

SIMONA SANFELICI (*)

**Numerical and analytic study
of a parabolic-ordinary system modelling cardiac activation
under equal anisotropy conditions (**)**

1 - Introduction

During a heart beat an *excitation wavefront* propagates throughout the myocardium; the wavefront is represented by a thin layer of cells which undergoes a *depolarization process*, i.e. a sudden temporal change of the transmembrane potential v , defined as the jump of the potential across the cellular membrane. This electrochemical perturbation (*action potential*) is brought about by the flow of ionic currents across the membrane. The depolarization process of the cardiac tissue may be described by means of the macroscopic *bidomain model* ([6]), with Hodgkin-Huxley type gating equations of the cellular membrane current.

In the bidomain model, the cardiac tissue is defined by two interpenetrating continuous media (i.e. the *intercellular i* and the *extracellular or interstitial e medium* respectively), connected everywhere by a continuous cardiac cell membrane. It's well known that the bidomain, under *equal anisotropy* conditions, reduces to a monodomain and that in this case the extracellular and intracellular potentials can be obtained as scaled versions of the transmembrane potential.

The introduction of non diagonal conductivity tensors for the *i* and *e* media allows to model the effect of the intramural fiber rotation (*rotational anisotropy*) on the current flow.

(*) Dip. di Matem., Univ. Parma, Via M. D'Azeglio 85, 43100 Parma, Italia.

(**) Received August 29, 1996. AMS classification 35 K 57.

In Section 2, we present the mathematical model in details and give existence and uniqueness results for its solution and other regularity results. In Sections 3 and 5, we propose and analyse a numerical method based on the finite element Galerkin space discretization. The associated nonlinear system of ordinary differential equations is solved by means of the Crank-Nicholson time-stepping, combined with a quasilinearization technique. Convergence of both space and time discretization is proved.

2 - Mathematical model

Let Ω be a bounded open domain in \mathbf{R}^3 with boundary $\partial\Omega$ of class $\mathcal{C}^{1+\alpha}$, $\alpha \in (0, 1)$, representing an insulated block of myocardium (the bidomain $\mathbf{i} + \mathbf{e}$).

Let σ_i^i and σ_i^e be the conductivity coefficients parallel and transverse, respectively, to cardiac fiber direction for the \mathbf{i} -medium, σ_e^i and σ_e^e be the conductivity coefficients for the \mathbf{e} -medium and let $\mathbf{a} = \mathbf{a}(\mathbf{x})$ be the unit vector tangent to the fiber at point $\mathbf{x} \in \Omega$. The conductivity tensors for the intracellular and extracellular media are given by:

$$\mathbf{M}_i = (\sigma_i^i - \sigma_i^e) \mathbf{a}\mathbf{a}^T + \sigma_i^e \quad \mathbf{M}_e = (\sigma_e^e - \sigma_e^i) \mathbf{a}\mathbf{a}^T + \sigma_e^i .$$

The tensor $\mathbf{M} = \mathbf{M}_i + \mathbf{M}_e$ characterizes the bulk conductivity of the composite medium $\mathbf{i} + \mathbf{e}$. In the following we assume that σ_i^i , σ_i^e , σ_e^i and σ_e^e are positive constants, with $\sigma_i^i \geq \sigma_i^e$ and $\sigma_e^e \geq \sigma_e^i$, i.e. conductivity is greater along than across fibers.

Let the *anisotropic ratios* be defined as follows:

$$\xi_i = \frac{\sigma_i^e}{\sigma_i^i} \quad \xi_e = \frac{\sigma_e^e}{\sigma_e^i} .$$

In this paper we suppose that *equal anisotropy* holds, i.e.:

$$(1) \quad \xi_i = \xi_e = r \text{ constant} .$$

Let $v(\mathbf{x}, t)$ be the transmembrane potential assumed to be a regular function in the region Ω . Under assumption (1), the activation process is described by the *reaction-diffusion* (RD) system (see [3], [4] and [6] for more details):

$$(2) \quad \begin{aligned} \chi C_m v_t - \frac{r}{1+r} \operatorname{div} \mathbf{M}_i \nabla v + \chi I_{\text{ion}}(v, m, h) &= 0 && \text{in } \Omega \times]0, T] \\ m_t = f_2(v, m) = -(\alpha_m(v) + \beta_m(v))m + \alpha_m(v) &&& \text{in } \Omega \times]0, T] \\ h_t = f_3(v, h) = -(\alpha_h(v) + \beta_h(v))h + \alpha_h(v) &&& \text{in } \Omega \times]0, T] \\ \mathbf{n}^T \mathbf{M}_i \nabla v &= 0 && \text{on } \partial\Omega \times]0, T] \\ v(\mathbf{x}, 0) = v_0 \quad m(\mathbf{x}, 0) = m_\infty(v_0) \quad h(\mathbf{x}, 0) = h_\infty(v_0) &&& \text{in } \Omega \end{aligned}$$

where \mathbf{n} is the outward unit normal vector to $\partial\Omega$, χ is the membrane surface area per unit volume of the tissue, C_m is the membrane capacitance per unit area and I_{ion} is the ionic current per unit area carried by the flow of ions across the membrane. Functions $m(\mathbf{x}, t)$ and $h(\mathbf{x}, t)$ are dimensionless gating variables responsible for the activation and inactivation of the sodium channel respectively and

$$m_\infty(v) = \frac{\alpha_m(v)}{\alpha_m(v) + \beta_m(v)} \quad h_\infty(v) = \frac{\alpha_h(v)}{\alpha_h(v) + \beta_h(v)}$$

are their steady-state values; $\alpha_m, \alpha_h, \beta_m$ and β_h are positive empirically determined functions of potential v , as described in [5]. These functions, in their original form, present some point discontinuities which can be eliminated easily by raccordng the jumps in such a way that the functions belong to $C^2(\mathbf{R})$, without affecting the numerical solution to system (2). Finally, the initial datum $v_0(\mathbf{x})$ is a Hölder continuous function such that $v_r \leq v_0 \leq v_{Na}$ in Ω .

The fast sodium current is the main cause of depolarization of cardiac cells during the excitation phase of the action potential. It is modeled by gating equations of the Hodgkin-Huxley type ([7]) and can be represented in the form

$$I_{\text{ion}}(v, m, h) = g_{Na} m^3 h (v - v_{Na}) + g_r (v - v_r)$$

where g_{Na} and g_r are the maximum conductances for sodium ions and repolarization currents, respectively, and v_{Na} and v_r are the equilibrium potentials for sodium and repolarizing currents.

For mathematical purposes, we can rewrite system (2) in the form

$$\begin{aligned} (3) \quad & v_t = \mathcal{L}v - \frac{g_r}{C_m} v + f_1(v, m, h) + f && \text{in } \Omega \times]0, T] \\ & m_t = f_2(v, m) \quad h_t = f_3(v, h) && \text{in } \Omega \times]0, T] \\ & \mathbf{n}^T \mathbf{M}_i \nabla v = 0 && \text{on } \partial\Omega \times]0, T] \\ & v(\mathbf{x}, 0) = v_0 \quad m(\mathbf{x}, 0) = m_\infty(v_0) \quad h(\mathbf{x}, 0) = h_\infty(v_0) && \text{in } \Omega \end{aligned}$$

where $f_1(v, m, h) = -\frac{1}{C_m} g_{Na} m^3 h (v - v_{Na})$ and $f = \frac{1}{C_m} g_r v_r$, $\mathcal{L}v = \sigma \operatorname{div} \mathbf{M}_i \nabla v$ and $\sigma = \frac{r}{\chi C_m (1+r)}$.

We have shown in [12] that the solution of system (3) exists and is unique. In fact, \mathcal{L} is a uniformly elliptic operator in $\overline{\Omega}$ since, for any ξ of \mathbf{R}^3

$$\begin{aligned} \xi^T \mathbf{M}_i \xi &= (\sigma_l^i - \sigma_j^i) \xi^T \mathbf{a} \mathbf{a}^T \xi + \sigma_l^i \xi^T \xi = (\sigma_l^i - \sigma_j^i) (\mathbf{a}^T \xi)^T (\mathbf{a}^T \xi) + \sigma_l^i \xi^T \xi \\ &= (\sigma_l^i - \sigma_j^i) |\mathbf{a}^T \xi|^2 + \sigma_l^i |\xi|^2 \geq \alpha_0 |\xi|^2 \end{aligned}$$

for every $\mathbf{x} \in \overline{\Omega}$, where $|\cdot|$ indicates the Euclidean norm and $\alpha_0 = \sigma_l^i$.

We assume that \mathcal{L} has Hölder continuous coefficients in $\overline{\Omega}$. Moreover functions f_i , $i = 1, 2, 3$, satisfy the *local Lipschitz condition*:

for each $R > 0$ there are constants $K_i = K_i(R)$ such that

$$|f_i(\mathbf{u}) - f_i(\mathbf{w})| \leq K_i |\mathbf{u} - \mathbf{w}|_1 \quad \text{when } |\mathbf{u}|_1 \leq R, \quad |\mathbf{w}|_1 \leq R$$

where $|\mathbf{u}|_1 = \sum_{i=1}^3 |u_i|$.

It is useful for the sequel to state the above mentioned result in this form

Theorem 1. *System (3) has a unique classical global solution $v \in \mathcal{C}(\overline{\Omega} \times [0, T]) \cap \mathcal{C}^{2,1}(\Omega \times]0, T])$; $m, h \in \mathcal{C}^{0,1}(\overline{\Omega} \times [0, T])$. Moreover this solution satisfies the limitations*

$$(4) \quad 0 \leq m \leq 1 \quad 0 \leq h \leq 1 \quad v_r \leq v \leq v_{Na} \quad \text{in } \overline{\Omega} \times [0, T].$$

Finally, it must be noted that the components m and h of the solution to (3) have continuous partial derivatives of the second order with respect to \mathbf{x} over $\Omega \times]0, T]$. This follows by composition argument, since the gating equations can be thought as ordinary differential equations in m and h , where the right member f_i ($i = 2, 3$) contains a parameter vector \mathbf{x} and, by composition, has continuous partial second order derivatives with respect to the components of (y, \mathbf{x}) , $y = m, h$. This yield ([2]) the desired regularity for m and h .

3 - Finite element approximation

Although system (3) has generally smooth solution, it represents a challenging numerical problem because it involves disparate time and space scales. Variables v and m have fast dynamics relative to h and the solution exhibits steep fronts, due to the short diffusion length and rapid variation of m . Formally speaking, this situation refers to the *stiffness* of the system of equations, which may impose severe constraints on the time and space discretization steps to achieve stability and accuracy.

In this work, system (3) is discretized first in space, by means of the finite element Galerkin method based on the following *weak formulation*:

Find $v \in \mathcal{L}^2(0, T; H^1(\Omega)) \cap \mathcal{C}^0(0, T; \mathcal{L}^2(\Omega))$; $m, h \in \mathcal{C}^0(0, T; \mathcal{L}^2(\Omega))$ such that

$$(5) \quad \begin{aligned} \frac{d}{dt} (v(t), \varphi) + \mathfrak{a}(v(t), \varphi) &= (f_1(v(t), m(t), h(t)), \varphi) + (f, \varphi) \\ \frac{d}{dt} (m(t), \varphi) &= (f_2(v(t), m(t)), \varphi) \\ \frac{d}{dt} (h(t), \varphi) &= (f_3(v(t), h(t)), \varphi) \\ v(0) = v_0 \quad m(0) = m_\infty(v_0) \quad h(0) &= h_\infty(v_0) \end{aligned}$$

$\forall \varphi \in H^1(\Omega)$, where d/dt is taken in the sense of distribution on Ω and we define

$$\mathfrak{a}(u, w) = \sigma \int_{\Omega} (\nabla w)^T \mathbf{M}_i \nabla u \, dx + \frac{g_r}{C_m} (u, w) \quad \forall u, w \in H^1(\Omega)$$

(\cdot, \cdot) being the scalar product in $\mathcal{L}^2(\Omega)$.

The Galerkin method consists in fixing a family of finite dimensional spaces $V_k \subset H^1(\Omega)$ of dimension N_k (tending to infinity as $k \rightarrow 0$) and in seeking approximate solutions $v_k, m_k, h_k : [0, T] \rightarrow V_k$ satisfying the following *ordinary differential equations system*:

$$(6) \quad \begin{aligned} \forall w_k \in V_k, \frac{d}{dt} (v_k(t), w_k) + \mathfrak{a}(v_k(t), w_k) &= (f_1(v_k(t), m_k(t), h_k(t)), w_k) + (f, w_k) \\ \frac{d}{dt} (m_k(t), w_k) &= (f_2(v_k(t), m_k(t)), w_k) \\ \frac{d}{dt} (h_k(t), w_k) &= (f_3(v_k(t), h_k(t)), w_k) \\ v_k(0) = v_{0,k} \quad m_k(0) = m_{0,k} \quad h_k(0) &= h_{0,k} \end{aligned}$$

where $v_{0,k}, m_{0,k}$ and $h_{0,k} \in V_k$ are suitable approximations of the initial data.

It is easy to verify that the bilinear form \mathfrak{a} is coercive and continuous, since $\mathfrak{a}(w, w) \geq \alpha \|w\|_1^2$, $\forall w \in H^1(\Omega)$, with $\alpha = \min(\alpha_0, \frac{g_r}{C_m})$, and $|\mathfrak{a}(\psi, \varphi)| \leq \gamma \|\psi\|_1 \|\varphi\|_1$, $\forall \varphi, \psi \in H^1(\Omega)$, where $\gamma = 9M + \frac{g_r}{C_m}$, $M = \max_{r,s} \|m_{r,s}\|_\infty$, $\mathbf{M}_i = (m_{r,s})$ and $\|\cdot\|_1$ is the H^1 -norm.

Functions $v_k(t), m_k(t)$ and $h_k(t)$ can be represented as linear combination of basis functions $\varphi_i, i = 1, \dots, N_k$, with time varying coefficients. Taking $w_k = \varphi_i, i = 1, 2, \dots, N_k$ in (6), we get the following system of nonlinear differential alge-

braic equations:

$$(7) \quad \mathbf{P} \frac{d\mathbf{v}}{dt}(t) + \mathbf{A}\mathbf{v}(t) = \mathbf{f}_1(\mathbf{v}(t), \mathbf{m}(t), \mathbf{h}(t)) + \mathbf{f}$$

$$(8) \quad \frac{d\mathbf{m}}{dt}(t) = \mathbf{f}_2(\mathbf{v}(t), \mathbf{m}(t))$$

$$(9) \quad \frac{d\mathbf{h}}{dt}(t) = \mathbf{f}_3(\mathbf{v}(t), \mathbf{h}(t))$$

$$\mathbf{v}(0) = \mathbf{v}_0 \quad \mathbf{m}(0) = \mathbf{m}_0 \quad \mathbf{h}(0) = \mathbf{h}_0$$

where $(\mathbf{y} = \mathbf{v}, \mathbf{m}, \mathbf{h})$:

$$\mathbf{y} = (y_1, \dots, y_{N_k})^T \quad y_k(t) = \sum_{j=1}^{N_k} y_j(t) \varphi_j \quad y_{0,k} = \sum_{j=1}^{N_k} y_{0,j} \varphi_j$$

$\mathbf{P} = (\int_{\Omega} \varphi_i \varphi_j dx; i, j = 1, \dots, N_k)$ is a positive definite matrix of dimension N_k (*mass matrix*); $\mathbf{A} = (a_{ij}), a_{ij} = \mathcal{C}(\varphi_j, \varphi_i), i, j = 1, \dots, N_k$, is a $N_k \times N_k$ symmetric definite positive matrix; $\mathbf{v}(t), \mathbf{m}(t), \mathbf{h}(t) \in \mathbf{R}^{N_k} \forall t \in [0, T]$. Functions $\mathbf{f}_1: \mathbf{R}^{3N_k} \rightarrow \mathbf{R}^{N_k}$ and $\mathbf{f}_{2,3}: \mathbf{R}^{2N_k} \rightarrow \mathbf{R}^{N_k}$ are still locally Lipschitz continuous.

For simplicity, we suppose in the sequel that Ω is a polygonal convex domain of \mathbf{R}^3 with Lipschitz continuous boundary and that (5) has still a unique bounded solution. The analysis below may be extended to more general (non polygonal) domains Ω by introducing the concept of isoparametric finite elements.

Let τ_k be a family of triangulations of $\overline{\Omega}$, associated to a reference polyhedron \widehat{F} by invertible affine maps T_F , for every $F \in \tau_k$, and suppose that $\tau_k, k > 0$ is *regular*, i.e there exists a constant $\sigma \geq 1$ such that

$$\max_{F \in \tau_k} \frac{k_F}{\varrho_F} \leq \sigma \quad \forall k > 0$$

where $k_F = \text{diam}(F)$ and $\varrho_F = \sup \{ \text{diam}(S) : S \text{ is a ball contained in } F \}$.

For instance, we can consider simplicial triangulations and the finite element spaces V_k of the continuous functions whose restrictions to each element of τ_k are linear polynomials

$$V_k = X_k^1 = \{ v_k \in \mathcal{C}^0(\overline{\Omega}) : v_{k|F} \in \mathcal{P}_1, \forall F \in \tau_k \}$$

where \mathcal{P}_1 is the space of polynomials that are of degree less than or equal to one. Note that $X_k^1 \subset H^1(\Omega)$.

If the integrals over an element are computed using a trapezoidal quadrature rule based on nodal values of the functions, then the mass matrix \mathbf{P} results of diago-

nal type, i.e. the mass matrix is *lumped*, and functions \mathbf{f}_s are defined as follows:

$$\begin{aligned} \mathbf{f}_1(\mathbf{v}, \mathbf{m}, \mathbf{h}) &= \mathbf{P}(f_1(v_1, m_1, h_1), \dots, f_1(v_{N_k}, m_{N_k}, h_{N_k}))^T \\ \mathbf{f}_s(\mathbf{v}, \mathbf{m}, \mathbf{h}) &= (f_s(v_1, m_1, h_1), \dots, f_s(v_{N_k}, m_{N_k}, h_{N_k}))^T, \quad s = 2, 3. \end{aligned}$$

4 - Semi-discrete error analysis

In this section, we are going to study the properties of the semidiscrete system (6), or equivalently of system (7)-(9). We point out that these results are independent of the particular choice of space V_k , which doesn't need to be X_k^1 .

The main results are the existence of a unique global solution to system (6) and the convergence of the Galerkin semidiscrete approximation to the exact solution of (5). The idea of the proof was used by Thomée [13] for general semilinear equations of parabolic type. Let's start with the following remark.

Remark 1. For fixed k , system (7)-(9) has a unique local solution $(\mathbf{v}, \mathbf{m}, \mathbf{h}): [0, T'] \rightarrow \mathbf{R}^{3N_k}$, for some $T' \leq T$, since \mathbf{P} is a positive definite matrix and functions $f_{s,r}(\mathbf{v}, \mathbf{m}, \mathbf{h})$, $s = 1, 2, 3$, $r = 1, \dots, N_k$ are Lipschitz continuous with respect to variables \mathbf{v}, \mathbf{m} and \mathbf{h} , in any bounded open subset B of \mathbf{R}^{3N_k} .

Moreover, the functions \mathbf{f}_s , $s = 1, 2, 3$ in (7)-(9) are $\mathcal{C}^2([0, T'] \times B)$, therefore the solution $(\mathbf{v}, \mathbf{m}, \mathbf{h}) \in \mathcal{C}^3([0, T'])$.

In the analysis developed below we shall show that, under usual assumptions on $\{V_k\}$ and for k sufficiently small, T' can be taken to be equal to T . Hence, global solvability of system (6) follows.

Definition 1. We say that the family $\{V_k\}$ is of class $\mathcal{S}_{s,\mu}$ (with $s \leq \mu$) if $V_k \subset H^s(\Omega)$ and

$$\forall w \in H^\mu(\Omega), \quad \inf_{w_k \in V_k} \sum_{j=0}^s k^j \|w - w_k\|_j \leq Ck^\mu \|w\|_\mu.$$

Here and below C denotes constants, not necessarily the same at different occurrences, which are independent of k and the functions involved. Similarly, c will denote constants which are independent of k but which may depend on the exact solution $\mathbf{u}(t) = (v(t), m(t), h(t))^T$ of (5).

In the following we assume that the family $\{V_k\}$ is of class $\mathcal{S}_{s,\mu}$ and satisfies the following *inverse inequality*

$$(10) \quad \|w_k\|_\infty \leq Ck^{-\nu} \|w_k\|_0 \quad \forall w_k \in V_k, k \leq k_1 \quad \text{for some } \nu \text{ and } k_1.$$

We introduce the *interpolation operator* $\pi_k : \mathcal{C}^0(\overline{\Omega}) \rightarrow V_k$

$$\pi_k(w) = \sum_{i=1}^{N_k} w(\mathbf{a}_i) \varphi_i$$

where \mathbf{a}_i are the nodes on $\overline{\Omega}$ and we assume that the following inequalities hold for any $w \in H^\mu(\Omega)$, $\mu \geq 2$:

$$(11) \quad \|w - \pi_k(w)\|_0 \leq Ck^\mu \|w\|_\mu \quad \|w - \pi_k(w)\|_\infty \leq Ck^{\mu-\nu} \|w\|_\mu.$$

Remark 2. The family $\{X_k^1\}$ of Section 3 is of class $S_{1,2}$ (see [10]). Moreover, suppose that the family of triangulations τ_k is quasi-uniform, i.e. it is regular and there exists a constant $\gamma > 0$ such that $\min_{F \in \tau_k} k_F \geq \gamma k$, $\forall k > 0$, and that the quantity $k = \max k_F$ approaches zero. Then the inverse assumption (10) and inequalities (11) hold with $\mu = 2$ and $\nu = \frac{n}{2}$, where n is the space dimension (see [1] and [10]).

Let Σ be the range of the exact solution $\mathbf{u} = (v, m, h)^T$ of (5)

$$\Sigma = \{\mathbf{u}(t, \mathbf{x}) : \mathbf{x} \in \overline{\Omega}, t \in [0, T]\}$$

$\mathbf{f} = (f_1, f_2, f_3)^T$ is Lipschitz continuous in the closed neighborhood Σ_δ of Σ , for any $\delta > 0$. Let's fix one of these δ , sufficiently large to include the initial datum $(v_k(0), m_k(0), h_k(0))^T$ in Σ_δ , and let $K_i > 0$ be constants such that the following Lipschitz conditions hold in Σ_δ :

$$|f_i(\mathbf{w}_1) - f_i(\mathbf{w}_2)| \leq K_i |\mathbf{w}_1 - \mathbf{w}_2|_1 \quad i = 1, 2, 3.$$

Heuristically, we might argue that, since the approximate solution $\mathbf{u}_k = (v_k, m_k, h_k)^T$ is always going to be close to \mathbf{u} , it belongs to Σ_δ . In order to show that this is the case, we have to provide maximum norm estimates for the approximation error, since closeness in the sense of \mathcal{L}^2 or H^1 does not automatically imply that \mathbf{u}_k belongs to Σ_δ for small k .

Now we state the main theorem of this section.

Theorem 2. Let $\{V_k\}$ be of class $S_{1,\mu}$ ($\mu \geq 2$), with $V_k \subset H^1(\Omega)$, and assume that (10), (11) and (14) below hold for some $\nu < \mu$. Then, if the solution $\mathbf{u}(t) \in (H^\mu(\Omega))^3$, $\forall t \in [0, T]$ and $m_t(t), h_t(t) \in H^\mu(\Omega)$ almost everywhere in $[0, T]$, there is a k_0 such that, for $k \leq k_0$, the solution \mathbf{u}_k of (6) exists for $t \leq T$, and for these t we have

$$\|(v_k(t), m_k(t), h_k(t))^T - (v(t), m(t), h(t))^T\|_0 \leq ck^\mu.$$

The error analysis between \mathbf{u}_k and \mathbf{u} can be accomplished comparing the Galerkin solution $v_k(t)$ to the *elliptic projection* $W(t)$ of the exact solution $v(t)$ of (5) onto V_k , defined by

$$(12) \quad \mathfrak{a}(W(t), w_k) = \mathfrak{a}(v(t), w_k) \quad \forall w_k \in V_k.$$

As it is well known [8], the error in this projection $\varrho(t) = v(t) - W(t)$ satisfies for $t \leq T$

$$(13) \quad \|\varrho(t)\|_j \leq ck^{\mu-j} \|v(t)\|_{\mu} \leq ck^{\mu-j} \quad j = 0, 1.$$

The rest of the error in v_k is then $\theta(t) = v_k(t) - W(t) \in V_k$.

We shall assume that the initial values are chosen in such a way that

$$(14) \quad \|(v_k - W)(0)\|_0 \leq ck^{\mu} \quad \|(m_k - m)(0)\|_0 \leq ck^{\mu} \quad \|(h_k - h)(0)\|_0 \leq ck^{\mu}$$

with $W(t)$ defined by (12). For instance, as a consequence of (13), (14) follows from the triangular inequality and from condition

$$(15) \quad \|(v_k - v)(0)\|_0 \leq ck^{\mu} \quad \|(m_k - m)(0)\|_0 \leq ck^{\mu} \quad \|(h_k - h)(0)\|_0 \leq ck^{\mu}.$$

Proof. Let t^k be the largest number less than or equal to T such that \mathbf{u}_k exists and belongs to Σ_{δ} for $t \leq t^k$: $t^k = \sup \{s \leq T : \mathbf{u}_k(t) \in \Sigma_{\delta} \forall t \leq s\}$.

Subtracting (5) from (6) and taking into account (12), we have for almost every $t \in [0, t^k)$

$$(16) \quad \begin{aligned} (\theta_t, w_k) - (\varrho_t w_k) + \mathfrak{a}(\theta, w_k) &= (f_1(v_k, m_k, h_k) - f_1(v, m, h), w_k) \\ &= (f_1(v_k, m_k, h_k) - f_1(W, m_k, h_k), w_k) + (f_1(W, m_k, h_k) - f_1(v, m_k, h_k), w_k) \\ &\quad + (f_1(v, m_k, h_k) - f_1(v, m, h_k), w_k) + (f_1(v, m, h_k) - f_1(v, m, h), w_k). \end{aligned}$$

Setting $w_k = \theta(t)$ in (16) and using the Cauchy-Schwarz inequality, the coerciveness of \mathfrak{a} , the Lipschitz condition for f_1 and the well known inequality $ab \leq \varepsilon a^2 + \frac{1}{4\varepsilon} b^2$, $\forall \varepsilon > 0$, we obtain for $t \in [0, t^k)$

$$\frac{1}{2} \frac{d}{dt} \|\theta\|_0^2 \leq \frac{1}{4\varepsilon} \|\varrho_t\|_0^2 + (4\varepsilon + K_1^2) \|\theta\|_0^2 + \frac{K_1^2}{4\varepsilon} \|\varrho\|_0^2 + \frac{K_1^2}{4\varepsilon} \|M\|_0^2 + \frac{K_1^2}{4\varepsilon} \|H\|_0^2$$

where we have set $M(t) = m_k(t) - m(t)$ and $H(t) = h_k(t) - h(t)$ for $t \in [0, t^k)$.

Since time differentiation commutes with elliptic projection, by (13) we have $\|\varrho_t\| \leq ck^{\mu}$. Hence

$$(17) \quad \frac{1}{2} \frac{d}{dt} \|\theta\|_0^2 < ck^{2\mu} + C_1 \|\theta\|_0^2 + C_2 \|M\|_0^2 + C_2 \|H\|_0^2$$

where $C_1 = 4\varepsilon + K_1^2 > 0$ and $C_2 = \frac{K_1^2}{4\varepsilon} > 0$.

Now, let's consider the first gating equation in its weak form; we have

$$(18) \quad (M_t(t), w_k) = (f_2(v_k, m_k) - f_2(v, m), w_k) \quad \forall w_k \in V_k.$$

Set $w_k = m_k(t) - \psi_k \in V_k$, where ψ_k is an arbitrary element of V_k . As before, this yields

$$\frac{1}{2} \frac{d}{dt} \|M\|_0^2 \leq C_3 \|M\|_0^2 + \frac{1}{2} \|M_t\|_0^2 + 2 \|m - \psi_k\|_0^2 + K_2^2 \|\theta\|_0^2 + K_2^2 \|\varrho\|_0^2.$$

If we choose $\psi_k = \pi_k(m(t))$ for all $t \in [0, t^k)$, then from (11) it follows that $\|m - \psi_k\| \leq ck^\mu$. Therefore

$$(19) \quad \frac{1}{2} \frac{d}{dt} \|M\|_0^2 \leq C_3 \|M\|_0^2 + \frac{1}{2} \|M_t\|_0^2 + K_2^2 \|\theta\|_0^2 + K_2^2 \|\varrho\|_0^2 + ck^{2\mu}.$$

Moreover, choosing $w_k = (m_k)_t - \pi_k(m_t(t)) \in V_k$ in (18), for almost any $t \in [0, t^k)$ we have

$$\|M_t\|_0^2 \leq 4\varepsilon \|M_t\|_0^2 + \left(\frac{1}{4\varepsilon} + 3\varepsilon\right) \|m_t - \pi_k(m_t)\|_0^2 + \frac{K_2^2}{2\varepsilon} \|\theta\|_0^2 + \frac{K_2^2}{2\varepsilon} \|\varrho\|_0^2 + \frac{K_2^2}{2\varepsilon} \|M\|_0^2$$

that is, for $\varepsilon < \frac{1}{4}$,

$$\|M_t\|_0^2 \leq ck^{2\mu} + C_3 \|\theta\|_0^2 + C_3 \|\varrho\|_0^2 + C_4 \|M\|_0^2.$$

The last inequality can be combined with (19), yielding

$$(20) \quad \frac{1}{2} \frac{d}{dt} \|M\|_0^2 \leq c' k^{2\mu} + C'_3 \|\theta\|_0^2 + C'_3 \|\varrho\|_0^2 + C'_4 \|M\|_0^2.$$

Similarly, the second gating equation yields

$$(21) \quad \frac{1}{2} \frac{d}{dt} \|H\|_0^2 \leq c' k^{2\mu} + C'_5 \|\theta\|_0^2 + C'_5 \|\varrho\|_0^2 + C'_6 \|H\|_0^2.$$

Now, we sum up (17), (20) and (21) and get

$$\frac{1}{2} \frac{d}{dt} (\|\theta\|_0^2 + \|M\|_0^2 + \|H\|_0^2) \leq C(\|\theta\|_0^2 + \|M\|_0^2 + \|H\|_0^2) + ck^{2\mu}.$$

Let's integrate over $[0, t]$, $t < t^k$

$$\|\theta(t)\|_0^2 + \|M(t)\|_0^2 + \|H(t)\|_0^2 \leq C \int_0^t (\|\theta(s)\|_0^2 + \|M(s)\|_0^2 + \|H(s)\|_0^2) ds + cTk^{2\mu}.$$

Gronwall's lemma and assumption (14) assure that

$$\begin{aligned} & \|\theta(t)\|_0^2 + \|M(t)\|_0^2 + \|H(t)\|_0^2 \\ & \leq (\|\theta(0)\|_0^2 + \|M(0)\|_0^2 + \|H(0)\|_0^2) + cTk^{2\mu} e^{cT} \leq ck^{2\mu}. \end{aligned}$$

Hence, for $t < t^k$, we have

$$\begin{aligned} \|\mathbf{u}_k(t) - \mathbf{u}(t)\|_0^2 &= \|v_k(t) - v(t)\|_0^2 + \|m_k(t) - m(t)\|_0^2 + \|h_k(t) - h(t)\|_0^2 \\ &\leq 2\|\theta\|_0^2 + 2\|\varrho\|_0^2 + \|M\|_0^2 + \|H\|_0^2 \leq ck^{2\mu} \end{aligned}$$

that is

$$(22) \quad \|(v_k(t), m_k(t), h_k(t)) - (v(t), m(t), h(t))\|_0 \leq ck^\mu .$$

Moreover, for $t < t^k$, we have

$$(23) \quad \begin{aligned} \|\mathbf{u}_k - \mathbf{u}\|_\infty &\leq \|(v_k, m_k, h_k) - (v, m, h_k)\|_\infty \\ &\quad + \|(v, m_k, h_k) - (v, m, h_k)\|_\infty + \|(v, m, h_k) - (v, m, h)\|_\infty \\ &= \|v_k - v\|_\infty + \|m_k - m\|_\infty + \|h_k - h\|_\infty . \end{aligned}$$

Let's consider the first term on the right-hand side of (23); set $w_k = \pi_k(v)$

$$\begin{aligned} \|v_k - v\|_\infty &\leq \|v_k - w_k\|_\infty + \|w_k - v\|_\infty \leq ck^{-\nu} \|v_k - w_k\|_0 + \|w_k - v\|_\infty \\ &\leq ck^{-\nu} \|v_k - v\|_0 + ck^{-\nu} \|w_k - v\|_0 + \|w_k - v\|_\infty \\ &\leq ck^{\mu-\nu} + ck^{\mu-\nu} + ck^{\mu-\nu} \leq \bar{c}k^{\mu-\nu} < \frac{\delta}{6} \quad k \leq k_1 \end{aligned}$$

since $\nu < \mu$. Similar results can be found for the other two terms in (23). Therefore, for $t < t^k$ we have

$$\|\mathbf{u}_k(t) - \mathbf{u}(t)\|_\infty < \frac{\delta}{2} \quad k \leq k_0$$

where k_0 is sufficiently small and independent of t^k . Hence we may conclude by continuity that t^k cannot be smaller than T , that is $t^k = T$ for $k \leq k_0$ and, by (22) $\|\mathbf{u}_k(t) - \mathbf{u}(t)\|_0 \leq ck^\mu, \forall t \leq T$.

In the case of linear finite elements, for suitable approximate initial data, if $\mathbf{u}(t) \in (H^2(\Omega))^3 \forall t \in [0, T]$ and $m_t(t), h_t(t) \in H^2(\Omega)$ almost everywhere in $[0, T]$, the assumptions of Theorem 2 are satisfied for $\mu = 2$ and $\nu = \frac{n}{2}$, thus yielding a second order \mathcal{L}^2 convergence of the Galerkin semi-discretization to the exact solution \mathbf{u} .

5 - Time discretization

As a consequence of the arguments above, we can replace $f_i, i = 1, 2, 3$, in (3) by smooth functions \tilde{f}_i , which coincide with f_i on a neighborhood $\Sigma_\delta, \supset \Sigma_\delta$ but

which are bounded in \mathbf{R}^{3N_k} together with their partial derivatives of order less than or equal to 2. This replacement doesn't affect either the exact solution \mathbf{u} or the discrete solution \mathbf{u}_k if k is sufficiently small.

Let's now turn to the time discretization of system (7)-(9), where the vector functions \mathbf{f}_i , $i = 1, 2, 3$, have to be modified into $\tilde{\mathbf{f}}_i$. The application of a convergent time-stepping assures that, for Δt sufficiently small, the totally discretized solution belongs to Σ_δ , where $\tilde{\mathbf{f}}_i \equiv \mathbf{f}_i$.

Since the elliptic term in the RD system (3) is small with respect to the reaction term, an explicit method turns out to be stable only for very small time steps. This can be avoided by using an implicit or semi-implicit time-stepping. Therefore, we discretize the semidiscrete system (7)-(9) by means of the Crank-Nicholson method

$$\begin{aligned}
 \mathbf{v}^{l+1} &= -\frac{\Delta t}{2} \mathbf{P}^{-1} \mathbf{A} \mathbf{v}^{l+1} + \frac{\Delta t}{2} \mathbf{P}^{-1} \tilde{\mathbf{f}}_1(\mathbf{v}^{l+1}, \mathbf{m}^{l+1}, \mathbf{h}^{l+1}) \\
 &\quad + (\mathbf{I} - \frac{\Delta t}{2} \mathbf{P}^{-1} \mathbf{A}) \mathbf{v}^l + \frac{\Delta t}{2} \mathbf{P}^{-1} \tilde{\mathbf{f}}_1(\mathbf{v}^l, \mathbf{m}^l, \mathbf{h}^l) + \Delta t \mathbf{P}^{-1} \mathbf{f} \\
 (24) \quad \mathbf{m}^{l+1} &= \frac{\Delta t}{2} \tilde{\mathbf{f}}_2(\mathbf{v}^{l+1}, \mathbf{m}^{l+1}) + \frac{\Delta t}{2} \tilde{\mathbf{f}}_2(\mathbf{v}^l, \mathbf{m}^l) + \mathbf{m}^l \\
 \mathbf{h}^{l+1} &= \frac{\Delta t}{2} \tilde{\mathbf{f}}_3(\mathbf{v}^{l+1}, \mathbf{h}^{l+1}) + \frac{\Delta t}{2} \tilde{\mathbf{f}}_3(\mathbf{v}^l, \mathbf{h}^l) + \mathbf{h}^l.
 \end{aligned}$$

Scheme (24) may be rewritten in the form:

$$(25) \quad \mathbf{x}^{l+1} = \Delta t \mathbf{g}(\mathbf{x}^{l+1}) + \mathbf{q}(\mathbf{x}^l) \quad l = 0, 1, \dots, M-1$$

where $\mathbf{x}^l = (\mathbf{v}^l, \mathbf{m}^l, \mathbf{h}^l)^T \in \mathbf{R}^{3N_k}$ are approximations of $(\mathbf{v}, \mathbf{m}, \mathbf{h})^T$ at time t^l and $\mathbf{g}, \mathbf{q}: \mathbf{R}^{3N_k} \rightarrow \mathbf{R}^{3N_k}$ are \mathcal{C}^2 Lipschitz continuous vector functions. A standard fixed point argument shows that, given the initial vector $\mathbf{x}^0 = (\mathbf{v}_0, \mathbf{m}_0, \mathbf{h}_0) \in \Sigma_\delta$, for Δt sufficiently small, equation (25) has a unique solution \mathbf{x}^{l+1} , for each $l = 0, 1, \dots, M-1$.

System (24) is solved, at each time step, by means of a *quasilinearization technique* [11] consisting of the first step of Newton's method. The continuous nature of the solution guarantees that, for small Δt , the value at the previous time step \mathbf{x}^l is close to the root \mathbf{x}^{l+1} being sought and the quadratic convergence of Newton's method implies that the vector computed by one iteration is a good approximation of \mathbf{x}^{l+1} .

When the integrals in (6) are computed by means of trapezoidal quadrature rule, we get the following iterative scheme ($l = 0, \dots, M - 1$):

$$(26) \quad \left(\frac{2}{\Delta t} \mathbf{P} + \mathbf{A} + \mathbf{J}^l \right) \mathbf{v}^{l+1} + \mathbf{D}_1^l \mathbf{m}^{l+1} + \mathbf{D}_2^l \mathbf{h}^{l+1} = \mathbf{b}_1^l \\ - \frac{\Delta t}{2} \mathbf{D}_3^l \mathbf{v}^{l+1} + \left(\mathbf{I} + \frac{\Delta t}{2} \mathbf{D}_4^l \right) \mathbf{m}^{l+1} = - \frac{\Delta t}{2} \mathbf{D}_3^l \mathbf{v}^l + \left(\mathbf{I} + \frac{\Delta t}{2} \mathbf{D}_4^l \right) \mathbf{m}^l + \Delta t \mathbf{b}_2^l \\ - \frac{\Delta t}{2} \mathbf{D}_5^l \mathbf{v}^{l+1} + \left(\mathbf{I} + \frac{\Delta t}{2} \mathbf{D}_6^l \right) \mathbf{h}^{l+1} = - \frac{\Delta t}{2} \mathbf{D}_5^l \mathbf{v}^l + \left(\mathbf{I} + \frac{\Delta t}{2} \mathbf{D}_6^l \right) \mathbf{h}^l + \Delta t \mathbf{b}_3^l$$

where

$$\mathbf{J}^l = \text{diag} \left(\frac{1}{C_m} \mathbf{P}_{ii} g_{Na} (m_i^l)^3 h_i^l \right) \quad \mathbf{D}_1^l = \text{diag} \left(\frac{1}{C_m} \mathbf{P}_{ii} (3g_{Na} (m_i^l)^2 h_i^l (v_i^l - v_{Na})) \right)$$

$$\mathbf{D}_2^l = \text{diag} \left(\frac{1}{C_m} \mathbf{P}_{ii} (g_{Na} (m_i^l)^3 (v_i^l - v_{Na})) \right)$$

$$\mathbf{D}_3^l = \text{diag} \left((1 - m_i^l) \frac{\partial \alpha_m (v_i^l)}{\partial v_i^l} - \frac{\partial \beta_m (v_i^l)}{\partial v_i^l} m_i^l \right) \quad \mathbf{D}_4^l = \text{diag} (\alpha_m (v_i^l) + \beta_m (v_i^l))$$

$$\mathbf{D}_5^l = \text{diag} \left((1 - h_i^l) \frac{\partial \alpha_h (v_i^l)}{\partial v_i^l} - \frac{\partial \beta_h (v_i^l)}{\partial v_i^l} h_i^l \right) \quad \mathbf{D}_6^l = \text{diag} (\alpha_h (v_i^l) + \beta_h (v_i^l))$$

$$\mathbf{b}_1^l = \left(\frac{2}{\Delta t} \mathbf{P} - \mathbf{A} + \mathbf{J}^l \right) \mathbf{v}^l + \mathbf{D}_1^l \mathbf{m}^l + \mathbf{D}_2^l \mathbf{h}^l + 2\tilde{\mathbf{f}}_1^l + 2\mathbf{f}$$

$$\mathbf{b}_2^l = (- (\alpha_m (v_1^l) + \beta_m (v_1^l)) m_1^l + \alpha_m (v_1^l), \dots, - (\alpha_m (v_{N_k}^l) + \beta_m (v_{N_k}^l)) m_{N_k}^l + \alpha_m (v_{N_k}^l))^T$$

$$\mathbf{b}_3^l = (- (\alpha_h (v_1^l) + \beta_h (v_1^l)) h_1^l + \alpha_h (v_1^l), \dots, - (\alpha_h (v_{N_k}^l) + \beta_h (v_{N_k}^l)) h_{N_k}^l + \alpha_h (v_{N_k}^l))^T$$

$$\tilde{\mathbf{f}}_1^l = \tilde{\mathbf{f}}_1 (\mathbf{v}^l, \mathbf{m}^l, \mathbf{h}^l).$$

Substituting \mathbf{m}^{l+1} and \mathbf{h}^{l+1} in (26), we get the single linear system

$$(27) \quad \mathbf{G}^l \mathbf{v}^{l+1} = \mathbf{E}^l \mathbf{v}^l + \mathbf{F}^l$$

where:

$$\mathbf{G}^l = \left(\frac{2}{\Delta t} \mathbf{P} + \mathbf{A} + \frac{\Delta t}{2} \mathbf{D}_1^l \left(\mathbf{I} + \frac{\Delta t}{2} \mathbf{D}_4^l \right)^{-1} \mathbf{D}_3^l + \frac{\Delta t}{2} \mathbf{D}_2^l \left(\mathbf{I} + \frac{\Delta t}{2} \mathbf{D}_6^l \right)^{-1} \mathbf{D}_5^l + \mathbf{J}^l \right)$$

$$\mathbf{E}^l = \left(\frac{2}{\Delta t} \mathbf{P} - \mathbf{A} + \frac{\Delta t}{2} \mathbf{D}_1^l \left(\mathbf{I} + \frac{\Delta t}{2} \mathbf{D}_4^l \right)^{-1} \mathbf{D}_3^l + \frac{\Delta t}{2} \mathbf{D}_2^l \left(\mathbf{I} + \frac{\Delta t}{2} \mathbf{D}_6^l \right)^{-1} \mathbf{D}_5^l + \mathbf{J}^l \right)$$

$$\mathbf{F}^l = 2\tilde{\mathbf{f}}_1^l + 2\mathbf{f} - \Delta t \mathbf{D}_1^l \left(\mathbf{I} + \frac{\Delta t}{2} \mathbf{D}_4^l \right)^{-1} \mathbf{b}_2^l - \Delta t \mathbf{D}_2^l \left(\mathbf{I} + \frac{\Delta t}{2} \mathbf{D}_6^l \right)^{-1} \mathbf{b}_3^l.$$

It's worth noticing that matrix \mathbf{G}^l turns out to be a Stieltjes matrix with respect to the discretization steps used in the applications, and this fact gives stability to our numerical scheme.

Now we analyse the above described time-stepping on a single non-linear ordinary differential equation of the form

$$(28) \quad y'(t) = \varphi(y(t)) \quad 0 \leq t \leq T \quad y(0) = y_0$$

where φ is a Lipschitz continuous function of class $\mathcal{C}^2(\mathbf{R})$. The Crank-Nicholson scheme applied to (28) yields the implicit finite difference scheme

$$\frac{1}{\Delta t} (y^{l+1} - y^l) = \frac{1}{2} \varphi(y^{l+1}) + \frac{1}{2} \varphi(y^l) \quad 0 \leq l \leq M-1.$$

The application of one step of Newton's method starting from y^l yields the scheme

$$(29) \quad \frac{1}{\Delta t} (y^{l+1} - y^l) = \varphi(y^l) + \frac{1}{2} \varphi'(y^l)(y^{l+1} - y^l)$$

which represents a correction of the forward Euler method. Scheme (29) turns out to be second order accurate with respect to Δt . In fact, by differentiating equation (28) with respect to t , we get

$$\frac{d^2 y}{dt^2} = \varphi'(y(t)) y'(t).$$

Therefore, the exact solution $y(t^{l+1})$ may be expressed as the Taylor expansion

$$(30) \quad y(t^{l+1}) = y^l + \varphi(y^l) \Delta t + \frac{\varphi'(y^l) y'(t^l)}{2} \Delta t^2 + O(\Delta t^3).$$

Taking into account (30), the local truncation error at point t^{l+1} is given by

$$\tau(t^l, y^l; \Delta t) = \frac{y(t^{l+1}) - y^l}{\Delta t} - \varphi(y^l) - \frac{1}{2} \varphi'(y^l)(y(t^{l+1}) - y^l) = O(\Delta t^2).$$

Therefore, the one-step method is consistent and second order accurate.

Moreover, for $\varphi \in \mathcal{C}^2(\mathbf{R})$ bounded together with its derivatives of order less than or equal to two, the consistency of scheme (29) implies its convergence and the order of this convergence is given by the local truncation error.

The absolute stability of method (29) may be analysed on the test equation

$$(31) \quad y'(t) = -\lambda y(t) \quad t \geq 0 \quad y(0) = y_0$$

where λ is a positive real number. Scheme (29) coincides with the Crank-Nicholson scheme when applied to the linear problem (31), therefore the scheme is A-stable.

References

- [1] P. G. CIARLET, *The finite element method for elliptic problems*, North-Holland, Amsterdam 1978.
- [2] E. A. CODDINGTON and N. LEVINSON, *Theory of ordinary differential equations*, McGraw-Hill, New York 1955.
- [3] P. COLLI FRANZONE and L. GUERRI, *Models of the spreading of excitation in myocardial tissue*, in High Performance Comput. Biomed. Res., ed T. C. Pilkington, Boca Raton, USA 1993.
- [4] P. COLLI FRANZONE and L. GUERRI, *Spreading of excitation in 3-D models of the anisotropic cardiac tissue. Validation of the eikonal model*, Math. Biosci. **113** (1993), 145-209.
- [5] L. EBIHARA and E. JOHNSON, *Fast sodium current in cardiac muscle. A quantitative description*, Biophys. J. **32** (1980), 779-790.
- [6] C. S. HENRIQUEZ, *Simulating the electrical behavior of cardiac tissue using the bidomain model*, Critical Rev. Biomed. Engrg. **21** (1993), 1-77.
- [7] A. L. HODGKIN and A. F. HUXLEY, *A quantitative description of membrane current and its application to conduction and excitation in nerve*, J. Physiol. **117** (1952), 500-544.
- [8] J. A. NITSCHKE, *Ein Kriterium für die Quasi-optimalität des Ritzschen Verfahrens*, Numer. Math. **11** (1968), 346-348.
- [9] J. M. ORTEGA and W. C. RHEINBOLDT, *Iterative solution of nonlinear equations in several variables*, Academic Press, New York 1970.
- [10] A. QUARTERONI and A. VALLI, *Numerical approximation of partial differential equations*, Springer, Berlin 1994.
- [11] B. L. RALL, *Computational solution of nonlinear operator equations*, Wiley, Chichester, UK 1969.
- [12] S. SANFELICI, *Numerical simulation of the depolarization process in anisotropic myocardium*, Quad. Dip. di Matem. **105**, Univ. Parma 1994.
- [13] V. THOMÉE and L. B. WAHLBIN, *On Galerkin methods in semilinear parabolic problems*, SIAM J. Numer. Anal. **12** (1975), 378-389.

Sommaro

In questo lavoro si considera un modello macroscopico del processo di eccitazione nel miocardio anisotropo. Il modello è descritto da un sistema di reazione-diffusione (RD) costituito da un'equazione parabolica semilineare accoppiata con due equazioni differenziali ordinarie che, sotto opportune ipotesi sui dati, ammette un'unica soluzione classica globale. Viene presentato un metodo numerico per l'approssimazione del sistema RD ottenuto combinando il metodo semidiscreto di Galerkin, il metodo di Crank-Nicolson e una tecnica di «quasilinearizzazione». Il principale risultato è l'analisi della convergenza e la stima dell'errore per le approssimazioni spaziale e temporale.

* * *

