The Cahn–Hilliard gradient theory for phase separation with non-smooth free energy Part II: Numerical analysis

J. F. BLOWEY and C. M. ELLIOTT

School of Mathematical and Physical Sciences, University of Sussex, Brighton BN1 9QH, UK

(Received 22 March 1991)

In this paper we consider the numerical analysis of a parabolic variational inequality arising from a deep quench limit of a model for phase separation in a binary mixture due to Cahn and Hilliard. Stability, convergence and error bounds for a finite element approximation are proven. Numerical simulations in one and two space dimensions are presented.

1 Introduction

This paper is concerned with the numerical analysis of a parabolic variational inequality introduced by Blowey & Elliott (1991) as a mathematical formulation of a model of phase separation in a binary alloy. We refer to Blowey & Elliott (1991) and the references cited therein for the physical background; for example, our work is relevant to that of Cerezo *et al.* (1989).

Let $\Omega \subset \mathbb{R}^d$ $(1 \leqslant d \leqslant 3)$ be a bounded domain with sufficiently smooth boundary. Given $\gamma > 0$, $u_0 \in K = \{\eta \in H^1(\Omega): \ -1 \leqslant \eta \leqslant 1\}$ with $m \equiv (u_0, 1) \in (-|\Omega|, |\Omega|)$ and $T > 0$, we consider the problems:

(***P***) Find $\{u, w\}$ such that $u \in H^1(0,T;(H^1(\Omega))') \cap L^\infty(0,T;H^1(\Omega))$, $u \in K$ for *a.e.* $t \in (0,T)$ and $w \in L^2(0,T;H^1(\Omega))$

$$\left\langle \frac{\partial u}{\partial t}, \eta \right\rangle_{H^1(\Omega))', H^1(\Omega)} + (\nabla w, \nabla \eta) = 0 \ \forall\, \eta \in H^1(\Omega), \quad a.e.\ t \in (0,T), \tag{1.1a}$$

$$\gamma(\nabla u, \nabla \eta - \nabla u) - (u, \eta - u) \geqslant (w, \eta - u) \ \forall\, \eta \in K, \quad a.e.\ t \in (0,T), \tag{1.1b}$$

and

$$u(0) = u_0. \tag{1.1c}$$

(***Q***) Find $u \in H^1(0,T;(H^1(\Omega))') \cap L^\infty(0,T;H^1(\Omega))$, $u \in K_m$ for *a.e.* $t \in (0,T)$ such that

$$\gamma(\nabla u, \nabla \eta - \nabla u) + \left(\mathscr{G}_N \frac{\partial u}{\partial t}, \eta - u\right) - (u, \eta - u) \geqslant 0 \ \forall\, \eta \in K_m, \quad a.e.\ t \in (0,T), \tag{1.2a}$$

and

$$u(0) = u_0, \tag{1.2b}$$

where $K_m = \{\eta \in K : (\eta, 1) = m\}$, $(.,.)$ is the $L^2(\Omega)$ inner-product and $\mathscr{G}_N$ is the inverse of $-\Delta$ with zero Neumann boundary data. Also, throughout this paper we denote the norm of $H^p(\Omega)$ $(p \geqslant 0)$ by $\|.\|_p$ and the semi-norm $\|D^p\eta\|_0$ by $|\eta|_p$.

https://doi.org/10.1017/S0956792500000759 Published online by Cambridge University Press

In Blowey & Elliott (1991) the following theorem concerning existence, uniqueness and regularity was established.

Theorem 1.1. *Problems* (*P*) *and* (*Q*) *have unique solutions and are equivalent. Furthermore, if* $\partial\Omega$ *is smooth or* Ω *is convex, then* $u \in L^2(0, T; H^2(\Omega))$.

In §2, a fully discrete finite element approximation is proposed. Existence and uniqueness results for a small enough time step independent of h, the space mesh parameter, are given, so that the scheme is well defined. It is shown in §3 that the scheme possesses a Lyapunov functional and satisfies certain stability estimates. Also, it is demonstrated that the scheme simulates the long time asymptotic behaviour of the underlying nonlinear equation. An error bound between the discrete and continuous solutions is given. In §4, an iterative method for solving the algebraic problem at each time step is suggested and shown to be convergent. Numerical simulations are described in §5. Various numerical tests are performed in one space dimension, and some interesting numerical simulations in one and two space dimensions are displayed which exhibit the behaviour expected of the physical problem.

For other papers concerning the numerical analysis and simulations of the Cahn–Hilliard equation with smooth free energy we refer to Elliott & French (1987), Elliott (1989), Elliott *et al.* (1989), French & Nicolaides (1989), Copetti & Elliott (1990), French (1990) and Elliott & Larsson (1991).

2 Discrete evolutionary problem

In this section we consider finite element approximations to (*P*) and (*Q*), denoted by (P^h) and (Q^h), respectively. For $d = 2, 3$, we assume that Ω either has a smooth boundary or is convex polyhedral, so that Theorem 1.1 holds.

2.1 Existence and uniqueness

For $T > 0$ and $M \in \mathbb{N}$, define $\Delta t = T/M$, $t^n = n\Delta t$, $(0 \leqslant n \leqslant M)$, $J^n = (t^{n-1}, t^n]$, $(1 \leqslant n \leqslant M)$ and

$$\partial\eta^n := \frac{\eta^n - \eta^{n-1}}{\Delta t}, \quad 1 \leqslant n \leqslant M, \tag{2.1}$$

for a given sequence $\{\eta^n\}_{n=0}^M$.

We will focus our attention on the discrete schemes (Q_l^h) and(P_l^h) defined as follows:

(P_l^h) Given $U^0 \in K_m^h$, for each $n \geqslant 1$ find $\{U^n, W^n\} \in K^h \times S^h$ such that

$$(\partial U^n, \chi)^h + (\nabla W^n, \nabla\chi) = 0 \ \forall \chi \in S^h, \tag{2.2a}$$

$$\gamma(\nabla U^n, \nabla\chi - \nabla U^n) - (U^n, \chi - U^n)^h \geqslant (W^n, \chi - U^n)^h \ \forall \chi \in K^h, \tag{2.2b}$$

and

$$U^0 = u_0^h. \tag{2.2c}$$

(Q_l^h) Given $U^0 \in K_m^h$, for each $n \geqslant 1$ find $\{U^n\} \in K_m^h$ such that $\forall \chi \in K_m^h$

$$\gamma(\nabla U^n, \nabla\chi - \nabla U^n) + (\mathcal{G}_N^h(\partial U^n), \chi - U^n)^h - (U^n, \chi - U^n)^h \geqslant 0, \tag{2.3a}$$

https://doi.org/10.1017/S0956792500000759 Published online by Cambridge University Press

and
$$U^0 = u_0^h, \tag{2.3b}$$

where u_0^h is an approximation to u_0,

$$S_m^h = \{\chi \in S^h : (\chi, 1)^h = (u_0^h, 1)^h = m\},$$
$$K^h = \{\chi \in S^h : -1 \leqslant \chi \leqslant 1\},$$
$$K_m^h = \{\chi \in S_m^h : -1 \leqslant \chi \leqslant 1\},$$

and we have introduced the following notation:

1. $\mathscr{T}^h$ is a regular family of triangulations for Ω (Ciarlet 1978), consisting of closed simplices τ, with maximum diameter not exceeding h, so that $\bar{\Omega} \equiv \bigcup_{\tau \in \mathscr{T}^h} \tau$. If $\partial\Omega$ is curved, then the boundary elements have at most one curved edge. Associated with $\mathscr{T}^h$ is the finite element space $S^h \subset H^1(\Omega)$
$$S^h = \{\chi \in C^0(\bar{\Omega}) : \chi|_\tau \text{ is linear for } \tau \in \mathscr{T}^h\}. \tag{2.4}$$
2. $(.,.)^h$ denotes an inner-product on $L^2(\Omega)$ or S^h (and $C(\bar{\Omega})$), which is an approximation to the $L^2(\Omega)$ inner-product using appropriate numerical quadrature.
3. The discrete Green's operator approximating the inverse of the Laplacian with Neumann boundary data defined by:
$$\mathscr{G}_N^h \in \mathscr{L}(\mathscr{F}^h, S_0^h),\ S_0^h = \{\chi \in S^h : (\chi, 1)^h = 0\}, \tag{2.5a}$$
$$(\nabla \mathscr{G}_N^h v, \nabla \chi) = (v, \chi)^h\ \forall\, \chi \in S^h, \tag{2.5b}$$
where $\mathscr{F}^h$ is either $\mathscr{F} \cap L^2(\Omega)$ or S_0^h when $(.,.)^h$ is, respectively, the $L^2(\Omega)$ inner-product or an approximation and $\mathscr{F} := \{f \in (H^1(\Omega))' : {}_{(H^1(\Omega))'}\langle f, 1\rangle_{H^1(\Omega)} = 0\}$. Note that $\chi \in \mathscr{F}^h$ satisfies
$$\|\chi\|_{-h}^2 \equiv |\mathscr{G}_N^h \chi|_1^2 = (\mathscr{G}_N^h \chi, \chi)^h = (\chi, \mathscr{G}_N^h \chi)^h. \tag{2.6}$$

The existence and uniqueness of $\mathscr{G}_N^h \chi$ solving (2.5a–b) follows from the discrete Poincaré inequality given in (2.9) below. We assume the following for $\eta, \chi \in S^h$:

$$|(\eta, \chi) - (\eta, \chi)^h| \leqslant C h^{1+r} \|\eta\|_1 \|\chi\|_r \quad r = 0, 1, \tag{2.7a}$$
$$C_1 |\chi|_h^2 \leqslant |\chi|_0^2 \leqslant C_2 |\chi|_h^2, \tag{2.7b}$$
$$(\eta, \chi)^h = (\eta\chi, 1)^h, \tag{2.7c}$$
$$(1, 1)^h = |\Omega|, \tag{2.7d}$$
$$\text{if} \quad \xi \geqslant 0 \quad \text{is continuous then} \quad (\xi, 1)^h \geqslant 0, \tag{2.7e}$$

where C_1 and C_2 are constants independent of h. From (2.7a, b) and the continuous Poincaré inequality, for $\xi \in H^1(\Omega)$,

$$|\xi|_0 \leqslant C_P(|\xi|_1 + |(\xi, 1)|); \tag{2.8}$$

it follows that for h sufficiently small, we obtain the discrete Poincaré inequality

$$((\chi, \chi)^h)^{\frac{1}{2}} \equiv |\chi|_h \leqslant \tilde{C}_P(|\chi|_1 + |(\chi, 1)^h|), \tag{2.9}$$

where $\tilde{C}_P$ is a constant independent of h.

Two possible choices for $(.,.)^h$ are

$$(\chi,\phi)^h = \int_\Omega \chi(x)\,\phi(x)\,\mathrm{d}x, \quad \text{and} \quad (\chi,\phi)^h = \int_\Omega I^h(\chi(x)\,\phi(x))\,\mathrm{d}x, \tag{2.10}$$

where $I^h : C(\bar{\Omega}) \to S^h$ is defined to the unique element in S^h such that $I^h v = v$ at each vertex of $\mathscr{T}^h$.

For $f \in \mathscr{F}^h$, it follows from the discrete Poincaré inequality that

$$|\mathscr{G}^h_N f|_1^2 = (f, G^h_N f)^h \leqslant |f|_h |\mathscr{G}^h_N f|_h \leqslant \tilde{C}_P |f|_h |\mathscr{G}^h_N f|_1,$$

so that

$$\|f\|_{-h} \equiv |\mathscr{G}^h_N f|_1 \leqslant \tilde{C}_P |f|_h. \tag{2.11}$$

Other numerical schemes which approximate $(\boldsymbol{P})$ are:

$(\boldsymbol{P}^h_2)$ Given $U^0 = u^h_0 \in K^h_m$, for each $n \geqslant 1$ find $\{U^n, W^n\} \in K^h \times S^h$ such that

$$(\partial U^n, \chi)^h + (\nabla W^n, \nabla \chi) = 0 \;\forall \chi \in S^h, \tag{2.12a}$$

$$\gamma(\nabla U^n, \nabla\chi - \nabla U^n) - (U^{n-1}, \chi - U^n)^h \geqslant (W^n, \chi - U^n)^h \;\forall \chi \in K^h. \tag{2.12b}$$

$(\boldsymbol{P}^h_3)$ Given $U^0 = u^h_0 \in K^h_m$, for each $n \geqslant 1$ find $\{U^n, W^n\} \in K^h \times S^h$ such that

$$(\partial U^n, \chi)^h + (\nabla W^n, \nabla \chi) = 0 \;\forall \chi \in S^h, \tag{2.13a}$$

$$\gamma(\nabla U^{n-\frac{1}{2}}, \nabla\chi - \nabla U^n) - (U^{n-\frac{1}{2}}, \chi - U^n)^h \geqslant (W^n, \chi - U^n)^h \;\forall \chi \in K^h, \tag{2.13b}$$

where $U^{n-\frac{1}{2}} = (U^n + U^{n-1})/2$.

Remarks Scheme $(\boldsymbol{P}^h_1)$ is fully implicit and $(\boldsymbol{P}^h_3)$ is the Crank–Nicolson scheme; also, scheme $(\boldsymbol{P}^h_2)$ has the advantage of having a unique solution for $\forall \Delta t > 0$ (see Corollary 2.2).

If we observe that $W^n = -\mathscr{G}^h_N(\partial U^n) + \lambda^n$, where $\lambda^n = (W^n, 1)^h / |\Omega|$, then it is clear that $\{U^n, W^n\}$ solving $(\boldsymbol{P}^h_1)$ immediately implies that $\{U^n\}$ solves $(\boldsymbol{Q}^h_1)$. The analogous result is also true for the associated problems $(\boldsymbol{P}^h_i)$ $(i = 2, 3)$. □

Theorem 2.1 *There exist sequences* $\{U^n\}$ *and* $\{U^n, W^n\}$ *satisfying* (2.3a, b) *and* (2.2a–c). *Furthermore, for* $\Delta t < \Delta t^* = 4\gamma$, *the sequence* $\{U^n\}$ *is uniquely defined.*

Proof Existence to $(\boldsymbol{Q}^h_1)$ is proved by consideration of the minimization problem: For each $1 \leqslant n \leqslant M$ fixed, find $U \in K^h_m$ such that

$$\mathscr{J}^h(U) = \min_{\chi \in K^h_m} \mathscr{J}^h(\chi) := \mathscr{E}^h(\chi) + \frac{1}{2\Delta t}\|\chi - U^{n-1}\|^2_{-h},$$

where

$$\mathscr{E}^h(\chi) := \frac{\gamma}{2}|\chi|_1^2 + \frac{|\Omega|}{2} - \frac{1}{2}|\xi|_h^2. \tag{2.14}$$

From the estimate $\mathscr{J}^h(\chi) \geqslant \frac{\gamma}{2}|\chi|_1^2$, a standard minimization argument yields the existence of U, a minimizer of $\mathscr{J}^h$ over K^h_m. The critical points of the minimizer $\mathscr{J}^h$ over K^h_m are given by

(2.3 a), thus proving existence of the sequence $\{U^n\}$. Given $\eta \in S_0^h$, using Young's inequality, $ab \leqslant \frac{1}{\delta}a^2 + \frac{\delta}{4}b^2$ $(\delta > 0)$, we obtain the estimate

$$(\eta, \eta)^h = (\nabla \mathcal{G}_N^h \eta, \nabla \eta) \leqslant |\mathcal{G}_N^h \eta|_1 |\eta|_1,$$
$$\leqslant \frac{1}{\delta}|\mathcal{G}_N^h \eta|_1^2 + \frac{\delta}{4}|\eta|_1^2 = \frac{1}{\delta}\|\eta\|_{-h}^2 + \frac{\delta}{4}|\eta|_1^2. \tag{2.15}$$

We now prove uniqueness. Let U^1 and U^2 be two possible solutions and define $\theta^U = U^1 - U^2$. By substituting $\chi = U_2$ into (2.3a), when U_1 is the solution, and vice-versa, it follows from addition of the resulting inequalities that

$$\gamma|\theta^U|_1^2 + \frac{1}{\Delta t}\|\theta^U\|_{-h}^2 - |\theta^U|_h^2 \leqslant 0,$$

so, setting $\eta = \theta^U$ and $\delta = \Delta t$ in (2.15) we obtain,

$$\left(\gamma - \frac{\Delta t}{4}\right)|\theta^U|_1^2 \leqslant 0,$$

and, since $(\theta^U, 1)^h = 0$, the discrete Poincaré inequality (2.9) implies that U is unique. So for each n, setting $U^n = U$, we have existence and uniqueness to $(\boldsymbol{Q}_l^h)$.

To prove existence to $(\boldsymbol{P}_l^h)$ we consider the problem:

$(\hat{\boldsymbol{S}}_\mu^h)$ Given $\mu \in [\mu_L, \mu_R]$, find $U_\mu \in K^h$ such that $\forall \eta \in K^h$

$$\gamma(\nabla U_\mu, \nabla \eta - \nabla U_\mu) + (U_\mu, \eta - U_\mu)^h \geqslant (f, \eta - U_\mu)^h + \mu(1, \eta - U_\mu)^h, \tag{2.16}$$

where $f = 2U^n - \mathcal{G}_N^h(\partial U^n)$, $\mu_L = -1 - \max_{x\in\Omega} f(x)$, $\mu_R = 1 - \min_{x\in\Omega} f(x)$ and U^n is the unique solution of $(\boldsymbol{Q}_l^h)$.

Since the bilinear form on the left of (2.16) is strictly coercive, there exists a unique solution to $(\hat{\boldsymbol{S}}_\mu^h)$. We define the mapping $\mathcal{M}^h : [\mu_L, \mu_R] \to \mathbb{R}$ by

$$\mathcal{M}^h(\mu) = (U_\mu, 1)^h.$$

Let $\mu_1, \mu_2 \in [\mu_L, \mu_R]$, then setting $\mu = \mu_1, \eta = U_{\mu_2}$ and $\mu = \mu_2, \eta = U_{\mu_1}$ in (2.16) and adding the resulting inequalities, we obtain

$$0 \leqslant \gamma|U_{\mu_1} - U_{\mu_2}|_1^2 + |U_{\mu_1} - U_{\mu_2}|_h^2 \leqslant (\mathcal{M}^h(\mu_1) - \mathcal{M}^h(\mu_2))(\mu_1 - \mu_2). \tag{2.17}$$

From the Cauchy Schwarz inequality and (2.17) we have

$$|\mathcal{M}^h(\mu_1) - \mathcal{M}^h(\mu_2)|^2/|\Omega| \leqslant |U_{\mu_1} - U_{\mu_2}|_h^2 \leqslant (\mathcal{M}^h(\mu_1) - \mathcal{M}^h(\mu_2))(\mu_1 - \mu_2),$$

so that $\mathcal{M}^h$ is monotone and continuous. Now, if we note that for all $x \in \Omega$, where $|\eta(x)| \leqslant 1$, the following trivial inequalities hold

$$(1 - \mu_R - f(x))(\eta(x) - 1) \geqslant 0 \quad \text{and} \quad (-1 - \mu_L - f(x))(\eta(x) + 1) \geqslant 0,$$

then it is clear from (2.7 c, e) that $U_{\mu_R} = 1$ and $U_{\mu_L} = -1$, so that $\mathcal{M}^h(\mu_R) = |\Omega|$ and $\mathcal{M}^h(\mu_L) = -|\Omega|$. It follows from the intermediate value theorem that there exists $\lambda \in [\mu_L, \mu_R]$

https://doi.org/10.1017/S0956792500000759 Published online by Cambridge University Press

such that $\mathscr{M}^h(\lambda) = (U_\lambda, 1)^h = m$. Setting $\chi = U_\lambda$ in (2.3a) and $\eta = U^n$ in (2.16) and adding the resulting inequalities yields

$$\gamma|U^n - U_\lambda|_1^2 + |U^n - U_\lambda|_h^2 \leqslant 0, \tag{2.18}$$

which proves that $U_\lambda = U^n$, and defining $W^n = -\mathscr{G}^h_{\mathscr{N}}(\partial U^n) + \lambda$ we have proved existence for $\{U^n, W^n\}$ solving $(\boldsymbol{P}_l^h)$.

We have already proved uniqueness for U^n, hence we conclude uniqueness for W^n up to the addition of a constant. □

Remark Under certain conditions (see (4.8a, b) below), it is clear that if there exists a point of the triangulation where $|U_i^n| < 1$, then W^n is unique. □

Corollary 2.2 *For schemes* $(\boldsymbol{P}_i^h)$ $(i = 2, 3)$, *there exist sequences* $\{U^n, W^n\}$ *satisfying the associated inequalities. Furthermore, for* $\Delta t < \Delta t^* = 4\gamma$ *and* $+\infty$, *for schemes* $(\boldsymbol{P}_i^h)$ *where* $i = 3$ *and* 2, *respectively, the sequence* $\{U^n\}$ *is uniquely defined.*

Proof The proof is analogous to Theorem 2.1 and is not included.

3 Stability and convergence

Theorem 3.1 *The sequences generated by* (2.2a–c) *and* (2.3a, b) *satisfy*

$$(U^n, 1)^h = (u_0^h, 1)^h, \tag{3.1a}$$

$$\mathscr{E}^h(U^n) - \mathscr{E}^h(U^{n-1}) + \frac{1}{2\Delta t}\|U^n - U^{n-1}\|_{-h}^2 + \left(\frac{\gamma}{2} - \frac{\Delta t}{8}\right)|U^n - U^{n-1}|_1^2 \leqslant 0. \tag{3.1b}$$

Proof It is enough to prove the estimate for $(\boldsymbol{P}_l^h)$. The conservation equation (3.1a) is a direct consequence of taking $\chi = 1$ in (2.2a). Taking $\chi = U^{n-1}$ in (2.3a) and noting that $2a(b-a) = b^2 - a^2 - (a-b)^2$, we obtain

$$-\mathscr{E}^h(U^n) + \mathscr{E}^h(U^{n-1}) - \frac{\gamma}{2}|U^n - U^{n-1}|_1^2 - \frac{1}{\Delta t}\|U^n - U^{n-1}\|_{-h}^2 + \frac{1}{2}|U^n - U^{n-1}|_h^2 \geqslant 0. \tag{3.2}$$

So from (2.15), setting $\eta = U^n - U^{n-1}$ and $\delta = \Delta t$, and (3.2), we obtain (3.1b). □

Corollary 3.2 *For* $\Delta t < \Delta t^*$ *scheme* $(\boldsymbol{P}_l^h)$ *satisfies the following stability estimates:*

$$\max_{n=0,\dots,M} \|U^n\|_1^2 + \sum_{n=1}^{M} \Delta t\|\partial U^n\|_{-h}^2 + \sum_{n=1}^{M} \|U^n - U^{n-1}\|_1^2 \leqslant C, \tag{3.3}$$

$$|(W^n, 1)^h| \leqslant C + \frac{C}{\Delta t^{\frac{1}{2}}}, \tag{3.4}$$

$$\sum_{n=1}^{M} \Delta t\|W^n\|_1^2 \leqslant C, \tag{3.5}$$

where all constants are independent of h and Δt.

Proof Summing (3.1 b) from $n = 1, \dots, k$ $(1 \leqslant k \leqslant M)$ and noting the Poincaré inequality (2.9) gives (3.3). If we note

$$|W^n|_1^2 \equiv \|\partial U^n\|_{-h}^2, \tag{3.6}$$

and use (3.3), we conclude that

$$\sum_{n=1}^{M} \Delta t |W^n|_1^2 = \sum_{n=1}^{M} \Delta t \|\partial U^n\|_{-h}^2 \leqslant C. \tag{3.7}$$

Setting $\chi = 1$ and -1 in (2.2 b) and rearranging we obtain

$$|(W^n, 1)^h| \leqslant |m| - 2\mathscr{E}^h(U^n) + (W^n, U^n)^h + |\Omega|.$$

Also $(W^n, U^n)^h = \left(W^n, U^n - \dfrac{m}{|\Omega|}\right)^h + \dfrac{m}{|\Omega|}(W^n, 1)^h$, so by using the definition of $\mathscr{G}_N^h$ and (2.11) upon rearrangement we obtain

$$\begin{aligned} |(W^n, 1)^h| &\leqslant \frac{|m| - 2\mathscr{E}^h(U^n) + |W^n|_1 \left\| U^n - \dfrac{m}{|\Omega|} \right\|_{-h} + |\Omega|}{1 - \dfrac{|m|}{|\Omega|}}, \\ &\leqslant \frac{|m| - 2\mathscr{E}^h(U^n) + \tilde{C}_P |W^n|_1 \left| U^n - \dfrac{m}{|\Omega|} \right|_h + |\Omega|}{1 - \dfrac{|m|}{|\Omega|}}. \end{aligned} \tag{3.8}$$

Hence, it follows from (3.3) and (3.6) that

$$|(W^n, 1)^h| \leqslant C + C\|\partial U^n\|_{-h} \leqslant C + \frac{C}{\Delta t^{\frac{1}{2}}}. \tag{3.9}$$

Also, using the discrete Poincaré inequality and (3.8) we obtain

$$\|W^n\|_1^2 \leqslant C + C|W^n|_1^2, \tag{3.10}$$

which together with (3.7) yields (3.5). □

Corollary 3.3 *For $\Delta t < \Delta t^*$ each of the numerical schemes (P_i^h) $(i = 2, 3)$ possess the Lyapunov functional $\mathscr{E}^h$. They also satisfy the following stability estimates:*

scheme (P_2^h) satisfies (3.3), (3.4) *and* (3.5)*;*
scheme (P_3^h) satisfies (3.4), (3.5) *and the estimate*

$$\max_{n=0,\dots,M} \|U^n\|_1^2 + \sum_{n=1}^{M} \Delta t \|\partial U^n\|_{-h}^2 \leqslant C.$$

Proof This is proven in an analogous manner to Corollary 3.2 □

https://doi.org/10.1017/S0956792500000759 Published online by Cambridge University Press

Remark The estimate

$$\sum_{n=1}^{M} \|U^n - U^{n-1}\|_1^2 \leqslant C,$$

is an important ingredient in the proof of the error bound in Theorem 3.5 to follow. ☐

Numerical schemes for solving a nonlinear evolution equation over a long time interval should simulate the asymptotic behaviour of the underlying equation. The critical points of $\mathscr{E}^h$ over K_m^h are given by: $U^h \in K_m^h$ such that

$$\gamma(\nabla U^h, \nabla\chi - \nabla U^h) - (U^h, \chi - U^h)^h \geqslant 0 \; \forall \chi \in K_m^h. \tag{3.11}$$

Using a similar technique to that which was used in the proof of Theorem 2.1, we may also consider the equivalent problem:

Find $\{U^h, \lambda^h\} \in K_m^k \times \mathbb{R}$ such that

$$\gamma(\nabla U^h, \nabla\chi - \nabla U^h) - (U^h, \chi - U^h)^h \geqslant (\lambda^h, \chi - U^h)^h \; \forall \chi \in K^h. \tag{3.12}$$

Note that (3.12) is the discrete analogue of the stationary problem considered in Blowey & Elliott (1991): ($\boldsymbol{S}$) Find $\{u, \lambda\} \in K_m \times \mathbb{R}$ such that for all $v \in K$

$$\gamma(\nabla u, \nabla v - \nabla u) - (u, v - u) \geqslant \lambda(1, v - u). \tag{3.13}$$

Theorem 3.4 *Let $\Delta t < \Delta t^*$ and $\{U^n, W^n\}$ be the uniquely defined sequence from* ($\boldsymbol{P}_i^h$) ($i = 1, 2, 3$). *Then there exists a subsequence $\{U^{n_p}, W^{n_p}\}$ which converges to $\{U^h, \lambda^h\}$ solving* (3.12).

Proof The proof will only be shown for ($\boldsymbol{P}_l^h$), as the proof for the other schemes is analogous. The result depends upon (3.1b), which shows that $\mathscr{E}^h(.)$ is a Lyapunov functional for the scheme ($\boldsymbol{P}_l^h$). Implicit use will be made of the fact that K_m^h is finite dimensional, so that bounded sets are compact and norms are equivalent. From Corollary 3.2,

$$\|U^n\|_1 \leqslant C, \quad \|(W^n, 1)^h| \leqslant C + C/\Delta t^{\frac{1}{2}} \quad \text{and} \quad \lim_{n\to\infty} |W^n|_1 = \lim_{n\to\infty} \|\partial U^n\|_{-h} = 0.$$

Hence there exists a subsequence $\{U^{n_p}, W^{n_p}\}$ converging to $\{U^h, \lambda^h\} \in K_m^h \times \mathbb{R}$. Therefore, $\lim_{n_p\to\infty} U^{n_p} = U^h$, $\lim_{n_p\to\infty} W^{n_p} = \lambda^h$ and $\lim_{n_p\to\infty} U^{n_p+1} - U^{n_p} = 0$. We may now pass to the limit in (2.2b) to obtain (3.12) and consequently (3.11). ☐

In the remainder of this subsection we assume that the discrete inner product satisfies

$$(\chi, 1)^h = (\chi, 1) \; \forall \chi \in S^h. \tag{3.14}$$

Two examples of this inner product are given by (2.10).

We introduce a little more notation; define

$$\eta^n(.) := \eta(., t^n), \quad \left(\bar{\eta}^n := \frac{1}{\Delta t}\int_{J^n} \eta(., s)\,\mathrm{d}s, \bar{\eta}^0 := \eta^0\right), \tag{3.15}$$

https://doi.org/10.1017/S0956792500000759 Published online by Cambridge University Press

($n = 0, 1, \ldots, M$) where η is any continuous (integrable) function in time defined on $\Omega \times (0, T)$. Define $P^h : L^2(\Omega) \to S^h$ to be the discrete L^2-projection operator onto S^h satisfying

$$(P^h \eta, \chi)^h = (\eta, \chi) \; \forall \chi \in S^h. \tag{3.16}$$

We note the following estimate from Ciarlet (1978); for $v \in H^2(\Omega)$

$$\|v - I^h v\|_s \leqslant Ch^{2-s} \|v\|_2, \quad s = 0, 1. \tag{3.17}$$

We define $\hat{\mathscr{G}}^h_N : \mathscr{F} \cap L^2(\Omega) \to S^h_0$ to be the discrete inverse Laplacian with zero Neumann boundary data satisfying

$$(\nabla \hat{\mathscr{G}}^h_N f, \nabla \chi) = (f, \chi) \; \forall \chi \in S^h, \tag{3.18}$$

where we have not used numerical integration. As $(\mathscr{T}^h)$ is regular and $\mathscr{G}_N$ is regular (Ciarlet 1978) for $f \in L^2(\Omega) \cap \mathscr{F}$

$$|(\hat{\mathscr{G}}^h_N - \mathscr{G}_N) f|_0 \leqslant Ch^2 |f|_0. \tag{3.19}$$

From (2.7a) and the Poincaré inequality, it follows that for $f \in S^h_0$,

$$\begin{aligned} |\hat{\mathscr{G}}^h_N f - G^h_N f|_1^2 &= (f, \hat{\mathscr{G}}^h_N f - \mathscr{G}^h_N f) - (f, \hat{\mathscr{G}}^h_N f - \mathscr{G}^h_N f)^h, \\ &\leqslant Ch^2 \|f\|_1 |\hat{\mathscr{G}}^h_N f - \mathscr{G}^h_N f|_1. \end{aligned} \tag{3.20}$$

Hence from (3.19), the Poincaré inequality and (3.20) we obtain

$$\begin{aligned} |\mathscr{G}^h_N f - \mathscr{G}_N f|_0 &\leqslant |\mathscr{G}^h_N f - \hat{\mathscr{G}}^h_N f|_0 + |\hat{\mathscr{G}}^h_N f - \mathscr{G}_N f|_0, \\ &\leqslant Ch^2 \|f\|_1 + Ch^2 |f|_0 = Ch^2 \|f\|_1. \end{aligned} \tag{3.21}$$

Note that from (3.19) for $f \in \mathscr{F} \cap L^2(\Omega)$, (3.21) trivially holds, that is (3.21) holds for all $f \in \mathscr{F}^h$.

Given $\eta \in L^2(\Omega)$, using the definition of P^h, (3.19) and (2.7a)

$$\begin{aligned} \|\eta - P^h \eta\|_{-1}^2 &= (\mathscr{G}_N(\eta - P^h \eta), \eta - P^h \eta), \\ &= ((\mathscr{G}_N - \hat{\mathscr{G}}^h_N)(\eta - P^h \eta), \eta - P^h \eta) \\ &\quad + (\hat{\mathscr{G}}^h_N(\eta - P^h \eta), P^h \eta)^h - (\hat{\mathscr{G}}^h_N(\eta - P^h \eta), P^h \eta), \\ &\leqslant Ch^2 |\eta - P^h \eta|_0^2 + Ch \|\hat{\mathscr{G}}^h_N(\eta - P^h \eta)\|_1 |P^h \eta|_0, \end{aligned}$$

it follows from (2.7b), the Poincaré inequality, the inequality $|\hat{\mathscr{G}}^h_N f|_1 \leqslant |\mathscr{G}_N f|_1$, Young's inequality and a kick-back argument that

$$\|\eta - P^h \eta\|_{-1} \leqslant Ch |\eta|_0. \tag{3.22}$$

Error bounds for parabolic variational inequalities have been calculated by Johnson (1976), Berger & Falk (1977), Scarpini & Vivaldi (1978), Colli & Verdi (1985), Fetter (1987) and Vuik (1990). In each they essentially consider the same Dirichlet problem but with a varying obstacle constraint. In Johnson (1976), Berger & Falk (1977), Fetter (1987) and Vuik (1990) they have a fixed single obstacle. In Colli & Verdi (1985) the obstacle constraint is only applied on part of $\partial\Omega$, and in Scarpini & Vivaldi (1978) the double obstacle is time dependent.

Theorem 3.5 *Let $\{u, w\}$ be the unique solution to $(\boldsymbol{P})$ and let $\{U^n, W^n\}$ be the sequence generated by the problem $(\boldsymbol{P}_I^h)$ where $\Delta t < \Delta t^*$. Define*

$$U_{h,\Delta t}(\,.\,,t) := \begin{cases} U^n & \text{for} \quad t \in J^n = (t^{n-1}, t^n], \\ U^0 & \text{if} \quad t = 0, \end{cases} \tag{3.23}$$

and
$$\epsilon_u(x,t) := u(x,t) - U_{h,\Delta t}(x,t). \tag{3.24}$$

If $U^0 = P^h u_0 \in K^h$ and $\Delta t < 2\gamma$, then

$$\|\epsilon_u\|^2_{L^\infty(0,T;(H^1(\Omega))')} + \|\epsilon_u\|^2_{L^2(0,T;H^1(\Omega))} \leqslant C\left(\frac{h^4}{\Delta t} + h^2 + \Delta t\right) =: \sigma(h, \Delta t). \tag{3.25}$$

Proof It is important to note that $\forall n \geqslant 0$

$$(U^n, 1) = (U^n, 1)^h = (U^0, 1)^h = (P^h u_0, 1)^h = (u_0, 1) = m.$$

Consider

$$\begin{aligned}(A) &:= \int_{J^n} \left\langle \frac{\partial u}{\partial t} - \partial U^n, \mathscr{G}_N(u - U^n) \right\rangle \mathrm{d}t \\ &\quad + \int_{J^n} \gamma(\nabla u - \nabla U^n, \nabla u - \nabla U^n)\,\mathrm{d}t - \int_{J^n} (u - U^n, u - U^n)\,\mathrm{d}t, \\ &\equiv (A_1) + (A_2) + (A_3).\end{aligned}$$

We introduce the notation

$$U_m^n := U^n - \frac{m}{|\Omega|} \quad \text{and} \quad u_m^n := u^n - \frac{m}{|\Omega|}\,(n = 0, 1, \ldots, M).$$

We estimate (A_1)

$$\begin{aligned}(A_1) &= \int_{J^n} \left\langle \frac{\partial u}{\partial t} - \partial U^n, \mathscr{G}_N(u - U^n) \right\rangle \mathrm{d}t, \\ &= \int_{J^n} \left\langle \frac{\partial u}{\partial t}, \mathscr{G}_N(u_m) \right\rangle \mathrm{d}t + \int_{J^n} (\partial U^n, \mathscr{G}_N(U_m^n))\,\mathrm{d}t \\ &\quad - (\langle u^n - u^{n-1}, \mathscr{G}_N(U_m^n)\rangle + (U^n - U^{n-1}, \mathscr{G}_N(\bar{u}_m^n))), \\ &= (A_1^1) + (A_1^2) + (A_1^3).\end{aligned}$$

Now

$$\begin{aligned}(A_1^1) &= \int_{J^n} \left\langle \frac{\partial u}{\partial t}, \mathscr{G}_N(u_m) \right\rangle \mathrm{d}t = \int_{J^n} \frac{1}{2}\frac{\partial}{\partial t} \|u_m\|^2_{-1}\,\mathrm{d}t = \frac{1}{2}\|u_m^n\|^2_{-1} - \frac{1}{2}\|u_m^{n-1}\|^2_{-1}, \\ (A_1^2) &= (U^n - U^{n-1}, \mathscr{G}_N(U_m^n)) = \tfrac{1}{2}(\|U_m^n\|^2_{-1} - \|U_m^{n-1}\|^2_{-1} + \|U^n - U^{n-1}\|^2_{-1}), \\ (A_1^3) &= -(\langle u^n - u^{n-1}, \mathscr{G}_N(U_m^n)\rangle + (U^n - U^{n-1}, \mathscr{G}_N(u_m^{n-1})) \\ &\quad + (U^n - U^{n-1}, \mathscr{G}_N(\bar{u}^n - u^{n-1}))), \\ &= -\langle u_m^n, \mathscr{G}_N(U_m^n)\rangle + (U_m^{n-1}, \mathscr{G}_N(u_m^{n-1})) - (U^n - U^{n-1}, \mathscr{G}_N(\bar{u}^n - u^{n-1})).\end{aligned}$$

Consequently, combining (A_1^1), (A_1^2) and (A_1^3) and noting the definition of $\epsilon_u(t)$

$$
\begin{aligned}
(A_1) &= \tfrac{1}{2}\|u^n - U^n\|_{-1}^2 - \tfrac{1}{2}\|u^{n-1} - U^{n-1}\|_{-1}^2 + \tfrac{1}{2}\|U^n - U^{n-1}\|_{-1}^2 \\
&\quad -(U^n - U^{n-1}, \mathscr{G}_N(\bar{u}^n - u^{n-1})), \\
&= \tfrac{1}{2}\|\epsilon_u(t^n)\|_{-1} - \tfrac{1}{2}\|\epsilon_u(t^{n-1})\|_{-1} + \tfrac{1}{2}\|U^n - U^{n-1}\|_{-1}^2 \\
&\quad -(U^n\, U^{n-1}, \mathscr{G}_N(\bar{u}^n - u^{n-1})).
\end{aligned}
$$

Again noting the definition $\epsilon_u(t)$

$$
(A_2) = \gamma \int_{J^n} |u - U^n|_1^2 \,\mathrm{d}t = \gamma \int_{J^n} |\epsilon_u(t)|_1^2 \,\mathrm{d}t,
$$

$$
(A_3) = -\int_{J^n} |u - U^n|_0^2 \,\mathrm{d}t = -\int_{J^n} |\epsilon_u(t)|_0^2 \,\mathrm{d}t.
$$

Hence combining (A_1), (A_2) and (A_3)

$$
\begin{aligned}
(A) = \tfrac{1}{2}\|\epsilon_u(t^n)\|_{-1}^2 - \tfrac{1}{2}\|\epsilon_u(t^{n-1})\|_{-1}^2 + \gamma \int_{J^n} |\epsilon_u(t)|_1^2 \,\mathrm{d}t - \int_{J^n} |\epsilon_u(t)|_0^2 \,\mathrm{d}t \\
+ \tfrac{1}{2}\|U^n - U^{n-1}\|_{-1}^2 - (U^n - U^{n-1}, \mathscr{G}_N(\bar{u}^n - u^{n-1})), \qquad (3.26)
\end{aligned}
$$

and summing from $n = 1, \ldots, k\,(1 \leqslant k \leqslant M)$ yields

$$
\begin{aligned}
\sum_{n=1}^{k} (A) &= \left\{\frac{1}{2}\|\epsilon_u(t^k)\|_{-1}^2 + \gamma \int_0^{t^k} |\epsilon_u(t)|_1^2 \,\mathrm{d}t - \int_0^{t^k} |\epsilon_u(t)|_0^2 \,\mathrm{d}t\right\} - \frac{1}{2}\|\epsilon_u(0)\|_{-1}^2 \\
&\quad + \left\{\frac{1}{2}\sum_{n=1}^{k} \|U^n - U^{n-1}\|_{-1}^2 - \sum_{n=1}^{k} (U^n - U^{n-1}, \mathscr{G}_N(\bar{u}^n - u^{n-1}))\right\}, \\
&= (I) + (II) + (III). \qquad (3.27)
\end{aligned}
$$

We estimate an upper bound for (A). Let $t \in J^n$ and set $\eta = U^n$ in (1.2a) and $\eta = I^h\bar{u}^n$ in (2.2b), so that we obtain the following inequalities

$$
\gamma(\nabla u, \nabla u - \nabla U^n) - (u, u - U^n) + (\mathscr{G}_N \frac{\partial u}{\partial t}, u - U^n) \leqslant 0,
$$

$$
\gamma(\nabla U^n, \nabla U^n - \nabla I^h\bar{u}^n) - (U^n, U^n - I^h\bar{u}^n)^h \leqslant (W^n, U^n - I^h\bar{u}^n)^h,
$$

which imply

$$
\begin{aligned}
(A) \leqslant \Delta t\{\gamma(\nabla U^n, \nabla I^h\bar{u}^n - \nabla\bar{u}^n) - (\partial U^n, \mathscr{G}_N(\bar{u}^n - U^n)) + (W^n, U^n - I^h\bar{u}^n)^h\} \\
+ \Delta t\{-(U^n, I^h\bar{u}^n - \bar{u}^n) + (U^n, I^h\bar{u}^n - U^n) - (U^n, I^h\bar{u}^n - U^n)^h\}. \qquad (3.28)
\end{aligned}
$$

In general $(I^h\bar{u}^n - U^n, 1) \neq 0$, so using the definition of $\mathscr{G}_N^h$, (2.2a) and (3.16) gives

$$
\begin{aligned}
(W^n, U^n - I^h\bar{u}^n)^h &= (W^n, U^n - P^h\bar{u}^h)^h + (W^n, P^h\bar{u}^n - I^h\bar{u}^n)^h, \\
&= (\nabla W^n, \nabla\mathscr{G}_N^h(U^n - P^h\bar{u}^n)) + (W^n, \bar{u}^n) - (W^n, I^h\bar{u}^n)^h, \\
&= -(\partial U^n, \mathscr{G}_N^h(U^n - P^h\bar{u}^n))^h + (W^n, \bar{u}^n) - (W^n, I^h\bar{u}^n)^h, \\
&= (\mathscr{G}_N^h(\partial U^n), \bar{u}^n - U^n) + (W^n, \bar{u}^n - I^h\bar{u}^n) \\
&\quad + (\mathscr{G}_N^h(\partial U^n), U^n) - (\mathscr{G}_N^h(\partial U^n), U^n)^h \\
&\quad + (W^n, I^h\bar{u}^n) - (W^n, I^h\bar{u}^n)^h, \qquad (3.29)
\end{aligned}
$$

and using (3.29) in (3.28) and summing from $n = 1, \ldots, k (1 \leqslant k \leqslant M)$ yields the estimate

$$\begin{aligned}\sum_{n=1}^{k} (A) \leqslant \gamma \sum_{n=1}^{k} \Delta t(\nabla U^n, \nabla I^h u^n - \nabla u^n) - \sum_{n=1}^{k} \Delta t(U^n, I^h u^n - u^n) \\ + \sum_{n=1}^{k} \Delta t(W^n, \bar{u}^n - I^h \bar{u}^n) + \sum_{n=1}^{k} \Delta t((\mathscr{G}_N^h - \mathscr{G}_N)\partial U^n, (u^n - U^n)) \\ + \sum_{n=1}^{k} \Delta t\{(\mathscr{G}_N^h(\partial U^n), U^n) - (\mathscr{G}_N^h(\partial U^n), U^n)^h\} \\ + \sum_{n=1}^{k} \Delta t\{(W^n, I^h \bar{u}^n) - (W^n, I^h \bar{u}^n)^h\} \\ + \sum_{n=1}^{k} \Delta t\{(U^n, I^h \bar{u}^n - U^n) - (U^n, I^h \bar{u}^n - U^n)^h\},\end{aligned} \tag{3.30}$$

which we write as

$$\sum_{n=1}^{k} (A) \leqslant (IV) + (V) + (VI) + (VII) + (VIII) + (IX) + (X). \tag{3.31}$$

Now we estimate terms (II)–(X) in turn, and note that if there is no numerical integration terms $(VIII)$–(X) disappear. We estimate (II) using (3.22)

$$|(II)| = \tfrac{1}{2}\|u^0 - U^0\|_{-1}^2 \leqslant Ch^2 \|u_0\|_0^2 \leqslant Ch^2. \tag{3.32}$$

Before estimating (III) we note that

$$\begin{aligned}\|\bar{u}^n - u^{n-1}\|_{-1} = \left\| \frac{1}{\Delta t} \int_{J^n} (u(s) - u(t^{n-1}))\, \mathrm{d}s \right\|_{-1} = \left\| \frac{1}{\Delta t} \int_{J^n} \int_{t^{n-1}}^{s} u_t(r)\, \mathrm{d}r\, \mathrm{d}s \right\|_{-1}, \\ \leqslant \Delta t^{\frac{1}{2}} \|u_t\|_{L^2(J^n;(H^1(\Omega))')}.\end{aligned} \tag{3.33}$$

So using the Cauchy–Schwarz inequality, the arithmetic-geometric mean inequality, (3.33) and the regularity result $u_t \in L^2(0, T; (H^1(\Omega))')$

$$\begin{aligned}|(III)| &\geqslant \frac{1}{2} \sum_{n=1}^{k} \|U^n - U^{n-1}\|_{-1}^2 - \sum_{n=1}^{k} \|U^n - U^{n-1}\|_{-1} \|\bar{u}^n - u^{n-1}\|_{-1}, \\ &\geqslant -\frac{1}{2} \sum_{n=1}^{k} \|\bar{u}^n - u^{n-1}\|_{-1}^2 \geqslant -\frac{\Delta t}{2} \sum_{n=1}^{k} \|u_t\|_{L^2(J^n;(H^1(\Omega))')}^2, \\ &\geqslant -\frac{\Delta t}{2} \|u_t\|_{L^2(0,T;(H^1(\Omega))')}^2 \geqslant -C\Delta t.\end{aligned} \tag{3.34}$$

We estimate (IV) by first decomposing it into two parts

$$\begin{aligned}(IV) &= \gamma \Delta t \sum_{n=1}^{k} (\nabla U^n - \nabla \bar{u}^n, \nabla I^h \bar{u}^n - \nabla \bar{u}^n) + \gamma \Delta t \sum_{n=1}^{k} (\nabla \bar{u}^n, \nabla I^h \bar{u}^n - \nabla u^n) \\ &= (IV_1) + (IV_2).\end{aligned}$$

https://doi.org/10.1017/S0956792500000759 Published online by Cambridge University Press

Noting $\int_{J^n} \epsilon_u(t)\,\mathrm{d}t = \Delta t\bar{u}^n - \Delta t U^n$, the Cauchy–Schwarz inequality for integrals and sums, (3.17) and Young's inequality, we estimate (IV_1) as follows

$$
\begin{aligned}
|(IV_1)|/\gamma &= \left|\sum_{n=1}^{k}\int_{J^n}(\nabla\epsilon_u(t),\nabla\bar{u}^n-\nabla I^h\bar{u}^n)\,\mathrm{d}t\right| \leqslant \sum_{n=1}^{k}\int_{J^n}|\epsilon_u(t)|_1\,|\bar{u}^n-I^h\bar{u}^n|_1\,\mathrm{d}t,\\
&\leqslant Ch\sum_{n=1}^{k}\int_{J^n}|\epsilon_u(t)|_1\,\mathrm{d}t\,\|\bar{u}^n\|_2\\
&\leqslant Ch\sum_{n=1}^{k}\Delta t^{\frac{1}{2}}\left(\int_{J^n}|\epsilon_u(t)|_1^2\,\mathrm{d}t\right)^{\frac{1}{2}}\frac{1}{\Delta t^{\frac{1}{2}}}\left(\int_{J^n}\|u(t)\|_2^2\,\mathrm{d}t\right)^{\frac{1}{2}},\\
&\leqslant Ch\left(\sum_{n=1}^{k}\int_{J^n}|\epsilon_u(t)|_1^2\,\mathrm{d}t\right)^{\frac{1}{2}}\left(\sum_{n=1}^{k}\int_{J^n}\|u(t)\|_2^2\,\mathrm{d}t\right)^{\frac{1}{2}},\\
&\leqslant \frac{Ch^2}{\nu}\|u\|^2_{L^2(0,T;H^2(\Omega))}+\frac{\nu}{2}\int_0^{t^k}|\epsilon_u(t)|_1^2\,\mathrm{d}t.
\end{aligned}
\tag{3.35}
$$

Using Theorem 1.1, integration by parts, (3.17) and the Cauchy–Schwarz inequality for sums and integrals, we estimate (IV_2) as follows

$$
\begin{aligned}
|(IV_2)|/\gamma &= \left|\sum_{n=1}^{k}\int_{J^n}\int_{\Omega}\nabla u(t)\,.\,\nabla(I^h\bar{u}^n-\bar{u}^n)\,\mathrm{d}x\,\mathrm{d}t\right|\\
&= \left.\sum_{n=1}^{k}\int_{J^n}\int_{\Omega}\Delta u(t)(\bar{u}^n-I^h\bar{u}^n)\,\mathrm{d}x\,\mathrm{d}t\right|,\\
&\leqslant \sum_{n=1}^{k}\int_{J^n}\|u(t)\|_2\,|\bar{u}^n-I^h\bar{u}^n|_0\,\mathrm{d}t\\
&\leqslant Ch^2\sum_{n=1}^{k}\Delta t^{\frac{1}{2}}\left(\int_{J^n}\|u(t)\|_2^2\,\mathrm{d}t\right)^{\frac{1}{2}}\frac{1}{\Delta t^{\frac{1}{2}}}\left(\int_{J^n}\|u(t)\|_2^2\,\mathrm{d}t\right)^{\frac{1}{2}},\\
&\leqslant Ch^2\|u\|^2_{L^2(0,T;H^2(\Omega))}\leqslant Ch^2,
\end{aligned}
\tag{3.36}
$$

hence

$$
|(IV)| \leqslant \frac{Ch^2}{\nu}+\frac{\gamma\nu}{2}\int_0^{t^k}|\epsilon_u(t)|_1^2\,\mathrm{d}t. \tag{3.37}
$$

We estimate (V) using (3.17), the Cauchy–Schwarz inequality for integral and sums, (3.3) and the regularity result $u\in L^2(0,T;H^2(\Omega))$

$$
\begin{aligned}
|(V)| &\leqslant \sum_{n=1}^{k}|U^n|_0|I^h\int_{J^n}u-\int_{J^n}u|_0 \leqslant Ch^2\sum_{n=1}^{k}|U^n|_0\left\|\int_{J^n}u\right\|_2,\\
&\leqslant Ch^2\Delta t^{\frac{1}{2}}\sum_{n=1}^{k}|U^n|_0\,\|u\|_{L^2(J^n;H^2(\Omega))},\\
&\leqslant Ch^2\left(\sum_{n=1}^{k}\Delta t|U^n|_0^2\right)^{\frac{1}{2}}\left(\sum_{n=1}^{k}\|u\|^2_{L^2(J^n;H^2(\Omega))}\right)^{\frac{1}{2}},\\
&\leqslant Ch^2\|u\|_{L^2(0,T;H^2(\Omega))}\leqslant Ch^2.
\end{aligned}
\tag{3.38}
$$

https://doi.org/10.1017/S0956792500000759 Published online by Cambridge University Press

We estimate (VI) using (3.17), the Cauchy–Schwarz inequality for integrals and sums, (3.5) and the regularity result $u \in L^2(0,T;H^2(\Omega))$

$$
\begin{aligned}
|(VI)| &\leqslant \Delta t \sum_{n=1}^{k} |W^n|_0 |\bar{u}^n - I^h \bar{u}^n|_0, \leqslant Ch^2 \Delta t \sum_{n=1}^{k} |W^n|_0 \, \|u^n\|_2, \\
&\leqslant Ch^2 \left(\sum_{n=1}^{k} \Delta t |W^n|_0^2 \right)^{\frac{1}{2}} \left(\sum_{n=1}^{k} \|u\|^2_{L^2(J^n;H^2(\Omega))} \right)^{\frac{1}{2}}, \\
&\leqslant Ch^2 \, \|u\|_{L^2(0,T;H^2(\Omega))} \leqslant Ch^2.
\end{aligned} \tag{3.39}
$$

We estimate (VII) using $\bar{u}^n - U^n = \dfrac{1}{\Delta t} \displaystyle\int_{J^n} \epsilon_u(t)\,\mathrm{d}t$, (3.21), the Cauchy–Schwarz inequality for sums and integrals, the Poincaré inequality and (3.3)

$$
\begin{aligned}
|(VII)| &\leqslant \frac{1}{\Delta t} \left| \sum_{n=1}^{k} \int_{J^n} ((\mathscr{G}^h_{\mathcal{N}} - \mathscr{G}_{\mathcal{N}})(U^n - U^{n-1}), \epsilon_u(t))\,\mathrm{d}t \right| \\
&\leqslant \frac{Ch^2}{\Delta t} \sum_{n=1}^{k} \|U^n - U^{n-1}\|_1 \Delta t^{\frac{1}{2}} \left(\int_{J^n} |\epsilon_u(t)|_1^2 \,\mathrm{d}t \right)^{\frac{1}{2}}, \\
&\leqslant \frac{Ch^2}{\Delta t^{\frac{1}{2}}} \left(\sum_{n=1}^{k} \|U^n - U^{n-1}\|_1^2 \right)^{\frac{1}{2}} \left(\int_0^{t^k} |\epsilon_u(t)|_1^2 \,\mathrm{d}t \right)^{\frac{1}{2}}, \\
&\leqslant \frac{Ch^4}{\gamma \nu \Delta t} + \frac{\nu\gamma}{2} \int_0^{t^k} |\epsilon_u(t)|_1^2 \,\mathrm{d}t.
\end{aligned} \tag{3.40}
$$

The terms $(VIII)$–(X) are simply due to the numerical integration, so if we note that for χ^n, $\phi^n \in S^h (1 \leqslant n \leqslant M)$, by using (2.7a) and the Cauchy–Schwarz inequality for sums

$$
\left| \sum_{n=1}^{k} \Delta t \{(\chi^n, \phi^n) - (\chi^n, \phi^n)^h\} \right| \leqslant Ch^2 \left(\sum_{n=1}^{k} \Delta t \|\chi^n\|_1^2 \right)^{\frac{1}{2}} \left(\sum_{n=1}^{k} \Delta t \|\phi^n\|_1^2 \right)^{\frac{1}{2}},
$$

so that noting the Poincaré inequality, (3.3), (3.5), $I^h u^n = u^n + (I^h u^n - u^n)$, (3.17) with $s = 1$, so that we are using the regularity $u \in L^2(0,T;H^2(\Omega))$, and

$$
\left(\sum_{n=1}^{k} \Delta t \|\bar{u}^n\|_i^2 \right)^{\frac{1}{2}} \leqslant \left(\sum_{k=1}^{k} \|u\|^2_{L^2(I^n;H^i(\Omega))} \right)^{\frac{1}{2}} \leqslant \|u\|_{L^2(0,T;H^i(\Omega))} \leqslant C, \quad i = 1,2,
$$

it follows that

$$
(VIII) + (IX) + (X) \leqslant Ch^2.
$$

Hence using (3.27), (3.31) and the estimates for (II)–(X), $(1 \leqslant k \leqslant M)$, we obtain

$$
\tfrac{1}{2}\|\epsilon_u(t^k)\|_{-1}^2 + \gamma \int_0^{t^k} |\epsilon_u(t)|_1^2 \,\mathrm{d}t - \int_0^{t^k} |\epsilon_u(t)|_0^2 \,\mathrm{d}t \leqslant \sigma(h, \Delta t) + \nu\gamma \int_0^{t^k} |\epsilon_u(t)|_1^2 \,\mathrm{d}t. \tag{3.41}
$$

As $(\epsilon_u, 1) = 0$, for all $\delta > 0$

$$
|\epsilon_u|_0^2 \leqslant \delta\gamma \, |\epsilon_u|_1^2 + \frac{1}{4\gamma\delta} \|\epsilon_u\|_{-1}^2,
$$

hence from (3.41) we obtain

$$\frac{1}{2}\|\epsilon_u(t^k)\|_{-1}^2+\gamma(1-\delta-\nu)\int_0^{t^k}|\epsilon_u(t)|_1^2\,\mathrm{d}t \leqslant \frac{1}{4\gamma\delta}\sum_{n=1}^{k}\int_{J^n}\|\epsilon_u(t)\|_{-1}^2\,\mathrm{d}t+\sigma(h,\Delta t). \quad (3.42)$$

For $t\in J^n$, noting the Poincaré inequality and Cauchy–Schwarz inequality for integrals

$$\begin{aligned}\|\epsilon_u(t^n)-\epsilon_u(t)\|_{-1}^2 &= \|u(t^n)-u(t)\|_{-1}^2,\\ &\leqslant \left\|\int_t^{t^n} u_t\,\mathrm{d}s\right\|_{-1}^2 \leqslant \Delta t\|u_t\|^2_{L^2(J^n;(H^1(\Omega))')}.\end{aligned} \quad (3.43)$$

So letting $\delta = 1-\nu$ in (3.42), using the inequality $(a+b)^2 \leqslant (1+\zeta^{-1})\,a^2+(1+\zeta)\,b^2\,(\zeta>0)$ and (3.43) yields

$$\begin{aligned}\frac{1}{2}\|\epsilon_u(t^k)\|_{-1}^2 &\leqslant \frac{1}{4\gamma\delta}\sum_{n=1}^{k}\int_{J^n}((1+\zeta^{-1})\,\|\epsilon_u(t^n)-\epsilon_u(t)\|_{-1}^2+(1+\zeta)\,\|\epsilon_u(t^n)\|_{-1}^2)\,\mathrm{d}t+\sigma(h,\Delta t),\\ &\leqslant \frac{(1+\zeta^{-1})\Delta t^2}{4\gamma\delta}\|u_t\|^2_{L^2(0,t^k;(H^1(\Omega))')}+\frac{(1+\zeta)\,\Delta t}{4\gamma\delta}\sum_{n=1}^{k}\|\epsilon_u(t^n)\|_{-1}^2+\sigma(h,\Delta t).\end{aligned} \quad (3.44)$$

Thus for $\Delta t < 2\gamma$, there exist δ close to 1 and ζ close to zero so that $C_{\zeta,\delta} = \dfrac{1}{2}-\dfrac{(1+\zeta)\Delta t}{4\gamma\delta} > 0$,

and hence

$$\|\epsilon_u(t^k)\|_{-1}^2 \leqslant C_{\zeta,\delta}^{-1}\left(\frac{(1+\zeta)\,\Delta t}{4\gamma\delta}\sum_{n=1}^{k-1}\|\epsilon_u(t^n)\|_{-1}^2+\sigma(h,\Delta t)\right). \quad (3.45)$$

Now using the discrete Grönwall inequality, it follows that

$$\|\epsilon_u(t^k)\|_{-1}^2 \leqslant \sigma(h,\Delta t)\exp\left(\frac{1+\zeta}{4\gamma\delta C_{\zeta,\delta}}\sum_{n=1}^{k-1}\Delta t\right) = \sigma(h,\Delta t)\ \forall\, 1\leqslant k\leqslant M. \quad (3.46)$$

For $t\in J^k$, using the arithmetic-geometric mean inequality and (3.43)

$$\begin{aligned}\|\epsilon_u(t)\|_{-1}^2 &\leqslant 2\|\epsilon_u(t^k)\|_{-1}^2+2\|\epsilon_u(t)-\epsilon_u(t^k)\|_{-1}^2,\\ &\leqslant \sigma(h,\Delta t)+2\Delta t\|u_t\|^2_{L^2(J^k;(H^1(\Omega))')} \leqslant \sigma(h,\Delta t)+C\Delta t,\\ &\leqslant \sigma(h,\Delta t).\end{aligned} \quad (3.47)$$

For $t\in J^k$, setting $\delta=\nu=\frac{1}{4}$ and using (3.47) yields

$$\begin{aligned}\gamma\int_0^t|\epsilon_u(s)|_1^2\,\mathrm{d}s \leqslant \gamma\int_0^{t^k}|\epsilon_u(s)|_1^2\,\mathrm{d}s &\leqslant \frac{2}{\gamma}\int_0^{t^k}\|\epsilon_u(s)\|_{-1}^2\,\mathrm{d}s+\sigma(h,\Delta t),\\ &\leqslant \sigma(h,\Delta t).\end{aligned}$$

□

Corollary 3.6 *For $\Delta t<2\gamma$ scheme ($\mathbf{P}_2^h$) satisfies the error bound* (3.25) *where* $\sigma(h,\Delta t) = C(h^4/\Delta t+h^2+\Delta t)$.

Proof If we note that

$$\Delta t(U^{n-1},U^n-I^h\bar{u}^n)^h = \Delta t\{(U^n,U^n-I^h\bar{u}^n)^h+(U^{n-1}-U^n,U^n-I^h\bar{u}^n)^h\},$$

then we may adapt the proof of Theorem 3.5 to yield the result. □

https://doi.org/10.1017/S0956792500000759 Published online by Cambridge University Press

Remark From the proof of Theorem 3.5, it is clear that the constant in (3.25) is of the form $C = c\exp^{1/\gamma}$, and since γ is a small positive parameter (see Blowey & Elliott 1991), this error bound is not effective unless h and Δt are sufficiently small. This may or may not be practical.

4 An iterative method for solving (P_1^h)

Many methods to solve algebraic problems arising from discretizations of variational inequalities can be found in Glowinski *et al.* (1981). We wish to use a method which exploits:

(i) the fact that the algebraic problem arises from a discretization of a partial differential equation,
(ii) the size and sparsity of the matrices involved.

We introduce $(\phi_i)_{i=1}^D$ as a piecewise linear basis for $\mathscr{T}^h$, where D is the number of vertices of the triangulation, with the property that $\phi_i(x_j) = \delta_{ij}\,(1 \leqslant i,j \leqslant D)$ where the x_j's are the vertices of the triangulation.

Define the matrices K and $M \in \mathbb{R}^{D\times D}$ to be

$$K_{ij} = (\nabla\phi_i, \nabla\phi_j) \quad \text{and} \quad M_{ij} = (\phi_i, \phi_j)^h,$$

and the vector $\underline{e} \in \mathbb{R}^D$ defined by

$$e_i = (\phi_i, 1)^h.$$

We assume that M_{ij} is a diagonal matrix, $M_{ii} > 0$, and write $\chi \in S^h$ as $\underline{\chi} \in \mathbb{R}^D$ where $\chi_i = \chi(x_i)$; we hope that no confusion will arise between the two equivalent notations.

We have found it convenient to use an iterative procedure for solving double obstacle problems based on the method of duality. A similar method for solving a single obstacle problem with an integral constraint was used by Chakrabarti (1988).

It is convenient to observe that the algebraic form of (2.5a, b) is as follows. Given $\underline{v} \in \mathbb{R}^D$ such that

$$\underline{e}^T\underline{v} = 0, \tag{4.1}$$

find $\hat{\underline{v}} \in \mathbb{R}^D$ satisfying

$$K\hat{\underline{v}} = M\underline{v}, \tag{4.2a}$$

and

$$\underline{e}^T\hat{\underline{v}} = 0. \tag{4.2b}$$

There exists a unique solution $\hat{\underline{v}}$ to (4.2a, b) and this implicitly defines the strictly coercive linear vector-space operator $G: \mathbb{S}_0 \to \mathbb{S}_0 = \{\underline{v} \in \mathbb{R}^D : \underline{e}^T\underline{v} = 0\}$ defined by

$$\hat{\underline{v}} \equiv G\underline{v}. \tag{4.3}$$

Since $\hat{v} = \mathscr{G}_N^h v$, (2.15) gives us a norm inequality for these matrices: – for $\delta > 0$

$$\underline{v}^T M\underline{v} \leqslant \frac{1}{\delta}\hat{\underline{v}}^T K\hat{\underline{v}} + \frac{\delta}{4}\underline{v}^T K\underline{v}. \tag{4.4}$$

If we note that $\chi \in S^h$ can be written as

$$\chi(x) = \sum_{i=1}^{D} \chi_i\phi_i(x),$$

and $W^n = -\mathscr{G}_N^h(\partial U^n) + \lambda^n$, where $\lambda^n = (W^n, 1)^h/|\Omega|$, then $(\boldsymbol{P}_l^h)$ may be written as:

Given $\underline{U}^0 \in \mathbb{R}^D$, where $|U_i^0| \leqslant 1\,(1 \leqslant i \leqslant D)$ and $\underline{e}^T \underline{U}^0 = m$, for each $n \geqslant 1$ find $\{\underline{U}^n, \lambda^n\} \in \mathbb{R}^D \times \mathbb{R}$ such that $\forall\, \underline{\eta} \in \mathbb{R}^D$, where $|\eta_i| \leqslant 1\,(1 \leqslant i \leqslant D)$,

$$\sum_{i,j=1}^{D} (\gamma U_i^n K_{ij} - \hat{w}_i^n M_{ij} - U_i^n M_{ij} - \lambda^n 1_i M_{ij})(\eta_j - U_j^n) \geqslant 0, \tag{4.5a}$$

$$\underline{\hat{w}}^n = -G\left(\frac{1}{\Delta t}(\underline{U}^n - \underline{U}^{n-1})\right), \tag{4.5b}$$

$$\underline{e}^T \underline{U}^n = m, \tag{4.5c}$$

$$|U_i^n| \leqslant 1 \quad \text{for} \quad (1 \leqslant i \leqslant D), \tag{4.5d}$$

where $\underline{1}$ is the vector with 1 in each component and $\underline{U}^0 = \underline{u}_0^h$.

It is convenient to set

$$\underline{R}^n := M^{-1}(\gamma K \underline{U}^n - M\underline{U}^n - M(\underline{\hat{w}}^n + \lambda^n \underline{1})); \tag{4.6}$$

it follows that (4.5a) may be rewritten as

$$(M\underline{R}^n)^T (\underline{\eta} - \underline{U}^n) \geqslant 0. \tag{4.7}$$

Fixing i and taking

$$\eta_j = \begin{cases} U_j^n, & \text{if} \quad j \neq i, \\ 2U_i^n - \operatorname{sgn}(U_i^n),\ \operatorname{sgn}(U_i^n) \quad \text{and} \quad U_i^n - \operatorname{sgn}(U_i^n), & \text{if} \quad j = i, \end{cases}$$

respectively, in (4.7) yields

$$M_{ii} R_i^n (U_i^n - \operatorname{sgn}(U_i^n)) \geqslant 0\ \forall\, i,$$
$$M_{ii} R_i^n (\operatorname{sgn}(U_i^n) - U_i^n) \geqslant 0\ \forall\, i,$$
$$-M_{ii} R_i^n \operatorname{sgn}(U_i^n) \geqslant 0\ \forall\, i,$$

which implies that

$$M_{ii} R_i^n (U_i^n - \operatorname{sgn}(U_i^n)) = 0\ \forall\, i, \tag{4.7a}$$

$$M_{ii} \operatorname{sgn}(U_i^n) R_i^n \leqslant 0\ \forall\, i. \tag{4.7b}$$

Further, if the vectors $\underline{R}^n$ and $\underline{U}^n$ satisfy (4.7a, b) and $|U_i^n| \leqslant 1$, then:

$$U_i^n = 1 \Rightarrow M_{ii} R_i^n (\eta_i - U_i^n) \geqslant 0\ \forall\, |\eta_i| \leqslant 1,$$
$$U_i^n = -1 \Rightarrow M_{ii} R_i^n (\eta_i - U_i^n) \geqslant 0\ \forall\, |\eta_i| \leqslant 1,$$
$$|U_i^n| < 1 \Rightarrow M_{ii} R_i^n = 0.$$

It follows that (4.5a–d) is equivalent to the following. Given $\underline{U}^0 \in \mathbb{R}^D$, $|U_i^0| \leqslant 1\ \forall\, i$, for each $n \geqslant 1$ find $\{\underline{U}^n, \underline{R}^n, \lambda^n\} \in \mathbb{R}^D \times \mathbb{R}^D \times \mathbb{R}$ such that

$$\underline{R}^n := M^{-1}(\gamma K \underline{U}^n - M\underline{U}^n - M(\underline{\hat{w}}^n + \lambda^n \underline{1})), \tag{4.8a}$$

$$M_{ii} R_i^n (U_i^n - \operatorname{sgn}(U_i^n)) = 0\ \forall\, i, \tag{4.8b}$$

$$M_{ii} \operatorname{sgn}(U_i^n) R_i^n \leqslant 0\ \forall\, i, \tag{4.8c}$$

$$\underline{\hat{w}}^n = -G\left(\frac{1}{\Delta t}(\underline{U}^n - \underline{U}^{n-1})\right), \tag{4.8d}$$

$$\underline{e}^T \underline{U}^n = m, \tag{4.8e}$$

$$|U_i^n| \leqslant 1 \quad \text{for} \quad (1 \leqslant i \leqslant D), \tag{4.8f}$$

where $\underline{U}^0 = \underline{u}_0^h$.

Proposition 4.1 *Let ρ be a given positive constant. Suppose that for n fixed $\{\underline{u}, \underline{r}, \lambda\} \in \mathbb{R}^D \times \mathbb{R}^D \times \mathbb{R}$ satisfy*

$$\gamma K\underline{u} - M\underline{u} - M(\underline{\hat{w}} + \lambda \underline{1}) = M\underline{r}, \tag{4.9a}$$

$$\underline{\hat{w}} = -G\left(\frac{1}{\Delta t}(\underline{u} - \underline{U}^{n-1})\right), \tag{4.9b}$$

$$\underline{e}^T \underline{u} = m, \tag{4.9c}$$

$$\underline{r} = (\underline{r} - \rho(\underline{u} - \underline{1}))^- + (\underline{r} - \rho(\underline{u} + \underline{1}))^+, \tag{4.9d}$$

where (4.9d) *is understood componentwise,* $(x)^- = \min(x, 0)$ *and* $(x)^+ = \max(x, 0)$. *Then* $\{\underline{U}^n, \underline{R}^n, \lambda^n\}$ *solves* (4.9a–d) *and* $\{\underline{u}, \underline{r}, \lambda\}$ *solves* (4.5a–d).

Proof Let $1 \leqslant i \leqslant D$ be fixed, if we note that either $(r_i - \rho(u_i - 1))^- = 0$ or $(r_i - \rho(u_i + 1))^+ = 0$, then it follows from

$$r_i = (r_i - \rho(u_i - 1))^- + (r_i - \rho(u_i + 1))^+ \tag{4.10}$$

that

$$\left.\begin{array}{lll} \text{if} \ \ r_i > 0 & \text{then} & u_i = -1, \\ \text{if} \ \ r_i < 0 & \text{then} & u_i = 1, \\ \text{if} \ \ r_i = 0 & \text{then} & |u_i| \leqslant 1, \end{array}\right\} \quad \text{and} \quad \left.\begin{array}{lll} \text{if} \ \ u_i = -1 & \text{then} & r_i \geqslant 0, \\ \text{if} \ \ u_i = 1 & \text{then} & r_i \leqslant 0, \\ \text{if} \ \ |u_i| < 1 & \text{then} & r_i = 0. \end{array}\right\} \tag{4.11a–f}$$

Hence

$$M_{ii} r_i (u_i - \operatorname{sgn}(u_i)) = 0, \tag{4.12a}$$

$$M_{ii} \operatorname{sgn}(u_i) r_i \leqslant 0, \tag{4.12b}$$

$$|u_i| \leqslant 1. \tag{4.12c}$$

Likewise, it is easy to check that if u_i and r_i satisfy (4.12a–c), u_i, r_i also satisfy (4.10). □

Remark From (4.2a) and (4.3)

$$KG\left(\frac{1}{\Delta t}(\underline{u} - \underline{U}^{n-1})\right) = M\left(\frac{1}{\Delta t}(\underline{u} - \underline{U}^{n-1})\right),$$

and since K has a simple eigenvalue of 0 with corresponding eigenvector $\underline{1}$, it is easy to show that if $\underline{u}$ solves

$$\gamma M^{-1} K M^{-1} K \underline{u} + \frac{1}{\Delta t}(\underline{u} - \underline{U}^{n-1}) - M^{-1} K \underline{u} = M^{-1} K \underline{r}, \tag{4.13a}$$

$$\underline{r} = (\underline{r} - \rho(\underline{u} - \underline{1}))^- + (\underline{r} - \rho(\underline{u} + \underline{1}))^+, \tag{4.13b}$$

then $\underline{u}$ solves (4.9a–b). Also, premultiplication of (4.13a) by $\underline{1}^T M = \underline{e}^T$ yields (4.9c). If we multiply (4.9a) by $M^{-1} K M^{-1}$ then we obtain (4.13a); that is, the two formulations are equivalent. □

https://doi.org/10.1017/S0956792500000759 Published online by Cambridge University Press

A natural iteration to solve (4.9a–d) is the following:

(A): Given $(\underline{U}^{n-1}, \underline{r}^0)$ for $k \geqslant 1$ find $\{\underline{u}^k, \underline{r}^k, \lambda^k\}$ such that

$$\gamma K\underline{u}^k - M\underline{\hat{w}}^k - M\underline{u}^k - \lambda^k M\underline{1} = M\underline{r}^{k-1}, \tag{4.14a}$$

$$\underline{\hat{w}}^k = -\frac{1}{\Delta t} G(\underline{u}^k - \underline{U}^{n-1}), \tag{4.14b}$$

$$\underline{e}^T \underline{u}^k = m, \tag{4.14c}$$

$$\underline{r}^k = (\underline{r}^{k-1} - \rho(\underline{u}^k - \underline{1}))^- + (\underline{r}^{k-1} - \rho(\underline{u}^k + \underline{1}))^+. \tag{4.14d}$$

We first show that the sequence is well defined. For $\Delta t < 4\gamma$ existence is proved by consideration of the unconstrained minimization problem:

Find $\underline{u}^k \in \mathbb{R}^D$ such that $\underline{e}^T \underline{u}^k = m$ and

$$J(\underline{u}^k) = \min_{\underline{\chi} \in \mathbb{R}^D : \underline{e}^T \underline{\chi} = m} J(\underline{\chi}), \tag{4.15}$$

where
$$J(\underline{\chi}) := \frac{\gamma}{2} \underline{\chi}^T K \underline{\chi} + \frac{1}{2\Delta t} (\underline{\hat{\chi}}^{n-1})^T K \underline{\hat{\chi}}^{n-1} - \frac{1}{2} \underline{\chi}^T M \underline{\chi} - (\underline{r}^{k-1})^T M \underline{\chi},$$

and $\hat{\chi}^{n-1} = G(\underline{\chi} - \underline{U}^{n-1})$. From (4.4) for $\Delta t < 4\gamma$, $J(.)$ is bounded below, hence the existence of a minimizer $\underline{u}^k$ is proved by standard arguments. The associated Euler–Lagrange equation is

$$\gamma K\underline{u}^k + \frac{1}{\Delta t} MG(\underline{u}^k - U^n) - M\underline{u}^k - \lambda^k M\underline{1} = M\underline{r}^{k-1}, \tag{4.16a}$$

where
$$\lambda^k = -\frac{1}{|\Omega|}(m + (r^{k-1}, 1)^h). \tag{4.16b}$$

Uniqueness follows for $\Delta t < 4\gamma$. So $\underline{u}^k$ and λ^k are well defined, and from (4.14d), $\underline{r}^k$ is well defined.

Theorem 4.2 *For* $0 < \rho < 2C$, *where* $C = \tilde{C}_P^{-2}(\gamma - \Delta t/4)$, *the sequence* $\{\underline{u}^k\}$ *defined by* (4.14a–d) *converges to* $\underline{u}$ *the solution to* (4.9a–d).

Proof We have adapted the proof for the method of duality from Glowinski *et al.* (1981). If $r_i r_i^k \geqslant 0$ then from (4.9d), (4.14d) and the inequalities

$$|(x)^+ - (y)^+| \leqslant |x - y|, \quad |(x)^- - (y)^-| \leqslant |x - y|,$$

it follows that

$$|r_i^k - r_i| \leqslant |r_i^{k-1} - r_i - \rho(u_i^k - u_i)|.$$

If $r_i r_i^k < 0$, let us take $r_i > 0$ and $r_i^k < 0$, then

$$\begin{aligned} 0 > r_i^k - r_i &= r_i^{k-1} - \rho(u_i^k - 1) - (r_i - \rho(u_i + 1)), \\ &= r_i^{k-1} - r_i - \rho(u_i^k - u_i) + 2\rho \geqslant r_i^{k-1} - r_i - \rho(u_i^k - u_i); \end{aligned}$$

https://doi.org/10.1017/S0956792500000759 Published online by Cambridge University Press

likewise, a similar inequality holds when $r_i < 0$ and $r_i^k > 0$. Hence

$$\begin{aligned}(\underline{r}-\underline{r}^k)^T M(\underline{r}-\underline{r}^k) &= \sum_i M_{ii}(r_i - r_i^k)^2,\\ &\leqslant \sum_i M_{ii}(r_i - r_i^{k-1} - \rho(u_i - u_i^k))^2,\\ &= \sum_i M_{ii}(r_i - r_i^{k-1})^2 - 2\rho \sum_i M_{ii}(u_i - u_i^k)(r_i - r_i^{k-1})\\ &\quad + \rho^2 \sum_i M_{ii}(u_i - u_i^k)^2,\\ &= (\underline{r}-\underline{r}^{k-1})^T M(\underline{r}-\underline{r}^{k-1}) - 2\rho(\underline{u}-\underline{u}^k)^T M(\underline{r}-\underline{r}^{k-1})\\ &\quad + \rho^2(\underline{u}-\underline{u}^k)^T M(\underline{u}-\underline{u}^k).\end{aligned}$$

Now noting (4.14d), (4.9a), using (4.4) with $\delta = \Delta t$ and $v = \underline{u}-\underline{u}^k$ and (4.2a) we obtain

$$\begin{aligned}(\underline{u}-\underline{u}^k)^T M(\underline{r}-\underline{r}^{k-1}) &= \gamma(\underline{u}-\underline{u}^k)^T K(\underline{u}-\underline{u}^k) + \frac{1}{\Delta t}(\underline{u}-\underline{u}^k)^T MG(\underline{u}-\underline{u}^k)\\ &\quad - (\underline{u}-\underline{u}^k)^T M(\underline{u}-\underline{u}^k),\\ &\geqslant \left(\gamma - \frac{\Delta t}{4}\right)(\underline{u}-\underline{u}^k)^T K(\underline{u}-\underline{u}^k), \end{aligned} \tag{4.17}$$

so the discrete Poincaré inequality (2.9) implies that

$$(\underline{u}-\underline{u}^k)^T M(\underline{u}-\underline{u}^k) \leqslant \tilde{C}_P^2(\underline{u}-\underline{u}^k)^T K(\underline{u}-\underline{u}^k),$$

and we deduce that

$$(\underline{r}-\underline{r}^k)^T M(\underline{r}-\underline{r}^k) \leqslant (\underline{r}-\underline{r}^{k-1})^T M(\underline{r}-\underline{r}^{k-1}) - (2C\rho - \rho^2)(\underline{u}-\underline{u}^k)^T M(\underline{u}-\underline{u}^k), \tag{4.18}$$

where $C = \tilde{C}_P^{-2}(\gamma - \Delta t/4)$. For $0 < \rho < 2C$, $(\underline{r}-\underline{r}^{k-1})^T M(\underline{r}-\underline{r}^{k-1})$ is a decreasing sequence which is bounded below, hence $r_i^k \to r_i^*$. So if we let $k \to \infty$ in (4.18), it follows that

$$(\underline{u}-\underline{u}^k)^T M(\underline{u}-\underline{u}^k) \to 0 \quad \text{as} \quad k \to \infty,$$

that is $u_i^k \to u_i$. Let $k \to \infty$ in (4.14d); then by continuity

$$\underline{r}^* = (\underline{r}^* - \rho(\underline{u}-\underline{1}))^- + (\underline{r}^* - \rho(\underline{u}+\underline{1}))^+,$$

and letting $k \to \infty$ in (4.14a), we obtain $\lambda^k \to \lambda^*$ such that

$$\gamma K\underline{u} - M\underline{\hat{w}} - M\underline{u} - \lambda^* M\underline{1} = M\underline{r}^*,$$

$$\underline{\hat{w}} = -\frac{1}{\Delta t} G(\underline{u} - \underline{U}^{n-1}).$$

Hence $\{\underline{u}, \underline{r}^*, \lambda^*\}$ solves (4.9a–d). □

Remarks If we take the previous remark into account, then it is easy to show that the algorithm (4.14a–d) is equivalent to: Given $(\underline{U}^{n-1}, \underline{r}^0)$ for $k \geqslant 1$ find $\{\underline{u}^k, \underline{r}^k\}$ such that

$$\gamma M^{-1}KM^{-1}K\underline{u}^k + \frac{1}{\Delta t}(\underline{u}^k - \underline{U}^{n-1}) - M^{-1}K\underline{u}^k = M^{-1}K\underline{r}^{k-1}, \tag{4.19a}$$

$$\underline{r}^k = (\underline{r}^{k-1} - \rho(\underline{u}^k - \underline{1}))^- + (\underline{r}^{k-1} - \rho(\underline{u}^k + \underline{1}))^+. \tag{4.19b}$$

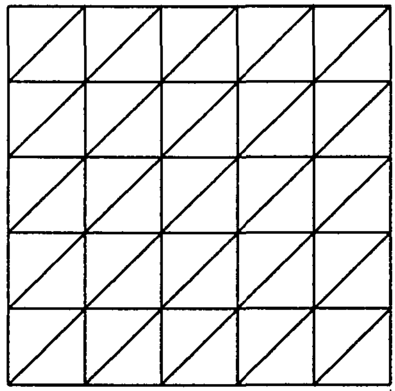

FIGURE 1. Right angled triangulation when Ω is a square.

After the completion of this work, the authors discovered a superior algorithm for the solution of (P_l^h) based on a splitting method for the sum of two nonlinear operators, due to Lions & Mercier (1979). □

5 Numerical simulations

5.1 Introduction

We sketch out a numerical method for solving (4.19a) in two space dimensions on a square in a single step; the method can also be adapted to one and three space dimensions. We have modified a method from French & Nicolaides (1989) which uses the discrete sine transform to solve a pair of coupled linear equations with Dirichlet boundary data.

Let $\Omega = (0,1)\times(0,1)$ and lay down on $N\times N$ uniform square mesh on Ω with vertices

$$(x_i, y_j) = (ih, jh), \quad i,j = 0,1,2,\dots,N, \quad h = 1/N.$$

We now choose the triangulation for Ω, in which each subsquare is bisected by the north east diagonal (see figure 1). For computational convenience the inner product $(.,.)^h$ is defined by

$$(\chi,\phi)^h = \int_\Omega \Pi^h(\chi(x)\,\phi(x))\,\mathrm{d}x,$$

where $\Pi^h v$ is the piecewise bilinear function on Ω which interpolates v at the nodes (x_i, y_j), (x_{i+1}, y_j), (x_i, y_{j+1}) and (x_{i+1}, y_{j+1}) $(i,j = 0,1,\dots,N-1)$; in one space dimension we use the piecewise linear interpolant. The resulting matrix $M^{-1}K$ has eigenvectors and eigenvalues given by $(\psi_{pq}, \lambda_{pq})$, where

$$\psi_{pq}(i,j) = \alpha(p,q)\cos(pi\pi h)\cos(qj\pi h),$$

$$\lambda_{pq} = (4 - 2\cos p\pi h - 2\cos q\pi h)/h^2,$$

$$\alpha(p,q) = \begin{cases} 1 & \text{if } \; p,q \not\equiv 0 \bmod N, \\ 1/2 & \text{if } \; p \equiv 0 \bmod N \quad \text{or} \quad q \equiv 0 \bmod N \quad \text{and not both}, \\ 1/4 & \text{if } \; p,q \equiv 0 \bmod N, \end{cases}$$

and the vectors $\{\psi_{pq}\}$ form a basis for $\mathbb{R}^{N+1}\times\mathbb{R}^{N+1}$. We may thus express $\underline{U}^{n-1}, \underline{r}^{k-1}$ and $\underline{u}^k$ as

$$\left.\begin{aligned} U_{ij}^{n-1} &= \sum_{p,q=0}^{N} c_{pq}^{n-1}\psi_{pq}(i,j),\\ r_{ij}^{k-1} &= \sum_{p,q=0}^{N} \mathrm{d}_{pq}^{k-1}\psi_{pq}(i,j),\\ u_{ij}^{k} &= \sum_{p,q=0}^{N} c_{pq}^{k}\psi_{pq}(i,j). \end{aligned}\right\} \tag{5.1}$$

Substituting (5.1) into (4.19a) yields

$$\sum_{p,q=0}^{N}\left(\frac{1}{\Delta t}(c_{pq}^{k}-c_{pq}^{n-1})+\gamma\lambda_{pq}^{2}c_{pq}^{k}-\lambda_{pq}c_{pq}^{k}\right)\psi_{pq}(i,j)=\sum_{p,q=0}^{N}\lambda_{pq}\mathrm{d}_{pq}^{k-1}\psi_{pq}(i,j), \tag{5.2}$$

so because ψ_{pq} forms a basis and $1+\gamma\Delta ts^2-\Delta ts>0$, for $\Delta t<4\gamma\,(s\in\mathbb{R})$, (5.2) can be rearranged to obtain

$$c_{pq}^{k}=\frac{c_{pq}^{n-1}+\Delta t\lambda_{pq}\mathrm{d}_{pq}^{k-1}}{1+\gamma\Delta t\lambda_{pq}^{2}-\Delta t\lambda_{pq}}. \tag{5.3}$$

So the implemented algorithm A_{IMP} comprises the following steps:

Given pointwise values of $(\underline{U}^{n-1}, \underline{r}^0)$ we calculate

1. The discrete cosine transform of $\underline{U}^{n-1}$.
2. The discrete cosine transform of $\underline{r}^{k-1}$.
3. Compute c_{pq}^k via the formula (5.3).
4. Apply the inverse discrete cosine transform to obtain $\underline{u}^k$.
5. Update $\underline{r}^{k-1}$ using (4.19b).
6. Repeat steps 2–5 until $\|\underline{r}^k-\underline{r}^{k-1}\|_\infty$ is smaller than a prescribed tolerance δ_{IT}.

The operation count to update $\underline{u}^k$ is $O(N^2\log_2 N)$; in one dimension the operation count to update $\underline{u}^k$ is $O(N\log_2 N)$. In the following three subsections we report on some numerical simulations in one and two dimensions which were carried out in double precision on a Sequent Symmetry S81; we did not use the parallel processing facility.

5.2 Numerical tests

We start by giving an account of some experiments to test the error estimate (3.25). We note that in any space dimension

$$(\chi,1)^h=\int_\Omega I^h(\chi(x))\,\mathrm{d}x=\int_\Omega \chi(x)\,\mathrm{d}x=(\chi,1).$$

In particular, this form of numerical integration is used in the algorithm A_{IMP} in one dimension, described in the previous subsection; hence the error estimate is valid for the one dimensional experiments which will follow.

As no exact solution is known, a comparison between solutions U^n on a coarse mesh with a solution u^n on a fine mesh is made at fixed time intervals. We take $\Delta t=O(h^2)<2\gamma$

Table 1 *Testing the error bound*

h	Test 1 error	Test 2 error
2^{-4}	0.077658	0.15530
2^{-5}	0.010215	0.02888
2^{-6}	0.002228	0.00678
2^{-7}	0.0004380	0.00137

so that $\sigma(h, \Delta t) = Ch^2$. The data used in the experiments is $\Omega = (0, 1), \gamma = 0.005$, $h = 1/N$ where $N = 2^j$, $(j = 4, 5, 6, 7)$, $T = 0.075$, $\delta_{IT} = 1.0 \times 10^{-7}$ and $\Delta t = 1.92h^2$ for the coarse meshes; for the fine mesh we take the same data, except $N = 2^8$. The initial data was chosen as follows: we took the solution of the problem $(\boldsymbol{P}^h)$ at $T = 0.40$ and 0.115 where $h = 2^{-8}$ and

$$u_0^h = \begin{cases} P^h(0.01 \cos \pi x - 0.6) & \text{in test 1,} \\ P^h(0.01 \cos 2\pi x - 0.6) & \text{in test 2,} \end{cases} \tag{5.4}$$

respectively. This solution was then interpolated onto the coarser mesh and used as initial data to initiate the error tests; in each of the error tests, on each of the meshes, after one time step there existed a mesh point where $|U_i^n| = 1$. We compute the quantity

$$\|u_{h,\Delta t} - U_{h,\Delta t}\|^2_{L^2(0,T;H^1(\Omega))} \approx c \sum_{k=1}^{M} \Delta t\, |u^k - U^k|_1^2 \approx \sigma(h, \Delta t),$$

where $U_{h,\Delta t}$ is given by (3.23) and $u_{h,\Delta t}$ is similarly defined for the sequence $\{u^k\}$. The values for this quantity are given in table 1 below. Notice that the results in the table agree with the proven error bound which, if sharp, implies a ratio of 4 in successive values.

On the basis of experimentation and lacking any theory, we used an empirical rule for choosing ρ which was found to be satisfactory. An automatic adjustment of the relaxation factor ρ was incorporated into the algorithm $\boldsymbol{A}_{IMP}$ the rule being as follows:

ρ is increased when $\boldsymbol{A}_{IMP}$ has converged only if the sequence $\{\|\underline{r}^k - \underline{r}^{k-1}\|_\infty\}$ was strictly decreasing and a large number of iterations were required;

ρ is reduced only if a few terms in the sequence $\{\|\underline{r}^k - \underline{r}^{k-1}\|_\infty\}$ are larger than the preceding term.

As the simulation is expected to take a large number of time steps to reach a stationary solution, it is easy to see that the number of computations become expensive as h becomes small.

It is worth while reviewing some of the results of Blowey & Elliott (1991, §3) concerning the nature of stationary solutions $\{u, \lambda\}$ satisfying (3.13). We briefly summarize some of the results. In one space dimension on the interval (0, 1) we can classify all of the stationary solutions. The solution is $C^{1,1}$ and is one of the following mutually exclusive types:

1. Piecewise $+1$ or -1 except on disjoint intervals $(x_L, x_L + \pi\sqrt{\gamma})$, $x_L \in (0, 1 - \pi\sqrt{\gamma})$, where $u(x) = \cos(x - x_L)/\sqrt{\gamma}$ and $\lambda = 0$;

https://doi.org/10.1017/S0956792500000759 Published online by Cambridge University Press

2. Without loss of generality, u is $+1$ and any combination of

 (i) $u(x) = (1+\lambda)\cos(x-x_L)/\sqrt{\gamma}-\lambda$ on $(x_L, x_L+2\pi\sqrt{\gamma})$, where $x_L \in (0, 1-2\pi\sqrt{\gamma})$.
 (ii) $u(x) = -(1+\lambda)\cos x/\sqrt{\gamma}-\lambda$ on $(0, \pi\sqrt{\gamma})$.
 (iii) $u(x) = (1+\lambda)\cos(1-\pi\sqrt{\gamma}-x)/\sqrt{\gamma}-\lambda$ on $(1-\pi\sqrt{\gamma}, 1)$.

 $\lambda < 0$ is determined from the mass constraint $(m > 0)$;
3. $u(x) \equiv m$ or if $\gamma = 1/(j^2\pi^2)$, $u(x) = m + a\cos x/\sqrt{\gamma}$ where $|a| < \min\{1-m, 1+m\}$.

It was also shown that the functional

$$\mathscr{E}(u) = \frac{1}{2}\int_0^1 (u_x^2 + 1 - u^2)\,\mathrm{d}x,$$

is minimized when the measure of the interval on which $|u(x)| < 1$ is minimized. In two space dimensions, radially symmetric stationary solutions were constructed using Bessel functions, where $\lambda \neq 0$.

Our choice of a spatial mesh size is guided by our expectation of the width of the interfacial region. We expect the width of a stationary solution with a single interface in one dimension to be $\pi\sqrt{\gamma}$. This suggests that for the simulations to approximate the interface of a stationary solution well, we need $h < C\sqrt{\gamma}$. In a recent paper by Bellettini *et al.* (1990), for $h = o(\gamma^{\frac{1}{2}})$ they prove the Γ-convergence of any discrete sequence of minimizers of

$$\mathscr{F}_{\gamma,h}(v) = \int_\Omega \left[\gamma^{\frac{1}{2}}|\nabla v|^2 + \frac{1}{\gamma^{\frac{1}{2}}} I^h(1-v^2) - \frac{\pi}{2} I^h(\kappa_\gamma v)\right]\mathrm{d}x,$$

over the set $\{v \in K^h : v = I^h g_\gamma \text{ on } \partial\Omega\}$, to a piecewise constant function whose interfaces have a prescribed mean curvature and contact angle with the boundary $\partial\Omega$ (κ_γ and g_γ are, respectively, functions related to a mean curvature and contact angle with the boundary $\partial\Omega$).

5.3 One-dimensional simulations

If the initial data is strictly less than one in magnitude, then the variational inequality reduces to the equation

$$u_t = -\gamma u_{xxxx} - u_{xx} \quad 0 < x < 1,$$

with boundary conditions

$$u_x(0,t) = u_x(1,t) = u_{xxx}(0,t) = u_{xxx}(1,t) = 0,$$

until the solution reaches one in magnitude. Thus for a short time

$$u(x,t) = m + \sum_{p=1}^{\infty} c_p \exp^{-\sigma_p t}\cos p\pi x,$$

where $c_p = 2\int_0^1 u_0 \cos p\pi x\,\mathrm{d}x$ and

$$\sigma_p = p^2\pi^2(\gamma p^2\pi^2 - 1). \tag{5.5}$$

We see from the dispersion relation (5.5) that the Fourier components with $p^2 > 1/(\gamma\pi^2)$ are damped out and the maximum growth rate is associated with $p_C^2 \approx 1/(2\gamma\pi^2)$. Similarly,

https://doi.org/10.1017/S0956792500000759 Published online by Cambridge University Press

Table 2 *Initial data taken in one dimensional simulations*

Simulation	*Initial data*	*Figure*	$m =$
1	$0.05\mathrm{rnd}\,(x)$	2	0
2	$0.05\mathrm{rnd}\,(x)-0.6$	3	-0.6
3	$P^h\,(0.01\,\Sigma_{k=1}^{16}\,(-1)^k \cos 2\pi k x - 0.6)$	4	-0.6
4	$P^h\,(-0.005 \cos 2\pi x - 0.95)$	5	-0.95
5	$P^h\,(0.005 \cos 2\pi x - 0.95)$	6	-0.95
6	$P^h\,(-0.005 \cos \pi x - 0.95)$	7	-0.95

denoting by u_i^n the value of the finite element approximation at the time level $t^n = n\Delta t$ and mesh point $x = ih\,(0 \leqslant i \leqslant N)$, we have the equation

$$\frac{u_i^n - u_i^{n-1}}{\Delta t} = \frac{\gamma}{h^4}\{-u_{i-2}^n + 4u_{i-1}^n - 6u_i^n + 4u_{i+1}^n - u_{i+2}^n\} + \frac{-u_{i-1}^n + 2u_i^n - u_{i+1}^n}{h^2},$$

where $u_{-1} = u_1, u_{-2} = u_2, u_{N+1} = u_{N-1}, u_{N+2} = u_{N-2}$, until the value at a mesh point breaks the magnitude one. Thus for small time

$$u_i^n = m + \sum_{p=1}^{N-1} \xi_p \left[\frac{\gamma \Delta t}{h^4} 16 \sin^4 \frac{ph\pi}{2} - \frac{\Delta t}{h^2} 4 \sin^2 \frac{ph\pi}{2} + 1\right]^{-n} \cos p\pi i h$$
$$+ \xi_N \left[16 \frac{\gamma \Delta t}{h^4} - 4 \frac{\Delta t}{h^2} + 1\right]^{-n} \frac{(-1)^i}{2}, \quad (5.6)$$

where $$\xi_p = \tfrac{1}{2}u_0^0 + \sum_{k=1}^{N-1} u_k^0 \cos k\pi p h + \tfrac{1}{2}u_N^0 (-1)^N \quad (p = 1, \ldots, N).$$

We see from the discrete dispersion relation (5.6) that the discrete Fourier components with $\sin \frac{p\pi h}{2} > (h/4\gamma)^{\frac{1}{2}}$ are damped out, and for h small enough the maximum growth rate is associated with $1/(2\gamma\pi^2)$; if applied directly when $\gamma = 0.005$, the maximum growth rate is for the discrete mode $\cos 3\pi i h$; for $p \geqslant 5$, the amplitude of the mode $\cos p\pi i h$ decays to zero.

In all of these simulations $\Omega = (0,1), \gamma = 0.005$, $N = 100$, $\Delta t = \gamma$, $T = 1.0$, $\delta_{IT} = 1.0 \times 10^{-7}$. In each experiment the initial data u_0^h was taken to be a perturbation of a uniform state (see table 2), $\mathrm{rnd}(x)$ generates a random number distributed between -1 and 1 at each node of the triangulation; the solutions were plotted out at time intervals of 0.2, that is every 40 time steps.

In the first two simulations, we expect the high frequency oscillations to be damped out before linear growth ensues, as discussed in the beginning of this subsection. In simulations 3–5, the initial data is even, and in simulation 6 the initial data is odd about the line $x = 0.5$. The initial data in simulation 3 has many high frequency modes which will decay whilst low modes, namely 2 and 4, grow and dominate all other modes; the large oscillations in

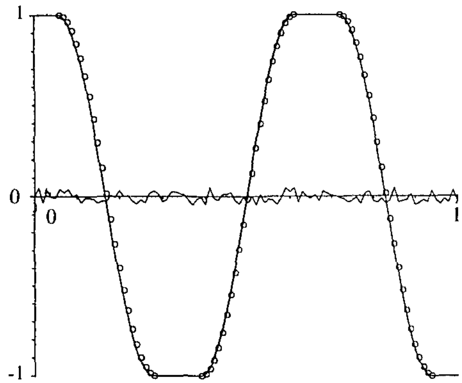

FIGURE 2. One dimensional simulation 1.

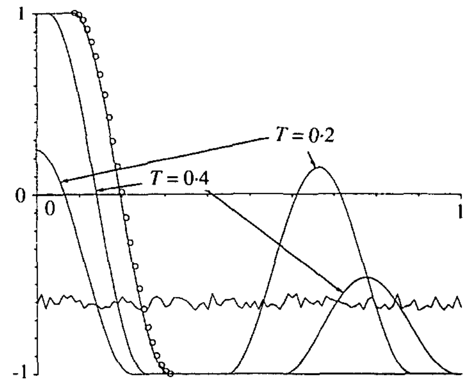

FIGURE 3. One dimensional simulation 2.

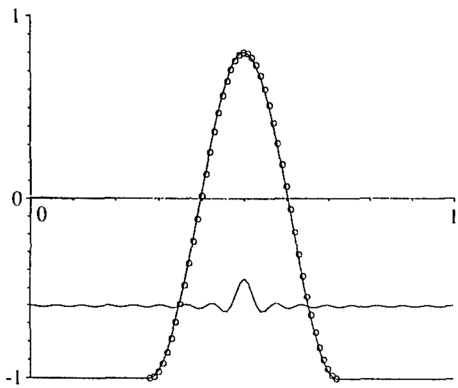

FIGURE 4. One dimensional simulation 3.

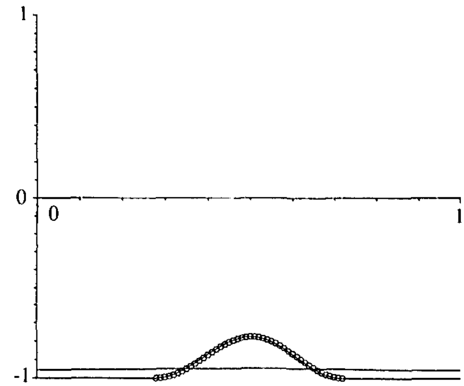

FIGURE 5. One dimensional simulation 4.

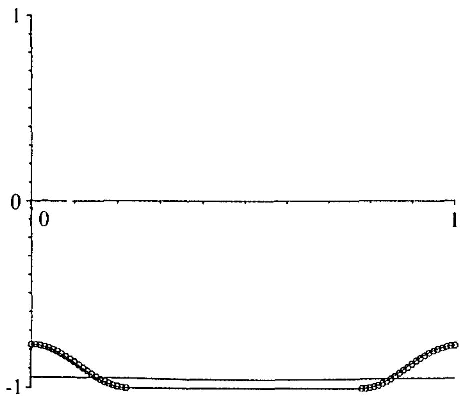

FIGURE 6. One dimensional simulation 5.

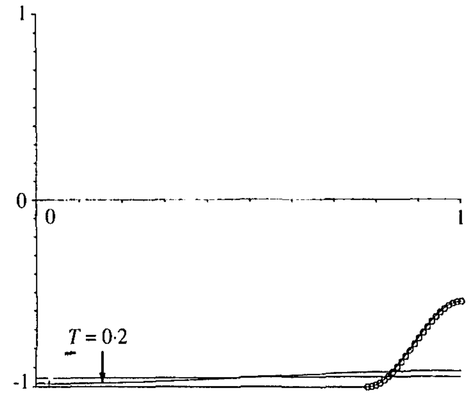

FIGURE 7. One dimensional simulation 6.

https://doi.org/10.1017/S0956792500000759 Published online by Cambridge University Press

the initial data play no part as they quickly decay. In simulations 4–6 there is only growth. For each simulation we have overlayed the discrete solution with an exact continuous stationary solution, as given in the previous subsection; ○ denotes the value at the mesh point for the continuous stationary solution.

We judged that we had a discrete stationary solution if we met all of the following criteria.

- Only one iteration was required to solve the algorithm A_{IMP} at each time step.
- The computed discrete chemical potential W^n was constant up to the prescribed tolerance δ_{IT}.
- The computed stationary solution persisted for a large number of time steps.

The final discrete stationary solution obtained in simulation 2 is an approximation of the minimizer of $\mathscr{E}$. In simulation 1, the computed stationary solution has three interfaces; the continuous analogue of this stationary solution is not the minimizer of $\mathscr{E}$. If the early stages are analysed more closely, then the cosine mode which dominates all others is the maximum growth mode, $\cos 3\pi x$. When $|U_i^n| = 1$ for some mesh point, the point expands to a domain which grows whilst the interfaces sharpen, which explains the stationary composition that is obtained.

The final stationary solutions obtained in simulations 3–5 are not approximations to the known global minimizer of $\mathscr{E}$; however, since the initial data taken is even, the solution is even for all time. The discrete stationary solutions obtained in these simulation appear to be of type 2(i), (ii), (iii), described in the previous subsection. The computed stationary solution observed in simulation 6 is close to the known minimizer of $\mathscr{E}$ for this value of m. Simulations 3–5 can be considered from an alternative point of view. Since the initial data is symmetric about the line $x = 0.5$, solving the problem on the half interval $(0, 0.5)$ is equivalent to solving the full problem on $(0, 1)$. In the continuous problem, the minimizer on the interval $(0, 0.5)$ is given by the monotone solution where the length of the interval on which $|u| < 1$ is $\pi\sqrt{\gamma}$. This is the case in simulations 3–5, where on the interval $(0, 0.5)$ the discrete stationary solution is monotone. Consequently, the computed stationary compositions obtained are close to the known even minimizers of $\mathscr{E}$.

A particular point of interest which arises from these simulations is that, however large m is taken to be, growth followed by phase separation always ensues. Then over a large period of time, the domains where $|U^n| = 1$ grow and shrink as the interface migrates and disappears. These two phenomena are also observed in the two dimensional simulations that follow.

A question which was not addressed in Blowey & Elliott (1991) was the nature of stability of the stationary solutions solving (3.13). A similar problem of finding $u \in K$ such that $\forall \eta \in K$

$$\gamma(\nabla u, \nabla \eta - \nabla u) - (u, \eta - u) \geqslant 0,$$

has been studied by Chen & Elliott (1991). They have obtained results about the stability of all of the one dimensional stationary solutions with respect to the evolutionary problem

$$u(t) \in K : (u_t, \eta - u) + \gamma(\nabla u, \nabla \eta - \nabla u) \geqslant (u, \eta - u) \ \forall \eta \in K.$$

In particular, steady state solutions of the form depicted in figure 2 were shown to be stable.

We have been able to prove that for most type 2 stationary solutions there are directions which decrease the energy of the Lyapunov functional $\mathscr{E}$. For instance, consider the continuous stationary solution

$$u(x) = \begin{cases} -\lambda-(1+\lambda)\cos\dfrac{x}{\sqrt{\gamma}} & \text{on} \quad (0, \pi\sqrt{\gamma}], \\ -\lambda+(1+\lambda)\cos\dfrac{x-x_L}{\sqrt{\gamma}} & \text{on} \quad (x_L, x_L+2\pi\sqrt{\gamma}], \\ 1 & \text{otherwise,} \end{cases} \tag{5.7}$$

where $\pi\sqrt{\gamma} \leqslant x_L \leqslant 1-2\pi\sqrt{\gamma}$, $-1 < u(x), \lambda < 0$ and $(u, 1) = m > 0$. Define

$$\eta(x) = \begin{cases} -2\left(\cos\dfrac{x}{\sqrt{\gamma}}+1\right) & \text{on} \quad (0, \pi\sqrt{\gamma}], \\ -\left(\cos\dfrac{x-x_L}{\sqrt{\gamma}}-1\right) & \text{on} \quad (x_L, x_L+2\pi\sqrt{\gamma}], \\ 0 & \text{otherwise,} \end{cases} \tag{5.8}$$

and take $|\mu| \leqslant \mu^* = \frac{1}{2}\min\{-\lambda, \lambda+1\}$ so that for all $|\mu| \leqslant \mu^*$

$$|u+\mu\eta| \leqslant 1, \quad (u+\mu\eta, 1) = (u, 1). \tag{5.9}$$

Then simple calculations reveal that for all $0 < |\mu| \leqslant \mu^*$

$$\mathscr{E}(u+\mu\eta) = \mathscr{E}(u)-3\mu^2\pi\sqrt{\gamma} < \mathscr{E}(u),$$

suggesting that this stationary solution is unstable. This instability was borne out in some numerical experiments, which we briefly describe. After some time, a discrete analogue of (5.7) was obtained. This persisted for a large number of time steps; each time step required a large number of iterations for A_{IMP} to converge. Eventually, the composition evolved into either a monotone discrete stationary solution or a discrete single humped stationary solution.

5.4 Two-dimensional simulations

We first discuss two simulations with data differing only in their initial conditions. In each case $\gamma = 0.0032$, $\Omega = (0,1)\times(0,1)$, $N = 64$, $\Delta t = \gamma$, $\delta_{IT} = 5.0\times 10^{-5}$, and the initial condition was taken to be a random perturbation, with values distributed between -0.05 and $+0.05$, of the uniform states $u_0 = m$, where $m = 0$ and $m = -0.704$, respectively. The results are graphically displayed in figures 8(*a–f*) and 9(*a–d*); at (x_i, y_j) we denote a $\bigcirc$ to mean that $U_{ij}^n = 1$, a $*$ where $U_{ij}^n = -1$, and a blank space where $|U_{ij}^n| < 1$. Each simulation was continued until we observed a computed stationary solution which fitted the following criteria. Only one iteration was required to solve A_{IMP} at each time step; the discrete chemical potential W^n was constant up to the prescribed tolerance δ_{IT}; the discrete stationary solution persisted for a large number of time steps. The choice of the tolerance δ_{IT} in the iteration was guided by the need to have reasonable computing time in order to

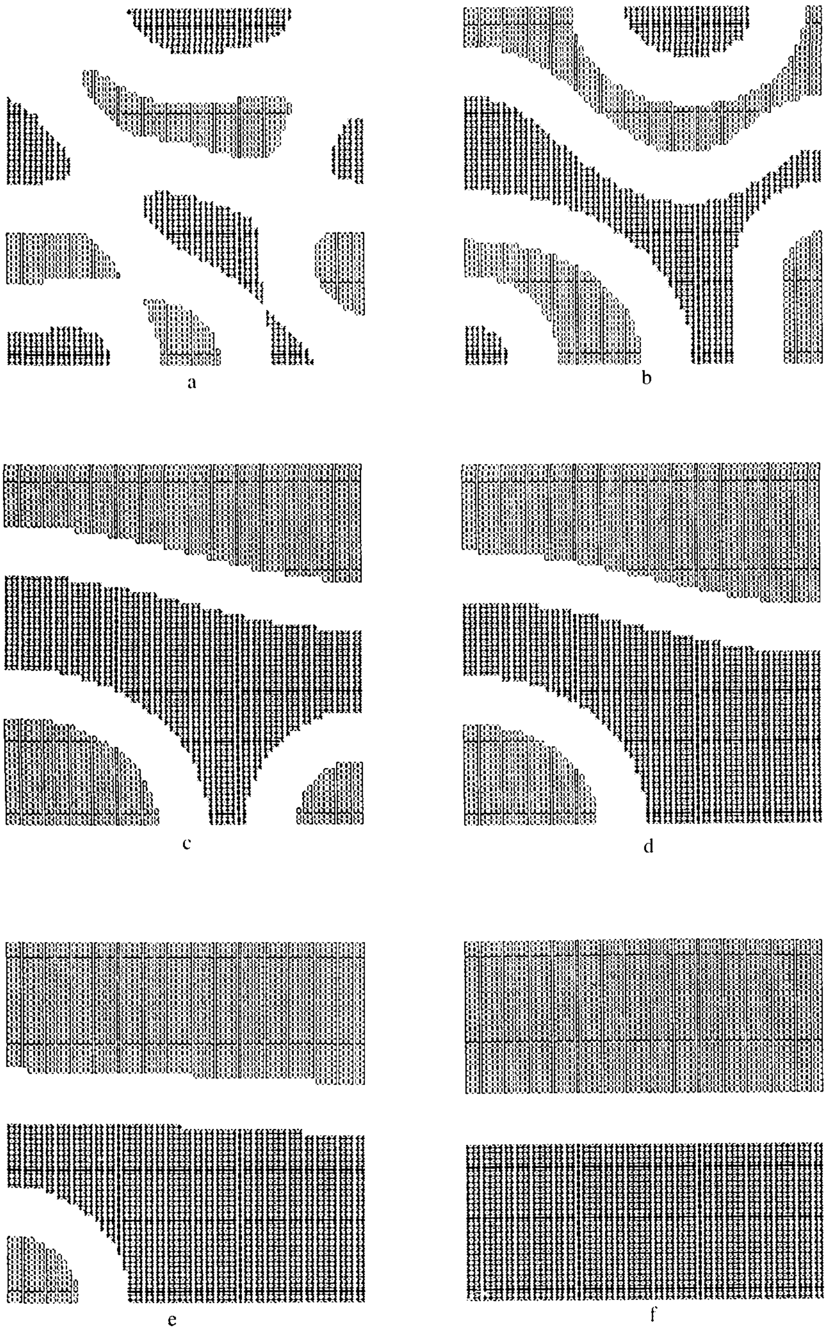

FIGURE 8. (*a*) $T = 0.08$, (*b*) $T = 0.16$; (*c*) $T = 0.4$; (*d*) $T = 0.8$; (*e*) $T = 1.28$; (*f*) $T = 3.2$.

reach the steady state. The use of a smaller δ_{IT} does not change the graphical representation of the simulation by a significant amount.

Both these simulations have the expected feature that the measure of the interfacial region (blank region where $|U_{ij}^n| < 1$) decreases in time. This is a consequence of the fact

https://doi.org/10.1017/S0956792500000759 Published online by Cambridge University Press

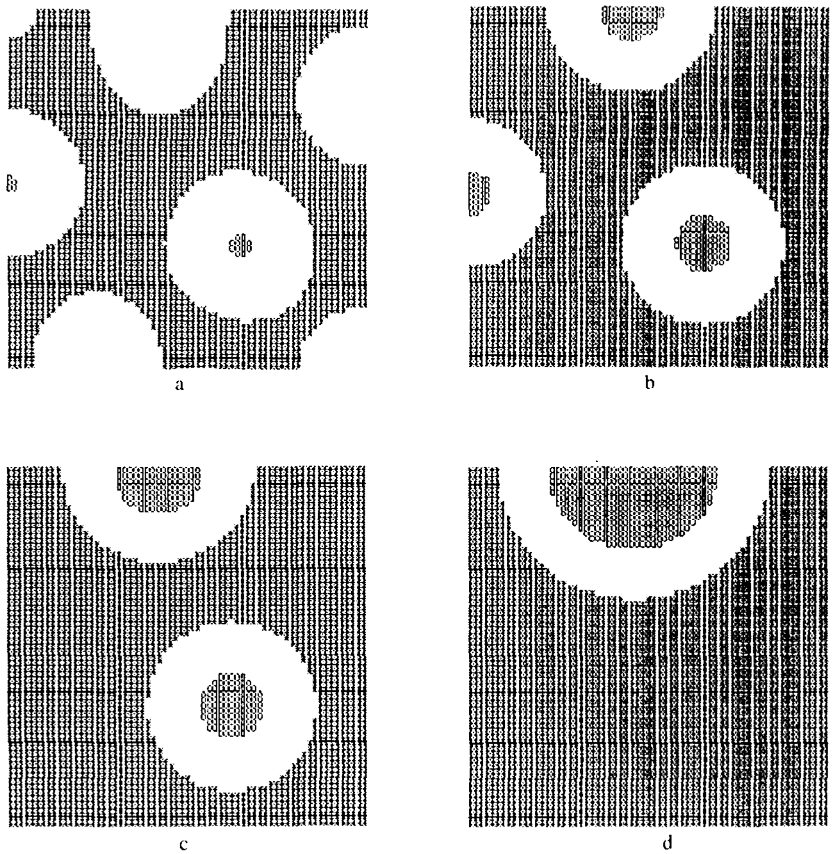

FIGURE 9. (a) $T = 0.08$; (b) $T = 0.16$; (c) $T = 0.4$; (d) $T = 1.6$.

that the energy $\mathscr{E}^h(.)$ is a Lyapunov functional, and that the term $\gamma|u|_1^2/2$ was introduced in this functional in order to account for interfacial energy.

In the early stages of the first simulation (figures 8(a–f)), a lamellar structure forms in which the domains where $|U^n| = 1$ are thin. In time these domains grow and shrink as the interfaces migrate and disappear until at $T = 3.2$ we observe a computed stationary composition which consists of strips, and is the two dimensional analogue of the one-dimensional monotone continuous stationary solution, described in the previous subsection. In fact, the stationary discrete chemical potential takes the constant value 1.931×10^{-4}, which is in good agreement with the exact steady chemical potential of constant value 0.

In the second simulation the morphology is completely different (figure 9(a–d)); now in the early stages circular domains are observed. In time these domains grow and shrink until at $T = 1.6$ we obtain a discrete stationary composition with a constant discrete chemical potential of 0.1458, which is in good agreement with the value 0.1460 given by Blowey & Elliott (1991, table 3.3) for the exact stationary chemical potential when $k = 1$; in fact, m was chosen from the same reference in table 3.4 when $k = 1$ and $n = 2$.

It is interesting to consider smaller values of γ. However, in order to have a finite computing time we are restricted in our choice of h and Δt. As a test we varied γ and the

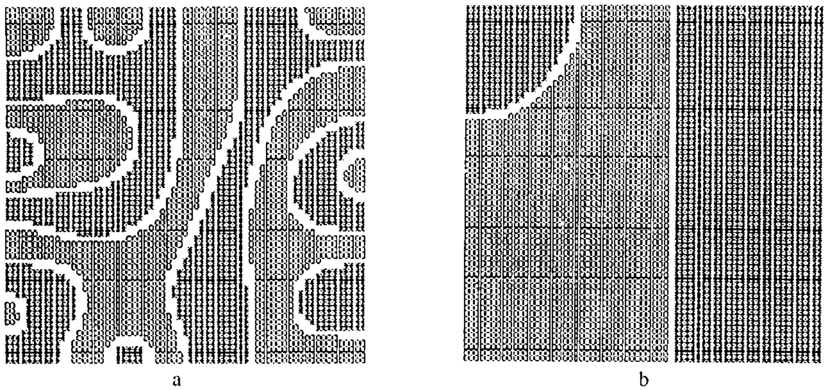

FIGURE 10. (*a*) $T = 0.08$; (*b*) $T = 8.0$.

remaining parameters were taken to be the same as above, except that at each node (x_i, y_j) $(0 \leqslant i,j \leqslant N)$ of the mesh, the initial data was given by

$$u_0^h(x_i, y_j) = \begin{cases} 1 & \text{if} \quad r \leqslant r_1, \\ 1-2(r-r_1)/(r_2-r_1) & \text{if} \quad r_1 < r < r_2, \\ -1 & \text{if} \quad r_2 \leqslant r, \end{cases} \tag{5.10}$$

where $r = ((x_i-0.5)^2+y_j^2)^{\frac{1}{2}}, r_1^2 = 0.32$ and $r_2^2 = 0.36$, so that $(u_0^h, 1) \approx 0$. We observe that when $m = 0$:

(i) as $\gamma > 0$ there always exists a continuous stationary solution consisting of strips;
(ii) the length of the interface is minimized by a strip solution.

We performed the test when $\gamma = 0.0002$, 0.0008 and 0.0032. The only discrete stationary solution which was a perfect strip was for the value $\gamma = 0.0032$; the other two discrete stationary solutions had an interface which had a small and differing amount of curvature, and appeared not to be approximations to any continuous stationary solution.

A further simulation was performed where the initial data taken was precisely the same as in the first simulation of this subsection (so that u_0^h was a random perturbation ± 0.05 of the uniform value $m = 0$), but γ was taken to be 0.0002. The behaviour in the early stages was similar to the previous simulation for larger γ (figure 10(*a*)), where a lamellar structure forms. However, the corresponding computed stationary solution was such that Ω_0^h, the domain where $|U^n| < 1$, consisted of a strip Ω_S^h, together with a quarter annulus Ω_A^h (figure 10(*b*)). We conjecture that this is a spurious solution in the sense that it is an artefact of the numerical scheme, and is not an approximation of an exact solution of the partial differential equation. It occurs because h is too large for the accurate resolution of the moving interfacial region. In one dimension the exact stationary solution with one interface is a cosine and $\lambda = 0$, whereas in two dimensions the exact radially-symmetric stationary solution with one interface is a Bessel function and $|\lambda| > 0$. Since λ is constant throughout Ω, we cannot have such exact solutions coexisting. This is compelling evidence that the computed stationary solution in figure 10(*b*) arising from the discretization with $\gamma = 0.0002$ is indeed spurious. Across the width of the interface, we note that there are only one or two

https://doi.org/10.1017/S0956792500000759 Published online by Cambridge University Press

mesh spacings. It is clear then in order to resolve the interfaces, we need sufficient mesh points in the interfacial region.

5.5 Concluding remarks

1. In the early stages of all experiments when $|U^n| < 1$, one iteration was required for the algorithm A_{IMP} to converge and solve the linear problem.
2. With our numerical experiments, no problems were experienced with respect to the convergence of the algorithm. In the earlier stages of the two dimensional simulations when $|U_i^n| = 1$ for some i, there were many variations in U^n, so the average number of iterations, > 100, was larger than in the later stages, 50–100, and in the very late stages very few iterations were required.
3. It was noted previously that if the initial data is non-constant, where $|m| < 1$, then phase separation always ensues. This contrasts with the Cahn–Hilliard equation, where for m outside the spinodal interval, initial data with nucleation centres are taken to obtain phase separation (Copetti & Elliott 1990). The reason for this difference is that in this model the whole interval $(-1, 1)$ is the spinodal interval.

It may be of interest to the general reader to view some real experimental data obtained by material scientists, thus we refer to Cerezo *et al.* (1989) for the spinodal decomposition of an alloy obtained using a position-sensitive atom probe.

References

Berger, A. & Falk, R. 1977 An error estimate for the truncation method for the solution of parabolic variational inequalities. *Math. Comp.* **31**, 619–628.

Bellettini, G., Paolini, M. & Verdi, C. 1990 Γ-convergence of discrete approximations to interfaces with prescribed mean curvature. *Atti. Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* **1**, 317–328.

Blowey, J. F. & Elliott, C. M. 1991 The Cahn–Hilliard gradient theory for phase separation with non-smooth free energy. Part I: Mathematical analysis. *Euro. J. Appl. Math.* **2**, 233–280.

Cerezo, A., Godfrey, T. J., Grovenor, C. R. M., Hetherington, M. G., Hoyle, R. M., Jakubovics, J. P., Liddle, J. A., Smith, G. D. W. & Worral, G. M. 1989 Material analysis with a position-sensitive atom probe. *J. Microscopy.* **154** (3), 215–225.

Chakrabarti, R. 1988 *Numerical Solution of some Free Boundary Problems.* PhD Thesis, Imperial College, UK.

Chen Xinfu & Elliott, C. M. 1991 *Asymptotics for a Parabolic Double Obstacle Problem.* Preprint.

Ciarlet, P. G. 1978 *The Finite Element Method for Elliptic Problems.* North-Holland.

Colli, P. L. & Verdi, C. 1985 Error estimates for an approximation of a problem of percolation in gently sloping beaches. *Calcolo* **22**, 383–390.

Copetti, M. I. M. & Elliott, C. M. 1990 Kinetics of phase decomposition processes: numerical solutions to the Cahn–Hilliard equation. *Mat. Sci. & Technol.* **6**, 273–283.

Elliott, C. M. 1989 The Cahn–Hilliard model for the kinetics of phase separation. In J. F. Rodrigues (ed.), *Mathematical Models for Phase Change Problems*, Birkhäuser.

Elliott, C. M. & French, D. A. 1987 Numerical studies of the Cahn–Hilliard equation for phase separation. *IMA J. Appl. Math.* **35**, 97–128.

Elliott, C. M., French, D. A. & Milner, F. 1989 A second order splitting method for the Cahn–Hilliard equation. *Numer. Math.* **54**, 575–590.

ELLIOTT, C. M. & LARSSON, S. 1991 On a finite element method for the Cahn–Hilliard equation (to appear in *Math. Comp.*).

FETTER, A. 1987 L^∞-error estimates for an approximation of a parabolic variational inequality. *Numer. Math.* **50**, 557–565.

FRENCH, D. A. 1990 Computations on a Cahn–Hilliard model of solidification. *Appl. Math. Comp.* **40** (1), 55–76.

FRENCH, D. A. & NICOLAIDES, R. A. 1989 Numerical results on a Cahn–Hilliard model of phase transition. Preprint.

GLOWINSKI, R., LIONS, J. L. & TRÉMOLIÈRES, R. 1981 *Numerical Analysis of Variational Inequalities.* North-Holland.

JOHNSON, C. 1976 A convergence estimate for an approximation of a parabolic variational inequality. *SIAM J. Numer. Anal.* **13**, 559–606.

LIONS, P. L. & MERCIER, B. 1979 Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.* **16** (6), 964–979.

SCARPINI, F. & VIVALDI, M. A. 1978 Evaluation de l'erreur d'approximation pour une inéquation parabolique rélative aux convex dépendent du temps. *Appl. Math. Opt.* **4**, 121–138.

VUIK, C. 1990 An L^2-error estimate for an approximation of a parabolic variational inequality. *Numer. Math.* **57**, 453–471.