

UNIVERSITÉ D'EVRY-VAL-D'ESSONNE

THÈSE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ D'EVRY-VAL-D'ESSONNE

Spécialité : sciences de l'ingénieur

préparée au Laboratoire de Mécanique et d'Énergétique d'Evry
dans le cadre de *l'École Doctorale SITEVRY*

présentée et soutenue publiquement par

Pierre-Olivier LAFFAY

le jeudi 4 décembre 2008

Sous-structuration de systèmes thermiques par modes de branche

JURY

M. Frédéric BOURQUIN	Ingénieur en chef, LCPC-GER	Président
M. Jean-Luc BATTAGLIA	Professeur, Université Bordeaux I	Rapporteur
M. Daniel PETIT	Professeur, ENSMA	Rapporteur
M. Zhi-Qiang FENG	Professeur, Université D'Evry Val-d'Essonne	Examineur
M. Alain NEVEU	Professeur, Université D'Evry Val-d'Essonne	Directeur de thèse
M. Olivier QUÉMÉNER	Maître de conférences, Université D'Evry Val-d'Essonne	Co-directeur de thèse

Remerciements

Mes remerciements s'adressent en premier lieu à mes directeurs de thèse, Alain Neveu et Olivier Quéméner : merci pour tout ce que vous m'avez appris, aussi bien dans le domaine de la recherche que dans celui de l'enseignement.

Mes remerciements s'adressent également à Jean-Luc Battaglia et à Daniel Petit qui ont accepté de rédiger un rapport détaillé sur mon travail.

J'exprime également toute ma gratitude aux autres membres du jury : Frédéric Bourquin qui a accepté de le présider et Zhiang-Qiang Feng pour l'intérêt qu'il a témoigné à ma thèse.

Je tiens à remercier l'ensemble du personnel que j'ai côtoyé à l'IUT de Brétigny et au LMEE et en particulier Frédéric, Etienne, Sylvie, Jean-François, Ayoob, Hakim, Mohamed, ...

Je remercie spécialement, les jeunes retraités qui m'ont fait l'honneur de venir spécialement pour ma soutenance : Raymond Uny et Olivier Daube. Mais aussi mes nouveaux collègues qui se sont déplacés à travers la région parisienne : Catherine Juramy et Pierre Huguet.

Pour finir, je remercie toute ma famille, mes amis et tous ceux que j'ai oublié.

Table des matières

Remerciements	3
Liste des figures	7
Liste des tableaux	9
Introduction	11
I Rappels préliminaires	13
1 Etat de l'art	15
2 Introduction à l'analyse modale	25
II Théorie et pratique du couplage par conduction	33
3 Théorie du couplage par conduction	35
4 Pratique du couplage par conduction en 2D	47
5 Pratique du couplage par conduction en 3D	73
Conclusion	95
Annexes	99
A Théorèmes de Green-Ostrogradsky	101
B Le théorème spectral	105
C Séparation des variables	109
D Analyse modale « classique »	119
E Base de branche	127
F Réduction modale	139
G Expressions matricielles	149
H Les cinq premiers modes du processeur	157
Bibliographie	160

Liste des figures

1.1	Paris sous la pluie	17
1.2	Figure 1 de l'article [39] présentant la frontière Γ_c	23
2.1	Domaine Ω avec sa frontière et sa normale	28
3.1	Domaine Ω constitué de deux sous domaines $\Omega^{(1)}$ et $\Omega^{(2)}$	37
4.1	Géométrie du processeur avec son radiateur (à l'échelle 2)	48
4.2	Loi de contrôle du ventilateur	49
4.3	Résumé des modèles éléments finis utilisés	51
4.4	Schéma général du pas de temps variable pour les modèles détaillés	52
4.5	Hystérésis de la fonction f	54
4.6	Effet du pas de discrétisation δx sur le résultat calculé par différences finies de 4.9	54
4.7	Effet du pas de discrétisation δx sur le résultat calculé par différences finies de 4.10	55
4.8	Evolutions thermiques maximales et minimales dans le processeur et dans le radiateur	55
4.9	Bilan sur les modèles détaillés de l'étude 2D	57
4.10	Erreur de projection pour le processeur	59
4.11	Erreur de projection pour le radiateur	60
4.12	Schéma général du pas de temps variable pour les modèles réduits	63
4.13	Cas test proposé pour le processeur	64
4.14	Cas test proposé pour le radiateur	66
4.15	Effet du cas test avec $\tilde{N} = 20$	67
4.16	Effet du cas test avec $\tilde{N} = 160$	68
4.17	Résultats pour $\tilde{N} = 20, 60, 160$ avec le cas test 1	71
4.18	Tracé de la fonction γ^* pour $\tilde{N} = 20, 60, 160$ avec le cas test 1	71
5.1	Photographie du bloc lors d'une expérience	75
5.2	Synoptique du dispositif expérimental	75
5.3	Résultats de l'analyse modale mono-corps en 3D.	76
5.4	Vue du 143ème mode de branche à l'aide de quatre tranches	77
5.5	Géométrie du bloc	78
5.6	Evolution temporelle de la puissance volumique dissipée dans les cartouches	79
5.7	Evolution de la température maximale et minimale pour le modèle MD_M	81
5.8	Erreur de projection en mono-corps	82
5.9	Evaluation de $\overline{\Delta T}_{saut,max}$ en fonction de la RTC par la formule 5.19	84
5.10	Températures minimale et maximale au cours du temps pour les deux modèles MD_M et MD_{3E}	85
5.11	Erreur de projection pour le bloc	86
5.12	Erreur de projection pour la cartouche de gauche	87
5.13	Erreur de projection pour la cartouche de droite	87
5.14	Températures minimale et maximale au cours du temps pour chacun des trois modèles éléments finis.	91
A.1	Domaine Ω avec sa frontière et sa normale	101
A.2	Domaine Ω constitué de deux sous domaines $\Omega^{(1)}$ et $\Omega^{(2)}$	102
C.1	Plaque de $[0, L_x] \times [0, L_y]$	110
C.2	Tracé des fonctions permettant de déterminer graphiquement les racines de C.32	115

D.1	Domaine Ω avec sa frontière et sa normale	120
E.1	Domaine Ω avec sa frontière et sa normale	128
E.2	Démarche de calcul des modes	137
G.1	Domaine Ω constitué de deux sous domaines $\Omega^{(1)}$ et $\Omega^{(2)}$	151
G.2	Zoom sur l'interface entre deux maillages non-conformes	153
G.3	Projection orthogonale de M sur $[AB]$	154
G.4	Zoom sur la projection du maillage 1 vers 2	155
G.5	Zoom sur la projection du maillage 2 vers 1	155
G.6	Illustration de projection 3D	155
H.1	Premier mode de branche du processeur	157
H.2	Deuxième mode de branche du processeur	158
H.3	Troisième mode de branche du processeur	158
H.4	Quatrième mode de branche du processeur	159
H.5	Cinquième mode de branche du processeur	159

Liste des tableaux

2.1	Comparaison entre la méthode modale de branche et la méthode modale classique . . .	31
3.1	Epaisseur équivalente en fonction de la conductivité du matériaux	37
4.1	Caractéristiques des matériaux et production volumique d'énergie	49
4.2	Paramètres du modèle MD_A	53
4.3	Paramètres du modèle MD_C	56
4.4	Paramètres des cas tests étudiés	66
5.1	Désignation des corps et des surfaces	78
5.2	Caractéristiques des matériaux	79
5.3	Résultats obtenus en analyse modale mono-corps	82
5.4	Comparaison MR_{3E} et MD_{3E}	89
5.5	Comparaison MR_{3E} et MD_M	90
5.6	Comparaison MR_{3S} (sans fonctionnelle de saut de flux) et MD_M	92
5.7	Comparaison MR_{3S} (avec un paramètre de pondération (β) optimal pour la fonctionnelle de saut de flux) et MD_M	92

Introduction

L'évolution des ordinateurs est très rapide, il est désormais possible d'utiliser un ordinateur portable pour effectuer des calculs complexes (mécanique des fluides, thermique, électromagnétisme, ...). Les utilisateurs des codes de calcul désirent utiliser des modèles de plus en plus précis. La prise en compte de géométries 3D complexes et de non-linéarités sur les propriétés physiques des matériaux est de plus en plus classique. Les méthodes graphiques (Binder-Schmidt ...) sont désormais classées dans la catégorie des méthodes historiques. De nouvelles thématiques voient actuellement le jour : analyse de sensibilité, comportement à long terme, ... Ainsi la course entre l'offre et la demande n'est pas prête de s'arrêter, la rapide progression des moyens informatiques n'est pas suffisante : la demande devance l'offre.

Un domaine où le temps de calcul prend une importance capitale est celui du contrôle commande : si le traitement des mesures effectuées sur un processus est suffisamment rapide, alors il est possible dans certains cas de contrôler le processus. En contrôle commande, les modèles sont forcément petits et s'adaptent mal à la demande d'une prise en compte de phénomènes de plus en plus complexes. La recherche de modèles dits réduits, qui malgré un faible nombre de paramètres sont capables d'être sensibles à ces phénomènes complexes, reste d'actualité. De nombreuses approches existent, l'une d'elles, l'approche modale qui permet une prise en compte de la réalité géométrique, se montre particulièrement efficace. Cette approche est basée sur l'utilisation judicieuse d'un nombre réduit de champs particuliers de température dans l'ensemble du domaine étudié, ces champs sont appelés modes. Selon la technique utilisée le modèle modal peut être calculé ou identifié à partir de mesures ou de simulations.

Ce travail de thèse utilise la méthode modale développée au Laboratoire de Mécanique et d'Energétique d'Evry (LMEE) qui utilise un modèle de connaissance détaillé pour le calcul des modes.

Les méthodes modales sont globalement performantes, toutefois il existe certaines limites à ces méthodes. Une des limitations est située au niveau de la détermination des modes, qui nécessite d'effectuer des calculs sur l'intégralité du domaine. Si l'ordre du modèle initial est grand, alors les besoins en mémoire RAM sont rapidement importants, ce qui limite la portée de la méthode. Une des solutions, étudiée dans cette thèse, consiste à découper le domaine physique en plusieurs sous-domaines de calcul. Une base modale est ensuite déterminée pour chaque sous-domaine, ces bases sont couplées lors du calcul global.

Ce manuscrit présente une méthode de couplage pour les modèles modaux étudiés au LMEE. Pour cela il est organisé de la manière suivante :

- Le chapitre 1 donne les éléments nécessaires sur la modélisation et sur les techniques de réduction.
- Le chapitre 2 présente la méthode modale de branche développée au LMEE, dans le cas où il y a un seul domaine.
- Le chapitre 3 donne les éléments théoriques sur la méthode de couplage envisagée.
- Les chapitres 4 et 5 sont destinés à présenter deux applications numériques de la méthode, respectivement en bidimensionnel et en tridimensionnel. De plus la méthode est présentée pour deux conditions particulières au niveau des frontières de couplage : avec une résistance de contact en bidimensionnel et sans résistance de contact en tridimensionnel.

Les détails mathématiques sont présentés en annexe ceci pour des raisons évidentes de concision.

Première partie

Rappels préliminaires

« Donnez-moi cent paramètres et je vous
ferai un éléphant. Donnez m'en un cent
unième et je lui ferai remuer la queue »

Jacques Hadamard

Contenu du chapitre

I.	Modèles : tentative de définition, propriétés, limites et principaux types	16
II.	Notion de représentation d'état	20
III.	Quelques modèles utilisés en thermique	21
IV.	Couplages de modèles réduits	22

Ce travail de thèse porte essentiellement sur l'obtention d'un modèle par représentation d'état d'un système thermique. Les modèles d'état considérés sont des modèles obtenus à partir de modes, les états sont alors les coefficients d'excitation de ces modes. Ces états ne sont donc pas des états thermodynamiques (température, pression), mais des états au sens de l'automatique. C'est ainsi à partir d'un modèle d'état que l'on effectue une réduction, c'est-à-dire une diminution du nombre de paramètres les définissant. Les modèles d'états qui sont étudiés dans cette thèse sont dits réduits ou d'ordre réduit. Bien que ces thématiques soient courantes dans certains domaines (mécanique, automatique, ...), elles sont relativement méconnues en thermique.

Ce chapitre a ainsi pour objectif d'éclaircir ces différents points. Il commencera par présenter de manière générale la notion de modèle, avec une application sur un cas de la vie courante. On définira alors la notion de modèle réduit, puis celle de la représentation d'état. La représentation d'état est une notion très générale, qui peut être appliquée au cas particulier de l'analyse modale. Les modes sont des champs particuliers de l'espace, l'ensemble des modes forment une base qui est capable de décomposer certains champs thermiques¹. Les modes sont temporellement invariants, ce sont les coefficients de la décomposition linéaire du champ de température dans l'espace modal (les états), qui évoluent au cours du temps.

Enfin ce chapitre se termine par un rapide tour d'horizon des modèles utilisés en thermique. Il précise notamment les différentes approches modales existantes (modèles identifiés ou modèles construits). Enfin, en fonction des différents modèles utilisés il passe en revue les différents travaux qui abordent la notion de couplage entre domaines.

Il est à noter que la formulation mathématique de l'analyse modale pour un problème thermique n'est pas abordée ici, mais fait l'objet d'un chapitre ultérieur.

¹On verra en annexe E que tous les champs thermiques peuvent être décomposés par une base de branche

I. Modèles : tentative de définition, propriétés, limites et principaux types

Nous allons débiter en définissant, ou au moins en tentant de le faire, la notion de modèle. Une fois que cela aura été fait, nous détaillerons les propriétés d'un modèle, puis nous donnerons des éléments sur les limites des modèles. Pour finir, concernant cette introduction, nous expliciterons les principaux types de modèles existants.

I.1. Notion de modèle

La définition d'un modèle n'est pas absolue, elle dépend du domaine (mathématiques, physique, sciences de l'ingénieur,...), des personnes, mais aussi de l'objectif recherché.

Le « Larousse » définit le modèle par : *Structure formalisée utilisée pour rendre compte d'un ensemble de phénomènes qui possèdent entre eux certaines relations.* Le même dictionnaire définit la structure formalisée par : *structure exprimée sous la forme d'un langage abstrait composé de symboles et de règles permettant de manipuler ces symboles.*

Minsky dans [33] définit le modèle par : *Pour un observateur O , un objet M est un modèle d'un objet A dans la mesure où O peut utiliser M pour répondre aux questions qui l'intéressent au sujet de A .*

Naslin dans la préface de [44] définit le modèle par : *D'une manière générale, un modèle d'un phénomène ou d'un processus est essentiellement un mode de représentation tel qu'il permette, d'une part de rendre compte de toutes les observations faites et, d'autre part, de prévoir le comportement du système considéré dans des conditions plus variées que celles qui ont donné naissance aux observations. Ainsi, le modèle généralise la validité des résultats expérimentaux et permet d'agir sur le processus dans un sens désiré. Les actions exercées fournissent de nouvelles observations, de sorte que la séquence observation-modélisation constitue en fait une boucle fermée qui assure l'amélioration progressive du modèle.*

Walliser dans [50] définit le modèle par : *Le modèle permet de comprendre un système complexe, difficile à connaître, en effectuant une simplification, une réduction, une sélection de certains aspects.*

Bachelard dans [2] définit le modèle par : *Un modèle est un objet réputé facile à connaître que l'on met en relation avec un objet difficile à connaître.*

On trouve, dans la thèse de Flament [18], la définition suivante de la modélisation : *Le but de toute modélisation est d'obtenir une image ou une représentation d'un phénomène réel.*

Au vu des différentes définitions (on pourrait en donner bien d'autres), il apparaît que la notion de modèle est très large et assez floue, ce qui est parfaitement normal étant donné l'importance et la variabilité des modèles.

I.2. Quelques propriétés d'un modèle

I.2.1. Le modèle : une représentation imparfaite de la réalité

L'exemple choisi est la photographie présentée en figure 1.1. Cette photographie est issue de la page internet suivante : « <http://www.meteo-paris.com/ile-de-france/climat.html> ».

Elle donne une représentation d'une réalité et à ce titre peut donc être qualifiée de modèle. Du premier coup d'œil, on reconnaît une vue d'une ville. Le décor, le style Hausmannien des immeubles laisse penser qu'il s'agit de Paris. Les différents types de véhicules nous permettent de situer approximativement l'époque (début du XX^{ème} siècle). Enfin l'existence de parapluies nous montre qu'il pleut.... Ainsi ce modèle nous apporte un certain nombre d'informations plus ou moins précises de ce qu'est la réalité.



FIG. 1.1 – Paris sous la pluie

Par contre l'image est en noir et blanc, bidimensionnelle, immobile et inodore, ce qui limite la perception de la réalité décrite. Ce modèle peut même fausser notre perception : pleut-il vraiment ? Comme nous le montre cet exemple, il existe un écart (ou biais) entre modèle et réalité modélisée. On parle aussi de dégradation de l'information, qui est une conséquence de la nécessaire simplification liée au processus de représentation.

La réalité représentée par cette photographie définie par le titre « Paris sous la pluie » pourrait faire l'objet de bien d'autres modèles : une description orale, un simple croquis, ou encore un enregistrement audio au coin d'un carrefour parisien sont également des modèles de la même réalité. De façon plus complète, un film, qui correspond à une superposition de différents modèles (défilement de photographies synchronisé à un enregistrement audio) en est un autre, plus précis. Ainsi pour une seule réalité il existe de nombreux modèles qui donneront chacun une vision différente et plus ou moins imparfaite.

Tous ces modèles peuvent être plus ou moins proches dans leur nature (physique ou abstraite) de la réalité modélisée : un prototype servant de modèle pour une réalisation en série sera très proche des nombreux objets ainsi fabriqués. A l'inverse un modèle visant à prédire la consommation énergétique d'un bâtiment ne sera ni un bâtiment, ni une énergie consommée...

Une caractéristique d'un modèle est dite **compatible** si elle est également caractéristique de la réalité modélisée : une maquette à l'échelle 1 :1 donnera en première approximation les dimensions exactes, pour celles qui sont représentées, de l'objet modélisé.

Une caractéristique du modèle est qualifiée de **formelle** si elle n'est pas une caractéristique de la réalité : le caractère bidimensionnel de la photographie 1.1 est une caractéristique formelle puisque la réalité est tridimensionnelle.

Enfin une caractéristique de la réalité est dite **réelle** si elle n'est pas une caractéristique du modèle : on peut ainsi reprendre l'exemple de la réalité tridimensionnelle, caractéristique qui disparaît avec le modèle de la photographie.

I.2.2. Nécessité de la définition de limite du modèle et notion de modèle adapté

La focale de l'appareil utilisé pour la photographie 1.1 a limité le champ d'enregistrement à un seul carrefour de la ville de Paris. Peut-on réellement utiliser ce modèle pour décrire l'ensemble de la cité comme le titre semble nous l'indiquer ? Ne faudrait-il pas plutôt en limiter son utilisation à certains types de quartiers, communément qualifiés de grands boulevards ? Cet exemple montre bien qu'un modèle reste limité à une petite partie finie de la réalité, et que plus ce domaine est étendu, plus le biais peut être important, ou alors le modèle doit être complexe. En sciences de l'ingénieur, les transferts thermiques par conduction sont modélisables par la loi de Fourier, à savoir $\varphi = -k \text{grad}T$. Cette loi n'est qu'un modèle imparfait puisqu'elle sous-entend que la vitesse de la propagation de la chaleur est infinie. Elle reste néanmoins utilisable dans de nombreux cas, excepté pour des situations de nano-thermique où le biais devient alors trop important pour que les résultats soient utilisables. Il importe ainsi de bien définir

les limites de chaque modèle, en deçà desquelles son utilisation reste acceptable.

D'après de qui vient d'être dit, pour une même réalité, il peut exister de nombreux modèles, plus ou moins complexes, et caractérisés par un biais plus ou moins grand. Il importe ainsi de choisir de façon judicieuse le modèle à utiliser, selon l'objectif visé : on parle ainsi de modèle adapté. Ainsi pour dimensionner une poutre en béton armé, un ingénieur en bâtiment préfère utiliser les modèles simples de la mécanique des structures (qui est aussi appelée résistance des matériaux), plutôt que d'utiliser l'ensemble des lois plus complètes de la mécanique des milieux continus associés à une discrétisation numérique de type éléments finis.

Un ingénieur utilisant un modèle adapté admet implicitement une dégradation plus ou moins importante de sa connaissance de la réalité représentée, mais qui n'aura pas d'impact sur l'objectif visé. Il est classique en thermique du bâtiment, lors d'une approche globale, de ne considérer qu'une température dans l'air de chaque pièce. La température étant une information statistique sur l'état de l'agitation moléculaire, qui considère que l'équilibre thermodynamique local est atteint. L'ingénieur a ainsi modélisé la réalité physique de l'agitation moléculaire par ce que l'on appelle la température, qui est définie localement. Comme cela n'était pas suffisant, il a encore tronqué, pour l'approche globale du bâtiment, sa connaissance de la température en ne considérant qu'une température moyenne par pièce. Il considéra alors cette température moyenne comme étant suffisamment représentative de ce qu'il recherche.

Tout l'art de la modélisation consistera à construire un modèle qui permettra d'obtenir un écart acceptable entre réalité modélisée et résultat du modèle. Une contrainte supplémentaire relative aux moyens à mettre en œuvre est généralement imposée par rapport aux objectifs visés : temps humain, temps CPU des ordinateurs, coût et délai des essais sur maquette... On comprend ainsi pourquoi la modélisation occupe une telle place en science.

I.3. Les différents types de modèles en sciences de l'ingénieur

Limitons notre propos aux sciences de l'ingénieur. Nous avons alors 5 grandes classes :

- modèles physiques
- modèles analogiques
- modèles heuristiques
- modèles de connaissance
- modèles de représentation

Le modèle le plus évident est le modèle physique de type maquette. Il s'agit d'un objet physique semblable en grande partie à l'objet à modéliser, dont la taille varie. En général, les échelles sont réduites comme par exemple les maquettes ferroviaires ou d'avion pour les études en soufflerie. Les souffleries automobiles utilisent au contraire le plus souvent des maquettes de taille réelle, qui sont en général des prototypes. L'étude de la cavitation dans les roulements à billes se fait presque toujours dans des maquettes plus grandes que la réalité.

Les modèles analogiques utilisent une analogie souvent d'équation pour permettre de construire un modèle d'un phénomène. Il est possible d'utiliser du papier conducteur d'électricité, connu sous le nom commercial « Teledeltos », pour modéliser des transferts de chaleur par conduction, ou des transferts d'eau dans un sol. Avant le développement actuel de l'ordinateur et du calcul numérique, cette méthode était classiquement utilisée en enseignement et dans les bureaux d'études.

Les modèles heuristiques sont des modèles basés sur la seule observation du phénomène à modéliser. Dans un modèle heuristique, on suppose ne rien connaître sur les causes de ce que l'on observe. La démarche heuristique est adoptée par l'extraction de connaissance connue sous son nom anglais « data-mining ». La démarche du data-mining consiste à récupérer un grand nombre d'informations et à tester au hasard un ensemble assez grand de modèles avec pour objectif d'en trouver un qui représentera correctement les données. Il est aussi éventuellement possible de travailler par interpolation entre les différentes données disponibles. Kepler a établi les lois qui portent son nom en appliquant la démarche heuristique aux données recueillies par Tycho Brahe. Les lois de Kepler ont été établies sur la période 1609-1618 bien avant le principe fondamental de la dynamique de Newton établi en 1687.

A l'opposé des modèles heuristiques, on trouve les modèles de connaissance. Un modèle de connaissance est basé sur la supposition que l'on est capable de modéliser l'ensemble de la physique du phénomène. L'ingénieur construit ses modèles de connaissance à partir des équations de la physique. Par exemple, un problème de conduction de la chaleur fondé sur l'équation de la chaleur et traité selon la méthode des éléments finis est un modèle de connaissance.

Les modèles de représentation appelés aussi modèles de comportement ou modèles identifiés sont des intermédiaires entre les modèles heuristiques et les modèles de connaissance. Dans un modèle de représentation, on suppose par connaissance la forme du modèle, on expérimente (physiquement ou numériquement) pour déterminer des données, et on cale le modèle de manière à représenter au mieux les données. Ainsi, contrairement au modèle de connaissance, nous avons besoin d'une information sur le comportement du système physique. Et comparativement au modèle heuristique, nous ne testons pas au hasard, mais émettons une hypothèse sur la forme du résultat escompté. Un exemple classique est la loi de Newton, $\Phi = h(T)\Delta T$, on suppose que le flux perdu par convection est proportionnel à la différence de température, le modèle de connaissance sous-jacent est figé, puis on identifie la loi donnant le coefficient de convection h en fonction de la température.

I.4. Notion de réduction de modèle

L'expression « réduction de modèle » peut être considérée comme superflue, puisque qu'un modèle est précisément une représentation réduite et imparfaite de la réalité. Cependant on a également vu qu'il peut exister de nombreux modèles pour une même réalité, certains étant plus complexes et lourds à utiliser que d'autres. Dans un objectif d'utilisation d'un modèle adapté, on comprend l'intérêt que peut apporter la démarche qui vise à diminuer de façon importante le nombre de paramètres définissant le modèle, tout en maintenant acceptable l'écart entre ce modèle et la réalité. C'est ce que recouvre le terme « modèle réduit ».

En sciences de l'ingénieur, différents types de réductions existent : lors d'une démarche expérimentale, il va s'agir de diminuer les points de mesure. Dans le cas d'une simulation numérique, on cherche alors à diminuer le coût (mémoire RAM et/ou temps CPU) du calcul.

Dans ce domaine de la simulation numérique qui correspond au présent travail de thèse, la notion de réduction de modèle est formalisée par l'ordre du modèle : on parle alors de modèle d'ordre réduit. De même les anglo-saxons utilisent l'abréviation ROM pour « Reduced Order Model ».

La notion d'ordre ou de dimension est une notion qui s'est particulièrement développée avec la croissance de l'outil informatique. En effet les ordinateurs bien que doués d'une puissance de calcul très élevée, ne sont pas capables de rendre compte d'une réalité continue. La mémoire des ordinateurs est discrète et se chiffre en multiples d'octets : kilo-octets, méga-octets, giga-octets, téra-octets. Cette mémoire va permettre de quantifier un problème physique par l'intermédiaire d'un nombre entier N de variables. L'ordre d'un modèle à N variables sera égal à N . Jusqu'ici, tout est simple, cette notion d'ordre est classiquement introduite en différences finies ou en éléments finis. Dans chacune de ces méthodes, on utilise une discrétisation du problème qui comporte N_{pts} points où l'on définit N_{dof} degrés de liberté (Degree Of Freedom en anglais). En conduction thermique dans un solide immobile, le nombre de degrés de liberté par nœud est égal à 1, il n'y a que la température. En mécanique des fluides isotherme et incompressible, ce nombre monte à 4 : 1 degré pour la pression, 3 pour la vitesse (suivant les 3 axes). L'ordre N d'un tel modèle est égal au produit $N_{pts} * N_{dof}$.

Les approches différences finies ou éléments finis font donc apparaître une notion d'ordre, cette notion est limitante pour deux raisons. La première raison étant le coût mémoire, ce coût est proportionnel à l'ordre N choisi et est proportionnel à la précision arithmétique requise (int, float, double, ...). L'ordre qui peut être simulé sur un ordinateur est limité par la quantité de mémoire présente sur cet ordinateur. La seconde raison est le temps de calcul : si la résolution de problème statique est relativement rapide même pour un ordre du modèle élevé, la résolution de problèmes dynamique peut rendre nécessaire la recherche de 10^6 (ou plus) pas de temps sur l'horizon temporel désiré.

L'ingénieur (de même pour le chercheur) désire utiliser des modèles de plus en plus précis et passer de moins en moins de temps à attendre les résultats d'une simulation. C'est dans ce cadre qu'intervient

la réduction de modèle.

II. Notion de représentation d'état

La représentation d'état est très courante en sciences de l'ingénieur, il est cependant classique de ne pas savoir que l'on utilise une représentation d'état. L'équation de la chaleur² sous forme matricielle (voir équation G.11) est un modèle d'état :

$$\mathbb{C}_{cap} \dot{\mathbf{T}} = -\mathbb{K}\mathbf{T} - \mathbb{B}\mathbf{T} + \mathbf{A} \quad (1.1)$$

La notion de variable d'état est classique en thermodynamique et en automatique. Les automaticiens ont formalisé la représentation d'état :

- Les variables d'état d'un système permettent de caractériser son état, elles peuvent avoir un sens physique direct (pression, température, vitesse, ...) mais cela n'est pas nécessaire. En général un nombre fini de variables d'état est retenu, leur nombre N est appelé ordre du système.
- La représentation d'état la plus générale est la suivante :

$$\begin{cases} \dot{\mathbf{X}} = \mathbb{A}\mathbf{X} + \mathbb{B}\mathbf{U} \\ \mathbf{Y} = \mathbb{C}\mathbf{X} + \mathbb{D}\mathbf{U} \end{cases} \quad (1.2)$$

Le vecteur \mathbf{X} de dimension N est le vecteur d'état. Le vecteur \mathbf{Y} de dimension Q est le vecteur des observables aussi appelés sorties ou mesures. Le vecteur \mathbf{U} de dimension P est le vecteur des sollicitations. La matrice \mathbb{A} comporte N lignes et N colonnes, elle caractérise la dynamique du système. La matrice de commande \mathbb{B} comporte N lignes et Q colonnes. La matrice d'observation \mathbb{C} comporte Q lignes et N colonnes. La matrice d'action directe \mathbb{D} comporte Q lignes et P colonnes, elle est très souvent nulle.

La représentation d'état n'est pas unique. Soit \mathbb{P} une matrice comportant N lignes et N colonnes inversible et indépendante du temps, il est possible d'écrire un modèle d'état pour la variable $\mathbf{Z} = \mathbb{P}\mathbf{X}$. La première équation du système 1.2 est multipliée à gauche par \mathbb{P} , soit :

$$\begin{cases} \dot{\mathbf{Z}} = \mathbb{P}\mathbb{A}\mathbf{X} + \mathbb{P}\mathbb{B}\mathbf{U} \\ \mathbf{Y} = \mathbb{C}\mathbf{X} + \mathbb{D}\mathbf{U} \end{cases} \quad (1.3)$$

Le vecteur \mathbf{X} est remplacé par $\mathbb{P}^{-1}\mathbf{Z}$, ce qui donne :

$$\begin{cases} \dot{\mathbf{Z}} = \mathbb{P}\mathbb{A}\mathbb{P}^{-1}\mathbf{Z} + \mathbb{P}\mathbb{B}\mathbf{U} \\ \mathbf{Y} = \mathbb{C}\mathbb{P}^{-1}\mathbf{Z} + \mathbb{D}\mathbf{U} \end{cases} \quad (1.4)$$

Dans le cas qui nous concerne, il existera deux vecteurs d'état :

- Le vecteur usuel des températures : \mathbf{T}
- Le vecteur dit d'état (modal) \mathbf{X} tel que : $\mathbf{T} = \mathbb{V}\mathbf{X}$

Remarque : les équations du système 1.2 sont du premier ordre, ce qui ne signifie pas qu'il soit impossible d'obtenir une représentation d'état pour un système d'ordre supérieur à 1. En effet, la plupart des systèmes différentiels d'ordre supérieur à 1 peuvent être transformés en un système différentiel d'ordre 1 suivant la procédure classique utilisée dans l'exemple de la page suivante.

²L'indice *cap* est utilisé pour distinguer la matrice de capacité thermique habituellement notée \mathbb{C} de la matrice d'observation (équation 1.2)

Soit une bille de masse m en chute libre, sa position verticale est notée z , le mouvement est régi par l'équation suivante :

$$\ddot{z} = -g \quad (1.5)$$

Il est possible de définir le vecteur d'état par :

$$X = \begin{pmatrix} z \\ v \end{pmatrix} \quad (1.6)$$

Avec $v = \dot{z}$ la vitesse verticale. Dans ces conditions, il est possible d'écrire 1.5 de la manière suivante :

$$\dot{\mathbf{X}} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \mathbf{X} + \begin{pmatrix} 0 \\ -g \end{pmatrix} \quad (1.7)$$

Ce qui correspond à la forme de la première équation du système 1.2.

III. Quelques modèles utilisés en thermique

L'équation de la chaleur est un modèle de connaissance capable de décrire systématiquement (ou presque : nano-thermique, ...) les transferts thermiques dans les solides.

Il est classique de compléter un cours de thermique en donnant des éléments sur les modèles les plus classiques de la thermique, soit les modèles basés sur une simplification liée à une ou des particularités du problème :

- Le modèle du mur en régime statique
- Le modèle de l'ailette mince
- Le modèle du corps à température uniforme (bille dans un four, thermocouple, ...)
- Le modèle du massif semi-infini soumis à : un échelon, une rampe, une sollicitation périodique
- ...

Ces modèles permettent d'introduire les notions suivantes :

- Résistance thermique
- Temps caractéristique de diffusion
- Profondeur de pénétration
- Atténuation et déphasage du « signal » thermique
- ...

Les thermiciens (ingénieurs et chercheurs) disposent actuellement de modèles basés sur :

- Une équivalence électro-thermique : méthode nodale
- Une discrétisation de l'équation de la chaleur : différences ou éléments finis
- Un modèle modal identifié (les modes sont inconnus) : MIM³ [40]
- Un modèle modal basé sur des modes identifiés : POD⁴ [14]
- Un modèle modal basé sur des modes calculés : BERM⁵ [43],...
- ...

Les modèles nodaux sont basés sur une équivalence électro-thermique. La capacité thermique est équivalente à une capacité électrique et la conduction thermique est modélisée par des résistances électriques. L'utilisation d'un modèle nodal est très simple et il existe des logiciels commerciaux adaptés aux modèles nodaux (ESACAP, ...). A l'inverse les modèles nodaux sont délicats à établir, une partie de leur développement étant souvent effectuée avec l'aide de l'intuition.

³MIM : Modal Identification Method ou Méthode d'Identification Modale

⁴POD : Proper Orthogonal Decomposition

⁵BERM : Branch Eigenmodes Reduction Method

Les méthodes de discrétisation spatiale (éléments finis, différences finies, volumes finis, ...) sont très largement répandues. De nombreux outils logiciels d'application systématique existent et leur utilisation est très courante.

Il existe principalement trois méthodes modales, toutes utilisent l'hypothèse qu'il existe une base modale telle que $\mathbf{T} = \mathbb{V}\mathbf{X}$:

- La méthode MIM identifie directement un modèle modal en utilisant ses propriétés mais sans déterminer les modes.
- La méthode POD utilise des données (expérimentales ou numériques) spatialement et temporellement bien définies pour déterminer des modes. Un modèle modal est construit à partir de ces modes.
- La méthode BERM définit un problème aux valeurs propres dont les solutions calculées numériquement déterminent une base modale. Un modèle modal est construit à partir des modes de cette base.

IV. Couplages de modèles réduits

Le couplage de modèles détaillés est présenté dans de très nombreux articles, de nombreuses méthodes sont actuellement étudiées. La situation est différente pour le couplage de modèles réduits, peu d'articles ont été écrits sur cette thématique :

- Le couplage de modèles modaux de type MIM est présenté dans [39]
- Le couplage de modèles modaux utilisant des modes classiques (voir annexe D) est présenté dans [16] et dans [32]

Il est par ailleurs possible de rajouter les idées développées pour la synthèse modale en thermique par Flament [18], idées qui seront utilisées dans cette thèse.

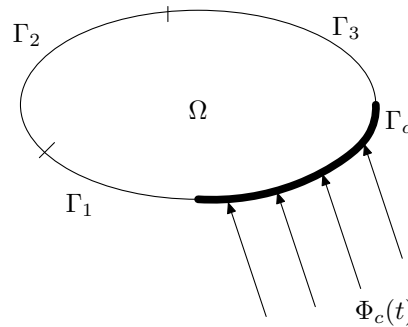
IV.1. Couplage de modèles réduits de type MIM

L'article [39] présente une méthodologie relativement générale pour le couplage d'un modèle modal de type MIM avec un autre modèle. Le système thermique envisagé est diffusif sans transport, il est représenté en figure 1.2. La frontière est séparée en quatre parties qui correspondent aux différentes conditions aux limites :

- Une condition aux limites de type Dirichlet est imposée sur Γ_1 .
- Une condition aux limites de type Neumann est imposée sur Γ_2
- Une condition aux limites de type Fourier ou Robin est imposée sur Γ_3
- Enfin la frontière dédiée au couplage est la frontière Γ_c .

La frontière de couplage Γ_c est discrétisée à l'aide de n_c points. Le modèle MIM est recherché en imposant des flux sur ces n_c points. Le modèle MIM est construit indépendamment de la condition qui sera appliquée sur cette interface. Plusieurs conditions aux limites peuvent être imposées sur la frontière Γ_c : couplage avec un modèle détaillé ou réduit de type diffusif, convection avec un coefficient de convection variable, rayonnement ...

Le couplage par conduction présenté dans [39] s'effectue à l'aide du flux traversant l'interface de contact.

FIG. 1.2 – Figure 1 de l'article [39] présentant la frontière Γ_c

IV.2. Couplage de modèles réduits construits à l'aide de modes classiques

Le couplage de modèles modaux utilisant des modes classiques est présenté dans [16]. La méthode est algébrique, et imposerait la continuité du flux et de la température s'il n'y avait pas de réduction modale. L'exemple traité est une partie (morceau de dalle et de mur) d'un bâtiment qui échange de l'énergie avec le sol. Le maillage global comporte 1024 éléments, les deux sous-domaines comportent 559 et 465 éléments, enfin il y a 38 éléments sur l'interface de contact. Le nombre de modes conservés par sous-domaine est relativement grand (43 et 50), ce qui permet une bonne conservation du flux. Cette approche est applicable lorsque les modes utilisés sont classiques car l'équation d'état est alors découplée (voir l'équation D.35). Son intérêt serait faible dans le cas des modes de branche car l'équation d'état est alors couplée (voir 2.16).

Le couplage de modèles modaux utilisant des modes classiques est aussi présenté dans [32]. La notion de couplage y est très large, l'exemple traité est une cellule cubique monozone. Cette cellule est représentative d'un petit bâtiment, les parois de la cellule échangent de l'énergie par convection et par rayonnement. Le modèle inclut du transport pour tenir compte du renouvellement d'air. Le couplage est alors un ensemble intégrant la convection, le rayonnement et le transport. Les travaux présentés dans [32] sont donc très éloignés de ceux qui seront présentés dans la suite de ce document.

IV.3. La synthèse modale (des modes classiques)

La synthèse modale est une technique de calcul des modes d'un domaine Ω qui utilise les modes des sous-domaines $\Omega^{(a)}$. Les méthodes de synthèse modale sont classiques en mécanique [24]. La synthèse modale a été développée en thermique dans le cas des modes classiques par Flament [18].

Un biais est admis lorsque l'on effectue une réduction modale, or les bases modales de chaque sous-domaine sont partielles lors d'une synthèse modale. Le biais est responsable de l'introduction de sauts de température et de flux au niveau des interfaces entre sous-domaines. Flament introduit deux fonctionnelles pour contrôler le saut de flux et le saut de température.

Les travaux de Flament ont directement influencé la démarche de couplage étudiée dans cette thèse. Cependant cette thèse est dédiée au couplage de modèles modaux de branche, alors que la thèse de Flament porte sur la synthèse des modes classiques.

Introduction à l'analyse modale

« Un brin de sagesse : tous les modèles sont
faux, certains sont utiles »

George Box

Contenu du chapitre

I.	Brève introduction à l'analyse modale	26
II.	Formulation variationnelle de l'équation de la chaleur	27
III.	Formulation modale	29
IV.	Réduction modale	30
V.	La méthode modale de branche confrontée à la méthode modale classique	31
VI.	Bilan des travaux numériques déjà réalisés sur la méthode modale de branche	32

Le présent chapitre aborde dans le détail l'outil fondamental qui a servi de base aux travaux réalisés dans cette thèse. Il s'agit donc de présenter la démarche utilisée pour créer un modèle modal d'ordre réduit, modèle basé sur les modes de branche réduits.

Une rapide introduction permet de faire un bref historique sur l'analyse modale en thermique. Celle-ci débute par la méthode de séparation des variables initiée par Fourier. La méthode modale classique est une extension aux géométries quelconques de la méthode de séparation des variables. La limite de cette technique est que le système thermique traité doit être linéaire et à paramètres stationnaires. En effet, les bases modales classiques sont par construction uniquement adaptées des conditions aux limites figées.

L'utilisation d'une nouvelle base modale à rendu possible le traitement des non-linéarités. Ce traitement a un prix : les dynamiques des modes deviennent couplées et ce même dans le cas d'un système linéaire. La nouvelle base modale que l'on nomme base de branche utilise une condition aux limites abstraite dite de Steklov. Cette condition aux limites dépend de la valeur propre du mode, elle n'est donc pas figée contrairement aux conditions aux limites utilisées pour construire les modes classiques.

Les outils mathématiques permettent de démontrer que la base de branche est une base pour tout système thermique linéaire ou non. Les expérimentations numériques montrent que la base de branche est réductible, il est ainsi possible de l'utiliser pour construire des modèles modaux réduits adaptés aux systèmes thermiques non-linéaires.

La base modale est déterminée numériquement à l'aide d'une discrétisation spatiale. Les méthodes classiques (différences finies, éléments finis, volumes finis) sont utilisables pour le calcul des modes. La méthode des éléments finis est utilisée pour cette thèse, elle est adaptée au traitement de géométries

complexes. Cette méthode est basée sur l'utilisation d'une formulation variationnelle, qui est compatible avec l'utilisation de modes. L'équation d'état est ainsi établie à l'aide de la formulation variationnelle.

Le passage de la base modale complète à d'une base modale réduite dans la formulation variationnelle est très simple. La méthode de réduction utilisée est présentée en annexe F.

Ce chapitre se termine par un tableau comparatif entre la méthode modale de branche et la méthode modale classique, puis par un bilan sur les travaux de recherche déjà effectués dans ce domaine.

I. Brève introduction à l'analyse modale

La méthode de séparation des variables (x,y,z,t) est connue depuis les travaux de Fourier. L'analyse modale classique est une généralisation de cette technique. De manière récente, une méthode d'analyse modale utilisant les modes de branche [36] à été proposée. Ces modes sont originaires des travaux des mécaniciens [20].

Quelques éléments vont être donnés sur la méthode de séparation des variables. L'analyse modale classique sera alors brièvement présentée. Le lecteur sera alors familiarisé avec certaines notions élémentaires d'analyse modale, ce qui permettra par la suite de mettre en valeur les points essentiels de l'analyse modale de branche.

I.1. Les premiers pas : la méthode de séparation des variables

La méthode de séparation des variables s'applique à un système thermique linéaire dont le domaine géométrique est simple. Il faut alors :

- Rechercher analytiquement une base de fonctions propres. Cette base étant en général déterminée par les racines d'une équation transcendante.
- Déterminer graphiquement les racines de l'équation transcendante.
- Expliciter des conditions d'orthogonalité et/ou un produit scalaire ad hoc.
- Déterminer les coefficients de la décomposition dans la base des fonctions propres

Cette méthode a été utilisée pendant une très longue période, de très nombreuses solutions ont été calculées¹. L'annexe C présente un exemple d'utilisation de la méthode de séparation des variables.

Cette démarche est limitée puisqu'il est (presque) impossible de traiter des géométries courantes. Cependant, la méthode de séparation des variables est applicable à la main sans le recours de l'ordinateur. Elle a été très utilisée à partir de sa découverte par Fourier et jusqu'à l'apparition de l'outil informatique.

I.2. L'analyse modale classique

L'outil informatique a profondément modifié les démarches de conception des ingénieurs. La méthode des différences finies a été rapidement utilisée dans des codes de calculs informatiques. L'idée d'utiliser une méthode de fonctions propres dans un code informatique est plus récente et a été (re)-découverte par de nombreuses personnes [3, 11, 46] dans la période 1975-1985². L'outil informatique permet alors de généraliser la méthode de séparation des variables.

Le domaine peut désormais être quelconque, les fonctions propres étant calculées numériquement. Les conditions d'orthogonalité sont systématiques.

Par contre le système thermique doit être linéaire et invariant. L'introduction de non-réciprocité est possible mais reste délicate [17].

¹Le lecteur intéressé pourra consulter [10] pour se donner une idée de l'ampleur des travaux qui ont été menés sur la méthode de séparation des variables.

²Dans une période plus récente, il est possible de citer [19].

L'annexe D présente de manière plus détaillée les notions fondamentales concernant l'analyse modale classique :

- Il faut séparer les régimes glissants et dynamiques.
- Le régime dynamique se décompose dans l'espace des fonctions propres.
- Les fonctions propres sont calculées numériquement.
- Le système thermique étant linéaire, le principe de superposition permet de montrer que les dynamiques modales sont découplées. Les équations d'état sont alors des équations différentielles ordinaires indépendantes les unes des autres.
- Il existe des dominances dans l'espace des modes qui permettent de réduire le système thermique.

La démarche est ainsi toujours limitée aux systèmes thermiques linéaires, mais elle informatisée et systématique. Son utilisation est alors moins laborieuse que la méthode de séparation des variables. Elle peut être très performante, par exemple le gain de temps est de 24000 par rapport à la méthode des éléments finis dans l'article [42].

I.3. L'idée des modes de branche

L'utilisation des modes de branche permet de s'affranchir des hypothèses de linéarité et d'invariance du système thermique. Il est alors possible d'étendre le domaine d'application de la méthode modale. La méthode modale de branche présente plusieurs particularités :

- La base de branche est capable de décomposer tous les champs, ainsi il n'y a plus de séparation entre le régime dynamique et le régime glissant.
- La prise en compte de non-linéarités supprime la possibilité d'utiliser le principe de superposition.
- Il existe toujours des dominances dans la base modale, ce qui permet d'envisager sa réduction.
- Les équations d'état sont couplées, même pour un système thermique linéaire.

II. Formulation variationnelle de l'équation de la chaleur

Les méthodes analytiques ne peuvent s'appliquer raisonnablement que dans le cas de géométries simples. Dans le cas de géométries quelconques, il faut utiliser une méthode numérique. Les méthodes numériques les plus courantes utilisent une discrétisation du domaine, c'est le cas des méthodes suivantes : différences finies, éléments finis, volumes finis. Chacune de ces méthodes est utilisable pour résoudre l'équation de la chaleur et pour le calcul des modes³.

La méthode des éléments finis est arbitrairement choisie comme méthode de référence pour la recherche d'une solution à l'équation de la chaleur. Elle sert aussi de support de discrétisation pour le calcul des modes de branche.

La méthode des éléments finis est basée sur l'utilisation d'une formulation variationnelle des équations du problème de base.

³Les différences finies sont utilisées dans [38]. Les éléments finis sont utilisés dans [43]. Les volumes finis sont utilisés dans [30].

II.1. Equation de la chaleur

Le problème thermique est un problème de diffusion thermique sans transport⁴ défini dans un domaine Ω . Les paramètres thermophysiques peuvent dépendre de l'espace, de la température et du temps⁵ : $k(M, T, t)$ et $c(M, T, t)$. La condition aux limites considérée est une condition de Fourier⁶ dont le coefficient de convection et la température extérieure sont variables : $h(M, T, T_{ext}, t)$ et $T_{ext}(M, t)$. Un terme source $\pi(M, t)$ est aussi pris en compte.

La frontière du domaine est notée $\partial\Omega$ et la normale unitaire extérieure est notée $\underline{n}_{\partial\Omega}$

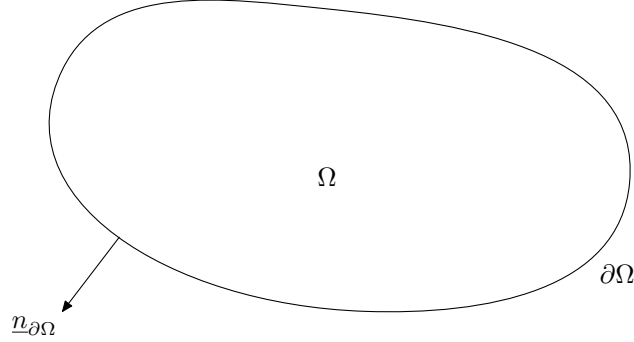


FIG. 2.1 – Domaine Ω avec sa frontière et sa normale

Les équations du problème sont donc :

$$\forall M \in \Omega \quad c \frac{\partial T}{\partial t} = \text{div}(k \underline{\nabla} T) + \pi \quad (2.1)$$

$$\forall M \in \partial\Omega \quad -k \underline{\nabla}(T) \bullet \underline{n}_{\partial\Omega} = h(T - T_{ext}) \quad (2.2)$$

$$\forall M \in \Omega \quad T(M, 0) = T_i(M) \quad (2.3)$$

Le champ de température initial T_i peut être quelconque. Contrairement à l'analyse modale classique (équations D.1, D.2 et D.3) il n'y a plus d'indice 0 car les propriétés thermophysiques ne sont pas linéarisées.

II.2. Etablissement de la formulation variationnelle

L'établissement d'une formulation variationnelle des équations 2.1 et 2.2 est un exercice classique. L'espace fonctionnel du champ de température T sous forme faible est l'espace $H^1(\Omega)$ [1].

L'établissement de la formulation variationnelle se fait en trois étapes. Dans un premier temps l'équation 2.1 est multipliée par une fonction v de $H^1(\Omega)$ et l'ensemble est intégré sur Ω . La fonction v est usuellement appelée fonction test.

$$\forall v \in H^1(\Omega) \quad \int_{\Omega} c \frac{\partial T}{\partial t} v \, d\Omega = \int_{\Omega} \text{div}(k \underline{\nabla} T) v \, d\Omega + \int_{\Omega} \pi v \, d\Omega \quad (2.4)$$

⁴Le lecteur intéressé par l'aspect transport est renvoyé à l'article [26].

⁵L'article [43] présente une application de la méthode en présence d'un changement de phase. Le changement de phase est pris en compte par l'intermédiaire d'une très forte variation de $c(M, T)$ sur un intervalle de température très serré.

⁶Il est parfaitement possible de prendre en compte un flux imposé, du rayonnement, ..., mais ces termes sont éliminés pour la clarté de l'exposé

Dans un deuxième temps, le théorème de Green A.4 permet de faire disparaître l'opérateur divergence du terme de droite de 2.4.

$$\forall v \in H^1(\Omega) \quad \int_{\Omega} c \frac{\partial T}{\partial t} v d\Omega = - \int_{\Omega} k \underline{\nabla}(T) \bullet \underline{\nabla}(v) d\Omega + \int_{\partial\Omega} v k \underline{\nabla}(T) \bullet \underline{n}_{\partial\Omega} d(\partial\Omega) + \int_{\Omega} \pi v d\Omega \quad (2.5)$$

Pour finir 2.5 est simplifiée à l'aide de la condition aux limites 2.2, ce qui établit la formulation variationnelle 2.6 du problème (équations 2.1, 2.2 et 2.3).

$$\left\{ \begin{array}{l} \forall v \in H^1(\Omega) \quad \int_{\Omega} c \frac{\partial T}{\partial t} v d\Omega = - \int_{\Omega} k \underline{\nabla}(T) \bullet \underline{\nabla}(v) d\Omega + \int_{\partial\Omega} v h (T_{ext} - T) d(\partial\Omega) + \int_{\Omega} \pi v d\Omega \\ \forall M \in \Omega \quad T(M, 0) = T_i(M) \end{array} \right. \quad (2.6)$$

II.3. Introduction des opérateurs intégraux

L'introduction d'opérateurs intégraux va permettre d'alléger la notation.

$$\left\{ \begin{array}{l} \mathcal{C}(f, g) = \int_{\Omega} c f g d\Omega \\ \mathcal{K}(f, g) = \int_{\Omega} k \underline{\nabla}(f) \bullet \underline{\nabla}(g) d\Omega \\ \mathcal{B}(f, g) = \int_{\partial\Omega} h f g d(\partial\Omega) \\ \mathcal{A}(g) = \int_{\partial\Omega} h T_{ext} g d(\partial\Omega) + \int_{\Omega} \pi g d\Omega \end{array} \right. \quad (2.7)$$

La formulation variationnelle 2.6 est alors équivalente à :

$$\left\{ \begin{array}{l} \forall v \in H^1(\Omega) \quad \mathcal{C} \left(\frac{\partial T}{\partial t}, v \right) = - (\mathcal{K}(T, v) + \mathcal{B}(T, v)) + \mathcal{A}(v) \\ \forall M \in \Omega \quad T(M, 0) = T_i(M) \end{array} \right. \quad (2.8)$$

Une remarque est ici nécessaire, les opérateurs $\mathcal{C}, \mathcal{K}, \mathcal{B}, \mathcal{A}$ ne sont pas invariants au cours du temps. Par exemple l'opérateur \mathcal{C} dépend de $c(M, T)$ or le champ de température est fonction du temps.

III. Formulation modale

Les modes de branche (z_i, V_i) sont présentés dans l'annexe E. La base de branche $\{z_i, V_i\}$ est composée de l'ensemble des modes de branche, elle forme une base de $H^1(\Omega)$. Le champ de température $T(M, t)$ se décompose donc dans l'espace des modes de branche :

$$T(M, t) = \sum_{i=1}^{\infty} x_i(t) V_i(M) \quad (2.9)$$

Contrairement à l'analyse modale classique, il n'y a plus de séparation entre le régime dynamique et le régime glissant. Les x_i sont les coefficients d'excitation ou états des fonctions propres V_i .

Les états initiaux sont obtenus à l'aide de E.44, d'où :

$$\forall i \in \mathbb{N}^* \quad x_i(0) = \mathcal{C}_b(T_i, V_i) \quad (2.10)$$

Il faut désormais établir une équation sur les états permettant de les déterminer au cours du temps.

III.1. Ecriture d'une équation d'état

L'équation d'état est obtenue facilement, il suffit de remplacer le champ de température par la décomposition 2.9 et d'utiliser les fonctions propres V_k comme fonction test dans 2.8.

Le champ T est donc remplacé par la décomposition 2.9 dans 2.8 (en oubliant provisoirement la condition initiale) :

$$\forall v \in H^1(\Omega) \quad \mathcal{C} \left(\frac{\partial \sum_{i=1}^{\infty} x_i V_i}{\partial t}, v \right) = - \left(\mathcal{K} \left(\sum_{i=1}^{\infty} x_i V_i, v \right) + \mathcal{B} \left(\sum_{i=1}^{\infty} x_i V_i, v \right) \right) + \mathcal{A}(v) \quad (2.11)$$

Les opérateurs $\mathcal{C}, \mathcal{K}, \mathcal{B}$ sont bi-linéaires, il est donc possible de sortir la sommation des opérateurs :

$$\forall v \in H^1(\Omega) \quad \sum_{i=1}^{\infty} \mathcal{C} \left(\frac{\partial x_i V_i}{\partial t}, v \right) = - \sum_{i=1}^{\infty} (\mathcal{K}(x_i V_i, v) + \mathcal{B}(x_i V_i, v)) + \mathcal{A}(v) \quad (2.12)$$

Les fonctions propres V_i sont uniquement fonction de l'espace ainsi $\frac{\partial x_i V_i}{\partial t} = \frac{dx_i}{dt} V_i$. La bi-linéarité des opérateurs $\mathcal{C}, \mathcal{K}, \mathcal{B}$ permet de sortir les états de ceux-ci, d'où :

$$\forall v \in H^1(\Omega) \quad \sum_{i=1}^{\infty} \left(\mathcal{C}(V_i, v) \frac{dx_i}{dt} \right) = - \sum_{i=1}^{\infty} ((\mathcal{K}(V_i, v) + \mathcal{B}(V_i, v)) x_i) + \mathcal{A}(v) \quad (2.13)$$

Pour finir les fonctions propres V_k sont utilisés comme fonction test⁷ v et la condition initiale est rappelée :

$$\left\{ \begin{array}{l} \forall k \in \mathbb{N}^* \quad \sum_{i=1}^{\infty} \left(\mathcal{C}(V_i, V_k) \frac{dx_i}{dt} \right) = - \sum_{i=1}^{\infty} ((\mathcal{K}(V_i, V_k) + \mathcal{B}(V_i, V_k)) x_i) + \mathcal{A}(V_k) \\ \forall i \in \mathbb{N}^* \quad x_i(0) = \mathcal{C}_b(T_i, V_i) \end{array} \right. \quad (2.14)$$

La formulation 2.14 comporte l'équation d'état et la condition initiale. Elle est théoriquement utilisable pour résoudre le problème thermique envisagé. En pratique, elle n'est pas utilisable car il y a une infinité de modes.

Les états sont découplés en analyse modale classique, ils sont couplés en analyse modale de branche car (en général) $\mathcal{C}(V_i, V_k) \neq 0$, les modes de branche n'étant orthogonaux que par rapport à \mathcal{C}_b et à \mathcal{A}_b (voir les équations E.25 et E.26).

IV. Réduction modale

La réduction modale qui consiste à ne plus utiliser la base de branche complète mais une base réduite $\{\tilde{z}_i, \tilde{V}_i\}$ permet de mener des calculs à partir d'une formulation modale. Les démarches de réduction de la base sont présentées dans l'annexe F.

La base réduite $\{\tilde{z}_i, \tilde{V}_i\}$ comporte un nombre fini de modes \tilde{N}_1 . Les méthodes présentées dans l'annexe F conservent l'orthogonalité par rapport à \mathcal{C}_b . Les fonctions \tilde{V}_i ne sont par contre plus forcément solution du problème de branche.

⁷La condition « pour toute fonction v appartenant à l'espace $H^1(\Omega)$ » est alors remplacée par la condition « pour toute fonction propre de la base de branche ». Ces deux conditions sont bien équivalentes car la base de branche est une base de $H^1(\Omega)$.

Le champ T n'est pas rigoureusement décomposable dans la base réduite. Une approximation est alors effectuée, le calcul dans l'espace modal ne détermine plus T mais le champ \tilde{T} tel que :

$$\tilde{T}(M, t) = \sum_{i=1}^{\tilde{N}_1} \tilde{x}_i(t) \tilde{V}_i(M) \quad (2.15)$$

Les états réduits sont déterminés par :

$$\left\{ \begin{array}{l} \forall k \in \llbracket 1, \tilde{N}_1 \rrbracket \quad \sum_{i=1}^{\tilde{N}_1} \left(\mathcal{C}(\tilde{V}_i, \tilde{V}_k) \frac{d\tilde{x}_i(t)}{dt} \right) = - \sum_{i=1}^{\tilde{N}_1} \left((\mathcal{K}(\tilde{V}_i, \tilde{V}_k) + \mathcal{B}(\tilde{V}_i, \tilde{V}_k)) \tilde{x}_i(t) \right) + \mathcal{A}(\tilde{V}_k) \\ \forall i \in \llbracket 1, \tilde{N}_1 \rrbracket \quad \tilde{x}_i(0) = \mathcal{C}_b(T_i, \tilde{V}_i) \end{array} \right. \quad (2.16)$$

Le problème 2.16 peut être résolu à l'aide de l'ordinateur. La démarche sera performante si les deux conditions suivantes sont vérifiées :

- L'écart entre le champ T et le champ réduit \tilde{T} reste faible.
- Le temps pris par la simulation modale est faible devant le temps d'une simulation de même précision par une autre méthode (éléments finis ou autre).

V. La méthode modale de branche confrontée à la méthode modale classique

Le tableau suivant résume les différences entre les approches modales de branche et les approches modales classiques :

Point de comparaison	méthode modale de branche	méthode modale classique
Nature du domaine	quelconque	quelconque
Propriétés thermophysiques	quelconques	le système thermique est linéaire ou linéarisé
Conditions aux limites	quelconques	dans le cas d'une condition de Fourier, le coefficient de convection est invariant
Condition initiale	tout champ initial T_i	Le champ initial T_i est presque quelconque, quelques restrictions sont néanmoins nécessaires
Champ thermique décomposé	le champ T	uniquement la partie dynamique T_d ce qui nécessite de séparer le régime glissant du régime dynamique
Equation d'état	tous les états sont couplés	les états sont découplés, le système thermique est alors « diagonalisé »

TAB. 2.1 – Comparaison entre la méthode modale de branche et la méthode modale classique

La méthode modale de branche permet donc de simuler des systèmes plus complexes que ceux qui sont simulables par la méthode modale classique. Par contre cela s'effectue au prix d'un couplage entre les états.

VI. Bilan des travaux numériques déjà réalisés sur la méthode modale de branche

La méthode modale de branche est déjà utilisée depuis quelques années, il est ainsi possible de donner quelques éléments pratiques sur celle-ci.

Le premier point est l'écart engendré par la réduction modale. Le champ de référence T est déterminé par la méthode des éléments finis. Le champ obtenu par la méthode modale \tilde{T} présente nécessairement un écart lié à la réduction modale. Cependant dans presque tous les cas la technique de l'amalgame modal découplé a montré sa performance. L'écart entre \tilde{T} et T reste faible, il est d'autant plus acceptable que la précision de calcul est limitée par l'imprécision⁸ des paramètres du problème de base.

Le second point est le gain en temps de calcul. Celui-ci est « faible » s'il est comparé au gain de temps de l'analyse modale classique. De manière générale il sera compris entre 2 et 80. Il existe cependant quelques réserves : la prise en compte des non-linéarités des propriétés thermophysiques (k et c) reste à améliorer. Il est possible qu'une méthode adaptée⁹ puisse améliorer les résultats actuels.

Le troisième point est la dimension du problème d'origine. La méthode des éléments finis discrétise l'espace, le nombre de degré de liberté dans le maillage est appelé ordre. Cet ordre est noté N , les algorithmes de calcul permettent de calculer des modes de manière performante pour des ordres relativement élevé. La mémoire RAM nécessaire au stockage des modes est proportionnelle à $N * \tilde{N}_0$, avec \tilde{N}_0 le nombre de modes calculés. L'utilisation conjointe d'un maillage fin et d'un grand nombre de modes posera un problème de mémoire RAM. La suite de ce manuscrit est dédié à la recherche d'une solution à ce problème. Elle consiste à découper le domaine en plusieurs sous-domaine, ce qui permet d'augmenter l'ordre des problèmes traités.

⁸Les propriétés thermophysiques et les coefficients d'échange (convection et rayonnement) ne sont connus que de manière approximative...

⁹En l'occurrence la méthode doit permettre un calcul rapide des termes $\mathcal{C}(V_i, V_k)$ et $\mathcal{K}(V_i, V_k)$. Actuellement la méthode utilisée est celle de l'annexe G (en particulier l'équation G.6). S'il y a des non-linéarités dans les propriétés thermo-physiques alors l'essentiel du temps de simulation est utilisé par le calcul des termes $\mathcal{C}(V_i, V_k)$ et $\mathcal{K}(V_i, V_k)$.

Deuxième partie

Théorie et pratique du couplage par conduction

Théorie du couplage par conduction

« Ce qu'on sait, savoir qu'on le sait, ce qu'on
ne sait pas, savoir qu'on ne le sait pas, c'est
savoir véritablement »

Confucius

Contenu du chapitre

I.	Notion de résistance de contact	36
II.	Obtention d'une formulation variationnelle	37
III.	Modes de branche et couplage par conduction	41
IV.	Ouverture sur la condition de contact parfait	45

Le présent chapitre présente la partie analytique de la méthode développée pour le couplage de modèles modaux de branche.

La méthode utilise des fonctionnelles de saut, ces fonctionnelles avaient été introduites dans le cadre de la synthèse modale [18]. La synthèse modale est une technique de calcul des modes d'un domaine Ω qui utilise les modes des sous-domaines $\Omega^{(q)}$. Un biais est admis lorsque l'on effectue une réduction modale, ce biais est responsable de l'introduction d'un saut de température et de flux au niveau des interfaces entre sous-domaines. Les sauts sont contrôlables, dans une certaine limite, à l'aide de fonctionnelles de saut.

La fonctionnelle de saut de température s'interprète physiquement de manière simple, elle correspond exactement à la prise en compte d'une résistance thermique de contact (*RTC*). Ce chapitre commence donc par présenter la notion de *RTC*. La formulation variationnelle utilisée pour les éléments finis est ensuite établie.

Puis une nouvelle base modale dite structurée utilisant les bases de branche locales est introduite. L'équation d'état est établie pour la base de branche structurée. Une méthode de réduction modale est alors proposée, elle implique une non-conservation du flux. Une fonctionnelle de saut de flux est alors proposée, l'amélioration apportée par cette fonctionnelle sera présentée dans le chapitre 4.

La présentation se limite aux cas traités dans la suite du manuscrit : il y a ainsi plusieurs corps et les systèmes thermiques sont purement diffusifs.

Le chapitre est clôturé par une ouverture au cas des coupures virtuelles, il s'agit alors d'approximer un champ continu par des fonctions discontinues. La théorie ne permet pas de conclure, ainsi ce point fera l'objet d'expériences numériques présentées au chapitre 5.

I. Notion de résistance de contact

Lorsque deux solides sont en contact, du fait de leurs rugosités et de la différence de forme de leurs surfaces, le contact ne s'effectue jamais sur toute la surface apparente, mais seulement en certaines zones de surface très faible devant la surface apparente. Le surface réelle de contact représente environ 1% de la surface apparente pour un contact entre deux surfaces métalliques.

L'espace interstitiel entre les zones de contact est en général un mauvais conducteur de la chaleur. Cet espace constitue un frein au transfert de chaleur, qui de ce fait passe de manière préférentielle au niveau des contact directs. Le champ de chaleur est ainsi fortement perturbé, il y a constriction des lignes de flux. Cette constriction peut être prise en compte par une résistance thermique de contact (RTC). Le phénomène de constriction est en effet localisé à des échelles microscopiques et la modélisation qui suit est suffisante au niveau macroscopique.

Dans le cas des régimes thermiques lentement variables (temps caractéristique supérieur à 50s), il est possible de négliger les effets d'inertie au niveau de l'interface de contact. La résistance de contact est alors identifiable à la résistance de contact du régime permanent. Dans cette approximation, la réalité physique de la constriction des lignes de flux est remplacée par l'apparition d'un saut de température au passage de l'interface de contact, le champ thermique étant alors considéré comme non perturbé.

Le contact entre deux corps se traduit alors par :

- la conservation du flux au passage de l'interface.
- une dépendance linéaire entre le flux et le saut de température.

La valeur de la résistance thermique dépend en pratique d'un nombre important de paramètres :

- état géométrique des surfaces avant assemblage
- répartition des surfaces en contact réel
- propriétés mécaniques et thermiques des matériaux
- température
- temps (oxydation des surfaces...)
- ...

La détermination de la valeur de la résistance de contact est hors de notre propos, puisque ce travail de thèse vise uniquement à développer des outils de calcul numérique.

Pour fixer les idées, on peut cependant donner des ordres de grandeur : Bardon ([4]) propose l'échelle suivante pour des contacts entre surfaces métalliques¹ :

- $RTC \leq 10^{-7} m^2.K.W^{-1}$: métaux soudés ou déposés par voie chimique.
- $RTC \in [10^{-6} \quad 10^{-5}] m^2.K.W^{-1}$: contact avec un matériau d'interface (graisse conductrice).
- $RTC \in [10^{-5} \quad 10^{-4}] m^2.K.W^{-1}$: surfaces rugueuses présentant une bonne coïncidence
- $RTC \in [10^{-4} \quad 10^{-3}] m^2.K.W^{-1}$: surfaces rugueuses présentant une mauvaise coïncidence

Il est possible de rendre cette échelle plus parlante en donnant des épaisseurs de matériaux équivalents. En régime statique, un matériau de conductivité k et d'épaisseur e présente une résistance thermique surfacique R qui vérifie :

$$R = \frac{e}{k} \quad (3.1)$$

¹Attention, ces valeurs sont données pour les pressions de contact relativement fortes (1MPa et plus)

Ce qui permet de dresser le tableau suivant :

Résistance ($m^2.K.W^{-1}$)	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
$k (W.m^{-1}.K^{-1})$	épaisseur équivalente en mm				
400	400	40	4	0.4	0.04
40	40	4	0.4	0.04	0.004
1	1	0.1	0.1	0.01	0.001
0.1	0.1	0.01	0.001	0.0001	0.00001

TAB. 3.1 – Epaisseur équivalente en fonction de la conductivité du matériaux

Dans les applications pratiques, la prise en compte d'une résistance de contact n'est ainsi pas forcément nécessaire. Par contre elle est primordiale lorsque les deux matériaux en contact sont de bons conducteurs de la chaleur, comme dans le cas de l'interface entre un composant électronique et son radiateur (voir chapitre 4).

II. Obtention d'une formulation variationnelle

Nous établissons ici, la formulation variationnelle d'un problème de diffusion thermique comportant au moins deux corps en contact.

II.1. Notations

Le domaine physique Ω (de frontière $\partial\Omega$) comporte N_Ω sous parties $\Omega^{(q)}$ (de frontières $\partial\Omega^{(q)}$) avec $q \in \llbracket 1, N_\Omega \rrbracket$.

Il existe N_Γ frontières de contact. La frontière commune à $\Omega^{(q)}$ et à $\Omega^{(p)}$ est notée $\Gamma_{qp}^{(k)}$ où $k \in \llbracket 1, N_\Gamma \rrbracket$ est le numéro de la frontière de contact.

On note Γ la réunion de l'ensemble des frontières de contact : $\Gamma \stackrel{def}{=} \bigcup_{k=1}^{N_\Gamma} \Gamma_{qp}^{(k)}$.

Les frontières internes $\Gamma_{qp}^{(k)}$ n'appartiennent pas à la frontière $\partial\Omega$ du domaine Ω . Ainsi $\partial\Omega \neq \bigcup_{q=1}^{N_\Omega} \partial\Omega^{(q)}$.

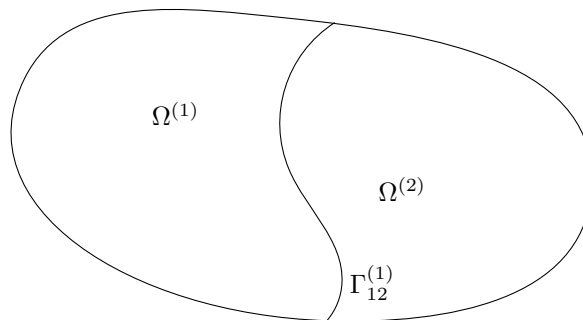


FIG. 3.1 – Domaine Ω constitué de deux sous domaines $\Omega^{(1)}$ et $\Omega^{(2)}$

On note $\underline{n}^{(q)}$ la normale extérieure à $\Omega^{(q)}$ et :

$$\forall M \in \Gamma_{qp}^{(k)} \quad \underline{n}_{\Gamma_{qp}^{(k)}} \stackrel{def}{=} \underline{n}^{(q)} = -\underline{n}^{(p)} = -\underline{n}_{\Gamma_{pq}^{(k)}} \quad (3.2)$$

T désigne le champ température sur l'ensemble du domaine Ω , sa restriction à $\Omega^{(q)}$ est notée $T^{(q)}$. Le champ T possède une double valuation au niveau des interfaces puisqu'il y est discontinu.

On note $[[f(M)]]_{\Gamma_{qp}^{(k)}}$ le saut de f (grandeur scalaire ou vectorielle) au point M à l'interface $\Gamma_{qp}^{(k)}$ entre $\Omega^{(q)}$ et $\Omega^{(p)}$:

$$\forall M \in \Gamma_{qp}^{(k)} \quad [[f(M)]]_{\Gamma_{qp}^{(k)}} \stackrel{def}{=} f^{(p)}(M) - f^{(q)}(M) \quad (3.3)$$

Le saut de f en passant la frontière de $\Omega^{(q)}$ vers $\Omega^{(p)}$ est donc la valeur à la frontière dans $\Omega^{(p)}$ moins la valeur à la frontière dans $\Omega^{(q)}$. Par définition, on a :

$$\forall M \in \Gamma_{qp}^{(k)} \quad [[f(M)]]_{\Gamma_{qp}^{(k)}} = -[[f(M)]]_{\Gamma_{pq}^{(k)}} \quad (3.4)$$

La loi de Fourier liant le gradient de la température au flux thermique s'écrit :

$$\forall M \in \Omega^{(q)} \quad \underline{\Phi}^{(q)} \stackrel{def}{=} -k^{(q)} \underline{\nabla} T^{(q)} \quad (3.5)$$

On note $\varphi^{(q)}$ le flux **algébrique sortant** à travers $\partial\Omega^{(q)}$:

$$\forall M \in \partial\Omega^{(q)} \quad \varphi^{(q)} \stackrel{def}{=} -k^{(q)} \underline{\nabla} T^{(q)} \bullet \underline{n}^{(q)} \quad (3.6)$$

II.2. Relations traduisant la résistance de contact

La prise en compte d'une résistance de contact se fait par le biais de deux équations correspondant à la conservation du flux d'une part et à la relation reliant le flux au saut de température d'autre part.

II.2.1. Condition de conservation du flux au passage de l'interface

La condition de conservation du flux s'écrit :

$$\forall M \in \Gamma_{qp}^{(k)} \quad \varphi^{(q)} + \varphi^{(p)} = 0 \quad (3.7)$$

Ce qui revient à dire que la somme des flux algébriques sortant est nulle à une interface de contact (conservation de l'énergie).

On peut écrire 3.7 en terme de relation de saut à l'interface :

$$\forall M \in \Gamma_{qp}^{(k)} \quad [[\underline{\Phi}]]_{\Gamma_{qp}^{(k)}} \bullet \underline{n}_{\Gamma_{qp}^{(k)}} = 0 \quad (3.8)$$

3.7 et 3.8 sont bien équivalentes puisque :

$$\forall M \in \Gamma_{qp}^{(k)} \quad [[\underline{\Phi}]]_{\Gamma_{qp}^{(k)}} \bullet \underline{n}_{\Gamma_{qp}^{(k)}} = \left(\underline{\Phi}^{(p)} - \underline{\Phi}^{(q)} \right) \bullet \underline{n}_{\Gamma_{qp}^{(k)}} \quad (3.9)$$

$$= \left(-k^{(p)} \underline{\nabla} T^{(p)} + k^{(q)} \underline{\nabla} T^{(q)} \right) \bullet \underline{n}_{\Gamma_{qp}^{(k)}} \quad (3.10)$$

$$= -k^{(p)} \underline{\nabla} T^{(p)} \bullet \underline{n}_{\Gamma_{qp}^{(k)}} + k^{(q)} \underline{\nabla} T^{(q)} \bullet \underline{n}_{\Gamma_{qp}^{(k)}} \quad (3.11)$$

$$= k^{(p)} \underline{\nabla} T^{(p)} \bullet \underline{n}^{(p)} + k^{(q)} \underline{\nabla} T^{(q)} \bullet \underline{n}^{(q)} \quad (3.12)$$

$$= -\left(\varphi^{(p)} + \varphi^{(q)} \right) \quad (3.13)$$

On fera **très attention** à l'égalité suivante :

$$\forall M \in \Gamma_{qp}^{(k)} \quad \llbracket \Phi \rrbracket_{\Gamma_{qp}^{(k)}} \bullet \underline{n}_{\Gamma_{qp}^{(k)}} = \llbracket \Phi \bullet \underline{n}_{\Gamma_{qp}^{(k)}} \rrbracket_{\Gamma_{qp}^{(k)}} \quad (3.14)$$

En effet :

$$\forall M \in \Gamma_{qp}^{(k)} \quad \llbracket \Phi \rrbracket_{\Gamma_{qp}^{(k)}} \bullet \underline{n}_{\Gamma_{qp}^{(k)}} = \left(\Phi^{(p)} - \Phi^{(q)} \right) \bullet \underline{n}_{\Gamma_{qp}^{(k)}} \quad (3.15)$$

$$= \left(-k^{(p)} \underline{\nabla} T^{(p)} + k^{(q)} \underline{\nabla} T^{(q)} \right) \bullet \underline{n}_{\Gamma_{qp}^{(k)}} \quad (3.16)$$

$$= -k^{(p)} \underline{\nabla} T^{(p)} \bullet \underline{n}_{\Gamma_{qp}^{(k)}} + k^{(q)} \underline{\nabla} T^{(q)} \bullet \underline{n}_{\Gamma_{qp}^{(k)}} \quad (3.17)$$

$$= \llbracket \Phi \bullet \underline{n}_{\Gamma_{qp}^{(k)}} \rrbracket_{\Gamma_{qp}^{(k)}} \quad (3.18)$$

Ce qui permet d'écrire 3.7 sous la forme suivante :

$$\forall M \in \Gamma_{qp}^{(k)} \quad \llbracket \Phi \bullet \underline{n}_{\Gamma_{qp}^{(k)}} \rrbracket_{\Gamma_{qp}^{(k)}} = 0 \quad (3.19)$$

II.2.2. Relation entre le flux et le saut de température au passage de l'interface

Dans le cas d'une résistance de contact, il faut écrire une relation entre le flux et le saut de température. La modélisation choisie considère la relation linéaire suivante :

$$\forall M \in \Gamma_{qp}^{(k)} \quad RTC(\diamond) \varphi^{(q)} = - \llbracket T \rrbracket_{\Gamma_{qp}^{(k)}} \quad (3.20)$$

La résistance thermique de contact RTC dépend d'une multitude de paramètres (\diamond). Le nombre de paramètres dépend du problème à modéliser et est externe à la méthode modale. La méthode développée peut être appliquée avec une résistance thermique fonction de l'espace, de la température, du temps...

II.3. Equations du problème thermique

Le problème thermique considéré est un problème de diffusion de la chaleur avec production volumique d'énergie dont les surfaces extérieures échangent de l'énergie par convection.

Les équations du problème sont alors :

$$\forall M \in \Omega \setminus \Gamma \quad c \frac{\partial T}{\partial t} = \text{div} (k \underline{\nabla} (T)) + \pi \quad (3.21)$$

$$\forall M \in \partial\Omega \quad -k \underline{\nabla} (T) \bullet \underline{n}_{\partial\Omega} = h(T - T_{ext}) \quad (3.22)$$

$$\forall M \in \Gamma_{qp}^{(k)} \quad \llbracket \Phi \bullet \underline{n}_{\Gamma_{qp}^{(k)}} \rrbracket_{\Gamma_{qp}^{(k)}} = 0 \quad (3.23)$$

$$\forall M \in \Gamma_{qp}^{(k)} \quad RTC(\diamond) \varphi^{(q)} = - \llbracket T \rrbracket_{\Gamma_{qp}^{(k)}} \quad (3.24)$$

$$\forall M \in \Omega \quad T(M, 0) = T_i(M) \quad (3.25)$$

II.4. L'espace fonctionnel $H^1(\Omega, \Gamma)$

Les méthodes variationnelles sont basées sur l'utilisation de fonctions test. Ces fonctions test doivent permettre de décrire l'espace fonctionnel dans lequel se trouve la solution. Le champ de température n'est pas continu, l'espace fonctionnel des fonctions test ne peut donc pas être l'espace $H^1(\Omega)$.

Les champs sur les $\Omega^{(q)}$ appartiennent aux espaces $H^1(\Omega^{(q)})$. L'espace des fonctions test sera donc la somme directe des $H^1(\Omega^{(q)})$, cet espace sera noté $H^1(\Omega, \Gamma)$:

$$H^1(\Omega, \Gamma) \stackrel{def}{=} \bigoplus_{q=1}^{N_\Omega} H^1(\Omega^{(q)}) \quad (3.26)$$

II.5. Formulation variationnelle du problème physique

L'équation 3.21 est multipliée par une fonction test $v \in H^1(\Omega, \Gamma)$ et est intégrée sur le domaine Ω :

$$\forall v \in H^1(\Omega, \Gamma) \quad \int_{\Omega} c \frac{\partial T}{\partial t} v d\Omega = \int_{\Omega} \text{div}(k \underline{\nabla}(T)) v d\Omega + \int_{\Omega} v \pi d\Omega \quad (3.27)$$

Le champ de température présente une discontinuité au niveau de l'interface, il appartient donc à $H^1(\Omega, \Gamma)$. Le théorème de Green-Ostrogradsky A.11 est appliqué à 3.27 :

$$\begin{aligned} \forall v \in H^1(\Omega, \Gamma) \quad \int_{\Omega} c \frac{\partial T}{\partial t} v d\Omega = & - \int_{\Omega} k \underline{\nabla}(v) \bullet \underline{\nabla}(T) d\Omega + \int_{\partial\Omega} v k \underline{\nabla}(T) \bullet \underline{n}_{\partial\Omega} d(\partial\Omega) \\ & - \sum_{k=1}^{N_{\Gamma}} \int_{\Gamma_{qp}^{(k)}} \llbracket v k \underline{\nabla}(T) \rrbracket_{\Gamma_{qp}^{(k)}} \bullet \underline{n}_{\Gamma_{qp}^{(k)}} d\Gamma_{qp}^{(k)} + \int_{\Omega} v \pi d\Omega \end{aligned} \quad (3.28)$$

La condition aux limites 3.22 permet de simplifier un terme :

$$\int_{\partial\Omega} v k \underline{\nabla}(T) \bullet \underline{n}_{\partial\Omega} d(\partial\Omega) = - \int_{\partial\Omega} v h(T - T_{ext}) d(\partial\Omega) \quad (3.29)$$

Pour le second terme, on utilise la conservation du flux (3.7) :

$$\forall M \in \Gamma_{qp}^{(k)} \quad \llbracket v k \underline{\nabla}(T) \rrbracket_{\Gamma_{qp}^{(k)}} \bullet \underline{n}_{\Gamma_{qp}^{(k)}} = v^{(p)} \varphi^{(p)} + v^{(q)} \varphi^{(q)} \quad (3.30)$$

$$= v^{(p)} \varphi^{(p)} - v^{(q)} \varphi^{(p)} \quad (3.31)$$

$$= \llbracket v \rrbracket_{\Gamma_{qp}^{(k)}} \varphi^{(p)} \quad (3.32)$$

$$= - \llbracket v \rrbracket_{\Gamma_{qp}^{(k)}} \varphi^{(q)} \quad (3.33)$$

Ainsi 3.28 devient :

$$\begin{aligned} \forall v \in H^1(\Omega, \Gamma) \quad \int_{\Omega} c \frac{\partial T}{\partial t} v d\Omega = & - \int_{\Omega} k \underline{\nabla}(v) \bullet \underline{\nabla}(T) d\Omega - \int_{\partial\Omega} v h(T - T_{ext}) d(\partial\Omega) \\ & + \sum_{k=1}^{N_{\Gamma}} \int_{\Gamma_{qp}^{(k)}} \llbracket v \rrbracket_{\Gamma_{qp}^{(k)}} \varphi^{(q)} d\Gamma_{qp}^{(k)} + \int_{\Omega} v \pi d\Omega \end{aligned} \quad (3.34)$$

On peut utiliser la relation 3.24 sur 3.34, ce qui permet d'obtenir la formulation faible du problème :

$$\left| \begin{aligned} \forall v \in H^1(\Omega, \Gamma) \quad \int_{\Omega} c \frac{\partial T}{\partial t} v d\Omega = & - \int_{\Omega} k \underline{\nabla}(v) \bullet \underline{\nabla}(T) d\Omega - \int_{\partial\Omega} v h(T - T_{ext}) d(\partial\Omega) \\ & - \sum_{k=1}^{N_{\Gamma}} \int_{\Gamma_{qp}^{(k)}} \frac{\llbracket v \rrbracket_{\Gamma_{qp}^{(k)}} \llbracket T \rrbracket_{\Gamma_{qp}^{(k)}}}{RTC(\diamond)} d\Gamma_{qp}^{(k)} + \int_{\Omega} v \pi d\Omega \end{aligned} \quad (3.35) \right.$$

$$\forall M \in \Omega \quad T(M, 0) = T_i(M)$$

II.6. Formulation en terme d'opérateurs

L'écriture des équations sous forme variationnelle est assez lourde, l'introduction d'opérateurs intégraux va permettre de l'alléger. Les opérateurs \mathcal{C} , \mathcal{K} , \mathcal{B} , \mathcal{J}_{Γ}^T (voir 3.36) sont symétriques, définis et positifs. \mathcal{A} est une forme linéaire.

$$\left\{ \begin{array}{l}
 \mathcal{C}(f, v) = \int_{\Omega} c f v d\Omega \\
 \mathcal{K}(f, v) = \int_{\Omega} k \nabla(v) \bullet \nabla(f) d\Omega \\
 \mathcal{B}(f, v) = \int_{\partial\Omega} v h f d(\partial\Omega) \\
 \mathcal{J}_{\Gamma}^T(f, v) = \sum_{k=1}^{N_{\Gamma}} \int_{\Gamma_{qp}^{(k)}} \frac{[v]_{\Gamma_{qp}^{(k)}} [f]_{\Gamma_{qp}^{(k)}}}{RTC(\diamond)} d\Gamma_{qp}^{(k)} \\
 \mathcal{A}(v) = \int_{\partial\Omega} v h T_{ext} d(\partial\Omega) + \int_{\Omega} v \pi d\Omega
 \end{array} \right. \quad (3.36)$$

La formulation variationnelle 3.35 est alors équivalente à :

$$\left\{ \begin{array}{l}
 \forall v \in H^1(\Omega, \Gamma) \quad \mathcal{C}\left(\frac{\partial T}{\partial t}, v\right) = -(\mathcal{K}(T, v) + \mathcal{B}(T, v) + \mathcal{J}_{\Gamma}^T(T, v)) + \mathcal{A}(v) \\
 \forall M \in \Omega \quad T(M, 0) = T_i(M)
 \end{array} \right. \quad (3.37)$$

III. Modes de branche et couplage par conduction

La formulation faible 3.37 est directement utilisable avec une méthode d'éléments finis. Il est possible d'utiliser une base modale dès lors que celle-ci forme une base de $H^1(\Omega, \Gamma)$. Une base de branche est fondamentalement composée de modes continus et ne peut ainsi pas être une base de $H^1(\Omega, \Gamma)$. Une nouvelle base dite base de branche structurée sera donc définie.

L'écriture de l'équation d'état dans la base de branche structurée sera naturelle. Une procédure de réduction de la base de branche structurée sera proposée. La réduction de la base de branche structurée ne permettra plus la conservation du flux au niveau des interfaces de contact. Un terme complémentaire abstrait sera introduit dans l'objectif de limiter cette non-conservation.

III.1. Base de branche structurée

Une base de branche ne pouvant pas former une base de $H^1(\Omega, \Gamma)$, il faut définir une nouvelle base modale. On rappelle que $H^1(\Omega, \Gamma)$ est égal à la somme directe des $H^1(\Omega^{(q)})$. Or la base de branche de $\Omega^{(q)}$ forme une base de $H^1(\Omega^{(q)})$. Il paraît donc intéressant de réutiliser les bases de branche locales pour former la nouvelle base.

La base des modes de branche structurés sera notée $\{z_{n_i^q}, \mathcal{V}_{n_i^q}\}$. Le n -ème vecteur de branche structuré $\mathcal{V}_{n_i^q}$ est égal au i -ème vecteur de branche de $\Omega^{(q)}$ prolongé par zéro sur le reste de Ω . La n -ème valeur propre $z_{n_i^q}$ est égale à la i -ème valeur propre de la base de branche de $\Omega^{(q)}$. De plus tous les modes de chacune des bases locales sont utilisés pour former la base de branche structurée.

La base de branche structurée ainsi formée est une base de $H^1(\Omega, \Gamma)$. Elle peut donc décomposer tous les champs de $H^1(\Omega, \Gamma)$. Le champ de température du domaine complet T appartenant à l'espace $H^1(\Omega, \Gamma)$, il est décomposable dans la base de branche structurée :

$$T = \sum_{n=1}^{\infty} x_{n_i^p} \mathcal{V}_{n_i^p} \quad (3.38)$$

III.2. L'équation d'état

L'équation d'état est obtenue facilement, il suffit de remplacer le champ de température par la décomposition 3.38 et d'utiliser les fonctions propres $\mathcal{V}_{k_j^q}$ comme fonction test dans 3.37

La formulation variationnelle 3.37 est (la condition initiale est temporairement omise) :

$$\forall v \in H^1(\Omega, \Gamma) \quad \mathcal{C} \left(\frac{\partial T}{\partial t}, v \right) = - (\mathcal{K}(T, v) + \mathcal{B}(T, v) + \mathcal{J}_\Gamma^T(T, v)) + \mathcal{A}(v) \quad (3.39)$$

T est remplacé par la décomposition 3.38 :

$$\begin{aligned} \forall v \in H^1(\Omega, \Gamma) \quad \mathcal{C} \left(\frac{\partial \sum_{n=1}^{\infty} x_{n_i^p} \mathcal{V}_{n_i^p}}{\partial t}, v \right) = \\ - \left(\mathcal{K} \left(\sum_{n=1}^{\infty} x_{n_i^p} \mathcal{V}_{n_i^p}, v \right) + \mathcal{B} \left(\sum_{n=1}^{\infty} x_{n_i^p} \mathcal{V}_{n_i^p}, v \right) + \mathcal{J}_\Gamma^T \left(\sum_{n=1}^{\infty} x_{n_i^p} \mathcal{V}_{n_i^p}, v \right) \right) + \mathcal{A}(v) \end{aligned} \quad (3.40)$$

Les opérateurs $\mathcal{C}, \mathcal{K}, \mathcal{B}, \mathcal{J}_\Gamma^T$ sont bi-linéaires, il est ainsi possible de sortir la sommation :

$$\begin{aligned} \forall v \in H^1(\Omega, \Gamma) \quad \sum_{n=1}^{\infty} \mathcal{C} \left(\frac{\partial x_{n_i^p} \mathcal{V}_{n_i^p}}{\partial t}, v \right) = \\ - \sum_{n=1}^{\infty} \left(\mathcal{K} \left(x_{n_i^p} \mathcal{V}_{n_i^p}, v \right) + \mathcal{B} \left(x_{n_i^p} \mathcal{V}_{n_i^p}, v \right) + \mathcal{J}_\Gamma^T \left(x_{n_i^p} \mathcal{V}_{n_i^p}, v \right) \right) + \mathcal{A}(v) \end{aligned} \quad (3.41)$$

Les fonctions propres ne dépendent que de l'espace ainsi :

$$\begin{aligned} \forall v \in H^1(\Omega, \Gamma) \quad \sum_{n=1}^{\infty} \mathcal{C} \left(\frac{\partial x_{n_i^p} \mathcal{V}_{n_i^p}}{\partial t}, v \right) = \\ - \sum_{n=1}^{\infty} \left(\mathcal{K} \left(x_{n_i^p} \mathcal{V}_{n_i^p}, v \right) + \mathcal{B} \left(x_{n_i^p} \mathcal{V}_{n_i^p}, v \right) + \mathcal{J}_\Gamma^T \left(x_{n_i^p} \mathcal{V}_{n_i^p}, v \right) \right) + \mathcal{A}(v) \end{aligned} \quad (3.42)$$

La bi-linéarité des opérateurs $\mathcal{C}, \mathcal{K}, \mathcal{B}, \mathcal{J}_\Gamma^T, \mathcal{A}$ permet de sortir les états :

$$\begin{aligned} \forall v \in H^1(\Omega, \Gamma) \quad \sum_{n=1}^{\infty} \mathcal{C} \left(\mathcal{V}_{n_i^p}, v \right) \frac{\partial x_{n_i^p}}{\partial t} = \\ - \sum_{n=1}^{\infty} \left(\mathcal{K} \left(\mathcal{V}_{n_i^p}, v \right) + \mathcal{B} \left(\mathcal{V}_{n_i^p}, v \right) + \mathcal{J}_\Gamma^T \left(\mathcal{V}_{n_i^p}, v \right) \right) x_{n_i^p} + \mathcal{A}(v) \end{aligned} \quad (3.43)$$

Pour finir les fonctions propres $\mathcal{V}_{k_j^q}$ sont utilisées comme fonction test v et la condition initiale est rappelée :

$$\left| \begin{aligned} \forall k \in \mathbb{N}^* \quad \sum_{n=1}^{\infty} \mathcal{C} \left(\mathcal{V}_{n_i^p}, \mathcal{V}_{k_j^q} \right) \frac{\partial x_{n_i^p}}{\partial t} = \\ - \sum_{n=1}^{\infty} \left(\mathcal{K} \left(\mathcal{V}_{n_i^p}, \mathcal{V}_{k_j^q} \right) + \mathcal{B} \left(\mathcal{V}_{n_i^p}, \mathcal{V}_{k_j^q} \right) + \mathcal{J}_\Gamma^T \left(\mathcal{V}_{n_i^p}, \mathcal{V}_{k_j^q} \right) \right) x_{n_i^p} + \mathcal{A} \left(\mathcal{V}_{k_j^q} \right) \\ \forall n \in \mathbb{N}^* \quad x_i(0) = \mathcal{C}_b(T_i, \mathcal{V}_{n_i^p}) \end{aligned} \right. \quad (3.44)$$

Le problème 3.44 est équivalent à 3.37, il comporte l'équation d'état et les conditions initiales. Tous les états sont couplés par l'équation d'état.

III.3. Réduction de la base de branche

L'utilisation de l'ensemble d'une base modale est possible dans une démonstration théorique². Cependant la sommation sur une infinité de modes rend impossible toute simulation numérique utilisant l'équation d'état 3.44. Pour pouvoir effectuer des simulations numériques, il faut utiliser une base modale réduite qui comporte un nombre fini de modes. Si le nombre de modes est faible, alors il est possible que le temps de calcul par la méthode modale soit faible par rapport à une méthode plus classique (éléments finis, ...).

La base réduite de la base de branche structurée est appelée base de branche structurée réduite et est notée $\{\tilde{z}_{n_i^q}, \tilde{\mathcal{V}}_{n_i^q}\}$. $\{\tilde{z}_{n_i^q}, \tilde{\mathcal{V}}_{n_i^q}\}$ comporte un nombre fini \tilde{N} de modes. $\{\tilde{z}_{n_i^q}, \tilde{\mathcal{V}}_{n_i^q}\}$ est une base de $H_{\tilde{N}}^1(\Omega, \Gamma)$ qui est un sous-espace de $H^1(\Omega, \Gamma)$.

$\{\tilde{z}_{n_i^q}, \tilde{\mathcal{V}}_{n_i^q}\}$ est formée à partir des bases de branche locales réduites :

- pour chaque domaine $\Omega^{(q)}$, les $\tilde{N}_0^{(q)}$ premiers modes de branche sont calculés et une réduction par amalgame modal [43] permet d'obtenir une base de branche réduite comportant $\tilde{N}_1^{(q)}$ modes
- les vecteurs propres des bases de branche locales réduites sont prolongés par zéro sur le reste du domaine et sont regroupés dans la base de branche structurée réduite. La valeur propre $\tilde{z}_{n_i^q}$ associée est égale à la valeur propre du mode réduit d'origine.

La dimension \tilde{N} de $\{\tilde{z}_{n_i^q}, \tilde{\mathcal{V}}_{n_i^q}\}$ vérifie l'égalité suivante : $\tilde{N} = \sum_{q=1}^{N_\Omega} \tilde{N}_1^{(q)}$

Le champ T n'est pas rigoureusement décomposable dans $\{\tilde{z}_{n_i^q}, \tilde{\mathcal{V}}_{n_i^q}\}$, cependant si la réduction est efficace alors le champ \tilde{T} est proche du champ T :

$$\forall t, \quad \forall M \in \Omega \quad T \approx \tilde{T} = \sum_{n=1}^{\tilde{N}} \tilde{x}_{n_i^j}(t) \tilde{\mathcal{V}}_{n_i^j} \quad (3.45)$$

L'équivalent de 3.44 pour les modes de branche structurés réduits est :

$$\left\{ \begin{array}{l} \forall k \in \llbracket 1, \tilde{N} \rrbracket \quad \sum_{n=1}^{\tilde{N}} \mathcal{C}(\tilde{\mathcal{V}}_{n_i^p}, \tilde{\mathcal{V}}_{k_j^q}) \frac{\partial \tilde{x}_{n_i^p}}{\partial t} = \\ \quad - \sum_{n=1}^{\tilde{N}} \left(\mathcal{K}(\tilde{\mathcal{V}}_{n_i^p}, \tilde{\mathcal{V}}_{k_j^q}) + \mathcal{B}(\tilde{\mathcal{V}}_{n_i^p}, \tilde{\mathcal{V}}_{k_j^q}) + \mathcal{J}_\Gamma^T(\tilde{\mathcal{V}}_{n_i^p}, \tilde{\mathcal{V}}_{k_j^q}) \right) \tilde{x}_{n_i^p} + \mathcal{A}(\tilde{\mathcal{V}}_{k_j^q}) \\ \forall n \in \llbracket 1, \tilde{N} \rrbracket \quad \tilde{x}_i(0) = \mathcal{C}_b(T_i, \tilde{\mathcal{V}}_{n_i^p}) \end{array} \right. \quad (3.46)$$

²Par exemple, la démonstration de la régularisation d'un problème inverse par une méthode de *Tikonov* [22]

III.4. Introduction d'une fonctionnelle de saut de flux

La réduction de la base de branche structurée s'effectue en admettant une approximation sur le champ de température. Le champ T est alors approximé par le champ \tilde{T} , ce qui implique une approximation sur le flux thermique. De plus les raccords sur les interfaces de contact deviennent incompatibles, c'est-à-dire qu'il ne sera plus possible d'assurer la conservation du flux au passage des interfaces. Or lors de l'écriture de 3.35, il a été supposé que le flux se conservait (passage de 3.30 à 3.31)

La discontinuité en flux n'est pas physique et n'existe qu'à cause de la réduction modale. En adoptant la démarche utilisée dans les méthodes de Galerkin discontinues ([9]), il est possible de rajouter dans la formulation variationnelle (et donc dans l'équation d'état) un terme supplémentaire qui vise à pénaliser la discontinuité.

Le terme supplémentaire vise à pénaliser une grandeur qui n'a pas d'existence physique, ainsi il n'existe pas un choix absolu mais un large panel de fonctionnelles de pénalisation possible. Cependant les opérateurs $\mathcal{C}, \mathcal{K}, \mathcal{B}, \mathcal{J}_\Gamma^T$ présents dans 3.37 sont bi-linéaires, symétriques, définis et positifs, il semble judicieux que le terme supplémentaire possède ces propriétés. Le choix le plus naturel est alors le choix de la fonctionnelle de saut de flux suivante :

$$\mathcal{J}_\Gamma^\Phi(f, v) = \sum_{k=1}^{N_\Gamma} \beta^{(k)} \int_{\Gamma_{qp}^{(k)}} \llbracket k \nabla(f) \bullet \underline{n}_{qp}^{(k)} \rrbracket_{\Gamma_{qp}^{(k)}} \llbracket k \nabla(v) \bullet \underline{n}_{qp}^{(k)} \rrbracket_{\Gamma_{qp}^{(k)}} d\Gamma_{qp}^{(k)} \quad (3.47)$$

Le facteur de pondération $\beta^{(q)}$ sert d'une part à assurer l'homogénéité dimensionnelle et d'autre part à équilibrer le poids numérique de ce nouveau terme par rapport aux termes déjà présents dans la formulation variationnelle.

La relation 3.46 avec la prise en compte de la fonctionnelle de saut de flux est alors :

$$\left\{ \begin{array}{l} \forall k \in \llbracket 1, \tilde{N} \rrbracket \quad \sum_{n=1}^{\tilde{N}} \mathcal{C}(\tilde{\mathcal{V}}_{n_i^p}, \tilde{\mathcal{V}}_{k_j^q}) \frac{\partial \tilde{x}_{n_i^p}}{\partial t} = \\ \quad - \sum_{n=1}^{\tilde{N}} \left(\mathcal{K}(\tilde{\mathcal{V}}_{n_i^p}, \tilde{\mathcal{V}}_{k_j^q}) + \mathcal{B}(\tilde{\mathcal{V}}_{n_i^p}, \tilde{\mathcal{V}}_{k_j^q}) + \mathcal{J}_\Gamma^T(\tilde{\mathcal{V}}_{n_i^p}, \tilde{\mathcal{V}}_{k_j^q}) + \mathcal{J}_\Gamma^\Phi(\tilde{\mathcal{V}}_{n_i^p}, \tilde{\mathcal{V}}_{k_j^q}) \right) \tilde{x}_{n_i^p} + \mathcal{A}(\tilde{\mathcal{V}}_{k_j^q}) \\ \forall n \in \llbracket 1, \tilde{N} \rrbracket \quad \tilde{x}_i(0) = \mathcal{C}_b(T_i, \tilde{\mathcal{V}}_{n_i^p}) \end{array} \right. \quad (3.48)$$

Le choix de $\beta^{(q)}$ devient le point clef de la méthode proposée, son choix étant le résultat d'un compromis. En effet si $\beta^{(q)}$ est trop faible alors son action ne sera pas sensible. Au contraire si sa valeur est trop élevée alors le problème initial sera écrasé par la fonctionnelle de saut de flux.

IV. Ouverture sur la condition de contact parfait

Dans le cas, où l'on voudrait traiter un problème de grande taille, il peut s'avérer judicieux de découper virtuellement un corps. Une coupure virtuelle correspond à une résistance de contact nulle, le champ de température est alors continu. Au niveau fonctionnel, ceci peut très facilement se prendre en compte en imposant la continuité des fonctions test sur l'interface de raccord. Par contre la réduction modale rendant les modes incompatibles sur l'interface de contact, il est impossible de prendre en compte rigoureusement une coupure virtuelle dans la méthode modale proposée.

Cependant il est possible de modifier un peu la physique en remplaçant le contact parfait par une résistance de contact très faible et d'utiliser la méthode proposée avec cette résistance de contact. Le choix numérique de la résistance de contact peut être effectuée à l'aide de 3.1, en considérant une épaisseur e négligeable par rapport au problème considéré.

Ceci sera mis en application dans l'exemple 3D (chapitre 5).

Pratique du couplage par conduction en 2D

« La seule certitude que j'ai, c'est d'être dans
le doute »

Pierre Desproges

Contenu du chapitre

I.	Présentation du problème	48
II.	Etude des modèles éléments finis	50
III.	Recherche des bases modales locales	58
IV.	Réduction de la base de branche structurée et étude <i>a posteriori</i> du paramètre β de la fonctionnelle de saut de flux	61
V.	Choix <i>a priori</i> du paramètre β de la fonctionnelle de saut de flux	69
VI.	Conclusion sur la méthode de couplage en 2D	70

Ce chapitre présente une première application permettant de valider la méthode de couplage présentée théoriquement dans le précédent chapitre. On cherche alors à s'affranchir le plus possible des lourdeurs et des difficultés techniques de mise en œuvre de ces outils sous forme numérique.

On choisit ainsi une configuration simple, caractérisée par deux corps modélisés en deux dimensions. De plus l'exemple considéré intègre naturellement une résistance thermique de contact (*RTC*). Ainsi la fonctionnelle de saut de température est physiquement déterminée, contrairement à la fonctionnelle de saut de flux qui est toujours un artifice numérique.

Dans un premier temps, le dispositif physique considéré est présenté : il s'agit d'un processeur accolé à son radiateur. Puis une deuxième partie aborde l'analyse du choix du modèle classique (éléments finis), que l'on nomme modèle détaillé (MD) et qui sert de référence. On verra alors que différents modèles détaillés sont utilisés selon l'objectif visé.

Enfin l'étude modale est menée. Elle montre l'intérêt de la méthode et analyse l'influence des différents paramètres. L'étude insiste tout particulièrement sur la difficulté du choix du cas test lors de la phase de réduction par amalgame. En effet la réduction est effectuée par sous domaine, il faut donc un cas test pour chaque sous-domaine. Or pour être efficace, ce cas test ne doit pas être trop éloigné du cas réel, qui lui est couplé. Pour l'application considérée, les résultats numériques montrent la faible sensibilité de ces paramètres d'amalgame. La réduction modale par la méthode de l'amalgame modal découplé [43] est donc adaptée à la méthode de sous-structuration proposée.

Pour terminer, une étude concernant le choix du paramètre de pondération de la fonctionnelle de saut de flux (β) est menée. Cette fois, les essais numériques font apparaître la forte sensibilité des résultats

à ce paramètre, dont la valeur optimale varie avec l'ordre de réduction du modèle réduit.

Aucun critère simple n'a été trouvé concernant le choix du paramètre β , à défaut une procédure permettant d'indiquer *a priori* la valeur optimale de β est proposée. Cette procédure nécessite le balayage du modèle réduit en fonction du paramètre β .

I. Présentation du problème

Le dispositif étudié est un microprocesseur avec son radiateur en 2D. La production volumique d'énergie est telle que le saut de température au niveau de la surface de contact entre le microprocesseur et le radiateur n'est plus négligeable. Le choix entre les deux vitesses de rotation du ventilateur est effectué par rapport à la température d'un point particulier du radiateur. Cette disposition est déjà mise en œuvre sur des ventilateurs commerciaux qui sont fournis avec une sonde de température.

I.1. Géométrie

Le processeur est constitué d'un support en céramique, d'un cœur à base de silicium et d'un élément en cuivre servant à répartir le flux thermique à évacuer. Le support en céramique à cause de sa faible conductivité thermique n'est pas modélisé, il est remplacé par une frontière adiabatique. Le radiateur est en aluminium.

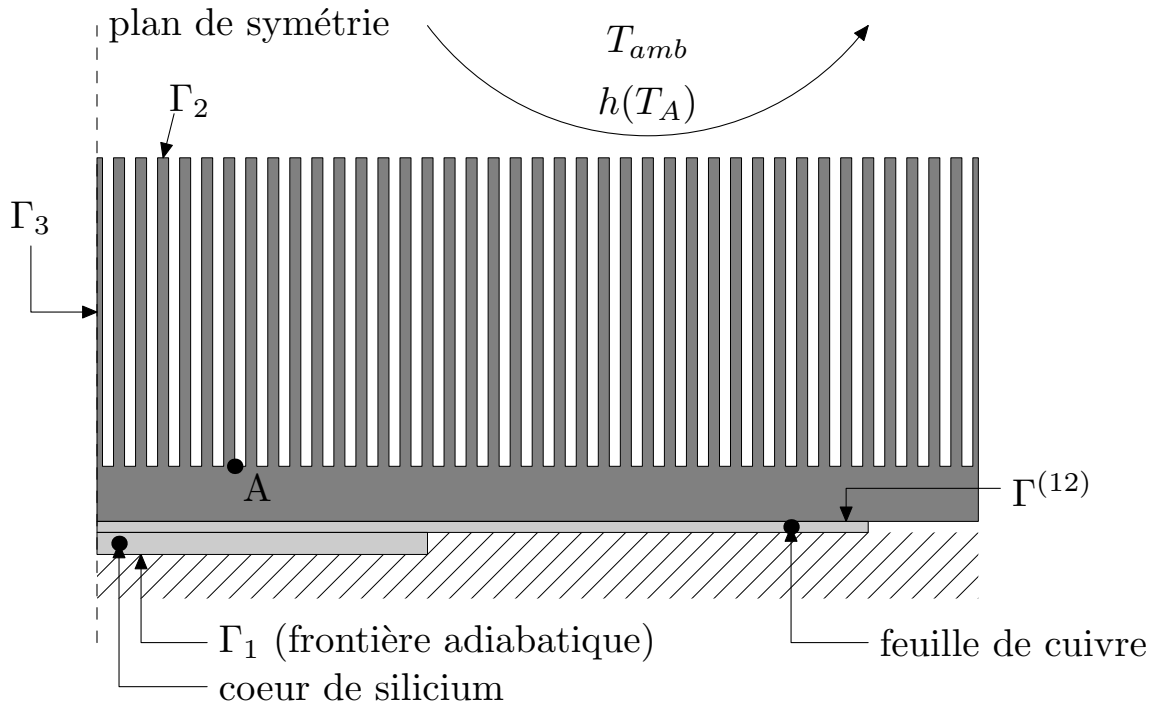


FIG. 4.1 – Géométrie du processeur avec son radiateur (à l'échelle 2)

La numérotation des corps est la suivante : $\Omega^{(1)}$ désigne le processeur et $\Omega^{(2)}$ désigne le radiateur.

I.2. Propriétés thermophysiques et puissances volumiques dégagées

Les propriétés thermophysiques des matériaux sont extraites de [27]. Les valeurs utilisées pour ces propriétés et pour les puissances volumiques π dégagées dans les différents éléments sont présentées dans le tableau suivant :

	k	c	π
	$W.m^{-1}.K^{-1}$	$J.m^{-3}.K^{-1}$	$W.m^{-3}$
aluminium	216	2340000	0
cuivre	380	3363000	0
silicium	145	1514500	$3.4 * 10^7$

TAB. 4.1 – Caractéristiques des matériaux et production volumique d'énergie

L'article [21] dresse une revue complète des matériaux d'interface pour les composants électroniques. Dans ce domaine la pression de contact reste relativement faible (de $0.02MPa$ à $0.3MPa$). La résistance thermique de contact est alors relativement forte, elle varie classiquement¹ dans l'intervalle $[1*10^{-5} \quad 6.5*10^{-5}] m^2.K.W^{-1}$ s'il y a un matériau d'interface (graisse de contact, ...). La valeur que l'on retient est une valeur moyenne sur l'ensemble des matériaux disponibles soit : $RTC = 2.5 * 10^{-5} m^2.K.W^{-1}$.

A titre de comparaison l'épaisseur équivalente à la résistance de contact considérée est de $5.4mm$ pour de l'aluminium et de $9.5mm$ pour du cuivre. La prise en compte de la résistance de contact est donc nécessaire dans la modélisation.

I.3. Loi de contrôle du coefficient de convection

Le contrôle de la vitesse s'effectue par rapport à la température du point A de la figure 4.1. La loi de contrôle présente un hystérésis :

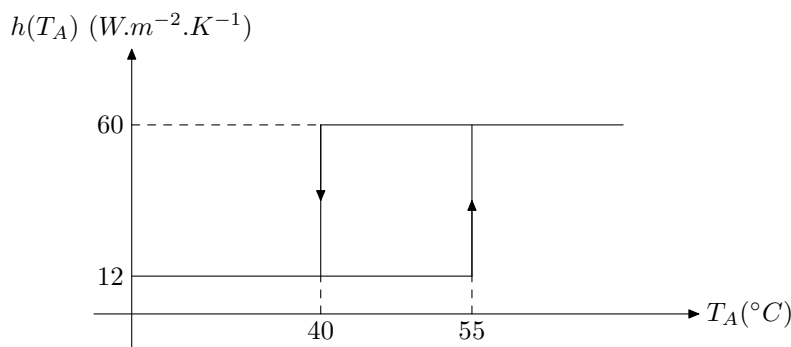


FIG. 4.2 – Loi de contrôle du ventilateur

I.4. Champ de température initial et température de l'ambiance

A l'instant initial, le champ de température est égal sur l'ensemble du domaine à la température de l'ambiance : $T_{amb} = 20^{\circ}C$.

¹Il existe quelques matériaux récents qui permettent de réduire un peu plus la résistance de contact mais ils sont peu durables.

I.5. Horizon temporel et sauvegarde

Le temps physique simulé va de 0s à 1000s. Les sauvegardes se font à pas de temps fixe $\delta_{t,sauvegarde}$. Ce pas de temps est calculé de manière à avoir $N_{sauv} = 1000$ temps de sauvegarde. Ainsi l'horizon temporel $\Delta t = 1000s$ et le pas de temps de sauvegarde sont reliés par :

$$\delta_{t,sauvegarde} = \frac{\Delta t}{N_{sauv} + 1} \quad (4.1)$$

II. Etude des modèles éléments finis

Les éléments finis sont très bien adaptés au cas des géométries complexes. Ils vont être utilisés pour rechercher une solution numérique de référence et pour le calcul des modes. En pratique, trois modèles éléments finis vont être nécessaires pour pouvoir mener l'analyse de ce chapitre.

Un premier modèle élément fini, noté MD_A ², utilise un maillage très fin et des paramètres de solveurs très contraignants. Le modèle MD_A est alors très précis, son champ de température servira ensuite de référence numérique. Par contre son temps de calcul est très important et ne peut pas servir pour une comparaison avec les modèles réduits.

L'amplitude thermique maximale ΔT_{max} , qui servira de critère de choix pour déterminer les autres modèles détaillés, est obtenue à partir du modèle MD_A (voir graphique 4.8 en page 55) :

$$\Delta T_{max} = \max_t (\max_{\Omega} (T_{MD_A})) - \min_t (\min_{\Omega} (T_{MD_A})) = 36.45K \quad (4.2)$$

Un second modèle élément fini, noté MD_B utilise les paramètres des solveurs du modèle MD_A avec un maillage plus grossier. Ce modèle est tel que $\epsilon_{AB,max}$ soit inférieur à 1% de ΔT_{max} avec :

$$\epsilon_{AB,max} = \max_t (\max_{\Omega} (abs (T_{MD_B} - T_{MD_A}))) \quad (4.3)$$

Le maillage du modèle MD_B sera qualifié d'adapté car il assure une précision acceptable sans être pour autant excessivement fin. Ce maillage sera ainsi utilisé pour la suite de l'étude. Le champ de température T_{MD_B} de ce modèle servira de référence pour déterminer la précision des différents modèles réduits.

Un troisième modèle élément fini, noté MD_C utilise le maillage du modèle MD_B avec des paramètres de solveurs tels que $\epsilon_{BC,max}$ soit inférieur à 1% de ΔT_{max} avec :

$$\epsilon_{BC,max} = \max_t (\max_{\Omega} (abs (T_{MD_C} - T_{MD_B}))) \quad (4.4)$$

Le temps de calcul du modèle MD_C servira de référence pour déterminer la rapidité de la méthode modale.

Le graphique 4.3 donne une vue synthétique des informations sur les modèles éléments finis utilisés.

²MD étant l'abréviation de modèle détaillé. La racine des modèles modaux sera MR (modèle réduit)

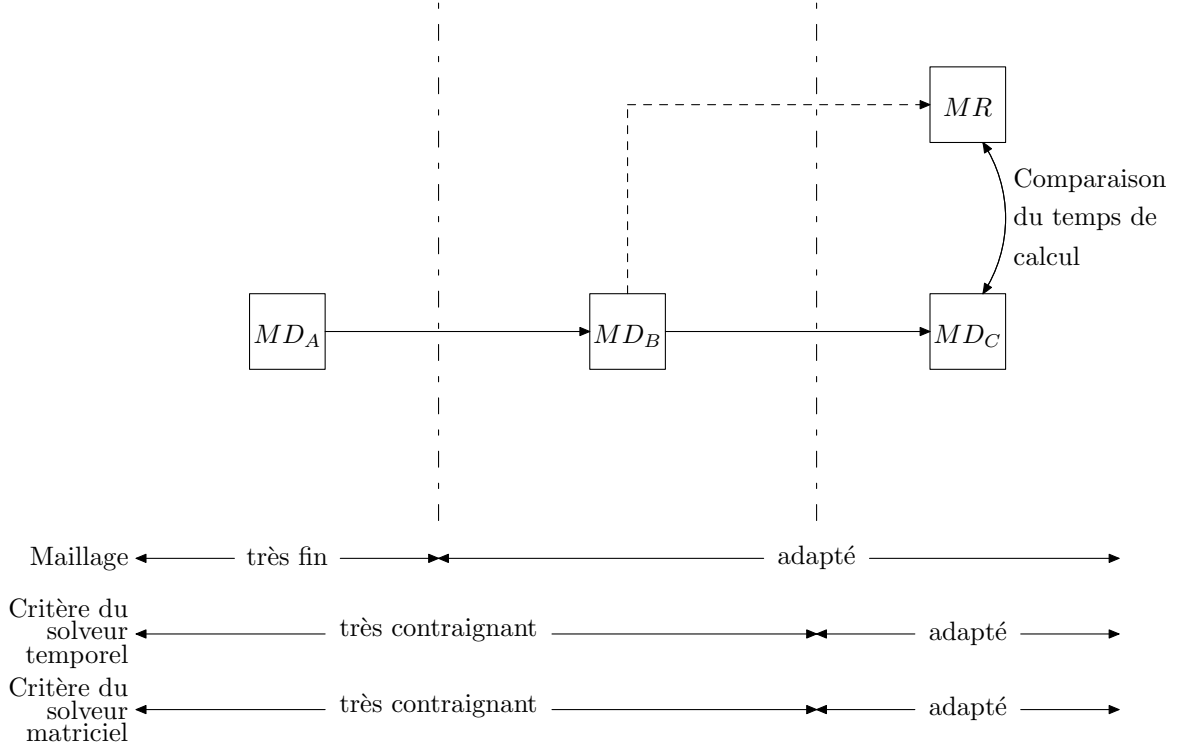


FIG. 4.3 – Résumé des modèles éléments finis utilisés

Les éléments finis utilisés sont des éléments P^1 , le schéma temporel est un schéma du premier ordre implicite avec un pas de temps variable. La résolution à chaque pas de temps du système matriciel est effectuée par l'algorithme du gradient conjugué préconditionné par la diagonale [28]. Le critère de convergence ϵ_{GC} s'effectue sur la norme du résidu :

$$\mathbf{x} \text{ est solution de } \mathbb{A}\mathbf{x} = \mathbf{B} \text{ si : } \|\mathbf{B} - \mathbb{A}\mathbf{x}\|_2 < \epsilon_{GC} \quad (4.5)$$

Le solveur étant à pas de temps variable, il faut utiliser un (ou des) critère(s) pour contrôler le pas de temps. La figure 4.4 illustre la gestion du pas de temps. Cette gestion est basée sur le maximum de l'évolution du vecteur représentant le champ de température lors du pas de temps courant (i) : $\epsilon^i = \|\mathbf{T}^{i+1} - \mathbf{T}^i\|_\infty$.

Si cet écart de température ϵ^i est supérieur au critère ϵ_{max} alors le pas de temps δ_t^i est divisé par un facteur choisi $f_{max} \geq 1$, et le calcul est repris à partir du temps initial de l'étape soit t^i . Cela permet d'assurer une bonne prise en compte des dynamiques rapides.

Sinon, le calcul passe à l'étape suivante. Dans ce cas, le pas de temps est éventuellement augmenté par multiplication par un facteur choisi $f_{min} \geq 1$ si l'écart de température ϵ^i est inférieur au critère ϵ_{min} .

De plus, il faut éviter une erreur trop importante dans la prise en compte de l'hystérésis du coefficient de convection. Une localisation temporelle du changement de valeur de $h(T_A)$ est alors effectuée par interpolation linéaire. Soit $T_{consigne}$ la température pour laquelle se produit l'hystérésis ($40^\circ C$ ou $55^\circ C$, voir figure 4.2). $T_{consigne}$ est compris entre T_A^i à T_A^{i+1} . Soit α le paramètre d'interpolation :

$$\alpha = \frac{T_{consigne} - T_A^i}{T_A^{i+1} - T_A^i} \quad (4.6)$$

La nouvelle valeur de t^{i+1} est égale à $t^i + \alpha * \delta_t^i$ et la nouvelle valeur de \mathbf{T}^{i+1} est égale à $\mathbf{T}^i + \alpha * (\mathbf{T}^{i+1} - \mathbf{T}^i)$.

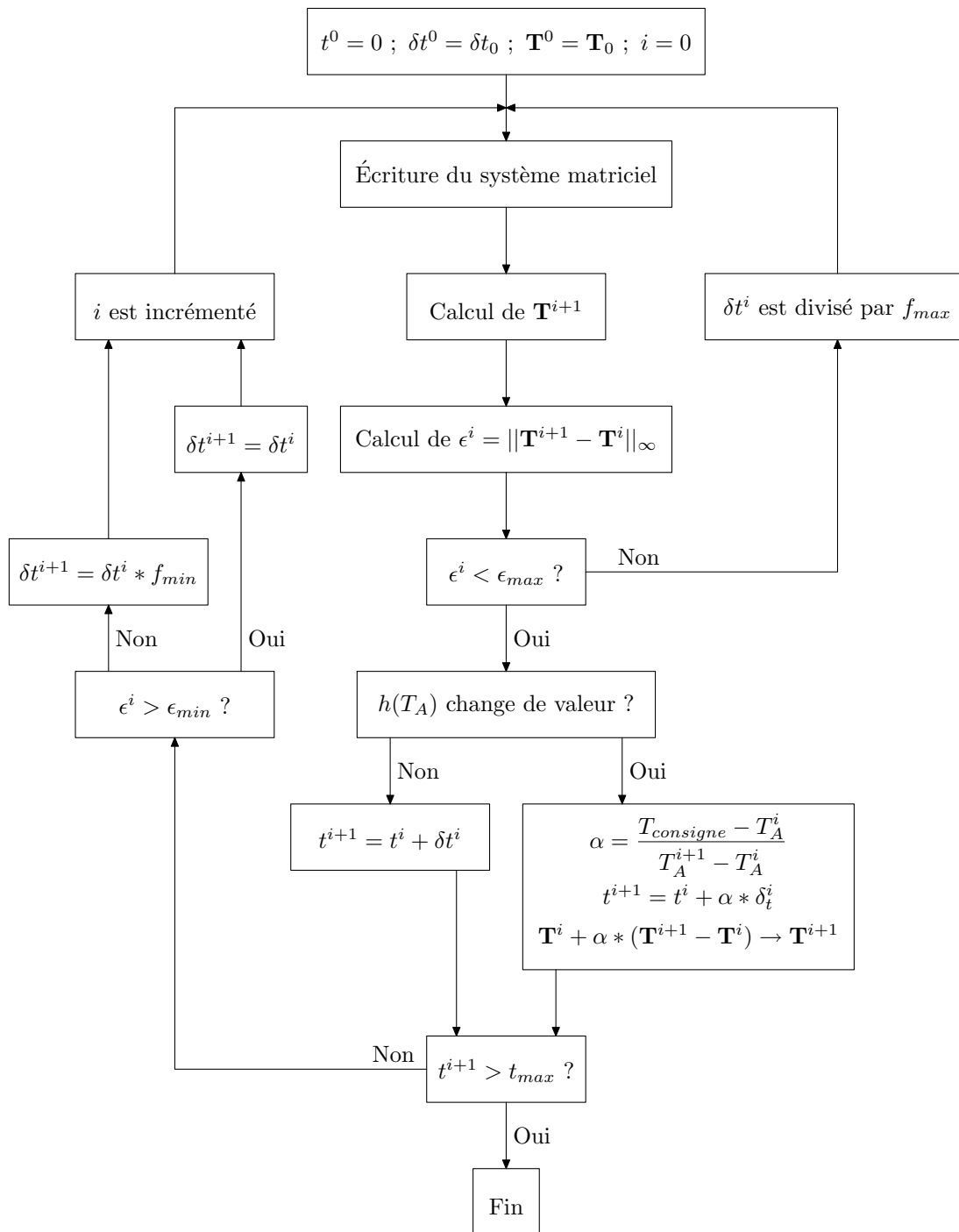


FIG. 4.4 – Schéma général du pas de temps variable pour les modèles détaillés

II.1. Recherche d'un modèle de référence : modèle MD_A

D'après ce qui a déjà été dit, le modèle détaillé MD_A est un modèle très précis qui sert de référence pour le processus visant à déterminer les deux modèles détaillés utilisés par l'étude modale : MD_B et MD_C .

Il n'est pas possible de prévoir à l'avance les paramètres (taille de maillage, critères des solveurs matriciel et temporel) qui vont définir le modèle MD_A . La recherche du modèle MD_A sera donc itérative.

Les itérations portent sur le maillage et sur les paramètres des solveurs. Les facteurs f_{min} et f_{max} , qui définissent l'évolution du pas de temps de calcul, sont fixés de manière semi-arbitraire aux valeurs suivantes : $f_{min} = 1.2$ et $f_{max} = 2$.

Le processus itératif sépare le maillage et les paramètres de solveur. Une série de modèles MD_{A_i} est construite, ces modèles utilisent des maillages de plus en plus fins. Le maillage du modèle MD_{A_0} comporte 5800 nœuds. Le maillage est raffiné géométriquement avec un facteur multiplicatif sur le nombre de nœuds proche de 1.3.

Les paramètres des solveurs de chacun des modèles MD_{A_i} sont recherchés de manière itérative. Pour chaque modèle MD_{A_i} , une série de modèles $MD_{A_{i,j}}$ utilisant des paramètres de plus en plus contraignants et le maillage du modèle MD_{A_i} est construite. La modification des critères est géométrique :

- soit le couple $(\epsilon_{min}, \epsilon_{max})$, qui caractérise l'évolution de champ de température entre deux pas de temps, est divisé par un facteur 1.5
- soit ϵ_{GC} , qui est le critère de convergence du solveur matriciel, est divisé par un facteur 10.

Par défaut c'est le couple $(\epsilon_{min}, \epsilon_{max})$ qui est diminué. ϵ_{GC} n'est modifié qu'une fois que l'influence du couple $(\epsilon_{min}, \epsilon_{max})$ est devenue suffisamment faible ($\epsilon_{A_{i,j-1}A_{i,j},max} \leq 0.01K$). Si ϵ_{GC} est modifié alors il est vérifié que la diminution $(\epsilon_{min}, \epsilon_{max})$ n'entraîne pas $\epsilon_{A_{i,j-1}A_{i,j},max} \geq 0.01K$, le processus est répété si ce n'est pas le cas.

Cette recherche itérative nécessite donc deux critères d'arrêt. Le premier critère porte sur les modèles $MD_{A_{i,j}}$ associés au modèle MD_{A_i} . Les paramètres du modèle MD_{A_i} sont fixés à la valeur du modèle $MD_{A_{i,j}}$ dès lors que $\epsilon_{A_{i,j-1}A_{i,j},max}$ est inférieur à $0.01K$ avec :

$$\epsilon_{A_{i,j-1}A_{i,j},max} = \max_t \left(\max_{\Omega} \left(\text{abs} \left(T_{MD_{A_{i,j}}} - T_{MD_{A_{i,j-1}}} \right) \right) \right) \quad (4.7)$$

Le second critère d'arrêt porte sur les modèles MD_{A_i} . Le raffinement du maillage est arrêté lorsque $\epsilon_{A_{i-1}A_i,max}$ est inférieur à $0.01K$ avec :

$$\epsilon_{A_{i-1}A_i,max} = \max_t \left(\max_{\Omega} \left(\text{abs} \left(T_{MD_{A_i}} - T_{MD_{A_{i-1}}} \right) \right) \right) \quad (4.8)$$

Le modèle MD_A retenu est alors identifié au dernier modèle MD_{A_i} , ce qui donne les paramètres suivants :

Paramètre	Unité	Valeur
Nombre de nœuds	-	28008
ϵ_{GC}	-	10^{-12}
ϵ_{max}	K	0.002
f_{max}	-	2
ϵ_{min}	K	0.0005
f_{min}	-	1.2

TAB. 4.2 – Paramètres du modèle MD_A

Les valeurs obtenues (tableau 4.2) pour le modèle MD_A sont inhabituellement faibles. Ceci s'explique par la forme particulière de la loi de contrôle du ventilateur.

Considérons l'équation différentielle suivante :

$$\begin{cases} y(0) = 0 \\ \frac{dy(x)}{dx} = f(y(x)) - y(x) \end{cases} \quad (4.9)$$

Où la fonction f présente un hystérésis :

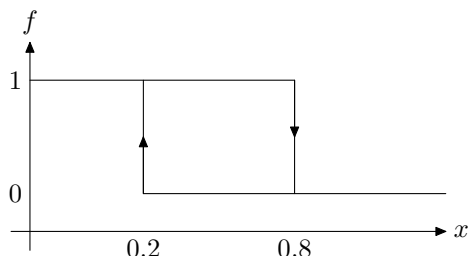


FIG. 4.5 – Hystérésis de la fonction f

L'équation différentielle 4.9 est résolue par un schéma aux différences finies avec un pas δ_x constant, et une localisation du saut de l'hystérésis. Les 4 courbes de la figure 4.6 montrent l'effet du pas de discrétisation δx sur le résultat du calcul : un faible décalage temporel entraîne un fort écart de température.

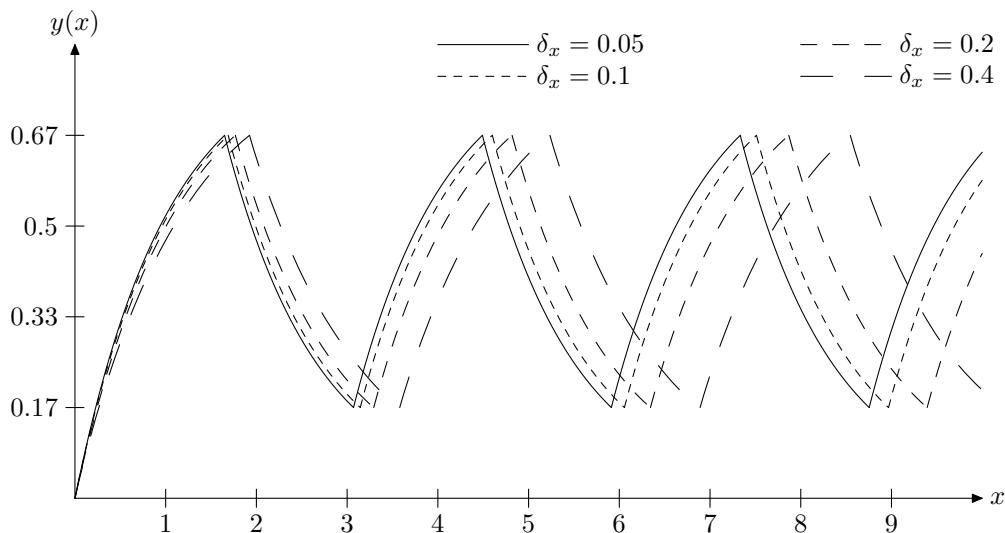


FIG. 4.6 – Effet du pas de discrétisation δx sur le résultat calculé par différences finies de 4.9

Pour mémoire, le pas de discrétisation δ_x est beaucoup moins pénalisant pour une équation différentielle plus classique :

$$\begin{cases} y(0) = 0 \\ \frac{dy(x)}{dx} = 0.5(1 + \cos(x)) - y(x) \end{cases} \quad (4.10)$$

L'équation différentielle 4.10 est résolue par un schéma aux différences finies avec un pas δ_x constant :

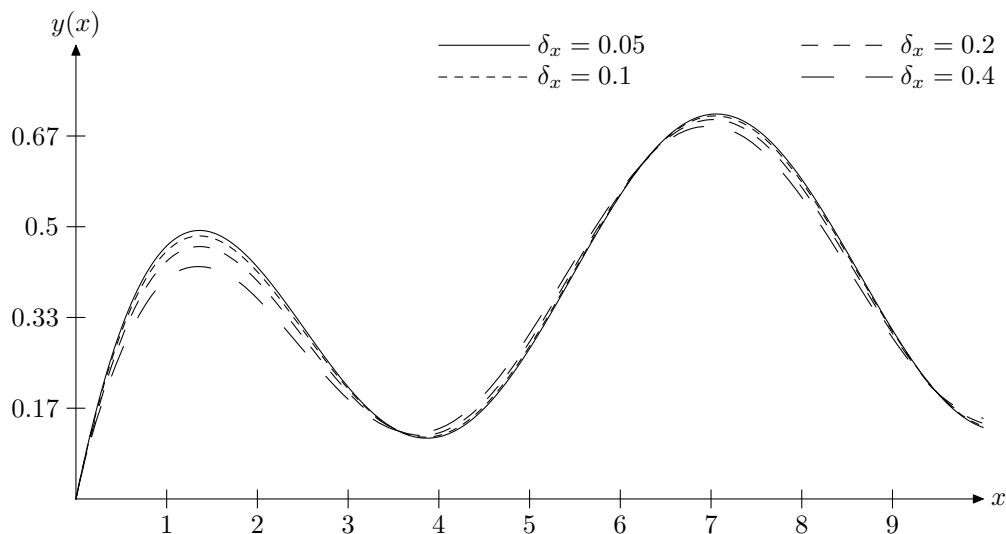


FIG. 4.7 – Effet du pas de discrétisation δx sur le résultat calculé par différences finies de 4.10

Le modèle MD_A défini par les paramètres du tableau 4.2 sert de référence pour déterminer les modèles MD_B et MD_C . La précision attendue est alors exprimée par rapport à l'amplitude thermique maximale ΔT_{max} (équation 4.2) :

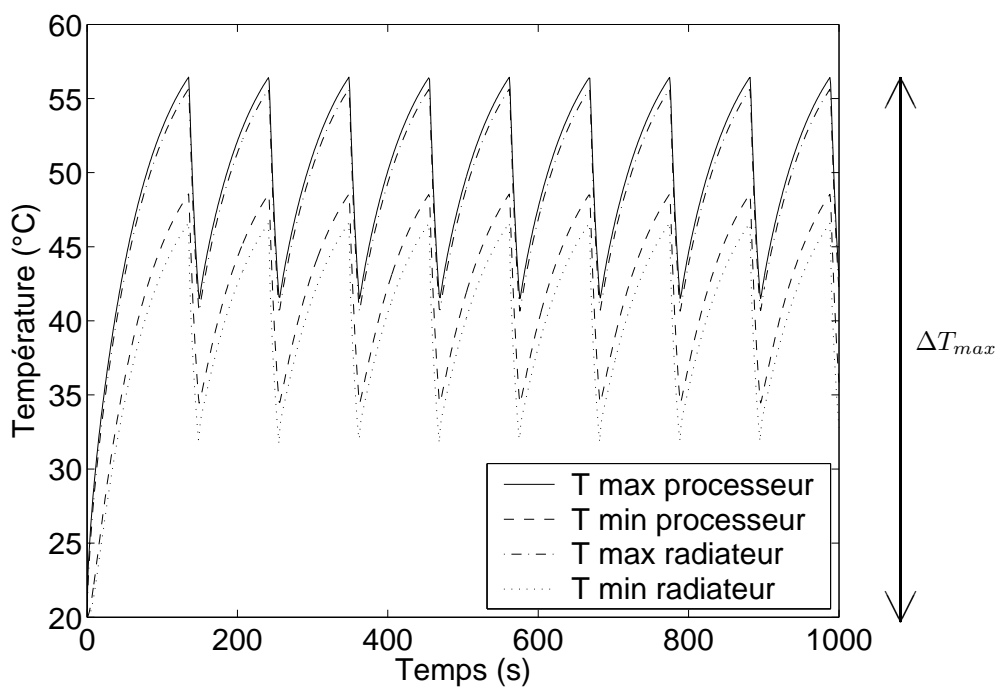


FIG. 4.8 – Evolutions thermiques maximales et minimales dans le processeur et dans le radiateur

II.2. Recherche d'un modèle à maillage adapté : modèle MD_B

Le modèle MD_B sert à déterminer un maillage adapté au problème physique. Le critère retenu est $\epsilon_{AB,max} < 0.01 * \Delta T_{max}$ (voir 4.2 et 4.3). MD_B est recherché de manière itérative à partir de la connaissance de MD_A . Les paramètres des solveurs matriciels et temporels sont inchangés par rapport au modèle MD_A . La recherche ne porte alors que sur le maillage.

Après une étude complète, le maillage retenu comporte $N^{(1)} = 1911$ nœuds dans le processeur et $N^{(2)} = 8561$ nœuds dans le radiateur. La dimension du modèle détaillé MD_B est donc $N = N^{(1)} + N^{(2)} = 10472$.

II.3. Recherche d'un modèle adapté : modèle MD_C

Le modèle MD_C utilise le maillage du modèle MD_B , mais les paramètres des solveurs sont ajustés de manière à obtenir le temps de calcul le plus faible possible. Il est imposé au modèle MD_C de vérifier la condition suivante : $\epsilon_{BC,max} < 0.01 * \Delta T_{max}$ (voir 4.2 et 4.4). MD_C est aussi recherché de manière itérative.

Les paramètres obtenus pour le modèle MD_C sont présentés dans le tableau 4.3.

Paramètre	Unité	Valeur
Nombre de noeuds	10472	
ϵ_{GC}	-	10^{-8}
ϵ_{max}	K	0.006
f_{max}	-	2
ϵ_{min}	K	0.003
f_{min}	-	1.2

TAB. 4.3 – Paramètres du modèle MD_C

II.4. Bilan sur les différents modèles détaillés

Les différents modèles sont résumés par cette figure :

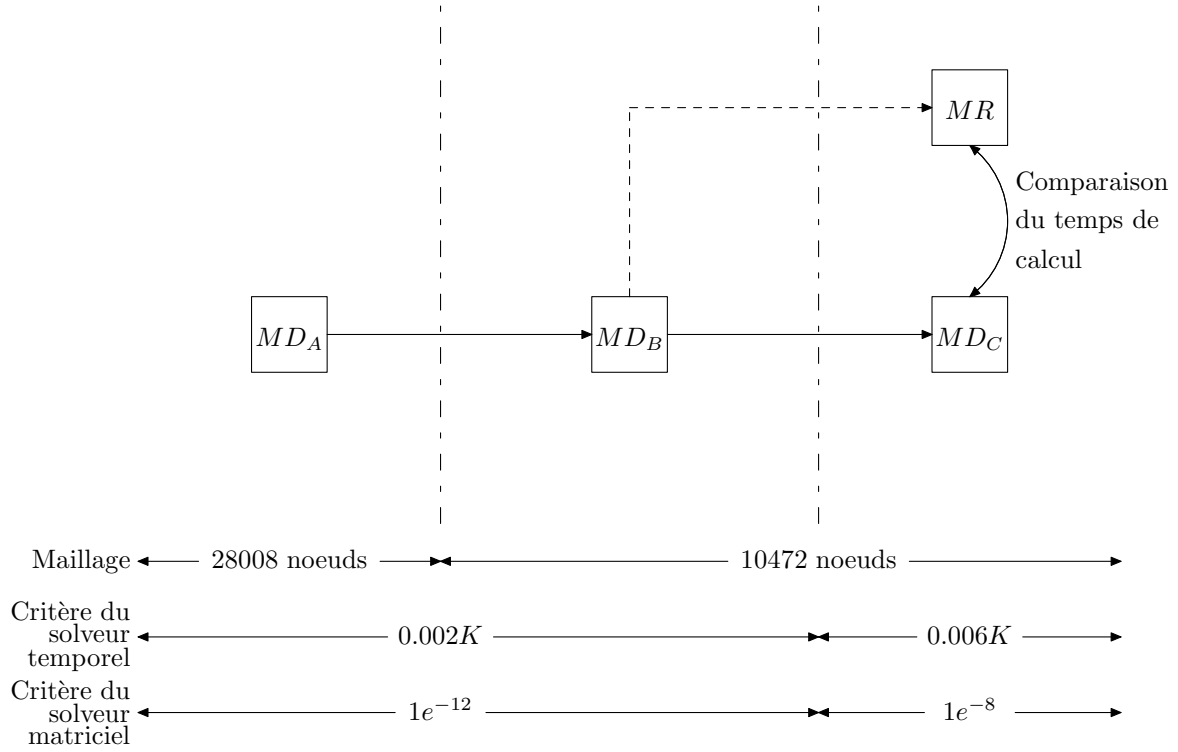


FIG. 4.9 – Bilan sur les modèles détaillés de l'étude 2D

Les temps de calculs des différents modèles sont : $T_{CPU,MD_A} = 332317s$, $T_{CPU,MD_B} = 48673s$ et $T_{CPU,MD_C} = 16408s$.

L'objectif est désormais de trouver un modèle réduit (MR) qui présente un écart $\epsilon_{BR,max}$ inférieur à 1% de ΔT_{max} avec un temps de calcul très inférieur à T_{CPU,MD_C} . L'écart $\epsilon_{BR,max}$ étant défini par :

$$\epsilon_{BR,max} = \max_t (\max_{\Omega} (abs(T_{MR} - T_{MD_B}))) \quad (4.11)$$

III. Recherche des bases modales locales

Une base de branche locale est formée par l'ensemble des solutions du problème de branche (équations E.3 et E.4). L'écriture d'une formulation variationnelle (E.14) du problème de branche permet d'utiliser les éléments finis pour calculer les modes. Le problème de branche devient alors un problème matriciel aux valeurs propres généralisé (E.52). Ce problème matriciel est alors résolu en utilisant le code ARPACK [47] suivant la démarche présentée en figure E.2.

Il y a deux bases de branche locales : la base de branche du processeur et la base de branche du radiateur.

Les nombres de Steklov (ζ) utilisés pour le calcul des bases de branche locales correspondent aux choix par défaut donnés par E.22.

Les maillages utilisés pour le calcul des modes sont formés à partir du maillage utilisé pour les modèles MD_B et MD_C . Le maillage du processeur comporte donc $N^{(1)} = 1911$ nœuds tandis que celui du radiateur comporte $N^{(2)} = 8561$ nœuds. Les bases de branche discrètes comporteront donc $N^{(1)}$ modes pour le processeur et $N^{(2)}$ modes pour le radiateur.

Les modes les plus importants sont usuellement ceux de constante de temps les plus élevées. Cette règle est approximative, elle est cependant acceptable la plupart du temps car les exceptions sont relativement peu fréquentes : sources impulsives, ... Il semble alors intéressant de ne calculer que les $\tilde{N}_0^{(q)}$ premiers modes de branche (triés par ordre de constante de temps décroissante), ce qui revient à pratiquer une troncature de Marshall [31].

Le code ARPACK servant à calculer les modes permet justement d'effectuer une telle troncature en calculant les modes dans l'ordre décroissant des constantes de temps.

Il n'existe pas de critère *a priori* permettant d'assurer que le nombre de modes calculé est suffisant pour atteindre une précision donnée. Cependant dans le cas présent, une solution de référence (modèle MD_B) est connue.

Il est alors possible d'étudier *a posteriori* l'erreur de décomposition par projection³ du champ T_{MD_B} dans la base. Ce point est traité dans les deux pages suivantes.

La décomposabilité dans la base tronquée est uniquement « un indicateur ». En effet l'écart obtenu par simulation est différent de l'écart obtenu par projection au sens de l'opérateur C_b :

- L'écart obtenu par simulation peut être plus faible que l'écart obtenu par projection, en effet la projection au sens de l'opérateur C_b ne donne pas les états qui minimiseraient l'écart.
- L'inverse est aussi possible car les données utilisées sont celles spécifiées par l'utilisateur (voir paramètre N_{sauv} en page 50). Le champ de température sauvegardé donne une vision globale du système physique au cours du temps, mais les transitoires rapides sont potentiellement écartés. Si ces transitoires sont mal représentés par la base modale alors un biais sera introduit lors de la simulation.

³au sens de l'opérateur C_b , voir E.44

III.1. Cas du processeur

Les $\tilde{N}_0^{(1)} = 1750$ premiers modes sont calculés pour le processeur. La matrice des températures sauvegardées pour ce domaine est notée $\mathbb{T}_{MD_B}^{(1)}$, elle comporte $N^{(1)}$ lignes et N_{sauv} colonnes. Les états de la décomposition modale tronquée (car $\tilde{N}_0^{(1)} < N^{(1)}$) sont déterminés grâce à l'opérateur $\mathbb{C}_b^{(1)}$ discrétisé (opérateur matriciel $\mathbb{C}_b^{(1)}$) :

$$\tilde{\mathbb{X}}^{(1)} = {}^t \tilde{\mathbb{V}}^{(1)} \mathbb{C}_b^{(1)} \mathbb{T}_{MD_B}^{(1)} \quad (4.12)$$

Où $\tilde{\mathbb{V}}^{(1)}$ est la matrice des modes tronqués du processeur ($\Omega^{(1)}$), elle comporte $N^{(1)}$ lignes et $\tilde{N}_0^{(1)}$ colonnes.

La matrice des températures obtenue avec les n premiers modes est alors telle que :

$$\tilde{\mathbb{T}}_n^{(1)}(i, j) = \sum_{k=1}^n \tilde{\mathbb{V}}^{(1)}(i, k) * \tilde{\mathbb{X}}^{(1)}(k, j) \quad (4.13)$$

L'écart maximum entre $\tilde{\mathbb{T}}_n^{(1)}$ et $\mathbb{T}_{MD_B}^{(1)}$ est $\epsilon_{MD_B, n}^{(1)}$:

$$\epsilon_{MD_B, n}^{(1)} = \text{max}_{\text{ligne}} \left(\text{max}_{\text{colonne}} \left(\text{abs} \left(\tilde{\mathbb{T}}_n^{(1)} - \mathbb{T}_{MD_B}^{(1)} \right) \right) \right) \quad (4.14)$$

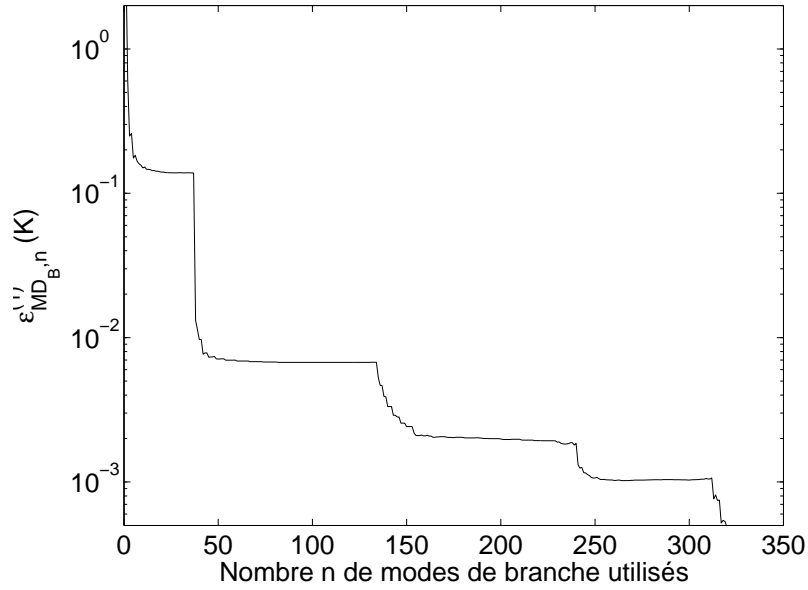


FIG. 4.10 – Erreur de projection pour le processeur

Au sens de l'opérateur $\mathbb{C}_b^{(1)}$ et dans le cas du processeur, il apparaît que les 50 premiers modes permettent d'assurer une très bonne décomposabilité du champ thermique de référence ($\mathbb{T}_{MD_B}^{(1)}$). $\epsilon_{MD_B, n}$ est inférieur à $10^{-3}K$ dès que $n > 312$.

III.2. Cas du radiateur

Les $\tilde{N}_0^{(2)} = 5000$ premiers modes sont calculés pour le radiateur. La matrice des températures sauvegardées pour ce domaine est notée $\mathbb{T}_{MD_B}^{(2)}$, elle comporte $N^{(2)}$ lignes et N_{sauv} colonnes. Les états de la décomposition modale tronquée (car $\tilde{N}_0^{(2)} < N^{(2)}$) sont déterminés grâce à l'opérateur $\mathcal{C}_b^{(2)}$ discrétisé (opérateur matriciel $\mathbb{C}_b^{(2)}$) :

$$\tilde{\mathbb{X}}^{(2)} = {}^t \tilde{\mathbb{V}}^{(2)} \mathbb{C}_b^{(2)} \mathbb{T}_{MD_B}^{(2)} \quad (4.15)$$

Où $\tilde{\mathbb{V}}^{(2)}$ est la matrice des modes tronqués du radiateur ($\Omega^{(2)}$), elle comporte $N^{(2)}$ lignes et $\tilde{N}_0^{(2)}$ colonnes.

La matrice des températures obtenue avec les n premiers modes est alors telle que :

$$\tilde{\mathbb{T}}_n^{(2)}(i, j) = \sum_{k=1}^n \tilde{\mathbb{V}}^{(2)}(i, k) * \tilde{\mathbb{X}}^{(2)}(k, j) \quad (4.16)$$

L'écart maximum entre $\tilde{\mathbb{T}}_n^{(2)}$ et $\mathbb{T}_{MD_B}^{(2)}$ est $\epsilon_{MD_B, n}^{(2)}$:

$$\epsilon_{MD_B, n}^{(2)} = \mathit{max}_{ligne} \left(\mathit{max}_{colonne} \left(\mathit{abs} \left(\tilde{\mathbb{T}}_n^{(2)} - \mathbb{T}_{MD_B}^{(2)} \right) \right) \right) \quad (4.17)$$

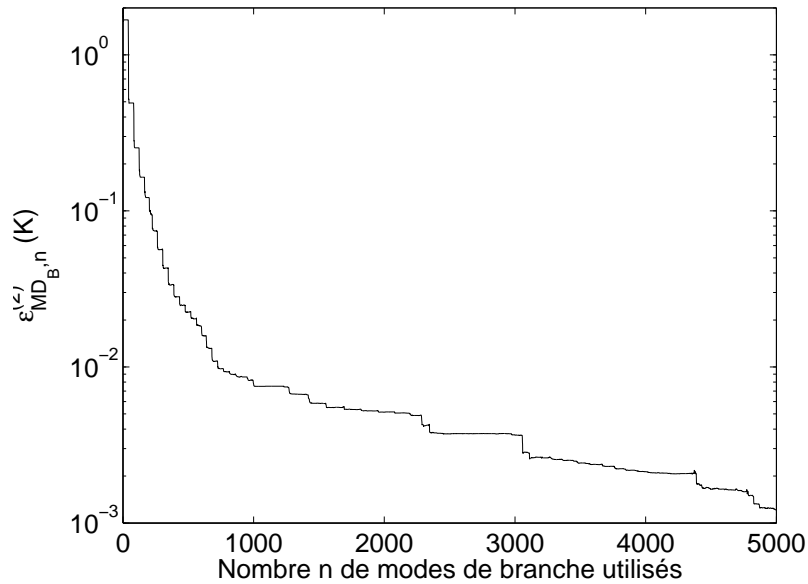


FIG. 4.11 – Erreur de projection pour le radiateur

Au sens de l'opérateur $\mathcal{C}_b^{(2)}$ et dans le cas du radiateur, il apparaît qu'il faut considérer plus de modes de branche que pour le processeur, 1500 modes seront nécessaires pour assurer une bonne décomposabilité du champ thermique de référence ($\mathbb{T}_{MD_B}^{(2)}$). On obtient $\epsilon_{MD_B, 5000} = 1.2 * 10^{-3} K$.

Les bases modales qui ont été déterminées (pour le processeur et pour le radiateur) peuvent donc décomposer les champs obtenus par le modèle MD_B . Il paraît alors raisonnable de considérer qu'elles sont suffisantes pour l'étude qui va suivre.

IV. Réduction de la base de branche structurée et étude *a posteriori* du paramètre β de la fonctionnelle de saut de flux

Les bases de branche locales ont été déterminées, il est théoriquement possible de les utiliser directement dans un modèle réduit.

En pratique le nombre de modes doit rester faible (compris entre 3 et 500), à cause de l'augmentation du temps de calcul (et de la mémoire nécessaire) avec le nombre de modes. Il faut donc réduire les bases modales locales.

Parmi les différentes méthodes de réduction, la méthode de l'amalgame modal découplé⁴ a montré sa performance. Elle est basée sur un problème linéaire et elle néglige les couplages entre modes. Avec ces approximations, la procédure d'amalgame s'effectue en quelques secondes.

Dans des applications mono-corps, il est relativement facile de choisir le cas test d'amalgame. Une difficulté supplémentaire apparaît en multi-corps pour le choix de la condition relative à l'interface de couplage.

Mais les résultats de la méthode modale proposée ne dépendent pas uniquement des bases modales, ils dépendent aussi du paramètre β de la fonctionnelle de saut de flux.

Une étude à trois jeux de paramètres est ainsi menée :

- les paramètres d'amalgame du processeur.
- les paramètres d'amalgame du radiateur.
- le paramètre β de la fonctionnelle de saut de flux.

La suite de cette section va commencer par la présentation du solveur modal. Ensuite les cas test envisagés seront présentés, l'expérimentation numérique sera alors menée. La section se terminera par une brève conclusion sur les résultats de l'analyse *a posteriori* des résultats.

⁴L'amalgame modal découplé a fait l'objet de l'article [43] et est présenté dans l'annexe F

IV.1. Présentation du solveur modal

Le solveur modal n'est pas fondamentalement différent du solveur éléments finis. Cependant il est nécessaire beaucoup plus de lignes de code c++ que le solveur éléments finis. Le nombre de lignes de code passe en effet de 800 lignes à 2200 lignes. Ceci est parfaitement compréhensible si l'on rappelle que le solveur éléments finis est contenu dans le solveur modal et qu'il faut un grand nombre de compléments : passage de l'espace des températures à l'espace modal (et vice-versa), critères des solveurs plus complexes, écriture de la fonctionnelle de saut de flux.

Le vecteur d'état initial est créé par projection du vecteur des températures dans l'espace modal, au sens de \mathcal{C}_b (voir équation F.59) dans chaque sous domaine.

L'équation 3.48 est traduite sous forme matricielle. Le système matriciel est carré de dimension $\tilde{N} = \tilde{N}_1^{(1)} + \tilde{N}_1^{(2)}$, avec $\tilde{N}_1^{(1)}$ le nombre de modes réduits du processeur et $\tilde{N}_1^{(2)}$ le nombre de modes réduits du radiateur. Dans la suite de l'étude \tilde{N} varie de 20 à 160.

Le système matriciel est donc de petite taille comparativement au système matriciel obtenu par la méthode des éléments finis (dimension $N = 10472$). Il est alors résolu par une méthode adaptée aux problèmes matriciels de petite taille : la méthode LD^tL [29].

Le choix des critères dans l'espace physique est relativement naturel, il est par contre problématique dans l'espace modal. Différents critères ont été codés, le critère qui donne le meilleur compromis précision/temps de calcul est le suivant :

- l'utilisateur définit un certain nombre de points d'observation M_j dans le domaine physique Ω .
- à chaque pas de temps, la température de ces points T_{M_j} est calculée à partir des états.
- un calcul partiel d'écart de température entre deux pas de temps est effectué à partir des T_{M_j} .
- les critères du modèle MD_B sont alors appliqués.

Le critère du solveur modal est donc situé dans l'espace physique des températures. Il est partiel car la reconstruction du champ de température sur l'ensemble du domaine à chaque pas de temps nécessiterait trop de temps de calcul.

La matrice des températures est créée à la fin du calcul par recombinaison modale.

Remarques sur le schéma général de la page suivante :

- Le vecteur d'état initial est créé à l'aide la formule E.43 (voir page 134).
- La matrice des températures sauvegardées (\mathbf{T}) est construite après le solveur temporel par recombinaison modale (grâce à la matrice des états sauvegardés (\mathbf{X})).

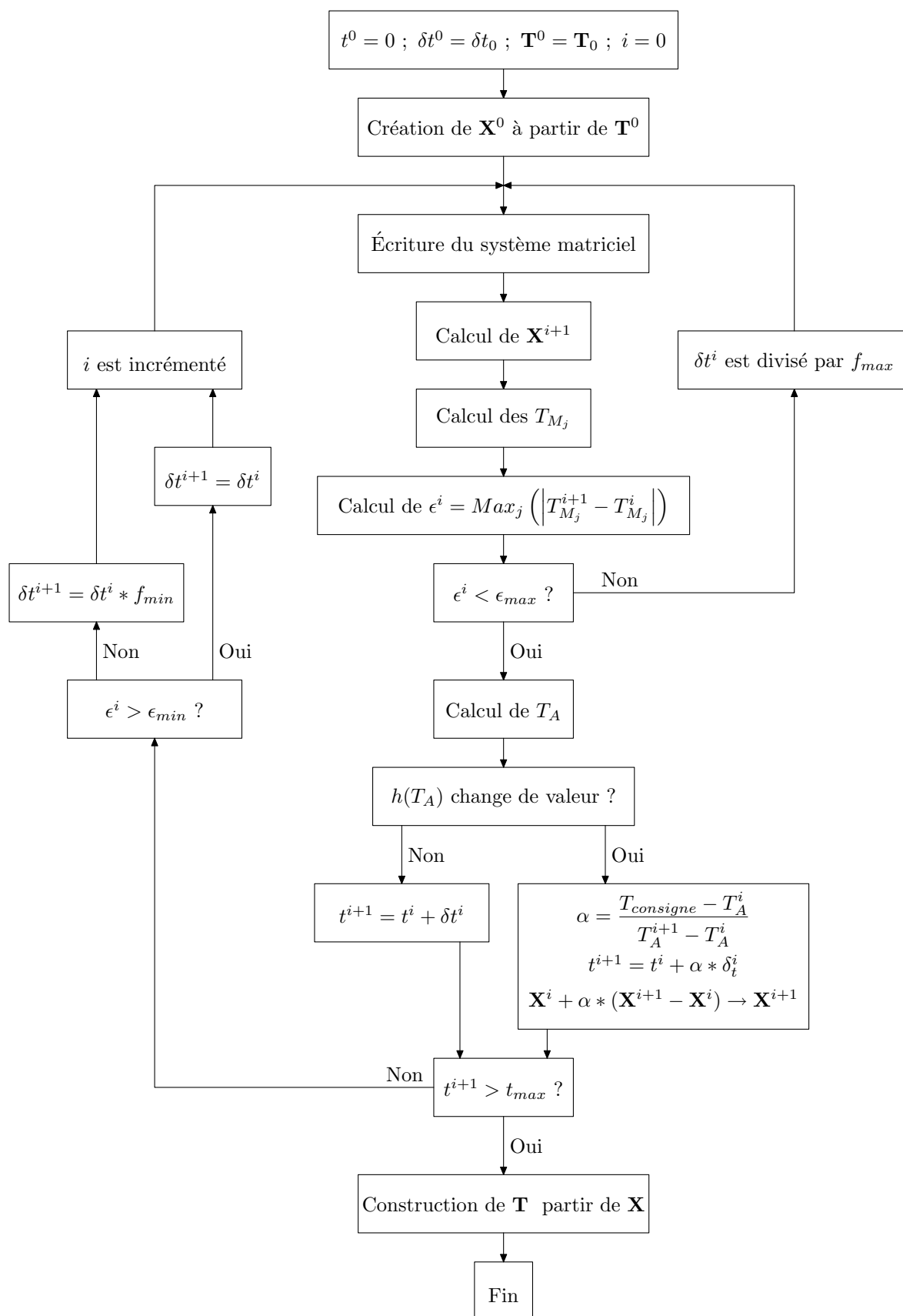


FIG. 4.12 – Schéma général du pas de temps variable pour les modèles réduits

IV.2. Présentation des cas tests

Les bases de branche sont réduites par corps avant d'être étendues sur le domaine complet. La procédure d'amalgame est basée sur un cas test (un par corps) qu'il faut définir. La méthode implique que chaque cas test soit aussi proche que possible du problème réel, tout en étant linéaire et indépendant du cas test de l'autre corps. Les cas tests ne peuvent donc pas reproduire la spécificité du couplage entre deux domaines. On choisit alors de tester différents choix à l'aide de simulations numériques. La condition de contact est remplacée par une condition convective, mais avec un coefficient de convection très grand (par rapport aux coefficients de convection habituels). L'influence réelle de cette condition ne pouvant être déterminée que numériquement, plusieurs cas tests sont envisagés et une étude numérique est ensuite menée pour chaque cas test.

IV.2.1. Cas tests du processeur

Le cas test du processeur est le suivant :

$$\forall M \in \Omega^{(1)} \quad c^{(1)} \frac{\partial T^{(1)}}{\partial t} = \text{div} \left(k^{(1)} \nabla \left(T^{(1)} \right) \right) + \pi^{(1)} \quad (4.18)$$

$$\forall M \in \Gamma_{12} \quad -k^{(1)} \nabla \left(T^{(1)} \right) \bullet \underline{n}^{(1)} = h_{eq1} \left(T^{(1)} - T_{eq1} \right) \quad (4.19)$$

$$\forall M \in \partial\Omega^{(1)} \setminus \Gamma_{12} \quad -k^{(1)} \nabla \left(T^{(1)} \right) \bullet \underline{n}^{(1)} = 0 \quad (4.20)$$

$$\forall M \in \Omega^{(1)} \quad T^{(1)}(M, 0) = T_0(M) \quad (4.21)$$

Les valeurs de $c^{(1)}$, $k^{(1)}$, $\pi^{(1)}$ et de $T_0^{(1)}$ sont celles du problème physique. Le plan de symétrie et la frontière adiabatique sont prises en compte par un flux nul. La condition de contact avec le radiateur est remplacée par une condition de convection (de coefficient h_{eq1} fixe) avec un fluide dont la température T_{eq1} est fixe.

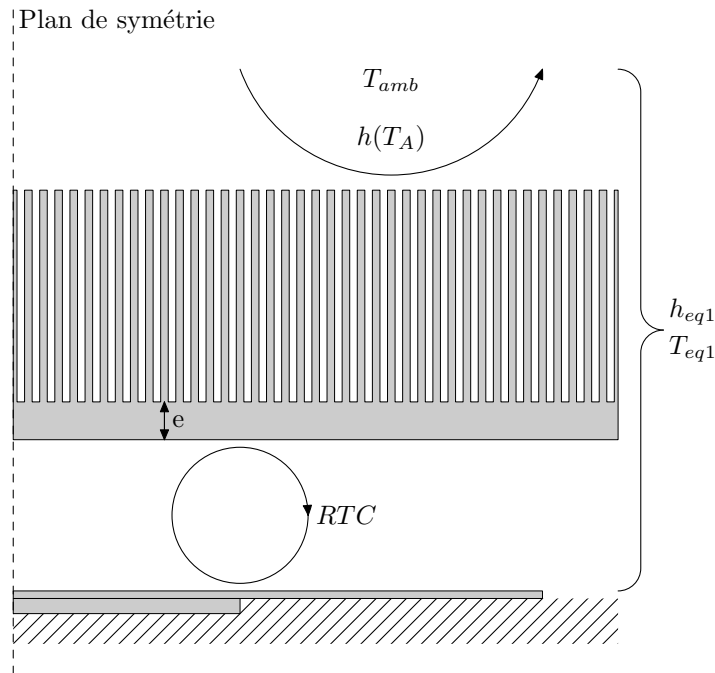


FIG. 4.13 – Cas test proposé pour le processeur

La condition de convection est censée être représentative du radiateur. La température T_{eq1} est alors choisie égale à celle de l'ambiance T_{amb} . Il reste alors le choix du coefficient de convection.

Un ordre de grandeur pour h_{eq1} est donné par l'analogie des résistances thermiques en série :

$$\frac{1}{h_{eq1}} = R_{eq1} = RTC + R_e + R_a \quad (4.22)$$

Où :

- RTC désigne la résistance thermique de contact
- R_e désigne la résistance apportée par l'épaisseur e du matériau qui est de l'ordre de grandeur de $\frac{e}{k}$.
- R_a désigne la résistance apportée par la convection thermique avec l'effet d'ailette.
 $R_a = \frac{S_{\Gamma_2}}{S_{\Gamma_{12}}} \frac{1}{h_{max}}$ en considérant une efficacité d'ailette unitaire et le coefficient de convection maximal.

Ainsi un ordre de grandeur de h_{eq1} est :

$$h_{eq1} = \frac{1}{RTC + \frac{e}{k} + \frac{S_{\Gamma_2}}{S_{\Gamma_{12}}} \frac{1}{h_{max}}} \approx 1000W.m^{-2}.K^{-1} \quad (4.23)$$

Si le choix $h_{eq1} = 1000W.m^{-2}.K^{-1}$ semble raisonnable, il paraît intéressant de tester le choix suivant :
 $h_{eq1} = RTC$ ($40000W.m^{-2}.K^{-1}$). Enfin une troisième valeur intermédiaire est arbitrairement choisie :
 $h_{eq1} = 10000W.m^{-2}.K^{-1}$

IV.2.2. Cas tests du radiateur

La démarche est liée à celle du processeur, le cas test du radiateur est alors :

$$\forall M \in \Omega^{(2)} \quad c^{(2)} \frac{\partial T^{(2)}}{\partial t} = \text{div} \left(k^{(2)} \underline{\nabla} \left(T^{(2)} \right) \right) \quad (4.24)$$

$$\forall M \in \Gamma_{12} \quad -k^{(2)} \underline{\nabla} \left(T^{(2)} \right) \bullet \underline{n}^{(2)} = h_{eq2} \left(T^{(2)} - T_{eq2} \right) \quad (4.25)$$

$$\forall M \in \Gamma_2 \quad -k^{(2)} \underline{\nabla} \left(T^{(2)} \right) \bullet \underline{n}^{(2)} = h_{amb} \left(T^{(2)} - T_{amb} \right) \quad (4.26)$$

$$\forall M \in \partial\Omega^{(2)} \setminus (\Gamma_{12} \cup \Gamma_2) \quad -k^{(2)} \underline{\nabla} \left(T^{(2)} \right) \bullet \underline{n}^{(2)} = 0 \quad (4.27)$$

$$\forall M \in \Omega^{(2)} \quad T^{(2)}(M, 0) = T_0(M) \quad (4.28)$$

Les valeurs de $c^{(2)}$, $k^{(2)}$ et de $T_0^{(2)}$ sont celles du problème physique. Le plan de symétrie et la frontière adiabatique sont prises en compte par un flux nul. La condition sur la frontière convective est rendue fixe (l'hystérésis est supprimé), le coefficient de convection est alors $h_{amb} = 60W.m^{-2}.K^{-1}$.

La condition de contact avec le processeur est remplacée par une condition de convection (de coefficient h_{eq2} fixe) avec un fluide dont la température T_{eq2} est fixe. Les valeurs de h_{eq2} qui seront considérées sont celles prises pour h_{eq1} . De manière arbitraire T_{eq2} prend la valeur suivante : $T_{eq2} = 40^\circ C$

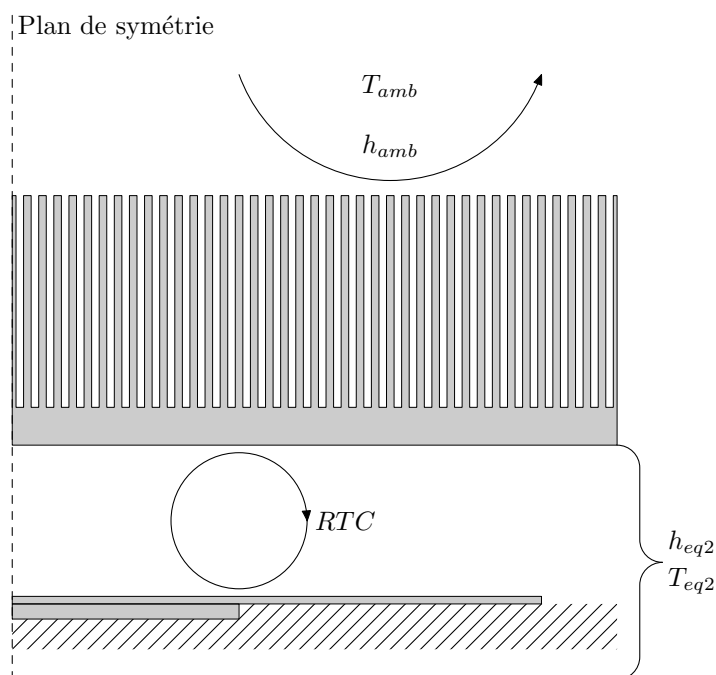


FIG. 4.14 – Cas test proposé pour le radiateur

IV.2.3. Les trois jeux de cas tests

Finaleme nt trois cas test vont être analysés, leurs paramètres respectifs sont résumés dans le tableau suivant :

Cas test	h_{eq1}	h_{eq2}	T_{eq1}	T_{eq2}
	$W.m^{-2}.K^{-1}$	$W.m^{-2}.K^{-1}$	$^{\circ}C$	$^{\circ}C$
1	1000	1000	20	40
2	10000	10000	20	40
3	40000	40000	20	40

TAB. 4.4 – Paramètres des cas tests étudiés

IV.3. Effet du cas test

Les cas tests sont désormais figés, il faut désormais effectuer des simulations numériques pour déterminer l'effet du cas test sur les résultats obtenus en fonction du nombre de modes conservés et de la valeur du coefficient β de la fonctionnelle de saut de flux.

Dans un soucis de concision, les résultats vont être présentés pour deux ordres de réduction : $\tilde{N} = 20$ et $\tilde{N} = 160$. Dans chaque cas, le nombre de modes réduits est le même pour chaque sous-domaine (donc $\tilde{N}_1^{(1)} = \tilde{N}_1^{(2)} = 10$ ou $\tilde{N}_1^{(1)} = \tilde{N}_1^{(2)} = 80$).

IV.3.1. Résultats avec $\tilde{N} = 20$ ($\tilde{N}_1^{(1)} = \tilde{N}_1^{(2)} = 10$)

L'écart⁵ entre le modèle détaillé MD_B et le modèle réduit MR est noté $\epsilon_{BR,max}$:

$$\epsilon_{BR,max} = \max_t (\max_{\Omega} (\text{abs}(T_{MR} - T_{MD_B}))) \quad (4.29)$$

Les résultats sont présentés en pourcentage de l'évolution thermique maximale $\Delta T_{max} = 36.45K$ (voir equation 4.2 et figure 4.8).

Les différentes simulations montrent que $\epsilon_{BR,max}$ ne varie pas de manière significative lorsque β varie de 0 à 10^{-7} . Une échelle logarithmique en β est alors utilisée sur l'intervalle $[10^{-7}, 10^{-3}]$:

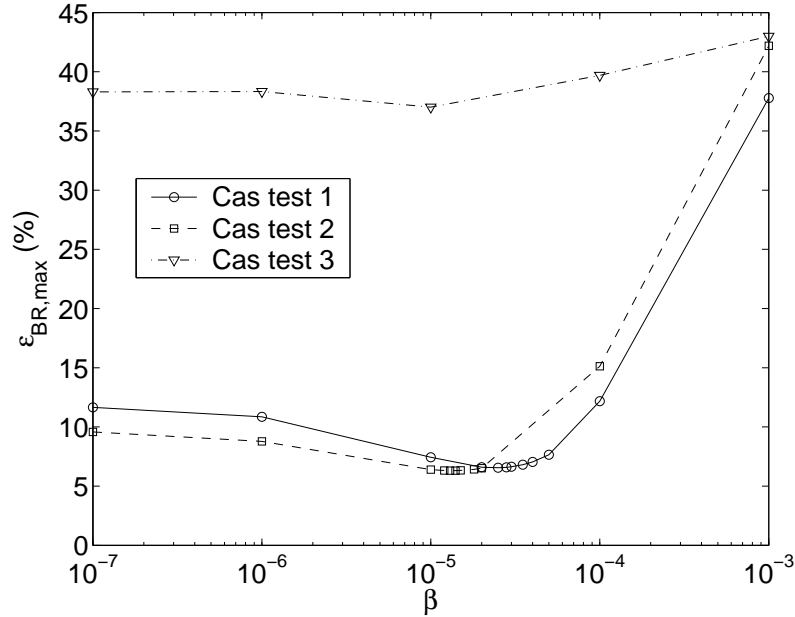


FIG. 4.15 – Effet du cas test avec $\tilde{N} = 20$

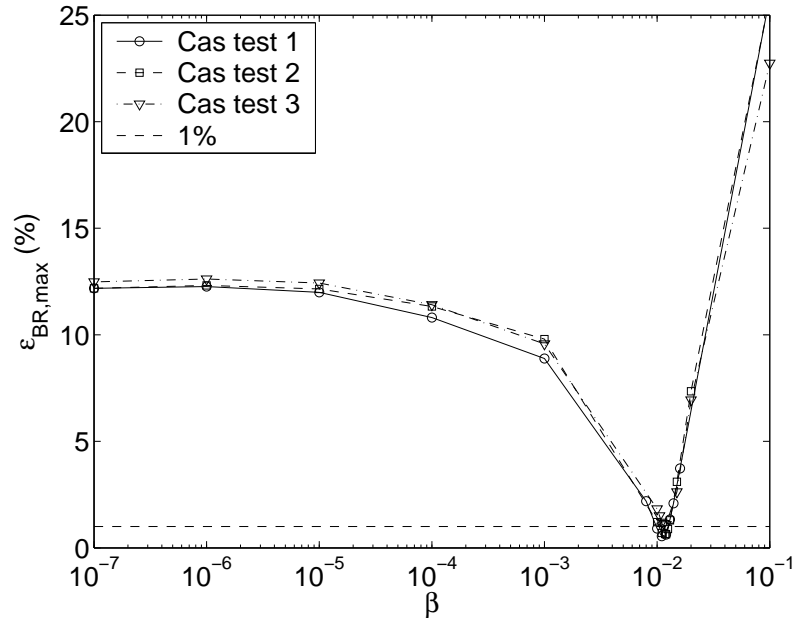
L'analyse des résultats du graphique 4.15 montre que :

- Les cas tests 1 et 2 donnent des résultats voisins qui sont meilleurs que ceux du cas test 3.
- Le cas test 3 ne permet pas d'obtenir un écart inférieur à 35% alors qu'avec les deux autres cas test un écart inférieur à 10% est obtenu. Le choix du cas test est donc important si $\tilde{N} = 20$.
- L'introduction d'une fonctionnelle de saut de flux permet effectivement d'améliorer les résultats. Il existe un optimum au paramètre β et une valeur trop importante de ce paramètre dégrade fortement les résultats.
- Le minimum obtenu pour $\epsilon_{BR,max}$ est sensiblement le même pour chacun des cas tests 1 et 2, cependant la valeur optimale de β est fonction du cas test.
- Pour $\tilde{N} = 20$, l'introduction de la fonctionnelle de saut de flux ne permet pas de réduire l'écart d'un facteur supérieur à 2.
- Le temps nécessaire à la simulation d'un modèle réduit est dans tous les cas supérieur à 23.4s et inférieur à 24.3s. Il est donc très faiblement dépendant du cas test et de la valeur de β .

⁵Il s'agit ici d'un rappel de la définition 4.11 donnée en page 57

IV.3.2. Résultats avec $\tilde{N} = 160$ ($\tilde{N}_1^{(1)} = \tilde{N}_1^{(2)} = 80$)

L'ordre du modèle réduit est augmenté, il passe de 20 à 160. Comme dans le cas où $\tilde{N} = 20$, $\epsilon_{BR,max}$ ne varie pas de manière significative lorsque β varie de 0 à 10^{-7} . Une échelle logarithmique en β est alors utilisée sur l'intervalle $[10^{-7}, 10^{-1}]$:


 FIG. 4.16 – Effet du cas test avec $\tilde{N} = 160$

L'analyse des résultats du graphique 4.16 montre que :

- Les résultats sont peu influencés par le cas test choisi.
- L'augmentation de la taille du modèle réduit de 20 à 160 n'améliore pas les résultats s'il n'y a pas de fonctionnelle de saut de flux.
- L'introduction de la fonctionnelle de saut de flux permet d'améliorer de manière importante les résultats. Pour chacun des cas tests $\epsilon_{BR,max}$ passe sous la barre des 1% (ce qui correspond au cas du modèle détaillé MD_C).
- Le temps nécessaire à la simulation d'un modèle réduit est, de nouveau, très faiblement dépendant du cas test et de la valeur de β , dans tous les cas il est supérieur à 198.1s et inférieur à 200s.
- Avec un choix optimal de β , il est possible d'obtenir une précision meilleure que celle du modèle détaillé MD_C avec un temps de simulation 82 fois plus faible (rappel : $T_{CPU,MD_C} = 16408s$ voir page 57).

IV.4. Conclusion sur la réduction de la base

De très nombreuses simulations numériques ont été menées. Les résultats des deux pages précédentes sont confirmés :

- L'augmentation de l'ordre du modèle réduit ne permet pas de réduire l'écart $\epsilon_{BR,max}$ s'il n'y a pas de fonctionnelle de saut de flux.
- Le temps de simulation du modèle réduit est très faiblement modifié par l'introduction d'une fonctionnelle de saut de flux ou par un changement de cas test.
- Lorsque l'ordre du modèle réduit est suffisamment important pour obtenir un résultat précis, le choix du cas test n'est pas primordial. Pour un modèle réduit d'ordre faible il est préférable d'éviter le cas test 3 qui correspond à un cas test extrême.
- Il existe une valeur optimale pour le paramètre β , cette valeur dépend de l'ordre de réduction et dépend du cas test. Dans certains cas, l'écart entre le modèle détaillé et le modèle réduit ($\epsilon_{BR,max}$) augmente avant de diminuer (voir graphique 4.17 page 71).
- La précision du modèle détaillé MD_C peut être obtenue avec un temps de calcul 82 fois plus faible.

La réduction de la base de branche structurée ne présente donc pas de grande difficulté. L'introduction d'une fonctionnelle de saut de flux est pleinement justifiée par les résultats obtenus. Il reste cependant une difficulté de taille : le choix de la valeur du paramètre β .

V. Choix *a priori* du paramètre β de la fonctionnelle de saut de flux

L'étude *a posteriori* des résultats montre qu'il existe une valeur optimale du paramètre β . Cette valeur est fonction du cas test utilisé pour l'amalgame et de l'ordre du modèle réduit. Aucune loi permettant de choisir la valeur de β ne semble se dégager à partir des données obtenues en étudiant l'écart $\epsilon_{BR,max}$. Ainsi, pour l'instant, il faut connaître une solution de référence pour pouvoir déterminer la valeur optimale du paramètre β .

Or l'utilisateur ne désire pas effectuer de longs calculs pour déterminer la valeur optimale de β , il désire déterminer rapidement le champ thermique au cours du temps à l'aide de la méthode modale. A défaut, un « indicateur » pour la détermination du paramètre β a donc été recherché. Le seul indicateur ayant donné satisfaction pour tous les modèles réduits (ordre et cas test) est celui qui va être présenté.

Une famille de modèles réduits est construite à partir d'un même modèle réduit qui ne comporte pas de fonctionnelle de saut de flux. Les différents modèles réduits ne diffèrent que par la valeur du paramètre β de la fonctionnelle de saut de flux. A chaque modèle modal (repéré par la valeur du paramètre β puisqu'il n'y a pas d'ambiguïté) est associé un nombre $\gamma(\beta)$. Ce nombre est défini par :

$$\gamma(\beta) = \max_n \left(\frac{\|\mathbf{X}^{n+1} - \mathbf{X}^n\|_2}{\delta t^n} \right) \quad (4.30)$$

La détermination de γ pour une valeur de β nécessite donc de simuler le modèle modal correspondant, puis de déterminer pour ce modèle le maximum, sur les n itérations du calcul, du ratio entre :

- la norme de l'évolution du vecteur d'état lors de l'itération n : $\|\mathbf{X}^{n+1} - \mathbf{X}^n\|_2$
- le pas de temps δt^n utilisé pour l'itération n .

La détermination de γ pour une valeur de β s'effectue donc sans la connaissance de la solution de référence.

Il est possible de tracer la courbe de γ en fonction de β , le point intéressant est le suivant :

La fonction $\gamma(\beta)$ présente un minimum pour une valeur de β proche de celle qui minimise l'écart $\epsilon_{BR,max}$. Il est ainsi possible **d'indiquer** le meilleur modèle réduit (*i.e* celui qui minimise $\epsilon_{BR,max}$) d'une famille de modèles réduits en traçant la courbe $\gamma(\beta)$ pour cette famille.

La valeur optimale du paramètre β n'est donc pas prédite à l'avance, mais elle est **indiquée** grâce à l'étude de la fonction $\gamma(\beta)$.

En pratique, la courbe de $\gamma(\beta)$ est relativement plate et ne varie que faiblement, ainsi la courbe tracée sera celle de la fonction « normalisée (*i.e* qui varie entre 0 et 1) » $\gamma^*(\beta)$ définie par :

$$\gamma^*(\beta) = \frac{\gamma(\beta) - \min_{\beta}(\gamma(\beta))}{\max_{\beta}(\gamma(\beta)) - \min_{\beta}(\gamma(\beta))} \quad (4.31)$$

Le graphique 4.17 est établi *a posteriori* avec la connaissance d'une solution de référence. Au contraire le graphique 4.18 est établi sans l'usage d'une solution de référence.

Pour une même réduction, le minimum de la courbe de $\gamma^*(\beta)$ indique bien le minimum de $\epsilon_{BR,max}$.

VI. Conclusion sur la méthode de couplage en 2D

La démarche de couplage proposée permet d'obtenir un bon compromis précision/temps de calcul. Si la précision recherchée est faible alors un modèle réduit utilisant un faible nombre de modes sans fonctionnelle de saut de flux est utilisable. L'augmentation du nombre de modes utilisés n'augmente pas la précision des résultats en l'absence d'une fonctionnelle de saut de flux. L'introduction de la fonctionnelle de saut de flux avec une valeur adaptée du paramètre β améliore la précision les résultats sans augmenter le temps de calcul. L'étude de la fonction γ^* permet d'indiquer la valeur optimale de β .

Note complémentaire sur les temps de calcul :

- Le calcul de la base de branche est assez long mais est effectué une fois pour toutes, le temps de calcul pour l'exemple 2D est de 891s pour le radiateur, 59s pour le processeur. La démarche sera ainsi plus intéressante si on doit effectuer une série importante de simulations.
- La procédure de l'amalgame modal découplé est très rapide de l'ordre de la seconde.
- La création de la matrice des températures à partir de la matrice des états dépend du nombre d'instantanés sauvegardés. Dans le cas présent avec 1000 sauvegardes, il faut environ 3s.

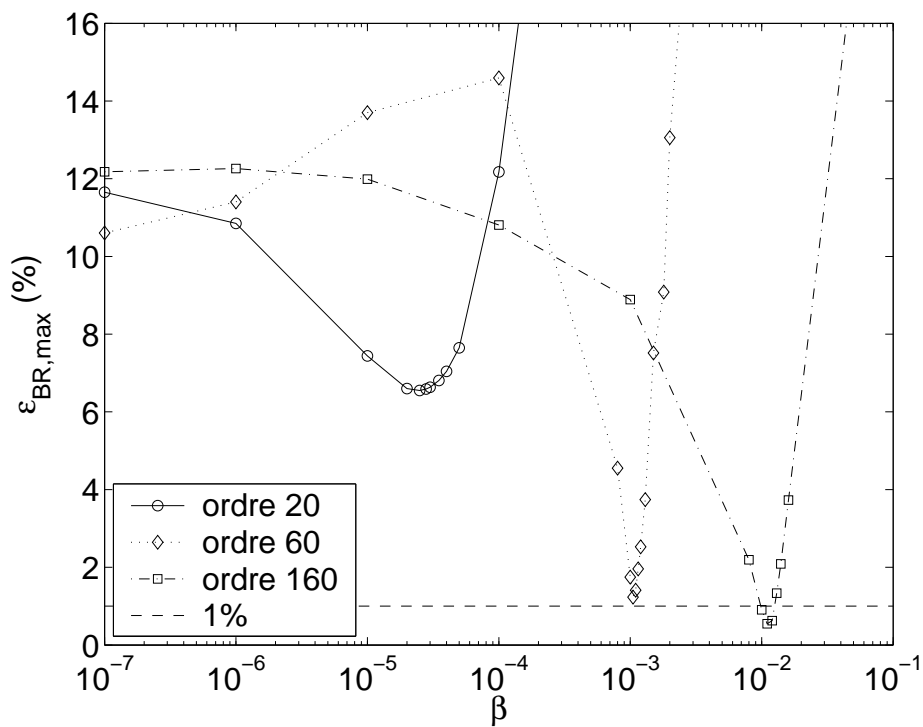


FIG. 4.17 – Résultats pour $\tilde{N} = 20, 60, 160$ avec le cas test 1

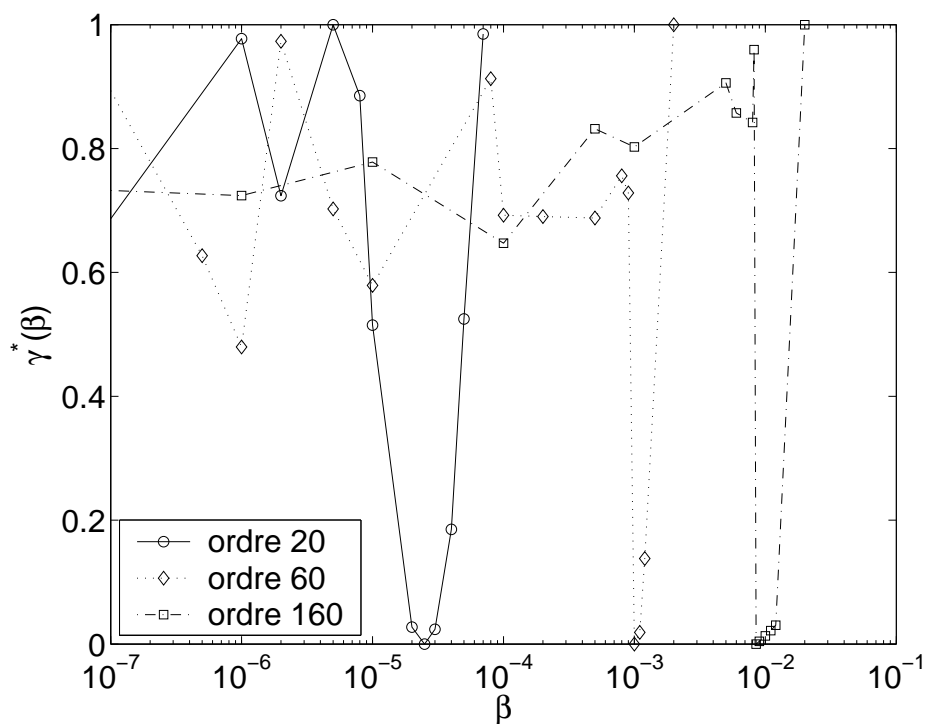


FIG. 4.18 – Tracé de la fonction γ^* pour $\tilde{N} = 20, 60, 160$ avec le cas test 1

Pratique du couplage par conduction en 3D

« On calcule aujourd'hui la même chose qu'hier, mais avec une plus grande précision et plus vite, et des choses qu'on envisageait même pas de calculer.

Ce que vous êtes capables de calculer aujourd'hui à grand-peine fera gentiment sourire vos enfants. »

J.C. Craveur

Contenu du chapitre

I.	Présentation des résultats de l'article [49]	74
II.	Présentation du support de l'étude	78
III.	Etude mono-corps	80
IV.	Etude tri-corps I : cas de maillages conformes	83
V.	Etude tri-corps II : cas de maillages non-conformes	91
VI.	Conclusion	93

L'application développée dans ce chapitre correspond à la principale problématique qui avait initié cette étude de couplage par sous-structuration. Il s'agit de permettre une réduction performante d'un modèle détaillé de grande taille. Ces modèles sont difficiles à traiter avec la méthode modale de branche, à cause de la quantité de mémoire qui est indispensable pour le calcul et l'utilisation de la base branche. En effet la base de branche est définie sur l'ensemble du domaine et le nombre de modes nécessaires augmente avec la complexité géométrique.

Ce chapitre présente ainsi une application de la technique de sous-structuration modale dans un cas tridimensionnel pour lequel le modèle détaillé comporte un grand nombre de nœuds.

Bien que le modèle détaillé ne comporte pas de résistance thermique de contact (*RTC*), le couplage rend nécessaire l'introduction d'une fonctionnelle de saut de température pour compenser l'erreur liée à la réduction des bases. Le coefficient de pondération de la fonctionnelle peut être interprété comme une *RTC*, sa valeur étant évidemment faible. Au delà de l'objectif de validation du code C++, on souhaite voir si l'introduction de cette *RTC* permet de contenir le saut de température dans des limites raisonnables au niveau des surfaces de coupure.

La configuration choisie est inspirée d'une étude expérimentale menée au LMEE, dans laquelle la méthode modale de branche est utilisée pour identifier, en temps réel et par mesure non intrusive, une dissipation volumique de flux de chaleur [49].

Le chapitre commence par présenter le système thermique étudié, ainsi que les résultats obtenus à l'aide d'un modèle détaillé, puis ceux d'un modèle réduit mono-corps.

Une démarche de sous-structuration est alors proposée. Deux cas sont étudiés :

- Les maillages sont identiques au niveau des frontières de contact. Dans ce cas, la fonctionnelle de saut de flux a peu d'influence sur les résultats. Les résultats obtenus par analyse modale sous-structurée sont meilleurs que ceux obtenus par analyse modale mono-corps.
- Les maillages ne sont pas identiques au niveau des frontières de contact. Dans ce cas, les opérateurs de projection introduisent une erreur de projection. Il apparaît alors que contrairement au cas précédent, les résultats des essais numériques sont très sensibles à la valeur du facteur de pondération de la fonctionnelle de saut de flux. Un choix optimal permet alors d'obtenir des résultats proches de ceux obtenus avec des maillages identiques sur les surfaces de contact.

I. Présentation des résultats de l'article [49]

Le problème inverse de chaleur est mal posé au sens de Hadamard [40]. Il est possible de stabiliser l'inversion en utilisant une méthode de régularisation¹. La méthode des pas de temps futurs de Beck [6] est utilisée conjointement avec la méthode modale classique dans les articles [5, 42] et avec la méthode modale de branche dans l'article [49].

Une base de branche mono-corps est utilisée dans l'article [49], des résultats satisfaisants en inversion sont obtenus avec des bases réduites d'ordre compris entre 5 et 100. Cependant l'analyse directe (présentée brièvement dans l'article) n'est pas précise (écart maximum supérieur à 30K dans certaines zones du système physique étudié).

I.1. Présentation du dispositif physique

Le dispositif physique est un bloc d'acier parallélépipédique (voir figure 5.5 en page 78) de base carrée ($9.8cm \times 9.8cm \times 16.4cm$). Deux canaux, de diamètre égal à $2,4cm$, dédiés à un refroidissement par huile ont été creusés longitudinalement dans le bloc. Deux cartouches chauffantes cylindriques, de diamètre égal à $2cm$ et de hauteur égale à $6cm$, sont placées verticalement dans le bloc. Les surfaces extérieures sont isolées, exception faite des deux grandes surfaces du bloc qui sont peintes en noir (voir figure 5.1).

La puissance dissipée dans chaque cartouche est commandée par un programme informatique. Une caméra infrarouge (modèle A40 de Flir) permet une acquisition du champ de température d'une des deux faces peintes en noir. L'acquisition des données thermographiques est effectuée en temps réel grâce au port haut débit Firewire (Norme IEEE 1394). L'inversion s'effectue avec un décalage temporel imposé par la méthode des pas de temps futurs. La détermination de la sollicitation à l'instant t_{k+1} s'effectuant avec la connaissance des températures aux pas de temps futurs t_{k+1+r} avec r compris entre 1 et n_{max} où n_{max} est le nombre de pas de temps futurs. Un décalage temporel supplémentaire est lié au temps de calcul de l'algorithme utilisé, cependant ce temps est faible (de $0.01s$ à $0.65s$ suivant la taille du modèle modal utilisé) devant le décalage imposé par les pas de temps futurs ($10s$ de pas de temps à multiplier par le nombre de pas de temps futurs).

Le synoptique général du banc expérimental est donné en figure 5.2. Des thermocouples permettent de mesurer les températures suivantes :

- Températures de l'eau et de l'huile aux entrées et sorties de l'échangeur de chaleur.
- Température de l'huile à la sortie du bloc d'acier.

¹Trois méthodes de régularisation sont classiques [40] : la régularisation par pénalisation (Tikhonov), la régularisation par troncature de spectre, la méthode de spécification de fonction ou méthode des pas de temps futurs (Beck).

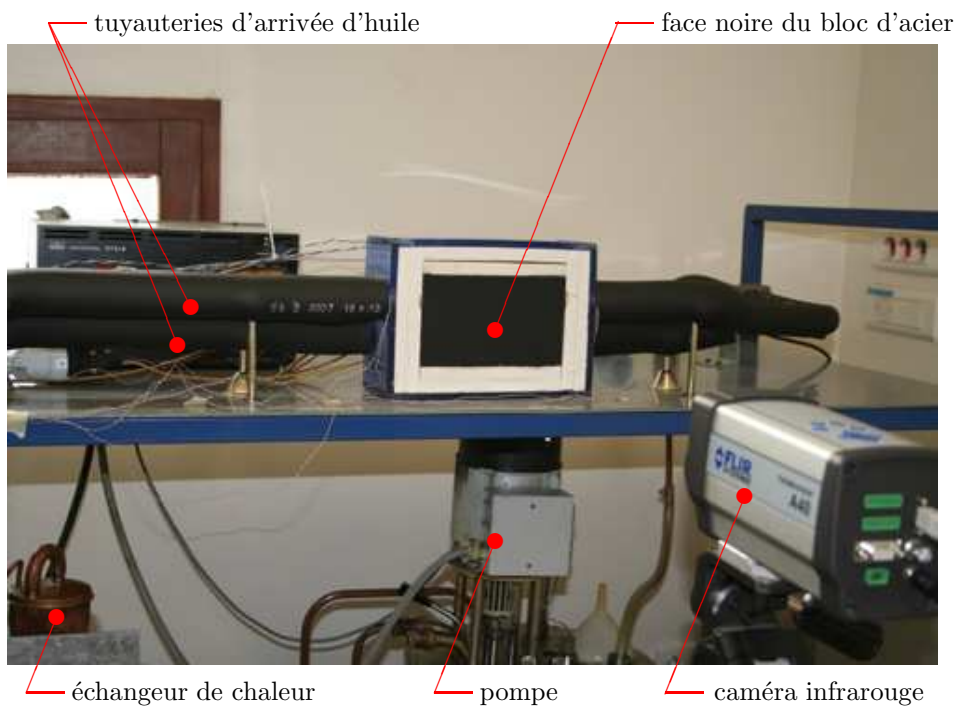


FIG. 5.1 – Photographie du bloc lors d'une expérience

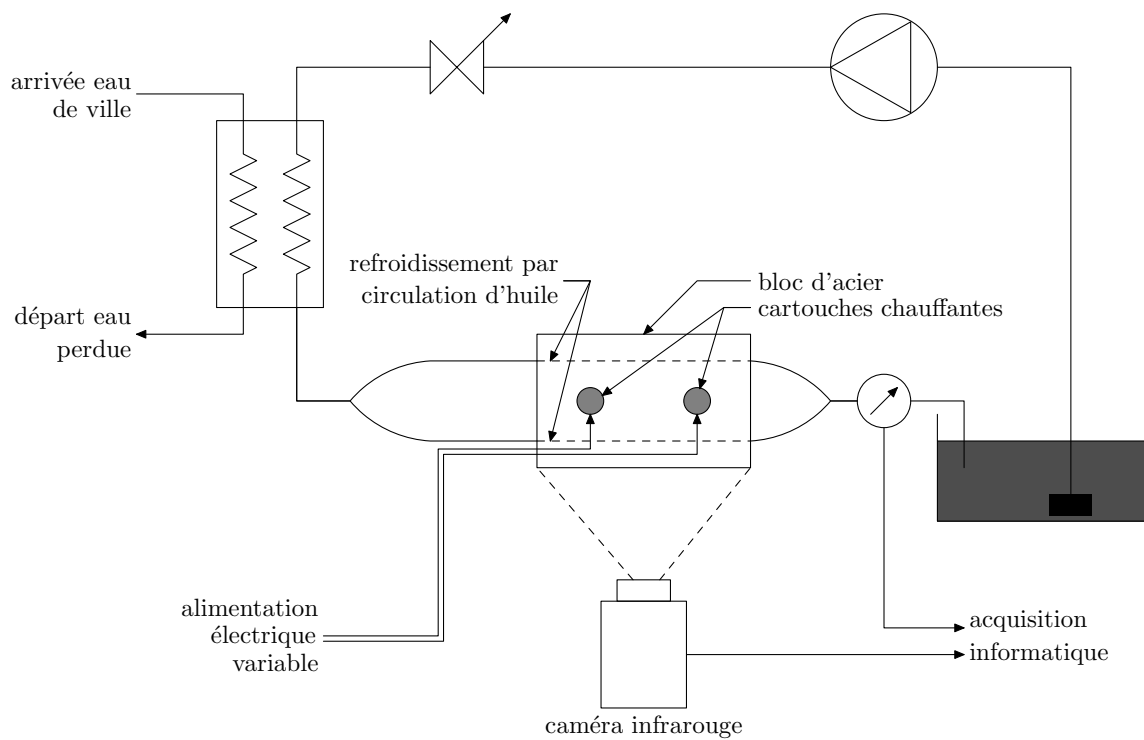


FIG. 5.2 – Synoptique du dispositif expérimental

I.2. Résultats de l'analyse modale directe

La détermination de la base modale est délicate, en effet la mémoire RAM de 2Go ne permet pas de déterminer plus de 4000 modes². La troncature de Marshall (voir page 140) qui est alors réalisée est trop importante : les 4000 modes sont insuffisants pour permettre une bonne décomposition du champ thermique au sens de \mathcal{C}_b (voir page 59), des modes dominants pour cette décomposition n'ont donc pas été calculés. Idéalement il faudrait pouvoir calculer les modes dans l'ordre de leur dominance (voir l'équation F.2), mais en pratique les modes sont calculés dans l'ordre de leur valeur propre.

Les différents modèles réduits conduiront toujours au même résultat : l'écart moyen (modal-éléments finis) sur l'ensemble du domaine est faible, mais le maximum de l'écart est toujours supérieur à 30K même avec un modèle réduit d'ordre 100. L'écart maximal est systématiquement observé au centre des cartouches chauffantes, ce qui n'est pas étonnant :

- Les cartouches chauffantes sont moins diffusives que le bloc d'acier.
- Les méthodes modales font intervenir un produit scalaire qui apporte une pondération par le volume. Les cartouches chauffantes sont donc moins prises en compte que le bloc dont le volume est prépondérant.
- La sollicitation du système (l'apport d'énergie par effet joule) est localisée dans les cartouches chauffantes.

Si l'écart maximum reste toujours élevé, l'écart moyen diminue fortement (voir figure 5.3) avec l'augmentation de l'ordre du modèle réduit. L'écart moyen est défini par :

$$\epsilon_{moy} = moy_{temps} \left(moy_{volume} \left(T - \tilde{T}_{\tilde{N}_1} \right) \right). \quad (5.1)$$

Avec T le champ de température du modèle élément fini, et $\tilde{T}_{\tilde{N}_1}$ le champ de température du modèle réduit à l'ordre \tilde{N}_1 .

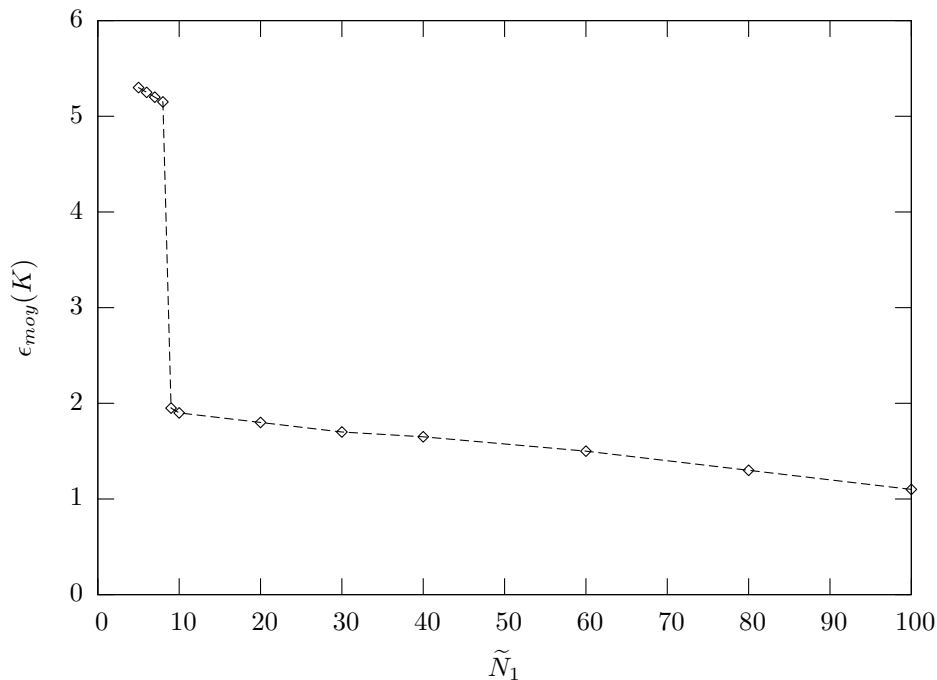


FIG. 5.3 – Résultats de l'analyse modale mono-corps en 3D.

²Le maillage comporte 23139 nœuds. La mémoire RAM disponible sur l'ordinateur, 2Go, impose alors le nombre de modes. La détermination de la base a été particulièrement laborieuse, la numérotation des nœuds a été spécifiquement optimisée car le code permettant l'inversion est mono-corps. Cette optimisation ne sera pas reprise par la suite.

Comme le montre la figure 5.3, l'écart moyen diminue brusquement avec le mode réduit numéro 9. La méthode de l'amalgame modal découplé (voir annexe F) a été utilisée pour la réduction : le mode maître associé au 9ème mode réduit est le 143ème mode de branche. La figure 5.4, présente quatre tranches du 143ème mode de branche. Ce mode est localisé³ en grande partie sur le centre des cartouches, ce qui explique naturellement son effet sur l'erreur moyenne.

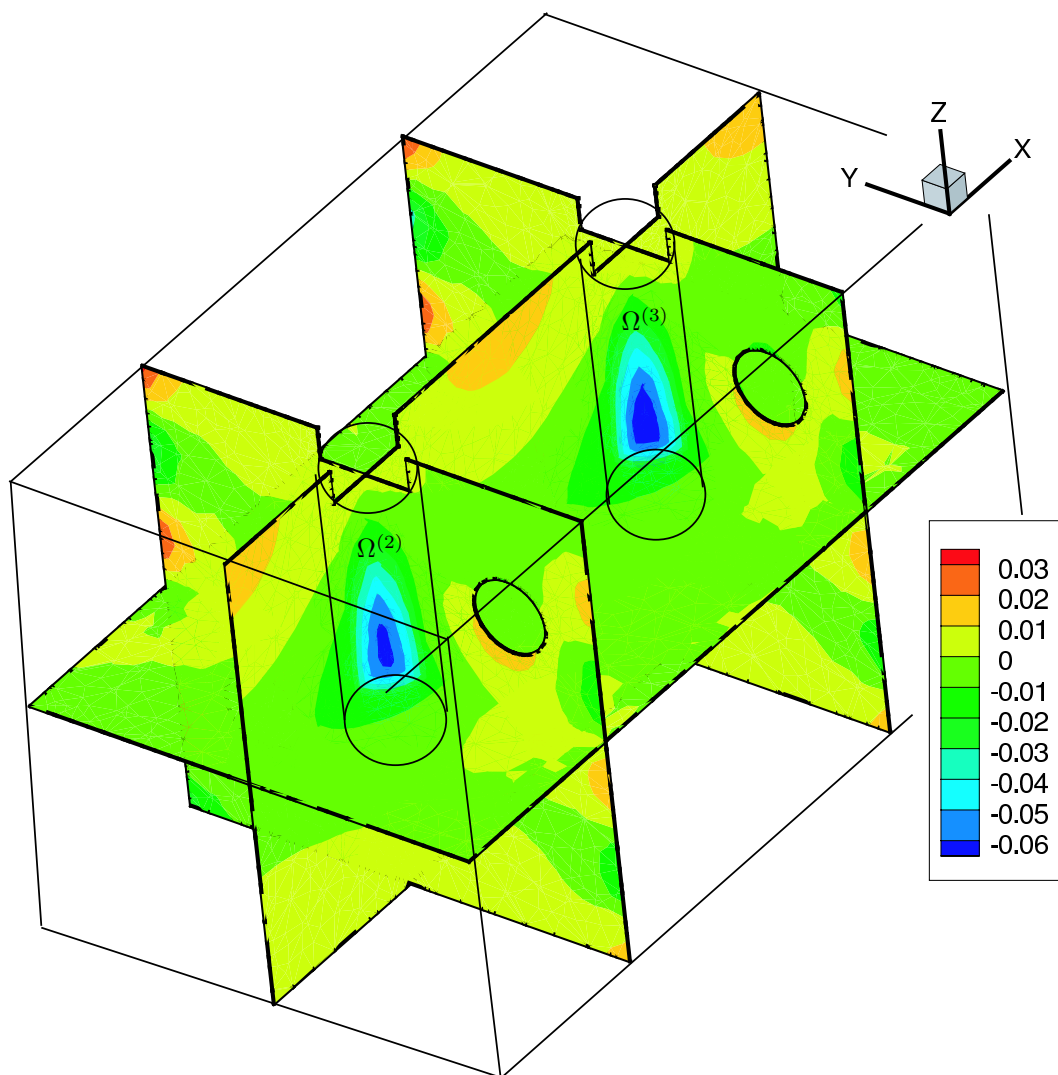


FIG. 5.4 – Vue du 143ème mode de branche à l'aide de quatre tranches

Remarque : dans la procédure d'amalgame (voir annexe F), les neuf premiers modes maîtres sont, dans l'ordre, les modes de branche suivants : {1 , 18 , 5 , 171 , 27 , 165 , 55 , 8 , 143}. Deux des neuf premiers modes maîtres sont localisés dans les cartouches chauffantes, il s'agit des modes de branche numéro 143 et 171. Le 143ème mode est donc le deuxième mode maître sélectionné par l'amalgame et localisé dans les cartouches. Mais le 171ème mode de branche est plus précisément localisé dans la partie supérieure des cartouches et non au centre des cartouches chauffantes.

³La définition de la localisation est, dans ce cas, la suivante : $Loc_{\Omega^{(2)} \cup \Omega^{(3)}} = \frac{\int_{\Omega^{(2)} \cup \Omega^{(3)}} V^2 d\Omega}{\int_{\Omega} V^2 d\Omega}$. Le mode est d'autant plus localisé dans les cartouches que $Loc_{\Omega^{(2)} \cup \Omega^{(3)}}$ est proche de 1.

II. Présentation du support de l'étude

II.1. Géométrie

La géométrie est celle du bloc de l'article [49] :

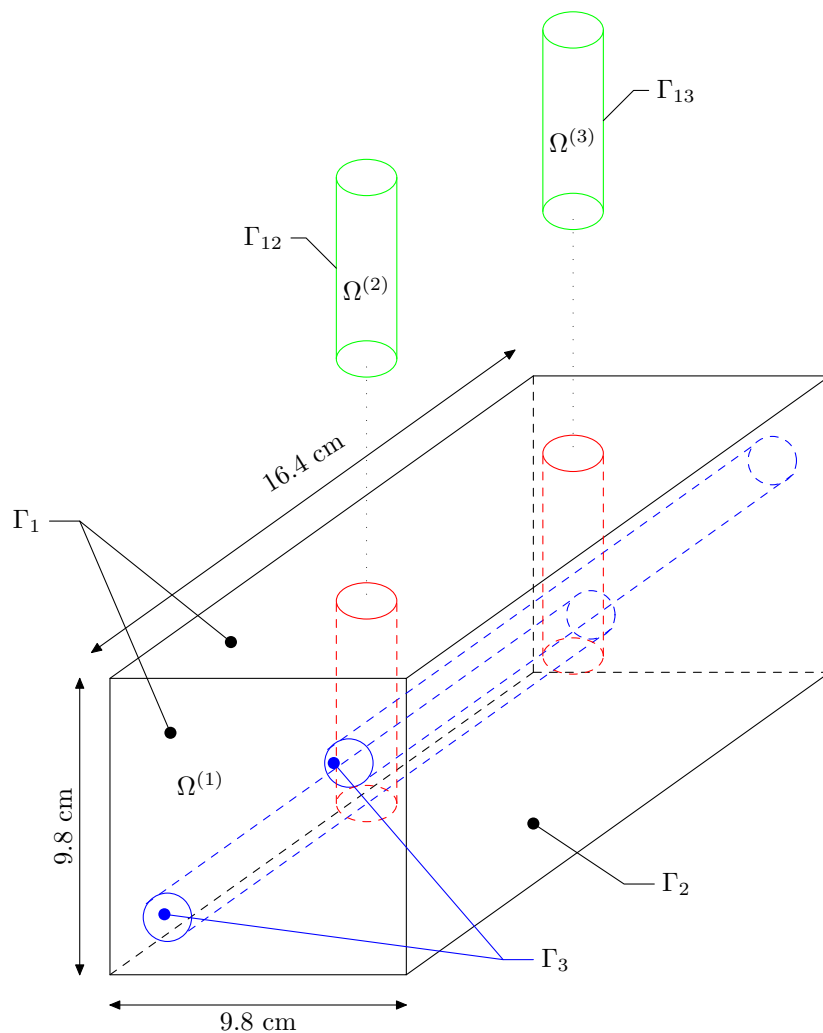


FIG. 5.5 – Géométrie du bloc

La désignation des corps et des surfaces est donnée dans le tableau suivant :

Notation	référence
$\Omega^{(1)}$	bloc d'acier
$\Omega^{(2)}$	cartouche chauffante gauche
$\Omega^{(3)}$	cartouche chauffante droite
Γ_1	surfaces extérieures isolées du bloc d'acier
Γ_2	surfaces extérieures non isolées du bloc d'acier
Γ_3	surfaces du bloc d'acier en contact avec huile
Γ_{12}	surface de contact entre $\Omega^{(1)}$ et $\Omega^{(2)}$
Γ_{13}	surface de contact entre $\Omega^{(1)}$ et $\Omega^{(3)}$

TAB. 5.1 – Désignation des corps et des surfaces

II.2. Propriétés thermophysiques

Les propriétés thermophysiques choisies sont données dans le tableau suivant :

	k	c
	$W.m^{-1}.K^{-1}$	$J.m^{-3}.K^{-1}$
bloc d'acier	52	3840000
cartouches chauffantes	8	4200000

TAB. 5.2 – Caractéristiques des matériaux

II.3. Lois d'échange et puissances volumiques dégagées

Le bloc est refroidi par la circulation de l'huile. La température de l'huile T_h et le coefficient de convection dans l'huile h_h sont supposés constants avec $T_h = 300K$ et $h_h = 40W.m^{-2}.K^{-1}$.

Les surfaces latérales sont considérées comme étant adiabatiques. Seules les deux surfaces les plus grandes du bloc échangent de l'énergie avec l'air de la pièce par convection et avec les surfaces solides par rayonnement. La température de l'ambiance T_{amb} et le coefficient de convection dans l'air h_a sont supposés constants avec $T_{amb} = 300K$ et $h_a = 8W.m^{-2}.K^{-1}$. Le rayonnement est estimé en considérant que les surfaces radiatives sont grises ($\epsilon = 0.9$) et que le bloc est petit devant les dimensions de la pièce.

L'apport volumique d'énergie par effet joule dans les cartouches chauffantes s'effectue suivant les lois données dans le graphique suivant :

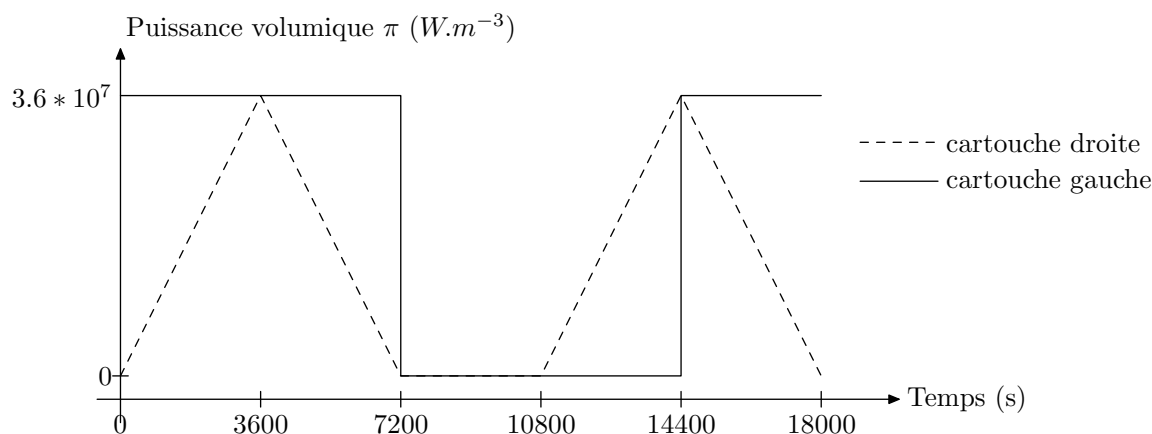


FIG. 5.6 – Evolution temporelle de la puissance volumique dissipée dans les cartouches

Le contact entre les cartouches et le bloc est supposé parfait, ce qui signifie que la résistance de contact est nulle.

II.4. Equations du problème

Le problème physique est décrit par les équations suivantes :

$$\forall M \in \Omega \setminus (\Gamma_{12} \cup \Gamma_{13}) \quad c \frac{\partial T}{\partial t} = \text{div}(k \underline{\nabla}(T)) + \pi \quad (5.2)$$

$$\forall M \in \Gamma_1 \quad -k^{(1)} \underline{\nabla}(T^{(1)}) \bullet \underline{n}^{(1)} = 0 \quad (5.3)$$

$$\forall M \in \Gamma_2 \quad -k^{(1)} \underline{\nabla}(T^{(1)}) \bullet \underline{n}^{(1)} = h_a (T^{(1)} - T_{amb}) + \epsilon \sigma_{sb} \left((T^{(1)})^4 - T_{amb}^4 \right) \quad (5.4)$$

$$\forall M \in \Gamma_3 \quad -k^{(1)} \underline{\nabla}(T^{(1)}) \bullet \underline{n}^{(1)} = h_h (T^{(1)} - T_h) \quad (5.5)$$

$$\forall M \in \Gamma_{12} \quad -k^{(1)} \underline{\nabla}(T^{(1)}) \bullet \underline{n}^{(1)} = k^{(2)} \underline{\nabla}(T^{(2)}) \bullet \underline{n}^{(2)} \quad \text{et} \quad T^{(1)} = T^{(2)} \quad (5.6)$$

$$\forall M \in \Gamma_{13} \quad -k^{(1)} \underline{\nabla}(T^{(1)}) \bullet \underline{n}^{(1)} = k^{(3)} \underline{\nabla}(T^{(3)}) \bullet \underline{n}^{(3)} \quad \text{et} \quad T^{(1)} = T^{(3)} \quad (5.7)$$

$$\forall M \in \Omega \quad T(M, 0) = T_{amb} \quad (5.8)$$

La constante de Stefan-Boltzmann est notée σ_{sb} pour éviter toute confusion avec le décalage spectral σ utilisé pour le calcul des modes de branche (voir page 136).

II.5. Horizon temporel et sauvegarde

Le temps physique simulé va de 0s à 18000s. Les sauvegardes se font à pas de temps fixe $\delta_{t,sauvegarde}$. Ce pas de temps est calculé de manière à avoir $N_{sauv} = 180$ temps de sauvegarde. Ainsi l'horizon temporel $\Delta t = 18000$ s et le pas de temps de sauvegarde sont reliés par :

$$\delta_{t,sauvegarde} = \frac{\Delta t}{N_{sauv} + 1} \quad (5.9)$$

III. Etude mono-corps

La solution du problème décrit par les équations 5.2 à 5.8 peut être recherchée à l'aide d'un modèle élément fini mono-corps. Le modèle que l'on note MD_M (Modèle Détaillé Mono-corps) détermine un champ de température sur l'ensemble du domaine : T_{MD_M} . Le maillage utilisé comporte $N_{MD_M} = 23503$ nœuds⁴, la taille des tétraèdres est aussi uniforme que possible.

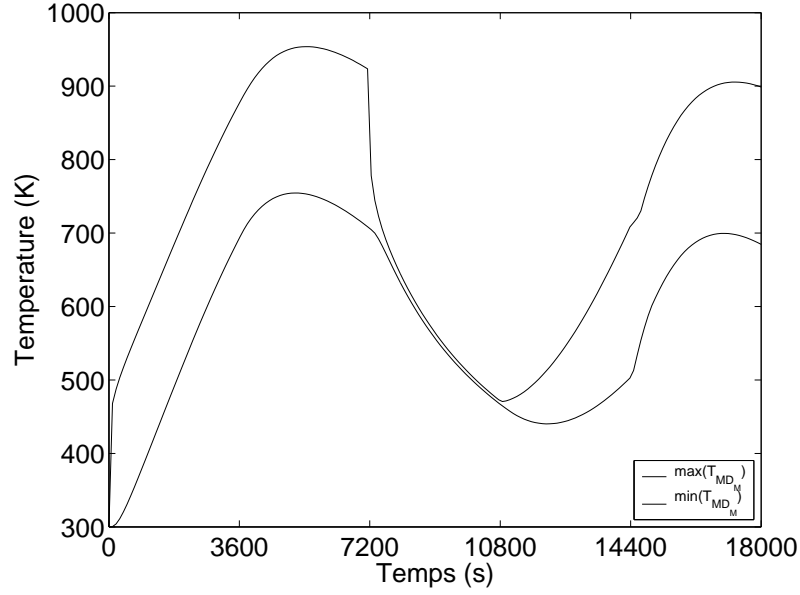
L'évolution de la température maximale et minimale sur l'ensemble du domaine est présentée dans la figure 5.7. L'amplitude thermique maximale est $\Delta T_{max} = 654K$.

La matrice des températures sauvegardées du modèle MD_M est notée \mathbb{T}_{MD_M} , elle comporte N_{MD_M} lignes et N_{sauv} colonnes.

La taille du modèle MD_M rend délicate l'analyse modale du problème, les besoins en mémoire RAM deviennent problématiques. Une base modale comportant⁵ les $\tilde{N}_0 = 3000$ premiers modes de branche est déterminée, la matrice des modes tronqués (car $\tilde{N}_0 < N$) est notée $\tilde{\mathbb{V}}$.

⁴Des oscillations liées au non-respect du principe du min-max sont inévitables (sauf si la matrice de capacité thermique est « condensée ») en 3D [41], ce qui limite la précision des simulations, la finesse du maillage doit être déterminée en conséquence.

⁵La mémoire RAM disponible sur l'ordinateur, 2Go, impose le nombre de modes. Il serait éventuellement possible de calculer quelques modes supplémentaires, au prix de beaucoup d'efforts car il faudrait pour cela optimiser la numérotation des nœuds dans le maillage.


 FIG. 5.7 – Evolution de la température maximale et minimale pour le modèle MD_M

En suivant les idées déjà développées en page 59, il est possible d'étudier l'erreur de projection dans la base modale. Les états de la décomposition modale tronquée sont déterminés grâce à l'opérateur \mathcal{C}_b discrétisé (opérateur matriciel \mathbb{C}_b) :

$$\tilde{\mathbf{X}} = {}^t \mathbf{V} \mathbb{C}_b \mathbb{T}_{MD_M} \quad (5.10)$$

La matrice des températures obtenue avec les n premiers modes est alors telle que :

$$\tilde{\mathbb{T}}_n(i, j) = \sum_{k=1}^n \tilde{\mathbb{V}}(i, k) * \tilde{\mathbf{X}}(k, j) \quad (5.11)$$

L'écart maximum entre $\tilde{\mathbb{T}}_n$ et \mathbb{T}_{MD_M} est $\epsilon_{MD_M, n}$:

$$\epsilon_{MD_M, n} = \max_{\text{ligne}} \left(\max_{\text{colonne}} \left(\text{abs} \left(\tilde{\mathbb{T}}_n - \mathbb{T}_{MD_M} \right) \right) \right) \quad (5.12)$$

La convergence dans la base modale est très lente comme l'atteste la figure 5.8. La réduction de la base est ainsi très délicate, les résultats obtenus sont donnés dans le tableau 5.3. Les écarts $\sigma_{M-M, \max}$ et $\sigma_{M-M, \text{moy}}$ entre le modèle modal mono-corps d'ordre \tilde{N}_1 (noté MR_M) et le modèle détaillé MD_M sont définis par :

$$\sigma_{M-M, \max}(\%) = \max_t \left(\max_{\Omega} \left(\text{abs} \left(T_{MR_M} - T_{MD_M} \right) \right) \right) \frac{100}{\Delta T_{\max}} \quad (5.13)$$

$$\sigma_{M-M, \text{moy}}(\%) = \frac{\int_t \int_{\Omega} \text{abs} \left(T_{MR_M} - T_{MD_M} \right) d\Omega dt}{\int_t dt \int_{\Omega} d\Omega} \frac{100}{\Delta T_{\max}} \quad (5.14)$$

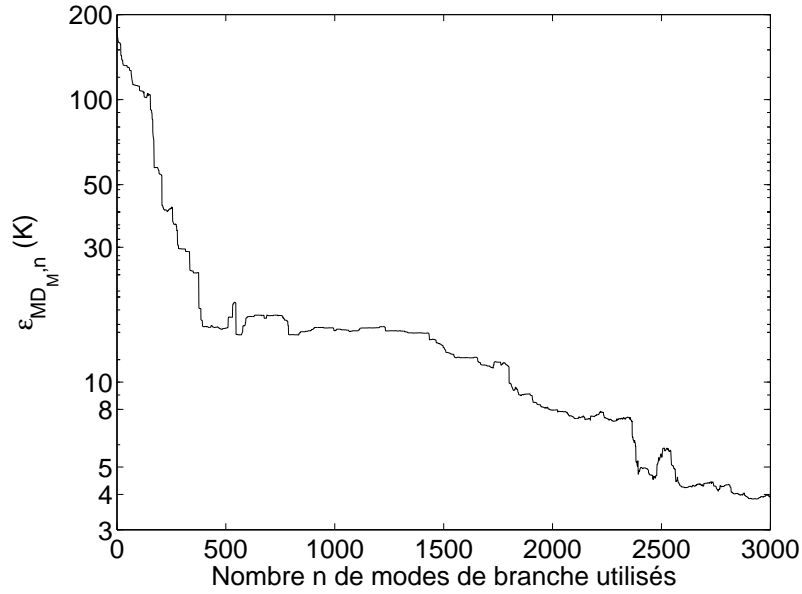


FIG. 5.8 – Erreur de projection en mono-corps

\tilde{N}_1	50	100	150	200	250	300
$\sigma_{M-M,max}(\%)$	14.1	10.9	9.2	7.0	5.7	4.9
$\sigma_{M-M,moy}(\%)$	1.57	0.88	0.80	0.78	0.75	0.71
T_{CPU} (s)	21	66	140	238	362	538

TAB. 5.3 – Résultats obtenus en analyse modale mono-corps

Le temps CPU nécessaire au calcul du modèle MD_M est de 1105s. Avec un écart maximal $\sigma_{M-M,max}$ inférieur à 5% le gain de temps est ainsi seulement de l'ordre de 2.

L'écart est indiquée en pourcentage de l'évolution thermique maximale, ici l'écart moyen est de 10.3K pour un modèle réduit comportant 50 modes. Or dans le graphique 5.3, l'écart moyen est inférieur à 2K pour le même nombre de modes. Comment peut-on expliquer cette différence dans les résultats? L'explication est très simple : une seule cartouche chauffante dissipe de l'énergie et cette dissipation est inférieure à $1.6 * 10^7 W.m^{-3}$ dans l'article [49].

Remarque importante : la méthode des éléments finis est très bien adaptée à la prise en compte de non-linéarités dans le volume (propriétés thermophysiques) et en surface (rayonnement, ...) : le temps de calcul est très faiblement affecté par le réassemblage des matrices éléments finis.

La méthode modale de branche est adaptée à la prise en compte des non-linéarités en volume et en surface, mais cela est très coûteux en temps de calcul. Un compromis entre la précision et le temps de calcul est obtenu en ne tenant pas compte des non-linéarités à chaque étape du calcul mais seulement tous les 100 pas de temps (cas des résultats du tableau 5.3).

Dans le cas présent la seule non-linéarité est celle introduite par le terme de rayonnement. Le biais introduit par les 99 pas de temps où l'on ne tient pas compte de l'évolution de la condition d'échange est faible (de l'ordre de 2K sur $\sigma_{M-M,max}$ soit 0.3% de ΔT_{max}).

IV. Etude tri-corps I : cas de maillages conformes

Le gain de temps pour une précision acceptable est très faible avec une méthode modale mono-corps. Il est possible que le gain de temps soit plus important pour une même précision avec une méthode modale multi-corps. L'étude de la sous-structuration est dans un premier temps effectuée avec des maillages conformes. L'étude avec des maillages non-conformes sera présentée dans la prochaine section soit à partir de la page 91.

Du fait de la réduction modale, il est nécessaire de contrôler les sauts de température et de flux. En effet la réduction modale pour chaque sous-domaine ne permet pas d'imposer rigoureusement la continuité du champ de température sur l'ensemble d'une interface de contact. Deux coefficients de pondération sont nécessaires dans la formulation variationnelle. Celui associé à la fonctionnelle de saut de température peut être interprété comme une résistance de contact. On retrouve le formalisme de couplage développé au chapitre 3. Le choix du coefficient associé à la fonctionnelle de saut de flux sera étudié par la suite.

IV.1. Le modèle MD_{3E}

Le modèle détaillé trois corps extraits MD_{3E} est obtenu par modification du modèle détaillé mono-corps MD_M :

- Les trois corps sont : le bloc d'acier et les deux cartouches chauffantes. Les maillages de ces trois corps sont **extraits** du maillage du modèle MD_M . Les maillages seront donc conformes au niveau des interfaces de raccord. Ceci implique que les nœuds sont identiques sur les interfaces de raccord, les projecteurs entre maillages (voir annexe G) n'introduisent alors aucune erreur, de plus $\mathbb{P}_{2 \rightarrow 1} \mathbb{P}_{1 \rightarrow 2} = \mathbb{Id}$ et $\mathbb{P}_{3 \rightarrow 1} \mathbb{P}_{1 \rightarrow 3} = \mathbb{Id}$.
- Les équations du modèle MD_{3E} sont en partie celles du modèle MD_M . Les équations 5.2, 5.3, 5.4, 5.5 et 5.8 sont conservées. Par contre les équations 5.6 et 5.7 sont modifiées pour introduire la RTC , elles sont remplacées par les équations 5.15 et 5.16.

Les nouvelles équations au niveau des interfaces de contact sont :

$$\forall M \in \Gamma_{12} \quad -k^{(1)} \underline{\nabla} \left(T^{(1)} \right) \bullet \underline{n}^{(1)} = k^{(2)} \underline{\nabla} \left(T^{(2)} \right) \bullet \underline{n}^{(2)} = -\frac{[T]_{\Gamma_{12}}}{RTC} \quad (5.15)$$

$$\forall M \in \Gamma_{13} \quad -k^{(1)} \underline{\nabla} \left(T^{(1)} \right) \bullet \underline{n}^{(1)} = k^{(3)} \underline{\nabla} \left(T^{(3)} \right) \bullet \underline{n}^{(3)} = -\frac{[T]_{\Gamma_{13}}}{RTC} \quad (5.16)$$

Un biais est admis lorsque l'on introduit la résistance thermique de contact, ce biais rend possible la réduction modale. Un compromis existe sur la valeur de la RTC (voir page suivante) :

- Si la RTC est trop grande (i.e supérieure à $RTC = 10^{-4} K.m^2.W^{-1}$) alors l'écart entre le modèle MD_{3E} et le modèle MD_M sera important. Même en admettant qu'il soit alors facile de réduire le modèle MD_{3E} , l'écart entre le modèle réduit MR_{3E} et le modèle MD_M sera important à cause de la valeur choisie pour la RTC .
- Si la RTC est très faible (i.e inférieure à $RTC = 10^{-6} K.m^2.W^{-1}$) alors l'écart entre le modèle MD_{3E} et le modèle MD_M sera très faible. Par contre la réduction modale devient très délicate, car le faible saut de température admis au niveau des interfaces de contact pénalise fortement les possibilités de réduction.

Ainsi de manière semi-arbitraire, la résistance thermique de contact qui sera utilisée est égale à $RTC = 10^{-5} K.m^2.W^{-1}$.

IV.2. Aide pour la détermination de la valeur de la RTC

Il est possible d'évaluer l'ordre de grandeur du biais dû à l'introduction d'une RTC . En effet l'ordre de grandeur de ce biais est donné par $\overline{\Delta T}_{saut,max}$ qui est la valeur moyenne du saut de température au travers de la RTC lorsque la puissance dissipée dans une cartouche est maximale et que le régime est permanent.

Soit $\pi_{max} = 3.6.10^7 W.m^{-3}$ la puissance volumique maximale dissipée dans une cartouche. En régime permanent, la conservation de l'énergie implique que $\overline{\Delta T}_{saut,max}$ vérifie :

$$\pi_{max}.V = \frac{S.\overline{\Delta T}_{saut,max}}{RTC} \quad (5.17)$$

Où V est le volume de la cartouche chauffante et S est la surface de la cartouche en contact avec le bloc. Il est possible d'exprimer V et S en fonction du rayon R et de la hauteur H de la cartouche :

$$V = \Pi.R^2.H \quad \text{et} \quad S = \Pi.R^2 + 2.\Pi.R.H \quad (5.18)$$

L'expression de $\overline{\Delta T}_{saut,max}$ est ainsi :

$$\overline{\Delta T}_{saut,max} = \frac{\pi_{max}.RTC.R.H}{R + 2.H} \quad (5.19)$$

Pour la valeur choisie, *i.e* $RTC = 10^{-5} K.m^2.W^{-1}$, la formule 5.19 donne : $\overline{\Delta T}_{saut,max} = 1.66 K$. Le graphe suivant permet de montrer que le choix d'une $RTC = 10^{-4} K.m^2.W^{-1}$ entraînerait un saut de température de l'ordre de 16K donc beaucoup trop important.

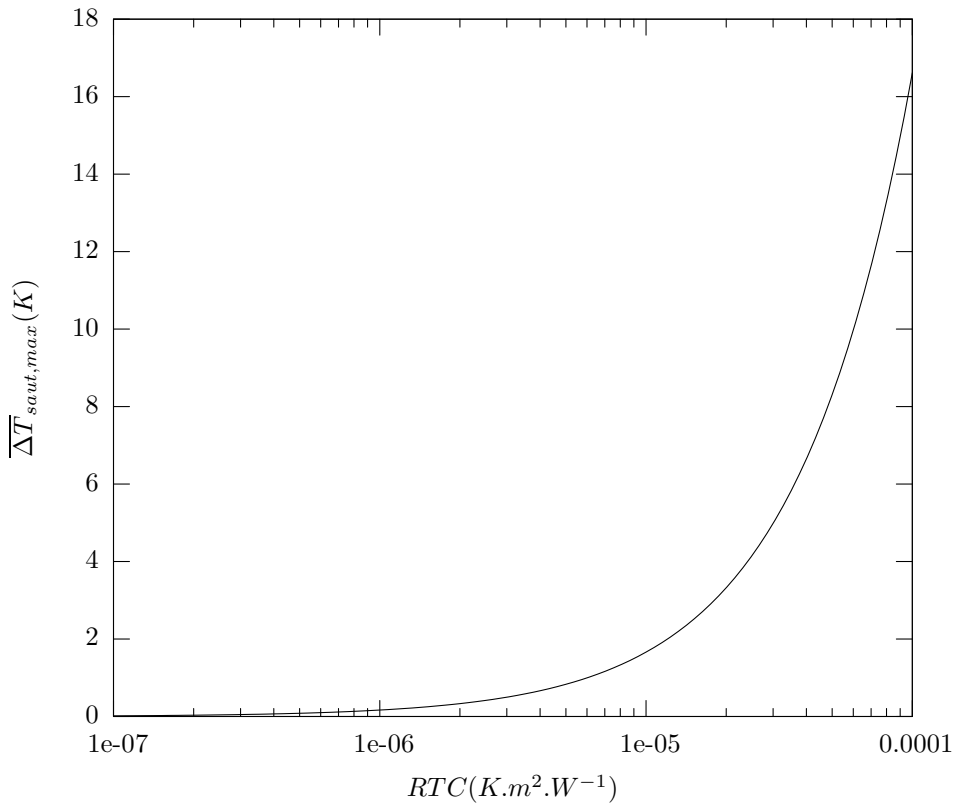


FIG. 5.9 – Evaluation de $\overline{\Delta T}_{saut,max}$ en fonction de la RTC par la formule 5.19

IV.3. Résultats du modèle détaillé MD_{3E}

Les températures minimale et maximale au cours du temps des deux modèles détaillés MD_M et MD_{3E} sont présentés dans la figure suivante :

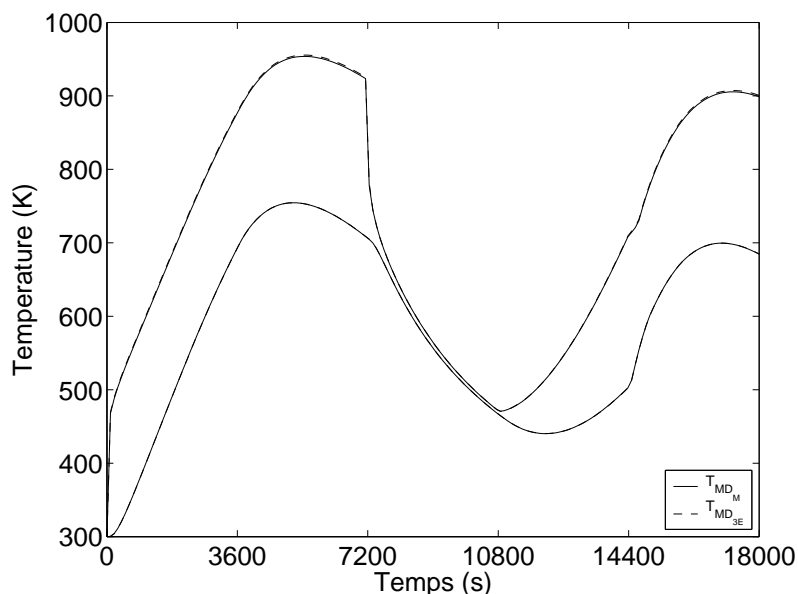


FIG. 5.10 – Températures minimale et maximale au cours du temps pour les deux modèles MD_M et MD_{3E}

Il y a bien 4 courbes (2 pour chaque modèle) dans le graphique 5.10, cependant l'écart entre les champs de température des modèles MD_M et MD_{3E} est suffisamment faible pour que les courbes du modèle MD_{3E} soient (presque) dans « l'épaisseur du trait » des courbes du modèle MD_M .

L'écart maximal entre les champs de température des modèles MD_M et MD_{3E} est très faible, l'écart maximum étant égal à $2.03K$ (l'estimation donnée $\overline{\Delta T}_{saut,max} = 1.66K$ est donc très acceptable).

L'introduction de la *RTC* introduit donc un écart qui est au plus de $2.03K$ soit 0.31% de l'évolution thermique maximale $\Delta T_{max} = 654K$.

Le maximum de l'écart est obtenu au centre de la cartouche gauche, ce qui est cohérent étant donné l'évolution temporelle de l'apport d'énergie (voir figure 5.6).

IV.4. Détermination des bases modales

La démarche qui suit a déjà été appliquée en page 59. L'idée est d'étudier *a posteriori* l'erreur de décomposition par projection (au sens de l'opérateur \mathcal{C}_b , voir E.44 en 134) du champ de référence dans la base modale. Les maillages utilisés sont ceux du modèle MD_{3E} . Les champs de températures des modèles modaux (notés *MR* pour Modèle Réduit) seront comparés au champ de température du modèle éléments finis sous-jacent donc le modèle MD_{3E} .

Les bases modales comportent⁶ les $\tilde{N}_0^{(1)} = 3000$, $\tilde{N}_0^{(2)} = 1470$, $\tilde{N}_0^{(3)} = 1476$ premiers modes de branche sur les $N^{(1)} = 22377$, $N^{(2)} = 1476$, $N^{(3)} = 1485$ modes déterminables (il s'agit du nombre de nœuds des maillages respectifs).

⁶La mémoire RAM disponible sur l'ordinateur, 2Go, impose le nombre de modes pour le bloc. Il serait éventuellement possible de calculer quelques modes supplémentaires, au prix de beaucoup d'efforts car il faudrait pour cela optimiser la numérotation des nœuds dans le maillage.

A noter : le nombre total de nœuds dans le modèle MD_{3E} , $N^{(1)} + N^{(2)} + N^{(3)} = 25338$, est supérieur au nombre de nœuds du modèle MD_M , $N_{MD_M} = 23503$, car les nœuds des interfaces de contact sont dupliqués.

La matrice des températures sauvegardées pour le domaine $\Omega^{(i)}$ est notée $\mathbb{T}_{MD_{3E}}^{(i)}$, elle comporte $N^{(i)}$ lignes et N_{sauv} colonnes. Les états de la décomposition modale tronquée (car $\tilde{N}_0^{(i)} < N^{(i)}$) sont déterminés grâce à l'opérateur $\mathcal{C}_b^{(i)}$ discrétisé (opérateur matriciel $\mathbb{C}_b^{(i)}$) :

$$\tilde{\mathbb{X}}^{(i)} = {}^t \tilde{\mathbb{V}}^{(i)} \mathbb{C}_b^{(i)} \mathbb{T}_{MD_{3E}}^{(i)} \quad (5.20)$$

Où $\tilde{\mathbb{V}}^{(i)}$ est la matrice des modes tronqués du domaine $\Omega^{(i)}$, elle comporte $N^{(i)}$ lignes et $\tilde{N}_0^{(i)}$ colonnes.

La matrices des températures du domaine $\Omega^{(i)}$ obtenue avec les n premiers modes est alors telle que :

$$\tilde{\mathbb{T}}_n^{(i)}(i, j) = \sum_{k=1}^n \tilde{\mathbb{V}}^{(i)}(i, k) * \tilde{\mathbb{X}}^{(i)}(k, j) \quad (5.21)$$

L'écart maximum entre $\tilde{\mathbb{T}}_n^{(i)}$ et $\mathbb{T}_{MD_{3E}}^{(i)}$ est $\epsilon_{MD_{3E},n}^{(i)}$:

$$\epsilon_{MD_{3E},n}^{(i)} = \max_{\text{ligne}} \left(\max_{\text{colonne}} \left(\text{abs} \left(\tilde{\mathbb{T}}_n^{(i)} - \mathbb{T}_{MD_{3E}}^{(i)} \right) \right) \right) \quad (5.22)$$

Les trois graphiques 5.11, 5.12 et 5.13 montrent que les bases de branche locales sont effectivement capables d'assurer une bonne décomposition. Au sens de l'opérateur \mathcal{C}_b , l'erreur de décomposition bien inférieure à $1K$ (soit 0.153% de l'évolution thermique maximale $\Delta T_{max} = 654K$) lorsque l'on garde l'intégralité des modes calculés, pour chacun des trois sous-domaines.

Dans le cas des deux cartouches, on note la nécessité de garder tous les modes avant d'effectuer l'amalgame modal. L'examen des figures 5.12 et 5.13 montre que la prise en compte des derniers modes dans le calcul de l'erreur maximale entraîne une chute brutale de celle-ci.

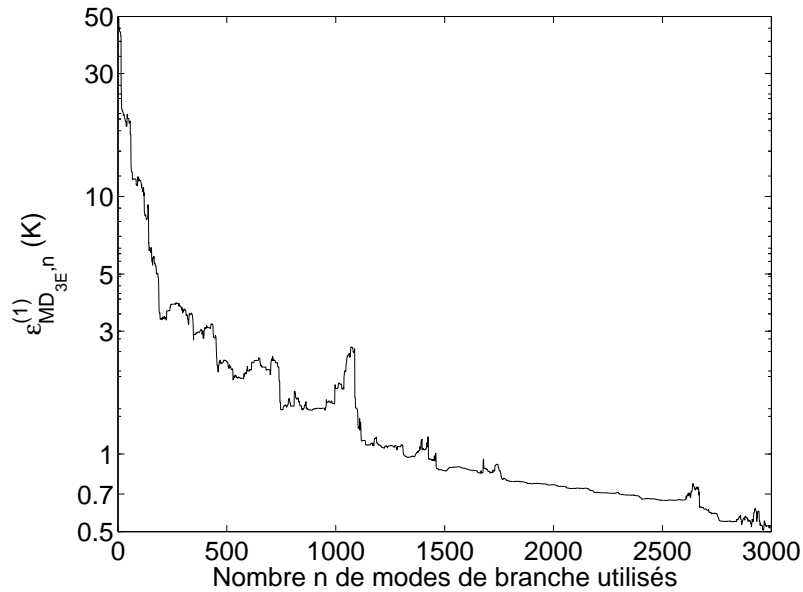


FIG. 5.11 – Erreur de projection pour le bloc

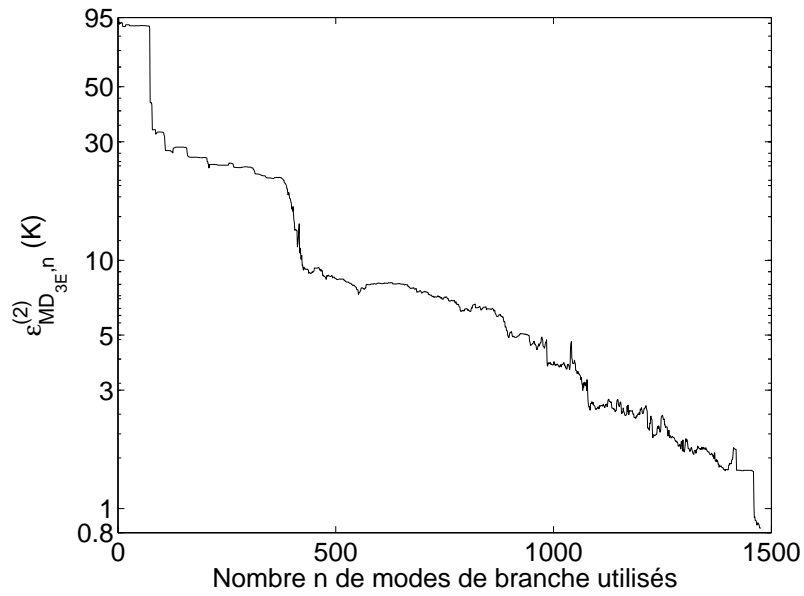


FIG. 5.12 – Erreur de projection pour la cartouche de gauche

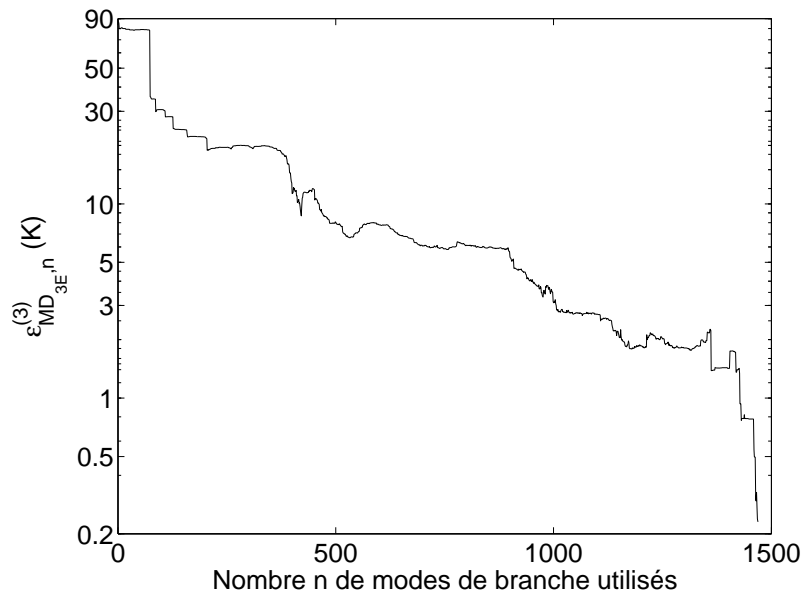


FIG. 5.13 – Erreur de projection pour la cartouche de droite

IV.5. Résultats de l'analyse modale

IV.5.1. Cas tests utilisés pour l'amalgame

A partir des bases de branche obtenues, composées respectivement des $\{3000, 1470, 1476\}$ premiers modes de branche, la réduction modale s'effectue par la méthode de l'amalgame modal découplé.

Il a été montré, dans le chapitre précédent, que la sensibilité au cas test est faible. Ainsi contrairement au chapitre précédent, la sensibilité au cas test ne sera pas approfondie.

Le cas test du bloc :

Pour le bloc, le cas test est défini par les équations suivantes :

$$\forall M \in \Omega^{(1)} \quad c^{(1)} \frac{\partial T^{(1)}}{\partial t} = \text{div} \left(k^{(1)} \underline{\nabla} \left(T^{(1)} \right) \right) \quad (5.23)$$

$$\forall M \in \Gamma_1 \quad -k^{(1)} \underline{\nabla} \left(T^{(1)} \right) \bullet \underline{n}^{(1)} = 0 \quad (5.24)$$

$$\forall M \in \Gamma_2 \quad -k^{(1)} \underline{\nabla} \left(T^{(1)} \right) \bullet \underline{n}^{(1)} = h_{2a} (T^{(1)} - T_{amb}) \quad (5.25)$$

$$\forall M \in \Gamma_3 \quad -k^{(1)} \underline{\nabla} \left(T^{(1)} \right) \bullet \underline{n}^{(1)} = h_h (T^{(1)} - T_h) \quad (5.26)$$

$$\forall M \in \Gamma_{12} \quad -k^{(1)} \underline{\nabla} \left(T^{(1)} \right) \bullet \underline{n}^{(1)} = h_{12a} (T^{(1)} - T_{12a}) \quad (5.27)$$

$$\forall M \in \Gamma_{13} \quad -k^{(1)} \underline{\nabla} \left(T^{(1)} \right) \bullet \underline{n}^{(1)} = h_{13a} (T^{(1)} - T_{13a}) \quad (5.28)$$

$$\forall M \in \Omega^{(1)} \quad T^{(1)}(M, 0) = T_{amb} \quad (5.29)$$

Dans le cas test du bloc :

- La frontière Γ_1 reste adiabatique.
- Pour la frontière Γ_2 , la condition aux limites est linéarisée⁷ et $h_{2a} = 30W.m^{-2}.K^{-1}$
- Pour la frontière Γ_3 , la condition aux limites est celle du problème physique.
- Pour les frontières de contact, deux conditions différentes sont utilisées pour tenir compte de la différence des apports volumiques d'énergie : $h_{12a} = 10000W.m^{-2}.K^{-1}$, $h_{13a} = 1000W.m^{-2}.K^{-1}$ et $T_{12a} = T_{13a} = 500K$. La valeur élevée du coefficient de convection donne du poids aux frontières de contact dont la surface est moindre (notons qu'une valeur 10 ou 100 fois plus faible permet d'obtenir des résultats comparables). Le rapport 10 entre h_{12a} et h_{13a} est justifié par la forte différence dans l'évolution temporelle des apports d'énergie (créneau contre rampe voir figure 5.6).

Le cas test des cartouches :

Il y a un cas test identique pour les deux cartouches (ainsi seul celui de la cartouche gauche $\Omega^{(2)}$ est donné), celui-ci est défini par les équations suivantes :

$$\forall M \in \Omega^{(2)} \quad c^{(2)} \frac{\partial T^{(2)}}{\partial t} = \text{div} \left(k^{(2)} \underline{\nabla} \left(T^{(2)} \right) \right) + \pi_{max} \quad (5.30)$$

$$\forall M \in \Gamma_{12} \quad -k^{(2)} \underline{\nabla} \left(T^{(2)} \right) \bullet \underline{n}^{(2)} = h_{21a} (T^{(2)} - T_{amb}) \quad (5.31)$$

$$\forall M \in \partial\Omega^{(2)} \setminus \Gamma_{12} \quad -k^{(2)} \underline{\nabla} \left(T^{(2)} \right) \bullet \underline{n}^{(2)} = 0 \quad (5.32)$$

$$\forall M \in \Omega^{(2)} \quad T^{(2)}(M, 0) = T_{amb} \quad (5.33)$$

⁷La sensibilité à la valeur de h_{2a} est très faible, en l'occurrence l'écart maximum entre deux modèles réduits comportant 30 modes pour le bloc (et 10 par cartouche) est inférieur à 0.12K pour h_{2a} variant de $14W.m^{-2}.K^{-1}$ à $58W.m^{-2}.K^{-1}$. Ces valeurs correspondent respectivement à celle obtenues par linéarisation de l'équation 5.4 pour $T^{(1)} = 300K$ et pour $T^{(1)} = 954K$.

Dans le cas test des cartouches :

- La valeur retenue pour l'apport volumique d'énergie est la valeur maximale : $\pi_{max} = 3.7.10^7 W.m^{-3}$
- La frontière $\partial\Omega^{(2)} \setminus \Gamma_{12}$ reste adiabatique.
- La frontière Γ_{12} est remplacée par une valeur acceptable⁸ : $h_{21a} = 50W.m^{-2}.K^{-1}$.

IV.5.2. Résultats concernant l'écart entre les MR_{3E} et les modèles détaillés

Les modèles modaux sous-structurés avec 3 corps et des maillages extraits (MR_{3E}) devraient présenter un écart faible avec le modèle détaillé MD_{3E} . Or le modèle détaillé MD_{3E} présente un écart faible avec le modèle détaillé mono-corps MD_M , celui-ci étant inférieur à 2.03K en valeur maximale. Ainsi les modèles MR_{3E} devraient présenter un écart faible avec le modèle détaillé MD_M .

Résultats de l'écart entre les MR_{3E} et le MD_{3E} :

Les écarts $\sigma_{3E-3E,max}$ et $\sigma_{3E-3E,moy}$ sont définis suivant le modèle des écarts $\sigma_{M-M,max}$ et $\sigma_{M-M,moy}$ (voir page 81).

$$\sigma_{3E-3E,max}(\%) = \max_t (\max_{\Omega} (abs(T_{MR_{3E}} - T_{MD_{3E}}))) \frac{100}{\Delta T_{max}} \quad (5.34)$$

$$\sigma_{3E-3E,moy}(\%) = \frac{\int_t \int_{\Omega} abs(T_{MR_{3E}} - T_{MD_{3E}}) d\Omega dt}{\int_t dt \int_{\Omega} d\Omega} \frac{100}{\Delta T_{max}} \quad (5.35)$$

Les résultats du tableau 5.4 sont obtenus sans fonctionnelle de saut de flux et en tenant compte des non-linéarités toutes les 100 étapes :

- L'ajout d'une fonctionnelle de saut de flux permet d'améliorer légèrement les résultats⁹. Cependant, contrairement au cas 2D, l'amélioration n'est pas assez importante pour justifier une étude de la valeur du paramètre de la fonctionnelle de saut de flux.
- La prise en compte des non-linéarités toutes les 100 étapes permet d'obtenir un bon compromis entre le temps de calcul d'une part et la précision d'autre part.

Ordre global	50	100	150	200	250	300
$\tilde{N}_1^{(1)}$	30	40	50	80	90	100
$\tilde{N}_1^{(2)}$	10	30	50	60	80	100
$\tilde{N}_1^{(3)}$	10	30	50	60	80	100
$\sigma_{3E-3E,max}(\%)$	7.40	3.17	3.05	2.96	2.79	2.78
$\sigma_{3E-3E,moy}(\%)$	1.22	0.87	0.81	0.77	0.76	0.76
T_{CPU} (s)	41	47	58	81	105	137

TAB. 5.4 – Comparaison MR_{3E} et MD_{3E}

Ces résultats ne sont pas directement comparables avec ceux du tableau 5.3, car le modèle modal est alors comparé au modèle détaillé mono-corps MD_M . Ainsi un deuxième jeu d'écarts va être défini.

⁸La sensibilité à la valeur de h_{21a} est très faible, en l'occurrence l'écart maximum entre deux modèles réduits comportant 10 modes par cartouche (et 30 pour le bloc) est inférieur à 0.42K pour h_{21a} variant de $50W.m^{-1}.K^{-1}$ à $5000W.m^{-1}.K^{-1}$.

⁹L'amélioration la plus importante a été obtenue pour le modèle d'ordre global 50. Pour ce modèle lorsque le paramètre de la fonctionnelle de saut de flux est optimal, les écarts sont : $\sigma_{3E-3E,max} = 7.34\%$ et $\sigma_{3E-3E,moy} = 1.19\%$

Résultats de l'écart entre les MR_{3E} et le MD_M :

Les écarts $\sigma_{3E-M,max}$ et $\sigma_{3E-M,moy}$ sont aussi définis suivant le modèle des écarts $\sigma_{M-M,max}$ et $\sigma_{M-M,moy}$ (voir page 81).

$$\sigma_{3E-M,max}(\%) = \max_t (\max_{\Omega} (abs(T_{MR_{3E}} - T_{MD_M}))) \frac{100}{\Delta T_{max}} \quad (5.36)$$

$$\sigma_{3E-M,moy}(\%) = \frac{\int_t \int_{\Omega} abs(T_{MR_{3E}} - T_{MD_M}) d\Omega dt}{\int_t dt \int_{\Omega} d\Omega} \frac{100}{\Delta T_{max}} \quad (5.37)$$

Ordre global	50	100	150	200	250	300
$\tilde{N}_1^{(1)}$	30	40	50	80	90	100
$\tilde{N}_1^{(2)}$	10	30	50	60	80	100
$\tilde{N}_1^{(3)}$	10	30	50	60	80	100
$\sigma_{3E-M,max}(\%)$	7.26	3.46	3.34	3.22	3.07	3.04
$\sigma_{3E-M,moy}(\%)$	1.18	0.86	0.82	0.79	0.79	0.79
T_{CPU} (s)	41	47	58	81	105	137

 TAB. 5.5 – Comparaison MR_{3E} et MD_M

Note : les résultats des tableaux 5.4 et 5.5 sont présentés pour les mêmes modèles réduits, ce qui justifie que les temps de calcul soient identiques.

Analyse des résultats obtenus

Il s'agit d'analyser les résultats des 5.3,5.4 et 5.5.

Concernant les tableaux 5.4 et 5.5 :

- Les erreurs moyennes sont très proches, car l'écart entre les modèles MD_M et MD_{3E} est très faible
- De même pour les erreurs maximales, la différence reste toujours inférieure à 0.29% de ΔT_{max} , ce qui est cohérent puisque l'écart maximal entre les MD_M et MD_{3E} est égal à 0.31% de ΔT_{max} .

L'analyse des résultats des tableaux 5.3 et 5.4 est celle qui nous intéresse le plus :

- L'amélioration de la réduction apportée par la sous-structuration modale permet d'obtenir, pour un même ordre global, $\sigma_{3E-M,max} < \sigma_{M-M,max}$. L'ordre de grandeur de l'erreur moyenne restant le même. La sous-structuration montre ainsi sa capacité dans un cas où la coupure est supposée fictive.
- Une erreur maximale inférieure à 5% est obtenue en 47s, le gain de temps est donc de 23.

V. Etude tri-corps II : cas de maillages non-conformes

La précédente section était dédiée au cas où les maillages étaient conformes sur les interfaces de contact. Or la méthode de couplage qui a été développée ne rend pas nécessaire la conformité des maillages au niveau des interfaces. De plus le code de calcul a été construit en intégrant des opérateurs de projection (voir annexe G). Cette section donne des éléments sur le traitement d'interfaces présentant une non-conformité.

Le traitement de maillages non conformes est particulièrement important pour l'ingénieur. Il est fréquent d'utiliser des maillages indépendants (car provenant de différents logiciels, de différentes entreprises, ...). Dans un tel cas, les maillages sont très probablement non-conformes sur les interfaces de contact.

Les trois corps (le bloc et les deux cartouches chauffantes) sont maillés séparément. Le modèle détaillé construit avec ces maillages est noté MD_{3S} . Il y a alors 23306 nœuds au total dans le maillage, ce qui est du même ordre de grandeur que le nombre de nœuds du modèle MD_M (23503) et du nombre de nœuds du modèle MD_{3E} (25338).

Les bases modales comportent les $\{3000, 1470, 1476\}$ premiers modes de branche comme dans le chapitre précédant. Les cas tests utilisés pour la réduction sont inchangés.

V.1. Cas du modèle détaillé MD_{3S}

Les courbes des température maximale et minimale au cours du temps du modèle MD_{3S} sont rajoutées à celles du graphique 5.10, ce qui donne le graphique suivant :

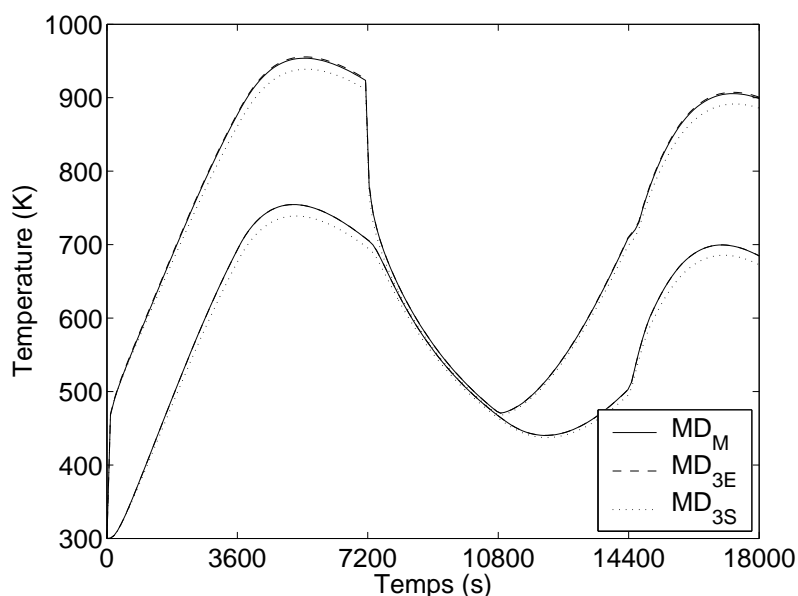


FIG. 5.14 – Températures minimale et maximale au cours du temps pour chacun des trois modèles éléments finis.

Si l'écart entre les modèles MD_{3E} et MD_M était faible (l'écart maximum est de 2.03K), l'écart entre les modèles MD_{3S} et MD_M est très important. L'écart élevé entre les modèles MD_{3S} et MD_M est dû pour une faible part à la *RTC* et pour une large part au biais introduit par les opérateurs de projection. Il serait éventuellement possible d'améliorer le modèle MD_{3S} en lui ajoutant une fonctionnelle de saut de flux¹⁰.

¹⁰Suivant la démarche adoptée ici dans le cas de l'analyse modale et comme proposé dans [9]

V.2. Résultats de l'analyse modale

Il est nécessaire d'utiliser une fonctionnelle de saut de flux dans les modèles réduits basés sur des maillages séparés (MR_{3S}). En effet sans fonctionnelle de saut de flux, les résultats des modèles réduits donnés dans le tableau 5.6, montrent que les écarts moyen et maximum entre les modèles MR_{3S} et MD_M sont très grand. A l'inverse lorsque le paramètre de pondération (β) de la fonctionnelle de saut de flux est optimum (β_{opt}) alors les écarts moyen et maximum entre les modèles MR_{3S} et MD_M obtenus sont proches des écarts obtenus entre les modèles MR_{3E} et MD_M .

La valeur optimale du paramètre β_{opt} est déterminée par une étude *a posteriori* qui nécessite de connaître la solution de référence (ici le modèle MR_{3E} qui est proche du modèle MD_M). Les courbes obtenues sont en tout point comparables à celles obtenues dans le cas bidimensionnel du chapitre précédant (voir les figures 4.15 et 4.16 des pages 67 et 68).

La recherche d'un indicateur, pour la détermination de β_{opt} , comparable à celui de l'étude bidimensionnelle (voir section V. en page 69) n'a pas encore été menée. L'indicateur du cas bidimensionnel n'a d'ailleurs pas encore été testé.

Les écarts présents dans les tableaux 5.6 et 5.7 sont définis par :

$$\sigma_{3S-M,max}(\%) = \max_t (\max_{\Omega} (abs(T_{MR_{3S}} - T_{MD_M}))) \frac{100}{\Delta T_{max}} \quad (5.38)$$

$$\sigma_{3S-M,moy}(\%) = \frac{\int_t \int_{\Omega} abs(T_{MR_{3S}} - T_{MD_M}) d\Omega dt}{\int_t dt \int_{\Omega} d\Omega} \frac{100}{\Delta T_{max}} \quad (5.39)$$

Ordre global	50	100	150	200	250	300
$\tilde{N}_1^{(1)}$	40	40	50	80	90	100
$\tilde{N}_1^{(2)}$	5	30	50	60	80	100
$\tilde{N}_1^{(3)}$	5	30	50	60	80	100
$\sigma_{3S-M,max}(\%)$	52.85	58.71	58.62	58.56	58.32	57.95
$\sigma_{3S-M,moy}(\%)$	35.71	41.28	41.01	41.74	41.94	40.97
T_{CPU} (s)	41	47	58	81	105	137

TAB. 5.6 – Comparaison MR_{3S} (sans fonctionnelle de saut de flux) et MD_M

Ordre global	50	100	150	200	250	300
$\tilde{N}_1^{(1)}$	40	40	50	80	90	100
$\tilde{N}_1^{(2)}$	5	30	50	60	80	100
$\tilde{N}_1^{(3)}$	5	30	50	60	80	100
β_{opt}	34.0	35.0	35.5	36.0	35.8	36
$\sigma_{3S-M,max}(\%)$	6.42	5.00	4.33	4.70	4.57	4.57
$\sigma_{3S-M,moy}(\%)$	1.57	0.86	0.71	0.71	0.68	0.73
T_{CPU} (s)	45	54	66	93	120	154

TAB. 5.7 – Comparaison MR_{3S} (avec un paramètre de pondération (β) optimal pour la fonctionnelle de saut de flux) et MD_M

Note : les temps de calcul affichés dans le tableau 5.6 sont identiques à ceux du tableau 5.5. Ceci malgré un nombre de nœuds légèrement différent.

En pratique dans le cas de la sous-structuration, le temps de calcul est fortement lié au temps pris pour la création des opérateurs de projection et des fonctionnelles de couplage. Ces trois temps dépendent du nombre de nœuds présents sur l'interface et non du nombre de nœuds présents dans le maillage.

Le code a été implémenté pour le cas général où les interfaces sont non-conformes. Il serait possible de gagner quelques secondes sur les simulations des modèles MD_{3E} et MR_{3E} en écrivant une table de correspondance des nœuds¹¹ (entre les nœuds de surfaces des deux maillages d'une même interface). Il a été fait le choix de rester le plus général possible et de ne pas coder cette table de correspondance.

De plus, le temps mis pour résoudre le système matriciel ne dépend que de la taille de celui-ci (donc l'ordre global du modèle). En effet, le solveur matriciel utilisé en modal est le solveur direct LD^tL [29]. Le nombre de pas de temps de calcul évoluant très peu, il est alors compréhensible que les temps de calcul soit très proches entre les différentes simulations (pour un même ordre global). La précision à la seconde est alors insuffisante pour différencier les temps de calcul.

Lorsque la fonctionnelle de saut de flux est introduite dans le modèle MR_{3S} (cas du tableau 5.7), le temps de calcul augmente (pour un même ordre global) car il faut déterminer les matrices qui décrivent la fonctionnelle de saut de flux.

VI. Conclusion

Nous avons montré comment traiter la conduction thermique dans un solide homogène découpé artificiellement en plusieurs parties. Ici l'exemple est tridimensionnel. Les bases étant incomplètes une correction de saut de température s'avère nécessaire même pour ce cas de contact parfait. Le coefficient de pondération de la fonctionnelle de saut de flux peut être vu comme une résistance thermique de contact fictive. On retrouve alors le formalisme déjà développé dans le chapitre 3. L'introduction de la résistance thermique induit nécessairement un biais. Il faut considérer une RTC faible pour limiter la valeur de ce biais, ce qui est logique du point de vue physique. L'évaluation *a priori* d'une valeur numérique pour cette RTC est possible suivant les idées développées en page 84.

L'introduction d'une fonctionnelle de saut de flux dans le modèle modal sous-structuré basé sur des maillages conformes ne s'avère pas nécessaire. Au contraire son utilisation dans le modèle modal sous-structuré basé sur des maillages non-conformes est une nécessité. Dans le cas où le paramètre de la fonctionnelle de saut de flux est optimal alors les résultats obtenus avec le modèle modal sous-structuré basé sur des maillages non-conformes sont comparables à ceux obtenus avec le modèle modal sous-structuré basé sur des maillages conformes.

L'ingénieur désirant utiliser des maillages non-conformes (utilisation de différents logiciels, ...) pourra donc utiliser la méthode de couplage proposée.

A ordre global donné, les résultats des modèles modaux multi-corps montrent qu'il est effectivement possible de réduire l'erreur maximale (la référence étant constituée par un modèle mono-corps, et l'erreur d'origine est celle du modèle modal mono-corps correspondant. Ceci malgré l'introduction d'une RTC fictive dans le modèle, cette RTC introduisant déjà un écart avec le modèle détaillé.

La sous-structuration permet alors d'augmenter de manière significative le gain de temps : pour un écart maximal inférieur à 5% le gain de temps est de l'ordre de 2 en mono-corps et de l'ordre de 23 en multi-corps.

¹¹En remplacement des opérateurs de projection.

Conclusion

Ces trois années de recherche ont permis d'aboutir aux objectifs fixés qui étaient de développer et d'analyser une technique de sous-structuration modale.

Un travail amont particulier a été mené pour rendre la formulation analytique de la technique de sous-structuration la plus concise possible. Nous avons aussi cherché à bien dissocier l'aspect fonctionnel de la méthode de sa transcription informatique qui passe par une discrétisation spatiale. En particulier nous n'avons imposé aux géométries des interfaces qu'un minimum de restrictions.

Nous attendions de la technique mise en œuvre des gains substantiels de temps de calcul et la réduction de l'espace mémoire nécessaire. C'est une des raisons d'être de la réduction de modèle. Cependant de nombreuses étapes sont ajoutées par rapport aux méthodes variationnelles classiques : détermination et réduction de la base modale, écriture de l'équation d'état et résolution de celle-ci, reconstruction des champs de température. Toutes ces étapes ont fait l'objet d'une attention particulière afin d'optimiser l'ensemble de la chaîne de calcul.

Le travail qui est présenté dans ce manuscrit comporte deux volets :

- Un volet analytique en partie basé sur des travaux antérieurs (menés dans le cadre de l'analyse modale classique) qui ont été largement adaptés à l'utilisation de la base de branche. C'est l'existence de la base de branche et plus particulièrement de sa condition de Steklov à la frontière qui rend possible les couplages d'interfaces multidimensionnelles complexes.
- Un volet d'expériences numériques qui montre la pertinence de l'approche adoptée. Elle permet aussi de mettre en lumière de façon pratique certains points théoriques qui peuvent apparaître délicats comme le choix des valeurs numériques des facteurs de pondération des fonctionnelles du saut de flux.

En ce qui concerne la démarche théorique :

Le chapitre 2 qui présente (partiellement) l'état de l'art en analyse modale montre les limites de l'utilisation des modes associés à des conditions limites physiques figées. Il était donc indispensable pour mettre en œuvre la sous-structuration de manière efficace d'utiliser de nouvelles conditions. C'est la raison du choix des bases de branche qui sont associées à une condition aux limites de Steklov. Le chapitre 3 développe la théorie du couplage dans une approche fonctionnelle. Une formulation variationnelle très générale, qui permet de s'affranchir de toute discrétisation spatiale est proposée.

Les modèles devant être réduits, les bases sont bien entendu incomplètes. Une correction est alors nécessaire. Elle s'opère par le biais de fonctionnelles de saut de température et de flux sur les interfaces. S'agissant de la fonctionnelle de saut de température, son apparition est naturelle lorsqu'il y a physiquement une résistance de contact entre deux corps. Le choix des fonctionnelles de saut est vaste [9]. Nous proposons dans ce mémoire des fonctionnelles dont les propriétés mathématiques garantissent une certaine cohérence avec le reste du modèle (symétrique, positive, ...).

Si l'utilisation de la base de branche promet la perspective de produire un « bon » raccord entre les éléments du système elle présente toutefois certains inconvénients. Il en est ainsi du couplage des états d'excitation des modes de branche. L'équation d'état perd le caractère diagonal strict associé aux modes classiques. On conjecture à ce stade que les dominances de certains modes produiront un découplage numérique des états, point qui sera confirmé sur des exemples pratiques. Ce couplage nécessite aussi d'adapter la technique de réduction. La technique d'amalgame propre à notre équipe de recherche a été conservée en y apportant des modifications. L'objectif était de ne pas générer un surcroît de temps de calcul. A ce stade « il ne restait plus » qu'à programmer la méthode.

En ce qui concerne la mise en œuvre de la démarche théorique :

Les travaux antérieurs menés au LMEE étaient effectués avec des logiciels commerciaux (MATLAB et FEMLAB). Ce choix avait été fait pour tester rapidement la pertinence des algorithmes et les idées théoriques sans avoir à supporter la contrainte d'un important développement logiciel. Les expériences numériques avaient alors confirmé les éléments théoriques présentés dans le chapitre 2. Cependant pour

aller plus loin, le choix des logiciels commerciaux trouvait ses limites (changement de version, codes peu ou pas documentés, interface avec nos propres programmes délicate et peu performante, ...). L'une de ces limites, la plus pénalisante, était imposées par MATLAB qui ne permettait pas d'utiliser plus de 400Mo sur les 2Go disponibles. Un comble pour la technique de sous-structuration qui ambitionne de traiter des grands systèmes qui génèrent d'importants maillages. Ainsi lorsque je suis arrivé au LMEE, il a été décidé que l'équipe¹ allait intégralement écrire un code adapté à la méthode modale de branche. Je m'y suis très largement investi. Ce code est actuellement fonctionnel. Il se charge de l'ensemble des calculs à deux exceptions près : le pré et post processeur. Autrement dit il faut fournir un maillage « éléments finis » et les résultats des calculs doivent être visualisés avec un logiciel graphique adapté². La gestion de la mémoire est désormais optimale, en tout état de cause l'intégralité de la RAM disponible sur une machine donnée peut être utilisée. Le code a servi à produire l'ensemble des simulations dont les résultats sont présentés dans ce manuscrit.

On trouvera peu d'éléments dans ce mémoire sur la mise en œuvre informatique. La lecture des 32425 lignes de code en C++ ne présente que peu d'intérêt. La description des nombreuses « astuces » de programmation pour optimiser le code risquerait de masquer l'essentiel³. Il faut toutefois en dire un mot. En effet la concision de la formulation analytique, qui se résume dans ce manuscrit à quelques pages très denses, contraste avec le temps passé à sa mise en œuvre numérique. La transcription de l'algorithme en code informatique a représenté une bonne part de ces trois années de thèse. Mais cette étape qui demande de la diplomatie⁴ lors du choix du langage, qui est parfois ingrate lorsqu'il s'agit de « déboguer » le code ou qui nécessite beaucoup de coordination entre les chercheurs puisque écrit à plusieurs, est absolument indispensable pour passer de l'idée à l'application. Sans elle l'étude des exemples numériques des chapitres 4 et 5 n'existerait pas.

En ce qui concerne la pertinence de notre approche :

La démarche théorique et sa mise en œuvre informatique doivent être validées sur des cas d'application pratiques. Ceci fait l'objet des chapitres 4 et 5 de ce mémoire. Nous avons montré qu'il est possible de traiter les couplages d'un système constitué de plusieurs corps en contact thermique imparfait à l'aide de la sous-structuration modale de branche. Les résultats obtenus montrent la pertinence de la méthode de couplage proposée en termes de précision et de gain de temps. Notre approche pouvait présenter cependant un inconvénient pour l'utilisateur : elle introduit un paramètre de pénalisation du saut de flux dont le choix est délicat. Une étude menée *a posteriori*, sur les cas que nous avons traités, a montré qu'il existe une plage de valeurs optimum pour ce paramètre. Pour que la méthode soit véritablement opérationnelle, nous avons proposé un indicateur permettant sa détermination *a priori*. La recherche de cet indicateur présente l'avantage de n'utiliser que le modèle modal réduit. Son calcul est donc très rapide. Il a été testé avec succès sur le cas traité en bidimensionnel (chapitre 4).

Le chapitre 5 développe sur l'exemple support d'une géométrie tridimensionnelle. L'utilisation de la méthode pour séparer un modèle homogène de grande taille en plusieurs sous-modèles de taille plus réduite. Il faut pour cela admettre un écart dans la modélisation au niveau des coupures fictives (interfaces de contact entre les différents sous-domaines). La fonctionnelle de saut de température est alors déterminée par l'utilisateur qui choisit la valeur de la résistance thermique de contact (virtuelle) qu'il juge admissible. L'évaluation *a priori* d'une valeur numérique pour cette *RTC* est possible suivant les idées développées en page 84. En tridimensionnel, si les maillages sont adaptés alors la fonctionnelle de saut de flux a peu d'effet. La méthode est alors plus efficace qu'une méthode modale mono-corps classique. Si nous nous plaçons du point de vue d'une utilisation dans le milieu industriel, il est fort probable que les maillages constituant les différents éléments du système ne soient pas concordants au niveau des frontières de raccordement. Dans ce cas l'interpolation apportée par les opérateurs de projection engendre un écart complémentaire à celui lié à l'introduction d'une résistance thermique de contact et à la réduction modale. Une fonctionnelle de saut de flux permet alors d'améliorer les résultats, pour finalement obtenir une précision équivalente au cas des maillages conformes.

¹Quatre enseignants chercheurs (1PU, 3MCF) plus moi même.

²Au sein du LMEE une mise en commun de ce type d'outils est envisagée avec l'équipe de Modélisation en Dynamique des Structures (Z.Q. Feng).

³Ces « astuces » contribuent toutefois pour beaucoup à ce que le gain de temps de calcul théorique escompté grâce à la réduction modale et à la sous-structuration ne soient pas annihilées par une mise en œuvre informatique sans finesse.

⁴Chaque programmeur ayant ses propres habitudes : C, C++, Fortran,...

Un bref bilan et quelques perspectives :

Les travaux effectués demandent bien sûr à être confirmés sur d'autres géométries. Il serait souhaitable de conforter la méthode de choix *a priori* du paramètre de la fonctionnelle de saut de flux sur de nouveaux exemples. Des progrès sont encore possibles pour réduire les modèles par amalgame. Le traitement d'enceintes thermiques devrait être envisagé. Ici les aspects du rayonnement entre parois devront être pris en compte et imbriqués dans le processus de réduction et de sous-structuration. Des tentatives analytiques ont été tentées mais le chemin est encore long avant de les voir se concrétiser dans le code de calcul.

La technique est cependant d'ores et déjà utilisable en conduction multi-corps bi ou tridimensionnel. Elle autorise le traitement effectif de systèmes comportant un très grand nombre d'éléments et de nœuds de discrétisation. Le code développé est compatible avec l'utilisation d'un ordinateur domestique actuel pour peu qu'il dispose d'assez de RAM (2Go et plus). L'origine des maillages est libre car il n'est pas nécessaire qu'ils soient conformes au niveau des interfaces de contact. L'utilisation de maillages générés par différents logiciels industriels est donc possible. La méthode apporte un gain de temps significatif pour les exemples traités, ouvrant ainsi d'intéressantes perspectives en identification et en commande temps réel.

Annexes

Théorèmes de Green-Ostrogradsky

Le théorème de Green-Ostrogradsky est bien connu des physiciens au moins dans le cas des fonctions de $H^1(\Omega)$. Il porte différents noms : théorème de Gauss, théorème de la divergence, théorème de Green... Cette annexe présente de manière très simple, les différents théorèmes de Green-Ostrogradsky utilisés.

I. Cas de l'espace $H^1(\Omega)$

Soit un ouvert borné Ω de frontière $\partial\Omega$ et de normale extérieure unitaire $\underline{n}_{\partial\Omega}$:

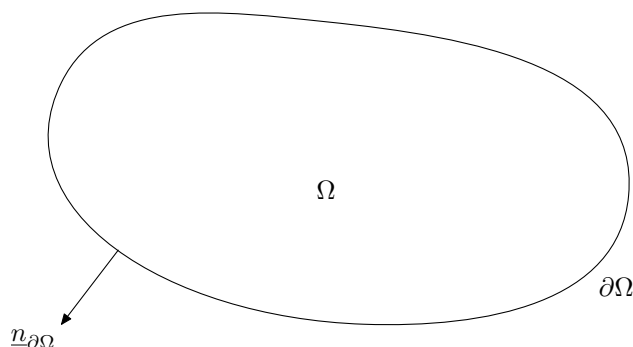


FIG. A.1 – Domaine Ω avec sa frontière et sa normale

Le théorème de Green-Ostrogradsky le plus simple est alors :

$$\forall \underline{u} \in (H^1(\Omega))^3 \quad \int_{\Omega} \operatorname{div}(\underline{u}) \, d\Omega = \int_{\partial\Omega} \underline{u} \cdot \underline{n}_{\partial\Omega} \, d(\partial\Omega) \quad (\text{A.1})$$

Il existe de très nombreuses variantes du théorème A.1. La version qui nous intéresse étant le théorème A.4. Pour l'établir, il faut utiliser la relation suivante :

$$\forall \underline{u} \in (H^1(\Omega))^3 \quad \forall v \in H^1(\Omega) \quad \operatorname{div}(v\underline{u}) = v \operatorname{div}(\underline{u}) + (\nabla(v)) \bullet \underline{u} \quad (\text{A.2})$$

A.2 est intégré sur le domaine Ω et les termes sont réorganisés :

$$\forall \underline{u} \in (H^1(\Omega))^3 \quad \forall v \in H^1(\Omega) \quad \int_{\Omega} \operatorname{div}(\underline{u}) v \, d\Omega = \int_{\Omega} \operatorname{div}(v\underline{u}) \, d\Omega - \int_{\Omega} (\nabla(v)) \bullet \underline{u} \, d\Omega \quad (\text{A.3})$$

A.3 est modifié à l'aide de A.1, pour obtenir le théorème de Green-Ostrogradsky qui sera utilisé pour établir les formulations variationnelles si les fonctions appartiennent à l'espace $H^1(\Omega)$.

$$\forall \underline{u} \in (H^1(\Omega))^3 \quad \forall v \in H^1(\Omega) \quad \int_{\Omega} \operatorname{div}(\underline{u}) v \, d\Omega = - \int_{\Omega} \underline{u} \bullet \underline{\nabla}(v) \, d\Omega + \int_{\partial\Omega} v \underline{u} \bullet \underline{n}_{\partial\Omega} \, d(\partial\Omega) \quad (\text{A.4})$$

II. Cas de l'espace $H^1(\Omega, \Gamma)$

Les théorèmes A.1 et A.4 ne sont valables que pour l'espace $H^1(\Omega)$, l'introduction d'une résistance de contact introduit une surface de discontinuité Γ . Il faut alors disposer d'un théorème de Green-Ostrogradsky pour l'espace $H^1(\Omega, \Gamma)$. Ce théorème est peu présent dans les ouvrages de mathématiques, il est surtout présenté dans des ouvrages de physique [8].

Avec les notations du paragraphe II.1. (page 37), le théorème de Green-Ostrogradsky A.1 réécrit pour l'espace $H^1(\Omega, \Gamma)$ est :

$$\forall \underline{u} \in (H^1(\Omega, \Gamma))^3 \quad \int_{\Omega} \operatorname{div}(\underline{u}) \, d\Omega = \int_{\partial\Omega} \underline{u} \bullet \underline{n}_{\partial\Omega} \, d(\partial\Omega) - \int_{\Gamma} \llbracket \underline{u} \rrbracket_{\Gamma} \cdot \underline{n}_{\Gamma} \, d\Gamma \quad (\text{A.5})$$

Dans cette relation $\llbracket \underline{u} \rrbracket_{\Gamma}$ représente le saut de \underline{u} au niveau de la surface Γ dans le sens de la normale à $\Gamma : \underline{n}_{\Gamma}$

A.5 se démontre à partir de A.1, la démarche est explicitée dans le cas où le domaine Ω est constitué de deux sous-domaines $\Omega^{(1)}$ et $\Omega^{(2)}$.

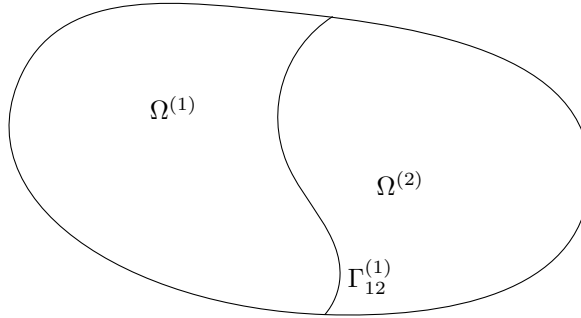


FIG. A.2 – Domaine Ω constitué de deux sous-domaines $\Omega^{(1)}$ et $\Omega^{(2)}$

Il faut commencer par écrire A.1 pour chacun des sous-domaines :

$$\forall \underline{u} \in \left(H^1 \left(\Omega^{(1)} \right) \right)^3 \quad \int_{\Omega^{(1)}} \operatorname{div}(\underline{u}) \, d\Omega^{(1)} = \int_{\partial\Omega^{(1)}} \underline{u} \bullet \underline{n}^{(1)} \, d \left(\partial\Omega^{(1)} \right) \quad (\text{A.6})$$

$$\forall \underline{u} \in \left(H^1 \left(\Omega^{(2)} \right) \right)^3 \quad \int_{\Omega^{(2)}} \operatorname{div}(\underline{u}) \, d\Omega^{(2)} = \int_{\partial\Omega^{(2)}} \underline{u} \bullet \underline{n}^{(2)} \, d \left(\partial\Omega^{(2)} \right) \quad (\text{A.7})$$

Il faut alors remarquer que la frontière de contact $\Gamma_{12}^{(1)}$ est égale à l'intersection de $\partial\Omega^{(1)}$ et de $\partial\Omega^{(2)}$: $\Gamma_{12}^{(1)} = (\partial\Omega^{(1)}) \cap (\partial\Omega^{(2)})$

On sépare les frontières des domaines en deux parties : $\partial\Omega^{(1)} = \left(\partial\Omega^{(1)} \setminus \Gamma_{12}^{(1)} \right) + \Gamma_{12}^{(1)}$ et $\partial\Omega^{(2)} = \left(\partial\Omega^{(2)} \setminus \Gamma_{12}^{(1)} \right) + \Gamma_{12}^{(1)}$. A.6 et A.7 deviennent alors respectivement A.8 et A.9.

$$\forall \underline{u} \in \left(H^1 \left(\Omega^{(1)} \right) \right)^3 \quad \int_{\Omega^{(1)}} \operatorname{div}(\underline{u}) \, d\Omega^{(1)} = \int_{\partial\Omega^{(1)} \setminus \Gamma_{12}^{(1)}} \underline{u} \bullet \underline{n}^{(1)} \, d(\partial\Omega^{(1)}) + \int_{\Gamma_{12}^{(1)}} \underline{u} \bullet \underline{n}^{(1)} \, d\Gamma_{12}^{(1)} \quad (\text{A.8})$$

$$\forall \underline{u} \in \left(H^1 \left(\Omega^{(2)} \right) \right)^3 \quad \int_{\Omega^{(2)}} \operatorname{div}(\underline{u}) \, d\Omega^{(2)} = \int_{\partial\Omega^{(2)} \setminus \Gamma_{12}^{(1)}} \underline{u} \bullet \underline{n}^{(2)} \, d(\partial\Omega^{(2)}) + \int_{\Gamma_{12}^{(1)}} \underline{u} \bullet \underline{n}^{(2)} \, d\Gamma_{12}^{(1)} \quad (\text{A.9})$$

Ensuite il faut sommer A.8 et A.9, en remarquant que $\Omega = \Omega^{(1)} + \Omega^{(2)}$ et que $\partial\Omega = \left(\partial\Omega^{(1)} \setminus \Gamma_{12}^{(1)} \right) + \left(\partial\Omega^{(2)} \setminus \Gamma_{12}^{(1)} \right)$, tout en précisant $\underline{u}^{(1)}$ (ou $\underline{u}^{(2)}$) lorsque cela est nécessaire :

$$\begin{aligned} \forall \underline{u} \in \left(H^1(\Omega) \right)^3 \quad \int_{\Omega} \operatorname{div}(\underline{u}) \, d\Omega = \\ \int_{\partial\Omega} \underline{u} \bullet \underline{n}^{(2)} \, d(\partial\Omega) + \int_{\Gamma_{12}^{(1)}} \underline{u}^{(1)} \bullet \underline{n}^{(1)} \, d\Gamma_{12}^{(1)} + \int_{\Gamma_{12}^{(1)}} \underline{u}^{(2)} \bullet \underline{n}^{(2)} \, d\Gamma_{12}^{(1)} \end{aligned} \quad (\text{A.10})$$

Il faut choisir l'orientation de \underline{n}_{Γ} : $\underline{n}_{\Gamma} = \underline{n}^{(1)}$, dans ce cas $[[\underline{u}]]_{\Gamma} = \underline{u}^{(2)} - \underline{u}^{(1)}$ (pour mémoire $n^{(2)} = -n^{(1)}$). Ce qui établit le théorème A.5, qui serait aussi obtenu avec le choix $\underline{n}_{\Gamma} = \underline{n}^{(2)}$ (et dans ce cas $[[\underline{u}]]_{\Gamma} = \underline{u}^{(1)} - \underline{u}^{(2)}$).

Le théorème recherché pour l'espace $H^1(\Omega, \Gamma)$ est obtenu en utilisant A.2 dans A.5 :

$$\begin{aligned} \forall \underline{u} \in \left(H^1(\Omega, \Gamma) \right)^3 \quad \forall v \in H^1(\Omega, \Gamma) \quad \int_{\Omega} \operatorname{div}(\underline{u}) \, v \, d\Omega = \\ - \int_{\Omega} (\nabla(v)) \bullet \underline{u} \, d\Omega + \int_{\partial\Omega} v \underline{u} \bullet \underline{n}_{\partial\Omega} \, d(\partial\Omega) - \int_{\Gamma} [[v\underline{u}]]_{\Gamma} \cdot \underline{n}_{\Gamma} \, d\Gamma \end{aligned} \quad (\text{A.11})$$

I. Le théorème spectral

Le théorème spectral est présenté dans un grand nombre d'ouvrages mathématiques. Il est cependant recommandé de consulter l'ouvrage [1] qui a été écrit pour des ingénieurs et qui est plus accessible que [15].

Le théorème spectral à utiliser est le théorème 7.3.2 page 219 de [1] :

Soit le problème variationnel de valeurs propres suivant (ou problème spectral) : trouver $\lambda \in \mathbb{R}$ et $u \in V \setminus \{0\}$ tels que

$$a(u, v) = \lambda \langle u, v \rangle_H \quad \forall v \in V \quad (\text{B.1})$$

Soit V et H deux espaces de Hilbert réels de dimension infinie. On suppose que $V \subset H$ avec injection compacte et que V est dense dans H . Soit $a(., .)$ une forme bilinéaire symétrique continue et coercive sur V . Alors les valeurs propres de B.1 forment une suite croissante $(\lambda_k)_{k \geq 1}$ de réels positifs qui tend vers l'infini, et il existe une base hilbertienne de H $(u_k)_{k \geq 1}$ de vecteurs propres associés, c'est-à-dire que

$$u_k \in V, \text{ et } a(u_k, v) = \lambda_k \langle u_k, v \rangle_H \quad \forall v \in V \quad (\text{B.2})$$

De plus, $\left(\frac{u_k}{\sqrt{\lambda_k}} \right)_{k \geq 1}$ est une base hilbertienne de V pour le produit scalaire $a(., .)$

II. Exemple d'utilisation du théorème spectral

L'utilisation rigoureuse du théorème spectral est très délicate. En grande partie à cause de la nécessité d'avoir une injection compacte de V dans H . Le théorème de Rellich permet de démontrer que l'injection de $H^1(\Omega)$ dans $L^2(\Omega)$ est compacte. Cependant les théorèmes de Green ne sont pas valides dans $L^2(\Omega)$. A la demande d'Alain Neveu, Frédéric Bourquin a effectué l'ensemble des démarches qui permettent d'utiliser le théorème spectral pour le problème des modes branche. Son texte est inclus dans les 3 pages suivantes, ce qui rend la paternité de la démonstration à son auteur.

Cela démontre que les modes de branche forment une base de $H^1(\Omega)$. De la même manière, il serait démontré que les modes de Fourier forment une base de $H_F^1(\Omega)$.

More on branch modes for heat transfer

Frédéric Bourquin
 French Public Works Research Laboratory, and Laboratoire Lagrange
 58 bd Lefebvre, 75015 Paris
 e-mail: bourquin@lcpc.fr

June 19, 2006

The so-called *branch modes* solve the Steklov eigenvalue problem

$$\begin{cases} -\Delta u + \alpha u = \lambda u & \omega \\ \frac{\partial u}{\partial \nu} + \alpha \rho u = \lambda \rho u & \gamma \end{cases} \quad (1)$$

here $\gamma = \partial\omega$, but this is no limitation on this point.

A tentative variational formulation of 1 would write

$$\int_{\omega} \nabla u \nabla v + \alpha uv + \int_{\gamma} \alpha \rho uv = \lambda \left(\int_{\omega} uv + \int_{\gamma} \rho uv \right) \quad (2)$$

let $u^{\gamma} = u|_{\gamma}$ and $\bar{u} = u - \mathcal{R}u^{\gamma}$, where \mathcal{R} stands for the harmonic extension operator in the sense of

$$\begin{cases} -\Delta \mathcal{R}u + \alpha \mathcal{R}u = 0 & \omega \\ \mathcal{R}u = u & \gamma \end{cases} \quad (3)$$

then 2 writes

$$\int_{\omega} \nabla \bar{u} \nabla \bar{v} + \alpha \bar{u} \bar{v} + \int_{\omega} \nabla \mathcal{R}u^{\gamma} \nabla \mathcal{R}v^{\gamma} + \alpha \mathcal{R}u^{\gamma} \mathcal{R}v^{\gamma} + \int_{\gamma} \alpha \rho u^{\gamma} v^{\gamma} = \lambda \left(\int_{\omega} (\bar{u} + \mathcal{R}u^{\gamma})(\bar{v} + \mathcal{R}v^{\gamma}) + \int_{\gamma} \rho u^{\gamma} v^{\gamma} \right) \quad (4)$$

which is of the standard form by just setting

$$\begin{aligned} a((\bar{u}, u^{\gamma}), (\bar{v}, v^{\gamma})) &= \int_{\omega} \nabla \bar{u} \nabla \bar{v} + \alpha \bar{u} \bar{v} + \int_{\omega} \nabla \mathcal{R}u^{\gamma} \nabla \mathcal{R}v^{\gamma} + \alpha \mathcal{R}u^{\gamma} \mathcal{R}v^{\gamma} + \int_{\gamma} \alpha \rho u^{\gamma} v^{\gamma} \\ b((\bar{u}, u^{\gamma}), (\bar{v}, v^{\gamma})) &= \int_{\omega} (\bar{u} + \mathcal{R}u^{\gamma})(\bar{v} + \mathcal{R}v^{\gamma}) + \int_{\gamma} \rho u^{\gamma} v^{\gamma} \end{aligned} \quad (5)$$

Ellipticity and continuity follow from standard arguments, and in particular from the regularity theory of elliptic equations. For the sake of completeness, an extended proof is given below.

Let us start with the bilinear form $a(\cdot, \cdot)$. is continuous on the space $V = H_0^1(\omega) \times H^{1/2}(\gamma)$ due to the usual trace theorem. It is coercive over the space V because it is bounded from below by the usual scalar product on the space V . Moreover, the quantity $\int_{\omega} \nabla \mathcal{R}u^{\gamma} \nabla \mathcal{R}v^{\gamma} + \alpha \mathcal{R}u^{\gamma} \mathcal{R}v^{\gamma}$ is equivalent to the standard norm over the space $H^{1/2}(\gamma)$.

As for the form $b(\cdot, \cdot)$, the right space involved is now $H = L^2(\omega) \times L^2(\gamma)$. Its continuity over H directly follows from the trace theorem in $H^1(\omega)$ and triangular inequality.

As for its coerciveness, we need the following lemma ensuring the continuity of the operator $\mathcal{R} : L^2(\gamma) \longrightarrow L^2(\omega)$. Actually we have much more (compactness).

Lemma 1 : *The operator $\mathcal{R} : L^2(\gamma) \longrightarrow L^2(\omega)$ is continuous whenever the solution w to the Dirichlet problem over ω*

$$\begin{aligned} -\nabla^2 w &= f & \omega \\ w &= 0 & \partial\omega \end{aligned} \quad (6)$$

satisfies $w \in H^{\frac{3}{2}+\epsilon}(\omega)$ whenever $f \in L^2(\omega)$.

Proof : Let $f \in L^2(\omega)$ denote an arbitrary function and consider a function $v \in H^{\frac{1}{2}}$. Then $\mathcal{R}v$ is defined in the usual sense as a function of the space $H^1(\omega)$. Moreover, a simple integration by parts shows that

$$\int_{\omega} \mathcal{R}v f = - \int_{\gamma} v \frac{\partial w}{\partial \vec{n}} \quad (7)$$

if \vec{n} denotes the unit outer normal vector to ω along $\partial\omega$.

Let

$$L(f) = - \int_{\gamma} v \frac{\partial w}{\partial \vec{n}} \quad (8)$$

This is clearly a linear form which is well defined for every $f \in L^2(\omega)$ because of the regularity assumption. Moreover, we have

$$|L(f)| \leq C \|v\|_{L^2(\gamma)} \|w\|_{H^{\frac{3}{2}+\epsilon}(\omega)} \quad (9)$$

where C denotes a generic positive constant.

Because of the variational continuity of w with respect to f and of the regularity assumption, the mapping $f \longrightarrow w$ is continuous from $L^2(\omega)$ to $H^{\frac{3}{2}+\epsilon}(\omega)$ due to the closed-graph theorem. Therefore, it follows from the previous estimate that

$$|L(f)| \leq C \|v\|_{L^2(\gamma)} \|f\|_{L^2(\omega)} \quad (10)$$

Therefore, using $\|\mathcal{R}v\|_{L^2(\omega)} = \sup_{f \in L^2(\omega), \|f\|_{L^2(\omega)}=1} \int_{\omega} \mathcal{R}v f$, we get

$$\|\mathcal{R}v\|_{L^2(\omega)} \leq C \|v\|_{L^2(\gamma)} \quad (11)$$

Since $H^{\frac{1}{2}}(\gamma)$ is densely embedded in $L^2(\gamma)$, the operator \mathcal{R} can be extended to a continuous operator from $L^2(\gamma)$ into $L^2(\omega)$.

QED.

We are now in a position to prove the coerciveness of $b(\cdot, \cdot)$. A simple calculus shows that

$$b((\bar{v}, v^\gamma), (\bar{v}, v^\gamma)) = \int_{\omega} \bar{v}^2 + \int_{\omega} (\mathcal{R}v^\gamma)^2 + \int_{\gamma} \rho(v^\gamma)^2 + 2 \int_{\omega} \bar{v} \mathcal{R}v^\gamma \quad (12)$$

hence, from Schwarz inequality,

$$b((\bar{v}, v^\gamma), (\bar{v}, v^\gamma)) \geq \int_\omega \bar{v}^2 + \int_\omega (\mathcal{R}v^\gamma)^2 + \int_\gamma \rho(v^\gamma)^2 - 2\sqrt{\int_\omega \bar{v}^2} \sqrt{\int_\omega (\mathcal{R}v^\gamma)^2} \quad (13)$$

Now let

$$a = \int_\omega \bar{v}^2 \quad b = \int_\gamma \rho(v^\gamma)^2 \quad d = \int_\omega (\mathcal{R}v^\gamma)^2 \quad (14)$$

and let $Q = b((\bar{v}, v^\gamma), (\bar{v}, v^\gamma))$. What follows aims at establishing that $Q \geq C(a^2 + b^2)$ for some positive constant C .

Thus 13 writes

$$Q \geq a^2 + d^2 + b^2 - 2ad \quad (15)$$

First recall the continuity of the operator $\mathcal{R} : L^2(\gamma) \longrightarrow L^2(\omega)$ if the boundary of the domain ω is smooth enough. Hence, $d \leq \alpha b$ for some positive α . From this estimate and from the obvious inequality $|ad| \leq \frac{1}{2} \left(\epsilon a^2 + \frac{d^2}{\epsilon} \right)$ for every positive ϵ , it follows that

$$Q \geq a^2(1 - \epsilon) + b^2 + d^2(1 - \frac{1}{\epsilon}) \quad (16)$$

and, for $\epsilon \leq 1$,

$$Q \geq a^2(1 - \epsilon) + b^2(1 + \alpha(1 - \frac{1}{\epsilon})) \quad (17)$$

Therefore, choosing $\epsilon = \frac{\alpha}{\alpha + \frac{1}{2}}$, it follows that $1 + \alpha(1 - \frac{1}{\epsilon}) = \frac{1}{2}$ and $1 - \epsilon = \frac{\frac{1}{2}}{\frac{1}{2} + \alpha}$, hence

$$Q \geq \min\left(\frac{1}{2}, \frac{\frac{1}{2}}{\frac{1}{2} + \alpha}\right)(a^2 + b^2) \quad (18)$$

QED.

Because the space V is continuously and compactly embedded into the space H , the standard abstract theory of self-adjoint compact operators shows that problem 4 possesses a countable family of solutions $(\lambda_k, \bar{u}_k, u_k^\gamma) \in \mathbb{R}^+ \times H_0^1(\omega) \times H^{1/2}(\gamma)$ such that the family (\bar{u}_k, u_k^γ) forms an orthogonal basis of both spaces $H_0^1(\omega) \times H^{1/2}(\gamma)$ and $L^2(\omega) \times L^2(\gamma)$. The orthogonality holds with respect to the scalar products $a(\cdot)$ and $b(\cdot)$.

Branch modes are the fields $\bar{u}_k + \mathcal{R}u_k^\gamma$. They do satisfy 2 as well as 1 in a weak sense.

Contenu du chapitre

I.	Présentation du problème	110
II.	Mise sous forme adimensionnelle	111
III.	Hypothèse des variables séparables	111
IV.	Recherche des couples $(f_n(x^*), g_n(y^*))$	112
V.	Détermination des constantes de $A_n, B_n, C_n, D_n, \rho_n$	112
VI.	Les fonctions sont orthogonales entre-elles	113
VII.	Détermination de β_n	114
VIII.	Complément I : Méthode graphique pour l'équation transcendante	115
IX.	Complément II : démonstration de l'orthogonalité C.36	116
X.	Complément III : Détail du calcul de β_n	117

La méthode de séparation des variables à fait l'objet de quelques lignes dans le chapitre 2. Cette annexe explicite la démarche sur un exemple. Il s'agit d'une volonté pédagogique vis-à-vis des lecteurs les plus jeunes qui n'ont pas forcément étudié cette méthode.

I. Présentation du problème

Il s'agit de chercher le régime permanent d'une plaque homogène de conductivité k constante :

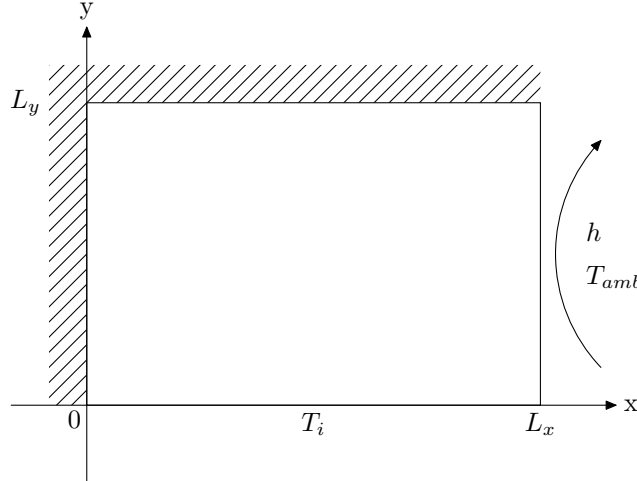


FIG. C.1 – Plaque de $[0, L_x] \times [0, L_y]$

Les équations du problème thermique sont :

$$\forall (x, y) \in [0, L_x] \times [0, L_y] \quad \Delta T(x, y) = 0 \quad (\text{C.1})$$

$$\forall x \in [0, L_x] \quad T(x, 0) = T_i \quad (\text{C.2})$$

$$\forall x \in [0, L_x] \quad \frac{\partial T}{\partial y}(x, L_y) = 0 \quad (\text{C.3})$$

$$\forall y \in [0, L_y] \quad \frac{\partial T}{\partial x}(0, y) = 0 \quad (\text{C.4})$$

$$\forall y \in [0, L_y] \quad -k \frac{\partial T}{\partial x}(L_x, y) = h(T(L_x, y) - T_{amb}) \quad (\text{C.5})$$

L'équation de la chaleur se limite alors à C.1 puisqu'il s'agit de trouver un régime permanent, les conditions aux limites sont de trois types :

- La condition C.2 est une condition de température imposée ou condition de Dirichlet
- C.3 et C.4 sont des conditions de flux imposé (ici nul) ou condition de Neumann
- C.5 est une condition de Fourier

La température imposée T_i est telle que $T_i < T_{amb}$, T_{amb} étant la température de l'ambiance. Il n'y a pas de condition initiale étant donné qu'il s'agit de la recherche d'un régime permanent.

II. Mise sous forme adimensionnelle

Le problème est adimensionné les nouvelles variables x^*, y^* et θ sont telles que :

$$x^* = \frac{x}{L_x} \quad (\text{C.6})$$

$$y^* = \frac{y}{L_y} \quad (\text{C.7})$$

$$\theta = \frac{T - T_{amb}}{T_i - T_{amb}} \quad (\text{C.8})$$

Les équations C.1-C.5 deviennent alors :

$$\forall (x^*, y^*) \in [0, 1] \times [0, 1] \quad \frac{1}{L_x^2} \frac{\partial^2 \theta}{\partial (x^*)^2} + \frac{1}{L_y^2} \frac{\partial^2 \theta}{\partial (y^*)^2} = 0 \quad (\text{C.9})$$

$$\forall x^* \in [0, 1] \quad \theta(x^*, 0) = 1 \quad (\text{C.10})$$

$$\forall x^* \in [0, 1] \quad \frac{\partial \theta}{\partial y^*}(x^*, 1) = 0 \quad (\text{C.11})$$

$$\forall y^* \in [0, 1] \quad \frac{\partial \theta}{\partial x^*}(0, y^*) = 0 \quad (\text{C.12})$$

$$\forall y^* \in [0, 1] \quad \frac{\partial \theta}{\partial x^*}(1, y^*) + Bi\theta(1, y^*) = 0 \quad (\text{C.13})$$

Le nombre de Biot présent dans l'équation C.13 est défini par :

$$Bi = \frac{hL_x}{k} \quad (\text{C.14})$$

III. Hypothèse des variables séparables

Les variables sont supposées séparables, il faut alors rechercher les couples de fonctions $(f_n(x^*), g_n(y^*))$ dont le produit $h_n(x^*, y^*) = f_n(x^*)g_n(y^*)$ vérifie C.9. Les fonctions trouvées devront vérifier les conditions aux limites C.11, C.12 et C.13.

La condition C.10 ne sera vérifiée que par la décomposition suivante :

$$\theta(x^*, y^*) = \sum_{n=1}^{\infty} \alpha_n f_n(x^*) g_n(y^*) \quad (\text{C.15})$$

Ce qui permettra de calculer les α_n

IV. Recherche des couples $(f_n(x^*), g_n(y^*))$

$h_n(x^*, y^*) = f_n(x^*)g_n(y^*)$ vérifie C.9 donc :

$$\forall (x^*, y^*) \in [0, 1] \times [0, 1] \quad \frac{1}{L_x^2} f_n''(x^*) g_n(y^*) + \frac{1}{L_y^2} f_n(x^*) g_n''(y^*) = 0 \quad (\text{C.16})$$

En supposant que $f_n(x^*)$ et que $g_n(y^*)$ ne s'annulent pas, l'expression C.16 peut s'écrire :

$$\forall (x^*, y^*) \in [0, 1] \times [0, 1] \quad -\frac{1}{L_x^2} \frac{f_n''(x^*)}{f_n(x^*)} = \frac{1}{L_y^2} \frac{g_n''(y^*)}{g_n(y^*)} \quad (\text{C.17})$$

Ainsi une fonction ne dépendant que de x^* est égale à une fonction ne dépendant que de y^* . Il n'existe qu'une possibilité, les deux fonctions sont égales à une constante.

On pose alors ρ_n tel que :

$$\forall (x^*, y^*) \in [0, 1] \times [0, 1] \quad -\frac{1}{L_x^2} \frac{f_n''(x^*)}{f_n(x^*)} = \frac{1}{L_y^2} \frac{g_n''(y^*)}{g_n(y^*)} = \rho_n^2 \quad (\text{C.18})$$

Remarque : il est aussi possible de supposer que la constante est négative et dans ce cas, il faut écrire $-\rho_n^2$ à la place de ρ_n^2 dans C.18.

Dans ces conditions $f_n(x^*)$ et $g_n(y^*)$ vérifient :

$$\forall x^* \in [0, 1] \quad f_n''(x^*) + (L_x \rho_n)^2 f_n(x^*) = 0 \quad (\text{C.19})$$

$$\forall y^* \in [0, 1] \quad g_n''(y^*) - (L_y \rho_n)^2 g_n(y^*) = 0 \quad (\text{C.20})$$

Les équations différentielles C.19 et C.20 sont classiques elles admettent les solutions générales suivantes :

$$\forall x^* \in [0, 1] \quad f_n(x^*) = A_n \sin(L_x \rho_n x^*) + B_n \cos(L_x \rho_n x^*) \quad (\text{C.21})$$

$$\forall y^* \in [0, 1] \quad g_n(y^*) = C_n \sinh(L_y \rho_n y^*) + D_n \cosh(L_y \rho_n y^*) \quad (\text{C.22})$$

V. Détermination des constantes de $A_n, B_n, C_n, D_n, \rho_n$

Les constantes $A_n, B_n, C_n, D_n, \rho_n$ sont déterminées à l'aide des conditions aux limites.

A_n doit être nul. En effet si l'on impose que $h_n(x^*, y^*) = f_n(x^*)g_n(y^*)$ vérifie C.12, alors :

$$\frac{\partial f_n(x^*) g_n(y^*)}{\partial x^*}(0, y^*) = A_n L_x \rho_n g_n(y^*) \quad (\text{C.23})$$

Or $g_n(y^*)$ n'est pas nulle sur $[0, 1]$. Donc $A_n = 0$.

C.11 implique une relation entre C_n et D_n , en effet :

$$\frac{\partial f_n(x^*) g_n(y^*)}{\partial y^*}(x^*, 1) = f_n(x^*) [C_n L_y \rho_n \cosh(L_y \rho_n) + D_n L_y \rho_n \sinh(L_y \rho_n)] \quad (\text{C.24})$$

Or $f_n(x^*)$ n'est pas nulle sur $[0, 1]$. Donc :

$$C_n L_y \rho_n \cosh(L_y \rho_n) + D_n L_y \rho_n \sinh(L_y \rho_n) = 0 \quad (\text{C.25})$$

Soit après simplification : $C_n = -D_n \tanh(L_y \rho_n)$

Ce qui permet de simplifier l'écriture de $g_n(y^*)$:

$$g_n(y^*) = C_n \sinh(L_y \rho_n y^*) + D_n \cosh(L_y \rho_n y^*) \quad (\text{C.26})$$

$$= -D_n \tanh(L_y \rho_n) \sinh(L_y \rho_n y^*) + D_n \cosh(L_y \rho_n y^*) \quad (\text{C.27})$$

$$= D_n [-\tanh(L_y \rho_n) \sinh(L_y \rho_n y^*) + \cosh(L_y \rho_n y^*)] \quad (\text{C.28})$$

$$= \frac{D_n [\cosh(L_y \rho_n) \cosh(L_y \rho_n y^*) - \sinh(L_y \rho_n) \sinh(L_y \rho_n y^*)]}{\cosh(L_y \rho_n)} \quad (\text{C.29})$$

$$= \frac{D_n \cosh(L_y \rho_n (1 - y^*))}{\cosh(L_y \rho_n)} \quad (\text{C.30})$$

Ainsi l'expression de $g_n(y^*)$ devient :

$$g_n(y^*) = \frac{D_n \cosh(L_y \rho_n (1 - y^*))}{\cosh(L_y \rho_n)} \quad (\text{C.31})$$

C.13 implique que ρ_n soit solution de l'équation transcendante suivante :

$$L_x \rho_n \tan(L_x \rho_n) = Bi \quad (\text{C.32})$$

En effet :

$$\frac{\partial f_n(x^*) g_n(y^*)}{\partial x^*}(1, y^*) + Bi f_n(1) g_n(y^*) = B_n g_n(y^*) [-L_x \rho_n \sin(L_x \rho_n) + Bi \cos(L_x \rho_n)] \quad (\text{C.33})$$

Or $B_n g_n(y^*)$ n'est pas nul donc $L_x \rho_n \sin(L_x \rho_n) = Bi \cos(L_x \rho_n)$ ce qui implique que ρ_n soit solution de C.32. Sans ordinateur C.32 peut être résolue graphiquement suivant la démarche proposée en section VIII.

Désormais la décomposition C.15 s'écrit ($\beta_n = \alpha_n B_n D_n$) :

$$\theta(x^*, y^*) = \sum_{n=1}^{\infty} \beta_n \cos(L_x \rho_n x^*) \frac{\cosh(L_y \rho_n (1 - y^*))}{\cosh(L_y \rho_n)} \quad (\text{C.34})$$

Les coefficients β_n seront déterminés à l'aide de C.10. Mais pour cela il faut montrer une propriété d'orthogonalité.

VI. Les fonctions sont orthogonales entre-elles

Un produit scalaire $\langle f, g \rangle$ est introduit, son expression est :

$$\langle f, g \rangle = \int_0^1 f(x^*) g(x^*) dx^* \quad (\text{C.35})$$

Si ρ_n et ρ_m deux solutions distinctes de C.32 alors $\cos(L_x \rho_n x^*)$ et $\cos(L_x \rho_m x^*)$ sont orthogonales vis-à-vis du produit scalaire C.35 :

$$\langle \cos(L_x \rho_n x^*), \cos(L_x \rho_m x^*) \rangle = \int_0^1 \cos(L_x \rho_n x^*) \cos(L_x \rho_m x^*) dx^* = 0 \quad (\text{C.36})$$

La démonstration de C.36 est un simple exercice de calcul intégral qui est effectué en section IX.

VII. Détermination de β_n

Il ne reste plus qu'à déterminer les β_n . La décomposition C.34 doit donc vérifier la condition aux limites C.10. Ce qui implique que :

$$\sum_{m=1}^{\infty} \beta_m \cos(L_x \rho_m x^*) = 1 \quad (\text{C.37})$$

L'utilisation judicieuse de la C.36 va permettre de conclure. C.37 est multipliée par $\cos(L_x \rho_n x^*)$ et est intégrée sur $[0, 1]$:

$$\int_0^1 \sum_{m=1}^{\infty} \beta_m \cos(L_x \rho_m x^*) \cos(L_x \rho_n x^*) dx^* = \int_0^1 \cos(L_x \rho_n x^*) dx^* \quad (\text{C.38})$$

Il est facile de reconnaître le produit scalaire C.35 :

$$\left\langle \sum_{m=1}^{\infty} \beta_m \cos(L_x \rho_m x^*), \cos(L_x \rho_n x^*) \right\rangle = \langle 1, \cos(L_x \rho_n x^*) \rangle \quad (\text{C.39})$$

Le terme de gauche dans C.39 se simplifie à l'aide de C.36 d'où :

$$\langle \beta_n \cos(L_x \rho_n x^*), \cos(L_x \rho_n x^*) \rangle = \langle 1, \cos(L_x \rho_n x^*) \rangle \quad (\text{C.40})$$

La valeur de β_n est donc donnée par :

$$\beta_n = \frac{\langle 1, \cos(L_x \rho_n x^*) \rangle}{\langle \cos(L_x \rho_n x^*), \cos(L_x \rho_n x^*) \rangle} \quad (\text{C.41})$$

Après les calculs présentés en section X., l'expression de β_n est :

$$\beta_n = \frac{2Bi}{((L_x \rho_n)^2 + Bi^2 + Bi) \cos(L_x \rho_n)} \quad (\text{C.42})$$

L'expression de $\theta(x^*, y^*)$ est alors :

$$\theta(x^*, y^*) = \sum_{n=1}^{\infty} \frac{2Bi \cos(L_x \rho_n x^*) \cosh(L_y \rho_n (1 - y^*))}{((L_x \rho_n)^2 + Bi^2 + Bi) \cos(L_x \rho_n) \cosh(L_y \rho_n)} \quad (\text{C.43})$$

Ce qui permet d'établir l'expression de T :

$$T = T_{amb} + 2(T_i - T_{amb}) \sum_{n=1}^{\infty} \frac{Bi \cos(L_x \rho_n x^*) \cosh(L_y \rho_n (1 - y^*))}{((L_x \rho_n)^2 + Bi^2 + Bi) \cos(L_x \rho_n) \cosh(L_y \rho_n)} \quad (\text{C.44})$$

VIII. Complément I : Méthode graphique pour l'équation transcendante

Sans ordinateur, il faut recourir à une méthode graphique pour résoudre l'équation transcendante C.32. Les ρ_n seront les abscisses des intersections entre les courbes des fonctions suivantes :

$$i_1 : z \rightarrow \tan(L_x z) \quad (\text{C.45})$$

$$i_2 : z \rightarrow \frac{Bi}{L_x z} \quad (\text{C.46})$$

Soit graphiquement pour les valeurs numériques suivantes $L_x = 3m$ et $Bi = 2$, :

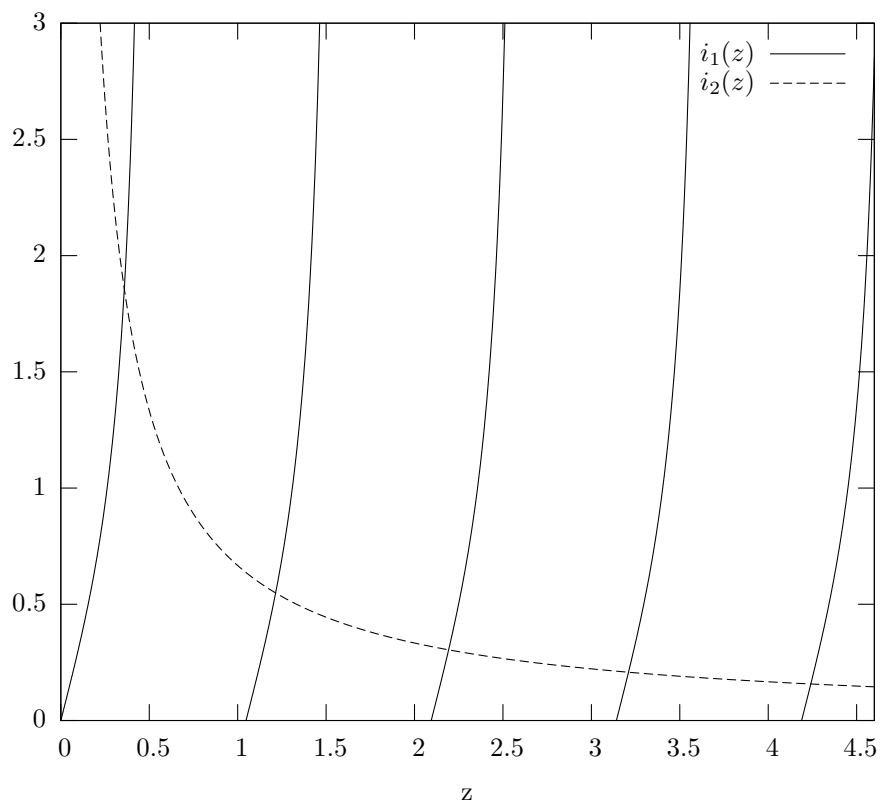


FIG. C.2 – Tracé des fonctions permettant de déterminer graphiquement les racines de C.32

Sur ce graphique, il est possible de déterminer les 5 premières racines de C.32 (pour $L_x = 3m$ et $Bi = 2$) :

- $\rho_1 \approx 0.4$
- $\rho_2 \approx 1.2$
- $\rho_3 \approx 2.2$
- $\rho_4 \approx 3.2$
- $\rho_5 \approx 4.2$

IX. Complément II : démonstration de l'orthogonalité C.36

Soit ρ_n et ρ_m deux solutions distinctes de C.32, il faut démontrer que

$$\langle \cos(L_x \rho_n x^*), \cos(L_x \rho_m x^*) \rangle = 0 \quad (\text{C.47})$$

Il s'agit d'un simple exercice de calcul intégral :

$$\langle \cos(L_x \rho_n x^*), \cos(L_x \rho_m x^*) \rangle = \int_0^1 \cos(L_x \rho_n x^*) \cos(L_x \rho_m x^*) dx^* \quad (\text{C.48})$$

$$= \frac{1}{2} \int_0^1 \cos(L_x(\rho_n - \rho_m)x^*) + \cos(L_x(\rho_n + \rho_m)x^*) dx^* \quad (\text{C.49})$$

$$= \frac{1}{2} \left[\frac{\sin(L_x(\rho_n - \rho_m)x^*)}{L_x(\rho_n - \rho_m)} + \frac{\sin(L_x(\rho_n + \rho_m)x^*)}{L_x(\rho_n + \rho_m)} \right]_0^1 \quad (\text{C.50})$$

$$= \frac{1}{2} \left[\frac{\sin(L_x(\rho_n - \rho_m))}{L_x(\rho_n - \rho_m)} + \frac{\sin(L_x(\rho_n + \rho_m))}{L_x(\rho_n + \rho_m)} \right] \quad (\text{C.51})$$

Il faut donc montrer que le numérateur de C.51 est nul soit que :

$$\rho_n [\sin(L_x(\rho_n - \rho_m)) + \sin(L_x(\rho_n + \rho_m))] + \rho_m [\sin(L_x(\rho_n - \rho_m)) - \sin(L_x(\rho_n + \rho_m))] = 0 \quad (\text{C.52})$$

Or d'après les formules d'addition :

$$\sin(L_x(\rho_n - \rho_m)) + \sin(L_x(\rho_n + \rho_m)) = 2 \sin(L_x \rho_n) \cos(L_x \rho_m) \quad (\text{C.53})$$

$$\sin(L_x(\rho_n - \rho_m)) - \sin(L_x(\rho_n + \rho_m)) = -2 \sin(L_x \rho_m) \cos(L_x \rho_n) \quad (\text{C.54})$$

C.52 est donc équivalent à :

$$\rho_n \sin(L_x \rho_n) \cos(L_x \rho_m) - \rho_m \sin(L_x \rho_m) \cos(L_x \rho_n) = 0 \quad (\text{C.55})$$

Or ρ_n et ρ_m sont deux solutions distinctes de C.32, ce qui implique les deux relations suivantes :

$$\rho_n \sin(L_x \rho_n) \cos(L_x \rho_m) = \frac{1}{L_x} \cos(L_x \rho_n) \cos(L_x \rho_m) \quad (\text{C.56})$$

$$\rho_m \sin(L_x \rho_m) \cos(L_x \rho_n) = \frac{1}{L_x} \cos(L_x \rho_m) \cos(L_x \rho_n) \quad (\text{C.57})$$

C.56 et C.57 impliquent C.55. Dès lors C.52 est vraie ce qui démontre C.47.

X. Complément III : Détail du calcul de β_n

Il s'agit ici d'effectuer les calculs menant de C.41 à C.42 :

$$\beta_n = \frac{\langle 1, \cos(L_x \rho_n x^*) \rangle}{\langle \cos(L_x \rho_n x^*), \cos(L_x \rho_n x^*) \rangle} \quad (\text{C.58})$$

$$= \frac{\int_0^1 \cos(L_x \rho_n x^*) dx^*}{\int_0^1 (\cos(L_x \rho_n x^*))^2 dx^*} \quad (\text{C.59})$$

$$= \frac{\left[\frac{\sin(L_x \rho_n x^*)}{L_x \rho_n} \right]_0^1}{\int_0^1 \frac{1 + \cos(2L_x \rho_n x^*)}{2} dx^*} \quad (\text{C.60})$$

$$= \frac{2 \sin(L_x \rho_n)}{L_x \rho_n \left[x^* + \frac{\sin(2L_x \rho_n x^*)}{2L_x \rho_n} \right]_0^1} \quad (\text{C.61})$$

$$= \frac{2 \sin(L_x \rho_n)}{L_x \rho_n \left(1 + \frac{\sin(2L_x \rho_n)}{2L_x \rho_n} \right)} \quad (\text{C.62})$$

$$= \frac{2 \sin(L_x \rho_n)}{L_x \rho_n + \sin(L_x \rho_n) \cos(L_x \rho_n)} \quad (\text{C.63})$$

ρ_n est racine de C.32 donc $L_x \rho_n \sin(L_x \rho_n) = Bi \cos(L_x \rho_n)$ dès lors :

$$\beta_n = \frac{2Bi}{\frac{L_x \rho_n}{\cos(L_x \rho_n)} (L_x \rho_n + \sin(L_x \rho_n) \cos(L_x \rho_n))} \quad (\text{C.64})$$

$$= \frac{2Bi}{\frac{(L_x \rho_n)^2}{\cos(L_x \rho_n)} + L_x \rho_n \sin(L_x \rho_n)} \quad (\text{C.65})$$

$$= \frac{2Bi}{\frac{(L_x \rho_n)^2 ((\sin(L_x \rho_n))^2 + (\cos(L_x \rho_n))^2)}{\cos(L_x \rho_n)} + Bi \cos(L_x \rho_n)} \quad (\text{C.66})$$

$$= \frac{2Bi}{Bi^2 \cos(L_x \rho_n) + (L_x \rho_n)^2 \cos(L_x \rho_n) + Bi \cos(L_x \rho_n)} \quad (\text{C.67})$$

$$= \frac{2Bi}{((L_x \rho_n)^2 + Bi^2 + Bi) \cos(L_x \rho_n)} \quad (\text{C.68})$$

Ce qui établit C.42.

Analyse modale « classique »

Contenu du chapitre

I.	Le système thermique envisagé	120
II.	Séparation glissant/dynamique	121
III.	Les modes de Fourier	122
IV.	Les équations d'états	123
V.	Résumés des points essentiels	125

Il s'agit ici de montrer rapidement la démarche de l'analyse modale classique dans le cas d'un système thermique purement conducteur avec des conditions aux limites de Fourier. La prise en compte des conditions aux limites de Neumann ou de Dirichlet est présentée dans [30], la prise en compte du transport est présenté dans [17]. Les méthodes de réduction modale sont ici totalement oubliées, une présentation complète est présenté dans [38].

I. Le système thermique envisagé

La forme du domaine physique n'est pas limitée. Le domaine est noté Ω , sa frontière est notée $\partial\Omega$ et la normale extérieure unitaire au domaine est notée $\underline{n}_{\partial\Omega}$:

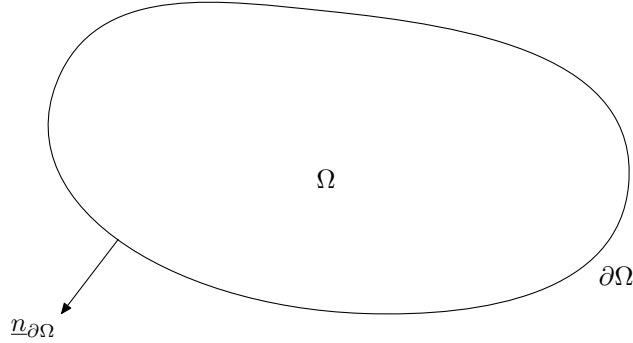


FIG. D.1 – Domaine Ω avec sa frontière et sa normale

Le système thermique envisagé est un système thermique conducteur, sans transport, linéaire. Les propriétés thermophysiques et le coefficient de convection ne dépendent que de l'espace : $k_0(M), c_0(M)$ et $h_0(M)$ ¹. Les propriétés thermophysiques ne sont donc pas fonction de la température et le coefficient de convection ne dépend pas du temps.

Les équations du problème sont donc :

$$\forall M \in \Omega \quad c_0 \frac{\partial T}{\partial t} = \text{div}(k_0 \underline{\nabla} T) + \pi \quad (\text{D.1})$$

$$\forall M \in \partial\Omega \quad -k_0 \underline{\nabla}(T) \bullet \underline{n}_{\partial\Omega} = h_0(T - T_{ext}) \quad (\text{D.2})$$

$$\forall M \in \Omega \quad T(M, 0) = T_i(M) \quad (\text{D.3})$$

Le coefficient de convection $h_0(M)$ ne peut pas dépendre du temps, par contre le champ de sollicitation du problème $T_{ext}(M, t)$ peut dépendre du temps. Le terme source $\pi(M, t)$ peut dépendre sans restriction du temps et de l'espace.

Le problème est donc classique, sa solution peut être recherchée à l'aide des méthodes classiques (différences finies, volumes finis, éléments finis) sans grande difficulté. Mais ce problème peut aussi être résolu à l'aide d'une méthode modale comme vont le montrer les lignes qui suivent.

Le champ initial est presque quelconque, pour plus de détail se référer à la remarque de la page suivante.

¹L'indice 0 est indiqué par soucis d'homogénéité avec l'annexe E sur les modes de branche.

II. Séparation glissant/dynamique

En analyse modale classique, il est nécessaire d'effectuer la séparation du champ thermique en deux termes : le champ dynamique T_d et le terme glissant T_g . Le champ glissant va être défini, puis le champ dynamique sera défini à l'aide du champ glissant.

II.1. Le champ glissant : $T_g(M, t)$

Le champ glissant est équivalent à tout instant au régime permanent qui serait atteint par le système si les sollicitations (π et T_{ext}) gardaient une valeur constante à partir de cet instant.

Une autre définition du champ glissant est la suivante : le champ glissant est le champ qui s'établirait à chaque instant si la capacité thermique était nulle.

Le champ glissant est donc défini par les deux équations suivantes :

$$\forall M \in \Omega \quad \text{div}(k_0 \underline{\nabla} T_g) + \pi = 0 \quad (\text{D.4})$$

$$\forall M \in \partial\Omega \quad -k_0 \underline{\nabla}(T_g) \bullet \underline{n}_{\partial\Omega} = h_0(T_g - T_{ext}) \quad (\text{D.5})$$

Il n'y a pas de troisième équation étant donné qu'il s'agit à chaque instant d'un régime permanent. Le champ glissant dépend du temps si une des sollicitations (T_{ext}, π) varie au cours du temps. La méthode modale ne pourra être performante que si la recherche du champ glissant est rapide. Il existe plusieurs situations où la détermination du champ glissant est rapide :

- Le problème thermique est un problème de relaxation thermique, T_g ne dépend pas du temps et une seule détermination sera nécessaire.
- Le champ glissant peut être déterminé analytiquement. C'est le cas si la géométrie est simple et si les sollicitations sont simples. De manière classique le champ glissant est alors déterminé par la méthode de séparation des variables.

II.2. Le champ dynamique : $T_d(M, t)$

Le champ dynamique est défini par :

$$\forall M \in \Omega \quad T_d(M, t) \stackrel{\text{def}}{=} T(M, t) - T_g(M, t) \quad (\text{D.6})$$

Ce qui permet d'établir les équations du champ dynamique :

$$\forall M \in \Omega \quad c_0 \frac{\partial T_d}{\partial t} = \text{div}(k_0 \underline{\nabla} T_d) - c_0 \frac{\partial T_g}{\partial t} \quad (\text{D.7})$$

$$\forall M \in \partial\Omega \quad -k_0 \underline{\nabla}(T_d) \bullet \underline{n}_{\partial\Omega} = h_0 T_d \quad (\text{D.8})$$

$$\forall M \in \Omega \quad T_d(M, 0) = T_i(M) - T_g(M, 0) \quad (\text{D.9})$$

Remarque :

Le champ dynamique se décompose dans la base des modes de Fourier (voir pages suivantes) si le champ dynamique initial est décomposable dans la base des modes de Fourier. En général cette restriction n'est pas problématique, car cela impose uniquement que $T_i(M) - T_g(M, 0)$ appartienne à l'espace $H_F^1(\Omega)$.

III. Les modes de Fourier

Le problème de Fourier va être défini. Ensuite il sera mis sous forme variationnelle et des opérateurs intégraux seront définis. L'application du théorème spectral (voir annexe B) permettra de montrer que les modes de Fourier forment une base de $H_F^1(\Omega)$. La définition de cet espace fonctionnel sera donnée. Enfin les propriétés d'orthogonalité vérifiées par les modes de Fourier seront rappelées.

III.1. Définition des modes de Fourier

Les modes à utiliser pour le problème thermique considéré sont les modes de Fourier. Les modes de Fourier (z_i^F, V_i^F) sont les solutions du problème aux valeurs propres de Fourier défini par les deux équations suivantes :

$$\forall M \in \Omega \quad \text{div}(k_0 \underline{\nabla} V_i^F) = z_i^F c_0 V_i^F \quad (\text{D.10})$$

$$\forall M \in \partial\Omega \quad -k_0 \underline{\nabla}(V_i^F) \bullet \underline{n}_{\partial\Omega} = h_0 V_i^F \quad (\text{D.11})$$

z_i^F est la valeur propre du mode tandis que $V_i^F(M)$ est la fonction propre ou vecteur propre du mode. La valeur propre est un scalaire et le vecteur propre est un champ du domaine Ω . Le problème va être mis sous forme variationnelle ce qui permettra d'utiliser le théorème spectral.

III.2. Formulation variationnelle du problème de Fourier

L'équation dans le domaine D.10 est multipliée par une fonction test $v \in H^1(\Omega)$ et est intégrée sur le domaine Ω :

$$\forall v \in H^1(\Omega) \quad \int_{\Omega} \text{div}(k_0 \underline{\nabla}(V_i^F)) v \, d\Omega = \int_{\Omega} z_i^F c_0 V_i^F v \, d\Omega \quad (\text{D.12})$$

Le théorème de Green-Ostrogradsky A.4 est appliqué à D.12 :

$$\forall v \in H^1(\Omega) \quad - \int_{\Omega} k_0 \underline{\nabla}(v) \bullet \underline{\nabla}(V_i^F) \, d\Omega + \int_{\partial\Omega} v k_0 \underline{\nabla}(V_i^F) \bullet \underline{n}_{\partial\Omega} \, d(\partial\Omega) = z_i^F \int_{\Omega} c_0 V_i^F v \, d\Omega \quad (\text{D.13})$$

La condition aux limites de Fourier D.11 alors utilisée pour simplifier le problème :

$$\forall v \in H^1(\Omega) \quad \int_{\Omega} k_0 \underline{\nabla}(v) \bullet \underline{\nabla}(V_i^F) \, d\Omega + \int_{\partial\Omega} h_0 v V_i^F \, d(\partial\Omega) = -z_i^F \int_{\Omega} c_0 V_i^F v \, d\Omega \quad (\text{D.14})$$

D.14 est la formulation variationnelle du problème de Fourier.

III.3. Introduction des opérateurs intégraux

Les opérateurs intégraux \mathcal{A}_F et \mathcal{C}_F sont définis à partir de la formulation variationnelle D.14, ils permettront d'utiliser le théorème spectral :

$$\begin{cases} \mathcal{A}_F(f, g) = \int_{\Omega} k_0 \underline{\nabla}(f) \bullet \underline{\nabla}(g) \, d\Omega + \int_{\partial\Omega} h_0 f g \, d(\partial\Omega) \\ \mathcal{C}_F(f, g) = \int_{\Omega} c_0 f g \, d\Omega \end{cases} \quad (\text{D.15})$$

Ces deux opérateurs sont bi-linéaires, symétriques, définis et positifs.

Le problème de Fourier défini par les équations D.10 et D.11 est donc équivalent au problème variationnel suivant :

$$\forall v \in H^1(\Omega) \quad \mathcal{A}_F(V_i^F, v) = -z_i^F \mathcal{C}_F(V_i^F, v) \quad (\text{D.16})$$

III.4. Espace fonctionnel des modes de Fourier

Les modes de Fourier forment une base de $H_F^1(\Omega)$ qui est défini par :

$$H_F^1(\Omega) = \{f \in H^1(\Omega) \mid \forall M \in \partial\Omega \quad -k_0 \underline{\nabla} f \bullet \underline{n}_{\partial\Omega} = hf\} \quad (\text{D.17})$$

Le lecteur intéressé pourra consulter l'annexe B.

III.5. Les conditions d'orthogonalité vérifiées par les modes de Fourier

Suivant les idées qui seront présentées en section III.4. page 133, il est possible d'obtenir les deux conditions d'orthogonalité pour les modes de Fourier.

Les modes de Fourier sont orthogonaux vis-à-vis des opérateur $\mathcal{C}_F(f, g)$ et $\mathcal{A}_F(f, g)$:

$$\mathcal{C}_F(V_i^F, V_j^F) = \delta_{ij} \quad (\text{D.18})$$

$$\mathcal{A}_F(V_i^F, V_j^F) = -z_i^F \delta_{ij} \quad (\text{D.19})$$

Ces deux propriétés vont permettre de montrer que les équations d'états sont découplées.

Le champ dynamique se décompose dans l'espace des modes de Fourier :

$$T_d(M, t) = \sum_{i=1}^{\infty} x_i^F(t) V_i^F(M) \quad (\text{D.20})$$

La condition d'orthogonalité D.18 permet de déterminer les états initiaux, en effet :

$$\mathcal{C}_F(T_d(M, 0), V_i^F) = \mathcal{C}_F\left(\sum_{k=1}^{\infty} x_k^F(0) V_k^F, V_i^F\right) \quad (\text{D.21})$$

$$= \sum_{k=1}^{\infty} x_k^F(0) \mathcal{C}_F(V_k^F, V_i^F) \quad (\text{D.22})$$

$$= x_i^F(0) \mathcal{C}_F(V_i^F, V_i^F) \quad (\text{D.23})$$

$$= x_i^F(0) \quad (\text{D.24})$$

III.6. Calcul des modes

Les modes peuvent être calculés numériquement d'une manière similaire au calcul des modes de branche (voir section IV. page 135).

IV. Les équations d'états

L'écriture des équations d'états s'effectue de la manière suivante :

- Les équations du champ dynamique sont mises sous forme variationnelle.
- Le champ dynamique est projeté dans l'espace des modes de Fourier.
- La fonction test v est remplacée par un mode de Fourier V_k^F .
- Les propriétés d'orthogonalité permettent de simplifier le problème.

IV.1. Formulation variationnelle des équations du champ dynamique

L'équation D.7 dans le domaine Ω est multipliée par une fonction test $v \in H_F^1(\Omega)$ et est intégrée sur le domaine Ω :

$$\forall v \in H_F^1(\Omega) \quad \int_{\Omega} c_0 \frac{\partial T_d}{\partial t} v d\Omega = \int_{\Omega} \text{div}(k_0 \nabla T_d) v d\Omega - \int_{\Omega} c_0 \frac{\partial T_g}{\partial t} v d\Omega \quad (\text{D.25})$$

Le théorème de Green-Ostrogradsky A.4 est appliqué à D.25 :

$$\begin{aligned} \forall v \in H_F^1(\Omega) \quad \int_{\Omega} c_0 \frac{\partial T_d}{\partial t} v d\Omega = \\ - \int_{\Omega} k_0 \nabla(v) \bullet \nabla(T_d) d\Omega + \int_{\partial\Omega} v k_0 \nabla(T_d) \bullet \underline{n}_{\partial\Omega} d(\partial\Omega) - \int_{\Omega} c_0 \frac{\partial T_g}{\partial t} v d\Omega \end{aligned} \quad (\text{D.26})$$

La condition aux limites de Fourier D.8 alors utilisée pour simplifier le problème :

$$\forall v \in H_F^1(\Omega) \quad \int_{\Omega} c_0 \frac{\partial T_d}{\partial t} v d\Omega = - \int_{\Omega} k_0 \nabla(v) \bullet \nabla(T_d) d\Omega - \int_{\partial\Omega} v h_0 T_d d(\partial\Omega) - \int_{\Omega} c_0 \frac{\partial T_g}{\partial t} v d\Omega \quad (\text{D.27})$$

Les opérateurs \mathcal{A}_F et \mathcal{C}_F sont reconnus dans D.27. La formulation variationnelle sera donc définie à l'aide des opérateurs déjà introduits.

La formulation variationnelle des équations du champ glissant est donc :

$$\left\{ \begin{array}{l} \forall v \in H_F^1(\Omega) \quad \mathcal{C}_F \left(\frac{\partial T_d}{\partial t}, v \right) = -\mathcal{A}_F(T_d, v) - \mathcal{C}_F \left(\frac{\partial T_g}{\partial t}, v \right) \\ \forall M \in \Omega \quad T_d(M, 0) = T_i(M) - T_g(M, 0) \end{array} \right. \quad (\text{D.28})$$

IV.2. Obtention des équations d'états

Le champ glissant T_g est remplacé par sa décomposition D.20 dans D.28 (en oubliant provisoirement la condition initiale).

$$\forall v \in H_F^1(\Omega) \quad \mathcal{C}_F \left(\frac{\partial \sum_{i=1}^{\infty} x_i^F V_i^F}{\partial t}, v \right) = -\mathcal{A}_F \left(\sum_{i=1}^{\infty} x_i^F V_i^F, v \right) - \mathcal{C}_F \left(\frac{\partial T_g}{\partial t}, v \right) \quad (\text{D.29})$$

Les opérateurs \mathcal{A}_F et \mathcal{C}_F sont bi-linéaires, il est donc possible de sortir la sommation des opérateurs :

$$\forall v \in H_F^1(\Omega) \quad \sum_{i=1}^{\infty} \mathcal{C}_F \left(\frac{\partial x_i^F V_i^F}{\partial t}, v \right) = - \sum_{i=1}^{\infty} \mathcal{A}_F (x_i^F V_i^F, v) - \mathcal{C}_F \left(\frac{\partial T_g}{\partial t}, v \right) \quad (\text{D.30})$$

Les fonctions propres V_i^F sont uniquement fonction de l'espace ainsi $\frac{\partial x_i^F V_i^F}{\partial t} = \frac{dx_i^F}{dt} V_i^F$. La bi-linéarité des opérateurs \mathcal{A}_F et \mathcal{C}_F permet de sortir les états de ceux-ci, d'où :

$$\forall v \in H_F^1(\Omega) \quad \sum_{i=1}^{\infty} \left(\mathcal{C}_F (V_i^F, v) \frac{dx_i^F}{dt} \right) = - \sum_{i=1}^{\infty} (\mathcal{A}_F (V_i^F, v) x_i^F) - \mathcal{C}_F \left(\frac{\partial T_g}{\partial t}, v \right) \quad (\text{D.31})$$

Pour finir les modes V_k^F sont utilisés comme fonction test² v .

$$\forall k \in \mathbb{N}^* \quad \sum_{i=1}^{\infty} \left(\mathcal{C}_F (V_i^F, V_k^F) \frac{dx_i^F}{dt} \right) = - \sum_{i=1}^{\infty} (\mathcal{A}_F (V_i^F, V_k^F) x_i^F) - \mathcal{C}_F \left(\frac{\partial T_g}{\partial t}, V_k^F \right) \quad (\text{D.32})$$

²En effet les V_k^F formant une base de $H_F^1(\Omega)$, toutes les fonctions v de cet espace sont décomposables dans la base des modes de Fourier. La condition $\forall v \in H_F^1(\Omega)$ est alors remplacée par la condition $\forall k \in \mathbb{N}^*$

L'équation D.32 se simplifie en utilisant les conditions d'orthogonalité D.18 et D.19, dans un premier temps la sommation est éliminée :

$$\forall k \in \mathbb{N}^* \quad \mathcal{C}_F(V_k^F, V_k^F) \frac{dx_k^F}{dt} = -\mathcal{A}_F(V_k^F, V_k^F) x_k^F - \mathcal{C}_F\left(\frac{\partial T_g}{\partial t}, V_k^F\right) \quad (\text{D.33})$$

Dans un second temps D.33 est simplifiée car $\mathcal{C}_F(V_k^F, V_k^F) = 1$ et $\mathcal{C}_F(V_k^F, V_k^F) = -z_k^F$:

$$\forall k \in \mathbb{N}^* \quad \frac{dx_k^F}{dt} = z_k^F x_k^F - \mathcal{C}_F\left(\frac{\partial T_g}{\partial t}, V_k^F\right) \quad (\text{D.34})$$

L'indice muet k est remplacé par l'indice i et la condition initiale est rappelée :

$$\left| \begin{array}{l} \forall i \in \mathbb{N}^* \quad \frac{dx_i^F}{dt} = z_i^F x_i^F - \mathcal{C}_F\left(\frac{\partial T_g}{\partial t}, V_i^F\right) \\ \forall i \in \mathbb{N}^* \quad x_i^F(0) = \mathcal{C}_F(T_d(M, 0), V_i^F) \end{array} \right. \quad (\text{D.35})$$

La formulation variationnelle D.35 est composée des équations d'états et des conditions initiales. La dynamique de chacun des états est indépendante de tous les autres, les équations d'états sont donc découplées. Par analogie avec les systèmes matricielles, l'équation du champ dynamique est alors qualifiée de diagonalisée.

V. Résumés des points essentiels

La méthode modale classique s'applique à une large gamme de systèmes thermiques linéaires. Les modes étant calculés numériquement, il est possible de traiter des géométries complexes. Les équations d'états sont découplées ainsi le calcul des états peut se faire état par état. Il est ainsi possible de calculer progressivement la solution. La méthode est très rapide pour peu que le calcul du champ glissant soit rapide.

Annexe E

Base de branche

Contenu du chapitre

I.	Définition des modes de branche	128
II.	Formulation variationnelle du problème de Branche et utilisation du théorème spectral	130
III.	Compléments sur les modes	132
IV.	Calcul des modes	135

Les points les plus essentiels sur la base de branche sont les suivants :

- L'ensemble des modes de branche du corps Ω forment une base de $H^1(\Omega)$.
- Les modes vérifient deux conditions d'orthogonalité.
- Le calcul numérique des modes d'un domaine Ω quelconque est possible.
- La base de branche présente des dominances, il est ainsi possible de la réduire.

Les propriétés des modes de branche font l'objet de cette annexe, l'annexe suivante présentera les différentes possibilités de réduction de la base de branche.

La base de branche va être présentée pour un système thermique sans transport, les transferts thermiques s'effectuent alors uniquement par conduction. La prise en compte du transport dans la base de branche est certes possible [34, 37, 26] mais n'a pas été utilisée dans cette thèse.

I. Définition des modes de branche

Les données fondamentales permettant de définir les modes de branche sont exposées avant l'introduction du problème de branche. Une comparaison avec le problème de Fourier sera menée.

I.1. Remarques préalables

Soit un domaine Ω de frontière $\partial\Omega$ et de normale extérieure unitaire $\underline{n}_{\partial\Omega}$:

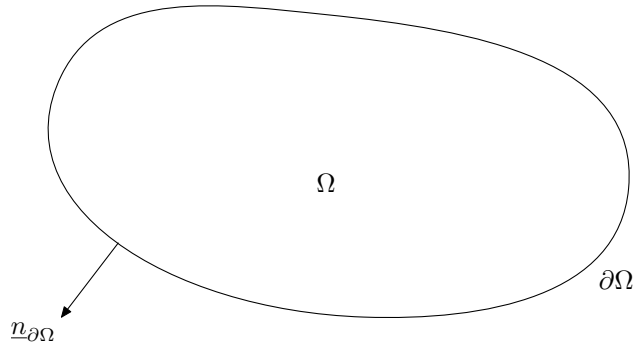


FIG. E.1 – Domaine Ω avec sa frontière et sa normale

Les transferts de chaleur s'effectuent uniquement par conduction sans transport, la géométrie du domaine Ω peut alors être quelconque¹.

Le domaine peut être hétérogène, il peut être constitué de plusieurs matériaux sous réserve que le contact entre les matériaux soit parfait.

Le problème de branche va être défini par rapport à des caractéristiques thermophysiques moyennes $(k_0(M), c_0(M))$ qui seraient celles considérées dans une linéarisation du domaine. Les caractéristiques thermophysiques $(k(M, T), c(M, T))$ réelles vérifient alors :

$$\forall M \in \Omega \quad k(M, T) = k_0(M) + k_1(M, T) \quad (\text{E.1})$$

$$\forall M \in \Omega \quad c(M, T) = c_0(M) + c_1(M, T) \quad (\text{E.2})$$

¹Dans le cas où il y aurait du transport solide ou fluide, il faut en plus que la frontière soit fixe [34], il y a alors certaines limites sur les systèmes qui peuvent être pris en compte.

I.2. Le problème de branche et la condition de Steklov

Le problème de branche associé au domaine Ω est constitué des deux équations suivantes :

$$\forall M \in \Omega \quad \text{div}(k_0 \underline{\nabla}(V_i)) = z_i c_0 V_i \quad (\text{E.3})$$

$$\forall M \in \partial\Omega \quad -k_0 \underline{\nabla} V_i \bullet \underline{n}_{\partial\Omega} = z_i \zeta V_i \quad (\text{E.4})$$

Le problème de branche est donc un problème aux valeurs propres généralisé. Les vecteurs propres $V_i(M)$ sont des champs spatiaux définis sur l'ensemble du domaine Ω . z_i est la valeur propre associée au vecteur propre $V_i(M)$. Le couple (z_i, V_i) est appelé mode de branche.

La condition aux limites E.4 est appelée condition de Steklov. Contrairement aux conditions aux limites classiques (Fourier, Dirichlet), la condition aux limites de Steklov n'est pas physique.

La condition aux limites de Steklov est donc abstraite, elle fait apparaître la valeur propre z_i du mode (z_i, V_i) . Le nombre de Steklov $\zeta(M)$ est une pondération qui assure l'homogénéité dimensionnelle et l'équilibre numérique entre E.3 et E.4. Le nombre de Steklov s'exprime en $J.m^{-2}.K^{-1}$ (ou encore en $kg.s^{-2}.K^{-1}$) et doit être positif.

Le choix du nombre de Steklov sera analysé en détail dans le paragraphe II.4. Il est cependant possible de signaler qu'en pratique :

$$\zeta(M) = \begin{cases} \zeta_c & \text{si la frontière n'est pas adiabatique} \\ 0 & \text{sinon} \end{cases} \quad (\text{E.5})$$

Avec ζ_c la valeur choisie pour le nombre de Steklov. La valeur 0 est utilisée pour les frontières adiabatiques du problème physique car ces frontières étant à flux nul, il est alors préférable d'imposer que les modes vérifient rigoureusement cette condition.

I.3. Les modes de Fourier et modes de branche

Il ne faut pas confondre le problème de branche avec le problème des modes de Fourier :

$$\forall M \in \Omega \quad \text{div}(k_0 \underline{\nabla}(V_i)) = z_i c_0 V_i \quad (\text{E.6})$$

$$\forall M \in \partial\Omega \quad -k_0 \underline{\nabla} V_i \bullet \underline{n}_{\partial\Omega} = h V_i \quad (\text{E.7})$$

Les équations E.3 et E.6 sont identiques. Par contre les équations E.4 et E.7 sont très différentes. La condition aux limites imposées dans le problème de Fourier ne dépend pas de la valeur propre, l'espace généré par les modes de Fourier est alors l'espace $H_F^1(\Omega)$:

$$H_F^1(\Omega) = \{f \in H^1(\Omega) \mid \forall M \in \partial\Omega \quad -k_0 \underline{\nabla} f \bullet \underline{n}_{\partial\Omega} = h f\} \quad (\text{E.8})$$

L'espace $H_F^1(\Omega)$ ne contient ainsi que les fonctions qui vérifient effectivement la condition aux limites E.6. Ce qui limite fortement le domaine d'application d'une base de Fourier.

A l'inverse la condition de Steklov fait intervenir la valeur propre du mode. Chaque mode vérifiera alors une condition dite de Fourier² mais différente de celle des modes qui auront une valeur propre différente. En conséquence l'espace généré par les modes de branche est l'espace $H^1(\Omega)$ sans aucune restriction si ζ n'est pas nul.

²Toutefois, il importe de signaler que $z_i \zeta$ n'est pas assimilable à un coefficient d'échange car les valeurs propres du problème de branche sont négatives ou nulles : $z_i \leq 0$ (voir équation E.23)

II. Formulation variationnelle du problème de Branche et utilisation du théorème spectral

Le théorème spectral présenté en annexe B est basé sur une formulation variationnelle du problème aux valeurs propres. Dans un premier temps la formulation variationnelle du problème de branche est établie. Dans un second temps les opérateurs \mathcal{C}_b et \mathcal{A}_b sont définis. L'utilisation du théorème spectral permet alors de conclure que les modes forment une base de $H^1(\Omega)$. Finalement le principe de moindre dégénérescence [13] appliqué aux opérateurs \mathcal{C}_b et \mathcal{A}_b permet de montrer qu'il existe un choix par défaut pour le paramètre de Steklov.

II.1. Etablissement de la formulation variationnelle

L'équation dans le domaine E.3 est multipliée par une fonction test $v \in H^1(\Omega)$ et est intégrée sur le domaine Ω :

$$\forall v \in H^1(\Omega) \quad \int_{\Omega} \text{div}(k_0 \underline{\nabla}(V_i)) v \, d\Omega = \int_{\Omega} z_i c_0 V_i v \, d\Omega \quad (\text{E.9})$$

Le théorème de Green A.4 est appliqué à E.9 :

$$\forall v \in H^1(\Omega) \quad - \int_{\Omega} k_0 \underline{\nabla}(v) \bullet \underline{\nabla}(V_i) \, d\Omega + \int_{\partial\Omega} k_0 v \underline{\nabla}(V_i) \bullet \underline{n}_{\partial\Omega} \, d(\partial\Omega) = \int_{\Omega} z_i c_0 V_i v \, d\Omega \quad (\text{E.10})$$

La condition aux limites de Steklov E.4 alors utilisée pour simplifier le problème :

$$\forall v \in H^1(\Omega) \quad - \int_{\Omega} k_0 \underline{\nabla}(v) \bullet \underline{\nabla}(V_i) \, d\Omega - \int_{\partial\Omega} z_i \zeta V_i v \, d(\partial\Omega) = \int_{\Omega} z_i c_0 V_i v \, d\Omega \quad (\text{E.11})$$

Les termes intégrant la valeur propre z_i sont placés du même côté et la valeur propre est sortie des intégrales, ce qui établit la formulation variationnelle du problème de branche.

$$\forall v \in H^1(\Omega) \quad \int_{\Omega} k_0 \underline{\nabla}(v) \bullet \underline{\nabla}(V_i) \, d\Omega = -z_i \left[\int_{\partial\Omega} \zeta V_i v \, d(\partial\Omega) + \int_{\Omega} c_0 V_i v \, d\Omega \right] \quad (\text{E.12})$$

II.2. Introduction des opérateurs \mathcal{A}_b et \mathcal{C}_b

Deux opérateurs sont définis à partir de E.12 :

$$\begin{cases} \mathcal{A}_b(f, g) = \int_{\Omega} k_0 \underline{\nabla}(f) \bullet \underline{\nabla}(g) \, d\Omega \\ \mathcal{C}_b(f, g) = \int_{\partial\Omega} \zeta f g \, d(\partial\Omega) + \int_{\Omega} c_0 f g \, d\Omega \end{cases} \quad (\text{E.13})$$

Le problème de branche s'écrit alors :

$$\forall v \in H^1(\Omega) \quad \mathcal{A}_b(V_i, v) = -z_i \mathcal{C}_b(V_i, v) \quad (\text{E.14})$$

Les opérateurs \mathcal{A}_b (et \mathcal{C}_b) sont bi-linéaires, symétriques, définis et positifs.

II.3. Les modes de branche forment une base de $H^1(\Omega)$

L'utilisation du théorème spectral est délicate, le lecteur intéressé est renvoyé à l'annexe B. Ce théorème démontre que l'ensemble des modes de branche forment une base de $H^1(\Omega)$ à la condition que $\zeta(M)$ soit non nul sur l'ensemble de la frontière.

Dans le cas où le nombre de Steklov est nul sur une partie de la frontière que l'on note $\partial\Omega_{nul}$ alors l'ensemble des modes de branche forment une base de l'espace suivant

$$\{v \in H^1(\Omega) \mid \forall M \in \partial\Omega_{nul} \quad \nabla(v) \bullet \underline{n}_{\partial\Omega} = 0\} \quad (\text{E.15})$$

Dans la méthode modale classique, les modes vérifiaient la condition aux limites homogènes du champ dynamique... En conséquence les équations d'états sont découplées, mais seuls les systèmes thermiques linéaires sont envisageables.

Les modes de branche ne vérifient aucune condition aux limites particulière (pour peu que $\zeta(M)$ soit non nul). Il y aura alors une seule équation d'état qui couplera la dynamique de tous les états. Par contre les systèmes thermiques non-linéaires pourront être traités.

II.4. Choix par défaut du paramètre de Steklov

Le choix du paramètre de Steklov s'effectue selon les idées du principe de moindre dégénérescence (voir [13] page 168) habituellement utilisé en analyse dimensionnelle.

Selon ce principe les deux termes intégraux de \mathcal{C}_b doivent avoir un poids *a priori* équivalent :

$$\mathcal{C}_b(f, g) = \int_{\partial\Omega} \zeta f g \, d(\partial\Omega) + \int_{\Omega} c_0 f g \, d\Omega \quad (\text{E.16})$$

Il est possible de définir une longueur de référence l_r par :

$$l_r = \frac{\int_{\Omega} d\Omega}{\int_{\partial\Omega} H(\zeta(M)) d(\partial\Omega)} \quad (\text{E.17})$$

La fonction H sert à supprimer la partie de la frontière où le nombre de Steklov s'annule :

$$H(x) = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases} \quad (\text{E.18})$$

La capacité thermique de référence est la moyenne volumique de la capacité thermique réelle :

$$c_r = \frac{\int_{\Omega} c_0 d\Omega}{\int_{\Omega} d\Omega} \quad (\text{E.19})$$

En pratique $\zeta(M)$ est soit constant soit prend deux valeurs (ζ_c est la valeur choisie pour ζ) :

$$\zeta(M) = \begin{cases} \zeta_c & \text{si la frontière n'est pas adiabatique} \\ 0 & \text{sinon} \end{cases} \quad (\text{E.20})$$

Le rapport des deux termes dans E.16 est en ordre de grandeur tel que :

$$\frac{\int_{\partial\Omega} \zeta f g \, d(\partial\Omega)}{\int_{\Omega} c_0 f g \, d\Omega} \approx \frac{\zeta_c}{c_r l_r} \quad (\text{E.21})$$

D'après le principe de moindre dégénérescence, il est préférable que $\frac{\zeta_c}{c_r l_r}$ soit de l'ordre de 1.

Ainsi le choix par défaut du paramètre de Steklov est ζ_d avec :

$$\zeta_d = c_r l_r = \frac{\int_{\Omega} c_0 d\Omega}{\int_{\partial\Omega} H(\zeta(M)) d(\partial\Omega)} \quad (\text{E.22})$$

Différentes études ont été menées avant d'écrire les articles [36, 35, 43]. Dans chacune de ces études, il a été montré que les meilleurs résultats étaient obtenus avec $\zeta_c \in [0.5\zeta_d, 2\zeta_d]$. Ainsi en pratique, la valeur choisie pour ζ est systématiquement ζ_d .

III. Compléments sur les modes

III.1. Signe de la valeur propre

Il est possible de déterminer le signe de la valeur propre à l'aide du quotient de Rayleigh. En effet \mathcal{A}_b et \mathcal{C}_b sont définis et positifs, ainsi $\mathcal{A}_b(V_i, V_i) > 0$ et $\mathcal{C}_b(V_i, V_i) > 0$ si V_i n'est pas nul. Or par définition la fonction propre d'un mode est non nulle. Ainsi z_i est tel que :

$$z_i = -\frac{\mathcal{A}_b(V_i, V_i)}{\mathcal{C}_b(V_i, V_i)} \leq 0 \quad (\text{E.23})$$

Le quotient présent dans E.23 est appelé quotient de Rayleigh.

III.2. Tri des modes

Les modes sont triés suivant leur valeur propre. Le premier mode est donc le mode plat³⁴ de valeur propre nulle. Ce choix assure que $z_j \leq z_i$ pour $j > i$.

III.3. Dimension de la valeur propre

La valeur propre s'exprime en s^{-1} . Il est alors classique de définir une constante de temps τ_i par :

$$\tau_i = -\frac{1}{z_i} \quad (\text{E.24})$$

Dans le cas de l'analyse modale classique, la valeur propre z_i^F apparaît directement dans les équations d'états : D.35. Dans le cas de l'analyse modale de branche, les fonctions propres interviennent seules dans l'équation d'état 2.14. Les valeurs propres sont ainsi assez peu utilisées explicitement en analyse modale de branche, elle peuvent toutefois servir de critère dans une troncature de Marshall (annexe F).

³Quelques modes de branche sont présentés graphiquement dans l'annexe H.

⁴Le mode plat n'est pas présent dans la base de Fourier hormis dans le cas où toutes les frontières sont adiabatiques.

III.4. Propriétés d'orthogonalité des modes

Il est possible de construire les bases modales de manière à obtenir les deux relations d'orthogonalité suivantes :

$$\forall i, j \in (\mathbb{N}^*)^2 \quad \mathcal{C}_b(V_i, V_k) = \delta_{i,j} \quad (\text{E.25})$$

$$\forall i, j \in (\mathbb{N}^*)^2 \quad \mathcal{A}_b(V_i, V_j) = -z_i \delta_{i,j} \quad (\text{E.26})$$

En effet, les relations E.25 et E.26 sont automatiquement vérifiées (après normalisation des modes) si $z_i \neq z_j$. Mais dans le cas contraire (*i.e* $z_i = z_j$), il faut rendre les modes orthogonaux car ceux-ci ne le sont pas forcément.

III.4.1. Démonstration dans le cas où les valeurs propres sont distinctes

Il s'agit de démontrer que :

$$\text{si } z_i \neq z_j \quad \text{alors } \mathcal{C}_b(V_i, V_j) = 0 \quad (\text{E.27})$$

$$\text{si } z_i \neq z_j \quad \text{alors } \mathcal{A}_b(V_i, V_j) = 0 \quad (\text{E.28})$$

En premier lieu, il faut démontrer E.27, pour cela le problème de branche sous forme variationnelle E.14 est écrit pour deux modes (z_i, V_i) , (z_j, V_j) tels que $z_i \neq z_j$:

$$\forall v \in H^1(\Omega) \quad \mathcal{A}_b(V_i, v) = -z_i \mathcal{C}_b(V_i, v) \quad (\text{E.29})$$

$$\forall v \in H^1(\Omega) \quad \mathcal{A}_b(V_j, v) = -z_j \mathcal{C}_b(V_j, v) \quad (\text{E.30})$$

E.29 est vraie pour toute fonction v de $H^1(\Omega)$, elle l'est donc pour V_j . De même E.30 est vraie avec V_i utilisé comme fonction test v . Ainsi E.29 et E.30 deviennent :

$$\mathcal{A}_b(V_i, V_j) = -z_i \mathcal{C}_b(V_i, V_j) \quad (\text{E.31})$$

$$\mathcal{A}_b(V_j, V_i) = -z_j \mathcal{C}_b(V_j, V_i) \quad (\text{E.32})$$

La différence des deux relations précédentes permet d'écrire :

$$\mathcal{A}_b(V_j, V_i) - \mathcal{A}_b(V_i, V_j) = -z_j \mathcal{C}_b(V_j, V_i) + z_i \mathcal{C}_b(V_i, V_j) \quad (\text{E.33})$$

En tenant compte de la symétrie des opérateurs \mathcal{A}_b et \mathcal{C}_b ($\mathcal{A}_b(V_i, V_j) = \mathcal{A}_b(V_j, V_i)$ et $\mathcal{C}_b(V_i, V_j) = \mathcal{C}_b(V_j, V_i)$), E.33 devient :

$$\mathcal{A}_b(V_j, V_i) - \mathcal{A}_b(V_j, V_i) = -z_j \mathcal{C}_b(V_j, V_i) + z_i \mathcal{C}_b(V_j, V_i) \quad (\text{E.34})$$

Dès lors :

$$0 = (z_i - z_j) \mathcal{C}_b(V_j, V_i) \quad (\text{E.35})$$

Or par hypothèse $z_i \neq z_j$ donc cela démontre E.27

Il reste à démontrer E.28, pour cela il suffit d'utiliser E.27 dans E.31.

III.4.2. Cas où plusieurs modes ont la même valeur propre

Dans le cas du carré avec des paramètres constants, de nombreuses valeurs propres sont d'ordre 4. Il existe alors 4 modes qui ont la même valeur propre. Le cas des valeurs propres multiples n'est donc pas purement théorique.

Les relations E.25 et E.26 ne seront vérifiées que si les modes ont été rendus orthogonaux (il est possible d'utiliser le procédé d'orthogonalisation de Schmidt).

En effet les modes de branche ne sont pas forcément orthogonaux deux-à-deux vis à vis de \mathcal{C}_b et de \mathcal{A}_b comme le démontre cet exemple :

Soit z_i une valeur propre d'ordre 2, et un couple de solutions $\{(z_i, V_i), (z_i, V_j)\}$ vérifiant les deux relations d'orthogonalité E.25 et E.26.

Chaque élément du couple $\{(z_i, V_i + V_j), (z_i, V_j)\}$ est aussi une solution de problème de branche, par contre ce couple ne vérifie pas les relations d'orthogonalité proposées.

Dans le reste de ce manuscrit, il est implicitement supposé que les modes de branche vérifient les deux relations d'orthogonalité E.25 et E.26. Si tel n'était pas le cas, le calcul des états initiaux devrait être revu (la relation E.44 ne serait plus vraie) et le couplage entre les états augmente.

III.4.3. Normalisation des modes

Pour finir E.25 et E.26 nécessitent que les fonctions propres V_i soient normalisées. En effet si (z_i, V_i) est une solution du problème de branche, alors $(z_i, \alpha V_i)$ l'est aussi pour tout $\alpha \in \mathbb{R}^*$.

Si la fonction propre V_i^1 ne vérifie pas $\mathcal{C}_b(V_i^1, V_i^1)$, alors il est possible de la normaliser :

$$V_i = \frac{V_i^1}{\sqrt{\mathcal{C}_b(V_i^1, V_i^1)}} \quad (\text{E.36})$$

En effet :

$$\mathcal{C}_b(V_i, V_i) = \mathcal{C}_b\left(\frac{V_i^1}{\sqrt{\mathcal{C}_b(V_i^1, V_i^1)}}, \frac{V_i^1}{\sqrt{\mathcal{C}_b(V_i^1, V_i^1)}}\right) = \frac{1}{\sqrt{\mathcal{C}_b(V_i^1, V_i^1)}} \mathcal{C}_b\left(V_i^1, \frac{V_i^1}{\sqrt{\mathcal{C}_b(V_i^1, V_i^1)}}\right) \quad (\text{E.37})$$

$$= \left(\frac{1}{\sqrt{\mathcal{C}_b(V_i^1, V_i^1)}}\right)^2 \mathcal{C}_b(V_i^1, V_i^1) = \frac{1}{\mathcal{C}_b(V_i^1, V_i^1)} \mathcal{C}_b(V_i^1, V_i^1) = 1 \quad (\text{E.38})$$

III.5. Calcul des états initiaux

La relation d'orthogonalité E.25 permet de déterminer les états initiaux à partir de la connaissance du champ de température initial T_i de $H^1(\Omega)$.

En effet tout champ de $H^1(\Omega)$ est décomposable dans la base de branche :

$$T_i = \sum_{k=1}^{\infty} x_k(0) V_k \quad (\text{E.39})$$

La quantité $\mathcal{C}_b(T_i, V_i)$ est alors calculée en utilisant la bi-linéarité de l'opérateur \mathcal{C}_b et E.39 :

$$\mathcal{C}_b(T_i, V_i) = \mathcal{C}_b\left(\sum_{k=1}^{\infty} x_k(0) V_k, V_i\right) \quad (\text{E.40})$$

$$= \sum_{k=1}^{\infty} \mathcal{C}_b(x_k(0) V_k, V_i) = \sum_{k=1}^{\infty} x_k(0) \mathcal{C}_b(V_k, V_i) \quad (\text{E.41})$$

$$= \sum_{k=1}^{\infty} x_k(0) \delta_{k,i} = x_i(0) \quad (\text{E.42})$$

Les états initiaux sont donc déterminés par la formule suivante :

$$x_i(0) = \mathcal{C}_b(T_i, V_i) \quad (\text{E.43})$$

Remarque : E.43 peut s'étendre sans difficulté pour permettre une détermination des états à n'importe quel instant :

$$x_i(t) = \mathcal{C}_b(T(t), V_i) \quad (\text{E.44})$$

Avec :

$$T(t) = \sum_{k=1}^{\infty} x_k(t) V_k \quad (\text{E.45})$$

E.44 sera utilisée pour étudier la décomposabilité des champs de température du processeur et du radiateur (voir pages 59 et 60). E.43 est utilisée pour déterminer les états initiaux de chaque simulation, elle est donc beaucoup plus utilisée que E.44.

IV. Calcul des modes

Le calcul des modes peut s'effectuer de manière analytique dans le cas de géométries simples, c'est le cas de l'article [36]. Mais la plupart du temps, il est nécessaire d'utiliser une méthode numérique. Les lignes qui suivent explicitent la démarche dans le cas des éléments finis.

IV.1. Obtention d'un problème matriciel aux valeurs propres généralisées

Le problème de branche sous forme variationnelle E.14 est :

$$\forall v \in H^1(\Omega) \quad \mathcal{A}_b(V_i, v) = -z_i \mathcal{C}_b(V_i, v) \quad (\text{E.46})$$

La méthode des éléments finis définit une base comportant N fonctions de forme φ_i à support local. Les φ_i sont linéairement indépendantes, l'espace vectoriel engendré par les φ_i est donc de dimension N et est noté $H_N^1(\Omega)$. L'approximation interne consiste alors à remplacer $H^1(\Omega)$ par le sous-espace $H_N^1(\Omega)$.

Dans le cadre de l'approximation interne, le problème de branche devient ($V_{i,N} \in H_N^1(\Omega)$) :

$$\forall v_N \in H_N^1(\Omega) \quad \mathcal{A}_b(V_{i,N}, v_N) = -z_{i,N} \mathcal{C}_b(V_{i,N}, v_N) \quad (\text{E.47})$$

Le problème E.46 possède une infinité de solutions (z_i, V_i) , tandis que le problème E.47 possède N solutions $(z_{i,N}, V_{i,N})$.

Il existe un effet lié au maillage qui influence les modes déterminés numériquement :

- Les solutions numériques $(z_{i,N}, V_{i,N})$ de E.47 ne sont pas (en général) solution de E.46 :

$$\exists h \in H^1(\Omega) \quad \mathcal{A}_b(V_{i,N}, h) = -z_{i,N} \mathcal{C}_b(V_{i,N}, h) \quad (\text{E.48})$$

Une telle fonction h n'appartient pas à $H_N^1(\Omega)$ car sinon il y aurait contradiction avec E.47.

- Les premières solutions (dans l'ordre des valeurs propres décroissantes) présenteront une bonne convergence (*i.e* $|z_i - z_{i,N}|$ et $\|V_i - V_{i,N}\|_{H^1(\Omega)}$ sont faibles). A l'inverse les dernières solutions sont uniquement des artefacts numériques.

Le problème E.47 peut se mettre sous forme matricielle. Pour cela, les matrices \mathbb{A}_b et \mathbb{C}_b sont définies par (l'indice i désigne la ligne et l'indice j désigne la colonne) :

$$(\mathbb{A}_b)_{ij} = \mathcal{A}_b(\varphi_i, \varphi_j) \quad (\text{E.49})$$

$$(\mathbb{C}_b)_{ij} = \mathcal{C}_b(\varphi_i, \varphi_j) \quad (\text{E.50})$$

Les matrices \mathbb{A}_b et \mathbb{C}_b sont donc définies, symétriques et positives.

$V_{i,N}$ est une fonction de $H_N^1(\Omega)$, elle peut s'exprimer en fonction des φ_k . Un vecteur \mathbf{V}_i est alors associé à la fonction propre $V_{i,N}$. La k -ième composante du vecteur \mathbf{V}_i étant le coefficient de φ_k :

$$\forall M \in \Omega \quad V_{i,N}(M) = \sum_{k=1}^N (\mathbf{V}_i)_k \varphi_k(M) \quad (\text{E.51})$$

E.47 est alors équivalent au problème suivant :

$$\mathbb{A}_b \mathbf{V}_i = -z_{i,N} \mathbb{C}_b \mathbf{V}_i \quad (\text{E.52})$$

E.52 est un problème matriciel aux valeurs propres généralisées. Des algorithmes numériques performants permettent de résoudre ce problème.

IV.2. Le code ARPACK et son utilisation

Le code ARPACK [47] librement disponible sur internet fournit les outils spécifiques au calcul des solutions de E.52. Les éléments théoriques utilisés par ARPACK sont détaillés dans les références suivantes : [47, 12].

Le code ARPACK utilise la méthode des sous-espaces d'Arnoldi sur le problème suivant :

$$(\mathbb{A}_b + \sigma \mathbb{C}_b)^{-1} \mathbb{C}_b \mathbf{V}_j = \mu_j \mathbf{V}_j \quad (\text{E.53})$$

L'utilisation judicieuse du décalage spectral σ permet de calculer les modes par lots. Les vecteurs propres de E.53 sont des vecteurs propres de E.52. Les valeurs propres de E.52 sont reliées à (σ, μ_j) par :

$$z_{i,N} = \sigma - \frac{1}{\mu_j} \quad (\text{E.54})$$

Le décalage spectral est tel que les indices i et j ne sont pas égaux en général. M appels à ARPACK vont permettre de déterminer M lots de N_{etape} vecteurs propres. Si un décalage optimal est utilisé à chaque étape alors les $M N_{\text{etape}}$ premiers modes sont alors déterminés car ARPACK est capable de calculer les modes par ordre de valeur propre décroissante. A chaque étape ARPACK effectue une orthonormalisation de Schmidt, une orthonormalisation finale n'est donc pas nécessaire en général.

Le code ARPACK a été écrit de manière à imposer le minimum de contraintes à l'utilisateur. Le code ARPACK utilise une interface de communication qui ne nécessite que des passages de vecteurs. Deux routines sont alors nécessaires pour fournir à ARPACK le résultat des opérations : $(\mathbb{A}_b + \sigma \mathbb{C}_b)^{-1} \mathbf{Y}$ et $\mathbb{C}_b \mathbf{X}$. L'utilisateur est alors totalement libre sur le choix du stockage des matrices et sur l'algorithme permettant la recherche de la solution \mathbf{X} de $(\mathbb{A}_b + \sigma \mathbb{C}_b) \mathbf{X} = \mathbf{Y}$.

Les matrices \mathbb{A}_b et \mathbb{C}_b sont stockées de manière optimale sous forme creuse [23] (sparse matrix en anglais). L'inversion est basée sur l'algorithme LU, ce choix est justifié par le fait qu'un nombre relativement important d'inversions seront effectuées.

Deux registres mémoires complémentaires servent à stocker les vecteurs propres en cours de calcul et les vecteur d'Arnoldi calculés. Le nombre de vecteurs d'Arnoldi est spécifié par l'utilisateur, il est fixé de manière arbitraire à $3N_{\text{etape}}$.

La figure E.2 présente la démarche utilisée dans le code c++. Le décalage initial σ_0 importe peu tant qu'il est strictement positif, généralement $\sigma_0 = 0.1$. Le critère de convergence ϵ_{ARPACK} est interne à ARPACK, il porte sur les valeurs propres. En général ϵ_{ARPACK} est choisi égal à 10^{-12} .

Remarque importante : dans l'objectif de simplifier les notations, l'indice « , N » caractérisant l'approximation interne d'ordre N disparaît. Les modes propres calculés par la méthode des éléments finis ne sont donc plus notés $(z_{i,N}, V_{i,N})$ mais sont notés (z_i, V_i)

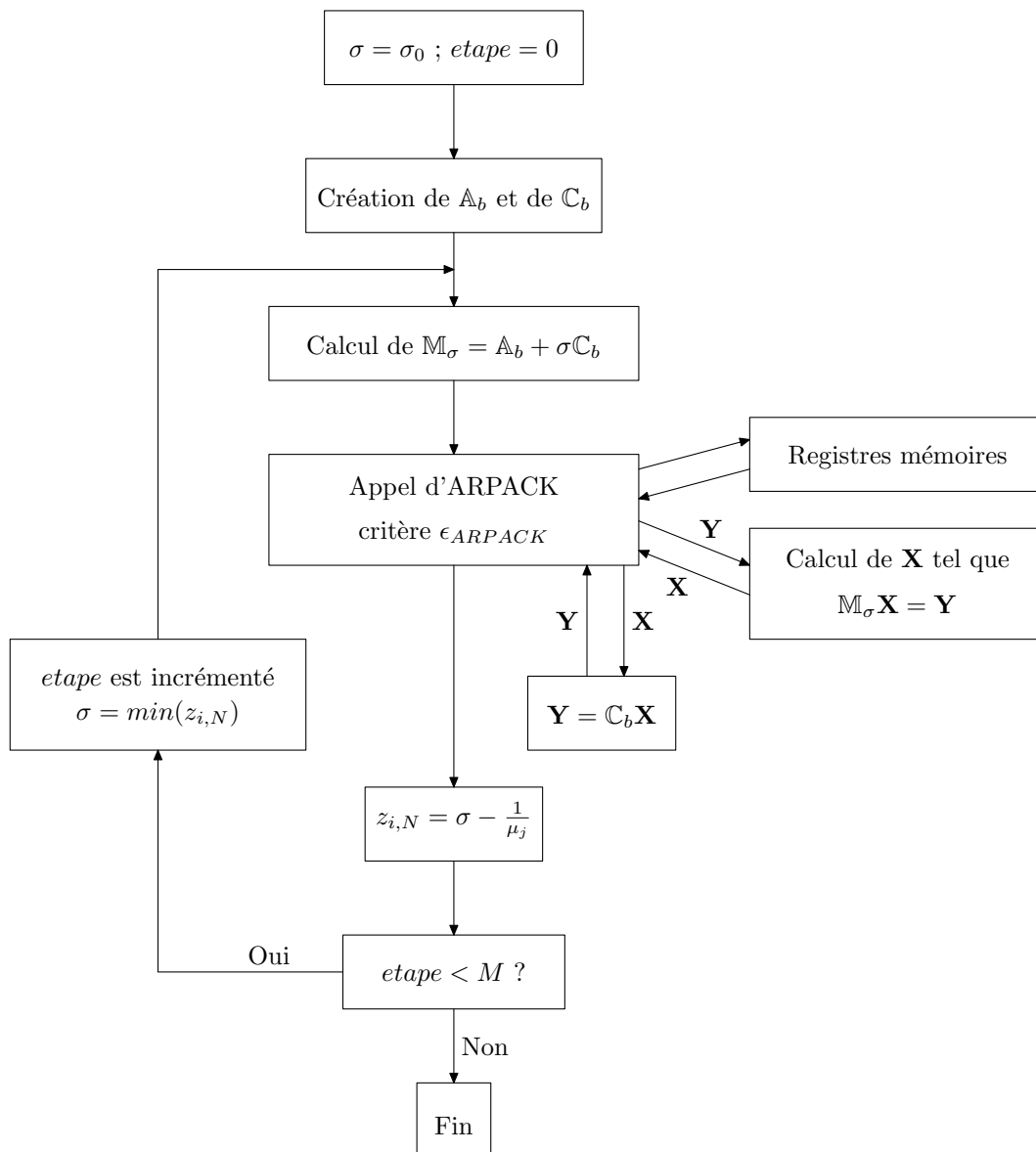


FIG. E.2 – Démarche de calcul des modes

Contenu du chapitre

I.	La base de branche calculée et la troncature de Marshall	140
II.	Notion de dominance	141
III.	L'amalgame découplé	142

L'équation d'état 2.14 couple tous les états dans le cas de l'analyse modale de branche, il est donc irréaliste de simuler un système thermique à l'aide d'une base modale comportant un grand nombre de modes. La réduction modale qui consiste à ne plus utiliser la base de branche complète mais une base réduite $\{\tilde{z}_i, \tilde{V}_i\}$ permet de mener des calculs à partir de la formulation modale.

Dans le cas de la méthode modale classique, les équations d'états D.35 sont découplées. L'utilisation d'une base modale complète est alors possible, mais l'usage d'une base réduite améliore le temps de calcul.

Les méthodes de réduction dans l'espace d'état ont été développées depuis longtemps par les automaticiens [7]. Le lecteur intéressé par les méthodes de réduction en analyse modale classique consultera les documents suivants : [48, 30].

Les méthodes existantes concernant la réduction de la base de branche sont : la troncature de Marshall, l'amalgame modal couplé et l'amalgame modal découplé. La procédure d'amalgame couplé nécessite une simulation du système, elle est ainsi très lente et n'est plus utilisée.

Cette annexe va détailler les deux méthodes de réduction utilisées dans le cas de la base de branche : la troncature de Marshall et l'amalgame modal découplé (qui sera désormais désigné par ses initiales : AMD).

I. La base de branche calculée et la troncature de Marshall

Marshall[31] a proposé la réduction qui porte son nom en 1966. La troncature de Marshall considère que les modes les plus importants sont ceux qui ont les constantes de temps les plus élevées. Un critère temporel τ_{crit} est alors considéré et seul les modes tels que $\tau_i > \tau_{crit}$ sont conservés. La troncature de Marshall est donc très simple mais présente l'inconvénient de conserver (en général) des modes peu énergétiques (l'énergie est définie par l'équation F.2) et d'éliminer potentiellement des modes très énergétiques dont la constante de temps serait inférieure à τ_{crit} .

I.1. Troncature de Marshall imposée par le maillage

Le calcul numérique des modes par la méthode des éléments finis est basé sur l'utilisation d'un maillage. Il est alors théoriquement possible de calculer N modes si le maillage est d'ordre N . Les modes calculés sont alors ceux de plus grande constante de temps. Une troncature de Marshall est ainsi imposée par le maillage.

I.2. Troncature de Marshall liée au choix de l'utilisateur

En pratique seul les \tilde{N}_0 premiers modes sont calculés à l'aide d'*ARPACK* sur les N modes possibles. Une troncature de Marshall complémentaire à celle imposée par le maillage est alors réalisée.

II. Notion de dominance

La troncature de Marshall est uniquement basée sur les constantes de temps, à l'opposé, l'AMD utilise un critère basé sur l'énergie des modes.

L'énergie de la réponse dynamique du système est notée \mathcal{D} et s'exprime par :

$$\mathcal{D} \stackrel{def}{=} \int_0^{t_{max}} \mathcal{C}_b(T, T) dt \quad (\text{F.1})$$

L'énergie ou dominance du mode (z_i, V_i) est notée \mathcal{D}_i et s'exprime par :

$$\mathcal{D}_i \stackrel{def}{=} \int_0^{t_{max}} (x_i)^2 dt \quad (\text{F.2})$$

L'énergie est ici une définition d'automaticien sans lien direct avec une énergie thermodynamique. L'horizon temporel usuellement utilisé en automatique est $[-\infty, \infty]$ ou $[0, \infty]$, la définition présentée (avec un horizon temporel égal à $[0, t_{max}]$) est liée à l'utilisation particulière qui sera faite de l'énergie.

\mathcal{D} s'exprime en fonction des \mathcal{D}_i :

$$\mathcal{D} = \sum_{i=1}^{\infty} \mathcal{D}_i \quad (\text{F.3})$$

F.3 se démontre en projetant le champ T dans l'espace modal et en tenant compte de la bi-linéarité de \mathcal{C}_b et de la condition d'orthogonalité E.25.

$$\mathcal{D} = \int_0^{t_{max}} \mathcal{C}_b(T, T) dt \quad (\text{F.4})$$

$$= \int_0^{t_{max}} \mathcal{C}_b \left(\sum_{i=1}^{\infty} x_i V_i, \sum_{k=1}^{\infty} x_k V_k \right) dt \quad (\text{F.5})$$

$$= \sum_{i=1}^{\infty} \left(\sum_{k=1}^{\infty} \left(\int_0^{t_{max}} x_i x_k \mathcal{C}_b(V_i, V_k) \right) \right) \quad (\text{F.6})$$

$$= \sum_{i=1}^{\infty} \left(\sum_{k=1}^{\infty} \left(\int_0^{t_{max}} x_i x_k \delta_{ik} \right) \right) \quad (\text{F.7})$$

$$= \sum_{i=1}^{\infty} \left(\int_0^{t_{max}} (x_i)^2 \right) \quad (\text{F.8})$$

$$= \sum_{i=1}^{\infty} \mathcal{D}_i \quad (\text{F.9})$$

Remarque : les états sont couplés par l'équation d'état, ainsi des modes ayant une dominance faible pourraient être nécessaires dans une simulation modale.

II.1. Energie de l'écart de la réponse dynamique

Soit T (\tilde{T}) le champ de température d'un modèle modal comportant un grand (petit) nombre de modes. L'énergie de l'écart de la réponse dynamique est notée \mathcal{E} et s'exprime par :

$$\mathcal{E} = \int_0^{t_{max}} \mathcal{C}_b(T - \tilde{T}, T - \tilde{T}) dt \quad (\text{F.10})$$

III. L'amalgame découpé

L'amalgame modal a été adapté à l'analyse modale classique par Oulefki [38]. L'AMD adapté aux modes de branche est le fruit des auteurs de l'article [43].

III.1. Introduction à l'amalgame

L'AMD s'applique à une base modale calculée numériquement. Du fait de la troncature de Marshall effectuée lors du calcul de la base, celle-ci comporte les \tilde{N}_0 premiers modes de branche (z_i, V_i) .

L'utilisateur définit un cas test linéaire et choisit la dimension \tilde{N}_1 de la base des modes amalgamés $(\tilde{z}_i, \tilde{V}_i)$. L'AMD utilise le cas test linéaire en négligeant¹ les couplages entre modes ce qui permet de déterminer analytiquement les états $x_i(t)$.

L'AMD détermine \tilde{N}_1 modes maîtres parmi les \tilde{N}_0 modes de la base d'origine. Les $\tilde{N}_0 - \tilde{N}_1$ modes restants sont alors appelés modes esclaves. Le p -ième mode maître est noté $(z_{p,1}, V_{p,1})$.

Un sous-espace propre est alors affecté à chaque mode maître. Les modes esclaves sont affectés les uns à la suite des autres dans un unique sous-espace propre. Les p sous-espaces propres comportent alors n_p modes dont 1 mode maître et $n_p - 1$ modes esclaves. Les modes esclaves associés au mode maître $(z_{p,1}, V_{p,1})$ sont notés $(z_{p,k}, V_{p,k})$ avec $k \in \llbracket 2, n_p \rrbracket$.

Les \tilde{N}_0 modes de branche sont affectés à un unique sous-espace propre, ainsi :

$$\tilde{N}_0 = \sum_{p=1}^{\tilde{N}_1} n_p \quad (\text{F.11})$$

Le p -ième mode réduit $(\tilde{z}_p, \tilde{V}_p)$ est construit par combinaison linéaire des vecteurs propres du sous-espace associé au p -ième mode maître.

$$\tilde{z}_p = z_{p,1} \quad (\text{F.12})$$

$$\tilde{V}_p = \sum_{k=1}^{n_p} \alpha_{p,k} V_{p,k} \quad (\text{F.13})$$

Les coefficients $\alpha_{p,k}$ sont déterminés en considérant que :

- L'état du mode amalgamé \tilde{x}_p est égal à l'état du mode maître correspondant $x_{p,1}$ lorsque la base n'est pas réduite.
- Les coefficients $\alpha_{p,k}$ sont tels que l'énergie de l'écart \mathcal{E} soit minimisée. Il y a alors un choix à effectuer sur la valeur à considérer pour t_{max} , ce point sera revu ultérieurement.

Les modes amalgamés sont ensuite normalisés par rapport à \mathcal{C}_b , ils sont orthogonaux deux-à-deux par rapport à \mathcal{C}_b . **Attention** les modes amalgamés ne sont pas des modes de branche et ne sont pas orthogonaux par rapport à \mathcal{A}_b .

III.2. Éléments nécessaires pour la procédure d'amalgame

La procédure d'amalgame sera présentée en détail dès que les éléments nécessaires à la procédure d'amalgame auront été présentés.

¹Si les couplages entre modes ne sont pas négligés alors l'amalgame est dit couplé et il faut déterminer numériquement les états au cours du temps. L'amalgame couplé est très lent et n'est plus utilisé.

III.2.1. Expression de l'énergie de l'écart

Le champ de température T déterminé par le modèle modal utilisant les \tilde{N}_0 premiers modes de branche (z_i, V_i) est :

$$T = \sum_{i=1}^{\tilde{N}_0} x_i V_i \quad (\text{F.14})$$

Il existe une bijection entre les (p, k) et i , il est donc possible d'écrire sans ambiguïté que :

$$T = \sum_{p=1}^{\tilde{N}_1} \left(\sum_{k=1}^{n_p} x_{p,k} V_{p,k} \right) \quad (\text{F.15})$$

Le champ de température \tilde{T} déterminé par le modèle modal utilisant les \tilde{N}_1 modes amalgamés $(\tilde{z}_p, \tilde{V}_p)$ est :

$$\tilde{T} = \sum_{p=1}^{\tilde{N}_1} \tilde{x}_p \tilde{V}_p \quad (\text{F.16})$$

Il est possible de remplacer \tilde{V}_p par son expression F.13 :

$$\tilde{T} = \sum_{p=1}^{\tilde{N}_1} \left(\tilde{x}_p \sum_{k=1}^{n_p} \alpha_{p,k} V_{p,k} \right) = \sum_{p=1}^{\tilde{N}_1} \left(\sum_{k=1}^{n_p} \tilde{x}_p \alpha_{p,k} V_{p,k} \right) \quad (\text{F.17})$$

La procédure d'AMD fait l'hypothèse que la dynamique (ou état) \tilde{x}_p du p -ième mode amalgamé est égale à la dynamique $x_{p,1}$ du mode maître lorsqu'il n'est pas amalgamé d'où :

$$\tilde{T} = \sum_{p=1}^{\tilde{N}_1} \left(\sum_{k=1}^{n_p} \alpha_{p,k} x_{p,1} V_{p,k} \right) \quad (\text{F.18})$$

L'écart entre le champ T et le champ \tilde{T} s'exprime alors par :

$$T - \tilde{T} = \sum_{p=1}^{\tilde{N}_1} \left(\sum_{k=1}^{n_p} (x_{p,k} - \alpha_{p,k} x_{p,1}) V_{p,k} \right) \quad (\text{F.19})$$

La relation d'orthogonalité E.25 est équivalente à :

$$\forall (p, q) \in \llbracket 1, \tilde{N}_1 \rrbracket \quad \forall k \in \llbracket 1, n_p \rrbracket \quad \forall l \in \llbracket 1, n_q \rrbracket \quad \mathcal{C}_b(V_{p,k}, V_{q,l}) = \delta_{p,q} \delta_{k,l} \quad (\text{F.20})$$

Il est possible de montrer que l'énergie de l'écart \mathcal{E} s'exprime alors par :

$$\mathcal{E} = \sum_{p=1}^{\tilde{N}_1} \sum_{k=1}^{n_p} \mathcal{E}_{p,k} \quad (\text{F.21})$$

L'énergie de l'écart du k -ième mode du p -ième sous-espace étant définie par :

$$\mathcal{E}_{p,k} = \int_0^{t_{max}} (x_{p,k} - \alpha_{p,k} x_{p,1})^2 dt \quad (\text{F.22})$$

La démonstration de F.21 est simple. Dans un premier temps, il faut remplacer $T - \tilde{T}$ par la décomposition F.19 dans la définition de l'énergie de l'écart F.10. La bi-linéarité de \mathcal{C}_b et la relation d'orthogonalité F.20 permettent de simplifier l'expression obtenue.

$$\mathcal{E} = \int_0^{t_{max}} \mathcal{C}_b \left(T - \tilde{T}, T - \tilde{T} \right) dt \quad (\text{F.23})$$

$$= \int_0^{t_{max}} \mathcal{C}_b \left(\sum_{p=1}^{\tilde{N}_1} \left(\sum_{k=1}^{n_p} (x_{p,k} - \alpha_{p,k} x_{p,1}) V_{p,k} \right), \sum_{q=1}^{\tilde{N}_1} \left(\sum_{l=1}^{n_q} (x_{q,l} - \alpha_{q,l} x_{q,1}) V_{q,l} \right) \right) dt \quad (\text{F.24})$$

$$= \sum_{p=1}^{\tilde{N}_1} \sum_{k=1}^{n_p} \sum_{q=1}^{\tilde{N}_1} \sum_{l=1}^{n_q} \int_0^{t_{max}} (x_{p,k} - \alpha_{p,k} x_{p,1}) (x_{q,l} - \alpha_{q,l} x_{q,1}) \mathcal{C}_b(V_{p,k} V_{q,l}) dt \quad (\text{F.25})$$

$$= \sum_{p=1}^{\tilde{N}_1} \sum_{k=1}^{n_p} \int_0^{t_{max}} (x_{p,k} - \alpha_{p,k} x_{p,1})^2 dt \quad (\text{F.26})$$

$$= \sum_{p=1}^{\tilde{N}_1} \sum_{k=1}^{n_p} \mathcal{E}_{p,k} \quad (\text{F.27})$$

III.2.2. Valeur optimale des coefficients d'amalgame

L'AMD cherche à minimiser l'énergie de l'écart F.21, or les $\mathcal{E}_{p,k}$ sont positifs par définition. Il faut donc minimiser chacun des $\mathcal{E}_{p,k}$.

Il faut donc étudier les variations de $\mathcal{E}_{p,k}$ en fonction du coefficient d'amalgame $\alpha_{p,k}$.

L'expression F.22 de $\mathcal{E}_{p,k}$ est développée :

$$\mathcal{E}_{p,k} = \int_0^{t_{max}} (x_{p,k} - \alpha_{p,k} x_{p,1})^2 dt \quad (\text{F.28})$$

$$= \int_0^{t_{max}} (x_{p,k})^2 dt - 2\alpha_{p,k} \int_0^{t_{max}} x_{p,1} x_{p,k} dt + (\alpha_{p,k})^2 \int_0^{t_{max}} (x_{p,1})^2 dt \quad (\text{F.29})$$

La dérivée de $\mathcal{E}_{p,k}$ en fonction de $\alpha_{p,k}$ est :

$$\frac{d\mathcal{E}_{p,k}}{d\alpha_{p,k}} = -2 \int_0^{t_{max}} x_{p,1} x_{p,k} dt + 2\alpha_{p,k} \int_0^{t_{max}} (x_{p,1})^2 dt \quad (\text{F.30})$$

L'intégrale $\int_0^{t_{max}} (x_{p,1})^2 dt$ est strictement positive : il est évident qu'elle est positive, et de plus $x_{p,1}$ désigne l'état d'un mode maître².

Ainsi :

$$\lim_{\alpha_{p,k} \rightarrow -\infty} \mathcal{E}_{p,k} = \infty \quad \text{et} \quad \lim_{\alpha_{p,k} \rightarrow -\infty} \frac{d\mathcal{E}_{p,k}}{d\alpha_{p,k}} = -\infty \quad (\text{F.31})$$

$$\lim_{\alpha_{p,k} \rightarrow \infty} \mathcal{E}_{p,k} = \infty \quad \text{et} \quad \lim_{\alpha_{p,k} \rightarrow \infty} \frac{d\mathcal{E}_{p,k}}{d\alpha_{p,k}} = \infty \quad (\text{F.32})$$

²Dans la démarche de l'AMD, les états sont connus analytiquement (F.45, F.46). Pour qu'un mode ne soit pas excité, il faut que $x_i(0)$ et que $\mathcal{A}(V_i)$ soient nuls. En pratique $x_i(0)$ est souvent nul, car le champ initial est plat et dans ce cas seul $x_1(0)$ n'est pas nul. Par contre $\mathcal{A}(V_i)$ est très rarement nul. Il faudrait ainsi que la base réduite comporte un très grand nombre de modes pour que $\mathcal{A}(V_{p,1})$ puisse être nul. Or la dimension de la base réduite est faible, ainsi le cas pathologique où $\int_0^{t_{max}} (x_{p,1})^2 dt$ est nulle ne se rencontre jamais.

Le tableau des variations de $\mathcal{E}_{p,k}$ en fonction de $\alpha_{p,k}$ est :

$\alpha_{p,k}$	$-\infty$	$\alpha_{p,k,opt}$	∞
$\frac{d\mathcal{E}_{p,k}}{d\alpha_{p,k}}$	$-\infty$	$-$	0
		$+$	∞
$\mathcal{E}_{p,k}$	∞	\searrow	\nearrow
		$min(\mathcal{E}_{p,k})$	

Il existe donc un minimum pour $\mathcal{E}_{p,k}$ en fonction de $\alpha_{p,k}$. Ce minimum est atteint lorsque $\alpha_{p,k}$ prend la valeur optimale $\alpha_{p,k,opt}$:

$$\alpha_{p,k,opt} = \frac{\int_0^{t_{max}} x_{p,1} x_{p,k} dt}{\int_0^{t_{max}} (x_{p,1})^2 dt} = \frac{\int_0^{t_{max}} x_{p,1} x_{p,k} dt}{\mathcal{D}_{p,1}} \quad (\text{F.33})$$

Le coefficient d'amalgame d'un mode maître est ainsi égal à 1 : $\alpha_{p,1} = 1$. Et alors $\mathcal{E}_{p,1} = 0$, ce qui n'est pas étonnant.

III.2.3. Expression de $\mathcal{E}_{p,k,opt}$

Si $\alpha_{p,k}$ prend la valeur optimale $\alpha_{p,k,opt}$ alors l'expression F.22 de $\mathcal{E}_{p,k}$ se simplifie en :

$$\mathcal{E}_{p,k,opt} = \mathcal{D}_{p,k} - (\alpha_{p,k,opt})^2 \mathcal{D}_{p,1} \quad (\text{F.34})$$

En effet :

$$\mathcal{E}_{p,k,opt} = \int_0^{t_{max}} (x_{p,k} - \alpha_{p,k,opt} x_{p,1})^2 dt \quad (\text{F.35})$$

$$= \int_0^{t_{max}} (x_{p,k})^2 dt - 2\alpha_{p,k,opt} \int_0^{t_{max}} x_{p,1} x_{p,k} dt + (\alpha_{p,k,opt})^2 \int_0^{t_{max}} (x_{p,1})^2 dt \quad (\text{F.36})$$

$$= \int_0^{t_{max}} (x_{p,k})^2 dt - 2 \frac{\left(\int_0^{t_{max}} x_{p,1} x_{p,k} dt \right)^2}{\int_0^{t_{max}} (x_{p,1})^2 dt} + \frac{\left(\int_0^{t_{max}} x_{p,1} x_{p,k} dt \right)^2}{\int_0^{t_{max}} (x_{p,1})^2 dt} \quad (\text{F.37})$$

$$= \int_0^{t_{max}} (x_{p,k})^2 dt - \frac{\left(\int_0^{t_{max}} x_{p,1} x_{p,k} dt \right)^2}{\int_0^{t_{max}} (x_{p,1})^2 dt} \quad (\text{F.38})$$

$$= \int_0^{t_{max}} (x_{p,k})^2 dt - \alpha_{p,k,opt} \int_0^{t_{max}} x_{p,1} x_{p,k} dt \quad (\text{F.39})$$

$$= \int_0^{t_{max}} (x_{p,k})^2 dt - (\alpha_{p,k,opt})^2 \int_0^{t_{max}} (x_{p,1})^2 dt \quad (\text{F.40})$$

III.2.4. Détermination analytique des états

L'AMD est qualifié de découplé car il néglige les couplages entre modes dans l'équation d'état 2.16, qui devient alors :

$$\left\{ \begin{array}{l} \forall i \in \llbracket 1, \tilde{N}_0 \rrbracket \quad \mathcal{C}(V_i, V_i) \frac{dx_i(t)}{dt} = -(\mathcal{K}(V_i, V_i) + \mathcal{B}(V_i, V_i)) x_i(t) + \mathcal{A}(V_i) \\ \forall i \in \llbracket 1, \tilde{N}_0 \rrbracket \quad x_i(0) = \mathcal{C}_b(T_i, V_i) \end{array} \right. \quad (\text{F.41})$$

Le cas test est linéaire, les opérateurs $\mathcal{C}, \mathcal{K}, \mathcal{B}$ et \mathcal{A} sont donc fixes. De plus \mathcal{C} est un opérateur défini et V_i est non nul, il est alors possible de diviser par $\mathcal{C}(V_i, V_i)$. Le problème F.41 est ré-écrit :

$$\left\{ \begin{array}{l} \forall i \in \llbracket 1, \tilde{N}_0 \rrbracket \quad \frac{dx_i(t)}{dt} = m_i x_i(t) + n_i \\ \forall i \in \llbracket 1, \tilde{N}_0 \rrbracket \quad x_i(0) = \mathcal{C}_b(T_i, V_i) \end{array} \right. \quad (\text{F.42})$$

Les m_i, n_i sont des constantes dont l'expression est :

$$m_i = -\frac{\mathcal{K}(V_i, V_i) + \mathcal{B}(V_i, V_i)}{\mathcal{C}(V_i, V_i)} \quad (\text{F.43})$$

$$n_i = \frac{\mathcal{A}(V_i)}{\mathcal{C}(V_i, V_i)} \quad (\text{F.44})$$

La constante m_i est nécessairement négative car les opérateurs \mathcal{K}, \mathcal{B} et \mathcal{C} sont positifs. Il existe une solution analytique à F.42 :

$$\forall i \in \llbracket 1, \tilde{N}_0 \rrbracket \quad x_i(t) = a_i e^{m_i t} - \frac{n_i}{m_i} \quad (\text{F.45})$$

Avec :

$$\forall i \in \llbracket 1, \tilde{N}_0 \rrbracket \quad a_i = x_i(0) + \frac{n_i}{m_i} \quad (\text{F.46})$$

Les états sont désormais connus analytiquement.

III.2.5. Expression analytique de $\mathcal{E}_{p,k,opt}$ et de $\alpha_{p,k,opt}$

Il s'agit ici d'exprimer $\mathcal{E}_{p,k,opt}$ et $\alpha_{p,k,opt}$ en fonction de $m_{p,k}, n_{p,k}, a_{p,k}$ et de t_{max} . Il faut donc calculer les différentes intégrales qui interviennent dans F.33 et dans F.34. Soit les intégrales suivantes : $\int_0^{t_{max}} (x_i)^2 dt$ (\mathcal{D}_i) et $\int_0^{t_{max}} x_i x_j dt$.

L'énergie du i -ème mode s'exprime analytiquement par :

$$\mathcal{D}_i = \int_0^{t_{max}} (x_i)^2 dt \quad (\text{F.47})$$

$$= \int_0^{t_{max}} \left(a_i e^{m_i t} - \frac{n_i}{m_i} \right)^2 dt \quad (\text{F.48})$$

$$= \int_0^{t_{max}} a_i^2 e^{2m_i t} - \frac{2a_i n_i}{m_i} e^{m_i t} + \left(\frac{n_i}{m_i} \right)^2 dt \quad (\text{F.49})$$

$$= \left[\frac{a_i^2}{2m_i} e^{2m_i t} - \frac{2a_i n_i}{m_i^2} e^{m_i t} + \left(\frac{n_i}{m_i} \right)^2 t \right]_0^{t_{max}} \quad (\text{F.50})$$

L'expression analytique de \mathcal{D}_i est donc :

$$\mathcal{D}_i = \frac{a_i^2}{2m_i} (e^{2m_i t_{max}} - 1) - \frac{2a_i n_i}{m_i^2} (e^{m_i t_{max}} - 1) + \left(\frac{n_i}{m_i}\right)^2 t_{max} \quad (\text{F.51})$$

L'hypothèse de découplage des modes à rendu possible le calcul analytique des états et donc de l'énergie de ceux-ci. L'expression analytique de \mathcal{D}_i montre qu'il faut que t_{max} soit fini³. Il faut alors trouver un compromis acceptable sur la valeur de t_{max} qui permette un équilibre entre les différents termes de F.51. D'après les tests effectués pour l'article [43], le choix $t_{max} = 5\tau_2$ donne de bons résultats.

Le calcul analytique de $\int_0^{t_{max}} x_i x_j dt$ est effectué dans les lignes suivantes :

$$\int_0^{t_{max}} x_i x_j dt = \int_0^{t_{max}} \left(a_i e^{m_i t} - \frac{n_i}{m_i} \right) \left(a_j e^{m_j t} - \frac{n_j}{m_j} \right) dt \quad (\text{F.52})$$

$$= \int_0^{t_{max}} a_i a_j e^{(m_i+m_j)t} - \frac{a_i n_j}{m_j} e^{m_i t} - \frac{n_i a_j}{m_i} e^{m_j t} + \frac{n_i n_j}{m_i m_j} dt \quad (\text{F.53})$$

$$= \left[\frac{a_i a_j}{m_i + m_j} e^{(m_i+m_j)t} - \frac{a_i n_j}{m_i m_j} e^{m_i t} - \frac{n_i a_j}{m_i m_j} e^{m_j t} + \frac{n_i n_j}{m_i m_j} t \right]_0^{t_{max}} \quad (\text{F.54})$$

Soit :

$$\int_0^{t_{max}} x_i x_j dt = \frac{a_i a_j}{m_i + m_j} (e^{(m_i+m_j)t_{max}} - 1) - \frac{a_i n_j}{m_i m_j} (e^{m_i t_{max}} - 1) - \frac{n_i a_j}{m_i m_j} (e^{m_j t_{max}} - 1) + \frac{n_i n_j}{m_i m_j} t_{max} \quad (\text{F.55})$$

Le calcul analytique de $\alpha_{p,k}$ s'obtient en utilisant F.55 pour calculer le numérateur et F.51 pour calculer le dénominateur :

$$\alpha_{p,k,opt} = \frac{\int_0^{t_{max}} x_{p,1} x_{p,k} dt}{\mathcal{D}_{p,1}} \quad (\text{F.56})$$

Le calcul analytique de $\mathcal{E}_{p,k}$ s'obtient en utilisant F.51 pour calculer $\mathcal{D}_{p,k}$ et $\mathcal{D}_{p,1}$, le calcul de $\alpha_{p,k,opt}$ étant obtenu à l'aide de F.56 :

$$\mathcal{E}_{p,k,opt} = \mathcal{D}_{p,k} - (\alpha_{p,k,opt})^2 \mathcal{D}_{p,1} \quad (\text{F.57})$$

Les termes exponentiels ($e^{m_i t_{max}}$, ...) sont numériquement négligeables la plupart du temps, car m_i est négatif. La valeur de $e^{m_i t_{max}}$ est inférieure à 0.01 dès que $-4.5 > m_i t_{max}$. Cependant la plupart du temps ne signifie pas que c'est toujours le cas, il ne faut donc pas supprimer les termes exponentiels⁴.

³Dans le cas contraire les modes tels que $n_i \neq 0$ auront une énergie infinie.

⁴Contrairement à ce qui a été écrit dans [43] (equations 52 et 53)

III.3. La procédure d'amalgame

La procédure d'AMD débute sur la base d'une base de branche tronquée comportant \tilde{N}_0 modes de branche. L'ordre \tilde{N}_1 de la base réduite est fixé par l'utilisateur. L'utilisateur détermine aussi le cas test.

Etape 1 : détermination des coefficients m_i, n_i, a_i et choix de t_{max} .

L'ensemble des coefficients m_i, n_i, a_i sont calculés par rapport au cas test choisi grâce aux formules F.43, F.44 et F.46. Le temps d'intégration est fixé par défaut à $5\tau_2$, il est possible de changer cette valeur, mais ce choix est acceptable par défaut.

Etape 2 : détermination du premier mode maître.

Le premier mode maître est le mode qui possède la plus grande énergie (\mathcal{D}_i). Le calcul de l'énergie est déterminé à l'aide de la formule F.51

Etape 3 : détermination des $\tilde{N}_1 - 1$ autres modes maîtres.

Le processus qui suit est répété $\tilde{N}_1 - 1$ fois :

q modes maîtres sont déterminés. Les sous-espaces propres sont provisoirement formés comme proposé à l'étape 4. Le mode qui contribue le plus à l'énergie de l'écart est choisi comme $(q + 1)$ -ème mode maître.

Etape 4 : Création des \tilde{N}_1 sous-espaces propres.

Chaque mode esclave est affecté dans le sous-espace propre où il contribue le moins à l'énergie de l'écart. Pour cela il faut calculer la contribution à l'énergie de l'écart ($\mathcal{E}_{p,k,opt}$) pour chaque mode maître à l'aide de F.57.

Etape 5 : Calcul des $\tilde{N}_0 - \tilde{N}_1$ coefficients d'amalgame.

Les coefficients d'amalgame des modes esclaves sont calculés à l'aide de F.56. Le coefficient d'amalgame du mode maître est égal à 1.

Etape 6 : Calcul des modes amalgamés non normés

Les modes amalgamés non normés sont calculés à l'aide de la combinaison linéaire F.13.

Etape 7 : Normalisation des modes amalgamés

Les modes amalgamés sont normalisés par rapport à \mathcal{C}_b . La base amalgamée sera alors orthonormale par rapport à \mathcal{C}_b .

III.4. Propriété utile de la base amalgamée

Les modes amalgamés sont orthonormaux par rapport à \mathcal{C}_b ce qui implique qu'il existe un équivalent à E.44 pour les modes amalgamés. Si le champ \tilde{T} est décomposable dans la base des modes amalgamés alors :

$$\tilde{T} = \sum_{k=1}^{\tilde{N}_1} \tilde{x}_k \tilde{V}_k \quad (\text{F.58})$$

Les états (initiaux ou non) peuvent être déterminés à l'aide de \mathcal{C}_b :

$$\tilde{x}_i(t) = \mathcal{C}_b(\tilde{T}(t), \tilde{V}_i) \quad (\text{F.59})$$

Annexe G

Expressions matricielles

Contenu du chapitre

I.	Cas mono-corps	150
II.	Cas multi-corps	151
III.	Projection entre maillages	153

Cette annexe vise à donner les éléments utiles pour pouvoir implémenter une méthode modale de branche multi-corps dans un code éléments-finis.

I. Cas mono-corps

Dans un premier temps, il est utile de comprendre l'implémentation de la méthode modale dans le cas où il y a un seul corps.

I.1. Les matrices du modèle détaillé

Les matrices du modèle détaillé sont celles des opérateurs $\mathcal{C}, \mathcal{K}, \mathcal{B}$, elles sont associées au vecteur lié à \mathcal{A} (voir page 29). Le schéma temporel choisi est implicite ce qui donne :

$$\mathbb{C} \frac{\mathbf{T}^{n+1} - \mathbf{T}^n}{\delta t^n} = -\mathbb{K}\mathbf{T}^{n+1} - \mathbb{B}\mathbf{T}^{n+1} + \mathbf{A} \quad (\text{G.1})$$

\mathbf{T}^n est le vecteur des températures à l'étape n . Les matrices $\mathbb{C}, \mathbb{K}, \mathbb{B}$ ne sont pas invariantes du fait de la prise en compte de non-linéarités : elles sont évaluées par rapport à un champ de température. Pour des raisons de simplicité, les matrices sont évaluées en fonction de \mathbf{T}^n .

S'il y a N nœuds dans le maillage alors les vecteurs colonnes $\mathbf{T}^n, \mathbf{T}^{n+1}, \mathbf{A}$ comportent N lignes et les matrices $\mathbb{C}, \mathbb{K}, \mathbb{B}$ comportent N lignes et N colonnes.

Le système matriciel à résoudre (l'inconnue étant \mathbf{T}^{n+1}) est :

$$(\mathbb{C} + (\mathbb{K} + \mathbb{B})\delta t^n) \mathbf{T}^{n+1} = \mathbb{C}\mathbf{T}^n + \mathbf{A}\delta t^n \quad (\text{G.2})$$

I.2. Les matrices du modèle réduit

Soit \mathbb{V} la matrice des modes. Le champ de température $\tilde{\mathbf{T}}$ s'exprime par :

$$\tilde{\mathbf{T}}^n = \mathbb{V}\mathbf{X}^n \quad (\text{G.3})$$

Dans le cas où il y a \tilde{N} modes et N nœuds, la matrice des modes comporte \tilde{N} colonnes et N lignes.

\mathbf{T} est remplacé par $\tilde{\mathbf{T}}$ dans G.2 :

$$(\mathbb{C} + (\mathbb{K} + \mathbb{B})\delta t^n) \tilde{\mathbf{T}}^{n+1} = \mathbb{C}\tilde{\mathbf{T}}^n + \mathbf{A}\delta t^n \quad (\text{G.4})$$

L'utilisation de G.3 permet de faire apparaître le vecteur d'état :

$$(\mathbb{C} + (\mathbb{K} + \mathbb{B})\delta t^n) \mathbb{V}\mathbf{X}^{n+1} = \mathbb{C}\mathbb{V}\mathbf{X}^n + \mathbf{A}\delta t^n \quad (\text{G.5})$$

L'équation matricielle G.5 comporte plus d'équations que d'inconnues. Il est souhaitable de réduire le nombre d'équations au nombre d'inconnues \tilde{N} , pour cela G.5 est multiplié à gauche par ${}^t\mathbb{V}$, ce qui donne le système matriciel recherché (l'inconnue étant \mathbf{X}^{n+1}) :

$$({}^t\mathbb{V}(\mathbb{C} + (\mathbb{K} + \mathbb{B})\delta t^n)\mathbb{V}) \mathbf{X}^{n+1} = {}^t\mathbb{V}\mathbb{C}\mathbb{V}\mathbf{X}^n + {}^t\mathbb{V}\mathbf{A}\delta t^n \quad (\text{G.6})$$

L'équation matricielle G.6 comporte autant d'équations que d'inconnues. Les vecteurs colonnes $\mathbf{X}^n, \mathbf{X}^{n+1}, {}^t\mathbb{V}\mathbf{A}$ comportent \tilde{N} lignes, les matrices ${}^t\mathbb{V}(\mathbb{C} + (\mathbb{K} + \mathbb{B})\delta t^n), {}^t\mathbb{V}\mathbb{C}\mathbb{V}$ comportent \tilde{N} lignes et \tilde{N} colonnes.

La multiplication par ${}^t\mathbb{V}$ est comparable à une minimisation au sens des moindres carrés. L'écart lié à la réduction modale est réparti sur l'ensemble du domaine physique. De plus G.6 est une formulation matricielle de l'équation d'état 2.16 obtenue numériquement à l'aide d'une utilisation des matrices éléments finis.

II. Cas multi-corps

S'il y a plusieurs corps couplés par des interfaces de contact, alors il y a un maillage pour chaque corps. Lorsque le maillage s'effectue de manière indépendante (par rapport aux autres corps), les nœuds des frontières communes ne coïncident pas. Des opérateurs de projection permettront alors le transfert du champ de température d'un corps vers un autre au niveau des interfaces de contact. Ces opérateurs seront présentés en section III. La démarche de couplage va être exposée dans le cas où il y a deux corps en contact :

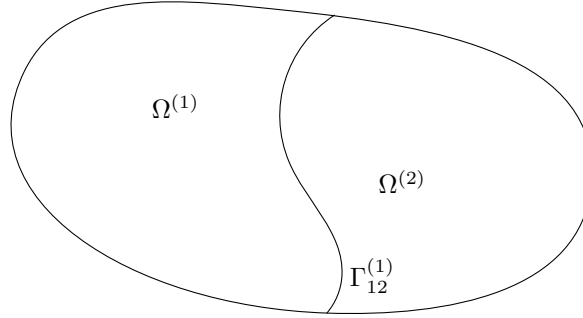


FIG. G.1 – Domaine Ω constitué de deux sous domaines $\Omega^{(1)}$ et $\Omega^{(2)}$

II.1. Les matrices du modèle détaillé

Il est possible de définir un problème mono-corps pour chaque corps, avec un flux lié à la condition de couplage par conduction :

$$(\mathbb{C}_1 + (\mathbb{K}_1 + \mathbb{B}_1)\delta t^n) \mathbf{T}_1^{n+1} = \mathbb{C}_1 \mathbf{T}_1^n + \mathbf{A}_1 \delta t^n + \mathbb{C}_{cd1}(\mathbf{T}_{2 \rightarrow 1}^{n+1} - \mathbf{T}_1^{n+1})\delta t^n \quad (\text{G.7})$$

$$(\mathbb{C}_2 + (\mathbb{K}_2 + \mathbb{B}_2)\delta t^n) \mathbf{T}_2^{n+1} = \mathbb{C}_2 \mathbf{T}_2^n + \mathbf{A}_2 \delta t^n + \mathbb{C}_{cd2}(\mathbf{T}_{1 \rightarrow 2}^{n+1} - \mathbf{T}_2^{n+1})\delta t^n \quad (\text{G.8})$$

L'indice 1,2 désigne respectivement le corps $\Omega^{(1)}$, $\Omega^{(2)}$. Les matrices \mathbb{C}_{cd1} et \mathbb{C}_{cd2} traduisent le couplage par conduction, elle sont en tout point similaires aux matrices éléments finis qui décrivent une condition convective. $\mathbf{T}_{2 \rightarrow 1}$ désigne le vecteur des températures du corps 2 projeté sur le corps 1, la dimension de ce vecteur est donc égale au nombre de nœuds du maillage du corps 1.

La condition de contact est prise en compte de manière implicite, comme pour les autres termes. Cependant une prise en compte implicite est obligatoire car dans le cas de résistances de contact très faibles ($T_{(i,P)}$ désigne la température du corps i au point P) :

- une prise en compte explicite conduit à :

$$\forall P \in \Gamma_{12}^{(1)} \quad T_{(1,P)}^{n+1} = T_{(2,P)}^n \quad \text{et} \quad T_{(2,P)}^{n+1} = T_{(1,P)}^n \quad (\text{G.9})$$

Ainsi la température au niveau de l'interface n'évolue pas.

- une prise en compte implicite conduit à :

$$\forall P \in \Gamma_{12}^{(1)} \quad T_{(1,P)}^{n+1} = T_{(2,P)}^{n+1} \quad (\text{G.10})$$

Ce qui est cohérent avec une condition de contact parfait.

ce qui est obligatoire si l'on désire pouvoir traiter le cas de résistances de contact très faibles.

Les projecteurs $\mathbb{P}_{1 \rightarrow 2}$ et $\mathbb{P}_{2 \rightarrow 1}$ (voir sous-section III.1. en page 153) sont tels que : $\mathbf{T}_{2 \rightarrow 1} = \mathbb{P}_{2 \rightarrow 1} \mathbf{T}_2$ et $\mathbf{T}_{1 \rightarrow 2} = \mathbb{P}_{1 \rightarrow 2} \mathbf{T}_1$.

Les équations G.7 et G.8 sont regroupées dans un seul système matriciel :

$$\begin{pmatrix} \mathbf{C}_1 + (\mathbb{K}_1 + \mathbb{B}_1 + \mathbf{C}_{cd1})\delta t^n & -\delta t^n \mathbf{C}_{cd1} \mathbb{P}_{2 \rightarrow 1} \\ -\delta t^n \mathbf{C}_{cd2} \mathbb{P}_{1 \rightarrow 2} & \mathbf{C}_2 + (\mathbb{K}_2 + \mathbb{B}_2 + \mathbf{C}_{cd2})\delta t^n \end{pmatrix} \begin{pmatrix} \mathbf{T}_1^{n+1} \\ \mathbf{T}_2^{n+1} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_1 \mathbf{T}_1^n + \mathbf{A}_1 \delta t^n \\ \mathbf{C}_2 \mathbf{T}_2^n + \mathbf{A}_2 \delta t^n \end{pmatrix} \quad (\text{G.11})$$

L'inconnue du système est le vecteur global des températures qui est composé par les vecteurs des températures des deux corps : \mathbf{T}_1^{n+1} et \mathbf{T}_2^{n+1} .

Les matrices $\mathbf{C}_{cd1}, \mathbf{C}_{cd2}$ traduisant le couplage par conduction sont estimées par rapport à \mathbf{T}^n comme pour les autres matrices. La matrice \mathbf{C}_{cd1} comporte N_1 lignes et N_1 colonnes, N_1 étant le nombre de nœuds du maillage du corps 1. De la même manière la matrice \mathbf{C}_{cd2} comporte N_2 lignes et N_2 colonnes, N_2 étant le nombre de nœuds du maillage du corps 2.

II.2. Les matrices du modèle réduit sans la fonctionnelle de saut de flux

Le système matriciel du modèle modal est obtenu en suivant la démarche proposée en mono-corps soit :

$$\begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{X}_1^{n+1} \\ \mathbf{X}_2^{n+1} \end{pmatrix} = \begin{pmatrix} \mathbf{N}_1 \\ \mathbf{N}_2 \end{pmatrix} \quad (\text{G.12})$$

Avec :

$$\begin{cases} \mathbf{M}_{11} = {}^t \mathbf{V}_1 (\mathbf{C}_1 + (\mathbb{K}_1 + \mathbb{B}_1 + \mathbf{C}_{cd1})\delta t^n) \mathbf{V}_1 \\ \mathbf{M}_{12} = -{}^t \mathbf{V}_1 \delta t^n \mathbf{C}_{cd1} \mathbb{P}_{2 \rightarrow 1} \mathbf{V}_2 \\ \mathbf{M}_{21} = -{}^t \mathbf{V}_2 \delta t^n \mathbf{C}_{cd2} \mathbb{P}_{1 \rightarrow 2} \mathbf{V}_1 \\ \mathbf{M}_{22} = {}^t \mathbf{V}_2 (\mathbf{C}_2 + (\mathbb{K}_2 + \mathbb{B}_2 + \mathbf{C}_{cd2})\delta t^n) \mathbf{V}_2 \\ \mathbf{N}_1 = {}^t \mathbf{V}_1 \mathbf{C}_1 \mathbf{V}_1 \mathbf{X}_1^n + {}^t \mathbf{V}_1 \mathbf{A}_1 \delta t^n \\ \mathbf{N}_2 = {}^t \mathbf{V}_2 \mathbf{C}_2 \mathbf{V}_2 \mathbf{X}_2^n + {}^t \mathbf{V}_2 \mathbf{A}_2 \delta t^n \end{cases} \quad (\text{G.13})$$

II.3. Les matrices du modèle réduit avec la fonctionnelle de saut de flux

La fonctionnelle de saut de flux est prise en compte de manière implicite et introduit 4 nouvelles matrices qui sont $\mathbf{C}_{flux11}, \mathbf{C}_{flux12}, \mathbf{C}_{flux21}$ et \mathbf{C}_{flux22} . Le système matriciel du modèle modal est :

$$\begin{pmatrix} \mathbf{M}_{11}^{flux} & \mathbf{M}_{12}^{flux} \\ \mathbf{M}_{21}^{flux} & \mathbf{M}_{22}^{flux} \end{pmatrix} \begin{pmatrix} \mathbf{X}_1^{n+1} \\ \mathbf{X}_2^{n+1} \end{pmatrix} = \begin{pmatrix} \mathbf{N}_1 \\ \mathbf{N}_2 \end{pmatrix} \quad (\text{G.14})$$

Les vecteurs \mathbf{N}_1 et \mathbf{N}_2 ne changent pas, par contre les matrices $\mathbf{M}_{11}, \mathbf{M}_{12}, \mathbf{M}_{21}, \mathbf{M}_{22}$ sont modifiées pour obtenir $\mathbf{M}_{11}^{flux}, \mathbf{M}_{12}^{flux}, \mathbf{M}_{21}^{flux}, \mathbf{M}_{22}^{flux}$:

$$\begin{cases} \mathbf{M}_{11}^{flux} = {}^t \mathbf{V}_1 (\mathbf{C}_1 + (\mathbb{K}_1 + \mathbb{B}_1 + \mathbf{C}_{cd1} + \mathbf{C}_{flux11})\delta t^n) \mathbf{V}_1 \\ \mathbf{M}_{12}^{flux} = {}^t \mathbf{V}_1 \delta t^n (\mathbf{C}_{flux12} - \mathbf{C}_{cd1}) \mathbb{P}_{2 \rightarrow 1} \mathbf{V}_2 \\ \mathbf{M}_{21}^{flux} = {}^t \mathbf{V}_2 \delta t^n (\mathbf{C}_{flux21} - \mathbf{C}_{cd1}) \mathbb{P}_{1 \rightarrow 2} \mathbf{V}_1 \\ \mathbf{M}_{22}^{flux} = {}^t \mathbf{V}_2 (\mathbf{C}_2 + (\mathbb{K}_2 + \mathbb{B}_2 + \mathbf{C}_{cd2} + \mathbf{C}_{flux22})\delta t^n) \mathbf{V}_2 \end{cases} \quad (\text{G.15})$$

Les matrices \mathbf{C}_{fluxij} sont obtenues par la méthode des éléments finis à partir de l'expression de la fonctionnelle de saut de flux 3.47. Le flux est estimé par interpolation linéaire, la valeur de la conductivité est déterminée à l'aide de $\tilde{\mathbf{T}}_1^n$ et de $\tilde{\mathbf{T}}_2^n$.

III. Projection entre maillages

Dans le cas où tous les nœuds des frontières entre les deux domaines sont communs, les projecteurs ne sont constitués que de 0 et de 1. Les deux projecteurs $\mathbb{P}_{2 \rightarrow 1}$ et $\mathbb{P}_{1 \rightarrow 2}$ sont alors transposés l'un de l'autre.

Dans le cas général, il faut transférer les champs de température au niveau des interfaces de contact. Les maillages sont générés sans imposer de contrainte de raccordement, ils ne comportent pas le même nombre de nœuds en surface, et possèdent des erreurs de discrétisation¹.

Voici un exemple de deux maillages qui ne se raccordent pas géométriquement et qui n'ont pas le même nombre de nœuds dans les zones où les frontières se raccordent :

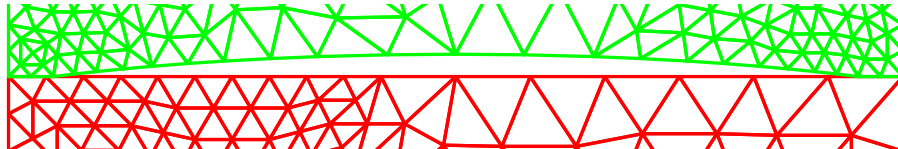


FIG. G.2 – Zoom sur l'interface entre deux maillages non-conformes

Ce problème n'est pas trivial, il ne possède d'ailleurs pas de solution exacte. La démarche retenue est fortement inspirée de la thèse de Jiao [25] : le problème est formulé de la manière suivante :

A chaque nœud $Node$ d'un maillage $Mesh_1$, on recherche le point P_{asso} (nœud ou point d'une arête) dans un maillage $Mesh_2$ qui se trouve à une distance minimale du nœud $Node$ considéré. Le point P_{asso} est alors qualifié de point associé à $Node$. La valeur à transférer pour le nœud $Node$ est alors la valeur du champ dans le maillage $Mesh_2$ au point P_{asso} .

Ce problème possède une solution unique.

III.1. Cas 2D

En 2D, la distance minimale d'un point à une arête est :

- La distance du point à son projeté orthogonal P sur l'arête si celui existe².
- La plus petite distance entre le point et un des sommets de l'arête.

L'algorithme permettant de déterminer le point associé P_{asso} au nœud $Node$ est :

- Détermination de la distance minimale entre le nœud $Node$ et toutes les arêtes du maillage $Mesh_2$.
- L'arête qui présente la distance calculée la plus faible est sélectionnée.
- Le point associé P_{asso} est alors le projeté orthogonal P s'il existe ou un des sommets de l'arête sélectionnée.

Il suffit donc de disposer d'une méthode permettant une détermination informatique du projeté orthogonal d'un point sur une droite pour pouvoir mettre en œuvre la méthode en 2D.

¹Ainsi, même si les deux maillages correspondent exactement à la même frontière géométrique, il est possible qu'ils ne soient pas en contact ou qu'ils se chevauchent légèrement.

²Le projeté orthogonal d'un point sur la droite (AB) existe toujours par contre il peut se trouver en dehors du segment $[AB]$

Le projecteur $\mathbb{P}_{2 \rightarrow 1}$ sera construit en recherchant le point associé de chacun des nœuds de la frontière du maillage $Mesh_1$. Chaque ligne de la matrice du projecteur comporte au plus 2 éléments non nuls :

- Dans le cas où le point associé est le sommet d'une arête, alors ce sommet est un nœud du maillage $Mesh_2$. La ligne de la matrice du projecteur est alors composée de zéro et d'un élément non nul dont la valeur est 1.
- Sinon le point associé est le projeté orthogonal sur une arête. Dans ce cas, deux éléments de la ligne de la matrice du projecteur sont non nuls, leur valeur est k (point A) ou $1 - k$ (point B) (voir G.16 et figure G.3).

III.1.1. Recherche du projeté orthogonal sur une arête

Soit P le projeté orthogonal sur le segment $[AB]$ du point M :

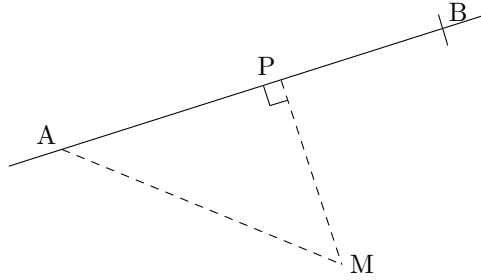


FIG. G.3 – Projection orthogonale de M sur $[AB]$

Le scalaire k est défini par :

$$\overrightarrow{AP} = k\overrightarrow{AB} \quad (\text{G.16})$$

La connaissance de k permet de déterminer P . Le produit scalaire est linéaire, ainsi :

$$\overrightarrow{AM} \bullet \overrightarrow{AB} = (\overrightarrow{AP} + \overrightarrow{PM}) \bullet \overrightarrow{AB} = \overrightarrow{AP} \bullet \overrightarrow{AB} + \overrightarrow{PM} \bullet \overrightarrow{AB} \quad (\text{G.17})$$

Or \overrightarrow{PM} et \overrightarrow{AB} sont orthogonaux, ainsi $\overrightarrow{PM} \bullet \overrightarrow{AB} = 0$, l'expression G.17 se simplifie en tenant compte de G.16 :

$$\overrightarrow{AM} \bullet \overrightarrow{AB} = \overrightarrow{AP} \bullet \overrightarrow{AB} = k\overrightarrow{AB} \bullet \overrightarrow{AB} \quad (\text{G.18})$$

Dès lors k peut se calculer à l'aide de la formule suivante :

$$k = \frac{\overrightarrow{AM} \bullet \overrightarrow{AB}}{\overrightarrow{AB} \bullet \overrightarrow{AB}} \quad (\text{G.19})$$

Le projeté orthogonal P appartient au segment $[AB]$ si et seulement si $k \in [0, 1]$.

III.2. Résultat de la projection

Il y a deux projections à effectuer :

- celle du maillage 1 vers le maillage 2 : figure G.4.
- celle du maillage 2 vers le maillage 1 : figure G.5.

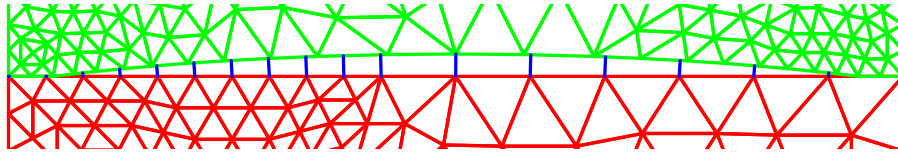


FIG. G.4 – Zoom sur la projection du maillage 1 vers 2

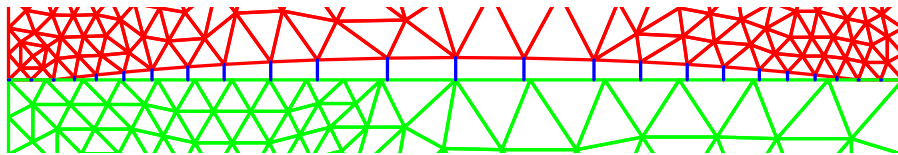


FIG. G.5 – Zoom sur la projection du maillage 2 vers 1

III.3. Cas 3D

Les idées sont les mêmes qu'en 2D :

- Recherche de l'élément frontière (triangle) le plus proche du point considéré.
- Détermination du point associé dans cet élément frontière.
- Calculs des coefficients des trois points.

Le point associé peut être un sommet, le projeté orthogonal dans le triangle s'il existe, ou l'un des projetés orthogonaux sur les arêtes s'ils existent. Graphiquement cela se traduit par :

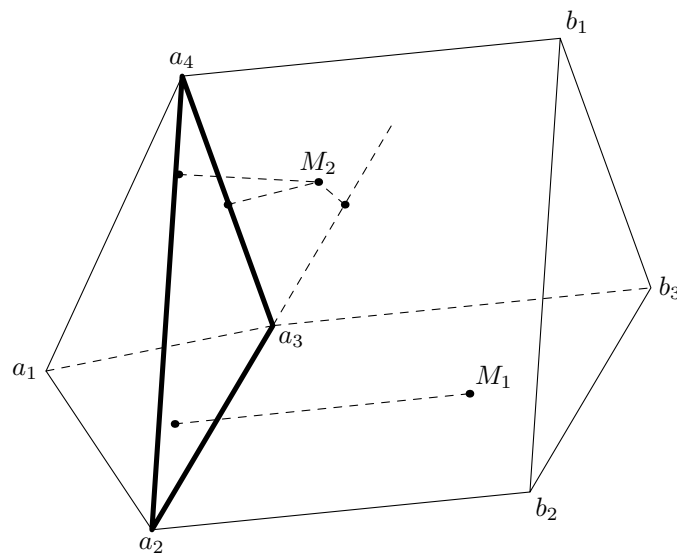


FIG. G.6 – Illustration de projection 3D

L'élément de surface considéré est le triangle $a_2a_3a_4$ du tétraèdre $a_1a_2a_3a_4$. Tous les points situés dans le volume du prisme droit $a_2a_3a_4b_1b_2b_3$ possèdent un projeté orthogonal compris dans l'élément de surface, c'est le cas du point M_1 . A l'inverse le projeté orthogonal du point M_2 sur l'élément de surface n'existe pas, de même pour son projeté sur l'arête a_2a_3 .

Par rapport au cas 2D, il ne manque qu'un seul outil permettant de déterminer le projeté orthogonal dans un triangle.

III.3.1. Recherche du projeté orthogonal sur un triangle

Soit P le projeté orthogonal du point M dans le plan défini par le triangle $a_2a_3a_4$. Soit \vec{n} la normale unitaire au triangle $a_2a_3a_4$, alors :

$$\vec{n} = \frac{\overrightarrow{a_2a_3} \wedge \overrightarrow{a_2a_4}}{|\overrightarrow{a_2a_3} \wedge \overrightarrow{a_2a_4}|} \quad (\text{G.20})$$

Soit la normale non normée \vec{n}_2 telle que $\vec{n}_2 = \overrightarrow{a_1a_2} \wedge \overrightarrow{a_1a_3}$:

$$\vec{n}_2 = \begin{pmatrix} (y_3 - y_2)(z_4 - z_2) - (y_4 - y_2)(z_3 - z_2) \\ (x_4 - x_2)(z_3 - z_2) - (x_3 - x_2)(z_4 - z_2) \\ (x_3 - x_2)(y_4 - y_2) - (y_3 - y_2)(x_4 - x_2) \end{pmatrix} \quad (\text{G.21})$$

La normale unitaire est déterminée par normalisation de \vec{n}_2 :

$$\vec{n} = \frac{\vec{n}_2}{|\vec{n}_2|} \quad (\text{G.22})$$

Les vecteurs \overrightarrow{PM} et \vec{n} sont colinéaires d'où

$$\overrightarrow{PM} = (\overrightarrow{PM} \bullet \vec{n}) \cdot \vec{n} = (\overrightarrow{Pa_2} \bullet \vec{n} + \overrightarrow{a_2M} \bullet \vec{n}) \cdot \vec{n} \quad (\text{G.23})$$

Or $\overrightarrow{Pa_2}$ et \vec{n} sont orthogonaux donc $\overrightarrow{Pa_2} \bullet \vec{n} = 0$, dès lors :

$$\overrightarrow{PM} = (\overrightarrow{a_2M} \bullet \vec{n}) \cdot \vec{n} \quad (\text{G.24})$$

Notons que a_2 peut être remplacé par n'importe quel point du plan ($a_2a_3a_4$) dans cette dernière expression. La détermination du projeté orthogonal sur le plan ($a_2a_3a_4$) est donc simple.

Il reste à déterminer si le projeté orthogonal P se situe dans le triangle $a_1a_2a_3$. Pour cela, il est préférable d'utiliser la méthode de l'élément de référence en 3D (méthode classique en éléments finis [45]). La transformation linéaire P qui permet de passer de l'élément de référence aux coordonnées physiques a pour expression :

$$P = \begin{cases} x = x_1 + (x_2 - x_1)\xi + (x_3 - x_1)\eta + (x_4 - x_1)\zeta \\ y = y_1 + (y_2 - y_1)\xi + (y_3 - y_1)\eta + (y_4 - y_1)\zeta \\ z = z_1 + (z_2 - z_1)\xi + (z_3 - z_1)\eta + (z_4 - z_1)\zeta \end{cases} \quad (\text{G.25})$$

Ainsi :

$$\begin{pmatrix} x - x_1 \\ y - y_1 \\ z - z_1 \end{pmatrix} = J_P \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} \quad (\text{G.26})$$

La matrice jacobienne est inversible puisque que son déterminant est égal à $6Volume(a_1a_2a_3a_4)$. Il est ainsi possible de calculer les valeurs de ξ, η, ζ à partir des coordonnées du projeté orthogonal P . Le projeté est dans le triangle si ξ, η et ζ sont compris entre 0 et 1.

Les cinq premiers modes du processeur

Un mode est formé d'une valeur propre et d'une fonction propre. La fonction propre est un champ spatial : il s'agit donc d'une donnée en tout point de l'espace. La représentation graphique des modes est relativement aisée dans le cas bidimensionnel. C'est pourquoi les cinq premières fonctions propres du processeur présenté au chapitre 4 en page 48 (figure 4.1) sont représentés graphiquement dans cette annexe. Les graphiques sont en élévation, l'élévation est localement égale à la fonction propre.

Le premier mode de Branche est toujours le mode dit plat (voir figure H.1). En effet la valeur propre est toujours nulle ($z_1 = 0$) et est associée à une fonction propre constante sur l'ensemble du domaine.

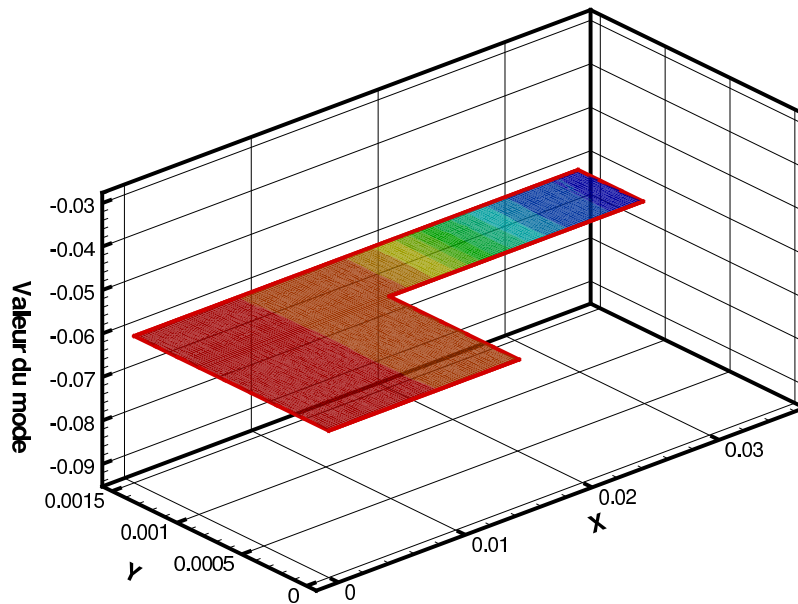


FIG. H.1 – Premier mode de branche du processeur

Les autres fonctions propres ne sont pas constantes sur le domaine, elles sont continues (pas de saut,...) et leur évolution spatiale est régulière. Au fur et à mesure de l'augmentation de l'ordre, les oscillations spatiales augmentent comme l'atteste les quatre figures des pages suivantes.

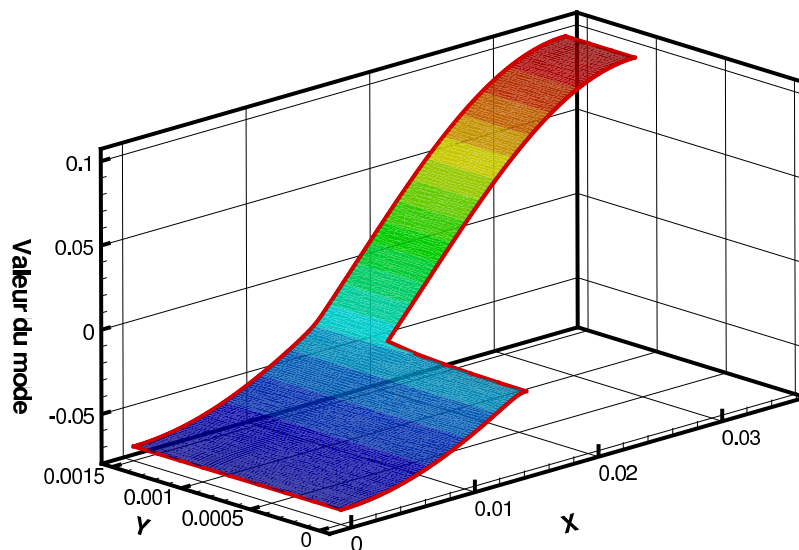


FIG. H.2 – Deuxième mode de branche du processeur

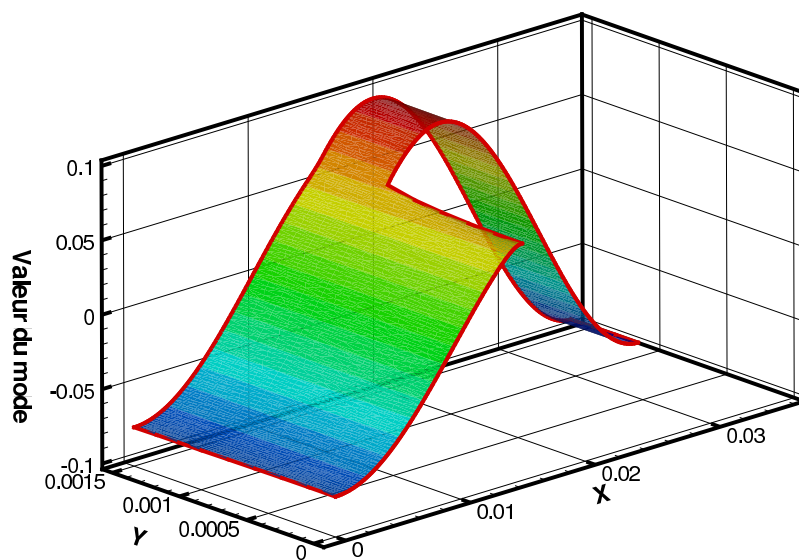


FIG. H.3 – Troisième mode de branche du processeur

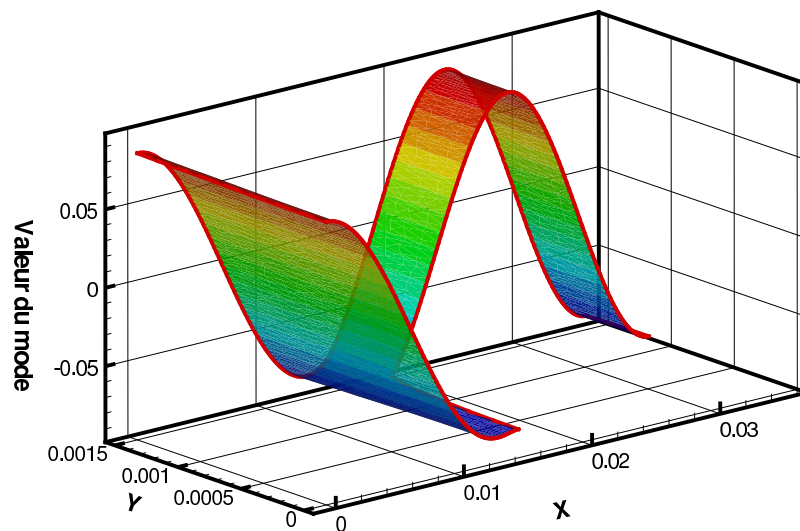


FIG. H.4 – Quatrième mode de branche du processeur

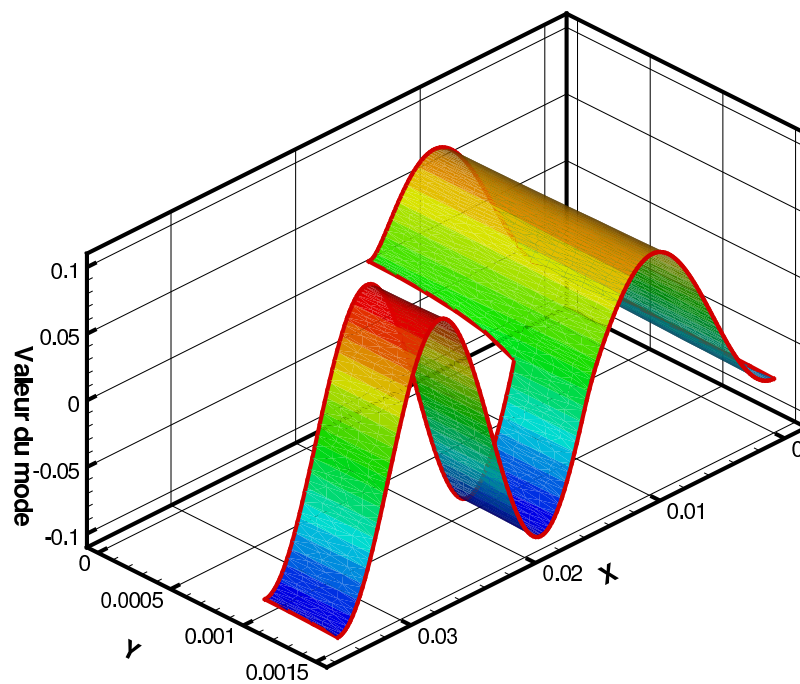


FIG. H.5 – Cinquième mode de branche du processeur

BIBLIOGRAPHIE

- [1] G. ALLAIRE : *Analyse numérique et optimisation*. Éditions de l'École Polytechnique, 2005.
- [2] G. BACHELAR : *La formation de l'esprit scientifique*. Vrin, 1967.
- [3] P. BACOT : *Analyse modale des systèmes thermiques*. Thèse de doctorat, Université Paris VI, Février 1984.
- [4] J.P BARDON : Heat transfer at solid interface : basic phenomenons-recent works. *In Proceedings of 4th Eurotherm*, Nancy, 1988.
- [5] J.L. BATTAGLIA : A modal approach to solve inverse heat conduction problems. *Inverse Problems in Engng*, 10(1):41–63, 2002.
- [6] J.V. BECK, B. BLACKWELL et C.R. SAINT CLAIR : *Inverse Heat Conduction : Ill-posed Problems*. WILEY, 1985.
- [7] P. BORNE, G. DAUPHIN-TANGUY, P. RICHARD, F. ROTELLA et I. ZAMBETTAKIS : *Modélisation et identification des processus Tome 2*. TECHNIP, 1992.
- [8] J.M. BRÉBEC : *Mécanique des fluides*. H-Prépa. HACHETTE, 2001.
- [9] F. BREZZI, B. COCKBURN, L.D. MARINI et E. SÜLI : Stabilization mechanisms in discontinuous Galerkin finite element methods. *Computer methods in applied mechanics and engineering*, 195:3293–3310, 2006.
- [10] H.S CARSLAW et J.C JAEGER : *Conduction of heat in solids*. OXFORD, 1959.
- [11] C. CARTER : A validation of the modal expansion method of modelling heat conduction in passive solar building. *Solar Energy*, 23:537–539, 1979.
- [12] F. CHATELIN : *Valeurs propres de matrices*. MASSON, 1988.
- [13] R. COMOLET : *Mécanique expérimentale des fluides*, volume II : Dynamique des fluides réels, turbomachines. MASSON, 4 édition, 2000.
- [14] L. CORDIER et M. BERGMANN : Réduction de dynamique par Décomposition Orthogonale aux Valeurs Propres (POD). Novembre 2006.
- [15] R. DAUTRAY et J.L. LIONS : *Analyse mathématique et calcul scientifique*, volume 5 : Spectre des opérateurs. MASSON, 1988.
- [16] A. EL BIYAALI, J.J. ROUX et C. INARD : Sous-structuration et couplage de modèles réduits appliqués aux transferts thermiques dans le sol. *Int. Comm. Haet Mass Transfer*, 23(4):575–586, 1996.
- [17] K. EL KHOURY : *Formulation modale de problèmes de diffusion thermique avec transport*. Thèse de doctorat, École des Mines de Paris, Octobre 1989.
- [18] B. FLAMENT : *Synthèse modale de systèmes thermiques*. Thèse de doctorat, École des mines de Paris, 1993.
- [19] Y.C. GERSTENMAIER et G. WACHUTKA : Time dependent temperature fields calculated using eigenfunctions and eigenvalues of the heat conduction equation. *Microelectronics Journal*, 32:801–808, 2001.

-
- [20] G.M.L. GLADWELL : Branch mode analysis of vibrating systems. *Journal of sound and vibrations*, 1:41–59, 1964.
- [21] J.P. GWINN et R.L. WEBB : Performance and testing of thermal interface materials. *Microelectronics journal*, 34:215–222, 2003.
- [22] P.C HANSEN : *Computational inverse problems in electrocardiology*, volume 5 de *Advances in computational bioengineering*, chapitre The L-curve and its use in the numerical treatment of inverse problems, pages 119–142. WIT Press, 2001.
- [23] F. HECH, I. DANAILA et O. PIRONNEAU : *Simulation numérique en C++*. DUNOD, 2003.
- [24] J.F IMBERT : *Analyse des structures par éléments finis*. CÉPADUÈS-ÉDITIONS, 1995. Présentation très complète des formulations variationnelles, matricielles, de la méthode des éléments finis, des problèmes de plaque et de coque, de l’analyse modale (dont la synthèse modale),...
- [25] X. JIAO : *Data transfer and interface propagation in multicomponent simulations*. Thèse de doctorat, University of Illinois, 2001.
- [26] F. JOLY, O. QUÉMÉNER et A. NEVEU : Modal reduction of an advection-diffusion model using a branch basis. *Numerical Heat Transfer Part B : Fundamentals*, 53(5):466–485, 2008.
- [27] F. KREITH et al : *The CRC Handbook of thermal engineering*. CRC Press, 2000.
- [28] P. LASCAUX et R. THÉODOR : *Analyse numérique matricielle appliquée à l’art de l’ingénieur 2 : méthodes itératives*. DUNOD, 1994.
- [29] P. LASCAUX et R. THÉODOR : *Analyse numérique matricielle appliquée à l’art de l’ingénieur 1 : méthodes directes*. DUNOD, 1998.
- [30] G. LEFÈBVRE : *La méthode modale en thermique*. DUNOD, 2007.
- [31] S.M. MARSHALL : An approximate method for reducing the order of linear system. *Control*, 10:642–643, 1966.
- [32] C. MÉNÉZO, J.J. ROUX et J. VIRGONE : Modelling heat transfers in building by coupling reduced-order models. *Building and Environment*, 37:133–144, 2002.
- [33] M. MINSKY : *La société de l’esprit*. Interéditions, 1988.
- [34] A. NEVEU : Approche modale pour les systèmes thermiques. 2005.
- [35] A. NEVEU et K. EL-KHOURY : Réduction d’un modèle thermique non linéaire par modes de branche. application a un câble chauffant. Janvier 2000.
- [36] A. NEVEU, K. EL-KHOURY et B. FLAMENT : Simulation de la conduction non linéaire en régime variable : décomposition sur les modes de branche. *International journal of thermal sciences*, 38:289–304, 1999.
- [37] A. NEVEU, O. QUÉMÉNER et É. VIDECOQ : Tentative d’écriture d’un formalisme modal pour les équations de diffusion-transport non linéaires et à paramètres instationnaires. Mars 2003.
- [38] A. OULEFKI : *Réduction de modèles thermiques par amalgame modal*. Thèse de doctorat, Ecole Nationale des Ponts et Chaussées, Février 1993.
- [39] D. PETIT et R. HACHETTE : Model reduction in linear heat conduction : use of interface fluxes for the numerical coupling. *International Journal of Heat and Mass Transfer*, 41:3177–3189, 1998.
- [40] D. PETIT et D. MAILLET : *Techniques inverses et estimation de paramètres. Partie 2*. Numéro AF 4516 de Techniques de l’ingénieur.
- [41] J.M PROIX : Diagonalisation de la matrice de masse thermique. Manuel de Référence de Code_Aster R3.06.07-A, EDF, 2000.
- [42] O. QUÉMÉNER, J.L. BATTAGLIA et A. NEVEU : Résolution d’un problème inverse par utilisation d’un modèle réduit modal. application au frottement d’un pion sur un disque en rotation. *International journal of thermal sciences*, 42:361–378, 2003.
- [43] O. QUÉMÉNER, A. NEVEU et É. VIDECOQ : A specific reduction method for the branch modal formulation : Application to a highly non-linear configuration. *International Journal of Thermal Sciences*, 46:860–907, 2007.
- [44] J RICHALLET, A RAULT et R POULIQUEN : *Théorie des systèmes, volume 4. Identification des processus par la méthode du modèle*. Gordon and Breach, 1971.
-

-
- [45] J-H SAIAC et B. MOHAMMADI : *Pratique de la simulation numérique*. DUNOD, 2003.
- [46] T.M. SHIH et J.T. SKLADANY : An eigenvalue method for solving transient heat conduction problems. *Numerical heat transfer*, 6:409–422, 1983.
- [47] D.C SORENSEN : *Implicit restarted Arnoldi/Lanczos methods for large scale eigenvalue calculations*.
- [48] É. VIDECOQ : *Problèmes inverses en diffusion thermique instationnaire : Résolution par représentation d'état et apport de la réduction de modèles*. Thèse de doctorat, Université de Poitiers, Novembre 1999.
- [49] É. VIDECOQ, O. QUÉMÉNER, M. LAZARD et A. NEVEU : Heat source identification and on-line temperature control by a Branch Eigenmodes Reduced Model. *International Journal of Heat and Mass Transfer*, 51:4743–4752, 2008.
- [50] B. WALLISER : *Systèmes et modèles. Introduction critique à l'analyse de systèmes*. Seuil, 1977.

Résumé

Cette étude concerne la simulation de systèmes thermiques comportant plusieurs domaines à l'aide d'une méthode d'ordre réduit adaptée au mono-corps.

Les modèles modaux réduits sont construits à partir de modes de branche. Ceux-ci sont déterminés à l'aide d'un modèle détaillé et permettent la prise en compte de non-linéarités. La réduction est effectuée par la technique de l'amalgame modal découplé qui permet de façon automatique et immédiate d'obtenir une base réduite qui ne comporte qu'un faible nombre de modes à partir de la base initiale.

Le couplage entre les différents domaines s'effectue par l'intermédiaire d'une résistance thermique de contact. On montre la nécessité de faire intervenir un terme supplémentaire de pénalisation de saut de flux qui vient améliorer les résultats.

Les simulations numériques effectuées sur des cas tests (microprocesseur et radiateur en 2D, bloc métallique avec cartouches chauffantes en 3D) montrent la pertinence de la méthode.

Mots clés : Calcul numérique, thermique, éléments finis, décomposition de domaines, méthodes modales, non-linéaire, résistance thermique de contact

Abstract

This study concerns the simulation of thermal systems with multiple fields with a reduced-order method suited to a single body.

Reduced models are constructed from modal branch eigenmodes. They are determined using a detailed model and allow the inclusion of non-linearities. The reduction is carried out by the simplified amalgam method which allows an automatic and immediate way to obtain a reduced basis which contains only a small number of modes from the original basis.

The coupling between the different areas is carried out through a thermal contact resistance. It shows the need to involve an additional flux jump penalty term to improve the results.

The numerical simulations carried out on test cases (microprocessor and radiator in 2D, metal block with hot cartridges in 3D) show the relevance of the method.

Keywords : scientific computing, heat science, finite elements, domain decomposition, modal methods, non-linear, thermal contact resistance

Titre en anglais : Substructuring thermal systems by branch eigenmodes