

**Titre:** Conception d'un schéma adaptatif d'intégration en temps appliqué à la discrétisation par éléments finis des équations de Maxwell pour l'étude de la supraconductivité  
**Title:**

**Auteur:** Damien Tageddine  
**Author:**

**Date:** 2018

**Type:** Mémoire ou thèse / Dissertation or Thesis

**Référence:** Tageddine, D. (2018). Conception d'un schéma adaptatif d'intégration en temps appliqué à la discrétisation par éléments finis des équations de Maxwell pour l'étude de la supraconductivité [Master's thesis, École Polytechnique de Montréal]. PolyPublie. <https://publications.polymtl.ca/3276/>  
**Citation:**

 **Document en libre accès dans PolyPublie**  
Open Access document in PolyPublie

**URL de PolyPublie:** <https://publications.polymtl.ca/3276/>  
**PolyPublie URL:**

**Directeurs de recherche:** Steven Dufour  
**Advisors:**

**Programme:** Maîtrise recherche en mathématiques appliquées  
**Program:**

UNIVERSITÉ DE MONTRÉAL

CONCEPTION D'UN SCHEMA ADAPTATIF D'INTÉGRATION EN TEMPS  
APPLIQUÉ À LA DISCRÉTISATION PAR ÉLÉMENTS FINIS DES ÉQUATIONS DE  
MAXWELL POUR L'ÉTUDE DE LA SUPRACONDUCTIVITÉ

DAMIEN TAGEDDINE  
DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL  
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION  
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES  
(MATHÉMATIQUES APPLIQUÉES)  
AOÛT 2018

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

CONCEPTION D'UN SCHÉMA ADAPTATIF D'INTÉGRATION EN TEMPS  
APPLIQUÉ À LA DISCRÉTISATION PAR ÉLÉMENTS FINIS DES ÉQUATIONS DE  
MAXWELL POUR L'ÉTUDE DE LA SUPRACONDUCTIVITÉ

présenté par : TAGEDDINE Damien

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

M. PRUDHOMME Serge, Ph.D., président

M. DUFOUR Steven, Ph.D., membre et directeur de recherche

M. LAFOREST Marc, Ph.D., membre

## DÉDICACE

*À mes parents et à ma soeur,  
inlassables soutiens.*

## REMERCIEMENTS

Il existe un paradoxe dans l'art de remercier, politesses méritées évidemment, et en même temps quelle avarice de mots ! Tant recevoir pour donner si peu. Pour autant, je ne vais pas me soustraire à l'exercice.

Je commence tout d'abord par remercier chaleureusement le Professeur Steven Dufour, pour sa patience et la confiance qu'il a su m'accorder tout au long de cette maîtrise et bien avant. Ces heures passées au laboratoire ont été une belle échappée mathématique dans des années de baccalauréat qui l'étaient parfois moins. Merci également d'avoir supervisé l'écriture de ce travail.

Je tiens aussi à remercier le maître d'enseignement Donatien N'dri de m'avoir permis d'enseigner le cours d'analyse numérique. Il est toujours agréable de partager sa passion, et lorsqu'on est rémunéré pour cela, c'est une victoire sur toute la ligne.

Ensuite, et comment l'oublier, je remercie tout particulièrement le Docteur Yann-Meing Law-Kam-Cio. D'abord chargé de cours puis ami, merci de m'avoir introduit aux mathématiques appliquées et à l'analyse numérique. J'ai beaucoup appris ces dernières années sous ta supervision bienveillante et tes conseils avisés.

Merci à mes collègues et amis du laboratoire, Daniel Lévesque, Wissem Maazoun, Kenan Kergrene et Antonin Paquette-Rufange, nos discussions mathématiques et métamathématiques, ainsi que nos longues divagations ont rendu ces années des plus agréables. Merci aussi à mes amis du génie physique, Sandryne David et Rodrigue Beaini pour nos longues pauses au CEGP, souvent méritées mais pas toujours.

Enfin, je tiens à remercier le Professeur Serge Prudhomme et le Professeur Marc Laforest d'avoir accepté d'être membre du jury et d'avoir eu la patience de lire et d'évaluer ce travail.

## RÉSUMÉ

Une nouvelle stratégie d'adaptation en temps grâce à un estimateur de l'erreur locale, basé sur la notion de variations  $\epsilon$ -voisines, est présentée. Une démonstration originale du lien entre cet estimateur et l'erreur locale est donnée. Une étude de la consistance, de l'efficacité et de la robustesse de la stratégie d'adaptation est menée sur un ensemble de problèmes raides standards. L'estimateur d'erreur est ensuite utilisé sur une quantité d'intérêt. Une démonstration de la consistance de l'algorithme de sélection de pas de temps qui en résulte est exposée. Une nouvelle méthode d'adaptation en quantité d'intérêt pour des schémas de différences finies est alors présentée. L'ensemble des stratégies d'adaptation développées est ensuite appliqué sur les discrétisations des équations de supraconductivité à haute température. La quantité d'intérêt est alors la perte de courant par effet Joule. Les résultats obtenus sont comparés avec des simulations obtenues avec une discrétisation à pas constants fins.

## ABSTRACT

A novel time step selection method based on an error estimate of the local error is given. The proposed error estimate is based on the concept of  $\epsilon$ -neighbour variations. An original proof leading to a link between the error estimate and the local error is presented. Numerical results support the analysis and explore the consistency, the robustness and the efficiency of the approach. A novel approach based on a so called quantity of interest is then introduced. A proof of the consistency of a time step selection algorithm is given leading to a new method of time steps control for finite difference schemes. Finally, the proposed method is applied to solve Maxwell's equations for high-temperature superconductor materials. The quantity of interest in this context is the so called AC loss.

## TABLE DES MATIÈRES

DÉDICACE . . . . .	iii
REMERCIEMENTS . . . . .	iv
RÉSUMÉ . . . . .	v
ABSTRACT . . . . .	vi
TABLE DES MATIÈRES . . . . .	vii
LISTE DES FIGURES . . . . .	ix
LISTE DES SIGLES ET ABRÉVIATIONS . . . . .	x
CHAPITRE 1 INTRODUCTION . . . . .	1
CHAPITRE 2 REVUE DE LITTÉRATURE . . . . .	4
2.1 Stratégies d'adaptation en temps en différences finies . . . . .	5
2.1.1 Méthode d'adaptation en temps par extrapolation de Richardson . . . . .	7
2.1.2 Méthodes de prédiction-correction . . . . .	8
2.1.3 Méthodes de Runge-Kutta emboîtées . . . . .	11
2.1.4 Méthode d'adaptation avec un contrôleur proportionnel-intégral (PI) . . . . .	13
2.2 Stratégies d'adaptation en temps basées sur une formulation variationnelle . . . . .	15
2.2.1 Estimateur d'erreur sur la solution . . . . .	16
2.2.2 Estimation d'erreur basée sur une quantité d'intérêt . . . . .	17
2.3 Problèmes raides . . . . .	18
2.3.1 Discussion sur les stratégies d'adaptation en temps . . . . .	18
2.3.2 Le cas des équations de Maxwell pour la supraconductivité . . . . .	20
CHAPITRE 3 ÉLÉMENTS DE THÉORIE . . . . .	22
3.1 Notations et définitions préliminaires . . . . .	23
3.2 Estimation de l'erreur sur la solution . . . . .	24
3.3 Estimation d'erreur sur une quantité d'intérêt . . . . .	31
3.4 Vérification de la stratégie d'adaptation en temps et résultats numériques . . . . .	38
CHAPITRE 4 APPLICATION AUX ÉQUATIONS DE SUPRACONDUCTIVITÉ . . . . .	48



4.1	Définition du problème . . . . .	49
4.2	Calcul de la perte AC . . . . .	50
4.3	Cas d'une plaque infinie supraconductrice . . . . .	50
4.4	Application à la modélisation d'un câble supraconducteur . . . . .	55
CHAPITRE 5 CONCLUSION . . . . .		58
RÉFÉRENCES . . . . .		60

## LISTE DES FIGURES

Figure 2.1	Adaptation du pas de temps modélisé par un système de contrôle proportionnel-intégral [28]. . . . .	14
Figure 3.1	Comparaison des stratégies d'adaptation pour des variations d'ordre $k = 0, 1, 2$ . (a) Précision de l'approximation. (b) Nombre de pas de temps en fonction de $\epsilon_R$ . . . . .	39
Figure 3.2	Discrétisation de l'équation de combustion avec la semi-norme $\Delta_2$ . (a) Solution discrète. (b) Pas de temps en fonction du temps. c) Ratio de contrôle. (d) L'erreur relative en fonction de $\epsilon_R$ . . . . .	40
Figure 3.3	Comparaison (a) de la précision et (b) de la régularité de l'approximation pour des adaptations en variations d'ordre $k = 0, 1, 2$ . . . . .	41
Figure 3.4	Discrétisation de l'équation de Van Der Pol avec $\Delta_2$ . (a) Solution discrète. (b) Ratio de contrôle. (c) Pas de temps en fonction du temps. (d) L'erreur relative en fonction du temps. . . . .	43
Figure 3.5	Discrétisation du système de Robertson avec $\Delta_2$ . (a) Solution discrète. (b) Ratio de contrôle. (c) Pas de temps en fonction du temps. (d) Erreur relative en fonction de $\epsilon_R$ . . . . .	45
Figure 3.6	Comparaison de la précision sur la solution discrète et sur la quantité d'intérêt $Q$ (a) selon une adaptation en fonction des variations de la solution et (b) en fonction de la quantité d'intérêt. . . . .	47
Figure 4.1	Discrétisation du problème de la plaque 1D avec une stratégie d'adaptation selon les variations de la perte AC. (a) Densité de courant $\mathbf{J}$ . (b) Perte AC. (c) Ratio de contrôle. (d) Pas de temps. . . . .	52
Figure 4.2	Discrétisation du problème de la plaque 1D avec une stratégie d'adaptation selon les variations des vecteurs $(\mathbf{E}, \mathbf{H})$ . (a) Densité de courant $\mathbf{J}$ . (b) Perte AC. (c) Ratio de contrôle. (d) Pas de temps. . . . .	53
Figure 4.3	Géométrie du problème de la modélisation câble supraconducteur. . .	55
Figure 4.4	Discrétisation du problème de câble supraconducteur 3D avec adaptation en temps selon la perte AC. (a) Ratio de contrôle et (b) pas de temps en fonction du temps. . . . .	56
Figure 4.5	Comparaison des pertes de courant AC avec une adaptation en temps selon les variations de la quantité d'intérêt et une résolution à pas constant $h = 10^{-4}$ . . . . .	57

## LISTE DES SIGLES ET ABRÉVIATIONS

### Abréviations

AC :	courant alternatif (« Alternating-current »)
BDF $k$ :	formule de différence arrière d'ordre $k$ (« Backward Difference Formula »)
DG :	Méthode de Galerkin discontinue (« Galerkin Discontinuous Method »)
FEM :	Méthode des éléments finis (« Finite Element Method »)
RK $k$ :	Méthode de Runge-Kutta d'ordre $k$
EDO :	Équation différentielle ordinaire
QoI :	Quantité d'intérêt (« Quantity of Interest »)

### Notations Mathématiques

$\mathbb{R}$ :	Corps des nombres réels
$\mathbb{C}$ :	Corps des nombres complexes
$\mathcal{M}_n(\mathbb{R})$ :	Espace des matrices carrées de taille $n$ à coefficients réels

### Géométrie

$\Omega$ :	Domaine du problème
$\partial\Omega$ :	Frontière du domaine
$h$ :	Pas de temps

### Variables

$J_c$ :	Densité de courant critique
$E_c$ :	Amplitude de courant électrique associée à $J_c$
$\rho$ :	Résistivité électrique
$\mu$ :	Perméabilité magnétique
$\mathbf{E}$ :	Champ électrique
$\mathbf{H}$ :	Champ magnétique
$\mathbf{J}$ :	Densité de courant

## CHAPITRE 1 INTRODUCTION

Les matériaux supraconducteurs sont des matériaux qui présentent deux propriétés caractéristiques en dessous d'une certaine température critique et d'un champ magnétique critique. La première est l'absence totale de résistance électrique et la seconde est l'expulsion du champ magnétique externe qui porte le nom d'effet Meissner. Selon leur comportement autour des températures critiques, deux types de supraconducteurs se distinguent. Les supraconducteurs de type I qui possèdent un unique champ magnétique critique et les supraconducteurs de type II qui en ont deux. Même si la théorie prévoit des températures critiques inférieures à 30 K, il est observé expérimentalement que certains supraconducteurs de type II possèdent un état supraconducteur à des températures pouvant aller jusqu'à 100 K. Il s'agit alors de supraconducteurs à haute température (HTS).

En théorie, ces supraconducteurs utilisés à haute température permettent d'augmenter le rendement des câbles électriques en transportant une plus grande quantité de courant sur une plus longue distance. Ceci facilite les connexions entre différentes infrastructures fournissant de l'électricité avec une consommation électrique plus intelligente et plus durable. De plus, les applications dans les domaines de l'ingénierie sont multiples, aussi bien dans les technologies d'IRM, en médecine, que pour la fabrication d'ordinateurs ultra-rapides ou encore de trains magnétiques. Les supraconducteurs présentent, pour ces raisons, un intérêt scientifique et industriel croissant, et c'est pourquoi cette étude s'intéresse à la discrétisation des équations de supraconductivité.

Puisque l'état supraconducteur dépend étroitement de la température du matériau, elle devient un paramètre de contrôle important. En effet, lorsqu'un composant supraconducteur est utilisé au-dessus de sa température critique, il perd progressivement son état supraconducteur entraînant, de facto, des pertes par effet Joule. La perte par effet Joule est la conversion indésirable, mais inévitable, d'énergie électrique en énergie thermique. Ces pertes énergétiques représentent une partie importante des coûts des infrastructures électriques, qu'il faut donc limiter. À cette fin, l'effet Joule est calculé à partir de la perte de courant alternatif. Cette quantité d'intérêt est donc importante pour aider les ingénieurs à concevoir des composants électriques performants. L'évaluation de la perte de courant alternatif nécessite le calcul de l'intégrale du produit scalaire entre le champ électrique et la densité de courant sur le domaine occupé par les matériaux conducteurs. En conséquence, la conception de composants supraconducteurs de pointe nécessite la simulation numérique des équations de Maxwell.

Le comportement supraconducteur peut être décrit par les équations de Maxwell avec la condition magnéto-quasistatique. Les équations de Maxwell, qui sont de nature hyperbo-

lique, peuvent ainsi être reformulées à l'aide de cette condition en un problème parabolique. Ces équations ne peuvent être résolues de manière générale. On a donc recours à la simulation numérique pour comprendre le comportement physique de ces matériaux.

Le développement de stratégies numériques pour simuler le comportement des composants supraconducteurs est complexe. À température constante, un conducteur standard possède une résistance électrique constante tandis qu'un supraconducteur à haute température dépend d'une résistivité électrique non linéaire souvent modélisée par une loi de puissance. Dans ce cas, les équations de Maxwell conduisent à une formulation appelée  $p$ -rotationnel. Le problème du  $p$ -rotationnel est un problème parabolique dégénéré et il est reconnu pour développer de forts gradients sur la densité de courant.

Une méthode numérique utilisée pour simuler ces équations est la méthode des éléments-finis avec une formulation dite de Galerkin. Cependant, un mauvais choix de discrétisation en temps des équations de Maxwell, dans le cadre de la méthode des éléments finis, peut conduire à une approximation de la solution qui est instable. En effet, les forts gradients développés par les équations et le mauvais conditionnement de la matrice d'assemblage obtenue conduisent à la discrétisation d'un problème dit raide en temps. Ceci nous contraint à choisir de petits pas de temps afin de garantir la stabilité du schéma numérique.

Dans le cadre de ce travail, un solveur éléments-finis tridimensionnel est mis à notre disposition. Il permet l'utilisation d'une formulation de Galerkin pour la discrétisation des équations elliptiques et paraboliques, ainsi que l'utilisation de la méthode de Galerkin discontinue pour la discrétisation des équations hyperboliques.

L'objectif principal de recherche est de développer une stratégie d'adaptation en temps basée sur un estimateur d'erreur pour la semi-discrétisation éléments-finis des équations de Maxwell pour l'étude de la supraconductivité.

De manière générale, une méthode numérique d'adaptation en temps est étudiée selon trois lignes directrices qui la caractérisent :

- La cohérence, lorsque l'erreur globale du schéma d'intégration en temps tend vers zéro à mesure que les pas de temps sélectionnés sont petits ;
- L'efficacité, qui prend en compte le temps et les ressources de calcul que la méthode d'adaptation en temps mobilise ;
- La stabilité, que nous entendrons ici comme étant la capacité de la méthode d'adaptation en temps à pouvoir s'appliquer à un large spectre d'équations différentielles.

À partir de ces points, on peut définir les objectifs spécifiques de cette étude :

- i) Démontrer un lien entre l'estimateur d'erreur et l'erreur de locale afin de prouver la cohérence de la stratégie d'adaptation en temps ;

- ii) Démontrer que la stratégie développée peut s'appliquer à une quantité d'intérêt physique pertinente pour une équation donnée ;
- iii) Vérifier la méthode développée sur des problèmes raides standards ;
- iv) Vérifier la stratégie d'adaptation en temps sur un problème d'électromagnétisme modélisant un matériau supraconducteur.

La méthodologie suivante permettra d'atteindre chacun des objectifs spécifiques :

- i) Une inégalité d'encadrement de l'erreur par l'estimateur d'erreur choisi est démontrée, donnant un cadre rigoureux à la stratégie d'adaptation ;
- ii) Un lien entre l'estimateur d'erreur sur la solution et l'estimateur d'erreur sur la quantité d'intérêt est exhibé, démontrant la cohérence de l'adaptation du pas de temps selon la quantité d'intérêt ;
- iii) Une liste non exhaustive d'équations raides, tirées de la littérature, permettent de vérifier la méthode d'adaptation en temps. Les caractéristiques pertinentes de la stratégie telles que l'efficacité et la stabilité sont mises en avant, permettant une éventuelle comparaison avec les méthodes déjà existantes ;
- iv) Le modèle mathématique sera vérifié à l'aide d'un problème d'induction d'un courant dans un câble électrique utilisant le diborure de magnésium ( $\text{MgB}_2$ ) comme supraconducteur à haute température.

## CHAPITRE 2 REVUE DE LITTÉRATURE

Dans ce chapitre, on donne une description détaillée mais non exhaustive des méthodes d'adaptation en temps pour la méthode des éléments finis. On choisit de présenter ce sujet à l'aide d'un exemple typique d'équation parabolique [40]. Ce choix n'est pas anodin car l'application étudiée au chapitre suivant peut être modélisée par une équation parabolique type. Un problème représentatif est l'équation de diffusion de la chaleur dans un milieu isotrope occupant un domaine  $\Omega \subset \mathbb{R}^n$  et s'écrit comme suit :

$$\begin{cases} \partial_t u + \mathcal{L}u = g & \text{dans } \Omega \times \mathbf{I}, \end{cases} \quad (2.1)$$

$$\begin{cases} u = 0 & \text{sur } \Gamma_1 \times \mathbf{I}, \end{cases} \quad (2.2)$$

$$\begin{cases} \mathbf{n} \cdot \nabla u = 0 & \text{sur } \Gamma_2 \times \mathbf{I}, \end{cases} \quad (2.3)$$

$$\begin{cases} u(x, 0) = u_0(x) & \forall x \in \Omega. \end{cases} \quad (2.4)$$

Ici,  $u(x, t)$  serait la température en un point  $x \in \Omega$  et à un temps  $t \in \mathbf{I} = ]0, T[$  où  $T$  est un temps final donné. L'opérateur  $\mathcal{L}$  est un opérateur différentiel linéaire, typiquement un laplacien. Le terme de droite  $g$  représente un terme source donné, tel que  $g \in L^2(\Omega)$  a priori. Les surfaces  $\Gamma_1$  et  $\Gamma_2$  sont des subdivisions de  $\partial\Omega$  telles que  $\overline{\Gamma_1} \cup \overline{\Gamma_2} = \partial\Omega$  et  $\Gamma_1 \cap \Gamma_2 = \emptyset$ , où  $\partial\Omega$  est la frontière de  $\Omega$ . Enfin,  $u_0$  est une condition initiale, qui représente ici la température à l'instant  $t = 0$ .

On considère, dans un premier temps, une semi-discrétisation de l'équation (2.1), que l'on ne discrétise qu'en espace à l'aide de la méthode des éléments finis [40]. On obtient alors un problème à valeur initiale pour un système d'équations différentielles ordinaires. Ce système peut être ensuite discrétisé à l'aide d'un schéma aux différences finies. Ceci peut donner lieu, dans certains cas, à un système dit raide dont on donnera une définition plus loin dans ce chapitre. L'étude de ce type de problème nécessite alors un choix adéquat de pas de temps pour obtenir une approximation numériquement stable. Cette problématique nous mène donc à introduire différentes stratégies d'adaptation en temps abordées dans la littérature.

Dans un second temps, on présente un autre type de discrétisation, en espace et en temps cette fois, en éléments finis. La discrétisation s'appuie sur une formulation variationnelle de l'équation (2.1) qui est un outil efficace pour construire des estimateurs d'erreur en temps. Cette seconde approche permet donc d'élaborer une toute autre variété de stratégies d'adaptation en temps à partir de ces estimateurs d'erreur obtenus en éléments finis.

Enfin, ce chapitre se conclut sur une comparaison des différentes approches d'adaptation en temps. On utilisera l'exemple des équations différentielles raides, qui est crucial dans la

suite des développements, afin de mener cette comparaison à terme. Cette partie a pour double objectif, de résumer, d'une part, les différentes approches présentées et, d'autre part, d'en dégager les avantages et les limitations face aux problèmes plus complexes tels que les problèmes raides.

## 2.1 Stratégies d'adaptation en temps en différences finies

La semi-discrétisation en espace de l'équation (2.1), par la méthode des éléments finis, permet d'approcher la solution  $u(\cdot, t) \in V = H^1(\Omega)$ , à  $t$  fixé dans  $I$ , dans un sous-espace discret  $V_h \subset V$ . On obtient donc une approximation éléments-finis en espace  $u_h(\cdot, t) \in V_h$ , de  $u$ . En fixant  $x \in \Omega$ , il nous reste alors à résoudre le problème de condition initiale :

$$\begin{cases} d_t u_h(t) + A(t)u_h(t) = g_h(t), & \forall t \in I \\ u_h(0) = u_0. \end{cases} \quad (2.5)$$

où  $A(t) \in \mathcal{M}_n(\mathbb{R})$  est la matrice d'assemblage (à  $t$  fixé) et  $g_h$  une projection du terme source sur l'espace  $V_h$ . On cherche alors  $t \mapsto u_h(t) \in \mathcal{C}^1([0, T])$  satisfaisant le problème (2.5)-(2.6). On est donc amené à résoudre un problème de condition initiale dont on rappelle la définition. On définit un problème de condition initiale d'ordre 1 comme la recherche d'une fonction  $u$  qui vérifie :

$$\begin{cases} F(t, u, \dot{u}) := \dot{u}(t) - f(t, u) = 0, \\ u(0) = u_0. \end{cases} \quad (2.7)$$

Afin de simplifier les notations et sans affecter le reste du raisonnement, nous supposons que la discrétisation en espace est exacte, c'est-à-dire que  $u = u_h$ . En d'autres termes, nous négligeons l'erreur éléments-finis dans un premier temps afin de ne présenter que les stratégies d'adaptation en temps. Lorsque  $f : V \subset \mathbb{R} \times \mathbb{R}^n \mapsto \mathbb{R}^n$  (où  $V$  est un ouvert de  $\mathbb{R} \times \mathbb{R}^n$ ) est continue et localement Lipschitz-continue par rapport à la seconde variable, d'après le théorème de Cauchy-Lipschitz [31], il existe U un ouvert de  $\mathbb{R}^n$  contenant  $u_0$ ,  $\tilde{I}$  un voisinage de 0, et une unique solution  $u \in \mathcal{C}^1(\tilde{I}, U)$ . On supposera dans la suite que  $I \subset \tilde{I}$ .

Bien que le théorème précédent nous donne l'existence et l'unicité d'une solution locale au problème (2.7), il est de manière générale difficile, voire impossible, d'exhiber une telle solution. On a donc recours, dans ce contexte, à une discrétisation à l'aide de schémas aux différences finies de (2.7). Il existe plusieurs types de schémas pour approximer la solution  $u$ . Par soucis de concision, nous ne garderons comme fil conducteur du développement dans ce chapitre, que la méthode d'Euler implicite ainsi que celles de Runge-Kutta. On peut cependant souligner que la plupart des techniques d'adaptation en temps s'étendent à des schéma numériques quelconques.



À partir de maintenant, considérons une subdivision régulière  $0 = t_0 < t_1 < \dots < t_N = T$  de l'intervalle fermé  $[0, T]$ . On notera pour tout  $n \in \{1, \dots, N\}$ ,  $u_n$  une approximation de  $u(t_n)$  sur l'intervalle de temps discret. Le schéma d'Euler implicite appliqué à l'équation (2.7) avec un pas de temps  $h := t_{n+1} - t_n$ , s'écrit

$$u_{n+1} = u_n + hf(t_{n+1}, u_{n+1}). \quad (2.8)$$

Les schémas d'intégration tel que le schéma d'Euler implicite fournissent une approximation de la solution exacte du problème (2.7). À chaque pas de temps, une erreur est donc commise entre la discrétisation de la solution et la solution du problème (2.7). On introduit alors les définitions des erreurs dites locale et globale.

**Définition 1** (Erreurs locale et globale).

*L'erreur d'un schéma de discrétisation est l'erreur due à l'approximation de l'équation (2.7) par une formule aux différences finies. On distingue alors*

- *l'erreur locale qui est l'erreur induite par un unique pas de temps à l'aide du schéma d'intégration numérique. On présuppose dans cette définition qu'aucune erreur d'approximation n'a été commise et que la solution exacte est connue aux pas de temps précédents. On désignera cette erreur par la lettre  $\tau$  ;*
- *L'erreur globale qui est la quantité totale d'erreur nécessaire pour obtenir l'approximation de la solution jusqu'au temps  $t_N$  :*

$$e_n = |u(t_N) - u_N|. \quad (2.9)$$

Dans ce qui suit, nous montrerons en quoi l'erreur locale constitue un élément clé des stratégies d'adaptation en temps.

Si la solution d'une équation différentielle ordinaire varie rapidement sur un certain intervalle de temps et plus lentement sur un autre intervalle, alors il est clair qu'utiliser un pas de temps constant est inefficace [53]. Pour ces méthodes, nous avons donc besoin d'un moyen d'estimer le pas de temps suivant qui permettra de contrôler l'erreur sur la solution. L'idée est que si l'on est capable d'estimer l'erreur due à la discrétisation du problème (2.7) au temps  $t_n$ , alors l'ordre de grandeur de cette erreur nous renseigne sur le choix du pas de temps suivant, i.e. s'il doit être accepté ou rejeté. Si le pas de temps est accepté, on recommence le même procédé avec le pas de temps suivant. Par conséquent, le but est d'élaborer des méthodes permettant d'obtenir un estimateur de l'erreur commise. Nous allons donc présenter dans cette section les différentes stratégies qui permettent de contrôler le pas de temps choisi à partir de prédictions de l'erreur commise par le schéma numérique choisi.

### 2.1.1 Méthode d'adaptation en temps par extrapolation de Richardson

Dans le cas d'une méthode d'adaptation en temps, le pas de temps  $h$  n'est plus constant et on note dans ce qui suit  $h_n$ , le  $n^{\text{ième}}$  pas de temps. Le premier algorithme de sélection automatique des pas de temps que l'on présente repose sur l'extrapolation de Richardson [53, 49]. L'idée à la base de l'extrapolation de Richardson est de calculer deux approximations différentes de la solution  $u$  dont les erreurs peuvent être développées en série de Taylor en fonction de  $h_n$ , à partir de l'équation (2.8). On recombine ensuite ces deux approximations pour en construire une nouvelle plus précise [19, 22].

On explore donc cette idée en comparant une approximation de  $u$  au temps  $t_{n+1}$ , en partant de la solution exacte au temps  $t_n$  avec un pas  $h_n$ , avec une autre approximation obtenue en faisant  $p$  pas de taille  $h_n/p$  à partir de  $t_n$ . Ici,  $p$  est un réel strictement supérieur à 1. Notre but est d'obtenir, à partir de ces deux approximations, une estimation de l'erreur locale et une estimation du pas de temps suivant  $h_{n+1}$ . Il faut ensuite déterminer un critère pour savoir si ce pas est accepté ou rejeté. Pour ce faire, on suppose que l'on dispose de la solution exacte au temps  $t_n$  et du pas de temps associé  $h_n$ . Dans un premier temps, on utilise le schéma d'Euler implicite pour obtenir une première approximation  $u_{n+1,1}$  avec le pas de temps  $h_n$ . Dans un second temps, on réutilise le même schéma en partant de  $t_n$  et en faisant  $p$  pas de temps de taille  $h_n/p$  pour construire une deuxième approximation  $u_{n+1,2}$ .

Lorsqu'on utilise le schéma d'Euler implicite, il existe une constante  $C > 0$  telle que la solution exacte vérifie l'équation de l'erreur locale

$$u(t_{n+1}) = u(t_n) + hf(t_{n+1}, u_{n+1}) + Ch_n^2 + \mathcal{O}(h_n^3). \quad (2.10)$$

On a donc pour la première approximation avec un pas  $h_n$ , l'erreur

$$u(t_{n+1}) - u_{n+1,1} = Ch_n^2 + \mathcal{O}(h_n^3), \quad (2.11)$$

et, pour la deuxième approximation avec un pas  $h_n/p$

$$u(t_{n+1}) - u_{n+1,2} = C \left( \frac{h_n^2}{p} \right) + \mathcal{O}(h_n^3). \quad (2.12)$$

En suivant la procédure de l'extrapolation de Richardson, on recombine ces deux solutions de manière à annuler les termes en  $h_n^2$  et à gagner un ordre de précision. On obtient une nouvelle approximation, au temps  $t_{n+1}$ , plus précise.

Supposons que l'utilisateur fixe une tolérance, arbitraire ou non, sur la variation maximale admissible de l'erreur, notée  $TOL$ . À partir de l'extrapolation de Richardson précédente, on

dispose en plus d'un indicateur  $r_n$  de cette variation, donné par le ratio

$$r_n = \frac{|u_{n+1,2} - u_{n+1,1}|}{h_n}. \quad (2.13)$$

On veut donc que  $r_n \leq TOL$ . Dans le cas où ce critère est respecté, on accepte l'approximation plus précise

$$u_{n+1}^* = \frac{pu_{n+1,2} - u_{n+1,1}}{p - 1}. \quad (2.14)$$

Dans le cas contraire, il faut réduire la taille de  $h_n$ . On estime le pas suivant à partir de la fraction actuelle de  $h_n$ , de l'estimateur d'erreur  $r_n$  et de la tolérance  $TOL$ . On peut également introduire un facteur multiplicatif de sécurité ad hoc,  $\gamma < 1$ . Une possibilité pour mettre à jour  $h_n$  s'écrit alors comme

$$h_{n+1} = \gamma \left( \frac{TOL}{r_n} \right) h_n, \quad \text{où } \gamma < 1. \quad (2.15)$$

De cette expression, on remarque que si la tolérance est plus petite que l'estimateur  $r_n$ , alors le pas de temps  $h_n$  est réduit, ce qui est cohérent avec l'objectif fixé.

### 2.1.2 Méthodes de prédiction-correction

On s'intéresse à présent aux algorithmes de prédiction-correction. Il s'agit d'une classe particulière de méthodes numériques développées, notamment par Gresho et al. [23, 24, 42], pour le calcul à pas de temps multiples et donnant accès à un estimateur de l'erreur à faible coût. Les algorithmes de prédiction-correction se décomposent, comme leur nom l'indique, en deux étapes.

- L'étape initiale dite de « prédiction », dans laquelle sont interpolées les valeurs déjà connues (ou du moins dont on dispose d'une approximation) de la fonction inconnue recherchée, ainsi que celles de ses dérivées. Le polynôme obtenu permet alors d'extrapoler la solution à un temps subséquent.
- La seconde étape, celle de la « correction », permet de raffiner l'approximation en utilisant le résultat à l'étape de prédiction et une autre méthode pour interpoler la solution au même temps subséquent.

Il existe principalement deux familles de méthodes numériques qui se prêtent aux techniques de prédiction-correction. La première stratégie utilise les formules d'Adams-Bashforth et celles d'Adam-Moulton [53]. L'idée est de combiner les premières avec les secondes. En effet, il s'agit simplement d'utiliser les schémas d'Adams-Bashforth pour obtenir une première approximation  $u_{n+1}^p$  de  $u_{n+1}$ , qui est l'étape de prédiction indiquée par la lettre p. On fait appel

ensuite aux formules d'Adams-Moulton pour corriger et éventuellement améliorer cette approximation. La seconde stratégie consiste à utiliser les formules de différences arrières, notée BDF (« *Backward-Difference Formula* ») dans la suite.

Ces deux familles de méthodes permettent de développer les mêmes techniques d'adaptation en temps [23, 53]. Pour des raisons de concision dans l'argumentation et afin d'éviter toute redondance dans la suite de cette section, nous ne développerons que les méthodes de différences arrières pour les techniques de prédiction-correction. En outre, ce sont les schémas BDF qui seront utilisés dans les développements des chapitres suivants.

Une des possibilités pour construire des approximations de haut ordre est donc d'utiliser des méthodes à pas multiples. Ces méthodes se prêtent également à l'adaptation en temps. Les schémas de différentiation arrière (BDF), notamment développés par Brenan et al. [10], font l'objet d'un intérêt particulier pour la discrétisation en temps d'équations différentielles. La construction de ces schémas repose sur l'utilisation du polynôme de Newton  $\pi_k$  de degré  $k$ , qui interpole les points  $(h_i, u_i)$  pour  $i \in \{n+1-k, \dots, n+1\}$ ,

$$\begin{aligned} \pi_k(t) = & u_{n+1} + (t - t_{n+1})[u_{n+1}, u_n] + (t - t_{n+1})(t - t_n)[u_{n+1}, u_n, u_{n-1}] + \dots \\ & + (t - t_{n+1}) \dots (t - t_{n+1-k})[u_{n+1}, \dots, u_{n+1-k}], \end{aligned} \quad (2.16)$$

où les coefficients du polynôme sont des différences divisées définies par la formule de récurrence :

$$[u_n] := u_n, \quad (2.17)$$

$$[u_{n+k}, \dots, u_n] := \frac{[u_{n+k}, \dots, u_{n+1}] - [u_{n+k-1}, \dots, u_n]}{t_{n+k} - t_n}. \quad (2.18)$$

On s'assure ensuite que le polynôme (2.16) vérifie l'équation différentielle (2.7) au temps  $t_{n+1}$  :

$$\dot{\pi}_k(t_{n+1}) = f(t_{n+1}, u_{n+1}). \quad (2.19)$$

On génère de cette manière les méthodes BDF d'ordre  $k$  pour approximer l'équation (2.7). Pour  $k = 1$ , la méthode BDF est identique au schéma d'Euler implicite [17]. Nous allons donc, dans la suite, nous focaliser sur le cas particulier des schémas BDF d'ordre 2. La stratégie d'adaptation présentée ci-après est cependant généralisable aux schémas BDF d'ordre  $k$  quelconque.

Pour utiliser le schéma BDF d'ordre 2 de façon à obtenir un algorithme d'adaptation en temps de type prédiction-correction, il faut décomposer la discrétisation de l'équation (2.19) en deux étapes. On rappelle que l'on cherche l'approximation  $u_{n+1}$  ainsi que le pas de temps  $h_{n+1}$  et que l'on suppose que les approximations aux pas de temps précédents sont connues.

La première étape consiste à construire le polynôme de prédiction  $\pi_P$  interpolant les points  $(u_n, u_{n-1}, u_{n-2})$  suivant :

$$\pi_P(t) := u_n + (t - t_n)[u_n, u_{n-1}] + (t - t_n)(t - t_{n-1})[u_n, u_{n-1}, u_{n-2}]. \quad (2.20)$$

En utilisant (2.20), une première approximation de  $u(t_{n+1})$  est alors obtenue en extrapolant au temps  $t_{n+1}$

$$\pi_P(t_{n+1}) = u_{n+1}^P. \quad (2.21)$$

La deuxième étape permet de corriger la valeur de  $u_{n+1}^P$  en utilisant le polynôme de correction qui interpole cette fois les trois valeurs  $(u_{n+1}, u_n, u_{n-1})$ . Ici,  $u_{n+1}$  est la valeur inconnue recherchée.

$$\pi_C(t) := u_{n+1} + (t - t_{n+1})[u_{n+1}, u_n] + (t - t_{n+1})(t - t_n)[u_{n+1}, u_n, u_{n-1}]. \quad (2.22)$$

Le polynôme de correction doit également vérifier l'équation (2.7) au temps  $t_{n+1}$  :

$$\dot{\pi}_C(t_{n+1}) = f(t_{n+1}, u_{n+1}). \quad (2.23)$$

En utilisant les relations (2.20) à (2.23), on peut montrer, en calculant la différence  $\dot{\pi}_C(t_{n+1}) - \dot{\pi}_P(t_{n+1})$  que la valeur de  $u_{n+1}$  peut être obtenue en résolvant directement l'équation suivante :

$$f(t_{n+1}, u_{n+1}) - \dot{\pi}_P(t_{n+1}) - \alpha_{n+1}(u_{n+1} - u_{n+1}^P) = 0, \quad (2.24)$$

où

$$\alpha_{n+1} = \frac{1}{t_{n+1} - t_n} + \frac{1}{t_{n+1} - t_{n-1}}. \quad (2.25)$$

Notre objectif principale est de déterminer un estimateur de l'erreur locale qui nous donne accès à une stratégie d'adaptation en temps. En notant  $\tau_n$  l'erreur locale après avoir effectué le pas de temps  $h_n$ , on peut montrer que le schéma BDF2 possède une erreur  $\tau_n$  d'ordre 3. On a donc :

$$\tau_n = |u(t_{n+1}) - u_{n+1}| = \mathcal{O}(h_n^3) \quad (2.26)$$

Le ratio des erreurs aux temps  $t_{n+1}$  et  $t_{n+2}$  peut donc être relié au ratio des pas de temps correspondants :

$$\frac{\tau_{n+1}}{\tau_n} \simeq \left( \frac{h_{n+1}}{h_n} \right)^3. \quad (2.27)$$

On suppose que l'utilisateur cherche à limiter l'erreur  $\tau_{n+1}$  en fonction d'une certaine tolérance, arbitraire ou non, notée  $TOL$ . Il ne nous reste plus qu'à évaluer la valeur de  $\tau_n$  afin

d'obtenir une relation entre les pas de temps à partir de l'équation (2.27). Bien que l'on ne dispose pas de la valeur exacte de  $u(t_{n+1})$ , c'est à ce niveau que l'algorithme de prédiction-correction entre en jeu. En effet, on dispose à présent de deux approximations de  $u(t_{n+1})$ , une valeur de prédiction  $u_{n+1}^p$  et une valeur de correction plus précise  $u_{n+1}$ . On peut alors légitimement approximer l'erreur  $\tau_n$  de la manière suivante :

$$\tau_n \simeq |u_{n+1} - u_{n+1}^p|. \quad (2.28)$$

Finalement, la relation qui sous-tend la stratégie d'adaptation sur les pas de temps pour les méthodes de prédiction-correction dans le cas du schéma BDF2 est donnée par

$$h_{n+1} = \left( \frac{TOL}{\tau_n} \right)^{\frac{1}{3}} h_n. \quad (2.29)$$

### 2.1.3 Méthodes de Runge-Kutta emboîtées

La stratégie d'adaptation sur les pas de temps avec les méthodes de Runge-Kutta (RK) emboîtées consiste à utiliser deux schémas différents (de la famille des méthodes RK) pour approximer la solution au temps  $t_{n+1}$ . La comparaison de ces deux approximations permet de construire une estimation de l'erreur locale. Si les résultats sont proches selon un critère prédéterminé, on est alors confiant sur le choix du pas de temps. Dans le cas contraire, le pas de temps est trop grand et il doit être réduit [53, 68].

Une combinaison populaire dans la littérature des méthodes RK est due à Fehlberg [20]. Dans cette section, nous allons la décrire, ainsi que l'adaptation en temps qui en découle. On définit d'abord les 6 constantes suivantes :

$$\begin{aligned} k_1 &= h_n f(t_n, u_n), \\ k_2 &= h_n f\left(t_n + \frac{1}{4}h_n, u_n + \frac{1}{4}k_1\right), \\ k_3 &= h_n f\left(t_n + \frac{3}{8}h_n, u_n + \frac{3}{32}k_1 + \frac{9}{32}k_2\right), \\ k_4 &= h_n f\left(t_n + \frac{12}{13}h_n, u_n + \frac{1932}{2197}k_1 - \frac{7200}{2197}k_2 + \frac{7296}{2197}k_3\right), \\ k_5 &= h_n f\left(t_n + h_n, u_n + \frac{439}{216}k_1 - 8k_2 + \frac{3680}{513}k_3 - \frac{845}{4104}k_4\right), \\ k_6 &= h_n f\left(t_n + \frac{1}{2}h_n, u_n - \frac{8}{27}k_1 + 2k_2 - \frac{3544}{2565}k_3 + \frac{1859}{4104}k_4 - \frac{11}{40}k_5\right). \end{aligned}$$

À l'aide de ces constantes  $k_i$ , on construit une première méthode de Runge-Kutta d'ordre 4 (RK4) qui nous donne une première approximation de  $u(t_{n+1})$  :

$$u_{n+1} = u_n + \left( \frac{25}{216}k_1 + \frac{1408}{2565}k_3 + \frac{2197}{4104}k_4 - \frac{1}{5}k_5 \right) = u_n + h_n \phi(t_n, u_n), \quad (2.30)$$

et une deuxième d'ordre 5 (RK5), qui nous donne une seconde approximation de  $u(t_{n+1})$  :

$$\tilde{u}_{n+1} = \tilde{u}_n + \left( \frac{16}{135}k_1 + \frac{6656}{12825}k_3 + \frac{28561}{56430}k_4 - \frac{9}{50}k_5 + \frac{2}{55}k_6 \right) = u_n + h_n \tilde{\phi}(t_n, u_n). \quad (2.31)$$

Les fonctions  $\phi$  et  $\tilde{\phi}$  sont introduites afin de simplifier les notations. On obtient donc deux approximations d'ordre différent de la même quantité, à savoir  $u(t_{n+1})$ . Cette combinaison de deux schémas a l'avantage de ne nécessiter que le calcul de six valeurs  $k_i$  au total, et ce, pour les deux approximations.

La méthode de référence sera RK4 et on utilise la méthode RK5 pour contrôler la valeur du pas de temps  $h_n$  et, en conséquence, pour contrôler l'erreur. Nous supposons, dans le reste du développement, que l'algorithme est initialisé à une valeur  $u_n$  exacte, c'est-à-dire  $u(t_n) = u_n$ . Cette hypothèse est importante car elle suppose qu'aucune erreur n'a été accumulée jusqu'au temps  $t_n$ . Puisqu'on a supposé que  $u_n = u(t_n)$ , on a alors que

$$\begin{aligned} u(t_{n+1}) - u_{n+1} &= u(t_{n+1}) - u_n - h_n \phi(t_n, u_n) \\ &= u(t_{n+1}) - u(t_n) - h_n \phi(t_n, u(t_n)) \\ &= h_n \tau_{n+1}(h_n) = \mathcal{O}(h_n^5), \end{aligned}$$

où  $\tau_{n+1}$  est l'erreur locale au temps  $t_{n+1}$  et on rappelle que  $\tau_{n+1} = \mathcal{O}(h_n^4)$  pour la méthode RK4. De même, on peut montrer que

$$u(t_{n+1}) - \tilde{u}_{n+1} = h_n \tilde{\tau}_{n+1}(h_n) = \mathcal{O}(h_n^6).$$

On rappelle également que  $\tilde{\tau}_{n+1} = \mathcal{O}(h_n^5)$  pour la méthode RK5. On en déduit alors :

$$u(t_{n+1}) - u_{n+1} = (u(t_{n+1}) - \tilde{u}_{n+1}) + (\tilde{u}_{n+1} - u_{n+1}),$$

et donc que

$$\tau_{n+1}(h_n) = \tilde{\tau}_{n+1}(h_n) + \frac{\tilde{u}_{n+1} - u_{n+1}}{h_n}. \quad (2.32)$$

En évaluant les ordres de grandeur de part et d'autre de l'égalité (2.32), on peut écrire que

$$\mathcal{O}(h_n^4) = \mathcal{O}(h_n^5) + \frac{\tilde{u}_{n+1} - u_{n+1}}{h_n}.$$

On peut alors faire l'approximation

$$\tau_{n+1}(h_n) \simeq \frac{\tilde{u}_{n+1} - u_{n+1}}{h_n} = \mathcal{O}(h_n^4),$$

qui est compatible avec l'égalité précédente.

Le terme dominant dans l'erreur commise dépend donc de la différence entre les approximations d'ordre 4 et 5. On note  $E$  cette approximation de l'erreur qui vérifie alors l'égalité

$$E = \frac{\tilde{u}_{n+1} - u_{n+1}}{h_n} = \frac{1}{h_n} \left( \frac{1}{360}k_1 - \frac{128}{4275}k_3 - \frac{2197}{75240}k_4 + \frac{1}{50}k_5 + \frac{2}{55}k_6 \right). \quad (2.33)$$

Cette valeur joue un rôle central dans le contrôle de l'erreur de locale. Il s'agit en effet de l'estimateur d'erreur que l'on obtient avec la méthode de Fehlberg. Supposons que l'utilisateur dispose d'une tolérance  $TOL$ , arbitraire ou non. Il suffit alors d'utiliser l'estimateur (2.33) et de modifier le pas  $h_n$  de manière à obtenir la précision recherchée. Pour y arriver, rappelons que

$$\tau_{n+1}(h_n) \simeq h_n^4 \simeq E.$$

On note  $\beta$  le paramètre de réduction (ou d'augmentation) du pas de temps  $h_n$ . En remplaçant alors le pas  $h_n$  par le pas  $\beta h_n$ , on constate que

$$\tau_{n+1}(\beta h_n) \simeq (\beta h_n)^4 = TOL.$$

Finalement, en isolant la valeur du paramètre  $\beta$ , on a

$$\beta \simeq \left( \frac{TOL}{E} \right)^{\frac{1}{4}}. \quad (2.34)$$

Ceci nous conduit à la relation d'adaptation sur les pas de temps :

$$h_{n+1} = \left( \frac{TOL}{E} \right)^{\frac{1}{4}} h_n. \quad (2.35)$$

#### 2.1.4 Méthode d'adaptation avec un contrôleur proportionnel-intégral (PI)

Les sections précédentes ont permis de montrer que les algorithmes standards de contrôle du pas de temps sont de la forme

$$h_{n+1} = \gamma \left( \frac{TOL}{r_n} \right)^{\frac{1}{k}} h_n. \quad (2.36)$$

Le paramètre  $h_n$  est le pas de temps,  $r_n$  est l'estimation de l'erreur,  $TOL$  est la tolérance spécifiée par l'utilisateur,  $\gamma$  est un facteur de sécurité réduisant le risque qu'un pas de temps soit rejeté et  $k$  est typiquement lié à l'ordre de précision du schéma d'intégration. Bien que ces algorithmes donnent des résultats satisfaisants en général, il existe des cas pathologiques. Ces cas apparaissent lorsque la stabilité, plutôt que la précision, devient le paramètre qui nous force à limiter le pas de temps. Ceci résulte souvent en des courbes d'évolution du pas de



temps qui oscillent fortement. En outre, cette perte de régularité peut entraîner une perte de précision au niveau de l'estimateur d'erreur, dans l'algorithme d'adaptation [28]. Ceci mène à des approximations imprécises. En améliorant l'algorithme de contrôle de pas temps, on cherche à contrôler l'apparition de ces artéfacts numériques, pour ainsi obtenir de meilleurs résultats lors des simulations.

La boucle de contrôle de temps peut être vue comme un contrôleur proportionnel-intégral (PI). On peut en effet la décomposer en une étape de contrôle et une étape dite d'intégration (cf. figure 2.1). Ce rapprochement a été fait par Gustafsson et al. et Soderlind [29, 30, 28, 64]. Le processus a une entrée : le pas de temps  $h_n$ , et deux sorties : la solution de l'équation différentielle  $u$  et l'estimateur d'erreur  $r$ . La partie de contrôle se doit de garder la sortie  $r$  aussi proche que possible de la tolérance  $TOL$ , fixée par l'utilisateur, en utilisant le pas de temps comme variable de contrôle. Une idée naturelle pour améliorer cet algorithme est de modifier le contrôleur pour le rendre plus efficace. En utilisant des techniques élémentaires

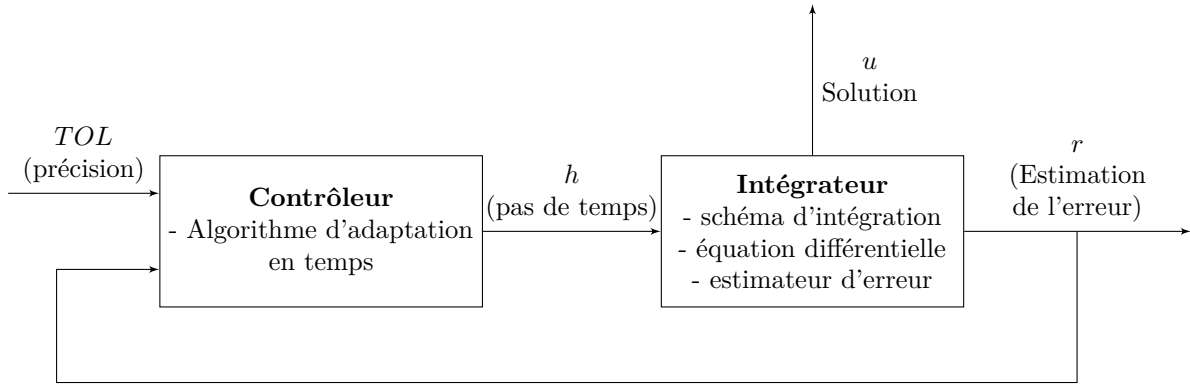


Figure 2.1: Adaptation du pas de temps modélisé par un système de contrôle proportionnel-intégral [28].

venant de la théorie du contrôle de systèmes dynamiques, on introduit un contrôleur de la forme

$$h_{n+1} = \left( \frac{TOL}{r_{n+1}} \right)^{k_I} \left( \frac{r_n}{r_{n+1}} \right)^{k_P} h_n, \quad (2.37)$$

où les paramètres  $k_I$  et  $k_P$  dépendent du schéma d'intégration choisi. Ainsi, ce contrôleur a pour effet de limiter le problème des oscillations de la variable dépendante en temps, mais il permet également d'obtenir des séquences de pas de temps beaucoup plus régulières. En conséquence, l'estimateur d'erreur gagne également en régularité.

De manière générale, les valeurs empiriques  $k_P = 0,075$  et  $k_I = 0,175$  semblent donner des résultats satisfaisants. Cependant, le choix des paramètres du contrôleur reste heuristique et ne peut être généralisé pour une EDO quelconque.

## 2.2 Stratégies d'adaptation en temps basées sur une formulation variationnelle

Comme évoqué au début du chapitre, une discrétisation éléments-finis en espace et en temps permet de développer d'autres types de stratégies d'adaptation en temps. En effet, la formulation variationnelle de l'équation (2.1) permet d'obtenir des estimateurs d'erreur en tirant avantage des inégalités issues du calcul des variations. C'est une des raisons pour laquelle cette voie est explorée par de nombreux auteurs [40, 39, 18, 66, 12, 15, 14, 58]. Nous avons notamment vu, dans les sections précédentes, que pour chaque estimateur d'erreur présenté, on peut associer une stratégie d'adaptation en temps.

Ainsi, le problème initial (2.1) peut être réécrit comme la formulation faible équivalente : trouver  $u : \mathbb{R}_+ \rightarrow H^1(\Omega)$  telle que :

$$\begin{cases} (\dot{u}(t), v) + (\mathcal{L}u(t), v) = (g(t), v), & \forall v \in H_0^1(\Omega), \\ u(0) = u_0, \end{cases} \quad (2.38)$$

où  $(\cdot, \cdot)$  dénote le produit scalaire usuel sur  $L_2(\Omega)$  et  $\dot{u} = \partial u / \partial t$ .

Pour la discrétisation en temps, on considère la méthode de Galerkin discontinue (voir [16, 35]) définie comme suit : soit  $0 = t_0 < t_1 < \dots < t_n < \dots$  une partition de l'axe du temps en sous-intervalles  $I_n = ]t_{n-1}, t_n]$  et soit  $q$  un entier positif non nul associé à l'espace des fonctions polynomiales par morceaux de degré au plus  $q$  en  $t$ , à valeurs dans  $H_0^1(\Omega)$  :

$$W = \left\{ v : v|_{I_n} = \sum_{i=0}^q a_{i,n} t^i, \ a_{i,n} \in H_0^1(\Omega), \text{ pour } i = 0, 1, \dots, q, \text{ et } n = 1, 2, \dots \right\}. \quad (2.39)$$

Avec la méthode de Galerkin discontinue, les fonctions tests et la solution éléments-finis, notée  $U \in W$ , sont continue sur chaque intervalle  $I_n$ . À chaque nœud  $t_n$ , on adopte les notations suivantes :

$$U_n = U(t_n) = \lim_{s \rightarrow t_n^-} U(s) \quad \text{et} \quad U_n^+ = \lim_{s \rightarrow t_n^+} U(s). \quad (2.40)$$

Les valeurs de saut aux interfaces sont notées par  $[U]_n = U_n(t_n^+) - U_n(t_n)$ .

Comme  $U$  est continue par morceaux, la dérivée en temps fait intervenir des distributions de Dirac aux points de discontinuité. Alors,  $U$  est également solution de l'équation :

$$\sum_{n=0}^{N-1} \int_{I_n} (\dot{U}(t), v) dt + \sum_{n=0}^{N-1} [U]_{n+1} v_{n+1}^+ + \sum_{n=0}^{N-1} \int_{I_n} (\mathcal{L}U(t), v) dt = \sum_{n=0}^{N-1} \int_{I_n} (g, v) dt, \quad (2.41)$$

pour tout  $v \in W$ . En pratique,  $U$  peut être déterminée localement sur chaque sous-intervalle  $I_n$  :

$$\begin{cases} \int_{I_n} (\dot{U}(t), v) dt + (U_{n+1}^+ - U_{n+1}) v_{n+1}^+ + \int_{I_n} (\mathcal{L}U(t), v) dt = \int_{I_n} (g, v) dt, \\ U(0^+) = u_0. \end{cases} \quad (2.42)$$

En particulier, pour  $q = 0$  et en utilisant un schéma d'Euler implicite en temps, le problème variationnel (2.42) se réécrit comme suit : Trouver  $U \in H^1(\Omega)$  telle que

$$\begin{cases} (U_n - U_{n-1}, v) + h_n (\mathcal{L}U_n, v) = \int_{I_n} (g, v) dt, \quad \forall v \in H_0^1(\Omega), \quad n = 1, 2, \dots \\ U(0^+) = u_0. \end{cases} \quad (2.43)$$

### 2.2.1 Estimateur d'erreur sur la solution

La formulation variationnelle (2.43) permet de construire un estimateur de l'erreur pour la méthode de Galerkin discontinue. Dans le cas où  $q = 0$ , i.e. pour la formulation variationnelle avec le schéma d'Euler implicite (2.43), sous des hypothèses adéquates sur la régularité de  $u$ , Eriksson et al. [18] montrent l'inégalité

$$\max_{t \in I} \|u(t) - U(t)\|_2 \leq C \left( \max_{n \leq N} h_n \|\dot{u}\|_\infty^{I_n} + \max_{t \in I} k^2 \|u(t)\|_{H^2(\Omega)} \right). \quad (2.44)$$

Le premier terme dans le membre de droite de l'inégalité (2.44) représente l'erreur de discrétisation en temps. Le deuxième terme représente quant à lui l'erreur de discrétisation en espace. On ne s'intéresse ici qu'à l'erreur en temps.

Supposons que l'on dispose d'une tolérance  $TOL$  et que l'erreur de discrétisation en temps de l'équation (2.43) doit être bornée par  $TOL$ , quelque soit  $t \in I$ . On en déduit d'après l'inégalité (2.44) que le terme d'erreur est contrôlé si

$$h_n \|\dot{u}\|_\infty^{I_n} \leq \frac{TOL}{C}, \quad \forall n \in \{1, \dots, N\}, \quad (2.45)$$

où  $C$  est la constante introduite dans l'équation (2.43). On constate donc, à l'aide de l'estimateur d'erreur, que la stratégie d'adaptation en temps est basée sur le terme  $\|\dot{u}\|_\infty^{I_n}$ . Bien sûr,  $\|\dot{u}\|_\infty^{I_n}$  est inconnue a priori. On peut cependant utiliser l'approximation issue d'une formule de différence arrière

$$h_n \|\dot{u}\|_\infty^{I_n} \approx \|U_n - U_{n-1}\|_2. \quad (2.46)$$

On dispose alors du critère de sélection de pas de temps

$$\|U_n - U_{n-1}\|_2 \approx \frac{TOL}{C}. \quad (2.47)$$

L'idée de la stratégie d'adaptation est alors de satisfaire à chaque pas de temps le critère (2.47) par essai-erreur. Supposons que l'on dispose de la valeur  $U_{n-1}$ . Une première hypothèse serait de prendre le même pas de temps qu'à l'étape précédente, c'est-à-dire  $h_n = h_{n-1}$ . On calcule ensuite le terme  $U_n$  correspondant. Si le critère  $\|U_n - U_{n-1}\|_2$  est suffisamment proche de la borne  $\frac{TOL}{C}$ , alors le pas de temps  $h_n$  est accepté. Dans le cas contraire,  $h_n$  est réduit ou augmenté de manière à obtenir  $\|U_n - U_{n-1}\|_2 \approx \frac{TOL}{C}$ . Une variante de cette procédure est également possible [39]. Par exemple, on peut choisir comme estimateur initial du pas de temps, le pas

$$h_n = h_{n-1} \left( \frac{TOL}{C\|U_{n-1} - U_{n-2}\|_2} \right), \quad (2.48)$$

que l'on obtient à partir de l'équation (2.46).

Sous des hypothèses adéquates de régularité sur  $U$  [18], on peut montrer que l'approximation (2.46) est valide. Il est donc possible de prouver que si le critère (2.47) est satisfait pour chaque pas de temps, alors l'erreur de discrétisation en temps de l'équation (2.43) est bornée par  $TOL$ .

### 2.2.2 Estimation d'erreur basée sur une quantité d'intérêt

Nous l'avons vu jusqu'à présent, les estimateurs présentés permettent d'obtenir un calcul de l'erreur de discrétisation pour une approximation particulière donnée [13]. Or, la précision du contrôle de l'erreur de discrétisation est un élément critique lors de simulations numériques. En effet, nous avons pu le constater dans les différentes stratégies d'adaptation présentées, toutes basées sur un estimateur de l'erreur. L'efficacité de ces stratégies dépend donc étroitement de la qualité de ce dernier.

Cependant, le but des simulations numériques n'est le plus souvent pas d'obtenir une approximation de la solution  $u$ , mais une approximation d'une quantité qui dépend de  $u$ . Cette quantité porte le nom de quantité d'intérêt et peut souvent être représentée comme une fonctionnelle linéaire de  $u$ ,

$$Q(u) = (u(T), \psi), \quad (2.49)$$

où  $\psi$  est élément d'un certain espace de Hilbert  $V$ . Le produit scalaire canonique sur  $V$  est alors dénoté par  $(\cdot, \cdot)$ .

L'idée à la base de l'adaptation en temps selon une quantité d'intérêt repose donc sur un changement de paradigme. Puisque la valeur pertinente recherchée est celle de la quantité d'intérêt, l'adaptation en temps se fait donc en fonction d'un estimateur de l'erreur sur cette quantité d'intérêt. L'objectif est alors de garantir une grande précision, non plus sur la solution, mais sur la quantité  $Q$  dépendante de la solution. Il en résulte des valeurs physiquement plus pertinentes et donc plus intéressantes du point de vue de l'application étudiée. Cette

approche génère pour ces raisons un intérêt croissant [13, 66, 43, 34, 3].

### 2.3 Problèmes raides

Il existe certaines classes d'équations plus difficiles à discrétiser numériquement que d'autres. De façon heuristique, on peut s'attendre à ce que pour une solution dont la variation est importante dans une région donnée, les pas de temps doivent être choisis plus petits. Si, en revanche, la solution change très peu, alors les pas de temps peuvent être plus grand pour plus d'efficacité. Cependant, certains problèmes nécessitent de très petits pas de temps, même si la solution semble a priori suffisamment lisse sur son domaine. Il n'existe pas de définition universellement admise de ce que l'on qualifie d'équation différentielle raide [32]. Un problème raide peut néanmoins être identifié à l'aide de certaines caractéristiques remarquables. Typiquement, le graphe de la solution change rapidement sur une certaine zone avant de tendre vers une solution à variations plus lentes. Une équation raide se distingue aussi par le fait que, lorsque l'on adapte en temps, ce n'est pas la précision qui domine pour choisir le pas de temps mais la stabilité du schéma numérique [32]. En conséquence, une autre caractéristique des équations différentielles raides est que, pour les discrétiser, les schémas implicites, en particulier les méthodes BDF, sont beaucoup plus efficaces que les schémas explicites.

Afin d'illustrer ce concept nous reprenons l'exemple du modèle de combustion donné par Shampine et al. [59]. Le problème de condition initiale s'écrit comme :

$$\begin{cases} u'(t) = u^2(1 - u), & 0 < t \leq \frac{2}{\delta}, \\ u(0) = \delta, \end{cases} \quad (2.50)$$

où  $\delta \ll 1$  est un paramètre donné. La solution  $u$  a comme condition initiale  $\delta$  et croît brusquement jusqu'à tendre vers la valeur maximale de 1. La variation rapide se produit sur l'intervalle  $[0, 1/\delta]$ , alors que sur l'intervalle  $[1/\delta, 2/\delta]$ , elle est approximativement égale à un. Plus généralement, pour les systèmes différentiels linéaires tels que le problème (2.5), la raideur se manifeste au niveau des valeurs propres de la matrice jacobienne  $A$ . Si pour cette matrice, il coexiste à la fois des valeurs propres avec des parties réelles de grande magnitude et d'autres avec des parties réelles de faible magnitude, alors le problème est dit raide. L'équation (2.5) devient encore plus difficile à discrétiser car la matrice  $A$  est de très grande taille avec, potentiellement, un mauvais conditionnement.

#### 2.3.1 Discussion sur les stratégies d'adaptation en temps

Dans cette section, nous revenons sur les stratégies d'adaptation en temps présentées afin de comparer leur efficacité lorsqu'elles sont utilisées pour la discrétisation de l'équation (2.5).

Nous avons expliqué plus haut en quoi cette équation (2.5) peut être classée parmi les problèmes dits raides. On étend alors naturellement les résultats et les analyses faites sur l'équation (2.5) à des problèmes raides. De plus, bien que le problème (2.50) soit différent en plusieurs points de l'équation différentielle (2.5), il est cependant un exemple simple d'équation raide. L'équation (2.50) constitue donc également un point de référence pour comparer les techniques d'adaptation en temps.

La plupart des méthodes d'adaptation que nous avons présentées reposent sur un estimateur de l'erreur. En effet, l'idée est de choisir chaque nouveau pas de temps de manière à contrôler l'erreur, ou tout du moins une approximation de cette erreur. C'est le cas entre autres des méthodes d'extrapolation de Richardson et plus généralement du contrôleur proportionnel-intégral. Bien que ces techniques d'adaptation ont fait leurs preuves, y compris pour des problèmes raides, elles possèdent une limitation importante. De manière générale, l'obtention d'un estimateur de l'erreur est souvent un problème très coûteux en terme de temps de calcul. En effet, dans le cas de l'extrapolation de Richardson, la méthode devient de plus en plus coûteuse à mesure que les pas de temps  $h_n$  deviennent petits. En outre, s'ajoute à cela que dans l'équation (2.5), la matrice  $A$  est de grande taille. Ceci rend alors impraticable le problème d'estimation de l'erreur locale à chaque pas de temps.

Le coût de calcul ne peut néanmoins pas être avancé comme seul facteur limitant des stratégies d'adaptation en temps. En effet, les méthodes de Runge-Kutta emboîtées fournissent un estimateur d'erreur précis pour un faible coût de calcul. Cependant, Shampine et al. montrent que la méthode de Runge-Kutta-Felbergh échoue à représenter la solution de l'équation (2.50) dans certains cas [48]. La solution est capturée sur l'intervalle  $[0, 1/\delta]$ , mais l'algorithme d'adaptation requiert des pas de temps de plus en plus petits sur l'intervalle  $[1/\delta, 2/\delta]$ .

Les méthodes de prédiction-correction, quant à elles, reposent sur l'hypothèse forte que la solution recherchée possède des dérivées suffisamment lisses. Plus précisément, elles nécessitent que les dérivées successives soient approximativement constantes sur un intervalle de temps suffisamment petit [68]. Cette hypothèse est mise à mal dans le cas des problèmes raides. En effet, on peut le constater à l'aide de l'exemple de l'équation de combustion, où les variations brusques de la solution, lorsque  $\delta \ll 1$ , forcent le pas de temps à être de plus en plus petit. Or, les algorithmes d'adaptation en temps disposent d'un pas de temps minimal  $h_{\min}$  et d'un pas de temps maximal  $h_{\max}$ . Ces bornes sur le pas de temps permettent d'éviter des instabilités ou des opérations risquées auxquelles sont sensibles les formules de différences finies [20]. Pour des problèmes raides, on est alors souvent confronté à des situations où le pas de temps nécessaire est plus petit que le pas de temps minimal pour que l'estimation de l'équation de prédiction respecte les critères de précision fixés.

La difficulté d'obtenir un estimateur de l'erreur pour des problèmes raides a conduit de nom-

breux auteurs à chercher des stratégies d’adaptation moins coûteuses [59, 38, 39, 37, 21]. La stratégie proposée par Johnson et al. [40] et plus récemment par Shampine et al. [59] repose sur un contrôle des variations de la solution. Un autre point de vue également exploré par de nombreux auteurs [34, 3, 13, 66, 43] est d’adapter le pas de temps en fonction de l’erreur sur une quantité d’intérêt.

### 2.3.2 Le cas des équations de Maxwell pour la supraconductivité

Les équations de supraconductivité, qui constituent l’intérêt principal de cette étude, sont un exemple particulier de problème raide. Dans le but d’utiliser des matériaux supraconducteurs pour des applications de transfert de puissance électrique, il est nécessaire de les concevoir de façon à minimiser les pertes d’énergie lorsque ces derniers sont soumis à des courants alternatifs. Pour caractériser une telle perte de courant, l’évolution des champs électromagnétiques à l’intérieur du HTS doit être connue de façon précise. Cependant, la simulation en 3D des composants supraconducteurs reste, à l’heure actuelle, très coûteuse du point de vue du temps de calcul. En particulier, les méthodes numériques développées dans le domaine, et basées sur un pas de temps constant [5, 9, 1, 54, 50, 62, 52, 27, 36, 7, 8, 46, 63], sont limitées à de très petites tailles de pas. En effet, la forte non linéarité des équations et la présence de fronts progressant dans le temps [61] rendent le problème raide. Cette importante limitation sur la discrétisation numérique des équations de supraconductivité a restreint l’analyse numérique à des géométries simples. Les géométries plus complexes, mais aussi plus réalistes de matériaux supraconducteurs, deviennent très vite impraticables compte tenu du coût et des ressources de calculs qu’elles mobilisent.

Dans le milieu de la recherche sur la supraconductivité, plusieurs solutions pour palier à ces difficultés sont mises au point. Il existe tout d’abord une communauté grandissante d’ingénieurs travaillant à rendre plus efficaces les discrétisations en éléments-finis en utilisant, entre autres, des approximations basées sur des modèles semi-analytiques [60]. D’autres, encore, ont recours à des logiciels commerciaux de simulation numérique pour obtenir des approximations des champs électriques des HTS. Ces logiciels proposent des méthodes d’adaptation en temps qui sont basées sur des approximations des erreurs locale du même type que celles présentées dans ce chapitre.

Pour construire des modèles numériques efficaces pour la conception de HTS avec des géométries complexes, un outil essentiel est l’estimation de l’erreur entre la solution numérique et la solution exacte des équations de Maxwell non linéaires. La recherche de nouveaux estimateurs d’erreur [11] en éléments finis présente donc un intérêt croissant [2], notamment dans le domaine de la simulation des matériaux supraconducteurs [65]. Quoi qu’il en soit, il apparaît primordial de recourir à des méthodes d’adaptation en temps pour surmonter les

difficultés de la discrétisation des équations de Maxwell en temps. Dans le prochain chapitre, on se propose ainsi de présenter une nouvelle stratégie de sélection des pas de temps basée sur un estimateur de l'erreur locale.



### CHAPITRE 3 ÉLÉMENTS DE THÉORIE

Dans le chapitre précédent, nous avons montré l'importance de l'adaptation du pas de temps pour la discrétisation d'un problème de condition initiale. L'idée principale sur laquelle repose une majorité des techniques de contrôle du pas de temps est celle de l'estimation de l'erreur locale. Le fait de pouvoir anticiper cette erreur donne l'avantage d'avoir une plus grande confiance quant à la précision de l'approximation numérique obtenue. Le contrôle de l'erreur nous assure que le pas de temps choisi soit suffisamment petit pour bien représenter les variations de la solution discrète et pour que le schéma numérique soit stable. Les méthodes d'adaptation en temps diffèrent alors selon l'estimateur de l'erreur qu'elles utilisent. Cependant, la précision du résultat obtenu par le contrôle de l'erreur locale n'est pas le seul critère dans le choix d'une stratégie d'adaptation en temps. Il faut aussi que la technique d'adaptation soit efficace, c'est-à-dire qu'elle puisse donner une solution numérique en un nombre raisonnable de pas de temps. En d'autres termes, il faut un pas de temps suffisamment petit pour bien capturer les variations de la variable dépendante et pour obtenir un schéma stable, mais assez grand pour minimiser les coûts de calcul. On cherche donc une précision maximale pour un coût minimal. Il faut enfin que cette méthode soit robuste, i.e. qu'elle permette l'obtention d'approximations stables pour le plus grand nombre de problèmes possibles.

L'exemple des problèmes raides introduits au chapitre précédent illustrent le type d'équations pour lesquelles la robustesse et l'efficacité de l'algorithme d'adaptation en temps sont des critères plus critiques que la précision. Shampine et al. [59] développent une méthode d'adaptation de faible précision et de faible coût pour les problèmes raides. L'idée à la base du schéma d'adaptation est de choisir le pas de temps de manière à contrôler les variations relatives de l'approximation numérique de la variable dépendante.

Dans ce chapitre, on se propose de développer une nouvelle stratégie d'adaptation sur les pas de temps basée sur un estimateur de l'erreur locale. Après avoir construit cet estimateur, nous démontrons une relation entre celui-ci et l'erreur locale. À partir de cette relation, nous en déduisons les propriétés de consistance, de robustesse et d'efficacité qui nous permettront de vérifier ce nouveau schéma. Nous montrerons que l'algorithme de choix de pas de temps peut être vu comme une généralisation de celui proposé par Shampine et al. Cet algorithme peut être utilisé directement pour la sélection de pas de temps selon une quantité d'intérêt qui dépend de la solution. Nous démontrerons que l'estimateur d'erreur ainsi construit peut être utilisé sur la quantité d'intérêt, donnant lieu à une nouvelle technique d'adaptation en temps. Enfin, nous utiliserons ces stratégies sur des problèmes raides et non raides afin de vérifier leur performance.

### 3.1 Notations et définitions préliminaires

Dans ce chapitre, nous utiliserons un certain nombre de notations et de définitions qu'il nous faut introduire au préalable. On considère  $i, j, k, n, p$  et  $q$  des nombres entiers positifs ou nuls. Dans ce qui suit  $(X, \|\cdot\|_X)$  et  $(Y, \|\cdot\|_Y)$  désigneront des espaces de Banach. On notera  $\mathcal{C}(X, Y)$  l'ensemble des fonctions continues de  $X$  à valeur dans  $Y$ . De façon similaire  $\mathcal{C}^p(X, Y)$  représente l'ensemble des fonctions  $p$ -fois différentiables de  $X$  dans  $Y$ . Soient  $U$  un ouvert de  $\mathbb{R}^n$  et  $I$  un intervalle ouvert de  $\mathbb{R}$ . Si  $u \in \mathcal{C}^p(I, U)$  alors  $u^{(q)}$  désigne la dérivée temporelle d'ordre  $q$  de  $u$  pour  $1 \leq q \leq p$ . On notera indifféremment  $u'$  ou  $u^{(1)}$  pour désigner la dérivée première de  $u$ . Enfin, on spécifie l'intervalle d'intégration  $I = ]0, T[$ , où  $T > 0$  est un réel représentant le temps final d'intégration.

Lorsqu'on s'intéresse à la discrétisation d'équations différentielles ordinaires, il est intéressant de rappeler le théorème général d'existence et d'unicité de problèmes de condition initiale.

**Théorème 1 (Cauchy-Lipschitz).** *Soit  $f : V \subset \mathbb{R} \times \mathbb{R}^n \mapsto \mathbb{R}^d$  (où  $V$  est un ouvert de  $\mathbb{R} \times \mathbb{R}^d$ ) une fonction continue, localement  $L$ -lipschitzienne par rapport à la seconde variable. Alors pour toute donnée initiale  $(t_0, u_0) \in I \times U$ , il existe un ouvert  $U$  de  $\mathbb{R}^d$  contenant  $u_0$ ,  $\tilde{I}$  un voisinage de  $t_0$ , et une unique solution  $u \in \mathcal{C}^1(\tilde{I}, U)$  au problème de condition initiale :*

$$\begin{cases} u'(t) = f(t, u(t)), \\ u(t_0) = u_0. \end{cases} \quad (3.1)$$

*On supposera dans la suite que  $I \subset \tilde{I}$ .*

Le théorème de Cauchy-Lipschitz nous informe sur la régularité de la solution  $u$  ainsi que sur celle de  $f$ . En outre, ceci nous permet d'introduire la constante  $L > 0$  qui sera utilisée dans la suite des développements. Pour intégrer numériquement l'équation (3.1), on a recourt à des schémas aux différences finies. Les solutions approchées ainsi obtenues possèdent une certaine précision qu'il faut caractériser en introduisant l'erreur locale. Cette erreur dépend du schéma numérique utilisé.

**Définition 2 (Erreur locale).** *On définit l'erreur locale comme l'erreur commise au temps  $t = t_{n+1}$  lorsqu'on effectue un seul pas de temps  $h_n$  à l'aide d'un schéma numérique. Autrement dit, on compare la solution de l'équation différentielle ordinaire*

$$\begin{cases} u'(t) = f(t, u(t)), \\ u(t_n) = u_n, \end{cases} \quad (3.2)$$

avec la valeur  $u_{n+1}$  obtenue avec le schéma au temps  $t_{n+1}$ . On la note alors

$$\tau(u, h_n, t_{n+1}) = u(t_{n+1}) - u_{n+1}. \quad (3.3)$$

Pour les besoins de l'analyse, on peut développer l'erreur locale en série de Taylor,

$$\tau(u, h_n, t_{n+1}) = \sum_{j=0}^{\infty} \mathcal{B}_j(u) h_n^j, \quad (3.4)$$

où les coefficients  $(\mathcal{B}_j)$  dépendent du schéma aux différences utilisé. Dans le cas du schéma BDF1 par exemple, les coefficients dans le développement de l'erreur sont

$$\mathcal{B}_0 = 0, \mathcal{B}_1 = 0, \text{ et } \mathcal{B}_j = \frac{1}{j!} u^{(j)}(t_{n+1}), \quad \forall j \geq 2. \quad (3.5)$$

Soit  $i_0$  un entier positif. Dans la suite des développements, on suppose que le schéma numérique traité possède les coefficients suivants dans l'expression de l'erreur locale associée :

$$\begin{cases} \mathcal{B}_j = 0, & \forall j < i_0, \\ \mathcal{B}_j = \alpha_j u^{(j)}(t_n), & \forall j \geq i_0. \end{cases} \quad (3.6)$$

où les coefficients  $\alpha_j \in \mathbb{R}$ .

L'approximation de la solution du problème (3.1) ainsi que l'erreur locale associée sont calculées en des points d'une grille qui forme une partition de l'intervalle d'intérêt  $I$ . On rappelle brièvement la définition d'une partition d'un segment de  $\mathbb{R}$ .

**Définition 3** (Partition). *Une partition d'un segment  $[a, b]$  de  $\mathbb{R}$  est une suite finie  $t_0, t_1, \dots, t_n$  de nombres réels tels que*

$$a = t_0 < t_1 < t_2 < \dots < t_n = b.$$

*Dans la suite, on désignera par  $\mathcal{P}$  une partition d'un intervalle donné, de cardinal  $|\mathcal{P}|$ .*

On comprend alors que trouver l'ensemble des pas de temps  $h_j = t_{j+1} - t_j$  d'un schéma d'intégration est équivalent à trouver une partition de l'intervalle  $I$ . On parlera dans la suite indifféremment de trouver une partition  $\mathcal{P}$  ou de trouver les valeurs  $h_j$ .

### 3.2 Estimation de l'erreur sur la solution

La stratégie d'adaptation en temps que nous présentons dans cette section repose sur une idée simple du contrôle des variations de la solution. L'approche générale, développée par Shampine et al. [59], consiste à choisir le pas de temps de manière à ce que les changements

dans la solution soient plus petits en amplitude qu'une certaine tolérance. Cette méthode a l'avantage de la simplicité et est applicable indépendamment du schéma d'intégration utilisé. Il est alors possible de choisir un petit pas de temps lorsque la solution varie fortement et un plus grand pas de temps dans le cas contraire. On construit ainsi, une étape après l'autre, une partition  $\mathcal{P}$  sur laquelle les variations de la solution approchée sont représentées avec une certaine précision fixée. D'autre part, on obtient une solution approchée possédant des propriétés de régularités (continuité, différentiabilité, etc...) semblables à celles de la solution analytique.

**Définition 4** (Variations  $\epsilon$ -voisines). *Soient  $\epsilon > 0$  fixé et  $u \in \mathcal{C}^m([0, T])$ , où  $m$  est un entier positif ou nul. On dira que  $u$  est à variations  $\epsilon$ -voisines d'ordre  $k$  ( $0 \leq k \leq m$ ), s'il existe une partition  $\mathcal{P} = \{t_0, \dots, t_n\}$  de  $I$  telle que*

$$\Delta_k(u, \mathcal{P}) = \sum_{j=0}^k \sup_{i \in \{0, \dots, n\}} |u^{(j)}(t_{i+1}) - u^{(j)}(t_i)| < \epsilon. \quad (3.7)$$

En posant  $h_{\max} = \max_i |t_{i+1} - t_i|$  pour une fonction  $u$  suffisamment lisse, on voit (par continuité des dérivées successives de  $u$ ) qu'en prenant  $h_{\max}$  assez petit, on obtient une partition  $\mathcal{P}$  de  $I$  telle que  $u$  soit à variations  $\epsilon$ -voisines d'ordre  $k$ . La valeur de  $k$  dépend alors étroitement de la régularité de la fonction  $u$ . Si  $u$  n'est pas suffisamment régulière, il peut n'exister aucune partition telle que  $f$  soit à variations  $\epsilon$ -voisines d'ordre  $k$ . Un exemple est celui de la fonction de Heaviside sur  $[-1, 1]$ , pour  $\epsilon \leq 2$  et  $k = 0$ . Bien sûr, la condition « $h_{\max}$  suffisamment petit» n'est qu'une condition suffisante pour avoir une telle partition de  $I$ . L'exemple simple de la fonction  $x \mapsto \sin(x)$  montre que celle-ci est à variations  $\epsilon$ -voisines pour  $\epsilon$  arbitrairement petit et ce lorsqu'on prend un pas  $h$  constant multiple impair de  $\frac{\pi}{2}$ . Enfin, on constate qu'à une partition  $\mathcal{P}$  et  $k$  fixés,  $\Delta_k(\cdot, \mathcal{P})$  forme une semi-norme sur  $\mathcal{C}^k(I)$ , i.e. elle vérifie les propriétés suivantes :

$$\forall \lambda \in \mathbb{R}, \forall u \in \mathcal{C}^k(I), \quad \Delta_k(\lambda u, \mathcal{P}) = |\lambda| \Delta_k(u, \mathcal{P}), \quad (3.8)$$

$$\forall u, v \in \mathcal{C}^k(I), \quad \Delta_k(u + v, \mathcal{P}) \leq \Delta_k(u, \mathcal{P}) + \Delta_k(v, \mathcal{P}). \quad (3.9)$$

On est donc à présent en mesure d'énoncer le théorème qui donne un lien entre la semi-norme  $\Delta_k$  et l'erreur locale.

**Théorème 2** (Estimateur d'erreur). *Soit  $u$  la solution de l'équation (3.1). Soit  $(u_j)$  l'approximation obtenue jusqu'au temps  $t_{n+1}$  à partir d'un schéma numérique, sur une grille formant une partition  $\mathcal{P}$  de l'intervalle  $[t_0, t_{n+1}]$ . On suppose que les coefficients  $\mathcal{B}_i$  associés au schéma numérique sont de la forme (3.6) pour tout  $t_j$  élément de  $\mathcal{P}$ . On suppose de plus*

que  $|\mathcal{P}| \approx \frac{1}{h_{\max}} + 1$  et que pour tout  $i \geq i_0$ ,  $|h_i - h_{i-1}| = \mathcal{O}(h_i^2)$ . Alors, pour  $m$  un entier positif donné, il existe un entier  $r$  et une constante  $C > 0$  tels que

$$|u(t_{n+1}) - u_{n+1}| \leq C \Delta_m(u, \mathcal{P}) + \mathcal{O}(h_{\max}^r). \quad (3.10)$$

*Démonstration.* On rappelle que  $i_0$  est définie dans (3.6) comme le premier entier  $j$  tel que  $\mathcal{B}_j \neq 0$ . Si  $m < i_0$ , le résultat s'obtient immédiatement à partir du développement de l'erreur (3.5) et en prenant  $r = i_0$ . Dans le cas contraire, la démonstration repose sur une réécriture de l'erreur locale en une somme télescopique :

$$\tau(u, h_n, t_{n+1}) = \sum_{j=0}^n \tau(u, h_j, t_{j+1}) - \tau(u, h_{j-1}, t_j). \quad (3.11)$$

On adopte les conventions  $h_{-1} := 0$  et  $t_{-1} := t_0$ , et on rappelle que l'erreur est nulle en  $t_0$  puisque  $u(t_0) = u_0$ . Le développement en série de Taylor de chacun des termes d'erreur locale jusqu'aux termes d'ordre  $m$  donne ensuite

$$\tau(u, h_n, t_{n+1}) = \sum_{j=0}^n \sum_{i=i_0}^m \alpha_i h_j^i u^{(i)}(t_j) - \alpha_i h_{j-1}^i u^{(i)}(t_{j-1}) + \mathcal{R}_m(h_0, \dots, h_n), \quad (3.12)$$

puisque seuls les termes d'indice  $i \geq i_0$  subsistent, où  $\mathcal{R}_m$  est le terme de reste du développement de Taylor. Dans ce reste, on retrouve les termes

$$\mathcal{R}_m(h_0, \dots, h_n) = \mathcal{O}(h_0^{m+1}) + \mathcal{O}(h_1^{m+1}) + \dots + \mathcal{O}(h_n^{m+1}). \quad (3.13)$$

Afin de regrouper les termes selon la même puissance de  $h_j$  lorsque  $i \geq i_0$ , on effectue le développement

$$h_{j-1}^i = (h_j + (h_{j-1} - h_j))^i = (h_j)^i + \mathcal{O}(h_j^{i-1}(h_j - h_{j-1})). \quad (3.14)$$

Quitte alors à renommer le reste en  $\tilde{\mathcal{R}}_m$ , on y rajoute les termes négligés à l'étape (3.14), tel que

$$\tilde{\mathcal{R}}_m(h_0, \dots, h_n) = \sum_{j=0}^n \sum_{i=i_0}^m \mathcal{O}(h_j^{i-1}(h_j - h_{j-1})) + \sum_{j=0}^n \mathcal{O}(h_j^{m+1}). \quad (3.15)$$

On obtient alors l'expression suivante de l'erreur locale

$$\tau(u, h_n, t_{n+1}) = \sum_{j=0}^n \sum_{i=i_0}^m \alpha_i h_j^i \left( u^{(i)}(t_j) - u^{(i)}(t_{j-1}) \right) + \tilde{\mathcal{R}}_m(h_0, \dots, h_n). \quad (3.16)$$

L'inégalité triangulaire appliquée à l'erreur locale nous donne la première majoration

$$|\tau(u, h_n, t_{n+1})| \leq \sum_{j=0}^n \sum_{i=i_0}^m |\alpha_i| |u^{(i)}(t_j) - u^{(i)}(t_{j-1})| h_j^i + |\tilde{\mathcal{R}}_m(h_0, \dots, h_n)|. \quad (3.17)$$

En rappelant que  $h_{\max} = \max_j |t_{j+1} - t_j|$ , on peut donc majorer le terme d'erreur en utilisant la propriété de borne supérieure de  $\Delta_m$ . De plus, en majorant chaque puissance de  $h_j$  par une puissance de  $h_{\max}$  dans l'expression précédente, on obtient

$$|\tau(u, h_n, t_{n+1})| \leq \sum_{j=0}^n \Delta_m(u, \mathcal{P}) \sum_{i=i_0}^m |\alpha_i| h_{\max}^i + |\tilde{\mathcal{R}}_m(h_0, \dots, h_n)|. \quad (3.18)$$

Puisque par hypothèse, on a que  $(n+1) \approx \frac{1}{h_{\max}}$ , on peut alors définir la constante  $C = \sum_{i=i_0}^m |\alpha_i| h_{\max}^{i-1}$  pour obtenir l'inégalité

$$|\tau(u, h_n, t_{n+1})| \leq C \Delta_m(u, \mathcal{P}) + |\tilde{\mathcal{R}}_m(h_0, \dots, h_n)|. \quad (3.19)$$

Pour démontrer le résultat escompté, il ne reste plus qu'à majorer le reste. Une analyse des termes dans  $\tilde{\mathcal{R}}_m$  montre que les termes dominants sont d'ordre  $i_0$ . En majorant chacune de ces puissances et en sommant comme à l'étape précédente, on obtient :

$$|\tau(u, h_n, t_{n+1})| \leq C \Delta_m(u, \mathcal{P}) + \mathcal{O}(h_{\max}^{i_0-1}). \quad (3.20)$$

□

Le résultat démontré (3.10) est un résultat clé de ce travail. En effet, il établit le lien entre l'erreur locale et le contrôle des variations de la solution  $u$  à l'aide de la semi-norme  $\Delta_k$ . Plus précisément, on montre que l'erreur locale diminue linéairement en fonction de la précision  $\epsilon$ . Ceci justifie l'utilisation d'un algorithme de contrôle en temps selon les variations  $\epsilon$ -voisines de la solution plutôt que le contrôle de l'erreur locale.

### Remarques.

- *L'ensemble des définitions et des raisonnements qui viennent d'être donnés permettent de justifier la consistance de l'adaptation en temps basée sur les variations de la solution  $u$ . Pour que l'erreur locale soit plus petite qu'une tolérance donnée, il suffit de choisir  $\epsilon$  assez petit.*
- *Par ailleurs, jusqu'à présent, la définition de l' $\epsilon$ -variation et la démonstration qui en a découlé ont été données pour une équation différentielle ordinaire. Cependant, la stratégie d'adaptation se généralise à des systèmes de la forme du système (3.1). Il suffit de remplacer les valeurs absolues par la norme désirée.*

- On voit dans la démonstration que la stratégie d'adaptation peut s'appliquer à une large variété de schémas linéaires à un pas. La linéarité des coefficients  $\mathcal{B}_j$  est importante dans le raisonnement et permet d'étendre la relation avec l'erreur locale pour des schémas à pas multiples linéaires.
- Les méthodes d'adaptation basées sur les variations de la solution soumises à une certaine précision  $\epsilon$  effectuent un nombre de pas de temps de l'ordre de  $1/\epsilon$  [59]. L'efficacité de ce type de méthode dépend donc de la précision  $\epsilon$ .
- L'hypothèse  $|\mathcal{P}| \approx \frac{1}{h_{\max}}$  est forte. En pratique, une hypothèse plus réaliste serait  $|\mathcal{P}| = \mathcal{O}\left(\frac{1}{h_{\max}}\right)$ .
- Dans le cas du schéma d'Euler explicite, on peut montrer que la constante  $C = \left(\frac{e^{h_{\max}} - h_{\max} - 1}{h_{\max}^2}\right)$ .
- Les constantes  $C$  et  $r$  dépendent du schéma numérique utilisé.

**Exemple 3.2.1.** Pour explorer davantage le résultat que l'on vient de démontrer, on considère l'exemple de l'équation différentielle type :

$$\begin{cases} u' = \lambda u, \\ u(0) = 1, \end{cases} \quad (3.21)$$

où  $\lambda \in \mathbb{C}$ , tel que  $\text{Re}(\lambda) < 0$ . On rappelle que la solution exacte de ce problème est donnée par

$$u(t) = e^{\lambda t}, \quad \forall t \in I. \quad (3.22)$$

D'un point de vue numérique, le comportement des schémas de discrétisation sur cette équation simple permet de tirer des conclusions assez générale.

Supposons que l'on dispose de la solution exacte obtenue à variations  $\epsilon$ -voisines d'ordre 0 sur une partition  $\tilde{\mathcal{P}}$  de  $[0, t_n]$ . Cherchons la condition sur  $h$  pour que la solution approchée, obtenue au temps  $t_{n+1}$  avec la méthode d'Euler explicite, demeure à variations  $\epsilon$ -voisines d'ordre 0 sur la nouvelle partition  $\mathcal{P}$  de  $[0, t_{n+1}]$ .

On raisonne par conditions nécessaires et suffisantes. Dans le cas où

$$\Delta(u, \mathcal{P}) < \epsilon,$$

alors, en particulier, on a que

$$|u_{n+1} - u(t_n)| < \epsilon.$$

Puisque l'approximation  $u_{n+1}$  est obtenue en effectuant un pas  $h$  de la méthode d'Euler

explicite, on peut écrire

$$|u(t_n) - h\lambda u(t_n) - u(t_n)| < \epsilon.$$

On en déduit alors qu'en particulier,

$$h < \frac{\epsilon}{|\lambda| \sup_n |e^{\lambda t_n}|} \leq \frac{\epsilon}{|\lambda|}.$$

Une condition nécessaire pour que la solution soit à variations  $\epsilon$ -voisine d'ordre 0 est que  $h < \epsilon/|\lambda|$ . On peut vérifier que c'est une condition suffisante. On trouve ainsi une condition qui permet de choisir le pas de temps de manière à représenter les variations de fréquence  $1/|\lambda|$  avec une précision  $\epsilon$ . On peut montrer que dans le cas général de l'équation (3.1), le coefficient  $L$  joue le rôle de  $\lambda$ .

En revanche, si l'on ne tient compte que de la consistance du schéma d'adaptation, la démonstration précédente nous montre également qu'il n'est pas nécessaire de prendre en compte toutes les dérivées jusqu'à l'ordre  $m$ . En effet, il pourrait être plus avantageux du point de vue de l'efficacité de la méthode que celle-ci prenne en compte la précision du schéma numérique adapté. On peut, pour ce faire, introduire la définition de variations  $\epsilon$ -voisines pondérées.

**Définition 5** (Variations  $\epsilon$ -voisines pondérées). *Soient  $\epsilon > 0$  fixé et  $u \in \mathcal{C}^m(I)$ , où  $m$  est un entier positif ou nul. On dira que  $u$  est à variations  $\epsilon$ -voisines pondérées d'ordre  $k$  ( $0 \leq k \leq m$ ) s'il existe  $\mathcal{P} = \{t_0, \dots, t_n\}$ , une partition de  $I$ , telle que*

$$\Delta_k(u, \mathcal{P}) = \sum_{j=0}^k \omega_j \sup_i |u^{(j)}(t_{i+1}) - u^{(j)}(t_i)| < \epsilon, \quad (3.23)$$

où les coefficients  $\omega_j \in \mathbb{R}_+$  dépendent du schéma numérique utilisé.

On pourrait alors prendre les coefficients  $w_j$  comme étant nuls lorsque la  $j$ -ième dérivée n'apparaît pas dans le développement de l'erreur et comme étant égal à 1 sinon. Ainsi on aurait un estimateur de l'erreur qui prendrait en compte la précision du schéma utilisé.

On expose à présent l'algorithme d'adaptation en temps selon le critère de variations  $\epsilon$ -voisines. Étant donné un pas de temps  $h_n$  et une valeur initiale  $u_n$  au temps  $t_n$ , le schéma d'intégration permet de construire l'approximation  $u_{n+1}$  au temps  $t_{n+1} = t_n + h_n$ . On est alors capable d'évaluer la semi-norme d'ordre  $k$ ,  $\Delta_k$ , à l'aide de formules aux différences arrières. Pour ce faire, on utilise la nouvelle approximation  $u_{n+1}$  calculée ainsi que les valeurs  $(u_j)$  précédentes. En effet, puisque l'on ne dispose pas de la solution exacte et de ses dérivées successives, la semi-norme  $\Delta_k$  nécessite d'être approximée pour les besoins de l'algorithme. Par ailleurs, telle que nous l'avons définie, la semi-norme prend en compte les variations absolues sur la solution discrétisée. Afin que l'adaptation en temps tienne compte des ordres



de grandeur intervenant dans le problème (3.1), il est nécessaire de définir un ratio de variation relative. On définit dans ce but la fonction de contrôle relative

$$r_n = \frac{\Delta_k(u)}{\epsilon_R |u_{n+1}| + \epsilon_A}. \quad (3.24)$$

Dans ce qui suit, on désignera par ratio de contrôle le ratio  $r_n$ . Le paramètre  $\epsilon_R$  est une tolérance relative sur les variations de l'approximation. La valeur de  $\epsilon_A$  est de l'ordre de grandeur de la précision machine afin que les variations de la solution discrète plus petites que l'erreur de représentation ne soient pas considérées par l'algorithme d'adaptation en temps. Afin d'obtenir une approximation de la solution de (3.1) à variations  $\epsilon$ -voisines, il suffit de s'assurer que le ratio de contrôle soit plus petit que 1. On peut alors contrôler les pas de temps de manière à conserver le ratio  $r_n$  entre les bornes  $0 < \eta \leq r_n \leq 1$ . Enfin, il est important de garder le pas de temps  $h_n$  entre une borne minimale  $h_{\min}$  et une borne maximale  $h_{\max}$ . Comme nous l'avons rappelé au chapitre 2, ces bornes sur le pas de temps permettent d'éviter des instabilités numériques auxquelles sont sensibles les formules de différences finies.

On peut maintenant exposer la structure logique de l'algorithme d'adaptation en temps. Soient  $\alpha < 1$  et  $\beta > 1$  deux paramètres fournis par l'utilisateur. Si le ratio de contrôle n'est pas respecté, i.e si  $r_n > 1$ , et que  $\alpha h_n > h_{\min}$ , l'approximation  $u_{n+1}$  est rejetée. Le pas de temps  $h_n$  est multiplié par le facteur  $\alpha$ , i.e.  $h_n = \alpha h_n$ , et le calcul est recommencé. La valeur de  $h_n$  est donc réduite afin de mieux capturer les variations. Si le ratio est inférieur à la borne  $\eta$ , et que  $\beta h_n < h_{\max}$ , l'approximation  $u_{n+1}$  est acceptée. Le pas de temps est multiplié par le facteur  $\beta$  pour être augmenté, i.e.  $h_{n+1} = \beta h_n$ , et on passe au temps suivant. Le but est toujours d'obtenir une précision maximale pour un coût en temps de calcul minimal. La borne  $\eta$  permet d'estimer les valeurs pour lesquelles  $h_n$  semble pessimiste, i.e. lorsqu'il est plus petit que nécessaire pour capturer les variations. Plus la valeur de  $\eta$  est proche de 1, la borne limite, plus l'algorithme d'adaptation va tenter d'améliorer le pas de temps en l'augmentant et ainsi essayer de réduire le coût en temps de calcul. Enfin, dans le cas où la valeur du ratio est encadrée, i.e.  $0 < \eta \leq r_n \leq 1$ , ou que  $\alpha h_n \leq h_{\min}$ , ou encore quand  $\beta h_n \geq h_{\max}$  le pas de temps conserve sa valeur, i.e.  $h_{n+1} = h_n$  et l'approximation  $u_{n+1}$  est acceptée. Ainsi, il peut arriver que le critère sur le ratio de contrôle ne soit pas respecté, lorsque  $\alpha h_n \leq h_{\min}$ , mais que la solution discrète soit validée. Ceci permet de conserver la robustesse de l'algorithme en évitant des boucles infinies. L'algorithme 1 illustre ce calcul sous forme de pseudo-code.

---

**Algorithm 1** Adaptation en temps sur la variable dépendente

---

```

1: Étant donné  $n_{max}$  un nombre maximal d'itérations.
2: Étant donné un temps initial  $t_0$  et un temps final  $t_{max}$ .
3: Étant donné une tolérance relative  $\epsilon_R$  et une tolérance absolue  $\epsilon_A$  sur l'erreur de représentation.
4: Étant donné un pas de temps maximal  $h_{max}$ , minimal  $h_{min}$  et initial  $h_0$ .
5: Étant donné les paramètres d'adaptation  $\alpha, \beta$  et  $\eta$ .
6: procedure ADAPTTEMPS( $n_{max}, t_0, t_{max}, \epsilon_R, \epsilon_A, h_{max}, h_{min}, h_0, \alpha, \beta, \eta$ )
7:   while ( $n \leq n_{max}$ ) and ( $t \leq t_{max}$ ) do
8:     Calculer la nouvelle valeur  $u_{new}$  à partir d'un schéma d'intégration
9:     Mettre à jour la valeur de  $\Delta_k$ 
10:    if ( $\Delta_k(u) \geq \epsilon_R |u_{new}| + \epsilon_A$ ) and ( $h_n \geq \alpha h_{min}$ ) then
11:       $h_n \leftarrow \alpha h_n$ 
12:       $t_n \leftarrow t_{n-1} + h_n$ 
13:    else if ( $\Delta_k(u) \leq \eta(\epsilon_R |u_{new}| + \epsilon_A)$ ) and ( $h_n \leq \beta h_{max}$ ) then
14:       $u_{n+1} \leftarrow u_{new}$ 
15:       $h_{n+1} \leftarrow \beta h_n$ 
16:       $t_{n+1} \leftarrow t_n + h_{n+1}$ 
17:       $n \leftarrow n + 1$ 
18:    else
19:       $u_{n+1} \leftarrow u_{new}$ 
20:       $h_{n+1} \leftarrow h_n$ 
21:       $t_{n+1} \leftarrow t_n + h_{n+1}$ 
22:       $n \leftarrow n + 1$ 
23:    end if
24:  end while
25: end procedure

```

---

Dans le cadre de ce travail, les valeurs  $\alpha = 0,5$ ,  $\beta = 2$  et  $\eta = 0,9$ , obtenues de façon empirique, ont fourni des résultats satisfaisants. Cet algorithme est à la base de la stratégie d'adaptation en temps développée dans ce travail. Dans la section suivante, nous montrons comment il peut être adapté pour contrôler les variations d'une quantité d'intérêt.

### 3.3 Estimation d'erreur sur une quantité d'intérêt

Dans le domaine de la simulation numérique de problèmes d'ingénierie, l'objectif n'est souvent pas d'obtenir une approximation de la solution du problème étudié mais plutôt de connaître

une certaine quantité physique qui dépend de cette solution. Cette quantité porte le nom de quantité d'intérêt. On peut citer comme exemples de quantités d'intérêt la vitesse moyenne d'un fluide dans une partie d'un domaine, une pression ou une température moyenne, la portance en aérodynamique, une énergie cinétique ou potentielle. Dans le cadre de cette étude, la quantité d'intérêt est la perte de courant par effet Joule dans des supraconducteurs. Il devient alors plus intéressant de minimiser l'erreur commise lors de la discrétisation, non plus par rapport à la solution exacte, mais par rapport à une quantité d'intérêt pertinente. La recherche d'estimateurs d'erreur sur une quantité d'intérêt en éléments finis connaît, pour ces raisons, un intérêt croissant [43, 66, 65].

Dans cette partie, on se propose de donner un cadre théorique à l'étude d'une quantité d'intérêt et de rappeler certains résultats mathématiques nécessaires aux développements qui vont suivre. On expose ensuite les bases d'une stratégie d'adaptation en temps en fonction des variations  $\epsilon$ -voisines d'une quantité d'intérêt.

**Définition 6** (Quantité d'intérêt). *Soient  $X$  et  $Y$  deux espaces de Banach,  $u$  un élément de  $X$  et  $Q$  une application de  $\mathcal{C}(X, Y)$ . On désignera par quantité d'intérêt (abrégé  $QoI$ ) l'image par  $Q$  de  $u$ , que l'on notera dorénavant  $Q(u)$ .*

#### Remarques.

- Dans le cas où  $Y = \mathbb{R}$  et  $Q \in \mathcal{L}(X, Y)$ ,  $Q$  est une forme linéaire sur  $X$  et la quantité d'intérêt est le scalaire  $Q(u)$ .
- De manière générale, une quantité d'intérêt représente une quantité physique pertinente pour le travail de l'ingénieur. Elle peut représenter une fonction dépendante de la température, une différence de pression, une portance ou encore une perte de courant comme nous le verrons dans le cas des supraconducteurs.

Pour pouvoir appliquer la stratégie d'adaptation en temps à une quantité d'intérêt, il faut néanmoins rajouter certaines hypothèses de régularité sur cette dernière. En effet, la semi-norme  $\Delta_k$  fait intervenir les dérivées successives de la fonction  $u$ . Il faut donc que  $Q$  soit Fréchet-différentiable. Avant de rappeler la définition de différentielle au sens de Fréchet, il nous faut introduire la notion d'opérateur linéaire borné.

**Définition 7** (Opérateur linéaire borné). *Soient  $(X, \|\cdot\|_X)$  et  $(Y, \|\cdot\|_Y)$  deux espaces de Banach et soit  $\mathcal{A} : X \mapsto Y$  un opérateur linéaire. On dira que  $\mathcal{A}$  est borné s'il existe un réel  $M \geq 0$  tel que pour tout  $x \in X$ ,*

$$\|\mathcal{A}x\|_Y \leq M\|x\|_X.$$

On peut alors définir la norme d'opérateur comme

$$\|\mathcal{A}\|_{\text{op}} = \sup_{x \in X \setminus \{0\}} \frac{\|\mathcal{A}x\|_Y}{\|x\|_X}.$$

On peut à présent introduire la différentiabilité selon Fréchet qui va nous permettre d'appliquer la semi-norme  $\Delta_k$  à une quantité d'intérêt.

**Définition 8** (Fréchet-différentiabilité). *Soient  $(X, \|\cdot\|_X)$  et  $(Y, \|\cdot\|_Y)$  deux espaces de Banach et  $U \subset X$  un ouvert de  $X$ . On dira qu'une fonction  $\mathcal{Q} : U \mapsto Y$  est Fréchet-différentiable en un point  $x \in U$  s'il existe un opérateur linéaire bornée  $\mathcal{A} : X \mapsto Y$  tel que pour tout  $h \in U$ ,*

$$\mathcal{Q}(x + h) = \mathcal{Q}(x) + \mathcal{A}h + o(\|h\|).$$

*S'il existe un tel opérateur  $\mathcal{A}$ , il est unique, et on le notera dans ce qui suit  $D\mathcal{Q}(x)$ .*

Plus généralement, nous aurons besoin que l'opérateur  $Q$  soit  $n$  fois différentiable. On notera  $D_u^n Q$  la  $n^{\text{ième}}$  différentielle. Cependant, nous ne travaillerons pas exactement avec l'opérateur  $Q$  mais plutôt avec la quantité d'intérêt  $Q(u)(t)$ . Nous avons donc besoin d'une identité reliant la différentielle  $t \mapsto D_t^n(Q(u))(t)$  avec la différentielle  $D_u^n Q$  et les dérivées successives  $u^{(n)}$ . Ce lien nous est fourni par le théorème de Faà di Bruno. Ce théorème fait intervenir une identité d'analyse combinatoire reliant la dérivée  $n^{\text{ième}}$  d'une composée de deux fonctions avec les dérivées successives de chacune des deux fonctions. Cette identité fait appel aux polynômes de Bell que nous définissons à présent.

**Définition 9** (Polynômes de Bell). *Un polynôme exponentiel de Bell est défini par la relation*

$$B_{n,k}(x_1, x_2, \dots, x_{n-k+1}) = \sum \frac{n!}{j_1! j_2! \dots j_{n-k+1}!} \left(\frac{x_1}{1!}\right)^{j_1} \left(\frac{x_2}{2!}\right)^{j_2} \dots \left(\frac{x_{n-k+1}}{(n-k+1)!}\right)^{j_{n-k+1}}, \quad (3.25)$$

*où la somme porte sur toutes les suites  $j_1, j_2, \dots, j_{n-k+1}$  d'entiers naturels telles que :*

$$j_1 + j_2 + \dots + j_{n-k+1} = k$$

*et*

$$j_1 + 2j_2 + 3j_3 + \dots + (n - k + 1)j_{n-k+1} = n.$$

Nous pouvons maintenant énoncer le théorème de Faà di Bruno dans le cas de fonctions à valeurs réelles.

**Théorème 3** (Formule de Faà di Bruno [41, 56]). *Soient  $f, g : \mathbb{R} \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^n$ . Alors pour tout point  $t \in \mathbb{R}$ , la formule de Faà di Bruno s'écrit*

$$\frac{d^n}{dt^n}(f \circ g)(t) = \sum_{k=1}^n f^{(k)}(u(t)) B_{j,k}(g^{(1)}(t), \dots, g^{(j-k)}(t)). \quad (3.26)$$

**Remarque.**

*Nous avons énoncé l'identité de Faà Di Bruno dans le cas des fonctions réelles. Cependant, celle-ci se généralise pour des opérateurs à valeur sur un espace de Banach [56].*

Nous nous plaçons à présent dans le cas simplifié où l'équation (3.1) est autonome, i.e. où  $u$  est solution du problème :

$$\begin{cases} u'(t) = f(u(t)), & \forall t \in I, \\ u(0) = u_0. \end{cases} \quad (3.27)$$

Cette hypothèse est importante dans la démonstration qui va suivre. Bien que nous nous limitons ici au cas des problèmes autonomes, il est possible que le résultat qui va suivre puisse être étendu aux équations non autonomes.

On est maintenant en mesure d'énoncer le second théorème qui fait le lien entre la semi-norme  $\Delta_k$  appliquée à une quantité d'intérêt et la semi-norme sur la solution  $u$  du problème (3.27).

**Théorème 4** (Estimateur d'erreur). *Soient  $(\mathcal{C}^m([0, T]), \|\cdot\|_\infty)$  et  $u$  la solution de (3.27) telle que  $u \in \mathcal{C}^m([0, T])$ . Soit  $(u_j)$  l'approximation obtenue jusqu'au temps  $t_{n+1}$ , à l'aide d'un schéma numérique, sur une grille formant une partition  $\mathcal{P}$  de l'intervalle  $I_n = [t_0, t_{n+1}]$ . On considère une quantité d'intérêt  $Q \in \mathcal{C}^{m+1}(X, \mathbb{R})$  de  $u$  qui vérifie*

$$\forall j \in \{0, m\}, \forall t \in I_n, \quad D_u^j Q(u(t)) > 0. \quad (3.28)$$

*Alors pour tout entier  $m$ , il existe des constantes  $C_1$  et  $C_2$  strictement positives telles que*

$$C_1 \Delta_m(Q(u), \mathcal{P}) \leq \Delta_m(u, \mathcal{P}) \leq C_2 \Delta_m(Q(u), \mathcal{P}) + \mathcal{O}(h_{\max}). \quad (3.29)$$

*Démonstration.* On commence par rappeler la définition de la semi-norme  $\Delta_m$  appliquée à la quantité d'intérêt  $Q(u)$  :

$$\Delta_m(Q(u), \mathcal{P}) = \sum_{j=0}^m \sup_i \left| D_t^j(Q(u))(t_{i+1}) - D_t^j(Q(u))(t_i) \right|.$$

Nous montrerons d'abord l'inégalité de gauche. Pour ce faire, on utilise la formule donnée par le théorème 3 pour développer la différentielle de  $Q(u)$ . Soient  $j \in \{0, m\}$  et  $i \in \{0, n\}$  fixés arbitrairement. En prenant la différence aux temps  $t_{i+1}$  et  $t_i$ , et à l'aide de l'inégalité triangulaire, on a

$$\begin{aligned} & \left| D_t^j(Q(u))(t_{i+1}) - D_t^j(Q(u))(t_i) \right| \\ & \leq \sum_{k=0}^j \left| D_u^k(Q)(u(t_{i+1})) B_{j,k}(u^{(1)}(t_{i+1}), \dots, u^{(j-k+1)}(t_{i+1})) - D_u^k(Q)(u(t_i)) B_{j,k}(u^{(1)}(t_i), \dots, u^{(j-k+1)}(t_i)) \right| \end{aligned} \quad (3.30)$$

On définit alors pour tout  $k \in \{0, \dots, j\}$ , la fonction  $G_k$  comme suit :

$$G_k : \begin{cases} \mathbb{R}^{j-k+2} \longrightarrow \mathbb{R}, \\ X_i \mapsto D_u^k(Q)(u(t_i)) B_{j,k}(u^{(1)}(t_i), \dots, u^{(j-k+1)}(t_i)), \end{cases}$$

où  $X_i$  est le vecteur  $(u(t_i) \ u^{(1)}(t_i) \ \dots \ u^{(j-k+1)}(t_i))^T$ . On muni  $\mathbb{R}^{j-k+2}$  de la norme  $\|\cdot\|_\infty$ . La différence (3.30) peut être réécrite à l'aide des fonction  $G_k$  comme

$$\left| D_t^j(Q(u))(t_{i+1}) - D_t^j(Q(u))(t_i) \right| \leq \sum_{k=0}^j \left| G_k(X_{i+1}^k) - G_k(X_i^k) \right|. \quad (3.31)$$

Puisque  $Q$  est de classe  $\mathcal{C}^{m+1}$ , les fonctions  $G_k$  sont au moins de classe  $\mathcal{C}^1$ . On peut alors utiliser l'inégalité des accroissements finis et déduire que

$$\left| D_t^j(Q(u))(t_{i+1}) - D_t^j(Q(u))(t_i) \right| \leq \sum_{k=0}^j \|D_X G_k\|_{\text{op}} \|X_{i+1}^k - X_i^k\|_\infty. \quad (3.32)$$

Par conséquent, puisque  $i$  est choisi arbitrairement, en utilisant la propriété de borne supérieure et par définition de la semi-norme  $\Delta_j$ , on obtient la majoration

$$\sup_i \left| D_t^j(Q(u))(t_{i+1}) - D_t^j(Q(u))(t_i) \right| \leq \sum_{k=0}^j \|D_X G_k\|_{\text{op}} \Delta_j(u, \mathcal{P}). \quad (3.33)$$

D'où, en posant la constante  $C_j = \sum_{k=0}^j \|D_X G_k\|_{\text{op}}$ , on a

$$\sup_i \left| D_t^j(Q(u))(t_{i+1}) - D_t^j(Q(u))(t_i) \right| \leq C_j \Delta_j(u, \mathcal{P}). \quad (3.34)$$

Comme  $j$  est choisi arbitrairement, on peut faire le même raisonnement et sommer sur les indices  $j \in \{0, m\}$ . En majorant alors à l'aide de la semi-norme d'ordre  $m$ , on montre qu'il

existe une constante  $C_1 > 0$  telle que

$$C_1 \Delta_m(Q(u), \mathcal{P}) \leq \Delta_m(u, \mathcal{P}) \quad (3.35)$$

où  $C_1 = \left(\sum_{j=0}^m C_j\right)^{-1}$ .

L'inégalité de droite nécessite des hypothèses plus fortes sur la quantité  $Q$ . La démonstration est cependant semblable puisqu'elle repose également sur l'identité de Faà Di Bruno. On commence par fixer arbitrairement  $j \in \{1, \dots, m\}$  ( $j \neq 0$ ) et  $i \in \{0, \dots, n\}$ . Puisque  $u$  est solution de l'équation différentielle autonome (3.27), on peut écrire au préalable les égalités

$$\begin{aligned} \left|u^{(j)}(t_{i+1}) - u^{(j)}(t_i)\right| &= h_i \left|u^{(j+1)}(t_i)\right| + \mathcal{O}(h_i) \\ &= h_i \left|(f \circ u)^{(j)}(t_i)\right| + \mathcal{O}(h_i). \end{aligned}$$

Encore une fois, l'élément essentiel de l'identité de Faà Di Bruno est de donner l'existence d'un polynôme, indépendant de  $f$ , à coefficients positifs. En effet, les coefficients de ce polynôme, que l'on note dans la suite  $p_n$ , sont issus de calculs combinatoires. On peut alors en déduire par monotonie

$$\begin{aligned} h_i \left|(f \circ u)^{(j)}(t_i)\right| + \mathcal{O}(h_i) &= h_i \left|p_n \left(f^{(0)}(u(t_i)), \dots, f^{(j)}(u(t_i))\right)\right| + \mathcal{O}(h_i) \\ &\leq h_i p_n \left(|f^{(0)}(u(t_i))|, \dots, |f^{(j)}(u(t_i))|\right) + \mathcal{O}(h_i). \end{aligned}$$

Par ailleurs,  $p_n$  vérifie  $p_n(0) = 0$ , alors s'il existe une constante  $C_j > 1$  telle que

$$\forall t_i \in \mathcal{P}, \quad C_j D_u^j Q(u(t_i)) \geq \left|f^{(j)}(u(t_i))\right|, \quad (3.36)$$

on en déduit encore une fois par monotonie et, en notant  $d = \deg(p_n)$  le degré du polynôme  $p_n$ , que :

$$\begin{aligned} \left|u^{(j)}(t_{i+1}) - u^{(j)}(t_i)\right| &\leq h_i C_j^d p_n \left(Q(u(t_j)), D_u Q(u(t_i)), \dots, D_u^j Q(u(t_j))\right) + \mathcal{O}(h_i) \\ &= h_i C_j^d \left|D_t^j(Q(u))(t_i)\right| + \mathcal{O}(h_i) \\ &\leq C_j^d \left|D_t^{j-1}(Q(u))(t_{i+1}) - D_t^{j-1}(Q(u))(t_i)\right| + \mathcal{O}(h_i). \end{aligned}$$

On pose ensuite  $C_{2,j} = C_j^d$ . Puisque  $i$  est fixé arbitrairement et que  $C_{2,j}$  est indépendant de  $i$ , en utilisant la borne supérieure de part et d'autre de l'inégalité précédente, on arrive à la relation

$$\sup_i \left|u^{(j)}(t_{i+1}) - u^{(j)}(t_i)\right| \leq C_{2,j} \sup_i \left|D_t^{j-1}(Q(u))(t_{i+1}) - D_t^{j-1}(Q(u))(t_i)\right| + \mathcal{O}(h_{\max}). \quad (3.37)$$

Puisque  $j$  est choisi arbitrairement, on peut faire le même raisonnement et sommer sur les indices  $j \in \{1, \dots, m\}$ .

Dans le cas où  $j = 0$ , par hypothèse, il existe un réel  $\ell > 0$  tel que  $Q^{(1)}(u)(t) \geq \ell > 0$  pour tout  $t \in I_n$ . On a donc l'inégalité d'accroissement

$$|u(t_{i+1}) - u(t_i)| \leq \frac{1}{\ell} |Q(u)(t_{i+1}) - Q(u)(t_i)| + \mathcal{O}(h_i). \quad (3.38)$$

Soit encore, en utilisant la propriété de borne supérieure, puisque  $i$  est choisi arbitrairement,

$$\sup_i |u(t_{i+1}) - u(t_i)| \leq \frac{1}{\ell} \sup_i |Q(u)(t_{i+1}) - Q(u)(t_i)| + \mathcal{O}(h_{\max}). \quad (3.39)$$

En regroupant les inégalités obtenues avec les indices  $j \in \{1, m\}$  avec celle donnée pour  $j = 0$ , on montre donc qu'il existe une constante  $C_2$  telle que :

$$\Delta_m(u, \mathcal{P}) \leq C_2 \Delta_m(Q(u), \mathcal{P}) + \mathcal{O}(h_{\max}). \quad (3.40)$$

Pour achever la démonstration, il ne reste plus qu'à montrer qu'une telle constante  $C_j$  existe, pour  $j$  fixé. Ceci est rendu possible puisqu'on a supposé que pour tout  $t \in I_n$ ,  $Q^{(j)}(u)(t) > 0$ . On peut alors poser

$$C_j = \alpha_j \frac{\|f^{(j)} \circ u\|_{\infty}^{I_n}}{\min_{t \in I_n} |D_u^j Q(u(t))|} \quad (3.41)$$

si  $f^{(j)}$  est non identiquement nulle sur  $\mathcal{P}$ . Dans le cas contraire, le choix de  $C_j$  est trivial. Le paramètre  $\alpha_j$  est choisi de sorte que  $C_j > 1$ .  $\square$

Le résultat démontré (3.29), combiné avec la relation (3.10), montre qu'il existe un lien entre la semi-norme  $\Delta_k$  appliquée à une quantité d'intérêt et l'erreur locale. Ce résultat est important puisqu'il justifie la consistance de l'adaptation en temps sur une quantité d'intérêt. Plus précisément, si l'on cherche à obtenir une quantité d'intérêt à variations  $\epsilon$ -voisines, avec une tolérance  $\epsilon$  de plus en plus petite, alors l'erreur locale tendra vers une valeur nulle.

### Remarque.

*L'hypothèse de positivité sur les dérivées successives de la quantité d'intérêt est restrictive. Cependant, il est courant en ingénierie de s'intéresser à des moyennes de quantités positives sur une période de temps [66]. On aurait donc en pratique la positivité des variations au moins jusqu'à l'ordre 1, ce qui peut être suffisant pour la stratégie d'adaptation.*



### 3.4 Vérification de la stratégie d'adaptation en temps et résultats numériques

Dans cette section, nous allons présenter les résultats numériques obtenus pour vérifier la stratégie d'adaptation en temps. Nous avons choisi plusieurs exemples de problèmes de condition initiale. Ces problèmes ont été sélectionnés pour mettre en avant les propriétés de consistance, d'efficacité et de robustesse, qui sont les critères de vérification définis dans la méthodologie. Il s'agit, pour la plupart de ces exemples, de problèmes raides standards rencontrés dans la littérature [32, 47]. Cette section a également pour objectif d'étudier la stratégie d'adaptation en temps proposée, basée sur les semi-normes  $\Delta_k$ . Des comparaisons entre différents choix de semi-normes seront faites. Nous vérifierons que les résultats numériques corroborent les théorèmes démontrés dans ce chapitre. Enfin, nous donnerons un exemple pour lequel l'adaptation en temps est faite en fonction des variations d'une quantité d'intérêt.

**Exemple 3.4.1** (Équation harmonique d'ordre 1). On s'intéresse encore une fois au cas simple de l'équation harmonique avec  $\lambda = -10$  :

$$\begin{cases} u' = \lambda u, \\ u(0) = u_0. \end{cases} \quad (3.42)$$

Cet exemple est choisi car la décroissance rapide de la solution en fait le problème raide le plus simple de la littérature [47]. On utilise le schéma BDF2 pour intégrer cette équation avec des pas non constants. L'algorithme d'adaptation en temps est appliqué avec les paramètres  $h_{\max} = 1/\lambda = 0,1$ ,  $h_{\min} = 1 \times 10^{-4}$  et  $\eta = 0,9$ .

À la figure 3.1a, on représente les courbes d'erreur entre l'approximation lorsque la stratégie d'adaptation temps est appliquée et la solution exacte. Nous utilisons successivement les semi-normes d'ordre 0 à 2. L'objectif est de vérifier, d'une part, que les pentes de convergence de l'erreur sont compatibles avec le résultat du théorème 2. D'autre part, nous cherchons à faire une première comparaison entre les différentes semi-normes. On observe ensuite une pente de 1 pour la convergence de la méthode en fonction de la tolérance relative  $\epsilon_R$ . Ce résultat est aussi en adéquation avec le théorème 2 pour ce qui est de la convergence. De plus, en comparant les comportements associés aux différentes semi-normes, on constate un écart entre les courbes de convergence. Cet écart est dû au fait que la constante  $C$  définie par la relation (3.10) dépend de l'ordre  $k$  de la semi-norme choisie. Par ailleurs, la figure 3.1b permet de comparer le nombre de pas de temps entre les différentes stratégies d'adaptation en temps. On constate que plus l'ordre  $k$  de la semi-norme est élevée, plus le coût en nombre de pas est grand. Ceci est prévisible car le choix des pas de temps est d'autant plus restrictif que la régularité des variations  $\epsilon$ -voisines est élevée. Dans le cas de la semi-norme d'ordre 2, un plateau est atteint lorsque le pas de temps minimal nécessaire devient plus petit que  $h_{\min}$ .

L'algorithme de sélection des pas de temps stagne alors à la valeur de  $h_{\min}$ .

En conclusion, ce premier exemple a permis de vérifier les résultats théoriques du chapitre 3 et d'identifier les premières différences entre les stratégies d'adaptation selon les semi-normes  $\Delta_k$ .

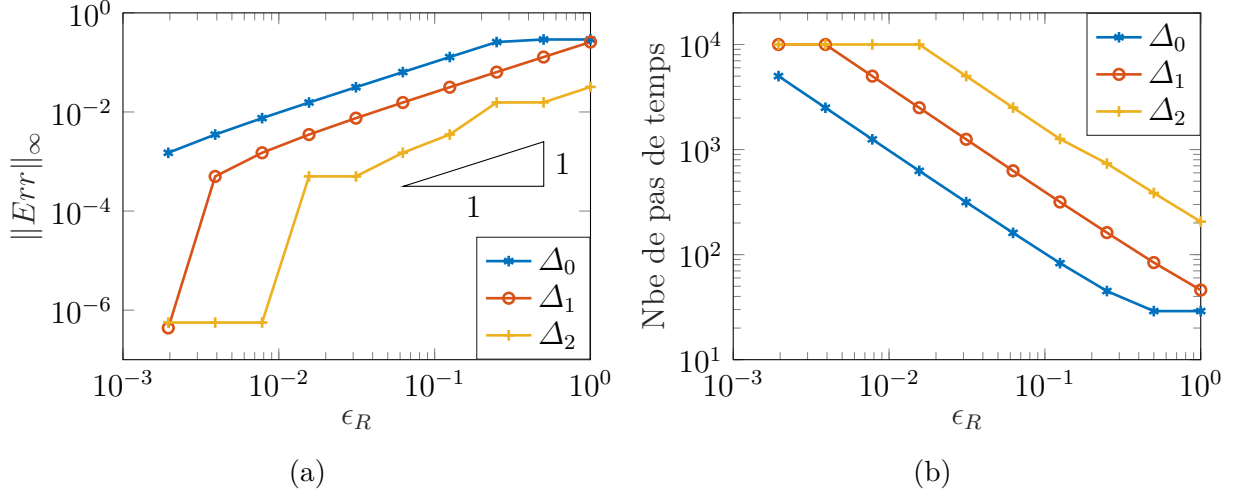


Figure 3.1: Comparaison des stratégies d'adaptation pour des variations d'ordre  $k = 0, 1, 2$ . (a) Précision de l'approximation. (b) Nombre de pas de temps en fonction de  $\epsilon_R$ .

**Exemple 3.4.2** (Équation de combustion). L'équation de combustion fournit un bon exemple d'équation raide. Cet exemple a été présenté pour la première fois par Shampine et al. [59] pour décrire des problèmes raides. L'équation décrit par exemple le comportement de flamme lorsqu'une allumette est craquée. La sphère de flamme connaît une expansion rapide avant d'atteindre une taille critique. Elle reste ensuite à cette taille puisque le niveau d'oxygène parvient à un niveau d'équilibre. La quantité d'oxygène consommée à l'intérieur de la sphère est égale à la quantité d'oxygène disponible sur sa surface. Un modèle simple est donné par :

$$\begin{cases} u' = u^2(1 - u), \\ u(0) = \delta. \end{cases} \quad (3.43)$$

On choisit une valeur  $\delta$  de  $10^{-4}$ . Au chapitre 2, nous avons expliqué que ce problème devient raide pour des valeurs de  $\delta \ll 1$ . Ce deuxième exemple va nous permettre de vérifier la précision de l'adaptation en temps sur un problème raide standard. L'efficacité de l'algorithme pourra également être analysée.

Le problème (3.43) est donc discrétisé à l'aide de la méthode BDF2 avec une tolérance fixe  $\epsilon_R = 1 \times 10^{-2}$ . Les paramètres d'adaptation pour l'algorithme de sélection des pas de temps

sont donnés par  $h_{\max} = 5000$ ,  $h_{\min} = 1 \times 10^{-5}$  et  $\eta = 0,9$ . On choisit de contrôler les variations jusqu'à l'ordre  $k = 1$ , i.e. avec la semi-norme  $\Delta_1$ , pour adapter en temps.

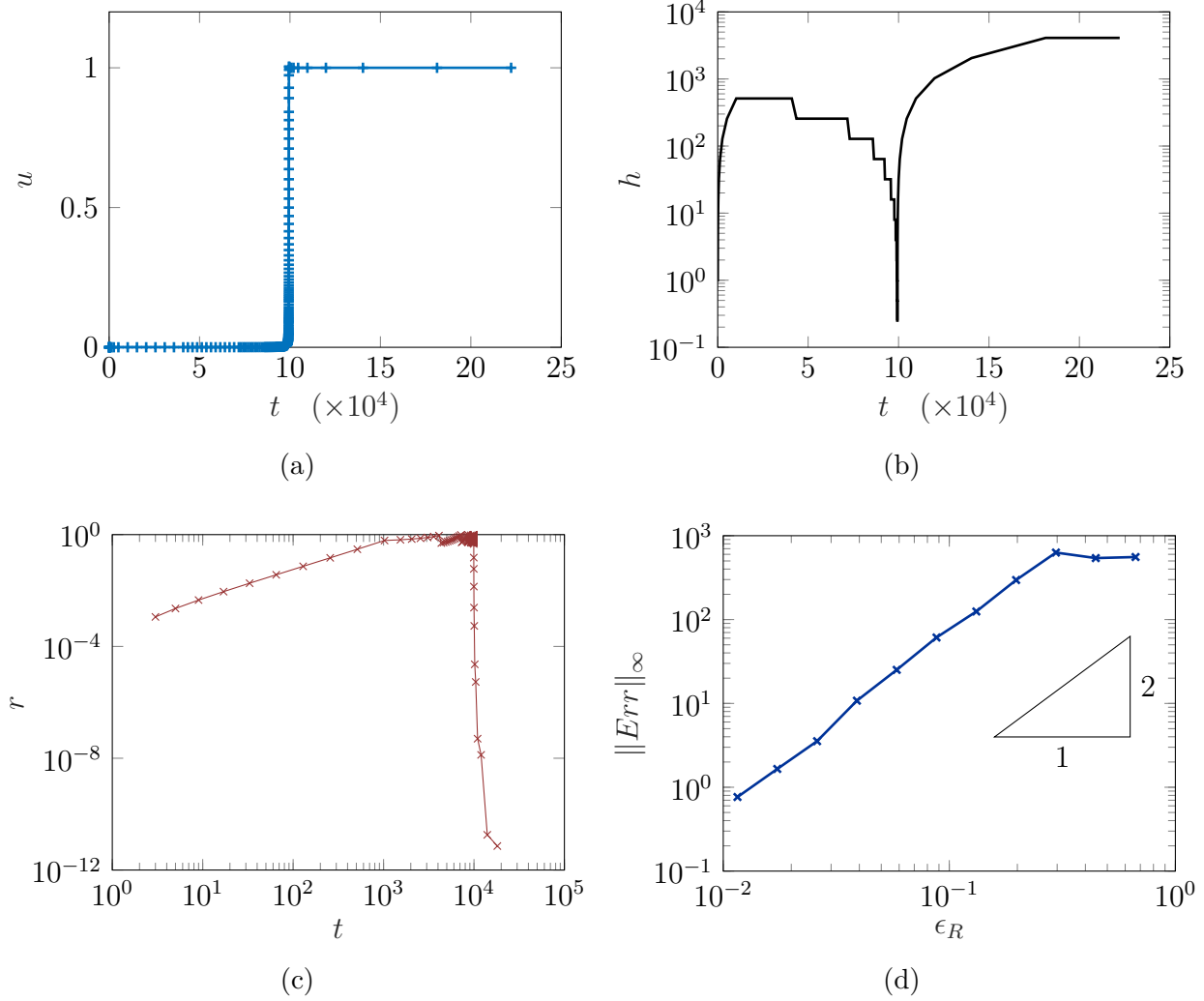


Figure 3.2: Discretisation de l'équation de combustion avec la semi-norme  $\Delta_2$ . (a) Solution discrète. (b) Pas de temps en fonction du temps. (c) Ratio de contrôle. (d) L'erreur relative en fonction de  $\epsilon_R$ .

On remarque à la figure 3.2a que l'algorithme d'adaptation produit de petits pas de temps dans la zone de transition, c'est-à-dire là où le comportement est le plus raide, et de grands pas de temps lorsque la solution varie très peu. Ceci illustre que la stratégie d'adaptation en temps permet de capturer des zones de front, i.e. des zones de variations brusques de la solution suivies de faibles variations. Cette caractéristique de l'algorithme est importante pour l'application du chapitre 4. La figure 3.2b représente la taille des pas de temps en fonction

du temps. Elle vient corroborer la sélection plus fine des pas  $h_n$  lorsque la solution varie de façon importante. Ce choix parcimonieux des pas de temps permet d'obtenir une stratégie d'adaptation en temps efficace. La figure 3.2c illustre la variation du ratio de contrôle  $r_n$  au cours du temps. L'algorithme de sélection des pas de temps parvient à maintenir le ratio inférieur à 1, ce qui est l'objectif recherché. Enfin, on observe à la figure 3.2d une convergence d'ordre 2 selon  $\epsilon_R$ . Il s'agit d'une meilleure convergence que celle attendue par la théorie, qui devrait être d'ordre 1. Ce gain de précision peut être dû à des annulations de termes dans le développement de l'erreur de locale. Ce résultat ne contredit cependant pas le théorème 2.

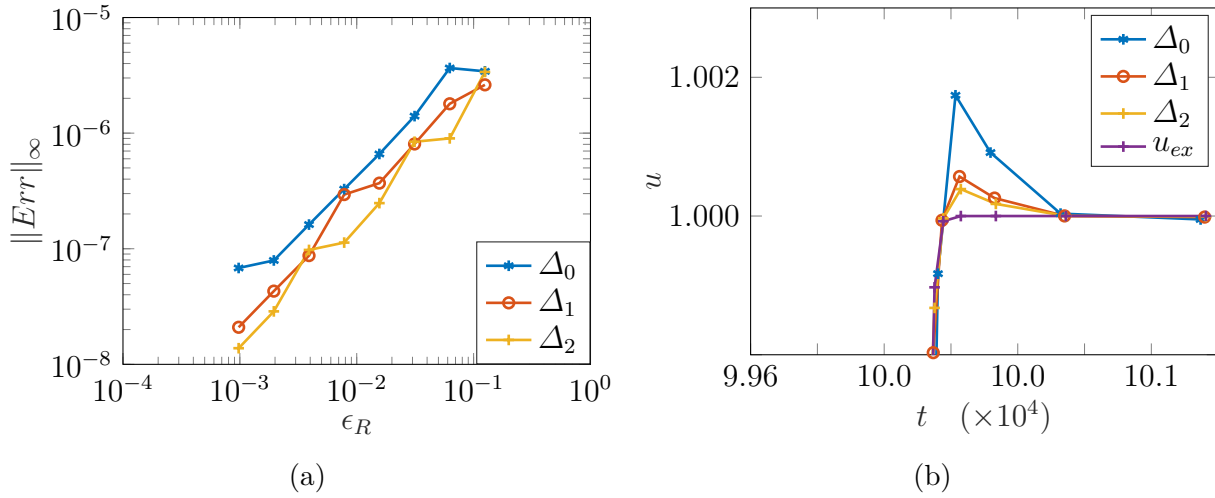


Figure 3.3: Comparaison (a) de la précision et (b) de la régularité de l'approximation pour des adaptations en variations d'ordre  $k = 0, 1, 2$ .

L'exemple de l'équation (3.43) permet également d'investiguer les propriétés de l'approximation obtenue en fonction de la semi-norme  $\Delta_k$ . La figure 3.3a compare la précision des solutions discrètes obtenues en fonction des variations contrôlées, pour une tolérance  $\epsilon_R = 0,9 \times 10^{-2}$  fixée. On constate que l'on obtient une erreur plus faible en choisissant une semi-norme d'ordre plus élevé. Ceci peut s'expliquer par la figure 3.3b qui illustre les approximations obtenues au niveau de la zone de transition. On constate que pour une même tolérance, et dans ce cas pour un même nombre de pas de temps, plus l'ordre de la semi-norme est élevé et plus la solution capturée à la zone de transition est régulière. Ainsi, choisir la semi-norme  $\Delta_2$  plutôt que  $\Delta_0$  permet de mieux contrôler la précision de l'approximation calculée.

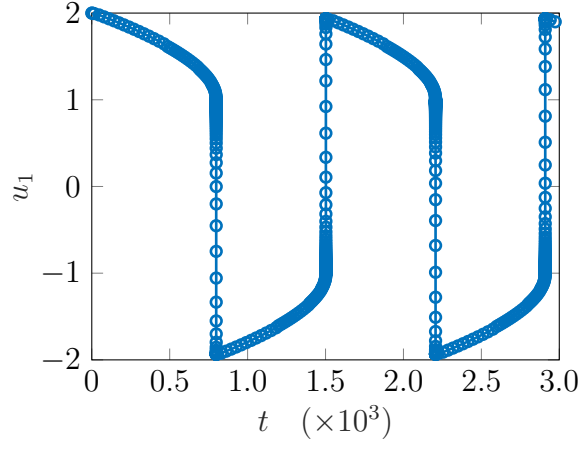
**Exemple 3.4.3** (Système dynamique de Van Der Pol). Un autre exemple de problème raide est fourni par l'équation de Van Der Pol qui décrit les phénomènes d'oscillations de relaxation

pour des circuits électriques :

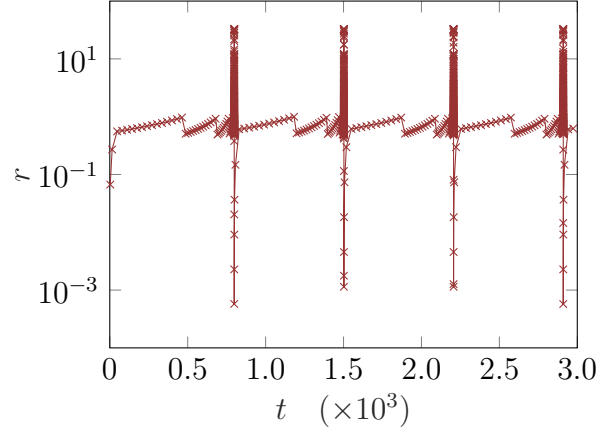
$$\begin{cases} u_1' = u_2, \\ u_2' = \mu(1 - u_1^2)u_2 - u_1, \\ u_1(0) = 2, \quad u_2(0) = 0. \end{cases} \quad (3.44)$$

La solution de l'équation de Van Der Pol, lorsqu'on se rapproche du cycle limite, possède des régions à variations brusques, où le problème n'est pas raide. Se mêlent à ces régions des zones où l'équation a un comportement raide mais où la solution possède des variations lentes. Le problème de Van Der Pol va nous permettre de vérifier la robustesse de notre stratégie d'adaptation en temps.

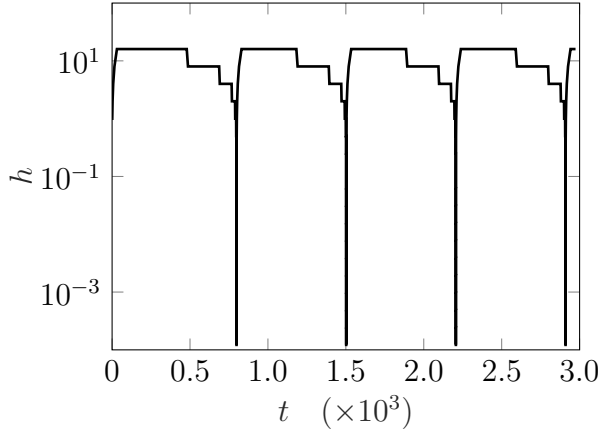
Le système (3.44) est discrétisé à l'aide de la méthode BDF2 et les paramètres d'adaptation pour l'algorithme de sélection des pas de temps sont  $h_{\max} = 100$ ,  $h_{\min} = 1 \times 10^{-4}$  et  $\eta = 0,9$ . On choisit de contrôler les variations jusqu'à l'ordre  $k = 2$ , donc avec la semi-norme  $\Delta_2$ , pour adapter en temps. La figure 3.4a illustre l'approximation obtenue avec un schéma BDF2 muni de la stratégie d'adaptation en temps proposée, pour une tolérance  $\epsilon_R = 0,9 \times 10^{-2}$  fixée. On constate que l'algorithme d'adaptation parvient à capturer les zones où le problème est raide, avec un pas de temps réduit. Des pas de temps plus larges sont utilisés dans des régions où la solution varie moins. On remarque cependant à la figure 3.4b que, dans les zones raides, l'algorithme d'adaptation ne parvient pas toujours à maintenir le ratio de contrôle en dessous de 1. Ceci est dû au fait que le pas de temps requis est plus petit que le pas minimal permis. Cette constatation est corroborée par l'observation de la figure 3.4c qui représente la taille des pas de temps en fonction du temps. Bien que dans ce cas la courbe d'erreur relative de la figure 3.4d présente des pics dans les zones raides, ce comportement peut être observé avec d'autres algorithmes d'adaptation tels que ode15s [47]. Notre algorithme d'adaptation en temps n'échoue cependant pas à poursuivre les calculs. Ceci confirme la robustesse de la stratégie mise en place.



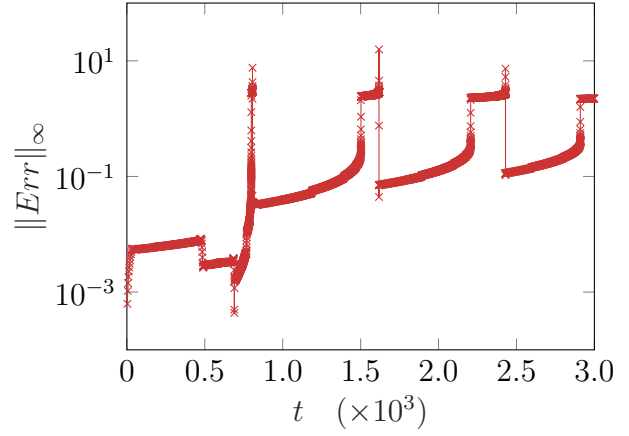
(a)



(b)



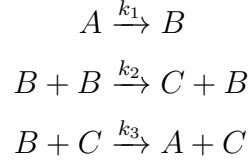
(c)



(d)

Figure 3.4: Discrétisation de l'équation de Van Der Pol avec  $\Delta_2$ . (a) Solution discrète. (b) Ratio de contrôle. (c) Pas de temps en fonction du temps. (d) L'erreur relative en fonction du temps.

**Exemple 3.4.4** (Réaction autocatalytique de Robertson). L'équation de Robertson décrit une réaction autocatalytique [55]. La structure de la réaction est décrite par les étapes :

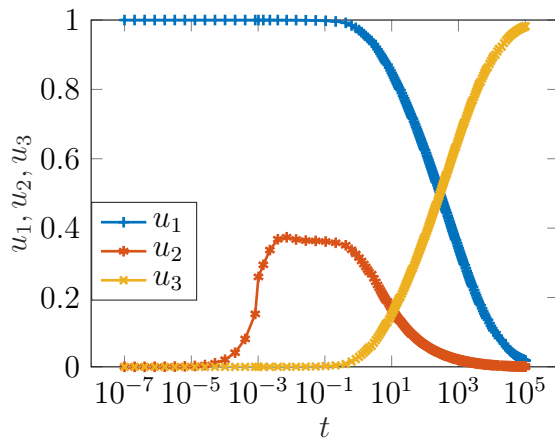


Les paramètres  $k_1, k_2$  et  $k_3$  sont les constantes de réactions et  $A, B$  et  $C$  sont les espèces chimiques impliquées. Sous certaines hypothèses de conservation, le système de Robertson s'écrit :

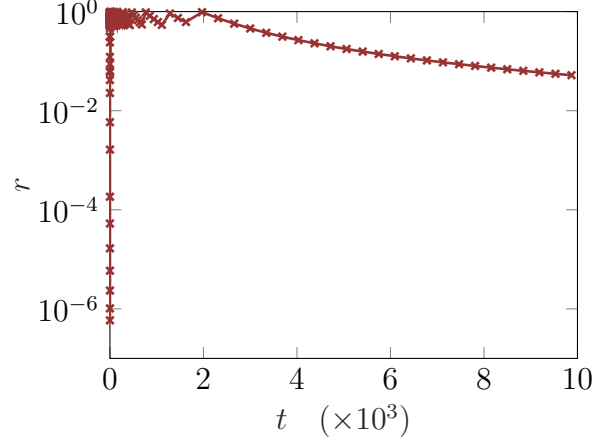
$$\begin{cases} u_1' = -k_1 u_1 + k_3 u_2 u_3, \\ u_2' = k_1 - k_3 u_2 u_3 - k_2 u_2^2, \\ u_3' = k_3 u_2^2, \\ u_1(0) = 1, \quad u_2(0) = 0, \quad u_3(0) = 0. \end{cases} \quad (3.45)$$

Ce système est populaire pour faire l'étude des intégrateurs d'équations raides. Le grand écart entre les valeurs numériques des constantes de réaction  $k_1 = 0,04$ ,  $k_2 = 3 \times 10^7$  et  $k_3 = 10^4$  vont permettre de vérifier la robustesse de l'algorithme. Comme pour plusieurs de réactions cinétiques, le système présente une phase transitoire très rapide suivie de variations plus lisses, où de plus grands pas de temps seraient plus appropriés.

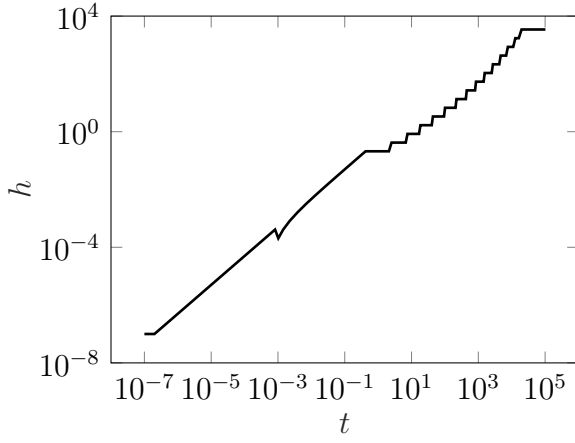
Le système (3.45) est discrétisé à l'aide de la méthode BDF2 pour une tolérance  $\epsilon_R = 0,9 \times 10^{-2}$  fixée. Les paramètres d'adaptation de l'algorithme sont  $h_{\max} = 5000$ ,  $h_{\min} = 1 \times 10^{-8}$  et  $\eta = 0,9$ . On choisit de contrôler les variations jusqu'à l'ordre 2, donc avec la semi-norme  $\Delta_2$ , pour adapter en temps. La figure 3.5a illustre la solution approximée du système obtenue. On remarque, en comparant avec la figure 3.5b, à l'aide du ratio de contrôle, que la phase de transition au début de la réaction correspond à la partie raide du problème. Cependant, l'algorithme d'adaptation en temps ne rencontre aucune difficulté à obtenir la solution discrète, comme en témoigne la figure 3.5c qui donne l'évolution du pas de temps. Ceci vient appuyer l'argument de robustesse de la stratégie d'adaptation. Enfin, la figure 3.5d montre une convergence d'ordre  $k = 1$  de l'erreur sur la solution du système (3.45) selon la tolérance  $\epsilon_R$ . Ce résultat vient également soutenir les propriétés de convergence de la stratégie d'adaptation en temps, démontrées au chapitre 3.



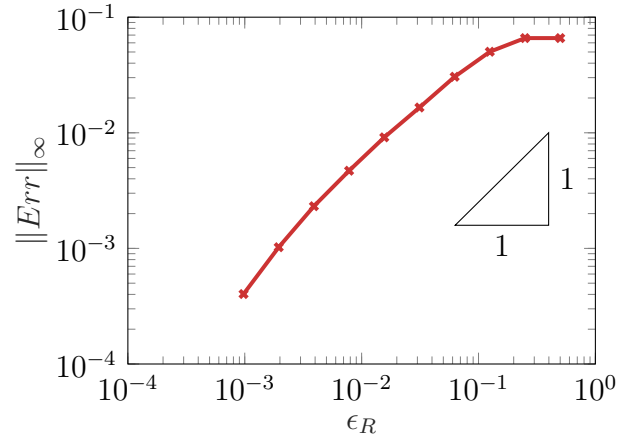
(a)



(b)



(c)



(d)

Figure 3.5: Discrétisation du système de Robertson avec  $\Delta_2$ . (a) Solution discrète. (b) Ratio de contrôle. (c) Pas de temps en fonction du temps. (d) Erreur relative en fonction de  $\epsilon_R$ .



**Exemple 3.4.5** (Oscillateur harmonique amorti). L'oscillateur harmonique amorti peut s'écrire comme un système autonome :

$$\begin{cases} u_1' = u_2, \\ u_2' = -\left(\frac{\gamma}{m}u_2 + \frac{k}{m}u_1\right), \\ u_1(0) = 1, \quad u_2(0) = 0. \end{cases} \quad (3.46)$$

où  $m$  est la masse d'un objet rattaché à un ressort de raideur  $k$ , avec un coefficient d'amortissement  $\gamma$ . On peut montrer (voir [67]) que la quantité

$$Q(u_1, u_2) := \frac{e^{\frac{\gamma}{m}T}}{2} (m(u_2(T))^2 + \gamma u_1 u_2(T) + k(u_1(T))^2) \quad (3.47)$$

est une quantité conservée, i.e. constante sur le flot de l'équation (3.46). Même si la fonction  $Q$  n'a aucune signification physique pertinente à notre connaissance, elle n'en demeure pas moins un exemple possible de quantité d'intérêt. En effet, rien a priori n'indique que les quantités conservées obtenues à partir de la solution exacte de (3.46) le sont encore pour l'approximation calculée. Il peut donc être intéressant d'adapter en temps pour s'assurer que cette quantité d'intérêt demeure conservée en tout temps. On notera par ailleurs que telle que définie, la quantité conservée ne vérifie pas les hypothèses du théorème 4. On montrera numériquement que cela n'entrave pas la convergence de l'algorithme.

Dans cette section on compare donc des stratégies d'adaptation en temps basées, d'une part, sur les variations de l'approximation  $u$  avec la semi-norme,  $\Delta_1(u)$ , et d'autre part, sur les variations de la quantité d'intérêt  $Q(u)$  avec la semi-norme  $\Delta_1(Q(u))$ . Le système (3.46) est discrétisé à l'aide de la méthode BDF2 et les paramètres d'adaptation pour l'algorithme de sélection de pas de temps sont  $h_{\max} = 5$ ,  $h_{\min} = 1 \times 10^{-5}$  et  $\eta = 0,9$ .

La figure 3.6a illustre l'erreur sur la solution et sur la quantité d'intérêt en fonction de la tolérance  $\epsilon_R$ . On adapte dans ce cas selon les variations de la solution  $u$ . On observe alors une convergence avec une pente de 1 pour l'erreur sur la solution. Ce résultat est en accord avec le théorème 2. On constate également une erreur d'ordre 1 sur la quantité d'intérêt.

La figure 3.6b représente l'erreur sur la solution et sur la quantité d'intérêt en fonction de la tolérance  $\epsilon_R$ . On adapte cette fois selon les variations de la quantité d'intérêt. On observe une convergence avec une pente de 1 pour l'erreur sur cette quantité. Ce résultat est en accord avec le théorème 4. On constate également une erreur d'ordre 1 sur la solution.

On note également que les résultats présentés des figures 3.6 sont obtenus avec des pas de temps du même ordre de grandeur. On a donc, dans ce cas, une adaptation en fonction de la solution et une adaptation en fonction de la quantité d'intérêt qui sont équivalentes du point

de vue de l'ordre de convergence et de l'efficacité. Cet exemple permet néanmoins de vérifier la stratégie de sélection du pas de temps selon les variations de  $Q$ .

En conclusion, on dispose d'une stratégie d'adaptation consistante, efficace et robuste, basée sur le concept de variations  $\epsilon$ -voisines de la solution discrète. Cet algorithme s'adapte aisément à une quantité d'intérêt du problème étudié. Au chapitre suivant, nous appliquons cette stratégie au problème de supraconductivité.

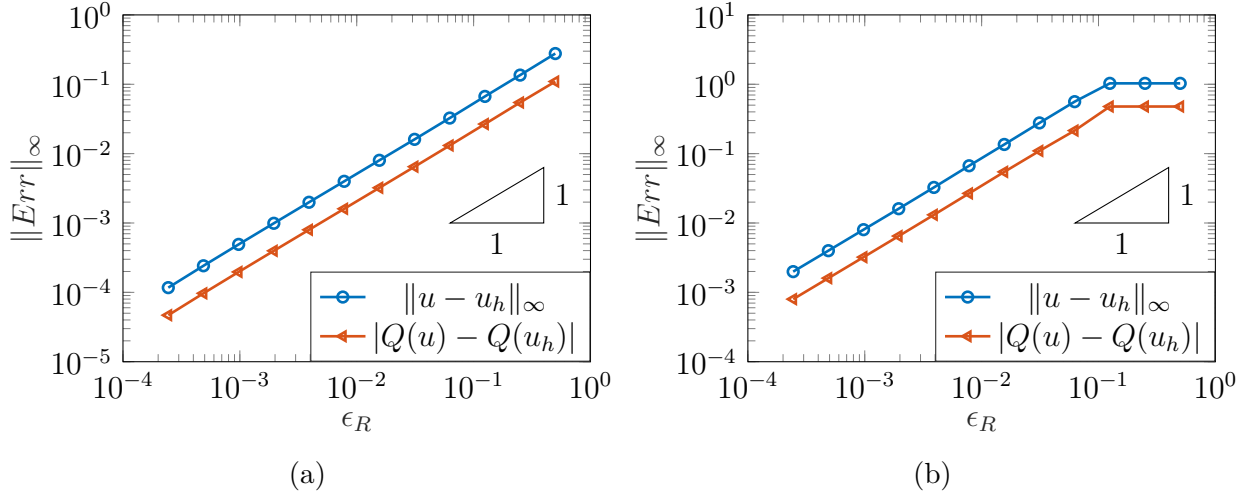


Figure 3.6: Comparaison de la précision sur la solution discrète et sur la quantité d'intérêt  $Q$  (a) selon une adaptation en fonction des variations de la solution et (b) en fonction de la quantité d'intérêt.

## CHAPITRE 4 APPLICATION AUX ÉQUATIONS DE SUPRACONDUCTIVITÉ

La dernière décennie a été marquée par l'intensification de la recherche portant sur les câbles supraconducteurs à haute température, tels que les filaments de  $\text{MgB}_2$ . Il s'agit d'augmenter le courant critique pouvant être transporté et produire des fils supraconducteurs toujours plus longs. Il existe de nombreuses utilisations des câbles supraconducteurs comme pour les transformateurs supraconducteurs [33] ou encore pour les câbles électriques à haute tension [4]. Le coût des composants supraconducteurs sont reliés aux pertes par dissipation thermique de l'effet Joule. Cette perte peut être mesurée à l'aide de la perte de courant alternatif (en anglais « AC-loss ») [69]. Les pertes AC à l'intérieur d'un composant peuvent être induites d'une part par un transport de courant alternatif, dite perte par transport de courant, et d'autre part par un champ magnétique instationnaire, dite alors perte par magnétisation.

La perte par magnétisation inclut les pertes hystérétiques, les pertes dues aux courants de Foucault et les pertes dues au couplage magnétique entre les filaments [51]. Les pertes hystérétiques dépendent de la géométrie du supraconducteur et sont générées par la pénétration du flux magnétique et ses variations à l'intérieur du HTS. Les pertes dues aux courants de Foucault se produisent dans la gaine du câble supraconducteur [25]. La gaine, la barrière de protection et l'enchevêtrement des filaments influencent donc les pertes AC dans un fil supraconducteur [57]. Il est donc intéressant d'avoir à sa disposition un modèle numérique de HTS afin d'estimer les pertes AC pour différentes structures de câbles.

Il existe différentes stratégies de modélisation numériques des HTS développées au sein de la communauté des ingénieurs électrique. La plupart d'entre elles reposent sur une réécriture des équations de Maxwell en basse fréquence sous la forme d'un problème parabolique. Le système parabolique, du point de vue du champ magnétique, ainsi obtenu est communément appelée «formulation en  $\mathbf{H}$ » [6]. Cette formulation peut être discrétisée en utilisant des méthodes adaptées aux équations paraboliques comme celles des éléments finis avec une formulation de Galerkin, comme c'est le cas dans certains codes commerciaux. Une autre approche est de discrétiser directement les équations de Maxwell en basse fréquence en se servant de méthodes numériques pour des systèmes hyperboliques telles que la méthode des éléments finis de Galerkin discontinue [45]. Les stratégies adaptées aux systèmes hyperboliques permettent de capturer avec précision les forts gradients du vecteur densité de courant lorsque les valeurs des exposants de la loi de puissance deviennent élevées. En outre, comme la perte AC est calculée en effectuant le produit scalaire entre le champ électrique et le vecteur densité de courant, notre hypothèse est que les méthodes adaptées aux équations hyperbo-

liques devraient permettre une meilleure estimation de cette perte AC.

Toutes les méthodes évoquées précédemment conduisent à un problème semi-discret en temps qui est raide. Ceci est en particulier dû à un conditionnement des matrices d'assemblage qui est souvent élevé mais aussi aux forts gradients observés. Ces systèmes raides nous forcent alors à choisir des pas de temps très petits. Il est donc plus avantageux de recourir à une stratégie d'adaptation en temps. De plus, le fait d'avoir une quantité d'intérêt pertinente, la perte AC dans notre cas, va permettre d'appliquer les différentes méthodes d'adaptation développées dans ce travail.

Le chapitre débute donc avec une introduction aux équations de Maxwell pour des conducteurs standards, pour des supraconducteurs, puis dans un milieu avec une faible conductivité comme l'air. On poursuit avec l'application des stratégies d'adaptation en temps dans le cas simple d'un plan supraconducteur. Enfin, on simule numériquement les équations de Maxwell de la supraconductivité pour un filament de  $\text{MgB}_2$  en 3D avec la méthode de Galerkin discontinue. On évalue les pertes AC dans ce cas plus réaliste avec une adaptation en quantité d'intérêt.

#### 4.1 Définition du problème

Dans l'hypothèse magnéto-quasistatique, les équations de Maxwell pour l'approximation en basse fréquence et pour un conducteur standard, s'écrivent [45] :

$$\begin{cases} \partial_t (\mu \mathbf{H}) + \nabla \times \mathbf{E} = 0, \end{cases} \quad (4.1)$$

$$\begin{cases} \mathbf{E} - \rho \nabla \times \mathbf{H} = 0, \end{cases} \quad (4.2)$$

$$\begin{cases} \nabla \cdot (\mu \mathbf{H}) = 0. \end{cases} \quad (4.3)$$

Ici,  $\mathbf{H}$  est le champ magnétique,  $\mathbf{E}$  est le champ électrique,  $\mu$  est la perméabilité magnétique et  $\rho$  est la résistivité électrique. Le vecteur de densité de courant s'évalue à l'aide de la loi d'Ampère,

$$\mathbf{J} = \nabla \times \mathbf{H}. \quad (4.4)$$

On rappelle que dans ce travail, l'hypothèse d'un milieu isotherme est considérée. Dans le cas d'un conducteur standard, la résistivité est constante. Pour un supraconducteur, en revanche, la résistivité électrique dépend du vecteur de densité de courant et peut être modélisé à l'aide de la loi de puissance

$$\rho(\nabla \times \mathbf{H}) = \frac{E_c}{J_c^p} \|\nabla \times \mathbf{H}\|_2^{p-1}. \quad (4.5)$$

Dans ce cas,  $J_c$  est la densité de courant critique,  $E_c$  est l'amplitude du champ électrique lorsque le courant critique est atteint et  $p$  est un paramètre.

À ce problème défini sur un domaine  $\Omega$ , on impose les conditions aux frontières :

$$\begin{cases} \mathbf{n} \times \mathbf{H} = g_1, & \text{sur } \Gamma \times \text{I}, \\ \mathbf{n} \times \mathbf{E} = g_2, & \text{sur } \Gamma \times \text{I}. \end{cases} \quad (4.6)$$

où  $\mathbf{n}$  est le vecteur normal extérieur à la frontière  $\Gamma$  de  $\Omega$  et  $\text{I} = [0, T]$  est l'intervalle de temps. Les conditions initiales sont  $\mathbf{H}(\mathbf{x}, 0) = \mathbf{H}_0(\mathbf{x})$  et  $\mathbf{E}(\mathbf{x}, 0) = \mathbf{E}_0(\mathbf{x})$ .

Lorsqu'on modélise l'air entourant le câble supraconducteur, la résistivité électrique est plusieurs ordres de grandeur plus grande que dans le matériau conducteur. Les équations de Maxwell qui régissent le comportement électromagnétique dans l'air prennent alors la forme [45] :

$$\begin{cases} \mu \partial_t \mathbf{H} + \nabla \times \mathbf{E} = 0, \\ \nabla \times \mathbf{H} = 0, \end{cases} \quad (4.8)$$

$$\begin{cases} \nabla \cdot \mathbf{H} = 0, \end{cases} \quad (4.9)$$

$$\begin{cases} \nabla \cdot \mathbf{E} = 0. \end{cases} \quad (4.10)$$

$$\begin{cases} \nabla \cdot \mathbf{E} = 0. \end{cases} \quad (4.11)$$

## 4.2 Calcul de la perte AC

La quantité d'intérêt naturelle pour ce problème est la perte AC. Celle-ci est définie comme

$$\int_{\Omega} \mathbf{E} \cdot \mathbf{J} \, d\Omega = \int_{\Omega} \rho (\nabla \times \mathbf{H}) \cdot (\nabla \times \mathbf{H}) \, d\Omega \quad (4.12)$$

Dans ce qui suit, nous considérerons un câble formé d'une partie conductrice et d'une partie supraconductrice. Dans ce cas particulier, la perte AC se décompose comme

$$\int_{\Omega} \mathbf{E} \cdot \mathbf{J} \, d\Omega = \int_{\Omega_{\text{HTS}}} \frac{E_c}{J_c^p} \|\mathbf{J}\|_2^p \, d\Omega + \int_{\Omega_C} \rho \|\mathbf{J}\|_2^2 \, d\Omega \quad (4.13)$$

où  $\Omega_C$  est le domaine du conducteur standard (barrière et gaine) et  $\Omega_{\text{HTS}}$  est le domaine correspondant au supraconducteur.

## 4.3 Cas d'une plaque infinie supraconductrice

Il est intéressant de vérifier au préalable l'adaptation en temps sur un cas simple de matériau supraconducteur. Si l'on considère la section droite d'un câble supraconducteur, celle-ci peut être vue en première approximation comme un plan infini. Les symétries de ce plan permettent de conclure que les champs électrique et magnétique ont les formes respectives  $\mathbf{E}(\mathbf{x}, t) = (0, 0, E(x, t))$  et  $\mathbf{H}(\mathbf{x}, t) = (0, H(x, t), 0)$ . La condition de divergence nulle est de plus toujours vérifiée pour des équations en une dimension avec une perméabilité magnétique constante.

On s'intéresse donc au champ magnétique d'une plaque supraconductrice infinie [26]. Dans ce cas, les équations de Maxwell (4.3) se simplifient et le champ magnétique est alors solution de

$$\partial_t(\mu H) - \partial_x(\alpha |\partial_x H|^{p-2} \partial_x H) = 0, \quad \forall (x, t) \in \Omega \times I. \quad (4.14)$$

Une solution analytique de l'équation (4.14) existe et est donnée dans [44] par

$$H(x, t) = \begin{cases} C t^{\frac{p-1}{p-2}} \left(1 - \frac{\gamma x}{\xi_0 t}\right)^{\frac{p-1}{p-2}} & \text{si } 0 \leq x \leq x_f(t), \\ 0 & \text{si } x > x_f(t), \end{cases}$$

où  $C$  est une constante et

$$\xi_0 = \left(\frac{p-1}{p-2}\right)^{1-\frac{1}{p}}, \quad \gamma = \frac{C^{\frac{2}{p}-1}}{\alpha^{\frac{1}{p}}}, \quad \text{et } \alpha = \frac{E_c}{\mu_0 J_c^{p-1}}.$$

La position du front du champ magnétique pénétrant dans le matériau est représentée par  $x_f(t) = \frac{\xi_0}{\gamma} t$ . On choisit comme domaines d'intégration  $\Omega = [0, 1]$  et  $I = [0, 0, 5]$ . Les paramètres de la solution analytique sont choisis tels que  $C = 1$ ,  $E_c = 1$ ,  $J_c = 1$  et  $\mu = 1$ .

L'objectif de cette section est d'utiliser la stratégie d'adaptation en temps développée et vérifiée au chapitre 3 pour discrétiser l'équation (4.14). Ce problème est une étape préliminaire avant l'étude du système plus réaliste d'un câble supraconducteur. La discrétisation en espace de l'équation (4.14) a été faite à l'aide de la méthode des éléments finis avec une taille de maillage de 0,05. L'intégration en temps de l'équation résultante s'est faite, quant à elle, à l'aide de la méthode BDF2. Les paramètres d'adaptation pour l'algorithme de sélection sont les pas de temps  $h_{\max} = 1 \times 10^{-2}$  et  $h_{\min} = 1 \times 10^{-5}$  et la borne inférieure  $\eta = 0,9$ .

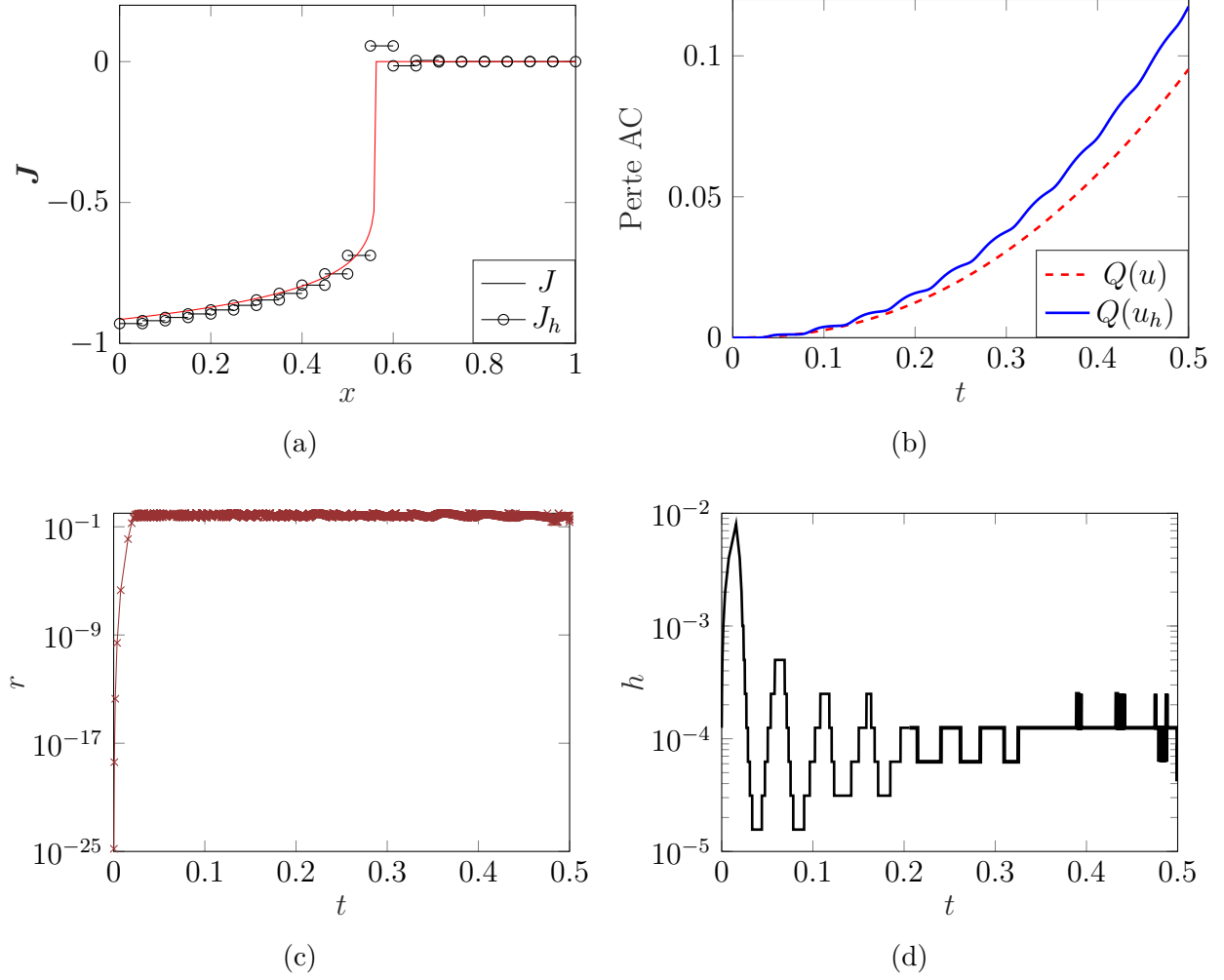


Figure 4.1: Discretisation du problème de la plaque 1D avec une stratégie d'adaptation selon les variations de la perte AC. (a) Densité de courant  $J$ . (b) Perte AC. (c) Ratio de contrôle. (d) Pas de temps.

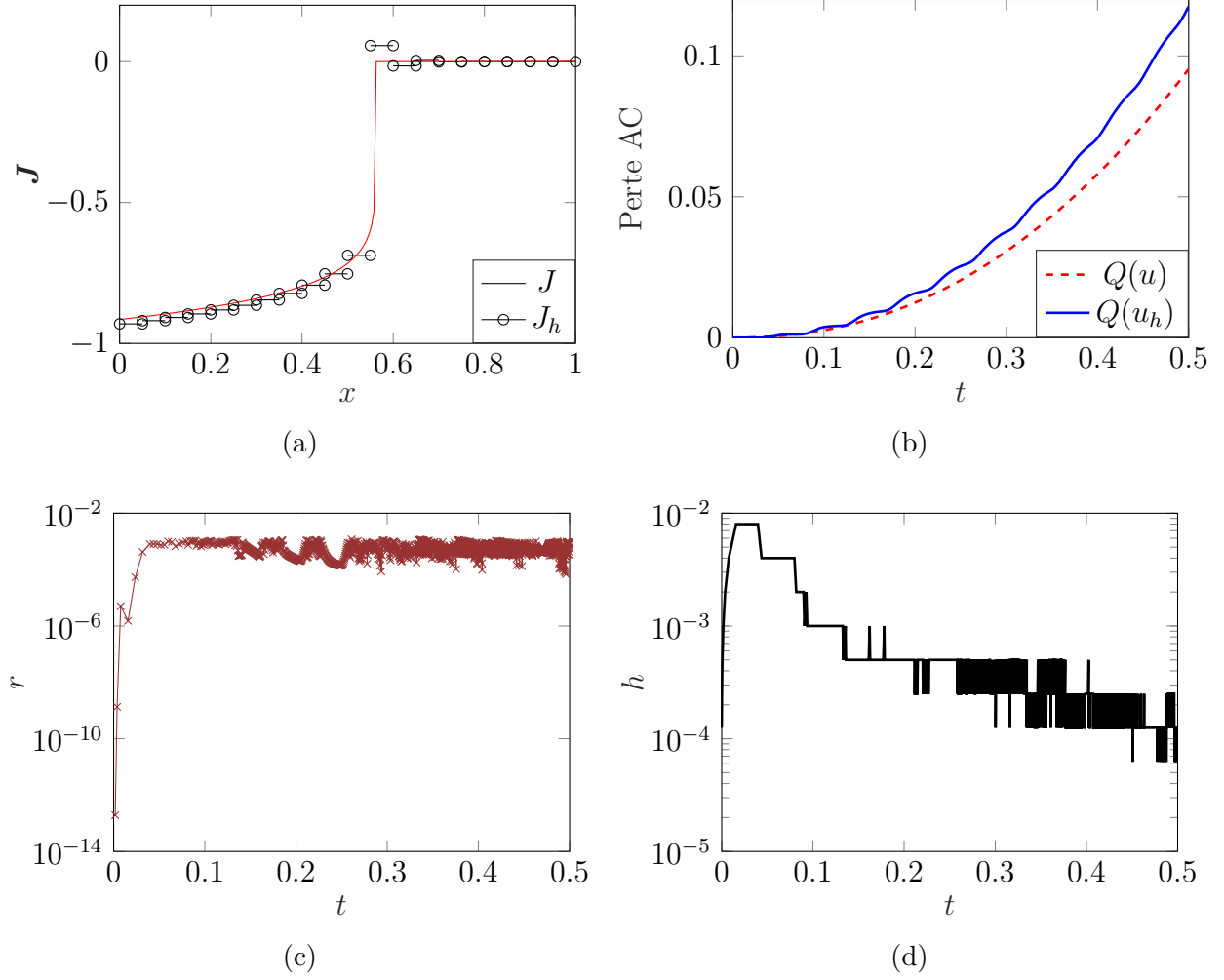


Figure 4.2: Discrétisation du problème de la plaque 1D avec une stratégie d'adaptation selon les variations des vecteurs  $(\mathbf{E}, \mathbf{H})$ . (a) Densité de courant  $\mathbf{J}$ . (b) Perte AC. (c) Ratio de contrôle. (d) Pas de temps.



On compare d'abord la discrétisation de l'équation (4.14) suivant une stratégie d'adaptation en temps basée, d'une part sur les variations de la solution  $(\mathbf{E}, \mathbf{H})$  et d'autre part sur les variations de la perte AC, qui est ici la quantité d'intérêt. Les figures 4.1a et 4.2a représentent l'évolution de la densité de courant obtenue le long du domaine  $\Omega$ , au temps  $t = 0,5$ . Ces approximations sont comparées avec la solution analytique de la densité de courant. On observe un front qui se déplace dans le matériau. Ce front n'est pas sans rappeler la solution obtenue pour le problème de combustion. C'est aussi l'une des motivations pour laquelle nous nous sommes intéressés aux problèmes raides au chapitre 3. Or, la stratégie de contrôle en pas de temps, développée dans ce travail, est adaptée aux équations raides. Les figures 4.1c et 4.2c donnent le ratio de contrôle pour l'adaptation en pas de temps selon la quantité d'intérêt et selon la solution. Elles montrent que l'algorithme parvient toujours à maintenir le critère  $r$  inférieur à 1. Les figures 4.1b et 4.2b représentent les courbes de pertes AC ainsi obtenues. Ces courbes sont comparées avec la solution analytique de la perte de courant. On constate que l'erreur par rapport à la solution analytique est non négligeable. Cependant, on ne peut évoquer l'erreur sur la perte AC sans prendre en compte l'erreur de discrétisation en éléments finis, qui est ici dominante. Afin de ne considérer que les effets des erreurs de la discrétisation en temps, on observe numériquement que la taille des éléments nécessaire est beaucoup plus petite que la valeur de 0,05 utilisée ici. Les coûts en temps de calculs deviennent dans ce cas plus importants que les ressources de calculs mobilisées.

La comparaison des stratégies de sélection des pas de temps des figures 4.1d et 4.2d illustrent les pas de temps obtenus à l'aide de chaque stratégie. Cependant, les deux méthodes de discrétisation conduisent à obtenir des valeurs de pertes AC semblables. En revanche, la comparaison du coût de calcul indique que l'adaptation en fonction de la perte AC a nécessité 6836 pas de temps contre 1513 pas de temps pour l'adaptation en fonction de la solution. Ainsi, même si l'adaptation selon la solution semble moins coûteuse que celle en quantité d'intérêt, elle a néanmoins rejeté plus de pas de temps, 1462, contre 306 pas de temps rejetés pour la seconde adaptation. Ceci peut également se voir sur la figure 4.2d. On peut expliquer ce comportement par le fait que la raideur du problème se retrouve au niveau de la densité de courant, qui est naturellement reliée à la perte AC. La solution  $(\mathbf{E}, \mathbf{H})$  présente quand à elle des variations plus régulières.

En conclusion, le modèle simplifier de câble supraconducteur nous a permis de vérifier que la stratégie d'adaptation basée sur la notion de variations  $\epsilon$ -voisines peut s'appliquer dans le cas des équations de supraconductivité. Des adaptations en temps en fonction de la perte AC et des variations des champs électromagnétiques ont pu être menées. Un gain de temps de calcul a été obtenu par rapport aux discrétisations à pas constants, qui nécessitent un pas  $h = 1 \times 10^4$ , donc 50000 pas de temps pour effectuer le calcul.

#### 4.4 Application à la modélisation d'un câble supraconducteur

Nous présentons maintenant la modélisation d'un câble supraconducteur ainsi que les résultats obtenus pour la perte AC lorsqu'on applique la stratégie d'adaptation en temps basée sur cette quantité d'intérêt. L'objectif est de simuler le comportement d'un câble supraconducteur plongé dans l'air afin de reproduire les conditions expérimentales dans lesquelles sont mesurées les pertes AC. On choisit pour la géométrie un domaine  $\Omega$  rectangulaire, qui représente le système global air et câble. Ce domaine est décomposé en deux sous-domaines  $\Omega_1$  et  $\Omega_2$  tels que  $\Omega_1 \cup \Omega_2 = \Omega$  et  $\Omega_1 \cap \Omega_2 = \emptyset$ . Le domaine  $\Omega_1$  est occupé par l'air et le champ électromagnétique y vérifie le système (4.8)-(4.11). Le domaine  $\Omega_2$  est occupé par le câble et les champs électromagnétiques y vérifient le système (4.3). Le domaine  $\Omega_2$  est lui-même subdivisé en 3 sous-domaines disjoints,  $\Omega_{\text{HTS}}$ ,  $\Omega_{\text{Barrière}}$  et  $\Omega_{\text{Gaine}}$ . En effet, le câble est formé par trois matériaux, le HTS, la barrière conductrice et la gaine, chacun occupant un sous-domaine tel qu'illustré à figure 4.3.

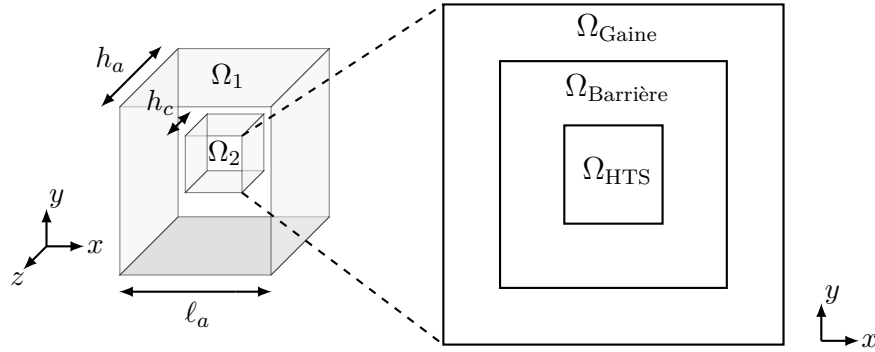


Figure 4.3: Géométrie du problème de la modélisation câble supraconducteur.

Les paramètres géométriques utilisés pour le domaine  $\Omega_1$  occupé par l'air sont, une longueur  $\ell_a = 2$  mm et une hauteur  $h_a = 1$  mm. Le domaine  $\Omega_2$  est cubique avec une longueur d'arête  $h_c = 0,5$  mm. L'ensemble du domaine  $\Omega$  est centré en  $(0, 0, 0)$ .

Dans cette partie, un solveur numérique est mis à notre disposition pour obtenir la semi-discrétisation en espace du système d'équations (4.3) via la méthode des éléments finis de Galerkin discontinue. L'intégration en temps se fait à l'aide du schéma BDF1 présent dans le code, auquel on ajoute la stratégie d'adaptation en temps développée dans ce travail. Les paramètres de contrôle sont  $h_{\max} = 1 \times 10^{-2}$ ,  $h_{\min} = 1 \times 10^{-8}$ ,  $\eta = 0,9$  et  $\epsilon_R = 0,9$ . La sélection du pas de temps est effectuée selon la semi-norme  $\Delta_1$  de la perte AC. En effet, comme évoquée en début de chapitre, il s'agit de la quantité physique pertinente pour ce problème de supraconductivité.

La figure 4.4a représente le ratio de contrôle et la figure 4.4b représente le pas de temps en

fonction du temps. On constate sur les figures 4.4a et 4.4b un comportement similaire à celui observé pour le système de Van Der Pol. La stratégie d'adaptation maintient le plus souvent le ratio en dessous de 1 avec de plus petits pas de temps dans les régions raides. Dans certains cas, le pas de temps requis est plus petit que le pas de temps minimal, ce qui conduit à des valeurs du ratio largement au dessus des valeurs permises. Néanmoins, lorsque l'on compare la perte AC obtenue avec la perte AC calculée en utilisant un pas constant  $h = 1 \times 10^{-4}$  à la figure 4.5, on note que cela n'entache pas de manière significative la qualité des résultats. Cependant, comme pour l'exemple du plan infini, l'erreur de discrétisation en espace doit être prise en compte. À la différence de la section précédente, on ne dispose pas de solution analytique pour effectuer une comparaison. La taille d'éléments nécessaire pour obtenir une courbe de perte AC suffisamment précise rend les calculs numériques impraticables avec les ressources en calcul disponibles.

Par ailleurs, la discrétisation est effectuée en 1920 pas de temps avec la stratégie d'adaptation en temps en quantité d'intérêt, contre 30 000 avec un pas de temps de  $h = 1 \times 10^{-4}$ . On passe alors d'un temps de simulation dépassant 1 semaine à seulement 41 heures avec la stratégie d'adaptation. En conclusion, la stratégie d'adaptation développée dans ce travail a permis de discrétiser les équations de supraconductivité de manière efficace tout en obtenant une approximation des pertes AC dans le matériau.

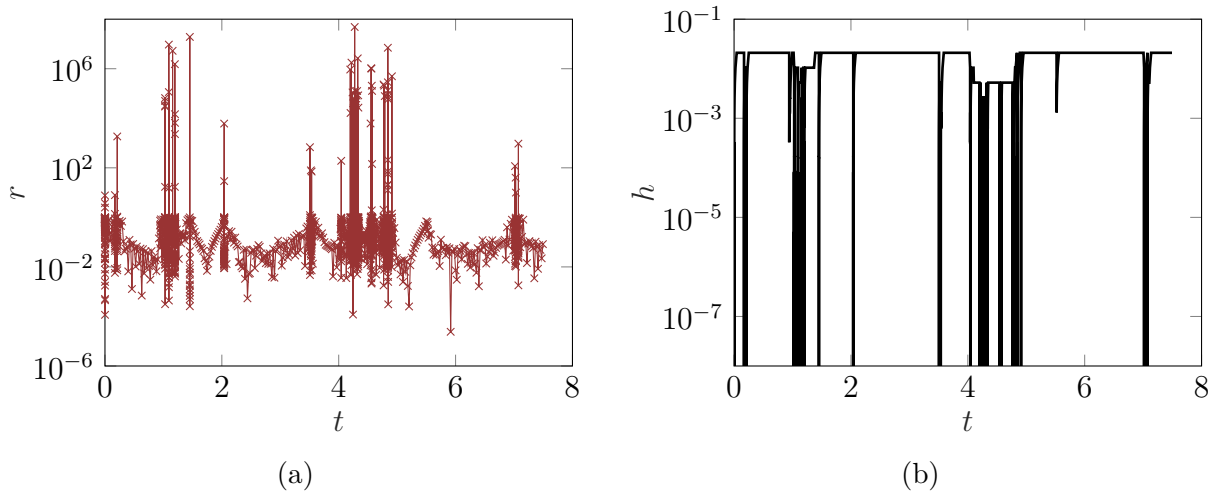


Figure 4.4: Discretisation du problème de câble supraconducteur 3D avec adaptation en temps selon la perte AC. (a) Ratio de contrôle et (b) pas de temps en fonction du temps.

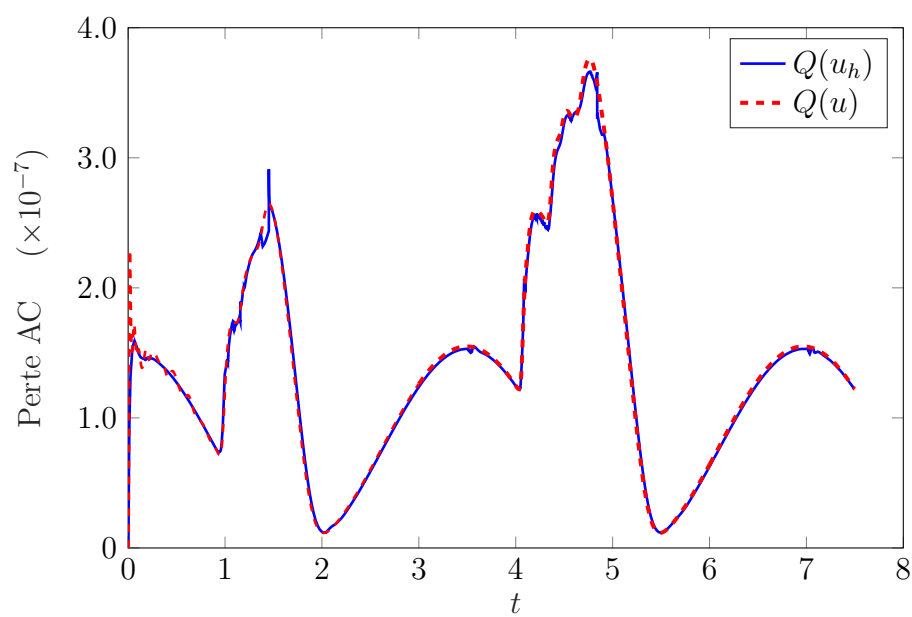


Figure 4.5: Comparaison des pertes de courant AC avec une adaptation en temps selon les variations de la quantité d'intérêt et une résolution à pas constant  $h = 10^{-4}$ .

## CHAPITRE 5 CONCLUSION

L'objectif principal de ce mémoire était de développer une stratégie d'adaptation en temps basée sur un estimateur d'erreur, pour la semi-discrétisation en éléments finis des équations de Maxwell pour la supraconductivité. Afin de synthétiser les travaux de recherche effectués, nous allons conclure sur chacun des objectifs spécifiques que nous nous sommes fixés :

- i) *Démontrer un lien entre l'estimateur d'erreur et l'erreur locale afin de prouver la cohérence de la stratégie d'adaptation en temps :*

Nous avons généralisé un estimateur d'erreur basé sur les variations de la solution approchée et démontré un résultat original reliant cet estimateur à l'erreur locale. Ceci nous a permis d'introduire le concept de solution à variations  $\epsilon$ -voisines. Nous avons ensuite montré que cette stratégie d'adaptation en temps dépend du schéma numérique utilisé et qu'elle prend donc en compte la précision de la méthode utilisée. Enfin, nous avons développé un algorithme basé sur le contrôle des variations  $\epsilon$ -voisines et discuté des résultats théoriques obtenus ;

- ii) *Démontrer que la stratégie développée peut s'appliquer à une quantité d'intérêt physique pertinente pour une équation donnée :*

Nous avons montré que le contrôle des variations d'une quantité d'intérêt permet d'aboutir également à une stratégie d'adaptation en temps. Pour ce faire, un résultat nouveau de convergence a été démontré à l'aide d'une inégalité entre l'estimateur d'erreur sur la solution discrète et sur la quantité d'intérêt, sous certaines conditions de régularité. Il est donc possible d'adapter l'algorithme de contrôle du pas de temps de manière à obtenir une précision satisfaisante sur la quantité d'intérêt choisie. Ceci résultat ouvre donc la voie pour le développement d'une nouvelle famille d'estimateurs d'erreurs basés sur une quantité d'intérêt pour les méthodes de différences finies.

- iii) *Vérifier la méthode développée sur des problèmes raides standards de la littérature :*

Après avoir défini la notion de problème raide, l'algorithme d'adaptation en temps a été testé sur un groupe de problèmes standards dans l'analyse de schéma de contrôle en temps. On a ainsi pu vérifié l'efficacité de la méthodologie développée pour les problèmes raides. La stratégies d'adaptation se montre permet de réduire significativement le nombre de pas de temps en comparaison avec des méthodes classiques comme RK45 pour des problèmes raides. De plus, il a été montré que la technique de contrôle proposée a démontré sa stabilité puisque l'ensemble des équations raides qui ont été discrétisées, en des temps acceptables. Enfin, une premier test, à l'aide du problème de l'oscillateur amorti, a permis de vérifier l'adaptation en quantité d'intérêt et de la comparer avec

l'adaptation basée sur la variable dépendante.

- iv) *Vérifier la stratégie d'adaptation en temps sur un problème d'électromagnétisme modélisant un matériau supraconducteur :*

L'approche proposée a été vérifiée en modélisant les équations de Maxwell pour estimer la perte de courant alternatif de câbles électriques. Deux géométries ont été utilisées, l'une en 1D pour avoir un cas simple du modèle de Bean et l'autre 3D afin de vérifier la méthode sur un cas réaliste de câble. Pour le modèle 3D, un agencement d'un conducteur standard et d'un supraconducteur à haute température, le diborure de magnésium ( $\text{MgB}_2$ ), a été considérés. La perte de courant alternatif a pu être estimée dans les deux géométries à l'aide de l'algorithme d'adaptation du pas de temps selon une quantité d'intérêt numérique. Les résultats numériques obtenues sont en accord avec ceux calculés en utilisant une discrétisation en temps et constante.

En conclusion, nous avons développé une stratégie d'adaptation en temps qui a pour avantage d'être peu coûteuse en nombre d'opérations et qui est adaptée aux problèmes raides. De plus cette méthode de sélection de pas de temps s'adapte aisément à des quantités d'intérêt. En revanche, le contrôle des variations  $\epsilon$ -voisines devient restrictif pour des ordre  $k \geq 3$ . De plus les hypothèses du théorème 4 du chapitre 3 semblent limiter l'adaptation en temps à une classe particulière de quantités d'intérêt.

Afin de poursuivre la vérification de l'adaptation en temps, en utilisant la méthode des variations  $\epsilon$ -voisines, il serait nécessaire de tester l'algorithme sur une variété plus large de quantités d'intérêt associées à des problèmes différents. De plus, il serait intéressant de montrer rigoureusement que le contrôle des variations d'ordre  $k$  élevé permet effectivement d'obtenir des solutions numériques plus régulières. Enfin, la prochaine étape de ce projet de recherche sur la modélisation des comportements supraconducteurs serait d'implémenter un schéma BDF2 à pas non constants afin d'améliorer la précision de la discrétisation en temps.

## RÉFÉRENCES

- [1] N. Ameniya, K. Miyamoto, N. Banno, and O. Tsukamoto. Numerical analysis of ac losses in high-t superconductors based on e-j characteristics represented with n-value. *Applied Superconductivity, IEEE Transactions on*, 1997.
- [2] I. Babuska and W. C. Rheinboldt. A-posteriori error estimates for the finite element method. *International Journal for Methods in Engineering*, 1978.
- [3] S. Balac and A. Fernandez. Comparison of adaptive step-size control strategies for solving the generalised non-linear schrodinger equation in optics by the interaction picture method. *Hal-archive*, 2012.
- [4] A. Ballarino. Development of superconducting links for the large hadron collider machine. *Superconductor Science and Technology*, 2014.
- [5] A. Bossavit. Numerical modelling of superconductors in three dimensions : a model and a finite element method. *Magnetics, IEEE Transactions on*, 1994.
- [6] A. Bossavit. *Computational Electromagnetism : Variational Formulations, Complementarity, Edge Elements*. Academic Press, 1998.
- [7] R. Brambilla, F. Grilli, and L. Martini. Development of an edge-element model for ac loss computation of high-temperature superconductors. *Superconductor Science and Technology*, 2007.
- [8] R. Brambilla, F. Grilli, L. Martini, and F. Sirois. Integral equations for the current density in thin conductors and their solution by the finite-element method. *Superconductor Science and Technology*, 2008.
- [9] E. H. Brandt. Universality of flux creep in superconductors with arbitrary shape and current-voltage law. *Physical Review Letters*, 1996.
- [10] K. E. Brenan, Campbell S. L., and Petzold L. R. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. Siam, 1996.
- [11] S. Brenner and L. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, 2007.
- [12] Y. Cao and L. Petzold. A posteriori error estimation and global error control for ordinary differential equations by the adjoint method. *SIAM*, 2004.
- [13] J. H. Chaudhry, J.B. Collins, and J. N. Shadid. A posteriori error estimation for multi-stage runge-kutta imex schemes. *arXiv*, 2015.

- [14] J.B. Collins, D. Estep, and S. Tavener. A posteriori error estimation for finite element methods with projection operators as applied to explicit time integration techniques. *BIT*, 2014.
- [15] J.B. Collins, D. Estep, and S. Tavener. A posteriori error estimation for the lax-wendroff finite difference scheme. *Journal of Computational and Applied Mathematics*, 2014.
- [16] A. D. Di Pietro and A. Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*. Springer, 2012.
- [17] S. Eckert, H. Baaser, D. Gross, and Scherf O. A bdf2 integration method with step size control for elasto-plasticity. *Computational Mechanics*, 2004.
- [18] K. Eriksson and C. Johnson. Error estimates and automatic time step control for non-linear parabolic problems. *Society For Industrial And Applied Mathematics*, 1987.
- [19] J. Feldman. Avariable step size methods, 1999.
- [20] A. Fortin. *Analyse Numérique pour Ingénieurs*. Presses Internationales Polytechnique, 2012.
- [21] D. Givoli and I. Henigsberg. A simple time-step control scheme. *Communications in Numerical Methods in Engineering*, 1993.
- [22] O. Golberg. Adaptative stepsize numerical methods for solving ordinary differential equations, May 2007.
- [23] P. Gresho. *Time Integration*. Siam, 2000.
- [24] P. M. Gresho, F. Griffiths, and Silvester D. J. Adaptive time-stepping for incompressible flow part i : Scalar advection-diffusion. *Siam Journal of Scientific Computing*, 2008.
- [25] F. Grilli, E. Pardo, A. Stenwall, D. N. Nguyen, W. Yuan, and F. Gomory. Computation of losses in hts under the action of varying magnetic fields and currents. *IEEE Transactions on Applied Superconductivity*, 2014.
- [26] F. Grilli, F. Sirois, M. Laforest, and Ashworth S. P. Periodic space-time formulation for numerical ac loss computation in superconductors. *IEEE Transactions on Applied Superconductivity*, 2009.
- [27] F. Grilli, S. Stavrev, Y. Le Floch, M. Costa-Bouzo, E. Vinot, I. Klutsch, G. Meunier, P. Tixador, and B. Dutoit. Finite element modeling of superconductors : from 2d to 3d. *Applied Superconductivity, IEEE Transactions on*, 2005.
- [28] K. Gustafsson. Control theoretic techniques for stepsize selection in runge-kutta methods. *ACM Transactions on Mathematical Software*, 1991.
- [29] K. Gustafsson and Söderlind G. Control strategies for the iterative solution of nonlinear equations in ode solvers. *Siam Journal of Scientific Computing*, 2004.



- [30] K. Gustafsson, M. Lundh, and Söderlind G. A pi stepsize control for the numerical solution of ordinary differential equations. *Siam Journal of Scientific Computing*, 1988.
- [31] E. Hairer, S. P. Norsett, and G. Wanner. *Solving Ordinary Differential Equation I, Nonstiff Problems*. Springer, 2008.
- [32] E. Hairer and G. Wanner. *Solving Ordinary Differential Equation II, Stiff and Differential-Algebraic Problems*. Springer, 1980.
- [33] Y. Hascicek, Y. Akin, T. Baldwin, M. Rindfleisch, J. Yue, M. Sumption, and M. Tomsic. A mgb2 12.5kva superconductor transformer. *Superconductor Science and Technology*, 2009.
- [34] A. M. Heidt. Efficient adaptive step size method for the simulation of supercontinuum generation in optical fibers. *Journal of Lightwave Technology*, 2009.
- [35] J. S. Hesthaven and T. Warburton. *Nodal Discontinuous Galerkin Methods*. Springer, 2008.
- [36] Z. Hong, A. M. Campbell, and T. A. Coombs. Numerical solution of critical state in superconductivity by finite element software. *Superconductor Science and Technology*, 2006.
- [37] G. M. Hulbert and I. Jang. Automatic time step control algorithms for structural dynamics. *Computer methods in applied mechanics and engineering*, 1994.
- [38] A. Jannelli and R. Fazio. Adaptive stiff solvers at low accuracy and complexity. *Journal of Computational and Applied Mathematics*, 2005.
- [39] C. Johnson. Error estimates and adaptive time-step control for a classe of one-step methods for stiff ordinary differential equations. *Siam*, 1988.
- [40] C. Johnson. *Numerical Solution of Partial Differential Equations by the finite Element Method*. Cambridge University Press, 1994.
- [41] W. P. Johnson. *The Curious History of Faà di Bruno’s Formula*. American Mathematical Monthly, 2002.
- [42] D. Kay, P. M. Gresho, F. Griffiths, and Silvester D. J. Adaptive time-stepping for incompressible flow part i : Navier-stokes equations. *Siam Journal of Scientific Computing*, 2008.
- [43] K. Kergrene, S. Prudhomme, L. Chamoin, and Laforest M. A new goal-oriented formulation of the finite element method. *Siam Journal of Scientific Computing*, 2008.
- [44] Y. Law and M. Laforest. A nonlinear relaxation formulation of the p-curl problem modelling high-temperature superconductors : A modified yee’s scheme. *Journal of Computational Physics*, 2018.

- [45] Y. Law-Kam-Cio. *Méthode de Galerkin Discontinue pour la Discrétisation en Éléments Finis des Équations de Maxwell pour la Modélisation de Problèmes d'Électromagnétisme en Basses Fréquences*. PhD thesis, École Polytechnique de Montréal, 2018.
- [46] G. P. Lousberg, M. Ausloos, C. Geuzaine, P. Dular, P. Vanderbemden, and B Vanderheyden. Numerical simulation of the magnetization of high-temperature superconductors : a 3d finite element method using a single time-step iteration. *Superconductor Science and Technology*, 2009.
- [47] Mathworks. ode15s. <https://www.mathworks.com/help/matlab/ref/ode15s.html>. Accessed : 2018-07-08.
- [48] Mathworks. Stiff differential equations. <https://www.mathworks.com/company/newsletters/articles/stiff-differential-equations.html>. Accessed : 2018-07-08.
- [49] M. Möller. Time stepping methods, 2015.
- [50] N. Nibbio, S. Stavrev, and B. Dutoit. Finite element method simulation of ac loss in hts tapes with b-dependent e-j power law. *Applied Superconductivity, IEEE Transactions on*, 2001.
- [51] M. P. Oomen, J. Rieger, V. Hussennether, and M. A. Leghissa. Ac loss in high-temperature superconducting conductors, cables and windings for power devices. *Superconductor Science and Technology*, 2004.
- [52] R. Pecher, M. D. McCulloch, S. J. Chapman, L. Prigozhin, and C. M. Elliott. D-modelling of bulk type-ii superconductors using unconstrained h-formulation. *EUCAS 2003 : 6th European Conference on Applied Superconductivity*, 2003.
- [53] J. Peterson. Numerical methods for differential equations, August 2017.
- [54] L. Prigozhin. Analysis of critical-state problems in type-ii superconductivity. *Applied Superconductivity, IEEE Transactions on*, 1997.
- [55] H. H. Robertson. *The Solution of a Set of Reaction Rate Equations*. Academic Press, 1966.
- [56] S. Roman. *The Formula of Faa di Bruno*. New York : Willey, 1958.
- [57] S. Safran, J. Souc, F. Gömöry, P. Kovac, and A. Gencer. Experimentally determined magnetization ac losses of mono and multifilamentary mgb2 wires. *Journal of Superconductivity and Novel Magnetism*, 2013.
- [58] D. Sarmany, M.A. Botchev, and J. J. W. Van Der Vegt. Time-integration methods for finite element discretisation of second-order maxwell equation. *Computers and Mathematics with Applications*, 1993.

- [59] L. F. Shampine and A. Witt. A simple step size selection algorithm for ode codes. *Journal of Computational and Applied Mathematics*, 1993.
- [60] F. Sirois and F. Grilli. Numerical consideration about using finite-element methods to compute ac losses in hts. *Applied Superconductivity, IEEE Transactions on*, 2008.
- [61] F. Sirois and F. Grilli. Potential and limits of numerical modelling for supporting the development of hts devices. *Superconductor Science and Technology*, 2015.
- [62] S. Stavrev, Y. Yang, and B. Dutoit. Modelling and ac losses of bscco conductors with anisotropic and position dependent  $j_c$ . *Physica C*, 2002b.
- [63] A. Stenvall and T. Tarhasaari. Programming finite element method based hysteresis loss computation software using non-linear superconductor resistivity an t-i formulation. *Superconductor Science and Technology*, 2010b.
- [64] G. Söderlind. Automatic control and adaptive time-stepping. *Springer*, 2001.
- [65] A. Wan. *Adaptive Space-Time Finite Element Method in High Temperature Superconductivity*. PhD thesis, École Polytechnique de Montréal, 2014.
- [66] A. Wan and M. Laforest. A posteriori error estimation for the p-curl problem. *Society For Industrial And Applied Mathematics*, 2016.
- [67] A. T. S. Wan, Bihlo A., and Nave J.C. Conservative methods for dynamical systems. *SIAM*, 2016.
- [68] C. William Gear. *Numerical Initial Value Problems in Ordinary Differential Equations*. Prentice-Hall, 1971.
- [69] S. L. Wipf. Ac losses in superconductors. *Journal of Applied Physics*, 1968.