=-1\}} \big|\max(\theta_0,0)\big|^2 + \int_{\{|\chi_0|<1\}} \big|\theta_0\big|^2 \le A^2\,\varepsilon, \tag{2.4}$$

for $A > 0$ given. To express (2.4) in a more convenient form, we define $z_0 \in \Lambda(\chi_0)$ by

$$z_0 := \begin{cases} \theta_0 - \max(\theta_0,0), & \text{if } \chi_0 = -1 \\ 0, & \text{if } -1 < \chi_0 < 1 \\ \theta_0 - \min(\theta_0,0), & \text{if } \chi_0 = 1. \end{cases} \tag{2.5}$$

Then (2.4) reads equivalently

$$\|\theta_0 - z_0\|_{L^2(\Omega)} \le A\sqrt{\varepsilon}. \tag{2.6}$$

Existence and uniqueness of solutions of (1.5) and (1.6) are a consequence of monotonicity arguments [17], along with the typical a priori estimates for degenerate parabolic problems:

$$\begin{aligned} &\|\theta\|_{L^\infty(0,T;H^1(\Omega))} + \|\partial_t u\|_{L^\infty(0,T;H^{-1}(\Omega))} \\ &\qquad + \sqrt{\varepsilon}\|\partial_t\chi\|_{L^\infty(0,T;L^2(\Omega))} + \|\theta\|_{H^1(0,T;L^2(\Omega))} \le C. \end{aligned} \tag{2.7}$$

Here C depends on $\|\nabla\theta\|_{L^2(\Omega)}$ and A in (2.4). The symbol C will always indicate a positive constant which may vary at the various occurrences but is independent of the relevant parameters involved, namely ε, τ and δ (defined below). Bounds (2.7) can be derived via the following regularization argument, which will be useful throughout the paper.

We introduce the following Lipschitz monotone approximation of Λ

$$\Lambda_\delta(s) := \begin{cases} \frac{s+1}{\delta}, & s \le -1 \\ 0, & -1 < s < 1 \\ \frac{s-1}{\delta}, & s \ge 1, \end{cases}$$

and denote by $\theta_\delta, \chi_\delta$, and $u_\delta := \theta_\delta + \chi_\delta$ the classical solutions of the corresponding regularized problem with $\theta_\delta(\cdot,0) = \theta_0$ and $\chi_\delta(\cdot,0)$ to be chosen as follows. Set $\chi_\delta(\cdot,0) := \chi_0 + \delta z_0$ and note that $\Lambda_\delta(\chi_\delta(\cdot,0)) = z_0$ and, by virtue of (2.6),

$$\varepsilon^{1/2}\|\partial_t\chi_\delta(0)\|_{L^2(\Omega)} = \varepsilon^{-1/2}\|\theta_0 - z_0\|_{L^2(\Omega)} \le A. \tag{2.8}$$

Existence and uniqueness of θ_δ and χ_δ result from monotonicity arguments [17]. With (2.8) at hand, it is easy to see that (2.7) holds for $\theta_\delta, \chi_\delta$ and u_δ uniformly in δ, and thus

$$\|u_\delta(t) - u_\delta(0)\|^2_{H^{-1}(\Omega)} + \varepsilon\|\chi_\delta(t) - \chi_\delta(0)\|^2_{L^2(\Omega)} \leq C\tau^2 \qquad \forall\ t = \mathcal{O}(\tau). \tag{2.9}$$

Whenever the regularization procedure is used, a limit $\delta \downarrow 0$ will follow. Since passing to the limit $\delta \downarrow 0$ is standard in the theory of nonlinear PDEs, it will be always omitted. Since Λ_δ is Lipschitz, and both $\theta_\delta, \chi_\delta \in H^1(0,T; L^2(\Omega))$, we conclude that

$$\varepsilon\partial_t\chi_\delta = \theta_\delta - \Lambda_\delta(\chi_\delta) \in H^1(0,T;L^2(\Omega)) \subset C^0([0,T];L^2(\Omega));$$

$\|\partial_t\chi_\delta(t)\|_{L^2(\Omega)}$ is thus well defined for all $t \in [0,T]$, and so is $\|\nabla\theta_\delta(t)\|_{L^2(\Omega)}$. It makes sense then to differentiate the above equality with respect to t, multiply it by $\partial_t\chi_\delta$, and add the resulting expression with (2.1) for θ_δ, χ_δ and $\phi = \partial_t\theta_\delta$. After integration in time we end up with the following "monotonicity" property, which will play a key role in the sequel:

$$\alpha(t_2) - \alpha(t_1) \leq \int_{t_1}^{t_2} \left(\|\mathrm{div}\mathbf{b}(\theta_\delta)\|^2_{L^2(\Omega)} + \|f(\theta_\delta)\|^2_{L^2(\Omega)}\right)\, dt, \tag{2.10}$$

for all $0 < t_1 < t_2 \leq T$, where

$$\alpha(t) := \|\nabla\theta_\delta(t)\|^2_{L^2(\Omega)} + \varepsilon\|\partial_t\chi_\delta(t)\|^2_{L^2(\Omega)} \tag{2.11}$$

is a (regularized) Liapunov functional. Note that

$$\alpha(0) \leq \|\nabla\theta_0\|^2_{L^2(\Omega)} + A^2 \leq C, \tag{2.12}$$

and that $\alpha(t)$ is indeed *nonincreasing* if there is no convection nor reaction effects.

We next introduce the weak forms of the three time discretizations to be studied. Let $\tau > 0$ be the time step, $N := T/\tau$ be a positive integer, $t^n := n\tau$, and $I^n := (t^{n-1}, t^n]$ for $1 \leq n \leq N$. For any given sequence $\{y^n\}_{n=0}^N$, we set $\partial y^n := (y^n - y^{n-1})/\tau$. The time discretizations read as follows: *For any $1 \leq n \leq N$, find $\Theta^n \in H^1_0(\Omega), |X^n| \leq 1$ and $Z^n \in L^2(\Omega)$ such that $Z^n \in \Lambda(X^n)$ a.e. in Ω and*

$$\Theta^0 = \theta_0,\ \ X^0 = \chi_0, \tag{2.13}$$

$$\langle\partial\Theta^n,\phi\rangle + \langle\partial X^n,\phi\rangle + \langle\nabla\Theta^n,\nabla\phi\rangle = -\langle\mathbf{b}(\Theta^{n-1}),\nabla\phi\rangle + \langle f(\Theta^{n-1}),\phi\rangle, \tag{2.14}$$

$$\varepsilon\langle\partial X^n,\varphi\rangle + \langle Z^n,\varphi\rangle = \langle T^n,\varphi\rangle, \tag{2.15}$$

for all $\phi \in H^1_0(\Omega)$ and $\varphi \in L^2(\Omega)$, where T^n is defined in (1.11). In what follows, we sometimes write $\Lambda(X^n)$ in place of Z^n. Existence and uniqueness for the semi-explicit and extrapolation schemes are rather obvious in that (2.15) provides an explicit expression for X^n

$$X^n = (\mathbf{I} - \beta)(\tfrac{\tau}{\varepsilon} T^n + X^{n-1}),$$

and (2.14) thus becomes a coercive elliptic PDE for Θ^n. For the semi-implicit scheme these issues can be tackled as in [17] by resorting to monotone operator theory.

When Λ is replaced by Λ_δ we obtain the regularized semidiscrete problem and corresponding solutions Θ^n_δ, X^n_δ, and $U^n_\delta := \Theta^n_\delta + X^n_\delta$ satisfying $\Theta^0_\delta := \theta_0$ and $X^0_\delta := \chi_\delta(\cdot, 0)(= \chi_0 + \delta z_0)$. It will be convenient for our stability analysis to choose $\Theta^{-1}_\delta = \Theta^{-2}_\delta := \Theta^0_\delta$ and select X^{-1}_δ so that

$$\tfrac{\varepsilon}{\tau}(X^0_\delta - X^{-1}_\delta) + \Lambda_\delta(X^0_\delta) = T^0_\delta. \tag{2.16}$$

Then the constitutive relation $\varepsilon \partial X^n_\delta + \Lambda_\delta(X^n_\delta) = T^n_\delta$ holds for $n \geq 0$. Since $\Lambda_\delta(X^0_\delta) = z_0$, we have the semidiscrete analogue of (2.8):

$$\sqrt{\varepsilon}\|\partial X^0_\delta\|_{L^2(\Omega)} = \tfrac{\sqrt{\varepsilon}}{\tau}\|X^0_\delta - X^{-1}_\delta\|_{L^2(\Omega)} \leq A. \tag{2.17}$$

We finish this section by stating three elementary identities to be used below, which are valid for all a_n, $b_n \in \mathbb{R}^k$ ($k \geq 1$):

$$2\sum_{n=1}^{m} a_n(a_n - a_{n-1}) = |a_m|^2 - |a_0|^2 + \sum_{n=1}^{m}|a_n - a_{n-1}|^2, \tag{2.18}$$

$$2\sum_{n=1}^{m} a_n \sum_{i=1}^{n} a_i = \Big|\sum_{n=1}^{m} a_n\Big|^2 + \sum_{n=1}^{m}|a_n|^2, \tag{2.19}$$

$$\sum_{n=1}^{m} a_n(b_n - b_{n-1}) = a_m b_m - a_0 b_0 - \sum_{n=1}^{m}(a_n - a_{n-1})b_{n-1}. \tag{2.20}$$

Identity (2.19) is a consequence of (2.18) applied to $A_n = \sum_{i=1}^n a_i$. Identity (2.20) is a summation by parts formula.

3. Implicit Scheme without Convection and Reaction. In this section, we consider the implicit scheme for $F = 0$. To see that [12] applies to this semidiscrete model, we introduce the Hilbert space $\mathbf{V} := H^{-1}(\Omega) \times L^2(\Omega)$ endowed with the inner product

$$\langle (u, \chi), (v, \eta) \rangle_{\mathbf{V}} := \langle u, v \rangle_{H^{-1}(\Omega)} + \varepsilon \langle \chi, \eta \rangle_{L^2(\Omega)}.$$

The (mutivalued) operator $\mathcal{A} : D(\mathcal{A}) \to \mathbf{V}$ is next defined by

$$\mathcal{A}(u, \chi) := \big(-\Delta(u - \chi), \tfrac{1}{\varepsilon}(\Lambda(\chi) - (u - \chi))\big),$$

where $D(\mathcal{A})$ is the domain of $\mathcal{A}$, namely the subspace of $\mathbf{V}$ such that

$$\|\nabla(u-\chi)\|^2_{L^2(\Omega)} + \tfrac{1}{\varepsilon}\|\Lambda(\chi)-(u-\chi)\|^2_{L^2(\Omega)} < \infty.$$

Then $\mathcal{A}$ is the subgradient of the following convex lower semicontinous function $\varphi : \mathbf{V} \to R$

$$\varphi(u,\chi) := \int_\Omega \tfrac{1}{2}|u-\chi|^2 + L(\chi), \qquad L(\chi) := \begin{cases} \infty, & \text{if } |\chi| \geq 1 \\ 0, & \text{otherwise.} \end{cases}$$

Consequently

$$\partial_t(u,\chi) + \mathcal{A}(u,\chi) \ni (0,0),$$

and we immediately conclude that the implicit scheme (1.8)–(1.9) with initial data $\Theta^0 := \theta_0,\ X^0 := \chi_0$ gives rise to a rate of $\mathcal{O}(\tau)$ in $\mathbf{V}$ independent of ε, provided $u_0 := \theta_0 + \chi_0$ and χ_0 belong to the domain of $\mathcal{A}$, namely [12]

$$\|\mathcal{A}(u_0,\chi_0)\|^2_{\mathbf{V}} = \|\nabla\theta_0\|^2_{L^2(\Omega)} + \tfrac{1}{\varepsilon}\|\theta_0 - z_0\|^2_{L^2(\Omega)} \leq C.$$

This is just a restatement of the regularity assumption on θ_0 and the compatibility condition (2.8). Therefore [12] yields the following important result.

THEOREM 3.1. *There exists $C > 0$, depending on $\|\nabla\theta_0\|_{L^2(\Omega)}$ and A in (2.4) but independent of $m \leq N$, such that $E(\varepsilon,\tau) \leq C\tau$.*

This first order error estimate is clearly *optimal* for the implicit Euler method and the regularity stated in (2.7). This improves upon the estimate by Crandall and Liggett of $\mathcal{O}(\sqrt{\tau}\,)$ for nonlinear semigroups of contractions in Banach spaces [5]. Letting $\varepsilon \downarrow 0$, the above estimates coincide with those of [12,13] for the Stefan problem without convection and reaction.

4. Stability. In this section we improve upon the a priori estimates of [15,16,17,18], that are optimal under minimal regularity of the initial data. We deal with the regularized solutions Θ^n_δ and X^n_δ of the semi-explicit scheme and derive a priori estimates uniformly in δ. We in fact need these estimates for $\delta > 0$ in the subsequent sections. We suppress the subscript δ for all functions involved except, however, for Λ_δ. No confusion will arise.

LEMMA 4.1. *There exists a constant $C > 0$ depending on A in (2.4), $\|\nabla\theta_0\|_{L^2(\Omega)}$, $L_{\mathbf{b}}$ and L_f such that the following strong stability estimate holds for the semi-explicit scheme provided $\tau \leq \varepsilon$*

$$\begin{aligned}\sum_{n=1}^N \tau\|\partial\Theta^n\|^2_{L^2(\Omega)} &+ \max_{1\leq n\leq N}\|\nabla\Theta^n\|^2_{L^2(\Omega)} + \sum_{n=1}^N\|\nabla(\Theta^n - \Theta^{n-1})\|^2_{L^2(\Omega)} \\ &+ \varepsilon\max_{1\leq n\leq N}\|\partial X^n\|^2_{L^2(\Omega)} + \varepsilon\sum_{n=1}^N\|\partial X^n - \partial X^{n-1}\|^2_{L^2(\Omega)} + \sum_{n=1}^N \tau\big\langle\partial X^n, \partial\Lambda_\delta(X^n)\big\rangle \leq C.\end{aligned} \tag{4.1}$$

Proof. We take $\varphi = \Theta^n - \Theta^{n-1} \in H_0^1(\Omega)$ in (2.14) to obtain, after summation from $n = 1$ to $n = m \le N$

$$
\begin{aligned}
&\tau\sum_{n=1}^{m}\|\partial\Theta^n\|^2_{L^2(\Omega)} + \tau\sum_{n=1}^{m}\langle\partial X^n, \partial\Theta^n\rangle + \sum_{n=1}^{m}\langle\nabla\Theta^n, \nabla(\Theta^n - \Theta^{n-1})\rangle \\
&\qquad = -\tau\sum_{n=1}^{m}\langle\mathbf{b}(\Theta^{n-1}), \nabla(\partial\Theta^n)\rangle + \tau\sum_{n=1}^{m}\langle f(\Theta^{n-1}), \partial\Theta^n\rangle.
\end{aligned}
\tag{4.2}
$$

For the third term, we apply the elementary identity (2.18) to get

$$
\sum_{n=1}^{m}\langle\nabla\Theta^n, \nabla(\Theta^n - \Theta^{n-1})\rangle = \tfrac{1}{2}\sum_{n=1}^{m}\|\nabla(\Theta^n - \Theta^{n-1})\|^2_{L^2(\Omega)} + \tfrac{1}{2}\|\nabla\Theta^m\|^2_{L^2(\Omega)} - \tfrac{1}{2}\|\nabla\Theta^0\|^2_{L^2(\Omega)}.
$$

With the aid of the regularized analogue of (2.15) with $T^n = \Theta^{n-1}$, the function $\Theta^n - \Theta^{n-1}$ also reads

$$
\Theta^n - \Theta^{n-1} = \tau\partial\Theta^n - \tau\partial\Theta^{n-1} + \varepsilon\big(\partial X^n - \partial X^{n-1}\big) + \tau\partial[\Lambda_\delta(X^n)].
$$

Therefore we can decompose the second term in (4.2) as follows:

$$
\begin{aligned}
\tau\sum_{n=1}^{m}\langle\partial X^n, \partial\Theta^n\rangle &= \tau\sum_{n=1}^{m}\langle\partial X^n, \partial\Theta^n - \partial\Theta^{n-1}\rangle \\
&\quad + \varepsilon\sum_{n=1}^{m}\langle\partial X^n, \partial X^n - \partial X^{n-1}\rangle + \tau\sum_{n=1}^{m}\langle\partial X^n, \partial[\Lambda_\delta(X^n)]\rangle.
\end{aligned}
\tag{4.3}
$$

The rightmost term is nonnegative due to the monotonicity of Λ_δ. The first term in the right hand side of (4.3) can be handled as follows via (2.20), Young's inequality and the fact that $\Theta^{-1} = \Theta^0$:

$$
\begin{aligned}
&\tau\Big|\sum_{n=1}^{m}\langle\partial X^n, \partial\Theta^n - \partial\Theta^{n-1}\rangle\Big| \\
&\qquad = \tau\Big|\langle\partial X^m, \partial\Theta^m\rangle - \langle\partial X^0, \partial\Theta^0\rangle - \sum_{n=1}^{m}\langle\partial X^n - \partial X^{n-1}, \partial\Theta^{n-1}\rangle\Big| \\
&\qquad \le \tfrac{2\tau}{3\varepsilon}\tfrac{\varepsilon}{2}\|\partial X^m\|^2_{L^2(\Omega)} + \tfrac{3}{4}\tau\sum_{n=1}^{m}\|\partial\Theta^n\|^2_{L^2(\Omega)} + \tfrac{2\tau}{3\varepsilon}\tfrac{\varepsilon}{2}\sum_{n=1}^{m}\|\partial X^n - \partial X^{n-1}\|^2_{L^2(\Omega)}.
\end{aligned}
$$

On using (2.18) once again, we arrive at

$$
\varepsilon\sum_{n=1}^{m}\langle\partial X^n, \partial X^n - \partial X^{n-1}\rangle = \tfrac{\varepsilon}{2}\|\partial X^m\|^2_{L^2(\Omega)} - \tfrac{\varepsilon}{2}\|\partial X^0\|^2_{L^2(\Omega)} + \tfrac{\varepsilon}{2}\sum_{n=1}^{m}\|\partial X^n - \partial X^{n-1}\|^2_{L^2(\Omega)}.
$$

For the convection term, we use Young's inequality to get

$$\begin{aligned}\Big|\sum_{n=1}^{m}\tau\langle \mathbf{b}(\Theta^{n-1}),\nabla(\partial\Theta^n)\rangle\Big| &= \Big|\sum_{n=1}^{m}\tau\langle \operatorname{div}\mathbf{b}(\Theta^{n-1}),\partial\Theta^n\rangle\Big| \\ &\le C\sum_{n=1}^{m}\tau\|\operatorname{div}\mathbf{b}(\Theta^{n-1})\|^2_{L^2(\Omega)}+\eta\sum_{n=1}^{m}\tau\|\partial\Theta^n\|^2_{L^2(\Omega)} \\ &\le CL_b\sum_{n=1}^{m}\tau\|\nabla\Theta^{n-1}\|^2_{L^2(\Omega)}+\eta\sum_{n=1}^{m}\tau\|\partial\Theta^n\|^2_{L^2(\Omega)},\end{aligned}$$

where $\eta>0$ is to be chosen. Similarly, applying Young's inequality together with Poincarè's inequality, we obtain

$$\tau\sum_{n=1}^{m}\big|\langle f(\Theta^{n-1}),\partial\Theta^n\rangle\big| \le CL_f\sum_{n=1}^{m}\tau\|\nabla\Theta^{n-1}\|^2_{L^2(\Omega)}+\eta\sum_{n=1}^{m}\tau\|\partial\Theta^n\|^2_{L^2(\Omega)}.$$

Combining the above estimates for $\eta=\frac{1}{16}$, invoking (2.17), and then the discrete Gronwall's inequality, the assertion follows immediately provided $\tau\le\varepsilon$. ∎

LEMMA 4.2. (4.1) *is valid for the semi-implicit scheme without restrictions on* τ.

Proof. We proceed as in Lemma 4.1 and note that (4.3) becomes

$$\sum_{n=1}^{m}\tau\langle\partial X^n,\partial\Theta^n\rangle=\varepsilon\sum_{n=1}^{m}\langle\partial X^n,\partial X^n-\partial X^{n-1}\rangle+\sum_{n=1}^{m}\tau\langle\partial X^n,\partial[\Lambda_\delta(X^n)]\rangle.$$

Since the missing term $\sum_n\langle\partial X^n,\partial\Theta^n-\partial\Theta^{n-1}\rangle$ is the only one responsible for the constraint $\tau\le\varepsilon$, the proof continues as that of Lemma 4.1. ∎

LEMMA 4.3. (4.1) *is valid for the extrapolation scheme provided* $\tau\le\varepsilon/2$.

Proof. Again the only difference is in dealing with (4.3). With the notation $\partial^2\Theta^n := (\partial\Theta^n-\partial\Theta^{n-1})/\tau$, (1.11) with $T^n=2\Theta^{n-1}-\Theta^{n-2}$ can be written equivalently as

$$\varepsilon\partial X^n+\Lambda_\delta(X^n)=\Theta^n-\tau^2\partial^2\Theta^n.$$

With the aid of (2.20), the first term on the right hand side of (4.3) thus becomes

$$\begin{aligned}&\tau^2\Big|\sum_{n=1}^{m}\langle\partial X^n,\partial^2\Theta^n-\partial^2\Theta^{n-1}\rangle\Big| \\ =&\tau^2\Big|\langle\partial X^m,\partial^2\Theta^m\rangle-\langle\partial X^0,\partial^2\Theta^0\rangle-\sum_{n=1}^{m}\langle\partial X^n-\partial X^{n-1},\partial^2\Theta^{n-1}\rangle\Big| \\ \le&\tfrac{4\tau}{3\varepsilon}\tfrac{\varepsilon}{2}\|\partial X^m\|^2_{L^2(\Omega)}+\tfrac{3}{8}\tau\sum_{n=1}^{m}\|\partial\Theta^n-\partial\Theta^{n-1}\|^2_{L^2(\Omega)}+\tfrac{4\tau}{3\varepsilon}\tfrac{\varepsilon}{2}\sum_{n=1}^{m}\|\partial X^n-\partial X^{n-1}\|^2_{L^2(\Omega)},\end{aligned}$$

because $\Theta^{-2} = \Theta^{-1} = \Theta^0$. Since $\sum_n \|\partial(\Theta^n - \Theta^{n-1})\|^2_{L^2(\Omega)} \le 2\sum_n \|\partial\Theta^n\|^2_{L^2(\Omega)}$, the constraint $\tau \le \varepsilon/2$ thus suffices to complete the proof. $\square$

The final time T can be taken to be $T = \infty$ in the above a priori bounds provided convection and reaction are not present. In fact in this particular case there is no need for Gronwall's inequality, which restricts T to be finite.

5. Semi-explicit Scheme. We first consider the simplest case $F = 0$ and derive error estimates that are uniform in the regularization parameter δ, thereby being valid in the limit $\delta \downarrow 0$. By regularization we avoid dealing with notions such as minimal section $\mathcal{A}^0$, right derivatives, and subgradients as in [12]. We next extend the analysis to the general case $F \neq 0$. Once again, we suppress the subscript δ from all functions except Λ_δ in this section. The approximation error $E(\varepsilon,\tau)$ is defined in (1.12).

THEOREM 5.1. *Suppose* $\mathbf{b} = \mathbf{0}$, $f = 0$. *Then there exists* $C > 0$, *depending on* $\|\nabla\theta_0\|_{L^2(\Omega)}$, A *in* (2.4) *and* $T < \infty$ *but independent of* ε, *and* τ, *such that* $E(\varepsilon,\tau) \le C\tau/\sqrt{\varepsilon}$ *provided* $\tau \le \varepsilon$.

Proof. Let $G : H^{-1}(\Omega) \to H^1_0(\Omega)$ stand for the Green's operator associated with $-\Delta$ and a vanishing Dirichlet boundary condition. Hence

$$G\varphi \in H^1_0(\Omega): \qquad \langle \nabla G\varphi, \nabla\psi\rangle = \langle \varphi, \psi\rangle, \qquad \forall \psi \in H^1_0(\Omega),\ \varphi \in H^{-1}(\Omega),$$

and

$$\|\varphi\|^2_{H^{-1}(\Omega)} = \|\nabla G\varphi\|^2_{L^2(\Omega)} = \langle \varphi, G\varphi\rangle. \tag{5.1}$$

We proceed now as in [12], by introducing first a periodic function g^τ on $\mathbb{R}$ defined by

$$g^\tau(t) := \tfrac{t^n - t}{\tau}, \qquad \text{for } t^{n-1} \le t < t^n, \tag{5.2}$$

and then the following piecewise linear interpolants of $\{U^n\}_{n=1}^N$ and $\{X^n\}_{n=1}^N$:

$$U(t) := \tfrac{t^n - t}{\tau} U^{n-1} + \tfrac{t - t^{n-1}}{\tau} U^n = U^n - \tau g^\tau(t)\partial U^n = U^n - \tau g^\tau(t)\Delta\Theta^n, \tag{5.3}$$

$$X(t) := \tfrac{t^n - t}{\tau} X^{n-1} + \tfrac{t - t^{n-1}}{\tau} X^n = X^n - \tau g^\tau(t)\partial X^n. \tag{5.4}$$

We obviously have $\partial_t U(t) = \partial U^n$ and $\partial_t X(t) = \partial X^n$ for $t \in I^n$. Subtracting (2.14) from (2.1) for $F = 0$, we deduce the first *error equation*

$$\langle \partial_t(u - U), \phi\rangle + \langle \nabla(\theta - \Theta^n), \nabla\phi\rangle = 0, \qquad \forall\ \phi \in H^1_0(\Omega),\ t \in I^n.$$

Taking $\phi := G(u - U) \in H^1_0(\Omega)$, and integrating from $t = 0$ to $t = t^m \le T$, we get

$$\sum_{n=1}^m \int_{I^n} \Big(\tfrac{1}{2}\tfrac{d}{dt}\langle u - U, G(u - U)\rangle + \langle \theta - \Theta^n, u - U\rangle \Big)\, dt = 0.$$

In view of (5.1) and (5.3), it is easy to see that

$$\tfrac{1}{2}\|(u-U)(t^m)\|^2_{H^{-1}(\Omega)} + \sum_{n=1}^{m}\int_{I^n}\|\theta-\Theta^n\|^2_{L^2(\Omega)}\,dt + \sum_{n=1}^{m}\int_{I^n}\langle\theta-\Theta^n,\chi-X^n\rangle\,dt$$
$$+\tau\sum_{n=1}^{m}\int_{I^n} g^\tau(t)\langle\nabla(\Theta^n-\theta),\nabla\Theta^n\rangle\,dt = \tfrac{1}{2}\|(u-U)(0)\|^2_{H^{-1}(\Omega)},$$

or equivalently, upon writing $2\Theta^n = (\Theta^n-\theta)+(\Theta^n+\theta)$,

$$\begin{aligned}
&\tfrac{1}{2}\|(u-U)(t^m)\|^2_{H^{-1}(\Omega)} + \sum_{n=1}^{m}\int_{I^n}\|\theta-\Theta^n\|^2_{L^2(\Omega)}\,dt\\
&+\sum_{n=1}^{m}\int_{I^n}\langle\theta-\Theta^n,\chi-X^n\rangle\,dt + \tfrac{\tau}{2}\sum_{n=1}^{m}\int_{I^n} g^\tau(t)\|\nabla(\Theta^n-\theta)\|^2_{L^2(\Omega)}\,dt\\
&= \tfrac{1}{2}\|(u-U)(0)\|^2_{H^{-1}(\Omega)}\\
&+\tfrac{\tau}{2}\sum_{n=1}^{m}\int_{I^n} g^\tau(t)\|\nabla\theta\|^2_{L^2(\Omega)}\,dt - \tfrac{\tau}{2}\sum_{n=1}^{m}\int_{I^n} g^\tau(t)\|\nabla\Theta^n\|^2_{L^2(\Omega)}\,dt.
\end{aligned}\tag{5.5}$$

In dealing with the third term on the left hand side of (5.5), we resort to the following *error equation* for the constitutive relation, obtained by subtracting the regularized analogue of (2.15) and (2.2), and next using $\partial_t X = \partial X^n$:

$$\varepsilon\langle\partial_t(\chi-X),\varphi\rangle + \langle\Lambda_\delta(\chi)-\Lambda_\delta(X^n),\varphi\rangle$$
$$= \langle\theta-\Theta^n,\varphi\rangle + \langle\Theta^n-\Theta^{n-1},\varphi\rangle \quad \forall\,\varphi\in L^2(\Omega),\ t\in I^n.$$

On choosing $\varphi=\chi-X^n$, and making use of (5.4) and the monotonicity of Λ_δ, we get

$$\begin{aligned}
&\sum_{n=1}^{m}\int_{I^n}\langle\theta-\Theta^n,\chi-X^n\rangle\,dt + \sum_{n=1}^{m}\int_{I^n}\langle\Theta^n-\Theta^{n-1},\chi-X^n\rangle\,dt\\
&=\sum_{n=1}^{m}\int_{I^n}\varepsilon\langle\partial_t(\chi-X),\chi-X^n\rangle\,dt + \sum_{n=1}^{m}\int_{I^n}\langle\Lambda_\delta(\chi)-\Lambda_\delta(X^n),\chi-X^n\rangle\,dt\\
&\geq \tfrac{\varepsilon}{2}\sum_{n=1}^{m}\int_{I^n}\tfrac{d}{dt}\|\chi-X\|^2_{L^2(\Omega)}\,dt - \varepsilon\tau\sum_{n=1}^{m}\int_{I^n} g^\tau(t)\langle\partial_t(\chi-X),\partial_t X\rangle\,dt\\
&= \tfrac{\varepsilon}{2}\|(\chi-X)(t^m)\|^2_{L^2(\Omega)} - \tfrac{\varepsilon}{2}\|(\chi-X)(0)\|^2_{L^2(\Omega)}\\
&+\tfrac{\varepsilon\tau}{2}\sum_{n=1}^{m}\int_{I^n} g^\tau(t)\Big(\|\partial_t(\chi-X)\|^2_{L^2(\Omega)} - \|\partial_t\chi\|^2_{L^2(\Omega)} + \|\partial_t X\|^2_{L^2(\Omega)}\Big)\,dt.
\end{aligned}\tag{5.6}$$

Replacing this back into (5.5), and noticing that $g^\tau(t) \geq 1/2$ on $[t^{n-1}, t^{n-1} + \tau/2]$, we deduce the estimate

$$
\begin{aligned}
&\|(u-U)(t^m)\|^2_{H^{-1}(\Omega)} + 2\sum_{n=1}^m \int_{I^n} \|\theta - \Theta^n\|^2_{L^2(\Omega)}\,dt + \varepsilon\|(\chi - X)(t^m)\|^2_{L^2(\Omega)}\\
&+ \tfrac{\tau}{2}\sum_{n=1}^m \int_{t^{n-1}}^{t^{n-1}+\frac{\tau}{2}} \Big(\|\nabla(\theta - \Theta^n)\|^2_{L^2(\Omega)} + \varepsilon\|\partial_t(\chi - X)\|^2_{L^2(\Omega)}\Big)\,dt\\
&\leq \|(u-U)(0)\|^2_{H^{-1}(\Omega)} + \varepsilon\|(\chi - X)(0)\|^2_{L^2(\Omega)}\\
&+ \tau\int_0^{t^m} g^\tau(t)\alpha(t)\,dt - \tfrac{\tau^2}{2}\sum_{n=1}^m \alpha^n + \sum_{n=1}^m \int_{I^n} \langle \Theta^n - \Theta^{n-1}, \chi - X^n\rangle\,dt,
\end{aligned}
\tag{5.7}
$$

where

$$
\alpha(t) := \|\nabla\theta(t)\|^2_{L^2(\Omega)} + \varepsilon\|\partial_t\chi(t)\|^2_{L^2(\Omega)}, \quad \alpha^n := \|\nabla\Theta^n\|^2_{L^2(\Omega)} + \varepsilon\|\partial X^n\|^2_{L^2(\Omega)}.
$$

We now point out that the fourth term on the left hand side of (5.7) does not quite yield an error estimate because the integration is restricted to half of I^n. To overcome such a difficulty we shift time by $\frac{\tau}{2}$, namely, we set

$$
\hat\theta := \theta(\cdot + \tfrac{\tau}{2}), \qquad \hat\chi := \chi(\cdot + \tfrac{\tau}{2}),
\tag{5.8}
$$

$\hat u := \hat\theta + \hat\chi$, and observe that these functions satisfy, because of the semigroup property,

$$
\begin{aligned}
\partial_t(\hat\theta + \hat\chi) - \Delta\hat\theta = 0, &\qquad \varepsilon\partial_t\hat\chi + \Lambda_\delta(\hat\chi) \ni \hat\theta,\\
\hat\theta(0) = \theta(\tfrac{\tau}{2}), &\qquad \hat\chi(0) = \chi(\tfrac{\tau}{2}).
\end{aligned}
$$

We now compare $\hat u, \hat\chi$ with U and X. In fact, applying (5.7), we have

$$
\begin{aligned}
&\tfrac{\tau}{2}\sum_{n=1}^m \int_{t^{n-1}}^{t^{n-1}+\frac{\tau}{2}} \Big(\|\nabla(\hat\theta - \Theta^n)\|^2_{L^2(\Omega)} + \varepsilon\|\partial_t(\hat\chi - X)\|^2_{L^2(\Omega)}\Big)\,dt\\
&\qquad\leq \|(\hat u - U)(0)\|^2_{H^{-1}(\Omega)} + \varepsilon\|(\hat\chi - X)(0)\|^2_{L^2(\Omega)}\\
&\qquad+ \tau\int_0^{t^m} g^\tau(t)\alpha(t + \tfrac{\tau}{2})\,dt - \tfrac{\tau^2}{2}\sum_{n=1}^m \alpha^n + \sum_{n=1}^m \int_{I^n} \langle \Theta^n - \Theta^{n-1}, \hat\chi - X^n\rangle\,dt,
\end{aligned}
$$

whence, owing to (5.8) and the change of variables $s = t + \frac{\tau}{2}$,

$$
\begin{aligned}
&\tfrac{\tau}{2}\sum_{n=1}^m \int_{t^{n-1}+\frac{\tau}{2}}^{t^n} \Big(\|\nabla(\theta - \Theta^n)\|^2_{L^2(\Omega)} + \varepsilon\|\partial_t(\chi - X)\|^2_{L^2(\Omega)}\Big)\,dt\\
&\qquad\leq \|(\hat u - U)(0)\|^2_{H^{-1}(\Omega)} + \varepsilon\|(\hat\chi - \chi)(0)\|^2_{L^2(\Omega)}\\
&\qquad+ \tau\int_{\frac{\tau}{2}}^{t^m+\frac{\tau}{2}} g^\tau(t - \tfrac{\tau}{2})\alpha(t)\,dt - \tfrac{\tau^2}{2}\sum_{n=1}^m \alpha^n + \sum_{n=1}^m \int_{I^n} \langle \Theta^n - \Theta^{n-1}, \hat\chi - X^n\rangle\,dt.
\end{aligned}
$$

Since $U(0) = u(0) = u_0$ and $X(0) = \chi(0)$, adding the above expression with (5.7) and using the fact that $g^\tau(t) + g^\tau(t - \frac{\tau}{2}) = \frac{1}{2} + g^{\tau/2}(t)$, we get

$$
\begin{aligned}
&\|u(t^m) - U^m\|^2_{H^{-1}(\Omega)} + \varepsilon\|\chi(t^m) - X^m\|^2_{L^2(\Omega)} + \tfrac{\tau\varepsilon}{2}\sum_{n=1}^{m}\int_{I^n} \|\partial_t(\chi - X)\|^2_{L^2(\Omega)}\, dt \\
&+2\sum_{n=1}^{m}\int_{I^n} \Big(\|\theta - \Theta^n\|^2_{L^2(\Omega)} + \tfrac{\tau}{4}\|\nabla(\theta - \Theta^n)\|^2_{L^2(\Omega)}\Big)\, dt \\
&\le \|u(0) - u(\tfrac{\tau}{2})\|^2_{H^{-1}(\Omega)} + \varepsilon\|\chi(0) - \chi(\tfrac{\tau}{2})\|^2_{L^2(\Omega)} + \tau\int_0^{t^m} \big(g^{\tau/2}(t) - \tfrac{1}{2}\big)\alpha(t)\, dt \qquad (5.9)\\
&+ \tau\int_0^{t^m} \alpha(t)\, dt - \tau^2\sum_{n=1}^{m}\alpha^n + \tau\int_0^{\tau/2} g^\tau(t - \tfrac{\tau}{2})\big(\alpha(t + t^m) - \alpha(t)\big)\, dt \\
&+ 2\sum_{n=1}^{m}\int_{I^n} \langle\Theta^n - \Theta^{n-1}, \chi - X^n\rangle\, dt - \sum_{n=1}^{m}\int_{I^n}\langle\Theta^n - \Theta^{n-1}, \chi - \hat{\chi}\rangle\, dt.
\end{aligned}
$$

We proceed now to examine the last six terms separately. We first notice that the integral term over $[0, \frac{\tau}{2}]$ is nonpositive, because of (2.10) with $F = 0$. This monotonicity property can be used again to argue as follows:

$$
\begin{aligned}
\tau\int_0^{t^m} \Big(g^{\tau/2}(t) - \tfrac{1}{2}\Big)\alpha(t)\, dt &= \tau\sum_{j=1}^{2m}\Big(\int_{(j-1)\frac{\tau}{2}}^{j\frac{\tau}{2}} g^{\tau/2}(t)\alpha(t)\, dt - \int_{(j-1)\frac{\tau}{2}}^{j\frac{\tau}{2}} \tfrac{1}{2}\alpha(t)\, dt\Big) \\
&\le \tau\sum_{j=1}^{2m}\Big(\alpha((j-1)\tfrac{\tau}{2})\int_{(j-1)\frac{\tau}{2}}^{j\frac{\tau}{2}} g^{\tau/2}(t)dt - \alpha(j\tfrac{\tau}{2})\int_{(j-1)\frac{\tau}{2}}^{j\frac{\tau}{2}} \tfrac{1}{2}dt\Big) \\
&\le \tfrac{\tau^2}{4}\sum_{j=1}^{2m}\Big(\alpha((j-1)\tfrac{\tau}{2}) - \alpha(j\tfrac{\tau}{2})\Big) \le \tfrac{\tau^2}{4}\alpha(0) \le C\tau^2;
\end{aligned}
$$

the bound $\alpha(0) = \|\nabla\theta_0\|^2_{L^2(\Omega)} + \sqrt{\varepsilon}\|\partial_t\chi(0)\|^2_{L^2(\Omega)} \le C$ results from (2.12). To handle the two summands involving $\alpha(t)$ and α^n, we introduce the *convex* functions

$$
\Phi(s) := \frac{s^2}{2}, \qquad \Psi(s) := \int_0^s \Lambda_\delta(r)dr.
$$

Since Λ_δ is Lipschitz, Ψ is a classical primitive. Otherwise Λ would have to be interpreted as a subdifferential. We have

$$
\begin{aligned}
\tau\int_0^{t^m}\alpha(t)\,dt &= \tau\int_0^{t^m}\langle\nabla\theta,\nabla\theta\rangle\,dt+\tau\varepsilon\int_0^{t^m}\langle\partial_t\chi,\partial_t\chi\rangle\,dt\\
&=-\tau\int_0^{t^m}\langle\partial_t(\theta+\chi),\theta\rangle\,dt+\tau\varepsilon\int_0^{t^m}\langle\partial_t\chi,\partial_t\chi\rangle\,dt\\
&=-\tau\int_0^{t^m}\langle\partial_t\theta,\theta\rangle\,dt-\tau\int_0^{t^m}\langle\partial_t\chi,\Lambda_\delta(\chi)\rangle\,dt\\
&=-\tau\int_\Omega\int_0^{t^m}\tfrac{d}{dt}\Phi(\theta(t))\,dt-\tau\int_\Omega\int_0^{t^m}\tfrac{d}{dt}\Psi(\chi(t))\,dt\\
&=\tau\int_\Omega\big(\Phi(\theta(0))+\Psi(\chi(0))\big)-\tau\int_\Omega\big(\Phi(\theta(t^m))+\Psi(\chi(t^m))\big).
\end{aligned}\tag{5.10}
$$

We now employ (2.14) and (2.15) with $F=0$, to obtain

$$
\begin{aligned}
-\tau^2\sum_{n=1}^m\alpha^n &=-\tau^2\sum_{n=1}^m\langle\nabla\Theta^n,\nabla\Theta^n\rangle-\tau^2\varepsilon\sum_{n=1}^m\langle\partial X^n,\partial X^n\rangle\\
&=\tau\sum_{n=1}^m\langle\Theta^n-\Theta^{n-1},\Theta^{n-1}\rangle+\tau\sum_{n=1}^m\|\Theta^n-\Theta^{n-1}\|^2_{L^2(\Omega)}\\
&\quad+\tau\sum_{n=1}^m\langle X^n-X^{n-1},\Lambda_\delta(X^{n-1})\rangle+\tau\sum_{n=1}^m\langle X^n-X^{n-1},\Theta^n-\Theta^{n-1}\rangle\\
&\quad+\tau\sum_{n=1}^m\langle X^n-X^{n-1},\Lambda_\delta(X^n)-\Lambda_\delta(X^{n-1})\rangle.
\end{aligned}\tag{5.11}
$$

Using the convexity of Φ and Ψ and Lemma 4.1, we get

$$
\begin{aligned}
-\tau^2\sum_{n=1}^m\alpha^n &\le C\tau^2+\tau\sum_{n=1}^m\langle\Theta^n-\Theta^{n-1},\Phi'(\Theta^{n-1})\rangle\\
&\quad+\tau\sum_{n=1}^m\langle X^n-X^{n-1},\Psi'(X^{n-1})\rangle+\tau\sum_{n=1}^m\langle X^n-X^{n-1},\Theta^n-\Theta^{n-1}\rangle\\
&\le C\tfrac{\tau^2}{\varepsilon}+\tau\int_\Omega\big(\Phi(\Theta^m)+\Psi(X^m)\big)-\tau\int_\Omega\big(\Phi(\Theta^0)+\Psi(X^0)\big).
\end{aligned}
$$

Consequently, exploiting again the monotonicity of both Φ and Ψ and (2.15), we end up with

$$
\begin{aligned}
\tau \int_0^{t^m} \alpha(t)\, dt - \tau^2 \sum_{n=1}^{m} \alpha^n & \\
&\le C\tfrac{\tau^2}{\varepsilon} + \tau \int_\Omega \Phi(\Theta^m) - \Phi(\theta(t^m)) + \Psi(X^m) - \Psi(\chi(t^m)) \\
&\le C\tfrac{\tau^2}{\varepsilon} + \tau\langle \Theta^m, \Theta^m - \theta(t^m)\rangle + \tau\langle \Lambda_\delta(X^m), X^m - \chi(t^m)\rangle \\
&\le C\tfrac{\tau^2}{\varepsilon} + \tau\langle \Theta^m, U^m - u(t^m)\rangle - \tau\varepsilon\langle \partial X^m, X^m - \chi(t^m)\rangle \\
&\qquad + \tau\langle \Theta^{m-1} - \Theta^m, X^m - \chi(t^m)\rangle \\
&\le C\tfrac{\tau^2}{\varepsilon} + \tfrac{\tau^2}{2}\alpha^m + \tfrac{1}{2}\|U^m - u(t^m)\|^2_{H^{-1}(\Omega)} + \tfrac{\varepsilon}{2}\|X^m - \chi(t^m)\|^2_{L^2(\Omega)} \\
&\qquad + \tau\langle \Theta^{m-1} - \Theta^m, X^m - \chi(t^m)\rangle \\
&\le C\tfrac{\tau^2}{\varepsilon} + \tfrac{1}{2}\|U^m - u(t^m)\|^2_{H^{-1}(\Omega)} + \tfrac{\varepsilon}{2}\|X^m - \chi(t^m)\|^2_{L^2(\Omega)} \\
&\qquad + \tau\langle \Theta^{m-1} - \Theta^m, X^m - \chi(t^m)\rangle.
\end{aligned}
\tag{5.12}
$$

Here we have also used the property $\alpha^m \le \alpha^0 \le C$. The last term above can be associated with the last two terms in (5.9). They all together can be handled by using Lemma 4.1 and the regularity (2.7) as follows:

$$
\begin{aligned}
&\sum_{n=1}^{m} \tau |\langle \Theta^n - \Theta^{n-1}, \chi(t^n) - X^n\rangle|\, dt + \sum_{n=1}^{m} \int_{I^n} |\langle \Theta^n - \Theta^{n-1}, \chi - \chi(t^n)\rangle|\, dt \\
&\qquad + \sum_{n=1}^{m} \int_{I^n} |\langle \Theta^n - \Theta^{n-1}, \chi - \hat\chi\rangle|\, dt \le C\tfrac{\tau^2}{\varepsilon} + \varepsilon \sum_{n=1}^{m} \tau \|\chi(t^n) - X^n\|^2_{L^2(\Omega)}.
\end{aligned}
\tag{5.13}
$$

In fact we note for instance that

$$
\|\chi(t) - \chi(t^n)\|_{L^2(\Omega)} \le \int_t^{t^n} \|\partial_t \chi(s)\|_{L^2(\Omega)}\, ds \le C\tfrac{\tau}{\sqrt{\varepsilon}} \quad \forall\ t^{n-1} < t < t^n. \tag{5.14}
$$

We finally take the limit $\delta \downarrow 0$ and replace the resulting estimates back into (5.9). This yields the asserted error bounds upon application of the discrete Gronwall's inequality. □

Theorem 5.1 shows that the explicit nature of the scheme degrades the asymptotic accuracy from $\mathcal{O}(\tau)$ for the implicit scheme (Theorem 3.1) to $\tau/\varepsilon^{1/2}$. This in turn is compensated by linearity of the resulting problem. *Optimality* of this order is a consequence of the error estimates for 1D semidiscrete traveling waves of [10].

We next consider the general case with convection $\mathbf{b} \neq \mathbf{0}$ and reaction $f \neq 0$.

THEOREM 5.2. *There exists $C > 0$, depending on $\|\nabla\theta_0\|_{L^2(\Omega)}$, A in (2.4), $L_{\mathbf{b}}, L_f$, and T but not on ε or τ, such that $E(\varepsilon,\tau) \le C\tau/\sqrt{\varepsilon}$ holds under the stability constraint $\tau \le \varepsilon$.*

Proof. We just sketch how to deal with the extra terms due to the presence of $\mathbf{b}$ and f. For the sake of simplicity we assume $f = 0$ since the terms involving $f(\theta)$ are in fact easier to handle than those for $\mathbf{b}(\theta)$. In deriving an analogue of (5.9), $\operatorname{div}\mathbf{b}(\theta)$ and $\operatorname{div}\mathbf{b}(\Theta^{n-1})$ contribute with the following extra terms:

$$
\begin{aligned}
\mathrm{I}+\mathrm{II}+\mathrm{III}+\mathrm{IV} := &-4\tau\sum_{n=1}^{m}\int_{I^n} g^\tau(t)\langle \operatorname{div}\mathbf{b}(\Theta^{n-1}), \theta-\Theta^n\rangle\, dt \\
&-4\sum_{n=1}^{m}\int_{I^n} \langle \mathbf{b}(\theta)-\mathbf{b}(\Theta^{n-1}), \nabla G(u-U)\rangle\, dt \\
&+2\tau\sum_{n=1}^{m}\int_{I^n} g^\tau(t)\langle \operatorname{div}\mathbf{b}(\Theta^{n-1}), \theta-\hat{\theta}\rangle\, dt \\
&+2\sum_{n=1}^{m}\int_{I^n} \langle \mathbf{b}(\theta)-\mathbf{b}(\Theta^{n-1}), \nabla G(u-\hat{u})\rangle\, dt.
\end{aligned}
\tag{5.15}
$$

In light of Lemma 4.1, we readily have

$$
\begin{aligned}
\mathrm{I} &\le C\tau^2\sum_{n=1}^{m}\int_{I^n} \|\nabla\Theta^{n-1}\|_{L^2(\Omega)}^2\, dt + \eta\sum_{n=1}^{m}\int_{I^n}\|\theta-\Theta^n\|_{L^2(\Omega)}^2\, dt \\
&\le C\tau^2 + \eta\sum_{n=1}^{m}\int_{I^n}\|\theta-\Theta^n\|_{L^2(\Omega)}^2\, dt,
\end{aligned}
$$

with $\eta>0$ to be chosen. Term II in (5.15) can be further written as

$$
\begin{aligned}
\mathrm{II} = &-4\sum_{n=1}^{m}\int_{I^n}\langle \mathbf{b}(\theta)-\mathbf{b}(\Theta^n), \nabla G(u(t^n)-U^n)\rangle\, dt \\
&-4\sum_{n=1}^{m}\int_{I^n}\langle \mathbf{b}(\theta)-\mathbf{b}(\Theta^n), \nabla G(u-u(t^n)-U+U^n)\rangle\, dt \\
&-4\sum_{n=1}^{m}\int_{I^n}\langle \mathbf{b}(\Theta^n)-\mathbf{b}(\Theta^{n-1}), \nabla G(u(t^n)-U^n)\rangle\, dt \\
&-4\sum_{n=1}^{m}\int_{I^n}\langle \mathbf{b}(\Theta^n)-\mathbf{b}(\Theta^{n-1}), \nabla G(u-u(t^n)-U+U^n)\rangle\, dt.
\end{aligned}
$$

On using (2.7), Lemma 4.1 and (5.3), these four terms can be handled as follows:

$$
\begin{aligned}
\mathrm{II} \le\ & \eta\sum_{n=1}^{m}\int_{I^n}\|\theta-\Theta^n\|_{L^2(\Omega)}^2\, dt + C\sum_{n=1}^{m}\tau\|u(t^n)-U^n\|_{H^{-1}(\Omega)}^2 \\
&+ C\sum_{n=1}^{m}\int_{I^n}\Big(\|u-u(t^n)\|_{H^{-1}(\Omega)}^2 + \|U-U^n\|_{H^{-1}(\Omega)}^2 + \tau^2\|\partial\Theta^n\|_{L^2(\Omega)}^2\Big)\, dt \\
\le\ & C\tau^2 + \eta\sum_{n=1}^{m}\int_{I^n}\|\theta-\Theta^n\|_{L^2(\Omega)}^2\, dt + C\sum_{n=1}^{m}\tau\|u(t^n)-U^n\|_{H^{-1}(\Omega)}^2.
\end{aligned}
$$

The other two terms in (5.15) can be handled in a similar way, namely,

$$\begin{aligned}
\mathrm{III}+\mathrm{IV} &\leq C\tau^2\sum_{n=1}^{m}\tau\big(\|\nabla\Theta^{n-1}\|^2_{L^2(\Omega)}+\|\partial\Theta^n\|^2_{L^2(\Omega)}\big)\\
&+\ C\sum_{n=1}^{m}\int_{I^n}\big(\|\theta-\hat{\theta}\|^2_{L^2(\Omega)}+\|u-\hat{u}\|^2_{H^{-1}(\Omega)}\big)\,dt\\
&+\ \eta\sum_{n=1}^{m}\int_{I^n}\|\theta-\Theta^n\|^2_{L^2(\Omega)}\,dt \leq\ C\tau^2+\eta\sum_{n=1}^{m}\int_{I^n}\|\theta-\Theta^n\|^2_{L^2(\Omega)}\,dt.
\end{aligned}$$

A key property in the previous error analysis of Theorem 5.1 is the monotonicity of the Liapunov functional $\alpha(t)$. This is no longer valid in the present situation, but instead we have (2.10):

$$\alpha(t_2)-\alpha(t_1)\leq\int_{t_1}^{t_2}\|\operatorname{div}\mathbf{b}(\theta)\|^2_{L^2(\Omega)}\,dt\qquad(t_2>t_1).$$

Thus for $t_1\leq t<t_2$, it holds

$$\alpha(t_2)-\int_{t_1}^{t_2}\|\operatorname{div}\mathbf{b}(\theta)\|^2_{L^2(\Omega)}\,dt\leq\alpha(t)\leq\alpha(t_1)+\int_{t_1}^{t_2}\|\operatorname{div}\mathbf{b}(\theta)\|^2_{L^2(\Omega)}\,dt.$$

For the third term on the right hand side of (5.9), we thus have

$$\begin{aligned}
\tau\int_0^{t^m}\big(g^{\tau/2}(t)-\tfrac{1}{2}\big)\alpha(t)\,dt &\leq\ \tau\sum_{j=1}^{2m}\Big(\alpha((j-1)\tfrac{\tau}{2})\int_{(j-1)\frac{\tau}{2}}^{j\frac{\tau}{2}}g^{\tau/2}(t)dt-\alpha(j\tfrac{\tau}{2})\int_{(j-1)\frac{\tau}{2}}^{j\frac{\tau}{2}}\tfrac{1}{2}dt\Big)\\
&+\tfrac{\tau^2}{2}\sum_{j=1}^{2m}\int_{(j-1)\frac{\tau}{2}}^{j\frac{\tau}{2}}\|\operatorname{div}\mathbf{b}(\theta)\|^2_{L^2(\Omega)}\,dt\leq\ \tfrac{\tau^2}{4}\alpha(0)+\tfrac{\tau^2}{2}\int_0^{t^m}\|\operatorname{div}\mathbf{b}(\theta)\|^2_{L^2(\Omega)}\,dt\\
&\leq\ \tfrac{\tau^2}{4}\alpha(0)+\tfrac{\tau^2}{2}L^2_{\mathbf{b}}\|\nabla\theta\|^2_{L^2(0,t^m;L^2(\Omega))}\leq C\tau^2,
\end{aligned}$$

because of (2.7) and (2.12). To estimate the term $\tau\int_0^{t^m}\alpha(t)\,dt-\tau^2\sum_{n=1}^{m}\alpha^n$ in (5.9), we proceed as in (5.10) and (5.11). The following additional terms have to be bounded:

$$\begin{aligned}
\mathrm{V}:=-\tau\int_0^{t^m}\langle\mathbf{b}(\theta),\nabla\theta\rangle\,dt+\tau^2\sum_{n=1}^{m}\langle\mathbf{b}(\Theta^{n-1}),\nabla\Theta^n\rangle &=-\tau\sum_{n=1}^{m}\int_{I^n}\langle\mathbf{b}(\theta)-\mathbf{b}(\Theta^n),\nabla\theta\rangle\,dt\\
-\tau\sum_{n=1}^{m}\int_{I^n}\langle\mathbf{b}(\Theta^n)-\mathbf{b}(\Theta^{n-1}),\nabla\theta\rangle\,dt &+\tau\sum_{n=1}^{m}\int_{I^n}\langle\operatorname{div}\mathbf{b}(\Theta^{n-1}),\theta-\Theta^n\rangle\,dt.
\end{aligned}$$

Here we have integrated by parts once. If we further use Lemma 4.1 and (2.7), then

$$\begin{aligned}
\mathrm{V}\leq\ \eta\sum_{n=1}^{m}\int_{I^n}\|\theta-\Theta^n\|^2_{L^2(\Omega)}\,dt+C\tau^2\sum_{n=1}^{m}\int_{I^n}\|\nabla\theta\|^2_{L^2(\Omega)}\,dt+C\tau^2\sum_{n=1}^{m}\tau\|\partial\Theta^n\|^2_{L^2(\Omega)}\\
+C\tau^2\sum_{n=1}^{m}\tau\|\nabla\Theta^{n-1}\|^2_{L^2(\Omega)}\leq\ \eta\sum_{n=1}^{m}\int_{I^n}\|\theta-\Theta^n\|^2_{L^2(\Omega)}\,dt+C\tau^2.
\end{aligned}$$

The remaining terms in (5.9) can be treated similarly or left unchanged. The proof can then be finished upon taking η sufficiently small and using the discrete Gronwall's inequality. $\square$

We realize that the explicit treatment of convection and reaction does not further degrade the order $\mathcal{O}(\tau/\sqrt{\varepsilon}\,)$. Moreover the extra contributions are all of order $\mathcal{O}(\tau)$. This will be exploited in the next section.

6. Semi-implicit Scheme. We turn our attention to the semi-implicit scheme which leads to a sequence of nonlinear elliptic problems. We would like to prove error estimates similar to those in §3, but for $F = 0$. The key observation is that the correction $\Theta^n - \Theta^{n-1}$, in $T^n = \Theta^n - (\Theta^n - \Theta^{n-1})$ of (1.11) for the semi-explicit method, is the only term responsible for the factor $1/\sqrt{\varepsilon}$ in the above error analysis. In fact, for $T^n = \Theta^n$ we realize that $\tau \sum_{n=1}^{m} \langle X^n - X^{n-1}, \Theta^n - \Theta^{n-1} \rangle$ in (5.11) as well as (5.13) are now missing. This yields the following *optimal* error estimates that extend those in §3 to $F \neq 0$.

THEOREM 6.1. *There exists $C > 0$, depending on $\|\nabla\theta_0\|_{L^2(\Omega)}$, A in (2.4), $L_{\mathbf{b}}, L_f$, and T but independent of ε and τ, such that $E(\varepsilon, \tau) \leq C\tau$.*

We now investigate the relation between approximability and regularity in the spirit of [2,13]. Estimates for the third and fifth terms in (1.12) are not justified in light of (2.7). However, the third and fifth terms in (4.1) hint on additional regularity of θ and χ. Next regularity result elucidates this issue, and will be instrumental in §7.

COROLLARY 6.1. *There exists $C > 0$ depending solely on $\|\nabla\theta_0\|_{L^2(\Omega)}$, A in (2.4), $L_{\mathbf{b}}, L_f$, and T but independent of ε and τ, such that*

$$\int_0^{T-\tau} \left(\varepsilon \|\partial_t\chi(t+\tau) - \partial_t\chi(t)\|^2_{L^2(\Omega)} + \|\nabla\theta(t+\tau) - \nabla\theta(t)\|^2_{L^2(\Omega)} \right) dt \leq C\tau. \tag{6.1}$$

Proof. By virtue of Lemma 4.2 and Theorem 6.1 we readily find out that

$$\begin{aligned}\varepsilon \int_0^{T-\tau} \|\partial_t\chi(t+\tau) - \partial_t\chi(t)\|^2_{L^2(\Omega)} &\leq C\varepsilon \sum_{n=1}^{N-1} \int_{I^n} \|\partial_t\chi(t+\tau) - \partial X^{n+1}\|^2_{L^2(\Omega)}\, dt \\ &+ C\varepsilon \sum_{n=1}^{N-1} \int_{I^n} \|\partial_t\chi(t) - \partial X^n\|^2_{L^2(\Omega)}\, dt + C\varepsilon \sum_{n=1}^{N-1} \tau \|\partial(X^n - X^{n-1})\|^2_{L^2(\Omega)} \leq C\tau.\end{aligned}$$

Similar reasoning for $\nabla\theta$ completes the proof. $\square$

Therefore both $\varepsilon\partial_t\chi$ and $\nabla\theta$ belong to the Besov space $B^{1/2}_{2,\infty}(0,T;L^2(\Omega))$, which is defined in terms of (6.1) being valid. This in turn yields the intermediate Sobolev regularity

$$\sqrt{\varepsilon}\partial_t\chi,\ \nabla\theta \in H^{1/2-\delta}(0,T;L^2(\Omega)) \qquad \forall\, \delta > 0.$$

7. Extrapolation Scheme. The semi-implicit scheme yields an optimal order τ of convergence, but requires the solution of a nonlinear elliptic problem. On the other hand, the semi-explicit scheme decouples X^n and Θ^n and leads to a simple algebraic correction for X^n plus a linear elliptic PDE for Θ^n, but at the expense of degrading the order of convergence from τ to $\tau/\sqrt{\varepsilon}$. The purpose of this section is to show that the extrapolation scheme exhibits the advantages of both schemes.

THEOREM 7.1. *There exists $C > 0$, depending on $\|\nabla\theta_0\|_{L^2(\Omega)}$, A in (2.4), $L_{\mathbf{b}}, L_f$, and T but independent of ε and τ, such that $E(\varepsilon,\tau) \le C\tau$ provided $\tau \le \varepsilon/2$.*

Proof. Once again we take $F = 0$ for the sake of simplicity, and examine the effect of extrapolation in (5.9), (5.11), and (5.12). We first notice that T^n in (1.11) can be written as

$$T^n = \Theta^n - \big(\Theta^n - (2\Theta^{n-1} - \Theta^{n-2})\big) = \Theta^n - \tau\partial(\Theta^n - \Theta^{n-1}).$$

If $\bar{\chi}^n := \frac{1}{\tau}\int_{I^n} \chi\, dt$ indicates the average of χ on I^n, then on appealing to (2.20) the penultimate term in (5.9) becomes

$$\tau^2 \sum_{n=1}^{m} \langle \partial\Theta^n - \partial\Theta^{n-1}, \bar{\chi}^n - X^n\rangle = \tau^2 \langle \partial\Theta^m, \bar{\chi}^m - X^m\rangle - \tau^3 \sum_{n=1}^{m} \langle \partial\Theta^{n-1}, \partial(\bar{\chi}^n - X^n)\rangle. \tag{7.1}$$

Since $\tau\|\partial\Theta^m\|^2_{L^2(\Omega)} \le \tau\sum_{n=1}^m \|\partial\Theta^n\|^2_{L^2(\Omega)} \le C$ and $\tau \le \varepsilon/2$, the first term can be bounded as follows with the aid of (5.14):

$$\begin{aligned}
\tau^2\langle \partial\Theta^m, \bar{\chi}^m - X^m\rangle &\le \tau^2\|\partial\Theta^m\|_{L^2(\Omega)}\|\bar{\chi}^m - \chi(t^m)\|_{L^2(\Omega)} + \tau^2\|\partial\Theta^m\|_{L^2(\Omega)}\|e_\chi^m\|_{L^2(\Omega)} \\
&\le C\tau^2 + C\tfrac{\tau^4}{\varepsilon}\|\partial\Theta^m\|^2_{L^2(\Omega)} + \tfrac{\varepsilon}{8}\|e_\chi^m\|_{L^2(\Omega)} \le C\tau^2 + \tfrac{\varepsilon}{8}\|e_\chi^m\|_{L^2(\Omega)}.
\end{aligned}$$

The last term above can be absorbed into the second term on the left hand side of (5.9). The same argument applies to the last term in (5.12), with the only difference that instead of $\partial\Theta^m$ we have $\partial\Theta^m - \partial\Theta^{m-1}$. For the remaining summand in (7.1), we point out that

$$\begin{aligned}
\|\partial(\bar{\chi}^n - X^n)\|_{L^2(\Omega)} &\le \tfrac{1}{\tau^2}\int_{I^n}\int_{t-\tau}^{t} \|\partial_t(\chi - X)\|_{L^2(\Omega)}\, dsdt \\
&\le \tfrac{1}{\tau}\int_{I^{n-1}} \|\partial_t(\chi - X)\|_{L^2(\Omega)}\, dt + \tfrac{1}{\tau}\int_{I^n} \|\partial_t(\chi - X)\|_{L^2(\Omega)} + \|\partial(X^n - X^{n-1})\|_{L^2(\Omega)}\, dt.
\end{aligned}$$

Consequently, in view of (4.1) and $\tau \le \varepsilon/2$, we see that

$$\begin{aligned}
-\tau^3 \sum_{n=1}^{m} \langle \partial\Theta^{n-1}, \partial(\bar{\chi}^n - X^n)\rangle &\le C\tfrac{\tau^4}{\varepsilon}\sum_{n=1}^{m}\|\partial\Theta^{n-1}\|^2_{L^2(\Omega)} + C\tau^2\varepsilon\sum_{n=1}^{m}\|\partial(X^n - X^{n-1})\|^2_{L^2(\Omega)} \\
&\quad + \tfrac{\tau\varepsilon}{4}\sum_{n=1}^{m}\int_{I^n}\|\partial_t(\chi - X)\|^2_{L^2(\Omega)}\, dt \le C\tau^2 + \tfrac{\tau\varepsilon}{4}\sum_{n=1}^{m}\int_{I^n}\|\partial_t(\chi - X)\|^2_{L^2(\Omega)}\, dt,
\end{aligned}$$

and that the last term can be hidden into the left hand side of (5.9). At the same time, on using (2.20) in conjunction with (2.7), (4.1), and (6.1), the last term in (5.9) becomes

$$\tau\sum_{n=1}^{m}\int_{I^n}\langle\partial\Theta^n-\partial\Theta^{n-1},\chi(t+\tfrac{\tau}{2})-\chi(t)\rangle\,dt=\tau\int_{I^m}\langle\partial\Theta^m,\chi(t+\tfrac{\tau}{2})-\chi(t)\rangle\,dt$$

$$+\tau\sum_{n=1}^{m-1}\langle\partial\Theta^n,\int_t^{t+\frac{\tau}{2}}(\partial_t\chi(s+\tau)-\partial_t\chi(s))ds\rangle\,dt\le C\tfrac{\tau^4}{\varepsilon}\sum_{n=0}^{m}\|\partial\Theta^n\|^2_{L^2(\Omega)}$$

$$+\tau^2\varepsilon\|\partial_t\chi\|^2_{L^2(\Omega)}+\tau\varepsilon\int_0^{T-\frac{\tau}{2}}\|\partial_t\chi(t+\tau)-\partial_t\chi(t)\|^2_{L^2(\Omega)}\,dt\le C\tau^2.$$

We finally resort to (2.20) and (4.1) to bound the penultimate term in (5.11) as follows:

$$\tau^2\sum_{n=1}^{m}\langle X^n-X^{n-1},\partial\Theta^n-\partial\Theta^{n-1}\rangle=\tau^3\langle\partial X^m,\partial\Theta^m\rangle-\tau^3\sum_{n=1}^{m}\langle\partial(X^n-X^{n-1}),\partial\Theta^{n-1}\rangle$$

$$\le\tau^2\varepsilon\|\partial X^m\|^2_{L^2(\Omega)}+\tau^2\varepsilon\sum_{n=1}^{m}\|\partial(X^n-X^{n-1})\|^2_{L^2(\Omega)}+C\tfrac{\tau^4}{\varepsilon}\sum_{n=1}^{m}\|\partial\Theta^n\|^2_{L^2(\Omega)}\le C\tau^2.$$

This completes the proof. ∎

8. Degenerate Diffusion with Convection and Reaction. We finish our analysis of time discretizations with a further discussion of the model we started with, namely the Stefan problem with temperature-dependent convection and reaction

$$\partial_t u-\Delta\beta(u)=\operatorname{div}\mathbf{b}(\beta(u))+f(\beta(u)). \tag{8.1}$$

This PDE appears in other models of interest but with different monotone Lipschitz function β. Relevant examples are petroleum reservoir and groundwater diffusion simulation [1], and modeling of reactive solute transport with an equilibrium adsorption process [3].

The semi-implicit discretization reads:

$$\tfrac{1}{\tau}(U^n-U^{n-1})-\Delta\beta(U^n)=\operatorname{div}\mathbf{b}(\beta(U^{n-1}))+f(\beta(U^{n-1})). \tag{8.2}$$

Our preceding stability and error estimates are solely based on monotonicity of Λ, and so apply to (8.2) upon letting $\varepsilon\downarrow 0$. We omit their proofs but point out that the compatibility condition becomes $\theta\in\Lambda(\chi)$ as $\varepsilon\downarrow 0$, which is the natural one for the Stefan problem. The case of enthalpy-dependent convection and reaction is also of interest but does not fit into this context.

THEOREM 8.1. *There exists $C>0$, depending on $\|\nabla\theta_0\|_{L^2(\Omega)}$, L_b, L_f, and T but independent of $m\le N$, such that*

$$\|u(t^m)-U^m\|_{H^{-1}(\Omega)}+\Big(\sum_{n=1}^{m}\int_{I^n}(\|\theta-\Theta^n\|^2_{L^2(\Omega)}+\tau\|\theta-\Theta^n\|^2_{H_0^1(\Omega)})\,dt\Big)^{\frac{1}{2}}\le C\tau.$$

This optimal result improves upon those in [1,3,8,11], which are of order $\mathcal{O}(\sqrt{\tau}\,)$.

Acknowledgement

We would like to thank N. Walkington for bringing Rulla's paper to our attention and for several enlightening discussions.

REFERENCES

[1] T. Argogast, M.F. Wheeler, and N.Y. Zhang, *A nonlinear mixed finite element method for a degenerate parabolic equation arising in flow in porous media*, SIAM J. Numer. Anal. (to appear).

[2] C. Baiocchi, *Discretization of evolution inequalities*, in Partial Differential Equations and the Calculus of Variations, F. Colombini, A. Marino, L. Modica, and S. Spagnolo eds, Birkäuser, Boston, 1989, pp. 59–92.

[3] J.W. Barrett and P. Knabner, *Finite element approximation of transport of reactive solutes in porous media. Part 2: Error estimates for equilibrium adsorption processes*, (to appear).

[4] H. Brezis, *Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations*, in Contributions to Nonlinear Functional Analysis, E. Zarantonello ed., Academic Press, New York (1971), pp. 101–156.

[5] M. G. Crandall and T. M. Liggett, *Generation of semi-groups of nonlinear transformations on general Banach spaces*, Amer. J. Math., 93 (1971), pp. 265–298.

[6] X. Jiang and R. H. Nochetto, *A P^1-P^1 finite element method for a phase relaxation model. Part I: Quasi-uniform mesh*, (to appear).

[7] E. Magenes, R. H. Nochetto and C. Verdi, *Energy error estimates for a linear scheme to approximate nonlinear parabolic problems*, RAIRO Model. Math. Anal. Numer., 21 (1987), pp. 655–678.

[8] R.H. Nochetto, *Error estimates for multidimensional singular parabolic problems*, Japan J. Appl. Math., 4 (1987), pp. 111–138.

[9] R. H. Nochetto, *Finite element methods for parabolic free boundary problems*, in Advances in Numerical Analysis, Vol I: Nonlinear Partial Differential Equations and Dynamical Systems, W. Light ed., 1990 Lancaster Summer School Proceedings, Oxford University Press, 1991, pp. 34–88.

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

[11] R.H. Nochetto and C. Verdi, *Approximation of degenerate parabolic problems using numerical integration*, SIAM J. Numer. Anal., 25 (1988), pp. 784–814.

[12] J. Rulla, *Error analysis for implicit approximations to solutions to Cauchy problems*, SIAM J. Numer. Anal. (to appear).

[13] G. Savare, *Weak solutions and maximal regularity for abstract evolution inequalities*, (to appear).

[14] C. Verdi, *Numerical aspects of parabolic free boundary and hysteresis problems*, in Modelling and Analysis of Phase Transition and Hysteresis Phenomena, A. Visintin (ed.), Springer-Verlag, Berlin, to appear.

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

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

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

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