



uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES
ET POSTDOCTORALES



FACULTY OF GRADUATE AND
POSTDOCTORAL STUDIES

Marc Ethier

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (mathématiques)

GRADE / DEGREE

Département de mathématiques et statistique

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Schémas de discrétisation en temps pour le modèle bidomaine

TITRE DE LA THÈSE / TITLE OF THESIS

Y. Bourgault

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

B. Dionne

D. Amundsen

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

SCHÉMAS DE DISCRÉTISATION EN TEMPS POUR LE MODÈLE BIDOMAINÉ

Par
Marc Ethier, B. Sc., M. Sc.
Juillet 2006

Thèse soumise à la
Faculté des études supérieures et postdoctorales
Université d'Ottawa
en vue de l'obtention de la maîtrise ès sciences¹

© Copyright 2006
par Marc Ethier, B. Sc., M. Sc., Ottawa, Canada

¹Le programme de M. Sc. est un programme conjoint avec l'Université Carleton, administré par l'Institut d'études supérieures et de recherche en mathématiques et en statistique d'Ottawa-Carleton



Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN: 978-0-494-18415-8
Our file *Notre référence*
ISBN: 978-0-494-18415-8

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

Résumé

Le modèle bidomaine est un système d'équations aux dérivées partielles fréquemment utilisé pour représenter la propagation d'ondes de potentiel électrique dans le muscle cardiaque. Il est composé d'une équation parabolique couplée à une équation elliptique, auxquelles nous rattachons au moins une équation différentielle pour modéliser l'activité ionique du myocarde. La nature de ce système en rend la solution difficile.

Nous considérerons ici l'application de plusieurs méthodes numériques d'intégration en temps implicites, semi-implicites et explicites à la résolution de ce modèle. Nous démontrerons leur stabilité sous certaines conditions du pas de temps Δt , et nous comparerons les solutions numériques obtenues par ces méthodes pour deux cas tests : un en 1D, le développement d'une onde de potentiel qui se déplace le long du domaine, et un autre en 2D, le développement d'une onde spirale.

Les schémas que nous recommandons sont celui de Crank-Nicolson-Adams-Bashforth et le schéma SBDF de 2^e ordre. Ces méthodes semi-implicites produisent une bonne solution, leur stabilité ne dépend pas de la finesse du maillage en espace comme les méthodes explicites, et elles ne requièrent pas la résolution d'un système non linéaire comme les méthodes implicites.

Abstract

The bidomain model is a system of partial differential equations frequently used to model the propagation of electrical potential waves in the heart muscle. It contains a coupled parabolic and elliptic equation, as well as at least one ordinary differential equation to model the ion activity of the myocardium. The nature of this system makes it especially hard to solve.

We will consider here several implicit, semi-implicit and explicit time-stepping methods applied to the solution of this model. We will show that they remain stable under certain conditions for the time step Δt , and we will compare the numerical solutions produced by these methods for two test cases : one in 1D, the propagation of a potential wave across the domain, and another in 2D, the formation of a spiral wave.

We recommend using either the Crank-Nicolson–Adams-Bashforth method or the second order SBDF method. These semi-implicit methods produce a good numerical solution, unlike the explicit methods, their stability does not depend on how fine the spatial mesh is, and unlike the implicit methods, they do not require the resolution of a system of nonlinear equations.

Remerciements

Je désire remercier mon directeur de thèse, Yves Bourgault. Il m'a fourni de bons conseils au cours de ma recherche et de la rédaction de cette thèse, et m'a consacré beaucoup de temps. Sans lui, je n'aurais pas pu obtenir les résultats que j'ai obtenus.

Je voudrais aussi remercier le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG), qui a subventionné les deux premières années de ma maîtrise.

Finalement, je voudrais remercier mes collègues étudiants diplômés, que j'ai côtoyés au travail et à l'extérieur au cours des dernières années, et avec qui j'ai discuté, travaillé, eu du plaisir, etc.

Dédicace

Je voudrais dédier ce travail à Sylvie. Elle m'a montré ce que c'est que d'aimer et d'être aimé, et m'a supporté au cours des derniers mois. Sans elle, ils auraient été beaucoup plus difficiles.

Table des matières

Résumé	ii
Abstract	iii
Remerciements	iv
Dédicace	v
1 Introduction	1
1.1 Le modèle bidomaine	1
1.2 But de la thèse	4
2 Méthodes numériques	6
2.1 Discrétisation en espace	6
2.2 Discrétisation en temps	7
2.2.1 Méthodes de 1 ^{er} ordre	8
2.2.2 Méthodes de 2 ^e ordre	9
2.2.3 Méthode de 3 ^e ordre	10
3 Analyse numérique des schémas	12
3.1 Stabilité de la méthode semi-discrète	12
3.2 Stabilité des méthodes totalement discrètes	15
3.2.1 Méthodes implicites	15
3.2.2 Méthodes semi-implicites	19
3.2.3 Méthodes explicites	24

4 Résultats numériques	33
4.1 Comparaison des schémas en 1D	33
4.1.1 Problème	33
4.1.2 Résultats	34
4.2 Comparaison des schémas en 2D	44
4.2.1 Problème	44
4.2.2 Résultats	45
5 Conclusion	57
5.1 Résumé	57
5.2 Recherche future	58
Bibliographie	60

Chapitre 1

Introduction

Dans ce chapitre, nous commençons par présenter le modèle bidomaine, qui s'avère être très populaire pour modéliser mathématiquement la propagation d'ondes de potentiel électrique dans le muscle cardiaque, et nous expliquons ensuite ce qui fait l'originalité de la recherche que nous présentons dans cette thèse, c'est-à-dire la comparaison de différentes méthodes d'intégration en temps dans le cadre de la résolution numérique des équations du modèle bidomaine.

1.1 Le modèle bidomaine

Le cœur est une pompe double faite de tissu musculaire qui sert à pomper le sang, d'abord vers les poumons dans le but de l'oxygéner, puis vers tous les organes du corps pour les irriguer. Pour ce faire, il possède un groupe de cellules appelé le *noeud sinoatrial* dont la fonction est de produire une pulsation électrique qui est ensuite transmise à travers des cellules dédiées à cette fonction, nommées les *fibres de Purkinje*. Les cellules musculaires qui reçoivent cette pulsation électrique entrent alors en contraction, ce qui permet au cœur de battre. Ce phénomène est bien expliqué dans [12].

Le modèle bidomaine est un modèle mathématique utilisé pour modéliser la propagation des ondes de potentiel électrique dans le muscle cardiaque. Le muscle cardiaque est un milieu discret composé de cellules, sur lequel on peut appliquer une

homogénéisation [12], ce qui nous permet de le voir comme deux domaines, le domaine intracellulaire et le domaine extracellulaire, qui s'étendent à travers tout le myocarde. À chaque point sur le domaine de calcul, on définit donc un potentiel intracellulaire u_i et un potentiel extracellulaire u_e . Une bonne explication de la base physique et mathématique du modèle bidomaine peut être trouvée dans [8], un article par Gese-lowitz qui a introduit ce modèle dans les années 1970. Il a été prouvé que le système d'équations du modèle bidomaine possède une solution unique [6].

Il existe diverses façons de représenter les équations du modèle bidomaine. Plusieurs d'entre elles sont présentées, et leurs mérites relatifs sont discutés, dans [10]. La formulation que nous utiliserons est la suivante :

$$\frac{\partial u}{\partial t} = \frac{1}{\epsilon} f(u, v) + \nabla \cdot (\sigma_i \nabla u) + \nabla \cdot (\sigma_e \nabla u_e) \quad (1)$$

$$\nabla \cdot (\sigma_i \nabla u + (\sigma_i + \sigma_e) \nabla u_e) = 0 \quad (2)$$

$$\frac{\partial v}{\partial t} = \epsilon g(u, v) \quad (3)$$

où $u = u_i - u_e$ est le potentiel d'action ou potentiel transmembrane, σ_i et σ_e sont des tenseurs d'ordre deux qui représentent, respectivement, la conductivité électrique du tissu intracellulaire et extracellulaire dans chacune des directions spatiales du domaine, et ϵ est un paramètre relié au rapport entre le taux de repolarisation et le taux d'excitation des tissus [16]. L'équation (3) ne fait pas à proprement parler partie du modèle bidomaine, mais sert à modéliser la concentration des ions qui sont responsables de l'activation électrique des tissus. Il existe plusieurs modèles mathématiques pour les concentrations ioniques, le plus connu étant celui de Hodgkin-Huxley, introduit à l'origine pour la propagation d'influx nerveux à l'intérieur des neurones, et duquel on a ensuite dérivé des modèles adaptés au cœur, comme le modèle de Noble et celui de Luo-Rudy [12]. Nous utiliserons les équations de FitzHugh-Nagumo, qui sont une simplification de celles de Hodgkin-Huxley [12], et qui sont assez fréquemment utilisées pour modéliser l'activité ionique de la membrane cellulaire dans le myocarde (voir, par exemple, [13, 16]). Les fonctions f et g s'écrivent alors comme

$$f(u, v) = u - \frac{u^3}{3} - v, \quad (4)$$

$$g(u, v) = u + \beta - \gamma v \quad (5)$$

où γ est un paramètre contrôlant le transport des ions, et β est un paramètre relié à l'excitabilité des cellules.

En raison de la complexité du modèle bidomaine, la majorité de la recherche qui a été effectuée dans le but de le résoudre a porté sur sa résolution numérique, avec quelques exceptions comme [15]. Plus précisément, dans la majorité des cas, les efforts ont été concentrés vers sa résolution à l'aide de méthodes de différences finies explicites [10] sur des grilles structurées. L'utilisation de méthodes explicites requiert de respecter des conditions strictes pour assurer la stabilité de la méthode, et l'utilisation de grilles structurées empêche d'effectuer des calculs sur des géométries plus complexes. C'est pourquoi une partie des efforts ont migré vers l'utilisation de méthodes semi-implicites ou totalement implicites [4, 13], qui permettent de relaxer les conditions de stabilité. Des méthodes de Runge-Kutta et des méthodes de partition d'opérateur de Godunov et de Strang ont aussi été utilisées pour résoudre les équations du modèle bidomaine [19].

Déterminer quelle méthode numérique devrait être utilisée pour résoudre un système d'équations aux dérivées partielles comme celui du modèle bidomaine n'est pas trivial et est d'une grande importance. L'utilisation de la méthode des éléments finis permet d'employer des grilles non structurées, ce qui ouvre la voie à des simulations avec des géométries plus réalistes [4]. Le choix de la méthode d'intégration en temps, quant à lui, influe beaucoup sur le temps de calcul et l'utilisation de la mémoire, ainsi que sur la qualité de la solution numérique. Les méthodes explicites ont l'avantage d'être simples à coder et de ne généralement pas requérir la résolution d'un système d'équations à chaque pas de temps, ce qui a pour effet qu'un pas de temps individuel avec ces méthodes est généralement peu onéreux. L'absence d'une matrice à conserver en mémoire les rend aussi peu coûteuses en espace mémoire pour un ordinateur. Cependant, elles ont de fortes restrictions de stabilité sur la taille du pas de temps Δt , avec comme conséquence qu'elles requièrent beaucoup de pas de temps pour arriver à la solution au temps voulu. Dans le cas particulier du modèle bidomaine, elles ont aussi comme désavantage majeur et inévitable de nécessiter la résolution d'un système linéaire à chaque pas de temps, en raison de l'absence d'un terme de dérivée en temps pour la variable u_e [13]. Les méthodes implicites, elles, peuvent

supporter un pas de temps beaucoup plus grand, mais quand le système à résoudre est non linéaire, comme le modèle bidomaine avec FitzHugh-Nagumo, elles requièrent la résolution d'un système non linéaire de grande taille à chaque pas de temps, ce qui est coûteux en temps de calcul, en mémoire, et parfois aussi en temps de programmation. Les méthodes semi-implicites (aussi appelées semi-explicites ou implicites-explicites), quant à elles, sont une combinaison de méthodes explicites et implicites : elles évaluent les termes linéaires de façon implicite et les termes non linéaires de façon explicite [2, 17]. Dans le cas des problèmes de réaction-diffusion comme le modèle bidomaine, les termes de réaction sont les termes non linéaires, et les termes de diffusion sont les termes linéaires. Les caractéristiques de ces méthodes sont donc à mi-chemin entre les méthodes explicites et implicites. Leur stabilité est plus grande que celle des méthodes explicites, mais moins grande que celle des méthodes implicites, et bien qu'elles nécessitent de résoudre un système d'équations à chaque pas de temps, il s'agit seulement d'un système linéaire. Des analyses de stabilité ont été faites pour des méthodes semi-implicites appliquées à des équations de réaction-diffusion [1], mais elles ne sont pas directement applicables au modèle bidomaine qui n'est pas une équation de réaction-diffusion standard.

1.2 But de la thèse

Le but de cette thèse est de présenter différentes méthodes d'intégration en temps qui peuvent être utilisées pour résoudre le modèle bidomaine. Nous utiliserons une variété de méthodes explicites, implicites et semi-implicites, que nous montrerons au chapitre 2. Nous nous appliquerons ensuite au chapitre 3 à démontrer que les méthodes utilisées sont stables pour un pas de temps Δt qui respecte certaines conditions, et à trouver quelles sont ces conditions, tandis que le chapitre 4 montrera, pour un même cas test donné, lesquelles des méthodes considérées sont les plus appropriées. Nous conclurons ensuite en décrivant la recherche qui demeure à effectuer.

L'originalité de notre démarche tient à ceci que la plupart de la recherche effectuée jusqu'à ce jour en rapport avec le modèle bidomaine n'a pas porté sur la comparaison de différentes méthodes numériques pour en calculer la solution. Certaines exceptions,

comme [13], existent, mais encore, plusieurs des méthodes numériques que nous utilisons, comme par exemple les méthodes semi-implicites d'ordre supérieur décrites dans [2, 17], ne semblent pas avoir déjà été appliquées spécifiquement au modèle bi-domaine. De plus, le fait d'analyser la stabilité de divers schémas numériques et de comparer les conditions qui en découlent constitue aussi une contribution originale à la recherche.

Chapitre 2

Méthodes numériques

Le but de ce chapitre est de présenter les différentes méthodes numériques que nous avons étudiées dans le but de résoudre les équations du modèle bidomaine. La grande majorité de nos tests ont été effectués avec une méthode d'éléments finis de premier ordre en espace, des conditions de Neumann aux frontières du domaine et différentes méthodes de différences finies pour itérer la solution numérique en temps.

2.1 Discrétisation en espace

Nous utiliserons une méthode d'éléments finis simpliciaux de Lagrange de premier ordre. Soit \mathcal{T} une partition du domaine Ω en simplexes (segments en 1D, triangles en 2D, tétraèdres en 3D) de diamètre maximal h , possédant $N + 1$ noeuds numérotés de x_0 à x_N , et soit $\{\Phi_0^h, \dots, \Phi_N^h\}$ un ensemble de fonctions ayant pour propriétés que $\Phi_i^h(x_i) = 1$ et $\Phi_i^h(x_j) = 0$ pour $i \neq j$, et que Φ_i^h est linéaire sur chacun des éléments de \mathcal{T} . L'ensemble $\{\Phi_0^h, \dots, \Phi_N^h\}$ est une base de l'espace P_h^1 des fonctions continues sur le domaine Ω dont la restriction à chacun des éléments de \mathcal{T} est linéaire. Nous cherchons à obtenir une solution $[u, u_e, v]^T \in (\mathcal{C}([0, T]; P_h^1))^3$, où $u(t) = \sum_{i=0}^N u_i \Phi_i^h$, $u_e(t) = \sum_{i=0}^N u_{e,i} \Phi_i^h$ et $v(t) = \sum_{i=0}^N v_i \Phi_i^h \forall t \in [0, T]$, du problème variationnel suivant :

$$\int_{\Omega} u_t \Phi_j^h \, d\mathbf{x} = \frac{1}{\epsilon} \int_{\Omega} f(u, v) \Phi_j^h \, d\mathbf{x} - \int_{\Omega} (\sigma_i \nabla(u + u_e)) \cdot \nabla \Phi_j^h \, d\mathbf{x} \quad (6)$$

$$\int_{\Omega} (\sigma_i \nabla u) \cdot \nabla \Phi_j^h \, d\mathbf{x} + \int_{\Omega} ((\sigma_i + \sigma_e) \nabla u_e) \cdot \nabla \Phi_j^h \, d\mathbf{x} = 0 \quad (7)$$

$$\int_{\Omega} v_i \Phi_j^h \, d\mathbf{x} = \epsilon \int_{\Omega} g(u, v) \Phi_j^h \, d\mathbf{x} \quad (8)$$

où $f(u, v) = u - \frac{u^3}{3} - v$ et $g(u, v) = u + \beta - \gamma v$. Compte tenu de la définition de u , u_e et v mentionnée ci-haut, nous pouvons réécrire ce problème sous forme matricielle :

$$M \mathbf{u}_t = \frac{1}{\epsilon} F(\mathbf{u}, \mathbf{v}) - A_i(\mathbf{u} + \mathbf{u}_e) \quad (9)$$

$$A_i \mathbf{u} + (A_i + A_e) \mathbf{u}_e = \mathbf{0} \quad (10)$$

$$M \mathbf{v}_t = \epsilon G(\mathbf{u}, \mathbf{v}) \quad (11)$$

où $\mathbf{u} = [u_0, \dots, u_N]$, $\mathbf{u}_e = [u_{e,0}, \dots, u_{e,N}]$, $\mathbf{v} = [v_0, \dots, v_N]$, A_i et A_e sont des matrices de rigidité obtenues par l'intégration des termes de second ordre, incluant les tenseurs σ_i et σ_e , respectivement, M est une matrice de masse telle que $M(i, j) = \int_{\Omega} \Phi_i^h \Phi_j^h \, d\mathbf{x}$, F et G sont des opérateurs obtenus par l'intégration des termes d'ordre zéro. On peut représenter l'opérateur G comme

$$G(\mathbf{u}, \mathbf{v}) = M(\mathbf{u} - \gamma \mathbf{v}) + \beta \left[\int_{\Omega} \Phi_0^h \, d\mathbf{x}, \dots, \int_{\Omega} \Phi_N^h \, d\mathbf{x} \right]^T.$$

L'opérateur non linéaire F , lui, peut être représenté de différentes façons. En général, nous utiliserons une forme de concentration de masse qui nous permettra d'écrire

$$F(\mathbf{u}, \mathbf{v}) = M \left(\mathbf{u} - \frac{\mathbf{u}^3}{3} - \mathbf{v} \right)$$

où $\mathbf{u}^3 = [u_0^3, \dots, u_N^3]^T$.

2.2 Discrétisation en temps

Pour résoudre numériquement des équations de réaction-diffusion comme celles du modèle bidomaine, plusieurs auteurs suggèrent (voir par exemple [2]) d'approximer la dérivée en temps avec des méthodes implicites-explicites. Ces méthodes consistent à utiliser une méthode implicite pour les termes linéaires de diffusion, et une méthode explicite pour les termes de réaction, qui sont possiblement non linéaires. L'avantage de ces méthodes est qu'elles combinent la stabilité des méthodes implicites tout en

ne nécessitant pas la résolution d'un système d'équations non linéaires, ce qui demanderait l'utilisation d'une méthode de Newton (ou similaire) et donc la résolution de plusieurs systèmes linéaires à chaque pas en temps. Pour notre problème, nous avons utilisé une variété de méthodes implicites-explicites. Ce sont pour la plupart des méthodes de 1^{er} et de 2^e ordre, ainsi qu'une méthode de 3^e ordre. Nous avons aussi testé quelques méthodes totalement explicites et totalement implicites, par mesure de comparaison. Dans les prochaines équations, M représente la matrice de masse décrite à la section précédente et Δt représente la grandeur du pas en temps constant. Il est à noter qu'en raison de l'équation (10), qui ne possède pas de dérivée en temps, même la méthode d'Euler « explicite » n'est pas réellement explicite, puisqu'il est nécessaire de résoudre un système d'équations pour obtenir la valeur de \mathbf{u}_e^{n+1} , connaissant \mathbf{u}^{n+1} , avec l'équation suivante :

$$A_i \mathbf{u}^{n+1} + (A_i + A_e) \mathbf{u}_e^{n+1} = \mathbf{0}$$

Puisque cette équation sera présente dans chacun de nos schémas en temps, elle ne sera mentionnée qu'une seule fois.

2.2.1 Méthodes de 1^{er} ordre

Euler explicite :

$$M \frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\Delta t} = \frac{1}{\epsilon} F(\mathbf{u}^n, \mathbf{v}^n) - A_i (\mathbf{u}^n + \mathbf{u}_e^n) \quad (12)$$

$$M \frac{\mathbf{v}^{n+1} - \mathbf{v}^n}{\Delta t} = \epsilon G(\mathbf{u}^n, \mathbf{v}^n)$$

Euler implicite-explicite :

$$M \frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\Delta t} = \frac{1}{\epsilon} F(\mathbf{u}^n, \mathbf{v}^n) - A_i (\mathbf{u}^{n+1} + \mathbf{u}_e^{n+1}) \quad (13)$$

$$M \frac{\mathbf{v}^{n+1} - \mathbf{v}^n}{\Delta t} = \epsilon G(\mathbf{u}^n, \mathbf{v}^n)$$

Crank-Nicolson–Euler explicite (Crank-Nicolson, CN) :

$$M \frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\Delta t} = \frac{1}{\epsilon} F(\mathbf{u}^n, \mathbf{v}^n) - \frac{1}{2} A_i (\mathbf{u}^{n+1} + \mathbf{u}_e^{n+1}) - \frac{1}{2} A_i (\mathbf{u}^n + \mathbf{u}_e^n) \quad (14)$$

$$M \frac{\mathbf{v}^{n+1} - \mathbf{v}^n}{\Delta t} = \epsilon G(\mathbf{u}^n, \mathbf{v}^n)$$

Gear 1^{er} ordre, semi-implicite :

$$M \frac{\frac{3}{2}\mathbf{u}^{n+1} - 2\mathbf{u}^n + \frac{1}{2}\mathbf{u}^{n-1}}{\Delta t} = \frac{1}{\epsilon} F(\mathbf{u}^n, \mathbf{v}^n) - A_i(\mathbf{u}^{n+1} + \mathbf{u}_e^{n+1}) \quad (15)$$

$$M \frac{\frac{3}{2}\mathbf{v}^{n+1} - 2\mathbf{v}^n + \frac{1}{2}\mathbf{v}^{n-1}}{\Delta t} = \epsilon G(\mathbf{u}^{n+1}, \mathbf{v}^{n+1})$$

Euler implicite :

$$M \frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\Delta t} = \frac{1}{\epsilon} F(\mathbf{u}^{n+1}, \mathbf{v}^{n+1}) - A_i(\mathbf{u}^{n+1} + \mathbf{u}_e^{n+1}) \quad (16)$$

$$M \frac{\mathbf{v}^{n+1} - \mathbf{v}^n}{\Delta t} = \epsilon G(\mathbf{u}^{n+1}, \mathbf{v}^{n+1})$$

2.2.2 Méthodes de 2^e ordre

Crank-Nicolson-Adams-Bashforth (CNAB) :

$$M \frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\Delta t} = \frac{1}{\epsilon} \left(\frac{3}{2} F(\mathbf{u}^n, \mathbf{v}^n) - \frac{1}{2} F(\mathbf{u}^{n-1}, \mathbf{v}^{n-1}) \right) - \frac{1}{2} A_i(\mathbf{u}^{n+1} + \mathbf{u}_e^{n+1}) - \frac{1}{2} A_i(\mathbf{u}^n + \mathbf{u}_e^n) \quad (17)$$

$$M \frac{\mathbf{v}^{n+1} - \mathbf{v}^n}{\Delta t} = \epsilon \left(\frac{3}{2} G(\mathbf{u}^n, \mathbf{v}^n) - \frac{1}{2} G(\mathbf{u}^{n-1}, \mathbf{v}^{n-1}) \right)$$

Crank-Nicolson-Adams-Bashforth modifiée (MCNAB) :

$$M \frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\Delta t} = \frac{1}{\epsilon} \left(\frac{3}{2} F(\mathbf{u}^n, \mathbf{v}^n) - \frac{1}{2} F(\mathbf{u}^{n-1}, \mathbf{v}^{n-1}) \right) - \frac{9}{16} A_i(\mathbf{u}^{n+1} + \mathbf{u}_e^{n+1}) - \frac{3}{8} A_i(\mathbf{u}^n + \mathbf{u}_e^n) - \frac{1}{16} A_i(\mathbf{u}^{n-1} + \mathbf{u}_e^{n-1}) \quad (18)$$

$$M \frac{\mathbf{v}^{n+1} - \mathbf{v}^n}{\Delta t} = \epsilon \left(\frac{3}{2} G(\mathbf{u}^n, \mathbf{v}^n) - \frac{1}{2} G(\mathbf{u}^{n-1}, \mathbf{v}^{n-1}) \right)$$

Différentiation rétrograde semi-implicite (SBDF) de 2^e ordre :

$$M \frac{\frac{3}{2}\mathbf{u}^{n+1} - 2\mathbf{u}^n + \frac{1}{2}\mathbf{u}^{n-1}}{\Delta t} = \frac{1}{\epsilon} (2F(\mathbf{u}^n, \mathbf{v}^n) - F(\mathbf{u}^{n-1}, \mathbf{v}^{n-1})) - A_i(\mathbf{u}^{n+1} + \mathbf{u}_e^{n+1}) \quad (19)$$

$$M \frac{\frac{3}{2}\mathbf{v}^{n+1} - 2\mathbf{v}^n + \frac{1}{2}\mathbf{v}^{n-1}}{\Delta t} = \epsilon (2G(\mathbf{u}^n, \mathbf{v}^n) - G(\mathbf{u}^{n-1}, \mathbf{v}^{n-1}))$$

Gear implicite :

$$M \frac{\frac{3}{2}\mathbf{u}^{n+1} - 2\mathbf{u}^n + \frac{1}{2}\mathbf{u}^{n-1}}{\Delta t} = \frac{1}{\epsilon} F(\mathbf{u}^{n+1}, \mathbf{v}^{n+1}) - A_i(\mathbf{u}^{n+1} + \mathbf{u}_e^{n+1}) \quad (20)$$

$$M \frac{\frac{3}{2}\mathbf{v}^{n+1} - 2\mathbf{v}^n + \frac{1}{2}\mathbf{v}^{n-1}}{\Delta t} = \epsilon G(\mathbf{u}^{n+1}, \mathbf{v}^{n+1})$$

2.2.3 Méthode de 3^e ordre

Différentiation rétrograde semi-implicite (SBDF) de 3^e ordre :

$$M \frac{\frac{11}{6}\mathbf{u}^{n+1} - 3\mathbf{u}^n + \frac{3}{2}\mathbf{u}^{n-1} - \frac{1}{3}\mathbf{u}^{n-2}}{\Delta t} = \frac{1}{\epsilon} \left(3F(\mathbf{u}^n, \mathbf{v}^n) - 3F(\mathbf{u}^{n-1}, \mathbf{v}^{n-1}) + F(\mathbf{u}^{n-2}, \mathbf{v}^{n-2}) \right) - A_i(\mathbf{u}^{n+1} + \mathbf{u}_e^{n+1}) \quad (21)$$

$$M \frac{\frac{11}{6}\mathbf{v}^{n+1} - 3\mathbf{v}^n + \frac{3}{2}\mathbf{v}^{n-1} - \frac{1}{3}\mathbf{v}^{n-2}}{\Delta t} = \epsilon \left(3G(\mathbf{u}^n, \mathbf{v}^n) - 3G(\mathbf{u}^{n-1}, \mathbf{v}^{n-1}) + G(\mathbf{u}^{n-2}, \mathbf{v}^{n-2}) \right)$$

Les méthodes de premier ordre présentées ci-haut sont pour la plupart des méthodes obtenues de façon relativement *ad hoc*, à l'exception peut-être de la méthode (13) qui est mentionnée dans [2]. Les méthodes implicites-explicites de deuxième et troisième ordre, de leur côté, ont été étudiées dans [2]. Quant aux méthodes totalement implicites (16) et (20), elles sont surtout connues pour leur grande stabilité.

Tel que mentionné ci-haut, toutes ces méthodes nécessitent la résolution d'un système d'équations linéaires, ou même non linéaires dans le cas des méthodes d'Euler et de Gear implicites, à chaque pas de temps. Le choix de la meilleure méthode

pour résoudre ce système n'est pas un problème trivial. Dans le cas 1D, l'utilisation d'une méthode directe comme la factorisation LU peut être considérée pour obtenir une solution rapidement, mais le passage au problème 2D complique le travail, en raison du fait que le nombre d'inconnues augmente alors de façon quadratique par rapport à la résolution du maillage dans chaque direction spatiale. Dans ce cas, la factorisation LU peut encore être utilisée avec les maillages les moins précis, mais l'augmentation de la résolution requiert très rapidement le passage à une méthode itérative (voir [9, 18] pour plus d'information sur les méthodes itératives de solution de systèmes d'équations), en raison des contraintes de mémoire vive. Les essais que nous avons effectués nous ont convaincus que la méthode du gradient conjugué est la plus appropriée dans ce cas. Comme préconditionneur, nous avons utilisé une factorisation LU incomplète. Avec un maillage modérément précis en 2D, nous avons pu utiliser cette factorisation avec deux niveaux de remplissage, mais il est probable que l'utilisation d'un maillage spatial ayant une plus grande résolution nécessiterait l'utilisation d'un seul ou même d'aucun niveau de remplissage. Ce serait à plus forte raison vrai dans le cas du problème 3D.

Notons que la présence de la deuxième équation du modèle, qui ne possède pas de dérivée en temps, nuit beaucoup au conditionnement du système. C'est la raison pour laquelle le choix d'un bon préconditionneur est difficile à faire.

Chapitre 3

Analyse numérique des schémas

Dans ce chapitre, nous analyserons de façon théorique les schémas numériques que nous avons décrits au chapitre 2. Notre but est principalement de démontrer que ces schémas sont stables, c'est-à-dire qu'ils génèrent une solution qui demeure bornée, pour des valeurs du pas de temps Δt suffisamment petites, et de déterminer des conditions sur ce pas de temps qui assurent la stabilité de la solution.

3.1 Stabilité de la méthode semi-discrète

Nous désignerons par le terme « méthode semi-discrète » la méthode qui consiste à appliquer au problème seulement une discrétisation en espace, tel que décrit à la section 2.1. Ceci produit un système différentiel-algébrique dans lequel chaque équation contient une fonction test $\Phi \in \langle \Phi_0^h, \dots, \Phi_N^h \rangle$. Les équations (6) à (8) représentent ce système avec $\Phi = \Phi_j^h$ dans chacune des trois équations. Prenons comme fonctions test dans (6) – (8) $u(t) = \sum_{i=0}^N u_i(t)\Phi_i^h$, $u_e(t) = \sum_{i=0}^N u_{e,i}(t)\Phi_i^h$ et $v(t) = \sum_{i=0}^N v_i(t)\Phi_i^h$, respectivement ; nous obtenons donc le système suivant :

$$\begin{aligned} \int_{\Omega} u_t u \, dx &= \frac{1}{\epsilon} \int_{\Omega} f(u, v) u \, dx - \int_{\Omega} (\sigma_i \nabla(u + u_e)) \cdot \nabla u \, dx \\ \int_{\Omega} (\sigma_i \nabla u) \cdot \nabla u_e \, dx + \int_{\Omega} ((\sigma_i + \sigma_e) \nabla u_e) \cdot \nabla u_e \, dx &= 0 \\ \int_{\Omega} v_t v \, dx &= \epsilon \int_{\Omega} g(u, v) v \, dx + \int_{\Omega} S v \, dx. \end{aligned} \tag{22}$$

Notons que la fonction g a été redéfinie dans les équations précédentes : avec FitzHugh-Nagumo comme modèle ionique, nous la prenons égale à $g(u, v) = u - \gamma v$ et nous posons $S = \epsilon\beta$. Ceci nous permet de nous assurer que le graphe des fonctions f et g passe par l'origine.

Additionnant les trois équations en (22), nous obtenons

$$\begin{aligned} \int_{\Omega} [u_t u + v_t v] \, dx + \int_{\Omega} (\sigma_i \nabla(u + u_e)) \cdot \nabla(u + u_e) \, dx + \int_{\Omega} (\sigma_e \nabla u_e) \cdot \nabla u_e \, dx = \\ \frac{1}{\epsilon} \int_{\Omega} \left(u - \frac{u^3}{3} - v \right) u \, dx + \epsilon \int_{\Omega} (u - \gamma v) v \, dx + \int_{\Omega} S v \, dx. \end{aligned} \quad (23)$$

Les termes diffusifs peuvent être bornés de la façon suivante :

$$\begin{aligned} \int_{\Omega} (\sigma_i \nabla(u + u_e)) \cdot \nabla(u + u_e) \, dx + \int_{\Omega} (\sigma_e \nabla u_e) \cdot \nabla u_e \, dx \geq m_i |u + u_e|_1^2 + m_e |u_e|_1^2 \\ \geq \bar{m} (|u|_1^2 + |u_e|_1^2) \end{aligned} \quad (24)$$

où m_i et m_e désignent l'infimum de la valeur propre minimale de σ_i et σ_e sur Ω , $|\cdot|_1$ désigne la semi-norme de l'espace de Sobolev H^1 , et \bar{m} est une constante qui dépend de m_i et m_e par équivalence de semi-normes. Maintenant, en utilisant le fait que

$$\int_{\Omega} [u_t u + v_t v] \, dx = \frac{1}{2} \frac{d}{dt} [\|u\|_0^2 + \|v\|_0^2],$$

où $\|\cdot\|_0$ désigne la norme de L^2 , ainsi qu'en utilisant l'inégalité de Cauchy-Schwartz, on obtient

$$\begin{aligned} & \frac{d}{dt} [\|u\|_0^2 + \|v\|_0^2] + 2\bar{m} (|u|_1^2 + |u_e|_1^2) \\ & \leq \frac{2}{\epsilon} \left(\int_{\Omega} u^2 \, dx - \frac{1}{3} \int_{\Omega} u^4 \, dx - \int_{\Omega} uv \, dx \right) + 2\epsilon \left(\int_{\Omega} uv \, dx - \gamma \int_{\Omega} v^2 \, dx \right) \\ & \quad + 2 \int_{\Omega} S v \, dx \\ & \leq \frac{2}{\epsilon} \|u\|_0^2 - 2\epsilon\gamma \|v\|_0^2 - \frac{2}{3\epsilon} \|u\|_{L^4(\Omega)}^4 + 2 \left| \frac{1}{\epsilon} - \epsilon \right| \|u\|_0 \|v\|_0 + 2\|S\|_0 \|v\|_0 \\ & \leq \frac{2}{\epsilon} \|u\|_0^2 - 2\epsilon\gamma \|v\|_0^2 + \frac{|\frac{1}{\epsilon} - \epsilon|}{k} \|u\|_0^2 + \left| \frac{1}{\epsilon} - \epsilon \right| k \|v\|_0^2 + \frac{1}{r} \|S\|_0^2 + r \|v\|_0^2 \\ & \leq \left(\frac{2}{\epsilon} + \frac{|\frac{1}{\epsilon} - \epsilon|}{k} \right) \|u\|_0^2 + \left(\left| \frac{1}{\epsilon} - \epsilon \right| k + r - 2\epsilon\gamma \right) \|v\|_0^2 + \frac{1}{r} \|S\|_0^2 \\ & \leq C_1 [\|u\|_0^2 + \|v\|_0^2] + C_2 \|S\|_0^2 \end{aligned} \quad (25)$$

où $k, r \geq 0$, et où $C_1 = \max\{2/\epsilon + |1/\epsilon - \epsilon|/k, |1/\epsilon - \epsilon|k + r - 2\epsilon\gamma\}$, $C_2 = 1/r$. Ces constantes sont obtenues en vertu du fait que pour $a, b \in \mathbb{R}$, et pour $k > 0$, $ab \leq a^2/2k + b^2k/2$. Utilisons maintenant le lemme suivant (voir [7], page 284) :

Lemme 1 (Gronwall) Soient $\alpha \in \mathbb{R}$, $\phi \in C^1([0, T]; \mathbb{R})$ et $f \in C^0([0, T]; \mathbb{R})$ tels que $\phi_t \leq \alpha\phi + f$. Alors

$$\forall t \in [0, T], \phi(t) \leq e^{\alpha t} \phi(0) + \int_0^t e^{\alpha(t-\tau)} f(\tau) d\tau.$$

Nous appliquerons le lemme 1 à l'inéquation (25), avec $\phi = \|u\|_0^2 + \|v\|_0^2$, $\alpha = C_1$ et $f = C_2 \|S\|_0^2 - 2\bar{m}(|u|_1^2 + |u_e|_1^2)$:

$$\begin{aligned} \|u(t)\|_0^2 + \|v(t)\|_0^2 &\leq e^{C_1 t} [\|u(0)\|_0^2 + \|v(0)\|_0^2] \\ &\quad + \int_0^t e^{C_1(t-\tau)} [C_2 \|S\|_0^2 - 2\bar{m}(|u|_1^2 + |u_e|_1^2)] d\tau \\ \Rightarrow \max_{t \in [0, T]} [\|u(t)\|_0^2 + \|v(t)\|_0^2] &\leq e^{C_1 T} [\|u_0\|_0^2 + \|v_0\|_0^2] + C_2 \int_0^T e^{C_1(T-\tau)} \|S\|_0^2 d\tau \end{aligned} \quad (26)$$

$$(27)$$

où u_0 et $v_0 \in P_h^1$ sont les données initiales pour le problème semi-discret. Ce résultat nous convainc que, pour un temps final T quelconque, les variables u et v demeurent bornées dans $L^\infty(0, T; L^2(\Omega))$, c'est-à-dire en norme L^∞ sur l'espace des fonctions de $[0, T]$ à valeur dans $L^2(\Omega)$, en autant que les conditions initiales u_0 et v_0 soient des fonctions de $L^2(\Omega)$ et que le terme source S demeure aussi dans $L^2(\Omega)$ pour tout $t \in [0, T]$. L'inéquation (26) implique aussi que

$$\begin{aligned} 2\bar{m} \int_0^T e^{C_1(T-\tau)} (|u|_1^2 + |u_e|_1^2) d\tau &\leq e^{C_1 T} [\|u_0\|_0^2 + \|v_0\|_0^2] + C_2 \int_0^T e^{C_1(T-\tau)} \|S\|_0^2 d\tau \\ \Rightarrow \int_0^T e^{-C_1 \tau} (|u|_1^2 + |u_e|_1^2) d\tau &\leq \frac{1}{2\bar{m}} [\|u_0\|_0^2 + \|v_0\|_0^2] + \frac{C_2}{2\bar{m}} \int_0^T e^{-C_1 \tau} \|S\|_0^2 d\tau \\ \Rightarrow e^{-C_1 T} \int_0^T (|u|_1^2 + |u_e|_1^2) d\tau &\leq \frac{1}{2\bar{m}} [\|u_0\|_0^2 + \|v_0\|_0^2] + \frac{C_2}{2\bar{m}} \int_0^T e^{-C_1 \tau} \|S\|_0^2 d\tau \\ \Rightarrow \int_0^T (|u|_1^2 + |u_e|_1^2) d\tau &\leq \frac{e^{C_1 T}}{2\bar{m}} [\|u_0\|_0^2 + \|v_0\|_0^2] + \frac{C_2}{2\bar{m}} \int_0^T e^{C_1(T-\tau)} \|S\|_0^2 d\tau. \end{aligned} \quad (28)$$

Cette dernière équation nous permet de voir que le gradient des fonctions u et u_e demeure aussi borné en norme L^2 en temps sous les mêmes conditions que précédemment. Les équations (22) ne contiennent u_e que sous la forme de sa dérivée en espace, ce qui est consistant avec le fait que cette fonction ne peut être déterminée que modulo une constante. Il n'est donc pas possible de borner u_e plus fortement que dans la semi-norme de $H^1(\Omega)$, c'est-à-dire dans $H^1(\Omega)/\mathbb{R}$. Nous pouvons donc dire, grâce à (28), que u et u_e sont bornées dans $L^2(0, T; H^1(\Omega)/\mathbb{R})$. Maintenant, puisque nous pouvons borner u en norme L^2 , soit $\|\cdot\|_0$, et en semi-norme H^1 , soit $|\cdot|_1$, nous pouvons aussi la borner en norme H^1 , soit $\|u\|_1^2 = \|u\|_0^2 + |u|_1^2$. Nous pouvons donc conclure, grâce à (26) – (28), que la solution de la formulation semi-discrète des équations du modèle bidomaine demeure bornée dans les espaces suivants :

$$\begin{aligned} u &\in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H^1(\Omega)) \\ u_e &\in L^2(0, T; H^1(\Omega)/\mathbb{R}) \\ v &\in L^\infty(0, T; L^2(\Omega)). \end{aligned} \tag{29}$$

Notons tout de même que la preuve qui précède n'est valide que quand nous utilisons le modèle ionique de FitzHugh-Nagumo. Nous analyserons maintenant la stabilité de quelques-unes des méthodes numériques totalement discrètes que nous avons présentées au chapitre 2.

3.2 Stabilité des méthodes totalement discrètes

3.2.1 Méthodes implicites

Euler implicite

Nous prendrons la méthode d'Euler implicite (16) comme une représentante des deux méthodes implicites que nous avons utilisées. Reprenons les équations (6) à (8) et appliquons-leur la discrétisation en temps d'Euler implicite. Prenons aussi comme fonctions test Φ , $u^{n+1} = \sum_{i=0}^N u_i^{n+1} \Phi_i^h$, $u_e^{n+1} = \sum_{i=0}^N u_{e,i}^{n+1} \Phi_i^h$ et $v^{n+1} = \sum_{i=0}^N v_i^{n+1} \Phi_i^h$, respectivement, où $u_0^{n+1}, \dots, u_N^{n+1}$, $u_{e,0}^{n+1}, \dots, u_{e,N}^{n+1}$ et $v_0^{n+1}, \dots, v_N^{n+1}$ sont les degrés de

liberté de la solution numérique à l'itération $n + 1$ au noeud $i = 0, \dots, N$, tel que montré à la section 2.2. Nous obtenons alors les équations variationnelles suivantes :

$$\begin{aligned} \int_{\Omega} \left[\frac{u^{n+1} - u^n}{\Delta t} - \frac{1}{\epsilon} f(u^{n+1}, v^{n+1}) \right] u^{n+1} \, dx + \int_{\Omega} (\sigma_i \nabla(u^{n+1} + u_e^{n+1})) \cdot \nabla u^{n+1} \, dx &= 0 \\ \int_{\Omega} (\sigma_i \nabla u^{n+1}) \cdot \nabla u_e^{n+1} \, dx + \int_{\Omega} ((\sigma_i + \sigma_e) \nabla u_e^{n+1}) \cdot \nabla u_e^{n+1} \, dx &= 0 \\ \int_{\Omega} \left[\frac{v^{n+1} - v^n}{\Delta t} - \epsilon g(u^{n+1}, v^{n+1}) \right] v^{n+1} \, dx &= \int_{\Omega} S^n v^{n+1} \, dx. \end{aligned} \quad (30)$$

Additionnons maintenant ces trois équations et bornons les termes diffusifs comme à l'équation (24). Appliquons aussi l'égalité

$$2(a^{n+1} - a^n)a^{n+1} = (a^{n+1})^2 + (a^{n+1} - a^n)^2 - (a^n)^2$$

au terme de dérivée en temps. En procédant de façon similaire à la méthode utilisée en (25) pour la formulation semi-discrète, avec FitzHugh-Nagumo comme modèle de l'activité ionique, nous obtenons le résultat suivant :

$$\begin{aligned} &\|u^{n+1}\|_0^2 + \|u^{n+1} - u^n\|_0^2 - \|u^n\|_0^2 + \|v^{n+1}\|_0^2 + \|v^{n+1} - v^n\|_0^2 - \|v^n\|_0^2 \\ &\quad + 2\Delta t \bar{m} (|u^{n+1}|_1^2 + |u_e^{n+1}|_1^2) \\ &\leq \frac{2\Delta t}{\epsilon} \left(\|u^{n+1}\|_0^2 - \frac{1}{3} \|u^{n+1}\|_{L^4(\Omega)}^4 - \int_{\Omega} u^{n+1} v^{n+1} \, dx \right) \\ &\quad + 2\Delta t \epsilon \left(\int_{\Omega} u^{n+1} v^{n+1} \, dx - \gamma \|v^{n+1}\|_0^2 \right) + 2\Delta t \int_{\Omega} S^n v^{n+1} \, dx \\ &\leq \Delta t \left(\frac{2}{\epsilon} + \frac{|\frac{1}{\epsilon} - \epsilon|}{k} \right) \|u^{n+1}\|_0^2 + \Delta t \left(\left| \frac{1}{\epsilon} - \epsilon \right| k + r - 2\epsilon\gamma \right) \|v^{n+1}\|_0^2 \\ &\quad + \frac{\Delta t}{r} \|S^n\|_0^2 \\ &\leq \Delta t C_1 [\|u^{n+1}\|_0^2 + \|v^{n+1}\|_0^2] + \Delta t C_2 \|S^n\|_0^2 \end{aligned} \quad (31)$$

où C_1 et C_2 sont définies comme elles l'étaient pour la formulation semi-discrète. Désignons maintenant le pas de temps final par M ($T = M\Delta t$), et sommions l'inéquation précédente pour n allant de 0 à $m - 1$, où m est un entier entre 1 et M . Nous

obtenons alors le résultat suivant :

$$\begin{aligned}
 & \|u^m\|_0^2 + \sum_{n=0}^{m-1} \|u^{n+1} - u^n\|_0^2 + \|v^m\|_0^2 + \sum_{n=0}^{m-1} \|v^{n+1} - v^n\|_0^2 + 2\Delta t \bar{m} \sum_{n=1}^m (|u^n|_1^2 + |u_e^n|_1^2) \\
 & \leq \|u^0\|_0^2 + \|v^0\|_0^2 + \Delta t C_1 [\|u^m\|_0^2 + \|v^m\|_0^2] + \Delta t C_1 \sum_{n=1}^{m-1} [\|u^n\|_0^2 + \|v^n\|_0^2] \\
 & \quad + \Delta t C_2 \sum_{n=0}^{m-1} \|S^n\|_0^2 \\
 & \Rightarrow (1 - \Delta t C_1) [\|u^m\|_0^2 + \|v^m\|_0^2] + 2\Delta t \bar{m} \sum_{n=1}^m (|u^n|_1^2 + |u_e^n|_1^2) \\
 & \leq \|u^0\|_0^2 + \|v^0\|_0^2 + \Delta t C_1 \sum_{n=0}^{m-1} [\|u^n\|_0^2 + \|v^n\|_0^2] + \Delta t C_2 \sum_{n=0}^{m-1} \|S^n\|_0^2 \quad (32)
 \end{aligned}$$

Utilisons maintenant le lemme suivant, qui est la version discrète du lemme 1 (voir [14], page 476) :

Lemme 2 (Gronwall discret) *Soient $\{k_n\}$, $\{p_n\}$ deux suites de réels non négatifs, Φ^m une fonction discrète à valeurs réelles et g_0 un nombre réel non négatif tel que $\Phi^0 \leq g_0$. Supposons aussi que, $\forall m \geq 1$,*

$$\Phi^m \leq g_0 + \sum_{n=0}^{m-1} p_n + \sum_{n=0}^{m-1} k_n \Phi^n.$$

La propriété suivante est alors vraie :

$$\Phi^m \leq \left(g_0 + \sum_{n=0}^{m-1} p_n \right) e^{\sum_{n=0}^{m-1} k_n}.$$

Supposons que $\Delta t < 1/C_1$. Dans ce cas, l'inéquation (32) implique que

$$\begin{aligned}
 \|u^m\|_0^2 + \|v^m\|_0^2 & \leq \frac{1}{1 - \Delta t C_1} [\|u^0\|_0^2 + \|v^0\|_0^2] + \frac{\Delta t C_1}{1 - \Delta t C_1} \sum_{n=0}^{m-1} [\|u^n\|_0^2 + \|v^n\|_0^2] \\
 & \quad + \frac{\Delta t C_2}{1 - \Delta t C_1} \sum_{n=0}^{m-1} \|S^n\|_0^2. \quad (33)
 \end{aligned}$$

Posons $\Phi^m = \|u^m\|_0^2 + \|v^m\|_0^2$, $g_0 = 1/(1 - \Delta t C_1) (\|u^0\|_0^2 + \|v^0\|_0^2)$, pour tout n , $k_n = \Delta t C_1/(1 - \Delta t C_1)$ et $p_n = \Delta t C_2/(1 - \Delta t C_1) \|S^n\|_0^2$, et appliquons le lemme 2 à la

dernière inéquation. Nous obtenons

$$\begin{aligned} \|u^m\|_0^2 + \|v^m\|_0^2 &\leq \left(\frac{1}{1 - \Delta t C_1} [\|u^0\|_0^2 + \|v^0\|_0^2] + \frac{\Delta t C_2}{1 - \Delta t C_1} \sum_{n=0}^{m-1} \|S^n\|_0^2 \right) e^{\sum_{n=0}^{m-1} \frac{\Delta t C_1}{1 - \Delta t C_1}} \\ &= \left(\frac{1}{1 - \Delta t C_1} [\|u^0\|_0^2 + \|v^0\|_0^2] + \frac{C_2}{1 - \Delta t C_1} \sum_{n=0}^{m-1} \Delta t \|S^n\|_0^2 \right) e^{\frac{m \Delta t C_1}{1 - \Delta t C_1}}. \end{aligned} \quad (34)$$

L'inégalité (34) est vraie pour $m = 1, \dots, M$, et donc nous pouvons finalement dire que

$$\begin{aligned} \max_{n=1, \dots, M} \{\|u^n\|_0^2 + \|v^n\|_0^2\} &\leq \left(\frac{1}{1 - \Delta t C_1} [\|u^0\|_0^2 + \|v^0\|_0^2] \right. \\ &\quad \left. + \frac{C_2}{1 - \Delta t C_1} \sum_{n=0}^{M-1} \Delta t \|S^n\|_0^2 \right) e^{\frac{TC_1}{1 - \Delta t C_1}}, \end{aligned} \quad (35)$$

en d'autres mots, les fonctions discrètes u et v sont bornées en norme L^∞ en temps sur $L^2(\Omega)$. Nous pouvons définir des fonctions, appelées \tilde{u} et \tilde{v} et définies sur $[0, T[$, qui sont telles que \tilde{u} et \tilde{v} sont constantes par morceaux en temps et $\tilde{u}(n\Delta t) = u^n$, $\tilde{v}(n\Delta t) = v^n$ pour $n = 0, \dots, M - 1$. Ces fonctions sont telles que $\tilde{u}, \tilde{v} \in L^\infty(0, T; L^2(\Omega))$, et par léger abus de notation nous écrirons aussi $u, v \in L^\infty(0, T; L^2(\Omega))$. Cette propriété est vraie en autant que $\sum_{n=0}^{M-1} \Delta t \|S^n\|_0^2$ demeure borné. Nous pouvons définir une fonction \tilde{S} comme précédemment, et nous trouvons que la condition précédente est vraie si $\tilde{S} \in L^2(0, T; L^2(\Omega))$. Nous écrirons donc aussi que $S \in L^2(0, T; L^2(\Omega))$. Revenons maintenant à l'inéquation (32). Nous pouvons nous en servir pour borner l'autre terme au membre de gauche, de la façon suivante :

$$\begin{aligned} 2\Delta t \bar{m} \sum_{n=1}^M (|u^n|_1^2 + |u_e^n|_1^2) &\leq \|u^0\|_0^2 + \|v^0\|_0^2 + \Delta t C_1 \sum_{n=0}^{M-1} [\|u^n\|_0^2 + \|v^n\|_0^2] \\ &\quad + \Delta t C_2 \sum_{n=0}^{M-1} \|S^n\|_0^2. \end{aligned}$$

$$\begin{aligned}
 \Rightarrow \sum_{n=1}^M \Delta t (|u^n|_1^2 + |u_e^n|_1^2) &\leq \frac{1}{2\bar{m}} (\|u^0\|_0^2 + \|v^0\|_0^2) + \frac{C_1}{2\bar{m}} \sum_{n=0}^{M-1} \Delta t \max_n \{\|u^n\|_0^2 + \|v^n\|_0^2\} \\
 &\quad + \frac{C_2}{2\bar{m}} \sum_{n=0}^{M-1} \Delta t \|S^n\|_0^2 \\
 &= \frac{1}{2\bar{m}} (\|u^0\|_0^2 + \|v^0\|_0^2) + \frac{TC_1}{2\bar{m}} \max_n \{\|u^n\|_0^2 + \|v^n\|_0^2\} \\
 &\quad + \frac{C_2}{2\bar{m}} \sum_{n=0}^{M-1} \Delta t \|S^n\|_0^2. \tag{36}
 \end{aligned}$$

Le terme au membre de gauche de cette dernière inégalité n'est nul autre, par le même raisonnement impliquant \tilde{u} et \tilde{u}_e que précédemment, que la norme L^2 des fonctions discrètes u et u_e à valeur dans $H^1(\Omega)$, modulo une constante. Donc, u_e est bornée dans $L^2(0, T; H^1(\Omega)/\mathbb{R})$ et u dans $L^2(0, T; H^1(\Omega))$, tel que vu au problème semi-discret. Les fonctions u , u_e et v existent donc, une fois étendues en \tilde{u} , \tilde{u}_e et \tilde{v} , dans les espaces en (29). Ce qui distingue la méthode d'Euler implicite de la méthode semi-discrète est que dans le cas totalement discret, nous avons une contrainte sur le pas de temps Δt . Nous avons vu que pour que la solution $[u, u_e, v]$ du problème discret demeure bornée, il faut que $\Delta t < 1/C_1$. Or, $C_1 = \max\{2/\epsilon + |1/\epsilon - \epsilon|/k, |1/\epsilon - \epsilon|k + r - 2\epsilon\gamma\}$, et donc $C_1 = O(1/\epsilon)$ pour $\epsilon \ll 1$. La stabilité du schéma d'Euler implicite est donc garantie pour $\Delta t = O(\epsilon)$. Notons que cette condition est indépendante de h .

3.2.2 Méthodes semi-implicites

SBDF de 2^e ordre

Nous considérons maintenant la stabilité de la discrétisation SBDF de 2^e ordre présentée en (19) appliquée aux équations du modèle bidomaine. Nous procéderons de la même façon que pour la méthode d'Euler explicite, pour obtenir les équations

variationnelles suivantes :

$$\begin{aligned}
 & \int_{\Omega} \left[\frac{3u^{n+1} - 4u^n + u^{n-1}}{2\Delta t} - \frac{1}{\epsilon} (2f(u^n, v^n) - f(u^{n-1}, v^{n-1})) \right] u^{n+1} \, dx \\
 & \quad + \int_{\Omega} (\sigma_i \nabla(u^{n+1} + u_e^{n+1})) \cdot \nabla u^{n+1} \, dx = 0 \\
 & \int_{\Omega} (\sigma_i \nabla u^{n+1}) \cdot \nabla u_e^{n+1} \, dx + \int_{\Omega} ((\sigma_i + \sigma_e) \nabla u_e^{n+1}) \cdot \nabla u_e^{n+1} \, dx = 0 \\
 & \int_{\Omega} \left[\frac{3v^{n+1} - 4v^n + v^{n-1}}{2\Delta t} - \epsilon (2g(u^n, v^n) - g(u^{n-1}, v^{n-1})) \right] v^{n+1} \, dx \\
 & \quad = \int_{\Omega} S^n v^{n+1} \, dx. \tag{37}
 \end{aligned}$$

Le fait que les termes de réaction sont pris de façon explicite pour ce schéma nous force à les borner. Nous ferons donc l'hypothèse que les fonctions f et g satisfont au critère de Lipschitz suivant :

$$\begin{aligned}
 \|f(u, v)\|_0 &\leq L_f [\|u\|_0 + \|v\|_0] \\
 \|g(u, v)\|_0 &\leq L_g [\|u\|_0 + \|v\|_0]
 \end{aligned} \tag{38}$$

où $L_f, L_g > 0$.

Remarque 1 Les fonctions $f(u, v) = u - \frac{u^3}{3} - v$ et $g(u, v) = u - \gamma v$ du modèle ionique de FitzHugh-Nagumo ne sont pas à proprement parler lipschitziennes. Cependant, si nous faisons l'hypothèse que $u \in L^\infty(\Omega)$, avec $\|u\|_\infty \leq \mathcal{C}$ pour $\mathcal{C} > 0$ bien choisi, nous trouvons que

$$\begin{aligned}
 \|f(u, v)\|_0 &\leq \|u\|_0 + \frac{1}{3}\|u^3\|_0 + \|v\|_0 \leq \|u\|_0 + \frac{1}{3}\|u\|_\infty^2 \|u\|_0 + \|v\|_0 \\
 &\leq \left(\frac{\mathcal{C}^2}{3} + 1 \right) \|u\|_0 + \|v\|_0 \\
 &\leq L_f [\|u\|_0 + \|v\|_0],
 \end{aligned}$$

où $L_f = \frac{\mathcal{C}^2}{3} + 1$. Nous pouvons aussi borner la fonction g :

$$\|g(u, v)\|_0 \leq \|u\|_0 + \gamma \|v\|_0 \leq L_g [\|u\|_0 + \|v\|_0]$$

où $L_g = \max\{1, \gamma\}$. Cela dit, nous observons numériquement que u demeure bornée dans $L^\infty(\Omega)$, mais il n'a pas été possible jusqu'à maintenant de démontrer un tel principe du maximum pour le modèle bidomaine avec les équations de FitzHugh-Nagumo.

En additionnant ensemble les équations du système (37) et en appliquant l'égalité

$$(6a^{n+1} - 8a^n + 2a^{n-1})a^{n+1} = (a^{n+1})^2 + (2a^{n+1} - a^n)^2 + (\delta_{tt}a^{n+1})^2 - (a^n)^2 - (2a^n - a^{n-1})^2$$

où $\delta_{tt}a^{n+1} = a^{n+1} - 2a^n + a^{n-1}$ au terme de dérivée en temps, en appliquant Cauchy-Schwartz ainsi que les inégalités de Lipschitz (38) au terme de réaction et l'équation (24) aux termes de diffusion, on obtient

$$\begin{aligned} & \|u^{n+1}\|_0^2 + \|2u^{n+1} - u^n\|_0^2 + \|\delta_{tt}u^{n+1}\|_0^2 - \|u^n\|_0^2 - \|2u^n - u^{n-1}\|_0^2 + \|v^{n+1}\|_0^2 \\ & \quad + \|2v^{n+1} - v^n\|_0^2 + \|\delta_{tt}v^{n+1}\|_0^2 - \|v^n\|_0^2 - \|2v^n - v^{n-1}\|_0^2 \\ & \quad + 4\Delta t \bar{m}(|u^{n+1}|_1^2 + |u_e^{n+1}|_1^2) \\ & \leq \frac{8\Delta t L_f}{\epsilon} [\|u^n\|_0 + \|v^n\|_0] \|u^{n+1}\|_0 + \frac{4\Delta t L_f}{\epsilon} [\|u^{n-1}\|_0 + \|v^{n-1}\|_0] \|u^{n+1}\|_0 \\ & \quad + 8\Delta t \epsilon L_g [\|u^n\|_0 + \|v^n\|_0] \|v^{n+1}\|_0 + 4\Delta t \epsilon L_g [\|u^{n-1}\|_0 + \|v^{n-1}\|_0] \|v^{n+1}\|_0 \\ & \quad + 4\Delta t \|S^n\|_0 \|v^{n+1}\|_0 \\ & \leq \frac{4\Delta t L_f}{\epsilon k_1} \|u^n\|_0^2 + \frac{4\Delta t L_f}{\epsilon k_2} \|v^n\|_0^2 + \frac{4\Delta t L_f}{\epsilon} (k_1 + k_2) \|u^{n+1}\|_0^2 + \frac{2\Delta t L_f}{\epsilon k_3} \|u^{n-1}\|_0^2 \\ & \quad + \frac{2\Delta t L_f}{\epsilon k_4} \|v^{n-1}\|_0^2 + \frac{2\Delta t L_f}{\epsilon} (k_3 + k_4) \|u^{n+1}\|_0^2 + \frac{4\Delta t \epsilon L_g}{k_5} \|u^n\|_0^2 \\ & \quad + \frac{4\Delta t \epsilon L_g}{k_6} \|v^n\|_0^2 + (4\Delta t \epsilon L_g)(k_5 + k_6) \|v^{n+1}\|_0^2 + \frac{2\Delta t \epsilon L_g}{k_7} \|u^{n-1}\|_0^2 + \\ & \quad \frac{2\Delta t \epsilon L_g}{k_8} \|v^{n-1}\|_0^2 + (2\Delta t \epsilon L_g)(k_7 + k_8) \|v^{n+1}\|_0^2 + \frac{2\Delta t}{r} \|S^n\|_0^2 + 2\Delta t r \|v^{n+1}\|_0^2 \\ & \leq \frac{2\Delta t L_f}{\epsilon} (2k_1 + 2k_2 + k_3 + k_4) \|u^{n+1}\|_0^2 + 2\Delta t (\epsilon L_g (2k_5 + 2k_6 + k_7 + k_8) \\ & \quad + r) \|v^{n+1}\|_0^2 + 4\Delta t \left(\frac{L_f}{\epsilon k_1} + \frac{\epsilon L_g}{k_5} \right) \|u^n\|_0^2 + 4\Delta t \left(\frac{L_f}{\epsilon k_2} + \frac{\epsilon L_g}{k_6} \right) \|v^n\|_0^2 \\ & \quad + 2\Delta t \left(\frac{L_f}{\epsilon k_3} + \frac{\epsilon L_g}{k_7} \right) \|u^{n-1}\|_0^2 + 2\Delta t \left(\frac{L_f}{\epsilon k_4} + \frac{\epsilon L_g}{k_8} \right) \|v^{n-1}\|_0^2 + \frac{2\Delta t}{r} \|S^n\|_0^2 \end{aligned} \tag{39}$$

où $k_1, \dots, k_8 > 0$, $r > 0$. Pour simplifier la notation, posons maintenant $C_1 = \frac{2L_f}{\epsilon} (2k_1 + 2k_2 + k_3 + k_4)$, $C_2 = 2(\epsilon L_g (2k_5 + 2k_6 + k_7 + k_8) + r)$, $C_3 = 4 \left(\frac{L_f}{\epsilon k_1} + \frac{\epsilon L_g}{k_5} \right)$, $C_4 = 4 \left(\frac{L_f}{\epsilon k_2} + \frac{\epsilon L_g}{k_6} \right)$, $C_5 = 2 \left(\frac{L_f}{\epsilon k_3} + \frac{\epsilon L_g}{k_7} \right)$, $C_6 = 2 \left(\frac{L_f}{\epsilon k_4} + \frac{\epsilon L_g}{k_8} \right)$, $C_7 = 2/r$, et sommons

l'inéquation précédente pour n allant de 0 à $m - 1$, où m est un entier entre 1 et M , tout comme pour le schéma d'Euler implicite. Nous verrons apparaître des valeurs u^{-1} et v^{-1} , valeurs qui sont utilisées durant la première itération de la méthode SBDF de 2^e ordre ; nous les prendrons égales respectivement à u^0 et v^0 . Nous obtenons alors le résultat suivant :

$$\begin{aligned}
 & (1 - \Delta t C_1) \|u^m\|_0^2 + (1 - \Delta t C_2) \|v^m\|_0^2 + \|2u^m - u^{m-1}\|_0^2 + \|2v^m - v^{m-1}\|_0^2 + \\
 & \quad + \sum_{n=0}^{m-1} (\|\delta_{tt} u^{n+1}\|_0^2 + \|\delta_{tt} v^{n+1}\|_0^2) - \|u^0\|_0^2 - \|v^0\|_0^2 - \|2u^0 - u^{-1}\|_0^2 \\
 & \quad - \|2v^0 - v^{-1}\|_0^2 + 4\Delta t \bar{m} \sum_{n=1}^m (|u^n|_1^2 + |u_e^n|_1^2) \\
 & \leq \Delta t (C_1 + C_3 + C_5) \sum_{n=0}^{m-1} \|u^n\|_0^2 + \Delta t (C_2 + C_4 + C_6) \sum_{n=0}^{m-1} \|v^n\|_0^2 \\
 & \quad + \Delta t C_5 \|u^{-1}\|_0^2 + \Delta t C_6 \|v^{-1}\|_0^2 + \Delta t C_7 \sum_{n=0}^{m-1} \|S^n\|_0^2 \\
 \Rightarrow & (1 - \Delta t C) [\|u^m\|_0^2 + \|v^m\|_0^2] + 4\Delta t \bar{m} \sum_{n=1}^m (|u^n|_1^2 + |u_e^n|_1^2) \\
 & \leq 2 (\|u^0\|_0^2 + \|v^0\|_0^2) + \Delta t (C_1 + C_3 + 2C_5) \sum_{n=0}^{m-1} \|u^n\|_0^2 \\
 & \quad + \Delta t (C_2 + C_4 + 2C_6) \sum_{n=0}^{m-1} \|v^n\|_0^2 + \Delta t C_7 \sum_{n=0}^{m-1} \|S^n\|_0^2 \\
 & \leq 2 (\|u^0\|_0^2 + \|v^0\|_0^2) + \Delta t K \sum_{n=0}^{m-1} (\|u^n\|_0^2 + \|v^n\|_0^2) + \Delta t C_7 \sum_{n=0}^{m-1} \|S^n\|_0^2 \quad (40)
 \end{aligned}$$

où $C = \max\{C_1, C_2\}$, $K = \max\{C_1 + C_3 + 2C_5, C_2 + C_4 + 2C_6\}$. Choisissons Δt de telle sorte que $\Delta t < 1/C$. Nous appliquerons le lemme 2 à l'inégalité suivante, obtenue à partir de (40) :

$$\begin{aligned}
 \|u^m\|_0^2 + \|v^m\|_0^2 & \leq \frac{2}{1 - \Delta t C} (\|u^0\|_0^2 + \|v^0\|_0^2) + \frac{\Delta t K}{1 - \Delta t C} \sum_{n=0}^{m-1} (\|u^n\|_0^2 + \|v^n\|_0^2) \\
 & \quad + \frac{\Delta t C_7}{1 - \Delta t C} \sum_{n=0}^{m-1} \|S^n\|_0^2. \quad (41)
 \end{aligned}$$

Nous posons $\Phi^m = \|u^m\|_0^2 + \|v^m\|_0^2$, $g_0 = 2/(1 - \Delta t C) (\|u^0\|_0^2 + \|v^0\|_0^2)$ et $k_n = \Delta t K/(1 - \Delta t C)$, $p_n = \Delta t C_7/(1 - \Delta t C) \|S^n\|_0^2$. Par application du lemme 2, nous obtenons :

$$\begin{aligned} \|u^m\|_0^2 + \|v^m\|_0^2 &\leq \left(\frac{2}{1 - \Delta t C} [\|u^0\|_0^2 + \|v^0\|_0^2] + \frac{\Delta t C_7}{1 - \Delta t C} \sum_{n=0}^{m-1} \|S^n\|_0^2 \right) e^{\sum_{n=0}^{m-1} \frac{\Delta t K}{1 - \Delta t C}} \\ &= \left(\frac{2}{1 - \Delta t C} [\|u^0\|_0^2 + \|v^0\|_0^2] + \frac{C_7}{1 - \Delta t C} \sum_{n=0}^{m-1} \Delta t \|S^n\|_0^2 \right) e^{\frac{m \Delta t K}{1 - \Delta t C}}. \end{aligned} \quad (42)$$

De la même façon que nous avons procédé pour l'inéquation (34), l'inéquation (42) implique que

$$\begin{aligned} \max_{n=1, \dots, M} \{ \|u^n\|_0^2 + \|v^n\|_0^2 \} &\leq \left(\frac{2}{1 - \Delta t C} [\|u^0\|_0^2 + \|v^0\|_0^2] \right. \\ &\quad \left. + \frac{C_7}{1 - \Delta t C} \sum_{n=0}^{M-1} \Delta t \|S^n\|_0^2 \right) e^{TK/(1 - \Delta t C)}. \end{aligned} \quad (43)$$

Revenant à l'inégalité (40), nous pouvons écrire que

$$\begin{aligned} 4\Delta t \bar{m} \sum_{n=1}^M (|u^n|_1^2 + |u_e^n|_1^2) &\leq 2 (\|u^0\|_0^2 + \|v^0\|_0^2) + \Delta t K \sum_{n=0}^{M-1} [\|u^n\|_0^2 + \|v^n\|_0^2] \\ &\quad + \Delta t C_7 \sum_{n=0}^{M-1} \|S^n\|_0^2 \\ \Rightarrow \sum_{n=1}^M \Delta t (|u^n|_1^2 + |u_e^n|_1^2) &\leq \frac{1}{2\bar{m}} (\|u^0\|_0^2 + \|v^0\|_0^2) + \frac{TK}{4\bar{m}} \max_n \{ \|u^n\|_0^2 + \|v^n\|_0^2 \} \\ &\quad + \frac{C_7}{4\bar{m}} \sum_{n=0}^{M-1} \Delta t \|S^n\|_0^2. \end{aligned} \quad (44)$$

Ceci achève de démontrer que les fonctions u , u_e et v demeurent bornées, tel que vu en (29), pour la méthode SBDF de 2^e ordre, en autant que $\Delta t < 1/C$. Ici, C dépend de C_1 et C_2 , et $C_1 = \frac{2L_f}{\epsilon} (2k_1 + 2k_2 + k_3 + k_4)$, $C_2 = 2(\epsilon L_g (2k_5 + 2k_6 + k_7 + k_8) + r)$, donc nous pouvons dire que $C = O(L_f/\epsilon, \epsilon L_g)$. La stabilité de la méthode est garantie avec $\Delta t = O(\epsilon/L_f, \frac{1}{\epsilon L_g}) = O(\epsilon/L_f)$ lorsque $\epsilon \ll 1$ ou $L_f \gg 1$.

3.2.3 Méthodes explicites

Euler explicite

L'utilisation de méthodes explicites nécessite de poser des conditions supplémentaires sur le pas de temps Δt , en raison du fait que les termes de diffusion sont choisis explicitement. Nous considérerons ici la méthode d'Euler explicite (12). Les équations variationnelles associées à cette méthode sont les suivantes :

$$\begin{aligned} \int_{\Omega} \left[\frac{u^{n+1} - u^n}{\Delta t} - \frac{1}{\epsilon} f(u^n, v^n) \right] u^{n+1} \, d\mathbf{x} + \int_{\Omega} (\sigma_i \nabla(u^n + u_e^n)) \cdot \nabla u^{n+1} \, d\mathbf{x} &= 0 \\ \int_{\Omega} (\sigma_i \nabla u^{n+1}) \cdot \nabla u_e^{n+1} \, d\mathbf{x} + \int_{\Omega} ((\sigma_i + \sigma_e) \nabla u_e^{n+1}) \cdot \nabla u_e^{n+1} \, d\mathbf{x} &= 0 \\ \int_{\Omega} \left[\frac{v^{n+1} - v^n}{\Delta t} - \epsilon g(u^n, v^n) \right] v^{n+1} \, d\mathbf{x} &= \int_{\Omega} S^n v^{n+1} \, d\mathbf{x}. \end{aligned} \quad (45)$$

Nous procéderons comme pour les méthodes vues précédemment, en additionnant ces trois équations pour obtenir

$$\begin{aligned} &\|u^{n+1}\|_0^2 + \|u^{n+1} - u^n\|_0^2 - \|u^n\|_0^2 + \|v^{n+1}\|_0^2 + \|v^{n+1} - v^n\|_0^2 - \|v^n\|_0^2 \\ &\quad + 2\Delta t \int_{\Omega} (\sigma_i \nabla(u^n + u_e^n - u^{n+1} - u_e^{n+1})) \cdot \nabla u^{n+1} \, d\mathbf{x} \\ &\quad + 2\Delta t \int_{\Omega} (\sigma_i \nabla(u^{n+1} + u_e^{n+1})) \cdot \nabla(u^{n+1} + u_e^{n+1}) \, d\mathbf{x} \\ &\quad + 2\Delta t \int_{\Omega} (\sigma_e \nabla u_e^{n+1}) \cdot \nabla u_e^{n+1} \, d\mathbf{x} \\ &\leq \frac{2\Delta t L_f}{\epsilon} [\|u^n\|_0 + \|v^n\|_0] \|u^{n+1}\|_0 + 2\Delta t \epsilon L_g [\|u^n\|_0 + \|v^n\|_0] \|v^{n+1}\|_0 \\ &\quad + 2\Delta t \|S^n\|_0 \|v^{n+1}\|_0 \\ \Rightarrow &\|u^{n+1}\|_0^2 + \|u^{n+1} - u^n\|_0^2 + \|v^{n+1}\|_0^2 + \|v^{n+1} - v^n\|_0^2 + 2\Delta t \bar{m} (|u^{n+1}|_1^2 + |u_e^{n+1}|_1^2) \\ &\quad - 2\Delta t M_i |u^{n+1} - u^n + u_e^{n+1} - u_e^n|_1 |u^{n+1}|_1 \\ &\leq \|u^n\|_0^2 + \|v^n\|_0^2 + \frac{\Delta t L_f}{\epsilon k_1} \|u^n\|_0^2 + \frac{\Delta t L_f k_1}{\epsilon} \|u^{n+1}\|_0^2 + \frac{\Delta t L_f}{\epsilon k_2} \|v^n\|_0^2 \\ &\quad + \frac{\Delta t L_f k_2}{\epsilon} \|v^{n+1}\|_0^2 + \frac{\Delta t \epsilon L_g}{k_3} \|u^n\|_0^2 + \Delta t \epsilon L_g k_3 \|v^{n+1}\|_0^2 + \frac{\Delta t \epsilon L_g}{k_4} \|v^n\|_0^2 \\ &\quad + \Delta t \epsilon L_g k_4 \|v^{n+1}\|_0^2 + \frac{\Delta t}{r} \|S^n\|_0^2 + \Delta t r \|v^{n+1}\|_0^2 \end{aligned}$$

$$\begin{aligned}
 &\Rightarrow \|u^{n+1}\|_0^2 + \|u^{n+1} - u^n\|_0^2 + \|v^{n+1}\|_0^2 + \|v^{n+1} - v^n\|_0^2 + 2\Delta t \bar{m} (|u^{n+1}|_1^2 + |u_e^{n+1}|_1^2) \\
 &\quad - 2\Delta t M_i [|u^{n+1} - u^n|_1 + |u_e^{n+1} - u_e^n|_1] |u^{n+1}|_1 \\
 &\leq \|u^n\|_0^2 + \|v^n\|_0^2 + \frac{\Delta t L_f}{\epsilon} (k_1 + k_2) \|u^{n+1}\|_0^2 + \Delta t (\epsilon L_g (k_3 + k_4) \\
 &\quad + r) \|v^{n+1}\|_0^2 + \Delta t \left(\frac{L_f}{\epsilon k_1} + \frac{\epsilon L_g}{k_3} \right) \|u^n\|_0^2 + \Delta t \left(\frac{L_f}{\epsilon k_2} + \frac{\epsilon L_g}{k_4} \right) \|v^n\|_0^2 \\
 &\quad + \frac{\Delta t}{r} \|S^n\|_0^2
 \end{aligned} \tag{46}$$

où $k_1, \dots, k_4 > 0$, $r > 0$, et où M_i désigne le supréum de la valeur propre maximale de σ_i sur Ω . Pour simplifier la notation, nous poserons $C_1 = \frac{L_f}{\epsilon} (k_1 + k_2)$, $C_2 = \epsilon L_g (k_3 + k_4) + r$, $C_3 = \frac{L_f}{\epsilon k_1} + \frac{\epsilon L_g}{k_3}$, $C_4 = \frac{L_f}{\epsilon k_2} + \frac{\epsilon L_g}{k_4}$ et $C_5 = 1/r$. Nous devons maintenant trouver une borne supérieure pour le terme de signe négatif au membre de gauche. Reprenons la seconde équation du système (45), nous pouvons la réécrire

$$\begin{aligned}
 \int_{\Omega} ((\sigma_i + \sigma_e) \nabla u_e^{n+1}) \cdot \nabla u_e^{n+1} dx &= - \int_{\Omega} (\sigma_i \nabla u^{n+1}) \cdot \nabla u_e^{n+1} dx \\
 \Rightarrow (m_i + m_e) |u_e^{n+1}|_1^2 &\leq M_i |u^{n+1}|_1 |u_e^{n+1}|_1 \\
 \Rightarrow |u_e^{n+1}|_1 &\leq \frac{M_i}{m_i + m_e} |u^{n+1}|_1 \\
 \Rightarrow |u_e^{n+1} - u_e^n|_1 &\leq \frac{M_i}{m_i + m_e} |u^{n+1} - u^n|_1.
 \end{aligned} \tag{47}$$

Le dernier passage est obtenu par linéarité de la seconde équation de (45). Utilisons aussi le lemme suivant (voir [5], page 142) :

Lemme 3 *Soit h_K le diamètre de l'élément K de la triangulation \mathcal{T} . Supposons que l'espace d'éléments finis $V_h \subset \mathcal{C}^0(\bar{\Omega}) \cap H^1(\Omega)$ et que \mathcal{T} satisfait à la condition de régularité suivante :*

$$\exists \nu \quad \text{tel que} \quad \frac{h}{h_K} < \nu.$$

pour ν indépendant de K . Alors la proposition suivante est vraie :

$$|v_h|_1 \leq \frac{\tilde{\mathcal{C}}}{h} \|v_h\|_0 \quad \forall v_h \in V_h.$$

L'espace P_h^1 dont font partie les fonctions u^n et u^{n+1} est un sous-espace de $\mathcal{C}^0(\bar{\Omega}) \cap H^1(\Omega)$. Il ensuit donc que $|u^{n+1} - u^n|_1 \leq \frac{\tilde{\mathcal{C}}}{h} \|u^{n+1} - u^n\|_0$. Si nous appliquons le lemme 3 et l'inégalité (47) à l'inégalité (46), nous obtenons la majoration suivante :

$$\begin{aligned}
 & \|u^{n+1}\|_0^2 + \|u^{n+1} - u^n\|_0^2 + \|v^{n+1}\|_0^2 + \|v^{n+1} - v^n\|_0^2 + 2\Delta t \bar{m} (|u^{n+1}|_1^2 + |u_e^{n+1}|_1^2) \\
 & \quad - \Delta t \frac{\tilde{\mathcal{C}}^2 M_i^2 \left(1 + \frac{M_i}{m_i + m_e}\right)^2}{h^2 s} \|u^{n+1} - u^n\|_0^2 - \Delta t s |u^{n+1}|_1^2 \\
 & \leq \|u^n\|_0^2 + \|v^n\|_0^2 + \Delta t C_1 \|u^{n+1}\|_0^2 + \Delta t C_2 \|v^{n+1}\|_0^2 + \Delta t C_3 \|u^n\|_0^2 \\
 & \quad + \Delta t C_4 \|v^n\|_0^2 + \Delta t C_5 \|S^n\|_0^2 \\
 \Rightarrow & \|u^{n+1}\|_0^2 + \left(1 - \Delta t \frac{\tilde{\mathcal{C}}^2 M_i^2 \left(1 + \frac{M_i}{m_i + m_e}\right)^2}{h^2 s}\right) \|u^{n+1} - u^n\|_0^2 + \|v^{n+1}\|_0^2 \\
 & \quad + \Delta t (2\bar{m} - s) (|u^{n+1}|_1^2 + |u_e^{n+1}|_1^2) \leq \|u^n\|_0^2 + \|v^n\|_0^2 + \Delta t C_1 \|u^{n+1}\|_0^2 \\
 & \quad + \Delta t C_2 \|v^{n+1}\|_0^2 + \Delta t C_3 \|u^n\|_0^2 + \Delta t C_4 \|v^n\|_0^2 + \Delta t C_5 \|S^n\|_0^2. \quad (48)
 \end{aligned}$$

Il suffit de poser comme conditions que $2\bar{m} - s > 0$, donc que $s < 2\bar{m}$, et que $1 - \Delta t \frac{\tilde{\mathcal{C}}^2 M_i^2 \left(1 + \frac{M_i}{m_i + m_e}\right)^2}{h^2 s} \geq 0$, ce qui implique que

$$\Delta t \leq \frac{h^2 s}{\tilde{\mathcal{C}}^2 M_i^2 \left(1 + \frac{M_i}{m_i + m_e}\right)^2} < \frac{2h^2 \bar{m}}{\tilde{\mathcal{C}}^2 M_i^2 \left(1 + \frac{M_i}{m_i + m_e}\right)^2}. \quad (49)$$

Posons $k = 2\bar{m} - s$, $C = \max\{C_1, C_2\}$, $K = \max\{C_1 + C_3, C_2 + C_4\}$ et sommons l'inéquation (48) pour $n = 0, \dots, m-1$:

$$\begin{aligned}
 & (1 - \Delta t C) [\|u^m\|_0^2 + \|v^m\|_0^2] + \Delta t k \sum_{n=1}^m (|u^n|_1^2 + |u_e^n|_1^2) \\
 & \leq \|u^0\|_0^2 + \|v^0\|_0^2 + \Delta t K \sum_{n=0}^{m-1} (\|u^n\|_0^2 + \|v^n\|_0^2) + \Delta t C_5 \sum_{n=0}^{m-1} \|S^n\|_0^2. \quad (50)
 \end{aligned}$$

Si $\Delta t < 1/C$, nous pouvons appliquer le lemme 2 à l'inégalité (50) de la même façon que nous l'avions appliqué aux autres méthodes, en posant $\Phi^m = \|u^m\|_0^2 + \|v^m\|_0^2$, $g_0 = 1/(1 - \Delta t C) (\|u^0\|_0^2 + \|v^0\|_0^2)$, $k_n = \Delta t K/(1 - \Delta t C)$ et $p_n = \Delta t C_5/(1 - \Delta t C) \|S^n\|_0^2$,

pour obtenir

$$\begin{aligned} \max_{n=1,\dots,M} \{ \|u^n\|_0^2 + \|v^n\|_0^2 \} &\leq \left(\frac{1}{1 - \Delta t C} [\|u^0\|_0^2 + \|v^0\|_0^2] \right. \\ &\quad \left. + \frac{C_5}{1 - \Delta t C} \sum_{n=0}^{M-1} \Delta t \|S^n\|_0^2 \right) e^{TK/(1-\Delta t C)}, \end{aligned} \quad (51)$$

$$\begin{aligned} \sum_{n=1}^M \Delta t (|u^n|_1^2 + |u_e^n|_1^2) &\leq \frac{1}{k} (\|u^0\|_0^2 + \|v^0\|_0^2) + \frac{TK}{k} \max_n \{ \|u^n\|_0^2 + \|v^n\|_0^2 \} \\ &\quad + \frac{C_5}{k} \sum_{n=0}^{M-1} \Delta t \|S^n\|_0^2. \end{aligned} \quad (52)$$

Ceci nous permet de conclure que les fonctions u , u_e et v demeurent bornées dans les espaces décrits en (29), si les conditions que nous avons imposées sur Δt sont respectées. Ces conditions sont que $\Delta t < 1/C$ où $C = O(L_f/\epsilon, \epsilon L_g)$, et la condition présentée en (49). Puisque Δt dépend entre autres de \bar{m} et $1/\tilde{\mathcal{C}}^2$, nous essaierons de faire apparaître une dépendance explicite pour ces deux paramètres.

Nous avons vu en (24) que

$$m_i |u + u_e|_1^2 + m_e |u_e|_1^2 \geq \bar{m} (|u|_1^2 + |u_e|_1^2).$$

Plaçons-nous dans le cadre général d'un espace vectoriel V doté d'un produit scalaire $\langle \cdot, \cdot \rangle$ qui induit une norme $|\cdot|$. L'espace $H^1(\Omega)/\mathbb{R}$ doté du produit scalaire

$$\langle u, v \rangle_1 = \int_{\Omega} \nabla u \cdot \nabla v \, dx$$

est un tel espace V . Dans V , la majoration (24) peut être réécrite

$$\begin{aligned} m_i |x + y|^2 + m_e |y|^2 &= m_i |x|^2 + 2m_i \langle x, y \rangle + (m_i + m_e) |y|^2 \\ &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} m_i & m_i \\ m_i & m_i + m_e \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &\geq \lambda_{\min} (|x|^2 + |y|^2) \end{aligned} \quad (53)$$

où $\lambda_{\min} > 0$ est la valeur propre minimale de la matrice $\begin{bmatrix} m_i & m_i \\ m_i & m_i + m_e \end{bmatrix}$. Nous pouvons donc prendre $\bar{m} = \lambda_{\min}$. Un calcul simple peut alors nous convaincre que

$$\bar{m} = \frac{2m_i + m_e - \sqrt{4m_i^2 + m_e^2}}{2}. \quad (54)$$

Considérons maintenant la constante $\tilde{\mathcal{C}}$ qui apparaît dans le lemme 3, plus précisément dans l'inégalité

$$|v_h|_1 \leq \frac{\tilde{\mathcal{C}}}{h} \|v_h\|_0.$$

Nous commencerons par évaluer $\tilde{\mathcal{C}}$ dans le cas 1D. Soit K l'élément linéaire sur l'intervalle $[0, h]$, et soit v_h une fonction de P_h^1 sur cet élément telle que $v_h(0) = a$ et $v_h(h) = b$, donc $v_h(x) = a + (b - a)x/h$ sur K . Nous calculerons $\|v_h\|_0$ et $|v_h|_0$ sur K :

$$\begin{aligned} \|v_h\|_0 &= \left(\int_K v_h^2 dx \right)^{1/2} \\ &= \left(\int_0^h \left(a + \frac{b-a}{h} x \right)^2 dx \right)^{1/2} \\ &= \left(\frac{h}{3} (a^2 + ab + b^2) \right)^{1/2}, \end{aligned} \tag{55}$$

$$\begin{aligned} |v_h|_1 &= \left(\int_K |\nabla v_h|^2 dx \right)^{1/2} \\ &= \left(\int_0^h \left(\frac{b-a}{h} \right)^2 dx \right)^{1/2} \\ &= \left(\frac{(a-b)^2}{h} \right)^{1/2}. \end{aligned} \tag{56}$$

Puisque $|v_h|_1 \leq \frac{\tilde{\mathcal{C}}}{h} \|v_h\|_0$, nous choisirons $\tilde{\mathcal{C}}$ tel que

$$\begin{aligned} \tilde{\mathcal{C}} &\geq h \frac{|v_h|_1}{\|v_h\|_0} \\ &= h \left(\frac{(a-b)^2/h}{h(a^2 + ab + b^2)/3} \right)^{1/2} \\ &= \left(\frac{3(a-b)^2}{a^2 + ab + b^2} \right)^{1/2}. \end{aligned} \tag{57}$$

Nous pouvons donc choisir $\tilde{\mathcal{C}}$ comme une borne supérieure pour la fonction rationnelle $(3(a-b)^2/(a^2 + ab + b^2))^{1/2}$. On découvre que $2\sqrt{3}$ est une borne supérieure pour

cette fonction. Le calcul suivant en fait la preuve :

$$\begin{aligned}
 3(a+b)^2 \geq 0 &\Rightarrow 3a^2 + 6ab + 3b^2 \geq 0 \Rightarrow a^2 - 2ab + b^2 \leq 4a^2 + 4ab + 4b^2 \\
 &\Rightarrow \frac{a^2 - 2ab + b^2}{a^2 + ab + b^2} \leq 4 \\
 &\Rightarrow \left(\frac{3(a-b)^2}{a^2 + ab + b^2} \right)^{1/2} \leq 2\sqrt{3}. \quad (58)
 \end{aligned}$$

Le terme $a^2 + ab + b^2$ est non négatif et n'égal 0 que lorsque $a = b = 0$, et dans ce cas $|v_h|_0 = \|v_h\|_0 = 0$. On note que la borne supérieure calculée en (58) est en fait la borne supérieure optimale, puisqu'elle est atteinte quand $b = -a$.

Considérons maintenant le cas 2D. Soit K l'élément linéaire triangulaire dont les noeuds sont situés aux points $(0, 0)$, $(0, h)$ et $(h/2, h/2)$. Il ne s'agit pas là du choix le plus général, puisque $\tilde{\mathcal{E}}$ dépendra de la forme de l'élément K , mais ce choix est néanmoins représentatif des maillages avec lesquels nous travaillerons au chapitre 4. Soit v_h une fonction de P_h^1 sur cet élément telle que $v_h(0, 0) = a$, $v_h(0, h) = b$ et $v_h(h/2, h/2) = c$. On calcule que $v_h(x, y) = a + (2c - (a + b))x/h + (b - a)y/h$ sur K . Calculons maintenant $\|v_h\|_0$ et $|v_h|_0$.

$$\begin{aligned}
 \|v_h\|_0 &= \left(\int_K v_h^2 \, d\mathbf{x} \right)^{1/2} \\
 &= \left(\int_0^{h/2} \int_x^{h-x} \left(a + \frac{2c - (a+b)}{h}x + \frac{b-a}{h}y \right)^2 \, dy \, dx \right)^{1/2} \\
 &= \left(\frac{h^2}{24} (a^2 + b^2 + c^2 + ab + ac + bc) \right)^{1/2}, \quad (59)
 \end{aligned}$$

$$\begin{aligned}
 |v_h|_1 &= \left(\int_K |\nabla v_h|^2 \, d\mathbf{x} \right)^{1/2} \\
 &= \left(\int_0^{h/2} \int_x^{h-x} \left(\left(\frac{2c - (a+b)}{h} \right)^2 + \left(\frac{b-a}{h} \right)^2 \right) \, dy \, dx \right)^{1/2} \\
 &= \left(\frac{a^2}{2} + \frac{b^2}{2} + c^2 - ac - bc \right)^{1/2}. \quad (60)
 \end{aligned}$$

Nous devons choisir $\tilde{\mathcal{E}}$ telle que

$$\begin{aligned}
 \tilde{\mathcal{E}} &\geq h \frac{|v_h|_1}{\|v_h\|_0} \\
 &= h \left(\frac{\left(\frac{a^2}{2} + \frac{b^2}{2} + c^2 - ac - bc \right)}{\frac{h^2}{24} (a^2 + b^2 + c^2 + ab + ac + bc)} \right)^{1/2} \\
 &= \left(\frac{12a^2 + 12b^2 + 24c^2 - 24ac - 24bc}{a^2 + b^2 + c^2 + ab + ac + bc} \right)^{1/2}. \tag{61}
 \end{aligned}$$

La valeur optimale pour $\tilde{\mathcal{E}}$ dans ce cas-ci est en fait $6\sqrt{2}$. On peut le prouver ainsi :

$$\begin{aligned}
 \left(\frac{12a^2 + 12b^2 + 24c^2 - 24ac - 24bc}{a^2 + b^2 + c^2 + ab + ac + bc} \right)^{1/2} &\leq 6\sqrt{2} \\
 \Leftrightarrow \frac{12a^2 + 12b^2 + 24c^2 - 24ac - 24bc}{a^2 + b^2 + c^2 + ab + ac + bc} &\leq 72 \\
 \Leftrightarrow 5a^2 + 5b^2 + 4c^2 + 6ab + 8ac + 8bc &\geq 0 \tag{62}
 \end{aligned}$$

si $a^2 + b^2 + c^2 + ab + ac + bc \geq 0$, ce qui est vrai pour tous $a, b, c \in \mathbb{R}$. Le terme à gauche du signe d'inégalité dans (62) est un polynôme de deuxième degré que nous pouvons réécrire comme la forme quadratique suivante :

$$\begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} 5 & 3 & 4 \\ 3 & 5 & 4 \\ 4 & 4 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}. \tag{63}$$

Or, la matrice symétrique de la forme quadratique (63) a comme valeurs propres 0, 2 et 12 et est donc semi-définie positive. Le polynôme $5a^2 + 5b^2 + 4c^2 + 6ab + 8ac + 8bc$ est donc toujours non négatif. Notons qu'il est en fait égal à 0 sur l'espace non trivial $\langle [1, 1, -2]^T \rangle$; en d'autres mots, $6\sqrt{2}$ est la valeur optimale de $\tilde{\mathcal{E}}$ pour l'élément K bidimensionnel que nous avons choisi.

Nous avons vu en (49) que

$$\Delta t < \frac{2h^2 \bar{m}}{\tilde{\mathcal{E}}^2 M_i^2 \left(1 + \frac{M_i}{m_i + m_e} \right)^2}.$$

Nous considérerons trois cas. Premièrement, supposons que $m_i \approx m_e$. Nous pouvons en fait prendre $m_i = m_e$ et $M_i = \lambda m_i$ pour un certain $\lambda > 1$. Dans ce cas, nous arrivons à

$$\begin{aligned} \Delta t &< \frac{(3 - \sqrt{5})h^2}{\tilde{\mathcal{C}}^2 \lambda^2 m_i \left(1 + \lambda + \frac{\lambda^2}{4}\right)} \\ \Rightarrow \Delta t &< \frac{\tilde{K} h^2}{\lambda^4 m_i} \\ &= \frac{\tilde{K} m_i^3}{M_i^4} h^2 \end{aligned}$$

où \tilde{K} est une constante qui dépend de $1/\tilde{\mathcal{C}}^2$. Dans ce cas, nous avons donc $\Delta t = O\left(\epsilon/L_f, \frac{1}{\epsilon L_g}, \frac{m_i^3}{M_i^4} h^2\right)$. Quand h devient petit, il s'agit du dernier terme qui domine. Deuxièmement, nous prendrons des valeurs qui se rapprochent des valeurs physiques du cœur, par exemple celles qui se trouvent dans [11], soient $m_i = 0,026\ 3$, $m_e = 0,108\ 7$ et $M_i = 0,263$. Alors, en 1D nous aurions $\Delta t < 0,002\ 56h^2$, et en 2D nous aurions $\Delta t < 0,000\ 167h^2$. Ce pas de temps est excessivement petit, ce qui rend une méthode explicite moins efficace en raison du grand nombre de pas de temps requis. Finissons en supposant que nous changeons la conductivité du domaine mais en la gardant proportionnelle à ce qu'elle était avant, c'est-à-dire que nous prendrons comme tenseurs de conductivité $\mu \sigma_i$ et $\mu \sigma_e$. Notre condition de stabilité sera alors

$$\begin{aligned} \Delta t &< \frac{h^2 \left(2\mu m_i + \mu m_e - \sqrt{4\mu^2 m_i^2 + \mu^2 m_e^2}\right)}{\tilde{\mathcal{C}}^2 \mu^2 M_i^2 \left(1 + \frac{\mu M_i}{\mu m_i + \mu m_e}\right)^2} \\ &= \frac{1}{\mu} \frac{h^2 \left(2m_i + m_e - \sqrt{4m_i^2 + m_e^2}\right)}{\tilde{\mathcal{C}}^2 M_i^2 \left(1 + \frac{M_i}{m_i + m_e}\right)^2}, \end{aligned}$$

en d'autres mots, augmenter la conductivité du domaine aura pour effet de diminuer la stabilité de la méthode d'Euler explicite.

Nous concluerons avec un tableau récapitulatif qui montre les différentes méthodes d'intégration en temps que nous avons analysées avec les conditions de stabilité sur le pas de temps Δt qui nous sont apparues. Le tableau 1 suppose que nous prenons $\epsilon \ll 1$. Nous n'y avons pas inclus la méthode semi-discrète, puisqu'elle ne possède pas de pas de temps Δt , n'étant pas une méthode numérique d'intégration en temps.

TAB. 1 – Conditions de stabilité sur Δt et sur le modèle utilisé pour différentes méthodes

Méthode	Δt	Modèle
Euler explicite	$O\left(\frac{\epsilon}{L_f}, \frac{m_i^3}{M_i^4} h^2\right)$	f, g lipschitziennes
SBDF 2 ^e ordre	$O\left(\frac{\epsilon}{L_f}\right)$	f, g lipschitziennes
Euler implicite	$O(\epsilon)$	FitzHugh-Nagumo

Chapitre 4

Résultats numériques

Ce chapitre contient le résultat de nos simulations destinées à résoudre numériquement les équations du modèle bidomaine. Il est divisé en deux parties : la première concerne le modèle 1D, et la seconde, le modèle 2D.

4.1 Comparaison des schémas en 1D

Nous avons utilisé pour le problème 1D un code d'éléments finis simple en Matlab. Les systèmes d'équations linéaires sont résolus avec la méthode de factorisation LU incluse avec Matlab. La facilité d'implémentation d'un code 1D, ainsi que la facilité d'utilisation de Matlab nous ont permis de tester différents schémas de différences finies.

4.1.1 Problème

Les équations du modèle bidomaine 1D peuvent s'écrire ainsi :

$$\frac{\partial u}{\partial t} = \frac{1}{\epsilon} f(u, v) + \frac{\partial}{\partial x} \left(\sigma_i \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial x} \left(\sigma_e \frac{\partial u_e}{\partial x} \right) \quad (64)$$

$$\frac{\partial}{\partial x} \left(\sigma_i \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial x} \left((\sigma_i + \sigma_e) \frac{\partial u_e}{\partial x} \right) = 0 \quad (65)$$

$$\frac{\partial u}{\partial t} = \epsilon g(u, v) \quad (66)$$

où $f(u, v) = u - \frac{u^3}{3} - v$ et $g(u, v) = u + \beta - \gamma v$, et des conditions aux limites de Neumann, c'est-à-dire $\partial_x u(0) = \partial_x u(L) = 0$ et $\partial_x u_e(0) = \partial_x u_e(L) = 0$. Nous avons déjà présenté aux sections 2.1 et 2.2 la manière selon laquelle nous discrétiserons ce problème. Nous avons déjà noté à la section 3.1 que u_e n'apparaît dans le problème que sous la forme de sa dérivée en espace, ce qui signifie que cette fonction n'est déterminée que modulo une constante. Pour éviter d'avoir à résoudre un système singulier, nous ajouterons la condition $u_e(0) = 0$. D'autres façons de pallier cette difficulté existent, comme par exemple de considérer la solution u_e comme existant dans l'espace quotient P_h^1/\mathbb{R} ; certaines sont mentionnées dans [3].

Tous nos tests ont été effectués sur le domaine $[0, L]$ avec quelques valeurs de L . Nous avons pris comme valeurs des paramètres $\epsilon = 0,1$, $\beta = 1,0$ et $\gamma = 0,5$, et comme valeurs des tenseurs de conductivité $\sigma_i = \sigma_e = 1,0$. Comme condition initiale pour le problème, nous avons choisi de prendre u et v constantes à la solution à l'équilibre du système $f(u, v) = 0$ et $g(u, v) = 0$, ainsi que u_e constante à 0, sauf sur l'intervalle $[0, \frac{1}{20}L]$ où nous fixons u à une valeur sur-critique, soit $u(x) = 2$. Ceci génère une pulsation qui, après avoir pris sa forme, se propage ensuite à travers le domaine à vitesse constante. La figure 1 présente l'évolution d'une solution de référence pour ce problème au cours du temps ($L = 70$, temps final $T = 30$).

4.1.2 Résultats

En l'absence d'une solution exacte du problème, nous avons obtenu une solution numérique très précise (résolution du maillage $h = 0,007$, pas en temps $\Delta t = 10^{-4}$) avec le schéma SBDF de 2^e ordre (19) que nous utiliserons comme point de comparaison avec nos autres méthodes numériques. Il convient de mentionner qu'il est relativement simple d'obtenir une solution numérique qui ressemble qualitativement à la solution modèle, mais pour laquelle l'erreur numérique est assez importante. Le problème est généralement qu'une résolution insuffisante du maillage spatial, ou un pas en temps trop grand, conduit à une onde mobile dont la vitesse de propagation est incorrecte. C'est pourquoi nous considérerons le pourcentage d'erreur entre la vitesse de l'onde modèle et la vitesse de l'onde numérique comme une bonne mesure de la

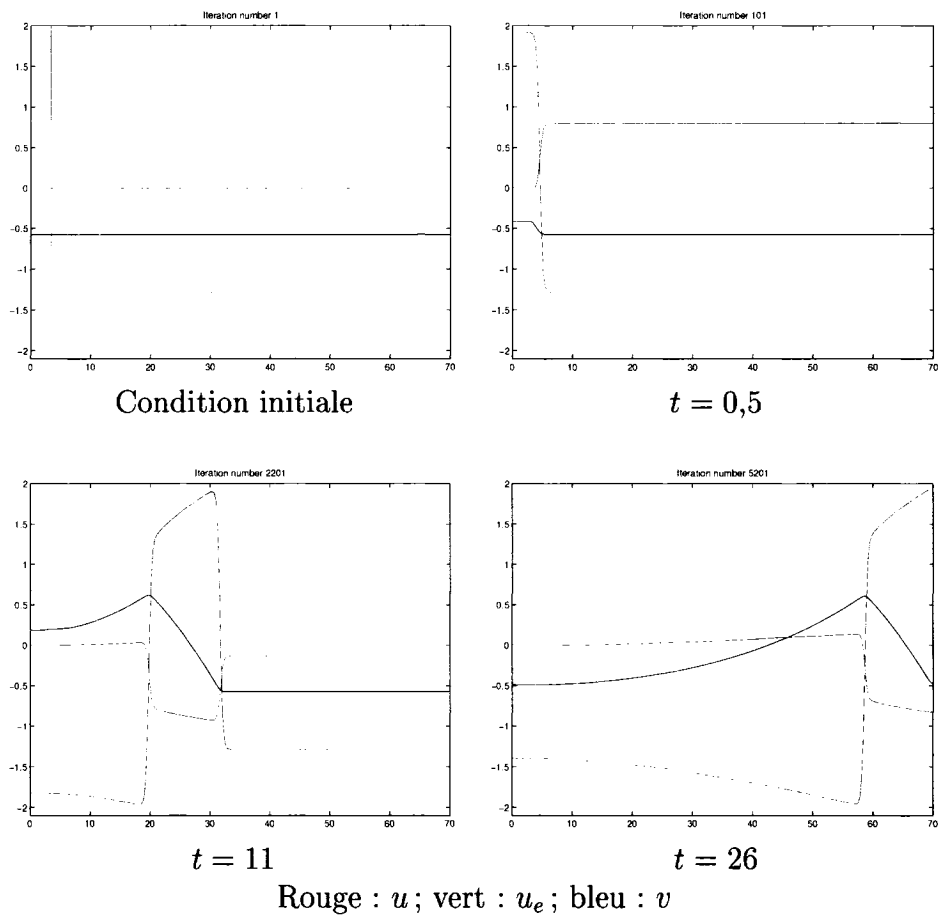


FIG. 1 – Solution des équations du modèle bidomaine 1D à différents temps

qualité de la solution, tout comme la norme L^2 ou L^∞ de l'erreur numérique.

Le tableau 2 donne la valeur du pas de temps critique (Δt_c), c'est-à-dire la valeur du pas de temps au-delà de laquelle la solution numérique devient non bornée, pour toutes nos méthodes et différentes résolutions spatiales dignes d'intérêt. Pour les tests présentés dans ce tableau, $L = 70$ et $T = 30$. Il est à noter que fréquemment, pour des valeurs de Δt légèrement inférieures à Δt_c , des oscillations apparaissent dans la solution mais tendent à se lisser d'elles-mêmes au fil du temps, à la différence de ce qui se produit avec des équations aux dérivées partielles linéaires comme l'équation de la chaleur, pour lesquelles des oscillations n'apparaissent que quand la solution est résolument instable, et croissent ensuite de façon polynomiale. Nous croyons que le terme de réaction non linéaire est responsable de cet effet, et qu'il s'agit en fait peut-être d'une perte de monotonie plus que d'une perte de stabilité.

TAB. 2 – Taille de Δt_c pour différentes méthodes numériques

Méthodes	$N = 500$	$N = 1000$	$N = 2000$
Euler explicite	0,003 206	0,000 813	0,000 204
Euler implicite-explicite	0,139 5	0,142 2	0,141 5
Crank-Nicolson-Euler explicite	0,129 9	0,104 4	0,083 86
Gear 1 ^{er} ordre, semi-implicite	0,159 4	0,159 4	0,159 4
Euler implicite	> 0,2	> 0,2	> 0,2
CNAB	0,067 72	0,049 51	0,040 27
MCNAB	0,075 38	0,075	0,075
SBDF de 2 ^e ordre	0,093 75	0,094 04	0,094 94
Gear implicite	> 0,25	> 0,25	> 0,25
SBDF de 3 ^e ordre	0,068 34	0,070 26	0,071 26

Il est évident, à la lecture du tableau 2, que la dépendance de Δt_c par rapport à la résolution du maillage n'est pas la même selon la méthode utilisée. La stabilité de la plupart des méthodes (en l'occurrence, Euler implicite, Gear 1^{er} ordre, MCNAB, ainsi que SBDF de 2^e et 3^e ordre) semble en fait très peu dépendre de la résolution du maillage en espace. Le pas de temps critique pour les méthodes de Crank-Nicolson-Euler explicite et CNAB, quant à lui, présente peut-être une légère dépendance sur la finesse du maillage en espace, mais une lecture du tableau montre que même pour

ces méthodes, l'ordre de cette dépendance en h est sous-linéaire. La stabilité de la méthode d'Euler explicite, elle, semble plutôt dépendre de h^2 , ce qui appuie l'analyse théorique que nous avons faite à la sous-section 3.2.3. Il est cependant important de noter que nos tests ne nous donnent pas d'information complète sur la dépendance asymptotique de Δt_c quand h tend vers 0. Par exemple, dans le cas de CNAB, la dépendance en apparence sous-linéaire n'apparaît que pour un N suffisamment grand (quand $N = 100$, $\Delta t_c = 0,072\ 64$ et quand $N = 200$, $\Delta t_c = 0,074\ 26$) et il est possible que pour des valeurs de N encore plus grandes, cette dépendance finisse par disparaître.

Tel que mentionné précédemment, nous avons utilisé comme mesures de la convergence de nos solutions numériques vers la solution modèle l'erreur L^2 et L^∞ entre ces deux solutions, ainsi que l'erreur sur la vitesse de l'onde de potentiel. Nous avons calculé la vitesse de l'onde pour notre solution modèle comme étant $c_{\text{modèle}} = 2,577\ 444$. Le tableau 3 présente l'erreur relative signée entre la vitesse de l'onde de potentiel numérique (c_{num}), définie comme étant la vitesse moyenne à laquelle se déplace la courbe de niveau $u = 1$ sur l'intervalle $[25, 50]$, et $c_{\text{modèle}}$ pour différentes méthodes, c'est-à-dire $E = (c_{\text{num}} - c_{\text{modèle}})/c_{\text{modèle}}$. Ici aussi, $L = 70$ et $T = 30$. N désigne le nombre d'éléments et Δt la taille du pas de temps. Le tableau présentant c_{num} pour la méthode d'Euler explicite (12) est différent des autres en raison du fait que la dépendance de Δt_c sur h est très forte, et qu'avec des valeurs de N plus grandes il faut prendre un pas de temps très petit, beaucoup plus petit que ce qu'il est nécessaire de prendre pour des valeurs de N plus petites. Nous avons donc décidé de donner la vitesse de l'onde obtenue pour un nombre de pas de temps nt approximativement égal au pas de temps nécessaire pour assurer la stabilité et la monotonie de la solution, puis pour un pas de temps 1,5 et 2 fois supérieur. La figure 2, quant à elle, montre la position de l'onde pour u à différents temps pour différentes méthodes. On voit sur les deux figures du bas que l'erreur de phase entre la solution modèle et la solution numérique croît avec le temps, ce qui justifie notre utilisation de l'erreur sur la vitesse de l'onde comme mesure de l'erreur numérique. La figure 3, finalement, montre la progression en temps de l'erreur L^2 entre la solution numérique et la solution modèle pour un N fixé à 500. Les deux images du haut comparent les méthodes

de 1^{er} ordre entre elles et avec les méthodes de 2^e ordre. On note que les tests pour les méthodes de Gear implicite et d'Euler explicite utilisent un nombre de pas de temps très différent de celui choisi pour les autres méthodes. La raison est que la méthode d'Euler explicite nécessite beaucoup plus de pas de temps que les autres, étant moins stable ; ces pas de temps sont cependant aussi moins coûteux. L'inverse est vrai pour la méthode de Gear implicite : elle est plus stable, et nécessite donc moins de pas de temps, mais chacun d'entre eux demande plus d'effort. Les deux images du bas comparant entre elles les méthodes de 2^e et de 3^e ordre pour deux valeurs de Δt , et nous nous en servons pour arriver à nos conclusions sur quelles méthodes sont les plus appropriées pour notre problème. On remarque que dans la majorité des cas, après une courte période d'ajustement, l'erreur L^2 augmente linéairement par rapport au temps, mais avec différentes pentes.

Nos résultats nous indiquent que des méthodes d'ordre supérieur à 1 sont absolument nécessaires pour ce problème. La figure 3 montre bien la différence entre une méthode de premier ordre, même une des meilleures comme Crank-Nicolson–Euler explicite, et une méthode d'ordre supérieur. On peut voir à la figure 2 que la différence est principalement due au fait que les méthodes de premier ordre ne parviennent pas à produire une solution dont la vitesse est suffisamment proche de la vitesse exacte, du moins pas pour un pas en temps de l'ordre de $\Delta t = 0,02$. Les données présentées dans le tableau 3 le confirment : quand $N = 500$, nous arrivons à obtenir une erreur de 1 % avec la solution modèle avec un pas de temps $\Delta t = 0,02$ pour SBDF de 2^e ordre, $\Delta t = 0,03$ pour SBDF de 3^e ordre, mais seulement $\Delta t = 0,005$ pour Crank-Nicolson. Ces résultats restent les mêmes quand la résolution du maillage est plus grande : pour un N fixé à 2000, aucune des méthodes implicites-explicites de 1^{er} ordre ne permet d'obtenir une erreur inférieure à 1 % même avec un pas de temps $\Delta t = 0,005$, tandis qu'avec la méthode CNAB, nous pouvons obtenir une telle erreur avec $\Delta t = 0,03$. Les figures 2 et 3 semblent indiquer que, malgré qu'elle soit seulement de premier ordre, la méthode d'Euler explicite mérite notre considération. En effet, elle donne des résultats proches de la solution modèle, elle est simple à coder et, malgré qu'elle ne soit pas entièrement « explicite », tel que mentionné à la section 2.2, elle est néanmoins relativement peu coûteuse. Cependant, comme nous l'avons déjà

TAB. 3 – Erreur relative (%) entre $c_{\text{modèle}}$ et c_{num}

Crank-Nicolson-Adams-Bashforth					
	$N = 100$	$N = 200$	$N = 500$	$N = 1000$	$N = 2000$
$\Delta t = 0,04$	9,702	1,567	-1,041	-1,591	-1,605
$\Delta t = 0,02$	10,43	2,406	-0,029 37	-0,176 5	-0,381 0
$\Delta t = 0,01$	10,43	2,512	0,072 94	-0,073 87	-0,073 87
$\Delta t = 0,005$	10,36	2,512	0,124 2	-0,022 50	-0,073 87

Crank-Nicolson-Adams-Bashforth modifié					
	$N = 100$	$N = 200$	$N = 500$	$N = 1000$	$N = 2000$
$\Delta t = 0,04$	9,702	1,152	-1,440	-1,866	-1,866
$\Delta t = 0,02$	10,43	2,195	-0,029 37	-0,381 0	-0,455 3
$\Delta t = 0,01$	10,43	2,512	0,072 94	-0,073 87	-0,073 87
$\Delta t = 0,005$	10,43	2,565	0,124 2	-0,022 50	-0,073 87

SBDF second ordre					
	$N = 100$	$N = 200$	$N = 500$	$N = 1000$	$N = 2000$
$\Delta t = 0,04$	8,284	-0,472 7	-3,005	-3,544	-3,544
$\Delta t = 0,02$	9,942	1,985	-0,638 9	-0,787 6	-0,787 6
$\Delta t = 0,01$	10,30	2,406	-0,029 37	-0,176 5	-0,176 5
$\Delta t = 0,005$	10,43	2,565	0,124 2	-0,073 87	-0,073 87

SBDF troisième ordre					
	$N = 100$	$N = 200$	$N = 500$	$N = 1000$	$N = 2000$
$\Delta t = 0,04$	11,16	3,260	1,003	0,161 7	0,161 7
$\Delta t = 0,02$	10,67	2,831	0,381 2	0,028 90	0,095 17
$\Delta t = 0,01$	10,55	2,618	0,175 5	0,028 90	0,028 90
$\Delta t = 0,005$	10,43	2,565	0,175 5	-0,022 50	-0,022 50

Crank-Nicolson					
	$N = 100$	$N = 200$	$N = 500$	$N = 1000$	$N = 2000$
$\Delta t = 0,04$	2,121	-6,128	-8,454	-8,620	-8,748
$\Delta t = 0,02$	5,999	-1,852	-4,146	-4,491	-4,491
$\Delta t = 0,01$	8,167	0,231 8	-2,131	-2,283	-2,381
$\Delta t = 0,005$	9,225	1,411	-0,991 1	-1,241	-1,241

Gear (premier ordre)					
	$N = 100$	$N = 200$	$N = 500$	$N = 1000$	$N = 2000$
$\Delta t = 0,04$	-5,201	-13,35	-15,13	-15,54	-15,54
$\Delta t = 0,02$	1,708	-6,305	-8,454	-8,620	-8,620
$\Delta t = 0,01$	5,776	-2,046	-4,520	-4,678	-4,678
$\Delta t = 0,005$	7,993	0,181 1	-2,229	-2,430	-2,430

Euler implicite-explicite					
	$N = 100$	$N = 200$	$N = 500$	$N = 1000$	$N = 2000$
$\Delta t = 0,04$	-1,483	-8,889	-11,12	-10,96	-11,21
$\Delta t = 0,02$	4,240	-3,571	-5,808	-5,969	-6,100
$\Delta t = 0,01$	7,243	-0,672 1	-3,005	-3,159	-3,159
$\Delta t = 0,005$	8,752	0,894 9	-1,490	-1,690	-1,690

Euler explicite					
	$N = 100$	$N = 200$	$N = 500$	$N = 1000$	$N = 2000$
nt	770	1900	10 000	37 300	148 000
Δt	0,039 0	0,015 8	0,003	0,000 804	0,000 203
% (nt)	-0,032 16	0,290 5	-0,304 6	-0,208 9	-0,102 3
% ($1,5 nt$)	1,542	0,827 4	-0,172 3	-0,170 4	-0,093 31
% ($2 nt$)	2,348	1,179	-0,090 67	-0,147 1	-0,088 81

mentionné, la taille de son Δt_c dépend très fortement de la résolution du maillage, ce qui la rend ultimement trop délicate d'utilisation.

Il est cependant utile de noter que l'utilisation de méthodes d'ordre très élevé n'est pas nécessairement un avantage. Par exemple, la vitesse de l'onde obtenue avec la méthode SBDF de 3^e ordre converge plus rapidement vers celle de la solution modèle que celle obtenue avec les différentes méthode de 2^e ordre. Cependant, cette méthode est aussi moins stable que les méthodes de 2^e ordre, ayant un Δt_c inférieur à celui des méthodes de 2^e ordre (voir le tableau 2). On peut aussi voir sur les deux images au bas de la figure 2 que même si la solution obtenue avec le SBDF de 3^e ordre reste bornée, nous avons besoin de plus de travail pour faire disparaître les instabilités de la solution. Quant à la figure 3, elle montre que cette méthode ne donne pas réellement de meilleurs résultats, à tout le moins quand $N = 500$. Nous concluons donc qu'utiliser une méthode de second ordre semble être le choix optimal. Parmi celles-ci, le schéma de Gear implicite se démarque par sa grande stabilité et la nécessité d'utiliser un très petit nombre de pas de temps pour obtenir une excellente solution, mais il demande aussi la résolution d'un système d'équations non linéaires à chaque pas de temps, ce qui, tel que mentionné à la section 2.2, demande beaucoup de travail. Nous nous concentrerons donc sur les méthodes implicites-explicites de 2^e ordre. Parmi celles-ci, la méthode SBDF de 2^e ordre nous semble être la plus intéressante. Elle est plus stable que les méthodes de Crank-Nicolson-Adams-Bashforth et de Crank-Nicolson-Adams-Bashforth modifiée (qui, en fait, génèrent des solutions qui diffèrent très peu l'une de l'autre), et s'il est vrai qu'elle préserve moins bien la vitesse de l'onde quand le pas de temps est trop grand (par exemple pour $\Delta t = 0,04$ sur la figure 2), elle produit une solution de meilleure qualité quand un pas de temps plus petit est utilisé. Elle a aussi l'avantage de nécessiter le calcul du terme de diffusion au pas de temps $n + 1$ seulement. Pour toutes ces raisons, c'est cette méthode que nous préconiserons.

Nous remarquerons aussi que parmi les méthodes implicites-explicites que nous avons utilisées, la méthode SBDF de 3^e ordre (21) est la seule qui produit une onde qui devance l'onde modèle. Toutes les autres méthodes implicites-explicites retardent. De plus, nos observations présentées au tableau 3 semblent indiquer que pour une valeur de N fixée, la vitesse de l'onde pour les méthodes SBDF de 3^e ordre et de 2^e ordre

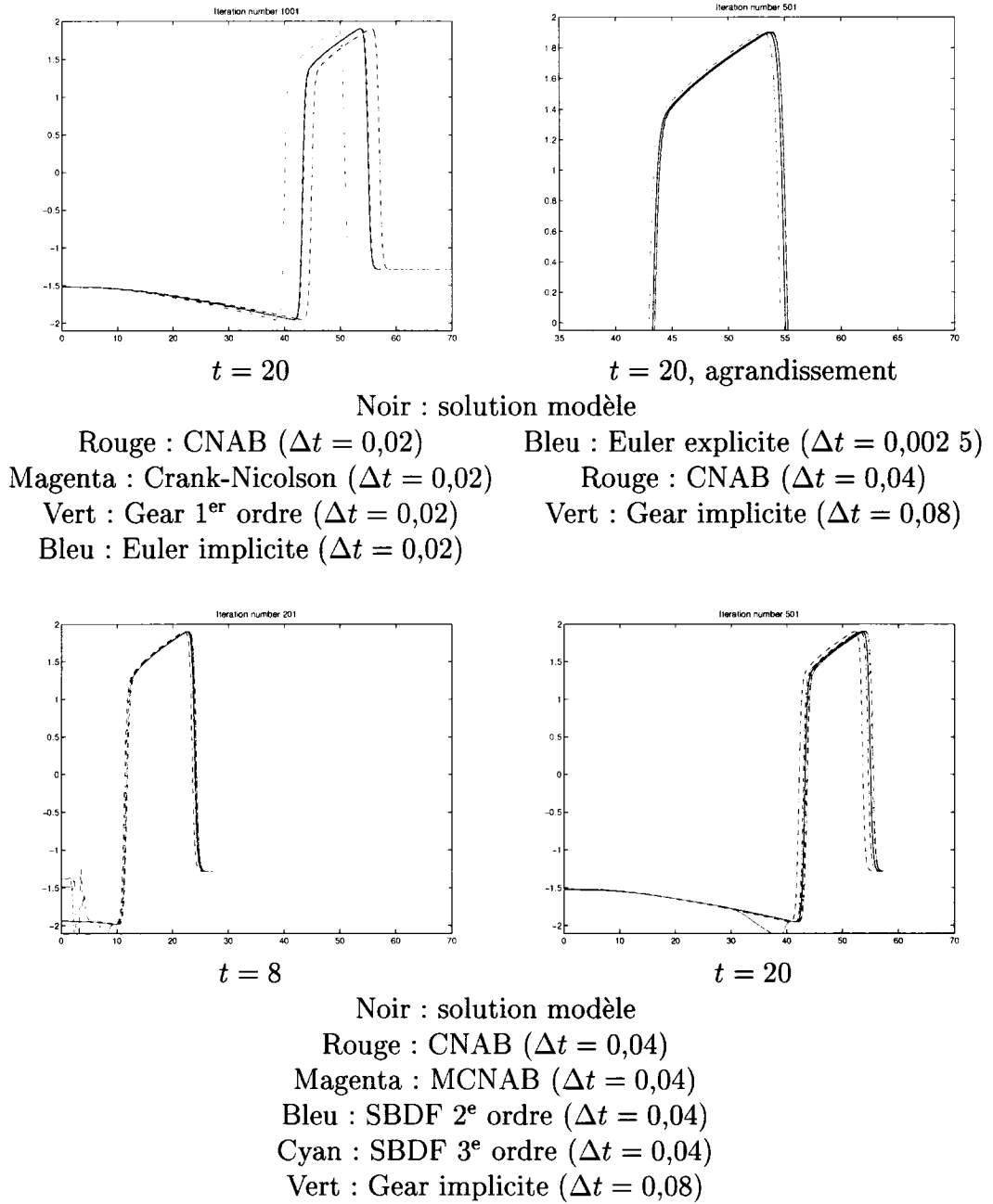
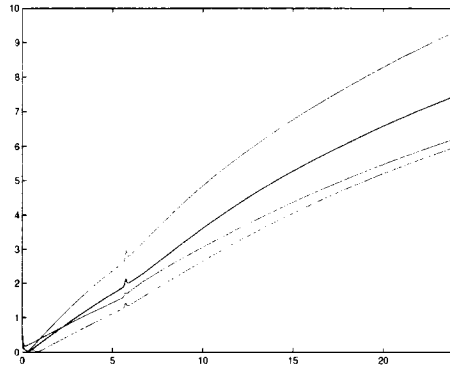
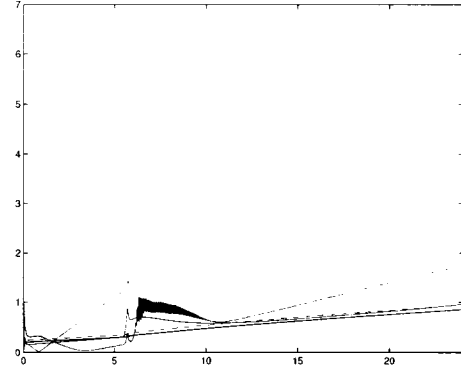


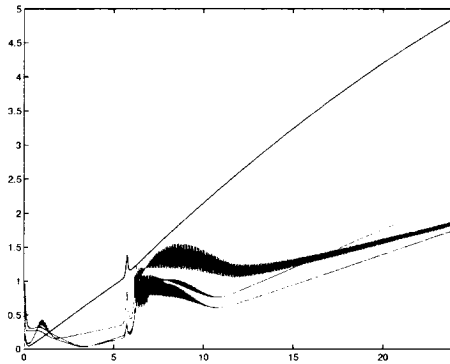
FIG. 2 – Position de l’onde pour u à différents t pour plusieurs méthodes ($N = 500$)



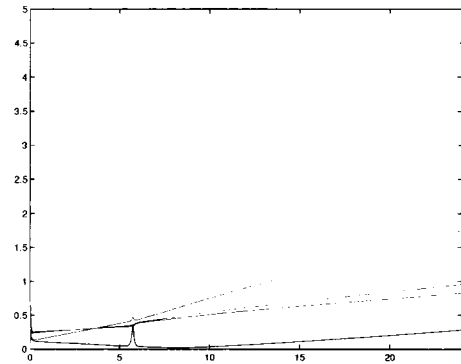
Vert : Gear 1^{er} ordre ($\Delta t = 0,02$)
 Bleu : Euler implicite-explicite ($\Delta t = 0,02$)
 Cyan : Euler implicite ($\Delta t = 0,02$)
 Rouge : Crank-Nicolson ($\Delta t = 0,02$)



Magenta : CN ($\Delta t = 0,02$)
 Rouge (pleine) : CNAB ($\Delta t = 0,04$)
 Rouge (tirets) : CNAB ($\Delta t = 0,02$)
 Vert : Gear implicite ($\Delta t = 0,08$)
 Bleu : Euler explicite ($\Delta t = 0,0025$)



Rouge : CNAB ($\Delta t = 0,04$)
 Magenta : MCNAB ($\Delta t = 0,04$)
 Bleu : SBDF 2^e ordre ($\Delta t = 0,04$)
 Cyan : SBDF 3^e ordre ($\Delta t = 0,04$)



Rouge : CNAB ($\Delta t = 0,02$)
 Magenta : MCNAB ($\Delta t = 0,02$)
 Bleu : SBDF 2^e ordre ($\Delta t = 0,02$)
 Cyan : SBDF 3^e ordre ($\Delta t = 0,02$)

FIG. 3 – Évolution en temps de l'erreur L^2 pour plusieurs méthodes ($N = 500$)

(19) converge vers la même valeur avec la diminution de Δt , mais de façon monotone croissante dans le cas du SBDF de 2^e ordre et de façon monotone décroissante dans le cas du SBDF de 3^e ordre. Nous n'avons cependant aucune preuve de cette conjecture.

4.2 Comparaison des schémas en 2D

Pour résoudre les équations du modèle bidomaine en deux dimensions, nous avons utilisé un code orienté objet en C++ utilisant la librairie de solveurs PETSc.¹ Cette librairie donne accès à plusieurs méthodes directes et itératives et à plusieurs préconditionneurs pour résoudre des systèmes d'équations linéaires. Nos essais effectués sur des maillages en espace relativement grossiers nous ont convaincus que la factorisation LU est idéale pour ce problème, mais étant une méthode directe, elle utilise beaucoup d'espace en mémoire, ce qui risque de dépasser les limites de la mémoire vive accessible si des maillages plus fins sont utilisés. Nous avons effectué des tests avec différents solveurs itératifs et préconditionneurs, et il appert que la méthode du gradient conjugué, avec comme préconditionneur la factorisation LU incomplète (voir [18]), utilise le moins de temps de calcul pour notre problème. Les tests présentés dans cette section ont été effectués avec ce solveur itératif, en utilisant comme préconditionneur la factorisation LU incomplète avec deux niveaux de remplissage, sauf mention contraire. Ce nombre de niveaux de remplissage est le plus grand qui pouvait être utilisé pour nos tests, compte tenu de nos limites de mémoire. Nous avons utilisé quelques méthodes de différences finies pour approximer la dérivée en temps, dont entre autres la méthode SBDF de second ordre (19), qui était celle qui s'était le plus démarquée dans nos tests 1D.

4.2.1 Problème

Le problème que nous considérons est la version 2D du problème présenté aux sections 2.1 et 2.2, incluant les conditions aux limites de Neumann. Tout comme pour le problème 1D, nous fixons la valeur de la fonction u_e quelque part sur le

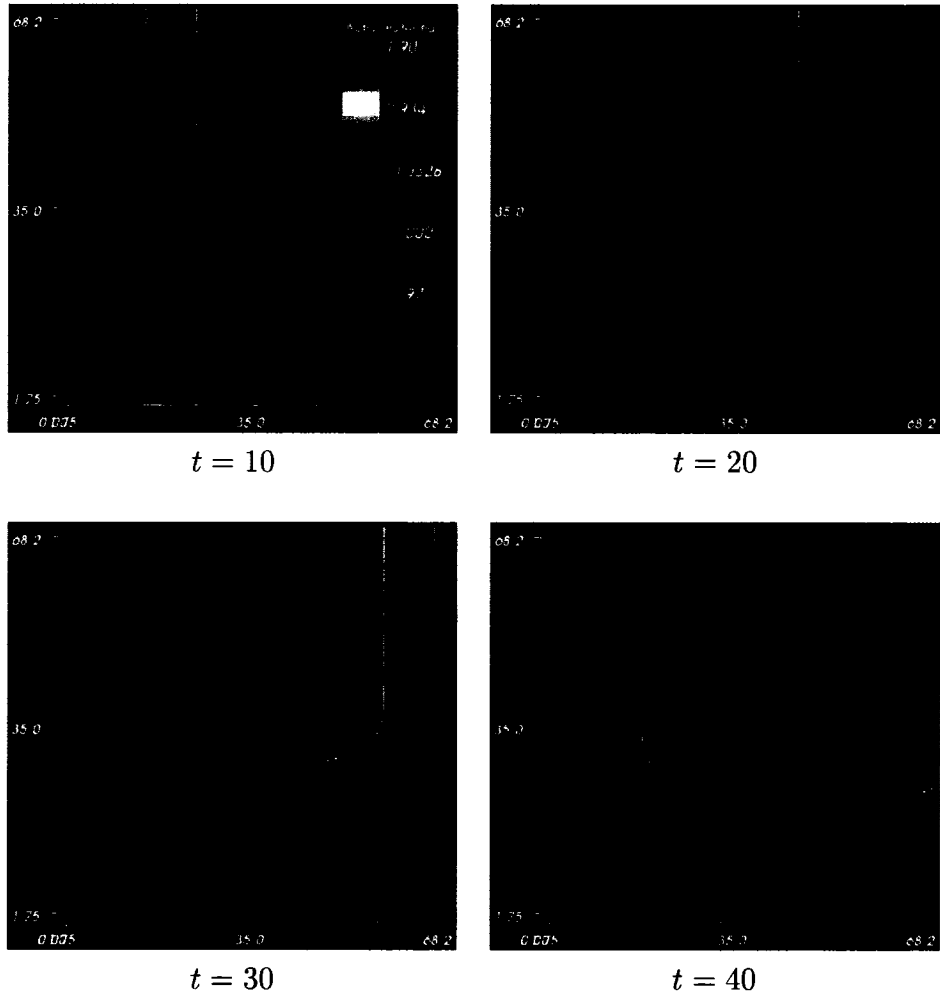
¹Voir la documentation à l'adresse [http : //www-unix.mcs.anl.gov/petsc/petsc-2/](http://www-unix.mcs.anl.gov/petsc/petsc-2/)

domaine pour obtenir un problème non singulier ; nous avons choisi $u_e(0,0) = 0$. Nous travaillons sur le domaine $[0, 70] \times [0, 70]$ divisé en triangles de la façon suivante : divisons le domaine en N intervalles le long de l'axe des x et de l'axe des y , et divisons ensuite chacun des carrés obtenus en quatre triangles congrus dont le centre du carré sera un sommet commun. Nous obtenons donc un maillage composé de $4N^2$ éléments triangulaires. Pour les tests présentés dans cette section, nous avons utilisé une valeur de N égale à 300 ; nous avons au préalable vérifié que la factorisation LU était la meilleure méthode de résolution du système linéaire pour $N = 100$. Nous avons utilisé comme paramètres $\epsilon = 0,1$, $\beta = 1,0$ et $\gamma = 0,5$. Comme condition initiale, nous avons fixé u et v à la solution du système $f(u, v) = 0$ et $g(u, v) = 0$, ainsi que u_e à 0, sauf sur $[0, 3,5] \times [0, 70]$, où nous avons fixé $u = 2$, tel qu'expliqué dans la sous-section 4.1.1, et sur $[31, 39] \times [0, 35]$, où nous fixons $v = 2$, ce qui crée une zone de recouvrement. Cela produit une onde plane de potentiel qui se déplace à travers le domaine à vitesse constante, jusqu'à ce qu'elle frappe la zone de recouvrement au centre du domaine, auquel point elle s'enroule autour de cette zone, formant une spirale. La figure 4 montre la façon dont cette spirale se forme pour u . Nous utilisons la méthode SBDF de 2^e ordre, une résolution de maillage $N = 300$, un pas de temps $\Delta t = 0,01$ et des tenseurs de conductivité $\sigma_i = 5 \text{ diag}\{0,263, 0,026\}$, $\sigma_e = 5 \text{ diag}\{0,263, 0,108\}$. Ces valeurs de conductivité sont proportionnelles à celles qui ont été utilisées par Hooks *et al.* [11].

4.2.2 Résultats

Nous utiliserons comme point de comparaison entre les différentes méthodes utilisées la trajectoire du bout de la spirale. Nous définissons le bout de la spirale comme le point où les courbes de niveau à valeur 0 pour u et v se rencontrent. Nos critères de comparaison seront si la trajectoire du bout de la spirale pour la solution numérique a la même forme que celle pour la solution modèle, et si la méthode cause une dérive du bout de la spirale.

Il est très difficile d'obtenir une solution de référence très précise pour le problème 2D étant donné que le nombre d'éléments augmente de façon quadratique par rapport

FIG. 4 – Solution en u des équations du modèle bidomaine 2D à différents temps

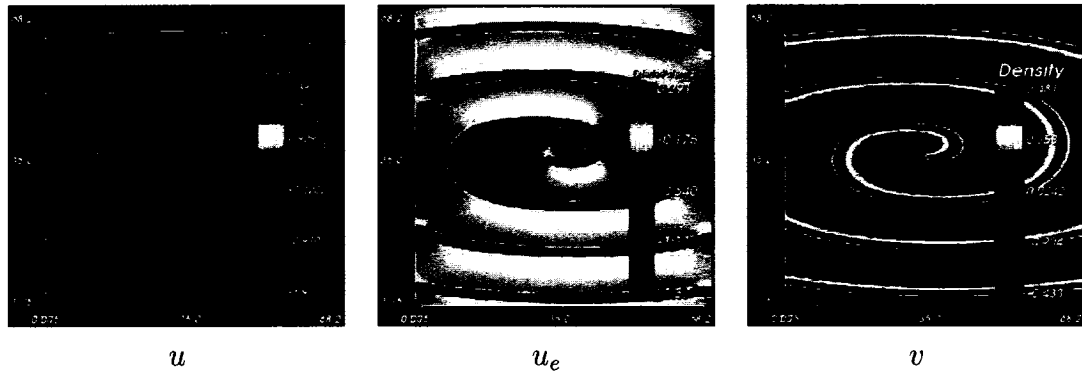


FIG. 5 – Condition initiale pour le problème 2D

à la résolution du maillage dans chaque direction spatiale. Nous avons utilisé $N = 300$ comme résolution du maillage, malgré le fait que cette valeur de N est à la limite de ce qui est nécessaire pour assurer que la solution soit correcte et que dans certains cas elle n'est pas suffisante, faisant apparaître des effets de maillage sur la trajectoire du bout de la spirale. La raison est que l'utilisation d'un maillage d'une plus grande finesse nécessiterait l'emploi de préconditionneurs moins gourmands en mémoire, et l'amélioration du conditionnement du système par l'utilisation de pas de temps plus petits, ce qui aurait un coût important en temps de calcul. Pour éviter l'apparition d'effets de maillage, nous modifierons donc certains paramètres : une plus forte conductivité nous permet d'obtenir une onde de potentiel dont le front et l'arrière sont plus larges et se trouvent donc sur davantage d'éléments, ce qui réduit les effets de maillage quand $N = 300$. Nos tests en 2D seront donc effectués avec les tenseurs de conductivité $\sigma_i = 25 \text{ diag}\{0,263, 0,026\}$ et $\sigma_e = 25 \text{ diag}\{0,263, 0,108\}$. Nous testerons les méthodes SBDF 2^e ordre (19) et Crank-Nicolson-Adams-Bashforth (17), c'est-à-dire les méthodes qui s'étaient démarquées dans nos tests 1D, ainsi que le Crank-Nicolson de 1^{er} ordre (14), pour pouvoir comparer des méthodes d'ordres différents. Nous prendrons comme condition initiale une onde spirale déjà formée, obtenue de la façon mentionnée à la sous-section 4.2.1. Nous présentons cette solution initiale à la figure 5.

Durant la résolution du système, nous essayons de conserver le moins de matrices possible en mémoire, en raison de la grande quantité d'espace mémoire que cela nécessite, malgré le fait que les matrices utilisées en éléments finis sont généralement creuses et qu'il existe des structures de données spéciales qui permettent de les représenter efficacement. Par exemple, considérons la matrice de masse M présentée à la section 2.1, soit $M(i, j) = \int_{\Omega} \Phi_i^h \Phi_j^h dx$. Sur un maillage comme celui que nous utilisons pour le problème 2D, qui est composé de $4N^2$ éléments triangulaires, chaque noeud est le voisin de huit autres noeuds, et chaque ligne de M , à l'exception de celles qui correspondent aux frontières du domaine, contiendra donc neuf entrées non nulles (en incluant le cas où $i = j$). Bien que creuse, cette matrice contient quand même une assez grande quantité d'éléments non nuls. Maintenant, étant donné que notre implémentation requiert de résoudre les trois équations du système simultanément, la matrice du système linéaire, qu'il est nécessaire de conserver en mémoire, contient trois fois le nombre de lignes et trois fois le nombre de colonnes que cette matrice M , qui possède elle-même un nombre de lignes et de colonnes égal au nombre de noeuds du domaine.

Nous avons besoin de conserver au moins deux matrices en mémoire pendant la résolution : la sus-mentionnée matrice du système linéaire, ainsi que la matrice du préconditionneur. Nous les calculons avant le premier pas de temps et elles ne sont pas modifiées par la suite, ce qui fait que le travail requis pour faire leur assemblage n'a lieu qu'une seule fois. L'implémentation que nous avons choisie de nos trois méthodes ne requiert pas de garder d'autres matrices en mémoire. Elles sont donc équivalentes sur ce point. Par contre, elles diffèrent sur les vecteurs qu'il faut assembler à chaque pas de temps. Le tableau 4 montre quels vecteurs doivent être gardés en mémoire pour chacune des méthodes. On suppose que le prochain pas de temps est celui qui calculera la solution au temps t_{n+1} .

Le tableau 4 montre que le terme de dérivée en temps pour la méthode SBDF de 2^e ordre demande de conserver en mémoire un vecteur de plus que les autres méthodes, soit la solution au temps t_{n-1} . De même, ce schéma ainsi que le schéma CNAB requièrent que soit conservée la valeur du terme de réaction à ce même temps, contrairement au Crank-Nicolson de 1^{er} ordre. Cependant, la méthode SBDF se reprend

TAB. 4 – Vecteurs utilisés par différents schémas numériques

	Dérivée en temps		Terme de réaction		Terme de diffusion
	t_n	t_{n-1}	t_n	t_{n-1}	t_n
CNAB	X		X	X	X
SBDF2	X	X	X	X	
CN	X		X		X

pour le terme de diffusion : elle n'en utilise l'approximation qu'au temps t_{n+1} , qui est intégrée dans la matrice du système linéaire, tandis que les deux autres méthodes demandent aussi d'en faire le calcul au temps t_n , ce qui nécessite de faire l'assemblage d'un vecteur supplémentaire à chacun des pas de temps. L'assemblage du membre de droite du système linéaire est donc un peu plus coûteux en temps de calcul, à chaque pas de temps, pour les méthodes de Crank-Nicolson-Adams-Bashforth et de Crank-Nicolson-Euler explicite que pour la méthode SBDF de 2^e ordre.

Pour les deux méthodes de 2^e ordre testées, nous avons utilisé trois pas de temps différents : $\Delta t = \frac{1}{50}$, $\Delta t = \frac{1}{30}$ et Δt près de Δt_c pour la méthode en question. Dans le cas de Crank-Nicolson-Adams-Bashforth, il s'agit de $\Delta t = \frac{1}{18}$, et dans le cas de SBDF 2^e ordre, $\Delta t = \frac{1}{14}$. La méthode de Crank-Nicolson-Euler explicite, quant à elle, n'a été testée qu'avec deux pas de temps différents : $\Delta t = \frac{1}{50}$ et $\Delta t = \frac{1}{30}$. Nous avons obtenu nos valeurs de Δt proches de la valeur critique en effectuant des essais avec des valeurs de Δt supérieures à Δt_c , mais décroissantes, jusqu'à ce que le schéma devienne stable. Nous pouvons donc dire que pour la solution initiale à la figure 5, $\frac{1}{12} > \Delta t_c \geq \frac{1}{14}$ pour SBDF de 2^e ordre et $\frac{1}{17} > \Delta t_c \geq \frac{1}{18}$ pour CNAB. Tout comme dans le cas 1D (voir le tableau 2), la méthode SBDF semble un peu plus stable que la méthode CNAB.

Nous remarquons premièrement que les valeurs de Δt les plus grandes que nous avons utilisées ne sont pas suffisantes pour produire une solution qui ne présente pas d'oscillations. Ces oscillations sont évidentes à la troisième ligne d'images sur les figures 6 et 7, soit celles où Δt est pris près de sa limite de stabilité. Nous avons aussi déjà noté (voir le tableau 3) que dans le cas 1D, l'onde de potentiel retarde quand

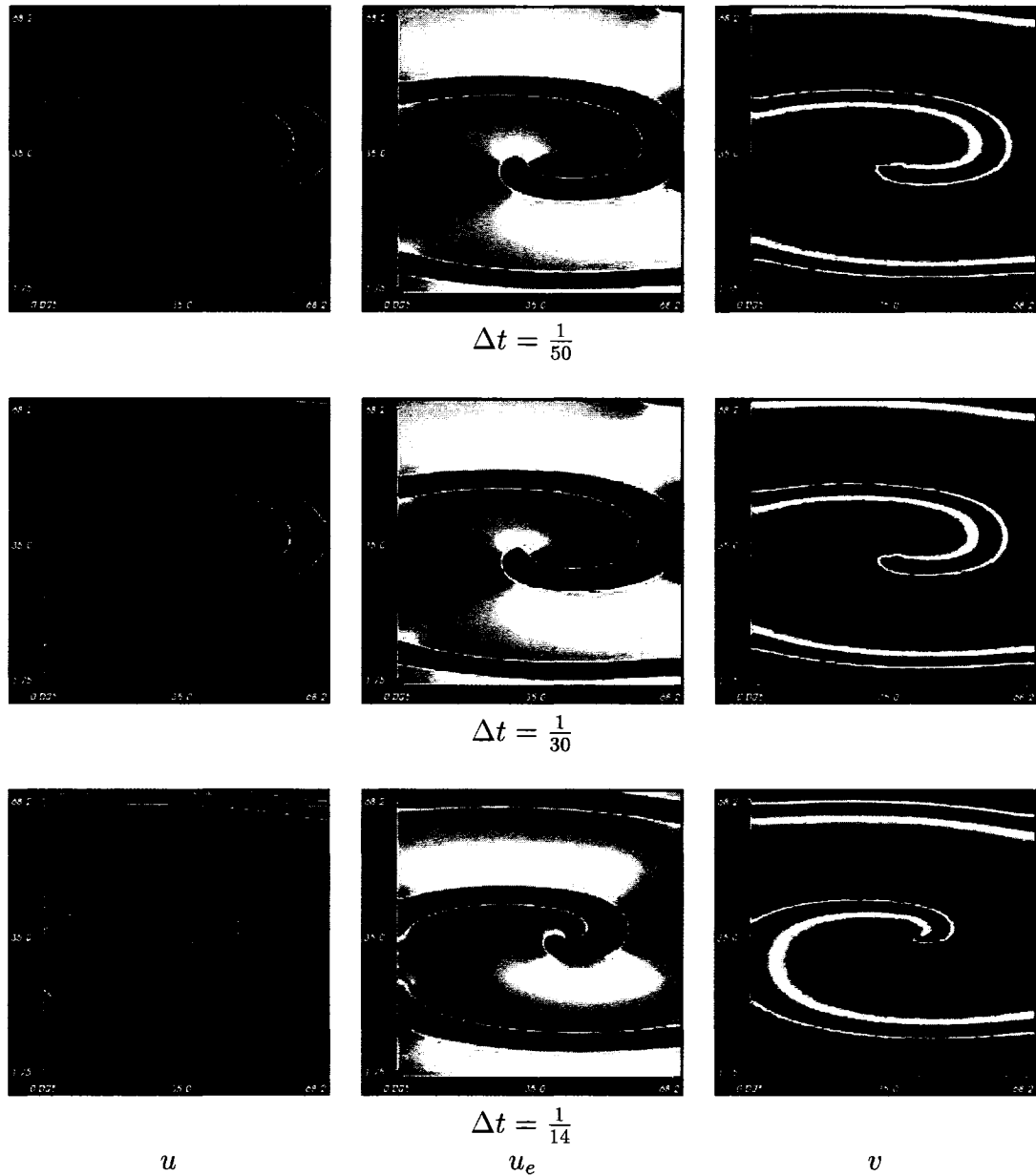


FIG. 6 – Solution à $t = 80$ obtenue avec SBDF 2^e ordre et différentes valeurs de Δt . La légende est la même que celle de la figure 5.

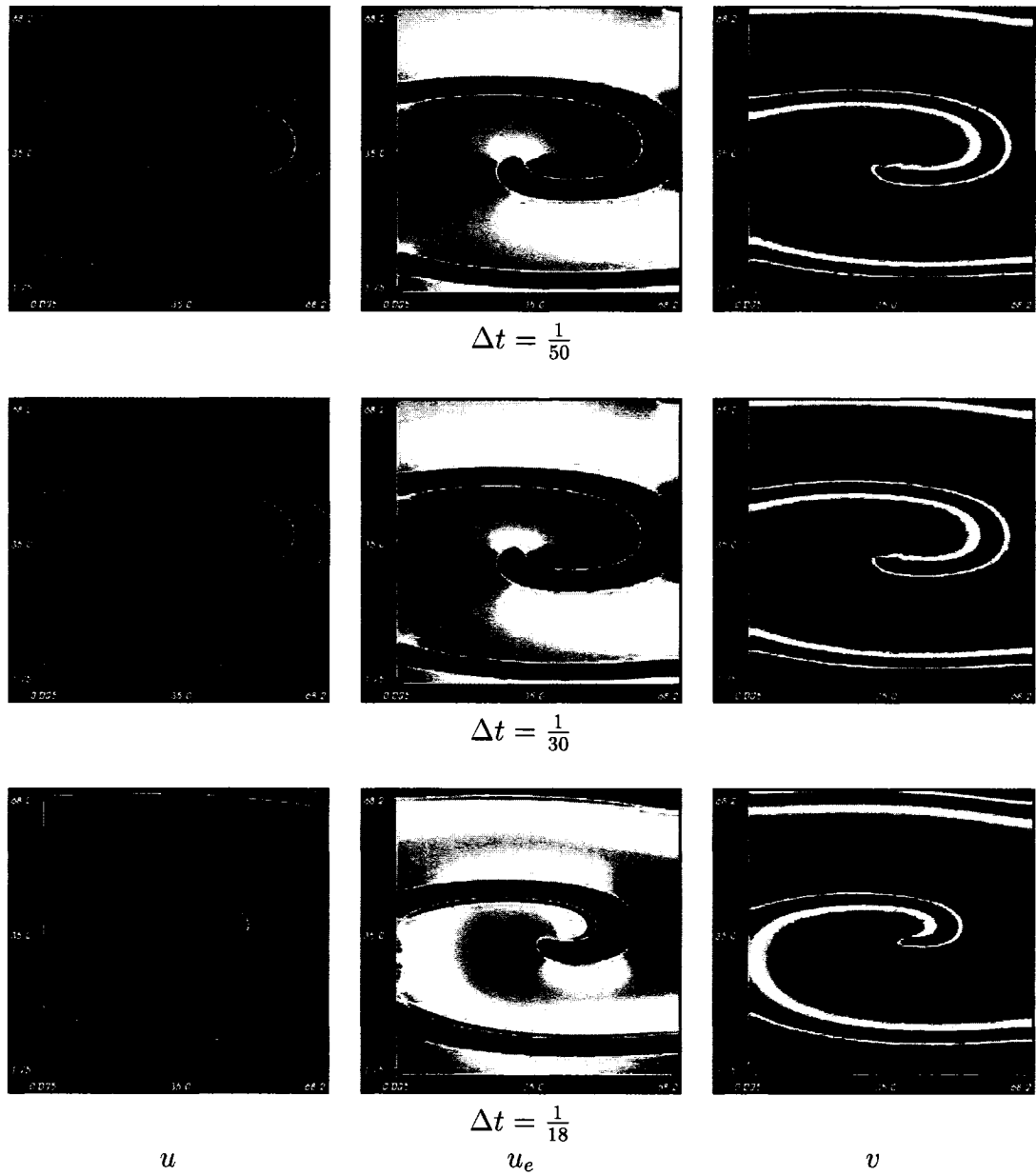


FIG. 7 – Solution à $t = 80$ obtenue avec CNAB et différentes valeurs de Δt . La légende est la même que celle de la figure 5.

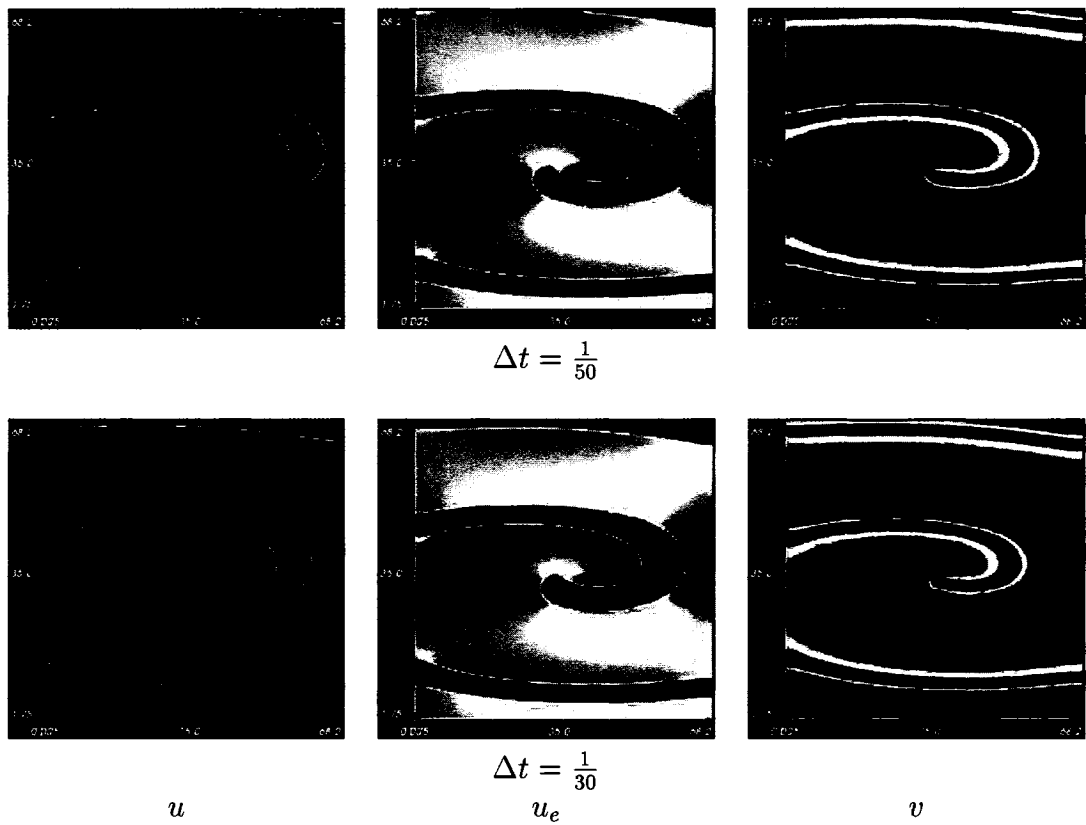


FIG. 8 – Solution à $t = 80$ obtenue avec Crank-Nicolson et différentes valeurs de Δt . La légende est la même que celle de la figure 5.

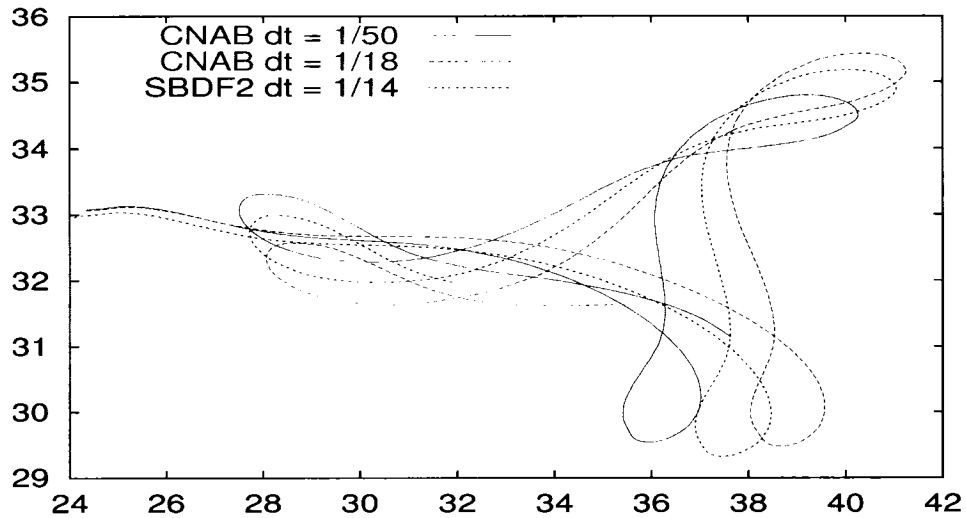


FIG. 9 – Trajectoire du bout de la spirale de $t = 10$ à $t = 55$ pour des méthodes de 2^e ordre et différentes valeurs de Δt

Δt est trop grand, pour les méthodes SBDF de 2^e ordre et CNAB. On peut voir aux figures 6 et 7 que cette propriété demeure vraie dans le cas 2D : les solutions qui utilisent une valeur de Δt trop grande ne semblent pas être aussi avancées. À titre d'exemple, on aperçoit une branche de la spirale au haut des images sur la troisième rangée de ces deux figures, tandis que sur les deux premières rangées, cette branche est presque entièrement sortie du domaine. Nous remarquons aussi finalement que la solution obtenue avec Δt trop grand présente une dérive importante par rapport à des solutions plus précises en temps. Nous pouvons voir cette dérive dans la trajectoire du bout de la spirale sur la figure 9. Les courbes verte et bleue (CNAB, $\Delta t = \frac{1}{18}$ et SBDF2, $\Delta t = \frac{1}{14}$, respectivement) s'éloignent de plus en plus de la courbe rouge (CNAB, $\Delta t = \frac{1}{50}$) au fur et à mesure que le temps avance. Le début de la trajectoire, sur les figures 9 à 11, se trouve à la gauche de l'image, sauf sur la deuxième image à la figure 11, où il se trouve en haut à droite.

Tout comme dans le cas du problème 1D, nous avons pu déterminer que des méthodes de 1^{er} ordre du type de la méthode de Crank-Nicolson-Euler explicite (14) ne sont pas appropriées pour ce genre de problème. Même dans le cas où le pas de

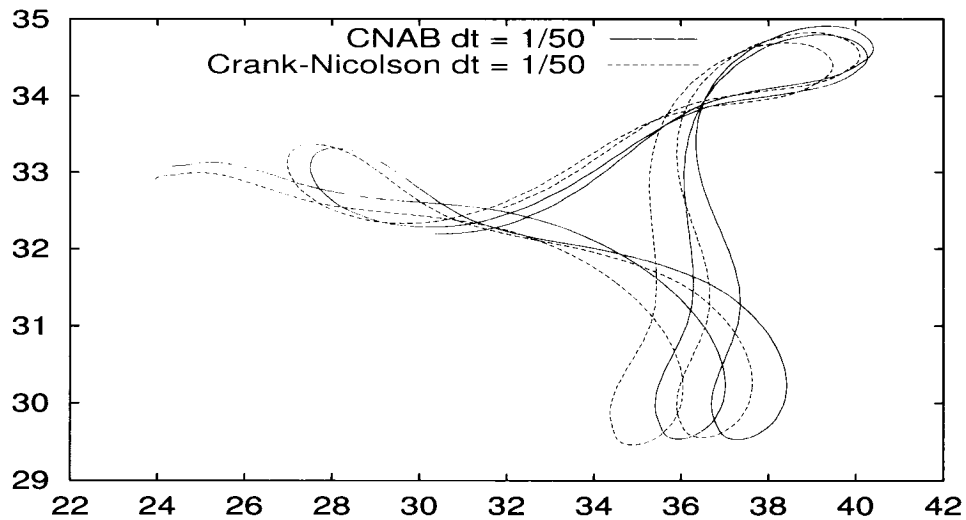


FIG. 10 – Trajectoire du bout de la spirale de $t = 10$ à $t = 80$ pour CNAB et Crank-Nicolson de 1^{er} ordre et $\Delta t = 1/50$

temps $\Delta t = \frac{1}{50}$, que l'on peut voir sur la première ligne de la figure 8, la solution présente déjà un retard de phase par rapport aux méthodes de 2^e ordre avec $\Delta t = \frac{1}{50}$ ou $\frac{1}{30}$, qui peut être remarqué dans le fait que le cœur de la spirale n'est pas aussi développé que pour les méthodes de 2^e ordre, et que le bras de la spirale au haut du domaine n'en est pas encore sorti. Cependant, ce retard de phase n'est pas aussi important que celui qui avait été constaté pour les méthodes de 2^e ordre avec le pas de temps Δt proche de sa valeur critique. On peut aussi voir en haut à droite sur la figure 10 le déplacement du bout de la spirale pour les schémas de Crank-Nicolson et CNAB avec $\Delta t = \frac{1}{50}$. Il est clair sur cette figure qu'une erreur est déjà apparue au cours des 10 unités de temps qui précèdent le début de l'image, et que cette erreur s'agrandit par la suite rapidement.

Les images sur la figure 11 montrent l'évolution du bout de la spirale pour les deux méthodes de 2^e ordre que nous avons testées, avec $\Delta t = \frac{1}{50}$ et $\Delta t = \frac{1}{30}$. On y voit que la solution obtenue est assez similaire peu importe la méthode utilisée. Cependant, on note que la méthode SBDF de 2^e ordre avec $\Delta t = \frac{1}{30}$ est un peu décalée par rapport

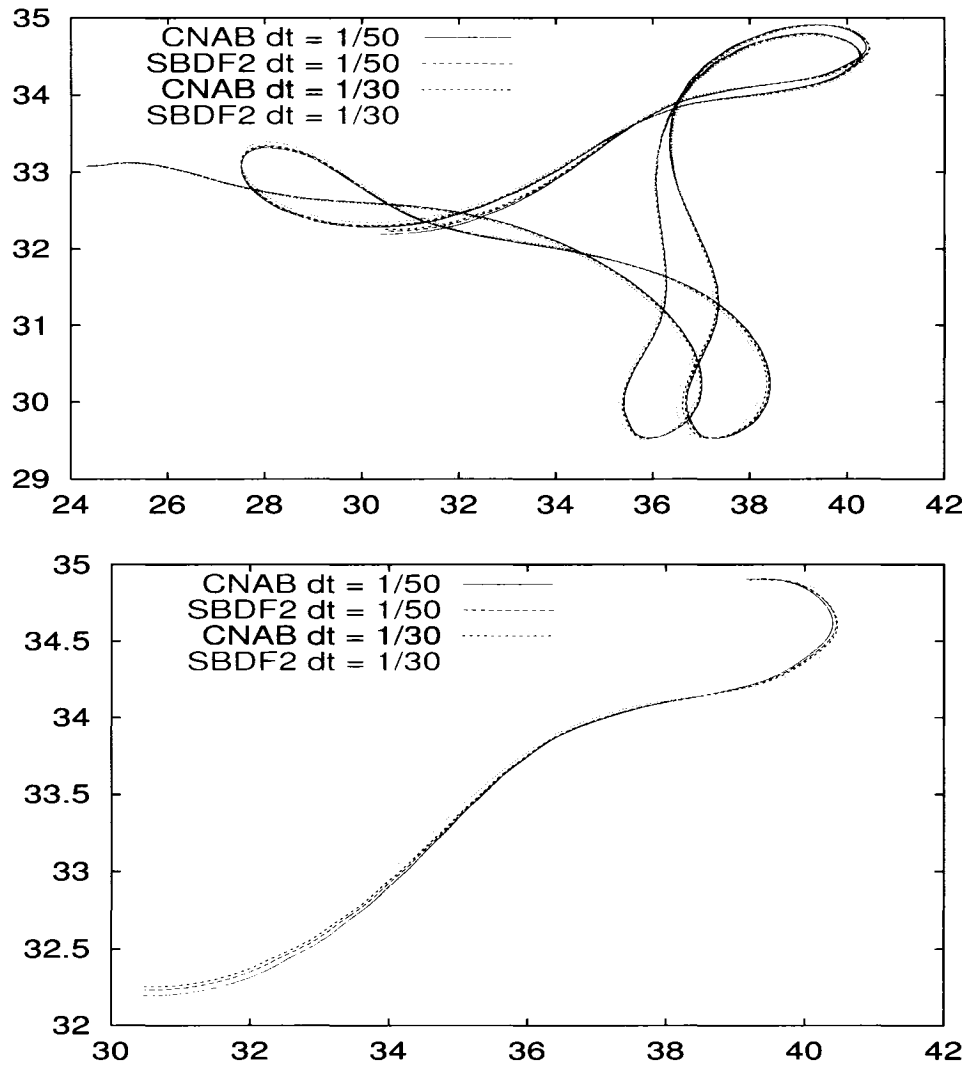


FIG. 11 - Trajectoire du bout de la spirale de $t = 10$ à $t = 80$ (haut) et de $t = 70$ à $t = 80$ (bas) pour des méthodes de 2^e ordre et différentes valeurs de Δt

aux autres, et ce décalage augmente au fil du temps. Cette situation est plus visible sur l'image du bas, au temps final qui se situe en bas à gauche de l'image. À l'inverse, la solution obtenue avec la méthode CNAB et $\Delta t = \frac{1}{30}$ suit de très près les deux solutions obtenues avec $\Delta t = \frac{1}{50}$. Souvenons-nous que dans le cas 1D, la méthode SBDF de 2^e ordre retarde davantage que la méthode CNAB pour des valeurs de Δt plus grandes, ce qui contribue à ce qu'elle produise une erreur numérique plus grande (cf. figure 3 en bas à gauche, et tableau 3, première ligne des sous-tableaux « Crank-Nicolson-Adams-Bashforth » et « SBDF second ordre », où $\Delta t = \frac{1}{25}$). Cependant, les pas de temps sont sensiblement plus coûteux pour un problème 2D que pour un problème 1D, et donc, il est probablement mieux d'utiliser une méthode pour laquelle la solution est meilleure pour des valeurs de Δt assez grandes. Dans ce cas-ci, il semble donc que la méthode Crank-Nicolson-Adams-Bashforth soit la plus appropriée, mais cependant, le schéma SBDF de 2^e ordre se classe presque aussi bien, il a l'avantage que, comme nous avons mentionné plus tôt, chacun de ses pas de temps coûte légèrement moins en temps de calcul que les autres méthodes de 2^e ordre, et il est probable que pour des valeurs de Δt de l'ordre de $\frac{1}{50}$ ou moins, il produise une meilleure solution. C'est du moins là le résultat que nous avons obtenu en 1D (voir la figure 3 en bas à droite, où $\Delta t = \frac{1}{50}$).

Chapitre 5

Conclusion

Dans ce chapitre, nous commencerons par résumer les résultats auxquels nous sommes arrivés dans cette thèse, avant de décrire la recherche qui demeure à effectuer.

5.1 Résumé

Dans le cadre de la recherche que nous avons effectuée, nous avons étudié plusieurs méthodes d'intégration en temps implicites, semi-implicites et explicites, pour résoudre les équations du modèle bidomaine, méthodes que nous avons présentées au chapitre 2. Au chapitre 3, nous avons décrit les conditions à poser sur le pas de temps Δt pour s'assurer de la stabilité de la solution. Nous avons ensuite, au chapitre 4, testé ces différentes méthodes pour des cas tests identiques. Nous noterons au passage que les valeurs du pas de temps critique calculé numériquement pour nos méthodes, que nous avons montrées au tableau 2, sont consistantes avec les analyses de stabilité du chapitre 3. En effet, le pas de temps critique pour la méthode d'Euler explicite dépendait aussi de h^2 dans nos simulations, tandis que celui des méthodes semi-implicites n'avait que pas ou très peu de dépendance en h , et que celui de la méthode d'Euler implicite était grand et n'avait aussi aucune dépendance sur h . Nous en sommes arrivés à la conclusion que les méthodes semi-implicites d'ordre supérieur sont bien appropriées à la résolution du modèle bidomaine. En effet, elles sont assez stables et leur stabilité ne dépend pas de h . Elles donnent aussi de bons résultats

numériques, même quand le pas de temps est assez grand, ce qui les rend supérieures aux méthodes semi-implicites de 1^{er} ordre. De plus, si le coût de chacun des pas de temps effectués avec ces méthodes dépend du choix du solveur et du préconditionneur, ce qui nécessite un travail pour déterminer lesquels sont les plus appropriés, elles ne sont quand même pas trop onéreuses en temps de calcul, n'ayant pas à résoudre de système d'équations non linéaires.

Nous avons aussi pu discriminer entre les différentes méthodes semi-implicites d'ordre supérieur que nous avons testées. Deux d'entre elles se sont particulièrement démarquées du lot : la méthode de Crank-Nicolson-Adams-Bashforth (CNAB) et la méthode SBDF de 2^e ordre. Nous étions arrivés à la conclusion que quoique la méthode SBDF de 2^e ordre soit plus stable que la méthode CNAB, ces deux schémas ne produisent pas une solution acceptable quand on utilise un pas de temps trop près de la valeur critique. Avec un pas de temps approximativement égal à la moitié de cette valeur critique, la solution produite par les deux méthodes est bonne, mais celle produite par la méthode SBDF de 2^e ordre retarde davantage que celle produite par la méthode CNAB, et possède donc une erreur plus grande. C'est quand on passe à des pas de temps plus petits que le schéma SBDF commence à produire une solution qui est plus proche de la solution exacte que celle produite par le schéma CNAB. Cela dit, le schéma SBDF est aussi légèrement moins coûteux en temps de calcul que le schéma CNAB, puisqu'il ne requiert pas d'assembler un vecteur pour le terme de diffusion à chaque pas de temps.

5.2 Recherche future

Nous n'avons pas présenté de résultats 3D dans cette thèse. La raison est que les résultats 1D et 2D que nous avons obtenus nous ont convaincus que pour obtenir une solution numérique décente, il faut employer un maillage qui contient au moins 300 éléments dans chaque direction spatiale, ce qui serait très coûteux en espace mémoire, ainsi qu'en temps de calcul, dans le cas de simulations 3D. Il serait néanmoins probablement possible d'en effectuer, à condition d'employer un préconditionneur qui ne demande pas beaucoup de mémoire, et d'améliorer le conditionnement du système

par des pas de temps plus petits. Il nous reste donc à effectuer ce genre de calculs, mais nous réalisons que pour obtenir des résultats précis, il faudra investir beaucoup de temps de calcul.

Cela dit, peut-être existe-t-il une façon de s'assurer que les maillages utilisés sont suffisamment fins pour avoir une bonne solution numérique, sans pour autant que la méthode de résolution soit trop coûteuse en mémoire. Utiliser un maillage uniforme, comme nous avons fait en 1D comme en 2D, n'est probablement pas la méthode optimale. En effet, la méthode des éléments finis permet d'utiliser des maillages non structurés, et puisque la solution que nous avons obtenue avec nos cas tests est une onde de pulsation en 1D et une onde spirale en 2D, il existe, à tout moment, de grandes parties du domaine où la solution varie très peu. Il ne semble pas optimal de résoudre les équations en utilisant autant de nœuds sur ces sections du domaine que sur celles où le gradient de la solution est très grand (les fronts avant et arrière de la pulsation en 1D ; les bords de la spirale en 2D). Il conviendrait mieux d'avoir des éléments plus fins dans les zones où le gradient est élevé et plus grands là où la solution varie peu. Il serait donc peut-être possible d'utiliser une méthode d'adaptation de maillage avec le modèle bidomaine. Ces méthodes consistent à régulièrement modifier le maillage en le raffinant aux endroits où les équations à résoudre requièrent une grande précision, et en le rendant plus grossier là où il n'est pas nécessaire d'avoir un maillage aussi fin. Cependant, ces méthodes sont généralement très coûteuses en temps de calcul, puisqu'elles requièrent de totalement recréer le maillage, généralement à chaque pas de temps. Il pourrait néanmoins être une bonne idée de tenter d'appliquer une méthode de ce type à la résolution du modèle bidomaine en 2D et en 3D.

Bibliographie

- [1] AKRIVIS, G., CROUZEIX, M., AND MAKRIDAKIS, C. Implicit-explicit multistep finite element methods for nonlinear parabolic problems. *Mathematics of Computation* 67, 222 (1998), 457–477.
- [2] ASCHER, U. M., RUUTH, S. J., AND WETTON, B. T. R. Implicit-explicit methods for time-dependent partial differential equations. *SIAM J. Numer. Anal.* 32, 3 (1995), 797–823.
- [3] BOCHEV, P., AND LEHOUCQ, R. B. On the finite element solution of the pure Neumann problem. *SIAM Review* 47, 1 (2005), 50–66.
- [4] BOURGAULT, Y., ETHIER, M., AND LEBLANC, V. G. Simulation of electrophysiological waves with an unstructured finite element method. *ESAIM : Mathematical Modelling and Numerical Analysis* 37, 4 (2003), 649–661.
- [5] CIARLET, P. G. *The Finite Element Method for Elliptic Problems*, vol. 4 of *Studies in Mathematics and its Applications*. North-Holland, 1978.
- [6] COLLI FRANZONE, P., AND SAVARÉ, G. Degenerate evolution systems modeling the cardiac electric field at micro and macroscopic level. In *Evolution Equations, Semigroups and Functional Analysis : in memory of Brunello Terreni*, A. Lorenzi and B. Ruf, Eds., vol. 50. Birkhauser, 2002, pp. 49–78.
- [7] ERN, A., AND GUERMOND, J.-L. *Theory and Practice of Finite Elements*, vol. 159 of *Applied Mathematical Sciences*. Springer-Verlag, 2004.
- [8] GESELOWITZ, D. B., AND MILLER, W. T. A bidomain model for anisotropic cardiac muscle. *Annals of Biomedical Engineering* 11, 3-4 (1983), 191–206.
- [9] GREENBAUM, A. *Iterative Methods for Solving Linear Systems*. SIAM, 1997.

- [10] HOOKE, N., HENRIQUEZ, C. S., LANZKRON, P., AND ROSE, D. Linear algebraic transformations of the bidomain equations : Implications for numerical methods. *Mathematical Biosciences* 120, 2 (1994), 127–145.
- [11] HOOKS, D. A., TOMLINSON, K. A., MARSDEN, S. G., LEGRICE, I. J., SMAILL, B. H., PULLAN, A. J., AND HUNTER, P. J. Cardiac microstructure : Implications for electrical propagation and defibrillation in the heart. *Circulation Research* 91, 4 (2002), 331–338.
- [12] KEENER, J., AND SNEYD, J. *Mathematical Physiology*, vol. 8 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, 1998.
- [13] KEENER, J. P., AND BOGAR, K. A numerical method for the solution of the bidomain equations in cardiac tissue. *Chaos* 8, 1 (1998), 234–241.
- [14] QUARTERONI, A., SACCO, R., AND SALERI, F. *Numerical Mathematics*, vol. 37 of *Texts in Applied Mathematics*. Springer-Verlag, 2000.
- [15] ROTH, B. J. Approximate analytical solutions to the bidomain equations with unequal anisotropy ratios. *Physical Review E* 55, 2 (1997), 1819–1826.
- [16] ROTH, B. J. Frequency locking of meandering spiral waves in cardiac tissue. *Physical Review E* 57, 4 (1998), R3735–R3738.
- [17] RUUTH, S. J. Implicit-explicit methods for reaction-diffusion problems in pattern formation. *Journal of Mathematical Biology* 34, 2 (1995), 148–176.
- [18] SAAD, Y. Iterative methods for sparse linear systems. Disponible sur le Web à l'adresse [http : //www-users.cs.umn.edu/~saad/books.html](http://www-users.cs.umn.edu/~saad/books.html), 2000.
- [19] SUNDNES, J. *Numerical Methods for Simulating the Electrical Activity of the Heart*. PhD thesis, University of Oslo, 2002.