

Réduction d'un problème thermique par sous structuration modale. Application à la modélisation d'ensembles électroniques complexes.

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université d'Evry-Val-d'Essonne

École doctorale n° 579 Sciences mécaniques et
énergétiques, matériaux et géosciences (SMEMAG)
Spécialité de doctorat: Thermique

Thèse présentée et soutenue à Évry, le 11 décembre 2018, par

Sébastien Grosjean

Composition du jury:

Jean-Luc BATTAGLIA PR, ENSAM — I2M	Président
Adrien TOUTANT HdR, Université de Perpignan Via Domitia — PROMES	Rapporteur
Etienne VIDEOCOQ HdR, ENSMA — Pprime	Rapporteur
Pierre Gilles LEMARIÉ-RIEUSSET PR, Université d'Evry-Val-d'Essonne — LaMME	Examineur
Martin RAYNAUD HdR, THALES Alenia Space	Examineur
Shihe XIN PR, INSA Lyon — CETHIL	Examineur
Alain NEVEU PR, Université d'Evry-Val-d'Essonne — LMEE	Directeur de thèse
Frédéric JOLY McF, Université d'Evry-Val-d'Essonne — LMEE	Co-Encadrant

Remerciements

Cette thèse est le fruit d'une collaboration entre la société Thales Communications & Security et le Laboratoire de Mécanique et d'Énergétique d'Évry via une convention CIFRE, j'adresse donc en premier mes remerciements à cette entreprise pour avoir rendu ce travail possible et m'avoir donné les moyens de l'accomplir dans les meilleures conditions. En particulier, merci à Jean-Pierre Aurez et à Martin Raynaud pour avoir cru au projet et l'avoir démarré.

J'adresse mes sincères remerciements et ma gratitude envers Alain Neveu, mon directeur de thèse, et Frédéric Joly, mon co-encadrant, qui m'ont fait confiance depuis mon stage de master 2 en se rendant toujours disponibles et en m'apportant leur expertise, que ce soit sur l'aspect théorique ou applicatif de mon travail, quand j'en avais besoin. Ils ont ainsi fait preuve d'un grand investissement tout au long de ces trois ans, et je réalise la chance de les avoir eu comme encadrement, me donnant l'assurance d'être un membre à part entière de l'équipe de recherche du laboratoire, venu apporter sa pierre à l'édifice. Encore merci à vous deux.

Merci à Adrien Toutant et à Étienne Videcoq pour avoir rapporté ma thèse et merci à Jean-Luc Battaglia pour avoir présidé ma soutenance. Je remercie également les examinateurs Pierre Gilles Lemarié-Rieusset, Shihe Xin, ainsi que Martin Raynaud, qui a ainsi fait preuve de son engagement tout le long du travail de thèse.

Je tiens à remercier Gilles Lacombe qui, avec Pierre Gilles Lemarié-Rieusset, a été d'une aide précieuse pour établir l'aspect théorique de ce travail.

Je remercie l'ensemble du personnel de l'IUT GTE de Brétigny et du LMEE pour leur accueil durant ces trois dernières années, et en particulier Jean-Michel Cros, directeur de l'IUT, et Olivier Quéméner, directeur du LMEE.

Je remercie également mes collègues de Thales : Karine Vera et Mehdi Gherib pour avoir repris le tutorat de ma thèse après le départ de Martin, François Turpin, Pierre Lonjou et Maxime Ladevez pour leur accueil, Elliot Broussard qui a commencé son doctorat en même temps que moi, et tous les autres...

Un grand merci à ceux qui ont partagé mon bureau durant toutes ces années, mes collègues thésards Stanislas, Sylvain, Benjamin, Ana Gabriela et Adarsh. Votre présence et votre camaraderie ont permis de construire une ambiance très agréable dans notre bureau, qui m'a indubitablement aidé et grandement motivé dans mon travail.

Enfin, merci à ma famille et mes amis pour leur soutien indéfectible depuis le début.

Table des matières

I	Position du problème	1
1	Introduction	3
1.1	De la nécessité de connaître le comportement thermique des cartes électroniques	3
1.2	Modélisation des circuits imprimés	5
1.3	Modélisation des composants	9
1.3.1	Les boîtiers SO	9
1.3.2	Les boîtiers QFN	10
1.3.3	Les boîtiers BGA	11
1.4	Vers la réduction de modèle	11
2	État de l’art des méthodes de sous-structuration	15
2.1	Problème étudié	15
2.2	Les origines de la sous-structuration	16
2.2.1	Méthodes avec recouvrement : méthodes de Schwarz	16
2.2.2	Méthode du complément de Schur	17
2.2.3	Deux philosophies pour décomposer Ω	18
2.3	Sous-structuration d’un système couplé	19
2.3.1	Formulation primale	22
2.3.2	Formulation duale	23
2.4	La création de bibliothèques de composants : les maillages non-conformes	24
2.5	La méthode des joints	26
2.5.1	Partition de Ω	26
2.5.2	Continuité de la solution à l’interface	29
2.5.3	Définition de $M_h(\delta)$ dans le cas 2D	30
2.5.4	Définition de $M_h(\delta)$ dans le cas 3D	31
2.5.5	Autres méthodes de raccordement associées	31

2.6	Méthodes de Galerkin discontinues (DG)	32
2.6.1	Origines des méthodes DG : les méthodes de pénalisation intérieure (IP)	32
2.6.2	Développement des méthodes DG	33
2.7	Choix de la méthode de raccordement	34
3	État de l'art des méthodes de réduction	35
3.1	De la nécessité de réduire	35
3.2	Modèle nodal	36
3.3	Réduction modale	40
3.3.1	Principe	40
3.3.2	SVD, POD et PGD	41
3.3.3	Modes de Krylov	43
3.3.4	Méthode modale	44
3.3.5	Méthode MIM	44
3.4	Choix de la méthode de réduction	45
II	Formulation sous-structurée réduite du problème thermique	47
4	Adaptation de la méthode des joints à un problème thermique sous-structuré	49
4.1	Formulation sous-structurée pour deux domaines	49
4.1.1	Équation de la chaleur	49
4.1.2	Formulation variationnelle	51
4.1.3	Discrétisation du problème sous-structuré : principe	53
4.1.4	Exemple avec une base de fonctions spatiales discrétisées	55
4.2	Généralisation à N domaines	57
4.2.1	Décomposition de Ω dans le cadre des cartes électroniques	57
4.2.2	Formulation variationnelle	58
4.2.3	Forme matricielle	59
4.3	Premier bilan sur la méthode de sous-structuration avec joints	62
5	Adaptation de la méthode de réduction modale à la sous-structuration par joints	65
5.1	Définition de la base utilisée	65
5.1.1	Base de Dirichlet	65
5.1.2	Base de Steklov	67
5.1.3	Base de Dirichlet-Steklov	70

5.2	Propriétés de la base de Dirichlet-Steklov	70
5.2.1	Définition d'un nouveau produit scalaire sur $H^1(\Omega)$	70
5.2.2	Propriétés d'orthogonalité	71
5.3	Formulation matricielle des bases de Dirichlet et Steklov	74
5.3.1	Modes de Dirichlet.	74
5.3.2	Modes de Steklov.	75
5.3.3	Exemples de modes	76
5.4	Réduction de la base	78
5.4.1	Principe de la méthode de réduction par amalgame modal	79
5.4.2	Écart de réduction et dominance	80
5.4.3	Algorithme d'amalgame	81
5.4.4	Vers un nouveau critère d'arrêt	83
5.5	Écriture modale du problème sous-structuré	84
5.5.1	Projection des équations sur la base modale.	85
5.5.2	Forme matricielle	86
5.6	Bilan sur la formulation modale sous-structurée	90
III Applications et exemples numériques		93
6	Modélisation par sous-structuration modale conforme : exemple de la carte d'évaluation d'un convertisseur de tension à découpage ISL8026	95
6.1	Description de la carte de test	95
6.2	Fonctionnement du convertisseur DC-DC. Estimation de la puissance thermique dissipée	97
6.2.1	Principe de fonctionnement d'un convertisseur DC-DC.	97
6.2.2	Origine de la dissipation thermique dans le hacheur.	100
6.3	Modélisation de la carte de test	102
6.3.1	Modélisation des composants	103
6.3.2	Modélisation du PCB	104
6.3.3	Résistance de contact	107
6.4	Réduction du modèle	108
6.4.1	Formulation du problème réduit modal	108
6.4.2	Choix du paramètre de Steklov	109
6.4.3	Calcul des modes	114
6.4.4	Réduction des bases	115
6.5	Résultats	117

6.5.1	Scénario de test	117
6.5.2	Résultats : comparaison avec une simulation détaillée	119
6.6	Bilan de cette première étude	122
7	Exemple d'application de la sous-structuration modale non-conforme avec joints	125
7.1	Validation du modèle éléments finis sous-structuré avec joints	126
7.1.1	Choix des fonctions mortiers	127
7.1.2	Avec un maillage conforme à l'interface	127
7.1.3	Avec un maillage non-conforme à l'interface	128
7.2	Réduction modale	129
7.2.1	Réduction modale des fonctions mortiers	129
7.2.2	Réduction modale en flux et en température	132
7.3	Modélisation d'un ensemble complexe	137
7.4	Validation du modèle sous-structuré détaillé	138
7.5	Réduction de modèle	141
7.5.1	Création des modèles réduits	141
7.5.2	Simulations réduites	143
7.5.3	Simulation d'une nouvelle configuration	149
7.6	Conclusions et perspectives de cette deuxième étude	152
8	Conclusion générale et perspectives	155
	Annexes	159
	Annexe A Décomposition d'une fonction de $H^1(\Omega)$ sur la base de Dirichlet-Steklov	161
A.1	Continuité et coercivité de l'opérateur de la chaleur.	162
A.2	Propriété de la base de Dirichlet.	169
A.2.1	Base de Dirichlet.	169
A.2.2	Orthogonalité des modes de Dirichlet.	169
A.3	Propriété de la base de Steklov.	170
A.3.1	Définition de la Base de Steklov [1] [2] [3] [4]	170
A.3.2	Orthogonalité	172
A.4	Propriétés de la base de Dirichlet-Steklov.	173
A.4.1	Orthogonalité entre les modes de Dirichlet et de Steklov.	173
A.4.2	Décomposition d'une fonction de $H^1(\Omega)$	175

A.4.3	Équivalence entre la norme $\ u\ _{H(\Omega)}$ et la norme usuelle dans $H^1(\Omega)$. . .	177
A.4.4	Complétude de la base de Dirichlet-Steklov	179
Annexe B	Résultats mathématiques généraux	191
B.1	Produit scalaire et norme.	191
B.2	Espace de Hilbert.	192
B.2.1	Suites de Cauchy.	192
B.3	Fonction à support borné.	193
B.4	Inégalité de Poincaré.	194
B.5	Notion de trace sur le bord d'un domaine.	197
B.6	Formules de Courant-Fischer.	202
B.7	Densité et Compacité.	207
B.8	Convergence.	208
B.9	Propriétés du tenseur de conductivité thermique.	208
Annexe C	Proposition d'un algorithme d'amalgame pour la base de Dirichlet-Steklov	213
Annexe D	Lien entre la norme $\ \cdot\ _{H(\Omega)}$ et la norme dans $L^2(\Omega)$	219
D.1	Majoration de la norme dans $L^2(\Omega)$	219
D.2	Encadrement de la constante $C_*(\Omega)$	222
D.2.1	Cas d'une boule de \mathbb{R}^d	222
D.2.2	Cas d'un pavé droit de \mathbb{R}^d	224
D.2.3	Cas d'un sphère creuse de faible épaisseur de \mathbb{R}^d	226
D.2.4	Cas d'un cylindre droit de \mathbb{R}^d	230
D.2.5	Cas d'un tube cylindrique droit de \mathbb{R}^d	232
D.3	Estimation d'erreur dans l'espace d'énergie.	233
D.4	Estimation d'énergie.	235
Annexe E	Fonctionnement de la carte ISL EVAL8026A	239
E.1	Fonctionnement du convertisseur DC-DC. Estimation de la puissance thermique dissipée.	239
E.1.1	Principe de fonctionnement d'un convertisseur DC-DC.	239
E.1.2	Origine de la dissipation thermique dans le hacheur.	246
Annexe F	Réduction modale sous-structurée avec des maillages non-conformes en 2D	249

Table des figures

1.1.1	Tous ces objets ont été remplacés par des smartphones (crédits : Thomas Frey) .	3
1.1.2	Évolution des performances des téléphones portables aux smartphones (crédit : Sandeep Bhat)	4
1.1.3	Schéma d'une carte électronique avec un composant sur un circuit imprimé . . .	6
1.1.4	Exemple de carte industrielle, ainsi que le routage à travers les différentes couches du circuit imprimé (12 couches sont superposées, chaque couleur correspondant à une couche différente).	6
1.1.5	Schéma d'un PCB multi-couches (6 couches de cuivre)	7
1.1.6	Carte ISL8026EVAL3Z (vue du dessus et des différentes couches métalliques (en couleur) du PCB).	8
1.1.7	Détail du PCB de la carte ISL8026EVAL3Z, avec maillage sur la partie zoomée .	8
1.1.8	Exemple de boîtier SO (Texas Instruments)	10
1.1.9	Exemples de boîtiers QFN	10
1.1.10	Exemple de package BGA : processeur graphique NVIDIA GeForce GTX 980M .	11
1.1.11	Modèle de BGA208 avec 60 fils d'or de connexion	12
1.2.1	Exemple de domaine Ω pour l'équation de Poisson.	15
1.2.2	Décomposition de domaine avec recouvrement (partie hachurée).	17
1.2.3	Décomposition de domaine sans recouvrement ; l'interface Γ est représentée en rouge sur le schéma de droite.	17
1.2.4	Différents types de décomposition de domaine.	19
1.2.5	Partition régulière du domaine Ω en deux sous domaines $\Omega^{(1)}$ et $\Omega^{(2)}$	20
1.2.6	Exemple de discrétisation de Ω	21
1.2.7	Exemple de discrétisation de Ω	23
1.2.8	Maillages conforme et non-conforme aux interfaces.	24
1.2.9	Comment exprimer le couplage lorsque l'on veut raccorder deux sous-structures dans un découpage <i>a priori</i> , maillées indépendamment ?	25

1.2.10	Maillage géométriquement non-conforme.	26
1.2.11	Différents types de décompositions.	27
1.2.12	Exemple de partition du squelette en mortiers (décomposition conforme).	28
1.2.13	Exemple de partition du squelette en mortiers (décomposition conforme à gauche et non conforme à droite).	28
1.2.14	Fonctions de base des multiplicateurs de Lagrange (2D).	30
1.3.1	Exemple de réseau nodal pour un composant dans un boîtier DFN8.	36
1.3.2	Exemples de réseaux nodaux pour un même composant : (a) modèle stationnaire (sans inertie) issu d'une analyse physique par volumes finis ; (b) modèle dynamique (l'inertie est représentée par une capacité sur le nœud interne T_J) issu d'une analyse physique par volumes finis ; (c) modèle dynamique qui prend en compte la conduction entre les faces : la détermination des résistances thermique entre les faces est délicate car elle fait intervenir des phénomènes tridimensionnels ; (d) modèle dynamique dont les résistances thermiques et les capacités doivent être identifiées (par exemple à l'aide d'un algorithme génétique) : on s'éloigne d'une interprétation physique directe des résistances et des capacités, en particulier la résistance qui relie les nœuds "top" (T_T) et "bottom" (T_B) directement sans passer par le nœud central ne respecte pas les lois du transfert de chaleur par conduction thermique (diffusion). De même une surface n'a théoriquement pas de capacité thermique.	38
1.3.3	Surfaces "inner" et "outer" des faces inférieure et supérieure.	40
2.4.1	Partition régulière du domaine Ω en deux sous domaines $\Omega^{(1)}$ et $\Omega^{(2)}$	50
2.4.2	Compatibilité des maillages. Exemple du raccordement de deux maillages incompatibles.	56
2.4.3	Deux stratégies de choix des frontières maîtres d'une carte électronique.	57
2.4.4	Schéma de carte "réaliste".	58
2.4.5	Maillage du contact entre le composant $\Omega^{(k)}$ et le circuit imprimé $\Omega^{(0)}$	64
2.5.1	Exemple de champ de température reconstruit à l'aide de modes de Dirichlet.	67
2.5.2	Exemple de champ de température reconstruit à l'aide de modes de Steklov.	69
2.5.3	Exemple de champ de température reconstruit à l'aide de modes de Dirichlet-Steklov.	73

2.5.4	Exemples de modes de Dirichlet calculés sur un pavé droit : les modes sont normalisés entre -1 et $+1$ pour la visualisation, et on a représenté l'isovaleur -0.2 en bleu et $+0.2$ en rouge. On note que pour les modes de Dirichlet, les isovaleurs sont fermées.	76
2.5.5	Exemples des premiers modes de Steklov calculés sur un pavé droit : les modes sont aussi normalisés entre -1 et $+1$ pour la visualisation, et on représente dix isovaleurs.	77
2.5.6	Exemples de modes de Steklov d'ordre plus élevé calculés sur un pavé droit.	77
2.5.7	Localisation des modes de Steklov pour le pavé droit.	77
2.5.8	Exemple d'états d'excitation d'une base constituée de trois modes.	79
3.6.1	Vues de la carte Intersil ISL8026EVAL3Z.	96
3.6.2	Schéma de principe d'un hacheur de courant élémentaire. Digramme temporel de tension et d'intensité.	98
3.6.3	Schéma de principe d'un hacheur de courant abaisseur de tension.	99
3.6.4	Phases de fonctionnement d'un hacheur de courant.	99
3.6.5	Diagramme temporel de l'évolution de la tension aux bornes des interrupteurs K_1 et K_2 et du courant les traversant.	100
3.6.6	Diagramme temporel montrant le défaut de commutation. A gauche schéma de principe (défaut amplifiés), à droite enregistrement des signaux sur un oscilloscope numérique.	101
3.6.7	Variation de la résistance d'une inductance. D'après document TDK.	102
3.6.8	Boitiers QFN16 et de l'inductance.	103
3.6.9	Composants Montés en Surface de la carte.	103
3.6.10	Maillages des composants.	105
3.6.11	Couches de la carte de test ISL8026EVAL3Z.	105
3.6.12	Couches métalliques de la carte de test.	106
3.6.13	Modélisation détaillée du PCB.	107
3.6.14	Projection de la "Metal Fraction" sur le maillage du PCB.	108
3.6.15	Géométrie traitée (à gauche) et résultat de la simulation (à droite).	110
3.6.16	Écart entre la simulation éléments finis et deux températures modales obtenues par projection et par simulation, en fonction de la résistance de contact.	111
3.6.17	Écart entre la simulation éléments finis et deux températures modales en fonction de la valeur du paramètre ζ sur la frontière de contact.	113
3.6.18	QFN16 : (a) & (b) exemples de modes de Dirichlet (vue du dessous), (c) & (d) modes de Steklov (vue du dessus).	114

3.6.19	Inductance : (a) & (b) exemples de modes de Dirichlet (vue du dessus), (c) & (d) modes de Steklov (vue du dessous)..	114
3.6.20	PCB : Exemples de modes de Steklov (vue du dessus)..	115
3.6.21	Simulation de référence pour l'amalgame.	116
3.6.22	Exemples de modes amalgamés pour le QFN16 et le PCB.. . . .	117
3.6.23	Scénario de test.	118
3.6.24	Évolution de $\max_{\Omega} T^{mod}(t)$ et $\varepsilon_{\max}(t)$ pour le QFN16 au cours du temps.	120
3.6.25	T^{mod} (gauche) et $ T^{mod} - T^{EF} $ (droite) à $t = 51s$, quand ε_{∞} est maximal.	120
3.6.26	T^{mod} (gauche) et $ T^{mod} - T^{EF} $ (droite) à $t = 51s$, quand ε_{∞} est maximal.	121
3.6.27	T^{mod} (gauche) et $ T^{mod} - T^{EF} $ (droite) à $t = 52s$	121
3.7.1	Description de la géométrie de test et champ de température obtenue à l'issue d'une simulation dynamique de 1000 s.	126
3.7.2	Maillage conforme de la géométrie de test et champ écart obtenu à l'issue d'une simulation dynamique de 1000 s.	128
3.7.3	Évolution de la température maximale dans la sous-structure 2. En rouge (échelle de droite) l'écart sur cette quantité entre le modèle éléments finis classique sur un maillage conforme et la méthode des joints sur un maillage non-conforme aux interfaces..	129
3.7.4	Simulation éléments finis avec la méthode des joints.	129
3.7.5	Exemples de modes de Neumann pour l'interface de contact..	131
3.7.6	Écart quadratique sur la température entre d'une part le modèle éléments finis sous-structuré avec les fonctions de formes éléments finis pour les fonctions mortiers et d'autre part le modèle éléments finis avec des modes de Neumann pour les fonctions mortiers. Le nombre de modes de Neumann conservé est en abscisse.	132
3.7.7	Écart moyen entre le champ issu de la simulation réduite avec 10 modes par sous-structure et champ de référence en fonction du nombre de mortiers utilisés à l'interface.	133
3.7.8	Écart moyen entre champ issu de la simulation réduite avec 10, 20, 50 et 100 modes par sous-structures et champ de référence en fonction du nombre de mortiers utilisés à l'interface.	135
3.7.9	Écart moyen entre le champ issu de la simulation réduite avec 10, 20, 50 et 100 modes par sous-structure et champ de référence, en fonction du nombre de mortiers utilisés à l'interface, les bases étant calculées avec $\zeta = 10\,000$ sur la frontière de contact.	136

3.7.10	Modèle de carte élémentaire avec les différents composants numérotés.	138
3.7.11	Maillage non conforme pour le modèle de test (avec zoom sur le composant 2). .	140
3.7.12	Champ de température à $t = 1000s$ de la simulation avec méthode des joints utilisant un maillage non conforme.	140
3.7.13	Évolution de la température maximale au cours du temps avec : en bleu, la température maximale au cours de la simulation monobloc, et en vert (ligne pointillée) la température maximale au cours de la simulation sous-structurée avec méthode des joints pour le maillage non conforme.	141
3.7.14	Assemblage de modes de Steklov des différentes sous-structures du modèle. . .	142
3.7.15	Exemples de modes de Neumann pour les interfaces de contact.	143
3.7.16	Champ de température à $t = 1000s$ de la simulation avec méthode des joints réduite.	144
3.7.17	Évolution de la température maximale atteinte par les composant 3 et 7 au cours de la simulation réduite (courbe noire), et écart (absolu) par rapport à la température maximale au cours de la simulation sous-structurée complète (courbe rouge).	145
3.7.18	Champ de température à $t = 1000s$ de la simulation avec méthode des joints réduite, avec une puissance dissipée par les composants doublée.	146
3.7.19	Évolution de la température maximale atteinte par les composant 3 et 7 au cours de la simulation réduite (courbe noire), et écart (absolu) par rapport à la température maximale au cours de la simulation sous-structurée complète (courbe rouge) lorsque ces deux composants dégagent une puissance doublée. . .	147
3.7.20	Champ de température à $t = 1000s$ de la simulation avec méthode des joints réduite, avec $h_{ext} = 50 W.m^{-2}.K^{-1}$	148
3.7.21	Évolution de la température maximale atteinte par les composant 3 et 7 au cours de la simulation réduite (courbe noire), et écart (absolu) par rapport à la température maximale au cours de la simulation sous-structurée complète (courbe rouge) avec $h_{ext} = 50 W.m^{-2}.K^{-1}$	149
3.7.22	Le composant 1 change de surface d'implémentation.	150
3.7.23	Champ de température à $t = 1000s$ de la simulation éléments finis avec méthode des joints de la nouvelle configuration.	150
3.7.24	Champ de température à $t = 1000s$ de la simulation réduite avec méthode des joints de la nouvelle configuration.	151
3.7.25	Évolution de la température maximale au cours du temps pendant les simulation complète (lignes pleines) et réduite (lignes pointillées) pour, de haut en bas, les composants 3, 7 et 1 dans la nouvelle configuration.	152

3.D.1	Tube mince d'épaisseur e	232
3.E.1	Schéma de principe d'un hacheur de courant élémentaire. Digramme temporel de tension et d'intensité.	240
3.E.2	Bloc diagramme du hacheur ISL8026A d'après [5]. Sur la figure de droite on a extrait du bloc diagramme la cellule de commutation.	241
3.E.3	Schémas de principe de la cellule de commutation d'un hacheur. Les deux premiers schémas proposent des réalisations pratiques. Sur la représentation de gauche l'interrupteur commandé à l'amorçage et au blocage est un IGBT (noté T). Ses commutations commandées entraîneront les commutations spontanées de la diode D. La figure centrale montre une réalisation utilisant deux transistors MOSFET, l'un de type P l'autre de type N, commandés de manière synchronisée. La figure de droite est un schéma général que nous utiliserons pour décrire le fonctionnement. Il comporte deux interrupteurs K_1 et K_2 à commande complémentaire (l'horloge qui cadence le découpage et la commande complémentaire ne sont pas représentées sur ce schéma générique).	242
3.E.4	Schéma de principe d'un hacheur de courant abaisseur de tension.	242
3.E.5	Phases de fonctionnement d'un hacheur de courant.	243
3.E.6	Digramme temporel de l'évolution de la tension aux bornes des interrupteurs K_1 et K_2 et du courant les traversant.	243
3.E.7	Diagramme temporel des signaux de sortie du hacheur et de l'inductance. Ici les ondulations on été très amplifiées par rapport à la réalité. Ces signaux peuvent être enregistrés sur la carte d'évaluation.	245
3.E.8	Diagramme temporel montrant le défaut de commutation. A gauche schéma de principe (défaut amplifiés), à droite enregistrement des signaux sur un oscilloscope numérique.	247
3.E.9	Variation de la résistance d'une inductance. D'après document TDK.	248

Liste des tableaux

3.6.1	Propriétés physiques des matériaux utilisés pour les composants	104
3.6.2	Propriétés physiques des matériaux utilisés pour le PCB	107
3.6.3	Différentes valeurs de R_c en fonction du contact [6]	108
3.6.4	Différences sur le QFN16 et le PCB, et temps de calcul en fonction de l'ordre de réduction	122
3.7.1	Écarts entre champ projeté et champ de référence en fonction du nombre de modes des bases réduites pour les sous-structures 1 et 2	133
3.7.2	Écarts entre champ projeté et champ de référence en fonction du nombre de modes des bases réduites pour les sous-structures 1 et 2 avec $\zeta = 10\,000$	137
3.7.3	Propriétés physiques des matériaux utilisés pour la carte de test (la couche 2 du PCB est la couche en contact avec les composants)	139
3.D.1	Valeur numérique pour la borne inférieure de la constante C_* en fonction de la dimension de l'espace \mathbb{R}^d . Cas d'un matériau homogène et isotrope et facteur de Steklov uniforme ($\alpha_o = 1$).	234

Première partie

Position du problème

Chapitre 1

Introduction

1.1 De la nécessité de connaître le comportement thermique des cartes électroniques



FIGURE 1.1.1 – Tous ces objets ont été remplacés par des smartphones (crédits : Thomas Frey)

Dans nos sociétés modernes, l'électronique est omniprésente. Prenons l'exemple du secteur de l'électronique grand public et des télécommunications. On pense aux communications dans le domaine militaire, où elles jouent un rôle crucial sur et au-delà des théâtres d'opération,

ou au domaine civil, la société étant plus connectée que jamais grâce à l'incroyable évolution des ordinateurs personnels et autres smartphones présents dans (quasiment) toutes les poches. La figure 1.1.1 illustre bien cette évolution, puisque radio, caméscope, appareil photo, réveil, calculatrice et même GPS sont maintenant disponibles dans un concentré d'électronique d'environ $150\text{mm} \times 100\text{mm}$.



FIGURE 1.1.2 – Évolution des performances des téléphones portables aux smartphones (crédit : Sandeep Bhat)

Cette évolution est rendue possible par la multiplication des fonctionnalités des systèmes électroniques, via une complexification des composants (venus remplacer les composants passifs) et par l'augmentation des performances des processeurs (voir figure 1.1.2). Cette croissance du nombre de fonctionnalités et de puissance est intimement liée à la miniaturisation et à l'évolution de la finesse de gravure (7nm pour les processeurs de dernière génération contre $10\ \mu\text{m}$ au début des années 1970).

Cette diminution des dimensions entraîne ainsi une augmentation des densités de puissance thermiques au sein même de ces composants, notamment dans les composants semi-conducteurs de puissance, qui subissent de nombreux cycles d'activation/désactivation. Or, les effets thermiques sur les composants ne sont pas anodins, que ce soit

- la détérioration des brasures entre composant et PCB¹ ;
- la dégradation des fils de bonding² ;

1. PCB : "Printed Circuit Board", ou circuit imprimé en français ; dans la suite, on n'utilisera que l'acronyme PCB, puisque le plus utilisé dans le domaine de l'électronique, francophone ou anglophone.

2. fil soudé entre un plot de connexion et la puce électronique, ou entre deux plots de connexion.

- l'apparition de courants de fuite ;
- la dégradation de l'intégrité physique des packages encapsulant les composants ;
- les performances électriques (limitées par une température maximale au-delà de laquelle le fonctionnement n'est plus garanti) ;
- la fatigue des matériaux assujettis à des cycles thermiques et qui sont soumis à des contraintes thermo-mécaniques, pouvant conduire à une rupture ;
- la fiabilité des composants en général...

Du fait d'une erreur de prévision des températures de fonctionnement, ces défaillances dues aux effets thermiques peuvent drastiquement réduire la durée de vie de ces composants, ce qui peut avoir des conséquences désastreuses, car de ces systèmes défaillants peuvent dépendre des vies (on peut penser aux passagers d'une voiture autonome devenue incontrôlable, à des systèmes médicaux sophistiqués ou encore au cas militaire, avec des soldats se trouvant dans l'incapacité d'être évacués d'une zone à risque, faute de moyens de communication à la suite d'une panne).

Il est donc crucial de pouvoir prédire finement le comportement thermique des cartes électroniques, et ce dès la phase de conception des équipements via des simulations *ad hoc*, pour réduire le plus possible le risque de dysfonctionnements et de défauts tels que ceux cités ci-dessus.

Pour étudier la dissipation thermique, on peut se focaliser sur plusieurs niveaux différents : au niveau puce, au niveau composant, au niveau carte et au niveau système. L'objectif de cette thèse est d'étudier le comportement thermique d'une carte complète, ce qui suppose donc de pouvoir la simuler au niveau PCB ainsi qu'au niveau de chaque composant, tout en tenant compte de leurs différentes interactions, chaque composant étant voisin d'autres composants, qui dissipent eux-même plus ou moins de chaleur via le PCB. Toutefois, la modélisation d'une carte électronique complète via une simulation détaillée reste un défi numérique, même si "le nombre de transistors des microprocesseurs sur une puce de silicium double tous les deux ans"³, ce qui permet une croissance exponentielle des capacités de calculs. Mais la complexité du système à simuler croît elle aussi exponentiellement : une simulation détaillée de la carte n'est donc sans doute pas la solution.

1.2 Modélisation des circuits imprimés

Sur la figure 1.1.3, on a représenté schématiquement un modèle de carte électronique, constituée d'un composant soudé sur un circuit imprimé. Si certains composants peuvent être

3. Deuxième Loi de Moore, énoncée en 1975.

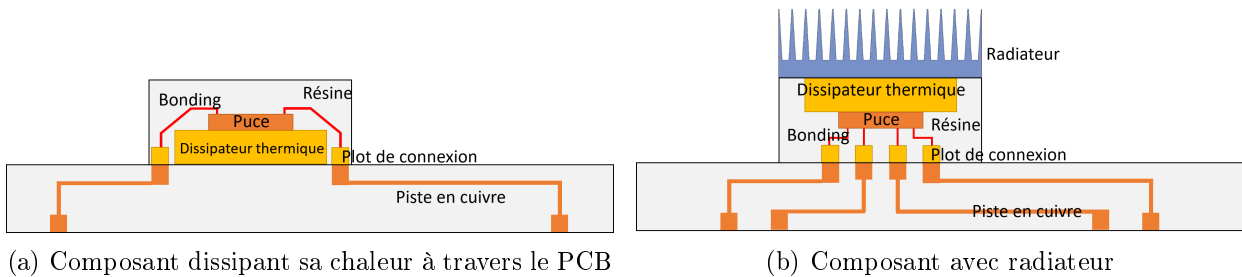


FIGURE 1.1.3 – Schéma d’une carte électronique avec un composant sur un circuit imprimé

refroidis via un radiateur (figure 1.1.3b) et l’action d’un ventilateur (comme les processeurs par exemple), beaucoup dissipent leur chaleur via le support de la puce, qui sert alors de dissipateur thermique (figure 1.1.3a). La chaleur peut aussi se dissiper à travers les bondings du composant et le long des pistes en cuivre du PCB.

Ainsi, si on s’intéresse d’abord à la modélisation au niveau carte, le premier défi se pose dans la modélisation du PCB, puisqu’il a un rôle essentiel de drainage et de répartition du flux thermique. Ce drainage s’effectue par les couches de cuivres du circuit imprimé, et par des “trous métallisés” placés sous les composants et qui traversent le PCB, de petite taille et parfois en très grand nombre. L’augmentation des fonctionnalités ainsi que des entrées et

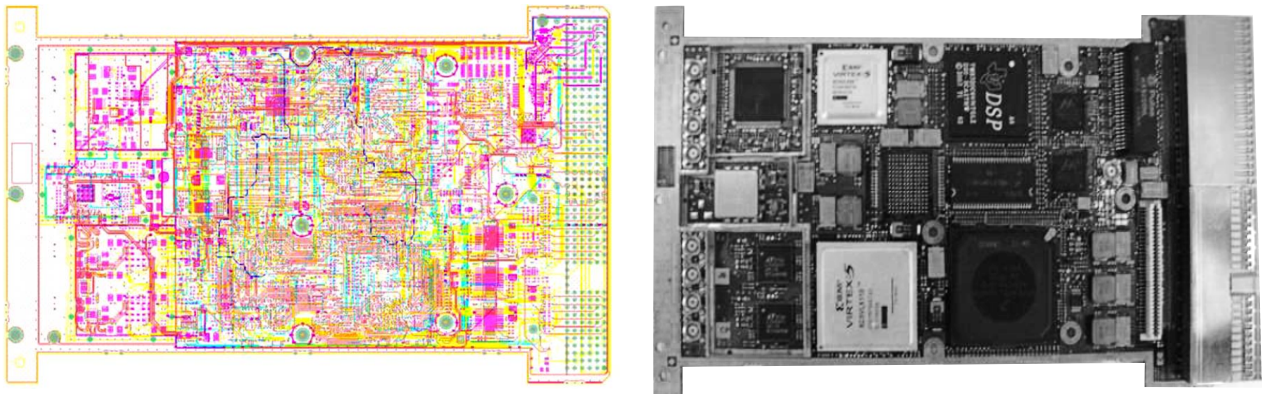


FIGURE 1.1.4 – Exemple de carte industrielle, ainsi que le routage à travers les différentes couches du circuit imprimé (12 couches sont superposées, chaque couleur correspondant à une couche différente)

sorties des composants se répercute sur la complexité des PCB eux-mêmes. La figure 1.1.4 montre un exemple de carte industrielle, de dimensions $160\text{ mm} \times 100\text{ mm} \times 1,6\text{ mm}$. 1675 composants sont implantés sur cette carte qui présente 20000 points de connexions. On a donc déjà affaire à une grande différence d’échelle, puisqu’il faut modéliser une carte de 16 cm de

longueur, alors que les points de connexion des composants ont un diamètre de l'ordre de la dizaine de μm .

De plus, pour effectuer le routage entre les différents composants de la carte, celle-ci est le plus souvent constituée d'un empilement de couches de différents matériaux. La figure 1.1.5

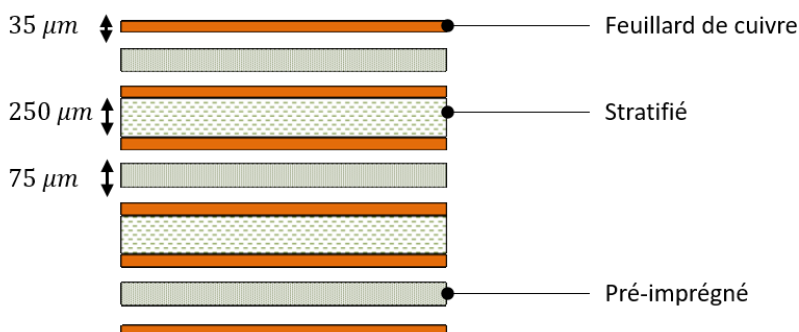


FIGURE 1.1.5 – Schéma d'un PCB multi-couches (6 couches de cuivre)

présente un exemple de PCB constitué de 6 couches d'interconnexion (feuillards de cuivre où est fait le routage). Ces feuillards sont séparés par des couches diélectriques : les stratifiés et les pré-imprégnés composés de résine époxy ou de polyimides.

Une difficulté supplémentaire de la modélisation du PCB réside dans la prise en compte (ou non) de ces différentes couches, qui présentent des propriétés physiques très différentes (la conductivité du cuivre étant environ 500 à 1000 fois supérieure à celle des résines).

Pour revenir à l'exemple de la figure 1.1.4, celle-ci est constituée de 12 couches d'interconnexion de 17,5 ou 35 μm , isolées par 11 couches diélectriques, d'une épaisseur allant de 75 à 250 μm .

On a donc une très grande complexité géométrique, et la modéliser avec des méthodes de type éléments finis est inenvisageable pour faire des simulations dans un temps raisonnable (*i.e.* des simulations qui prennent moins d'un jour de temps de calcul avec des moyens de calculs raisonnables, tels que des stations de travail pour ingénieur, et non pas avec des serveurs de calculs massivement parallélisés).

Pour l'illustrer, prenons un autre exemple : la carte Intersil ISL8026EVAL3Z (figure 1.1.6). Cette carte permet d'évaluer les performances d'un convertisseur DC/DC pouvant délivrer un courant continu d'une intensité de 3A depuis un courant d'entrée pouvant varier de 2,7V à 5,5V [5]. Cette carte est beaucoup plus simple que celle présentée en 1.1.4, puisqu'elle présente seulement 19 composants implantés et est constituée de 4 couches de cuivre, isolées par 3 couches de diélectrique.

Pourtant, le maillage nécessaire à la modélisation fine des pistes de cuivre et des différentes

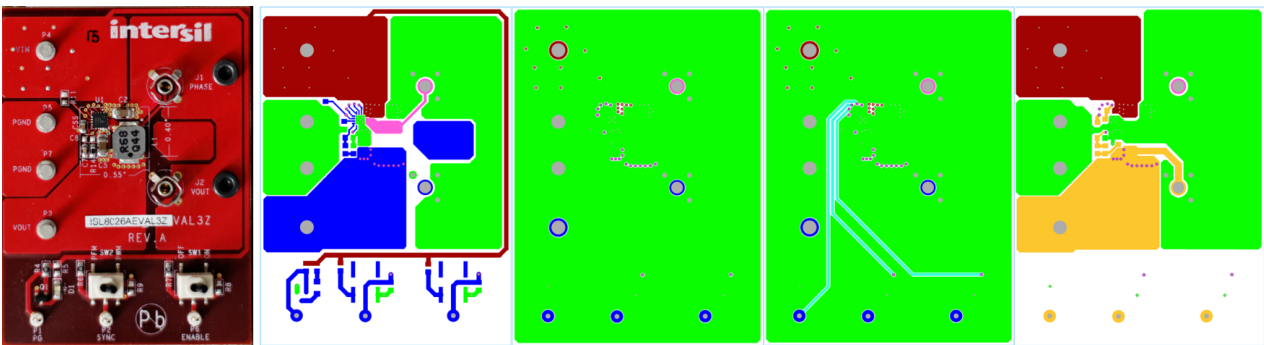


FIGURE 1.1.6 – Carte ISL8026EVAL3Z (vue du dessus et des différentes couches métalliques (en couleur) du PCB)

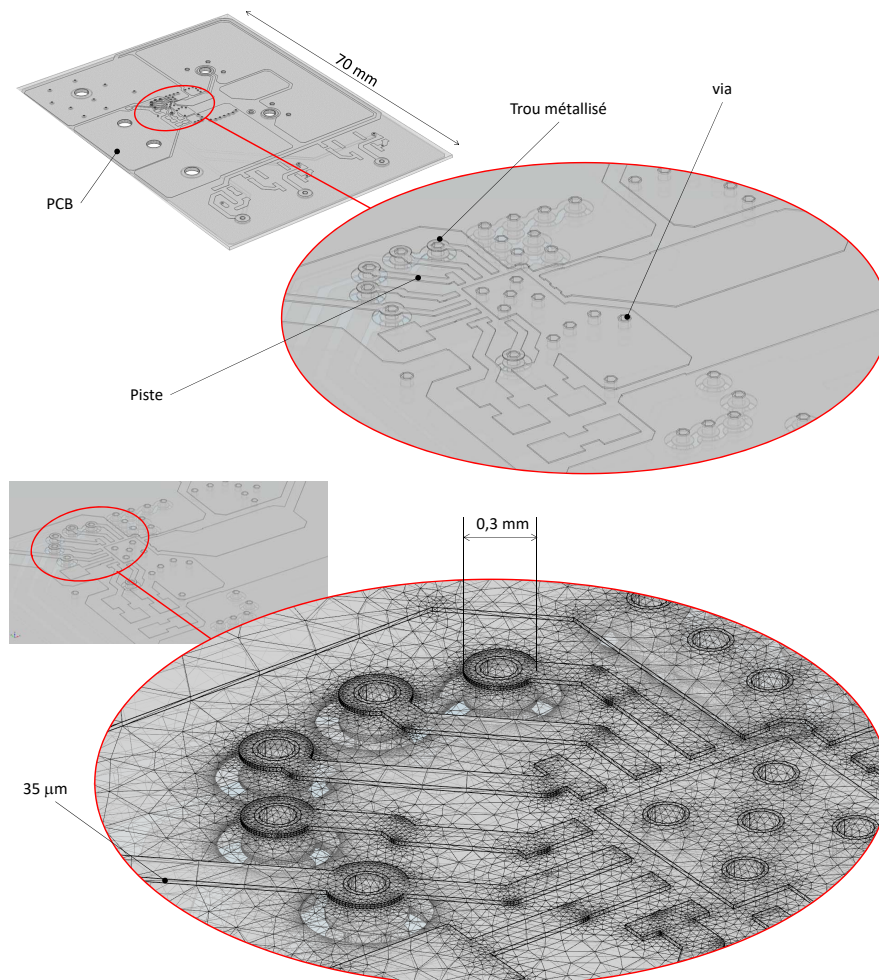


FIGURE 1.1.7 – Détail du PCB de la carte ISL8026EVAL3Z, avec maillage sur la partie zoomée

couches atteint le million de nœuds pour le PCB seul. Et il faut encore y ajouter les maillages des composants...

Une approche pour se passer de calculs numériques lourds est d'utiliser des modèles de cartes analytiques. Une grande majorité utilise des modèles de carte mono-couche [7, 8], où les différentes couches du PCB sont assimilées à un bloc orthotrope équivalent. D'autres utilisent des modèles multicouches, considérées être en contact parfait [9, 10] ou présentant une résistance thermique à l'interface [11, 12] dans des travaux plus récents. Dans tous les cas, ces modèles présentent beaucoup de simplifications, et ne permettent pas de représenter le PCB de façon détaillée.

1.3 Modélisation des composants

On s'intéresse maintenant à la modélisation au niveau composant. Ceux-ci sont majoritairement constitués d'une ou plusieurs puces encapsulées dans un boîtier électronique, ou "package". On les classe alors selon leur méthode de report sur le PCB en deux grandes catégories : les composants "à piquer", dont les connexions sont des pattes insérées puis soudées dans les trous métallisés de la carte (appelés PTH pour "Plated Through Hole"), et les Composants à Montage Surfacique (CMS, ou SMD pour "Surface Mounted Device" pour les anglo-saxons), implantés par un soudage dit à la vague.

Si les composants PTH représentaient plus de 80% du marché de l'électronique dans les années 80, ils ne représentent aujourd'hui même pas 5% des composants montés sur les cartes, on va donc s'intéresser aux CMS. On distingue plusieurs types de boîtiers pour ces derniers. Quelques exemples de ces types de boîtiers sont décrits ci-dessous.

1.3.1 Les boîtiers SO

Les boîtiers SO ("Small Outline Package") et SOT ("Small Outline Transistor") et autres SOP, SOJ, SSOP, TSOP, TSSOP... représentent les boîtiers les plus "populaires" utilisés depuis le début des années 90. Ces boîtiers sont constitués d'une puce, généralement collée à une plateforme en cuivre et reliée à deux rangées de broches par des fil en or. La plateforme et les broches sont réalisées en même temps par une attaque chimique et le tout est ensuite placé dans un moule pour être encapsulé dans une résine. Les broches sont enfin pliées en forme de L (figure 1.1.8).

Mais ces boîtiers sont en voie de disparition depuis quelques années, supplantés par les QFN, CSP et BGA.

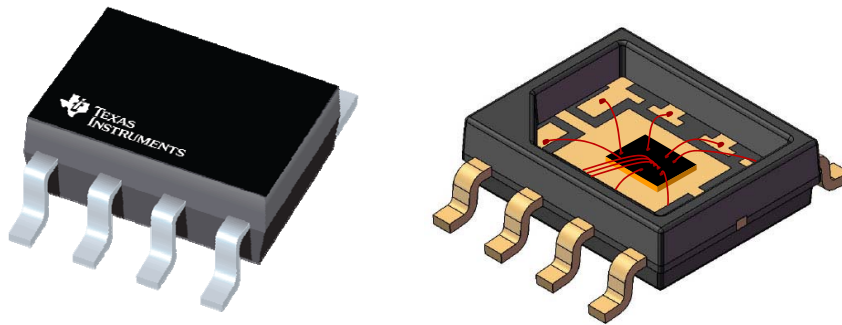
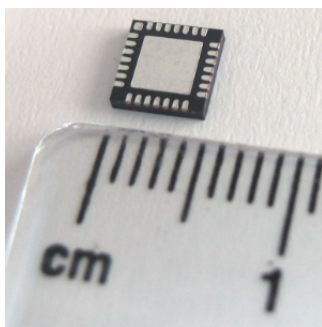


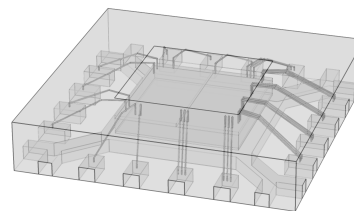
FIGURE 1.1.8 – Exemple de boîtier SO (Texas Instruments)

1.3.2 Les boîtiers QFN

Les boîtiers QFN (pour “Quad Flat No-Lead”) sont nés de la demande de miniaturisation évoquée plus haut, qui pousse à vouloir des pattes toujours plus fines. Or, un pas plus fin que 0.5mm fragilise la mise en forme de ces pattes. L’intégration de broches droites et courtes directement dans la résine d’encapsulation (d’où le “Flat No-Lead”) a permis de résoudre ce problème et de réduire énormément l’encombrement des composants, tout en multipliant le nombre de connexions possibles, correspondant au nombre toujours croissant d’entrées/sorties souhaitées.



(a) Photo d’un boîtier QFN28 de dimensions $4\text{mm} \times 4\text{mm} \times 0,2\text{mm}$



(b) Modèle 3D de QFN16

FIGURE 1.1.9 – Exemples de boîtiers QFN

La figure 1.1.9 présente ainsi deux exemples de boîtiers QFN, la figure 1.1.9a présentant un boîtier à 28 broches de connexion, tandis que la figure 1.1.9b représente un modèle DAO d’un boîtier QFN16, où l’on peut distinguer la puce et les différents fils de connexions reliant cette dernière aux 16 broches du package.

On peut aussi parler de boîtier DFN (“Dual Flat No-Lead”) pour des boîtiers construits

selon le même procédé mais avec seulement deux rangées de broches au lieu de quatre.

1.3.3 Les boîtiers BGA

Comme leur nom l'indique, les boîtiers BGA (pour "Ball Grid Array") sont des boîtiers dont les connexions à la carte sont assurées par des billes de 1 mm de diamètre disposées sous forme de matrice directement sur la surface inférieure du composant, et qui vont servir à la brasure. Ces boîtiers sont surtout utilisés pour les microprocesseurs ou les mémoires, qui nécessitent un grand nombre d'entrée-sorties pour un faible encombrement. En effet, ils permettent d'avoir de 156 à 1156 billes de connexion pour des boîtiers carrés de 15 à 35 mm de côté. La figure 1.1.10 montre un exemple de BGA, et notamment la matrice des billes de connexion tandis que la figure 1.1.11 illustre la complexité de modéliser en détail de tels composants.



FIGURE 1.1.10 – Exemple de package BGA : processeur graphique NVIDIA GeForce GTX 980M

1.4 Vers la réduction de modèle

Les sections précédentes permettent de mettre en évidence que des simulations détaillées de cartes électroniques seraient bien trop coûteuses à mettre en œuvre. En effet, en plus du PCB qui présente un grand niveau de détails, on doit pouvoir modéliser en même temps tous ses composants, ce qui représente des dizaines, des centaines voire des milliers de nouvelles structures à modéliser. Ces composants présentent aussi un grand niveau de détails et de grandes disparités en termes de propriétés thermiques, de par la variété des matériaux utilisés

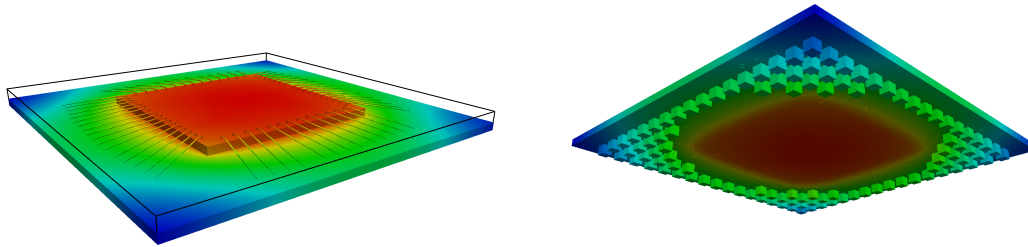


FIGURE 1.1.11 – Modèle de BGA208 avec 60 fils d’or de connexion

(cuivre, or, silicium, résine époxy, colles, ...). Ainsi, si on regarde le problème d’un point de vue global, on a une très grande différence d’échelles entre la carte, dont la longueur est de l’ordre de 20 cm , et les fils de connexion reliant les puces des composants et les broches brasées sur le PCB, d’un diamètre de l’ordre de $20\ \mu\text{m}$, soit un rapport 10^4 . Les maillages nécessaires pour une simulation avec des méthodes classiques de type différences finies, volumes finis ou éléments finis dépassent aisément la dizaine de millions de nœuds. Or, l’objectif de sociétés fabriquant ces cartes telles que Thales est de pouvoir faire ces simulations dès la phase de conception des cartes.

S’il est possible de faire une simulation de ce type sur un super-calculateur, les industriels ont besoin de faire plusieurs dizaines de simulations, correspondant aux différents scénarios d’utilisation de leurs produits qui peuvent être soumis à de larges variations de conditions aux limites. Faire des simulations détaillées à partir de ce type de modèle est alors inenvisageable pour tenir des délais de conception raisonnables. D’autant plus que la thermique ne constitue pas le cœur de métier des sociétés d’électronique, car, même si elle est un facteur limitant, elle reste un effet périphérique. D’autres simulations doivent aussi être effectuées en phase de conception, comme des simulations mécaniques, électroniques bien sûr, ... C’est pourquoi on va chercher à obtenir des modèles thermiques réduits de cartes électroniques.

La plupart des composants entrant dans la conception des cartes électroniques étant standardisés et susceptibles d’être utilisés pour de multiples configurations, il est pertinent de créer des bibliothèques de composants incluant leur modèle thermique réduit. On peut ainsi établir un cahier des charges pour créer cette bibliothèque ; les modèles réduits doivent :

1. Être sobres en puissance de calcul, puisque c’est l’aspect le plus important de notre problématique : réduire le temps de calcul ;
2. Accepter toutes conditions aux limites, pour qu’on puisse simuler les différents scénarios imposés par l’industriel et prendre en compte les multiples usages du composant, pour une même configuration ;

3. Présenter une grande finesse géométrique, afin que, malgré la réduction, on puisse accéder au champ de température en intégralité ;
4. Et enfin qu'ils soient réutilisables et compatibles avec n'importe quel PCB.

Pour déterminer la méthode que l'on va utiliser pour créer des modèles réduits répondant à ce cahier des charges, les chapitres 2 et 3 seront dédiés à un état de l'art des méthodes de sous-structuration et des méthodes de réduction. À la suite de cette étude, la deuxième partie de ce manuscrit sera la mise en application de la méthode retenue pour atteindre cet objectif : la sous-structuration avec joints couplée à la réduction modale. Le chapitre 4 servira à établir le problème couplé issu de la sous-structuration tandis que, dans le chapitre 5, on y appliquera la méthode de réduction modale. Enfin, dans la troisième partie de ce manuscrit, on présentera des exemples d'application pour éprouver la méthode développée dans la partie précédente.

Chapitre 2

État de l'art des méthodes de sous-structuration

Dans ce chapitre, on présente un historique des différentes méthodes de sous-structuration, avant de décrire la méthode utilisée en pratique pour coupler les sous-structures d'un problème thermique.

2.1 Problème étudié

L'équation modèle qui va servir à illustrer les différentes techniques est l'équation de Poisson

$$\begin{cases} -\Delta u = f & \text{dans } \Omega, \\ u = 0 & \text{sur } \partial\Omega, \end{cases} \quad (1.2.1)$$

où $f \in L^2(\Omega)$, avec Ω un ouvert borné et régulier de \mathbb{R}^d ($d \geq 2$) (comme par exemple sur la figure 1.2.1).

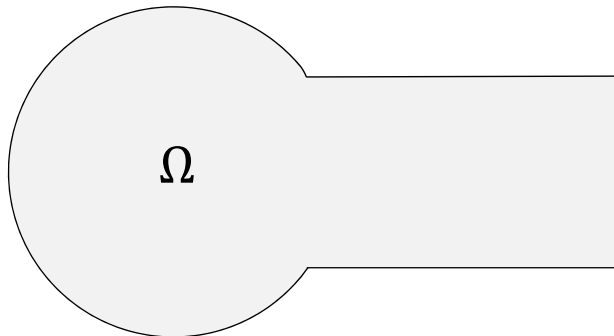


FIGURE 1.2.1 – Exemple de domaine Ω pour l'équation de Poisson.

La formulation faible de (1.2.1) est : trouver $u \in H_0^1(\Omega)$ tel que $\forall v \in H_0^1(\Omega)$,

$$\int_{\Omega} \nabla u \cdot \nabla v = \int_{\Omega} f v. \quad (1.2.2)$$

Une discrétisation spatiale du problème mène à la formulation matricielle suivante :

$$\mathbf{K}\mathbf{u} = \mathbf{f}. \quad (1.2.3)$$

Remarque 2.1.1 *Pour alléger l'écriture des équations dans tout le manuscrit, et quand il n'y a pas de risque d'ambiguïté, on adopte la convention suivante :*

$$\begin{aligned} \int_{\Omega} f(x) dx &\equiv \int_{\Omega} f dv \equiv \int_{\Omega} f, \\ \int_{\partial\Omega} f(\sigma) d\sigma &\equiv \int_{\partial\Omega} f da \equiv \int_{\partial\Omega} f. \end{aligned}$$

2.2 Les origines de la sous-structuration

Cette section et la suivante reposent principalement sur l'article de De Klerk, Rixen et Voormeeren [13] qui propose un historique, une revue ainsi qu'une classification des différentes méthodes de sous-structuration.

Historiquement, les méthodes de sous-structuration font partie de la famille des méthodes de décomposition de domaines. Le principe de ces méthodes est de considérer séparément le problème à résoudre sur les différents sous-domaines qui le constituent d'une part, et le problème à leur interface d'autre part. On peut classer ces méthodes en deux grandes catégories : les méthodes avec recouvrement, ou méthodes de Schwarz, et les méthodes sans recouvrement, ou méthodes du complément de Schur.

2.2.1 Méthodes avec recouvrement : méthodes de Schwarz

La méthode de décomposition de domaine avec recouvrement a été introduite par Hermann Schwarz en 1869-1870 [14] pour démontrer l'existence et l'unicité de la solution d'un problème de Laplace sur le domaine Ω (figure 1.2.1) considéré comme l'union d'un cercle et d'un rectangle qui se recouvrent, comme en figure 1.2.2.

Des problèmes locaux de Dirichlet sont résolus sur chaque sous-domaine, et le raccordement est assuré par la zone de recouvrement $\Omega^{(1)} \cap \Omega^{(2)}$, où la méthode itérative développée par Schwarz converge vers une solution unique. Les différentes méthodes de Schwarz correspondent aux différents algorithmes utilisés pour les itérations : méthode alternée [15, 16, 17] ou additive [18, 19] (particulièrement développée pour le calcul parallèle).

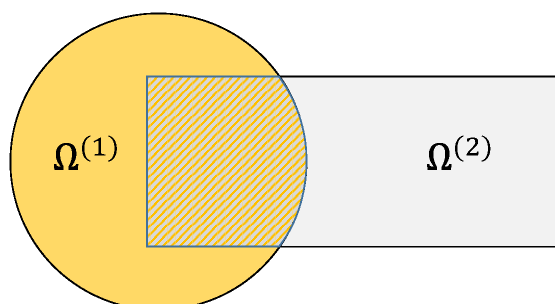
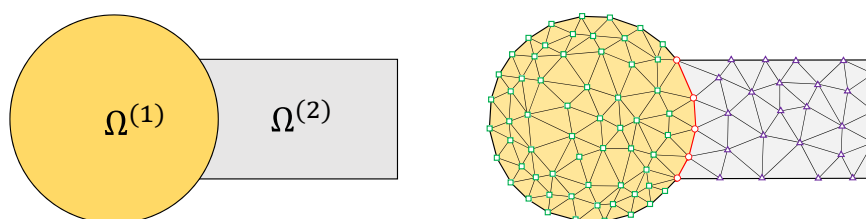


FIGURE 1.2.2 – Décomposition de domaine avec recouvrement (partie hachurée).

2.2.2 Méthode du complément de Schur

On fait un petit bond dans le temps, puisque cette méthode a été développée à partir des années 1960 ([20]). Cette fois-ci, les domaines ne se recouvrent pas (figure 1.2.3), et la solution est cherchée de manière directe et non plus itérative.

FIGURE 1.2.3 – Décomposition de domaine sans recouvrement ; l'interface Γ est représentée en rouge sur le schéma de droite.

La solution du problème est d'abord cherchée sur l'interface $\Gamma = \overline{\Omega^{(1)}} \cap \overline{\Omega^{(2)}}$ des deux sous-domaines grâce à la formation du complément de Schur, et la solution sur $\Omega^{(1)}$ et $\Omega^{(2)}$ en est déduite. On sépare donc les inconnues du problème en deux ensembles : d'une part les inconnues internes des sous-domaines (les carrés verts pour $\Omega^{(1)}$ et les triangles violets pour $\Omega^{(2)}$, voir la figure 1.2.3, à droite) et d'autre part les inconnues à l'interface de ces sous-domaines.

Cette méthode est particulière, puisqu'elle s'applique sur un problème déjà discrétisé et propose de le résoudre en réorganisant le système matriciel.

Pour l'illustrer, prenons le cas où on a discrétisé un problème à résoudre sur Ω pour obtenir le système matriciel suivant :

$$\begin{bmatrix} K_{11} & 0 & K_{1\Gamma} \\ 0 & K_{22} & K_{2\Gamma} \\ K_{\Gamma 1} & K_{\Gamma 2} & K_{\Gamma\Gamma} \end{bmatrix} \begin{bmatrix} u_{\square}^{(1)} \\ u_{\triangle}^{(2)} \\ u_{\circ}^{(\Gamma)} \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_{\Gamma} \end{bmatrix}, \quad (1.2.4)$$

où $u_{\square}^{(1)}$ est le vecteur d'inconnues sur $\Omega^{(1)} \setminus \Gamma$, $u_{\triangle}^{(2)}$ le vecteur d'inconnues sur $\Omega^{(2)} \setminus \Gamma$ et $u_{\circ}^{(\Gamma)}$ le vecteurs d'inconnues sur Γ .

Si K_{11} et K_{22} sont inversibles, le complément de Schur est alors la matrice définie par

$$S = K_{\Gamma\Gamma} - K_{\Gamma 1} K_{11}^{-1} K_{1\Gamma} - K_{\Gamma 2} K_{22}^{-1} K_{2\Gamma}. \quad (1.2.5)$$

En posant $g_{\Gamma} = f_{\Gamma} - K_{\Gamma 1} K_{11}^{-1} f_1 - K_{\Gamma 2} K_{22}^{-1} f_2$, la méthode consiste ensuite à résoudre

$$S u_{\circ}^{(\Gamma)} = g_{\Gamma}, \quad (1.2.6)$$

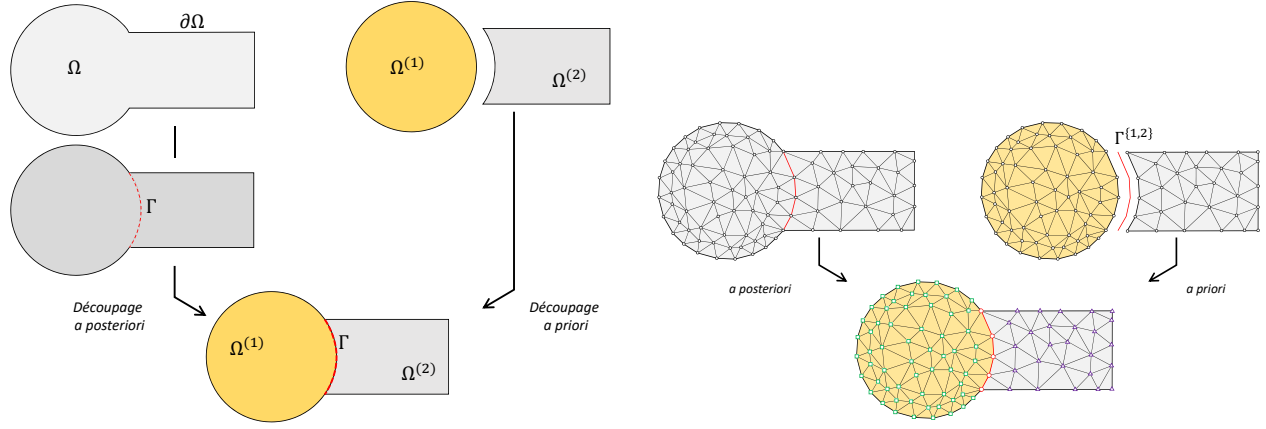
avec une méthode de type gradient conjugué pour obtenir u^{Γ} . On peut ensuite paralléliser le problème pour déduire $u^{(1)}$ et $u^{(2)}$.

2.2.3 Deux philosophies pour décomposer Ω

La méthode de Schur est une méthode de décomposition *a posteriori*, puisque l'on est parti d'un problème à résoudre sur le domaine Ω , que l'on a seulement ensuite découpé en deux sous-domaines $\Omega^{(1)}$ et $\Omega^{(2)}$ suivant une interface Γ (qui peut être fictive). Ce découpage permet de chercher la solution locale du problème, sur Γ , avant de la déduire sur les sous-domaines, et donc sur Ω .

Cette méthodologie n'est pas compatible avec notre problématique, puisque l'on cherche à résoudre l'équation de la chaleur sur plusieurs systèmes bien définis, que l'on veut ensuite coupler : le découpage de Ω est alors *a priori* (figure 1.2.4).

On va donc s'intéresser dans la suite aux méthodes permettant de résoudre un problème sous-structuré selon un découpage *a priori*, ce qui nécessite de faire apparaître des conditions de couplage entre les sous-structures pour assurer la continuité de la solution aux interfaces.



(a) Deux façon de découper Ω : *a posteriori* et *a priori*. (b) Découpage avec des domaines déjà discrétisés.

FIGURE 1.2.4 – Différents types de décomposition de domaine.

2.3 Sous-structuration d'un système couplé

On part donc cette fois de deux domaines disjoints $\Omega^{(1)}$ et $\Omega^{(2)}$, tels que $\Omega = \Omega^{(1)} \cup \Omega^{(2)}$ et $\Omega^{(1)} \cap \Omega^{(2)} = \emptyset$, que l'on va coupler par leur interface commune (voir figure 1.2.5). On note $\Gamma = \overline{\Omega^{(1)}} \cap \overline{\Omega^{(2)}}$ et $\Gamma^{1 \rightarrow 2} = \overline{\Omega^{(1)}} \cap \Gamma$ l'interface "du côté de $\Omega^{(1)}$ ". Similairement, on note $\Gamma^{2 \rightarrow 1} = \overline{\Omega^{(2)}} \cap \Gamma$.

Enfin, on définit la frontière avec le milieu extérieur $\Sigma^{(k)} = \partial\Omega^{(k)} \cap \partial\Omega$ et $\underline{n}_{k,l}$, la normale extérieure à $\Omega^{(k)}$ sur Γ , $k, l \in \{1, 2\}$, $k \neq l$.

Le problème couplé s'écrit alors :

$$\begin{cases} -\Delta u^{(i)} = f & \text{dans } \Omega^{(i)}, & (1.2.7a) \\ u^{(i)} = 0 & \text{sur } \Sigma^{(i)}, & (1.2.7b) \\ u^{(2)} - u^{(1)} = 0 & \text{sur } \Gamma, & (1.2.7c) \\ \varphi^{(2)} + \varphi^{(1)} = 0 & \text{sur } \Gamma, & (1.2.7d) \end{cases}$$

avec $u^{(i)} = u|_{\Omega^{(i)}}$, $i = 1, 2$ et $\varphi^{(i)} = \frac{\partial u^{(i)}}{\partial \underline{n}_i}$ le flux sortant de $\Omega^{(i)}$, \underline{n}_i étant la normale sortante de $\Omega^{(i)}$ par Γ . Les deux dernières équations traduisent respectivement la continuité de la température et du flux de chaleur à l'interface pour un contact parfait.

Pour chaque sous structure, les équations s'écrivent sous forme matricielle

$$\mathbf{K}^{(i)} \mathbf{u}^{(i)} = \mathbf{f}^{(i)} + \mathbf{\Phi}^{(i)}, \quad (1.2.8)$$

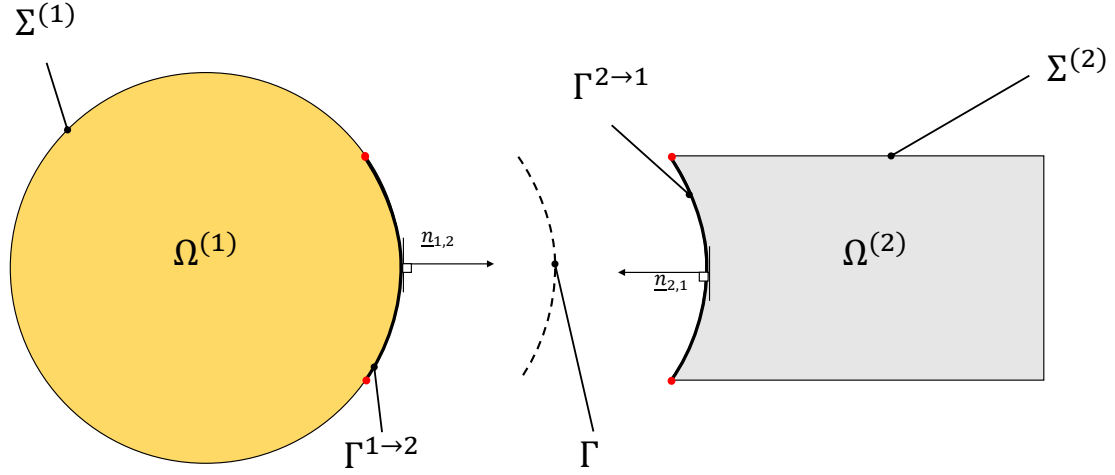


FIGURE 1.2.5 – Partition régulière du domaine Ω en deux sous domaines $\Omega^{(1)}$ et $\Omega^{(2)}$.

avec $\mathbf{u}^{(i)}$ le vecteur des inconnues (*i.e.* les nœuds du maillages dans le cas d'éléments finis), $\mathbf{f}^{(i)}$ le vecteur des sollicitations extérieures et $\Phi^{(i)}$ le vecteur des flux échangés avec les autres sous-structures.

Pour les deux sous-structures, le problème matriciel défini par les équations (1.2.7a) et (1.2.7b) s'écrit

$$\mathbf{K}\mathbf{u} = \mathbf{f} + \Phi, \quad (1.2.9)$$

avec $\mathbf{K} = \text{diag}(\mathbf{K}^{(1)}, \mathbf{K}^{(2)})$,

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}^{(1)} \\ \mathbf{u}^{(2)} \end{bmatrix}, \quad \mathbf{f} = \begin{bmatrix} \mathbf{f}^{(1)} \\ \mathbf{f}^{(2)} \end{bmatrix}, \quad \Phi = \begin{bmatrix} \Phi^{(1)} \\ \Phi^{(2)} \end{bmatrix}.$$

Pour obtenir le système matriciel couplé, on doit ajouter à (1.2.9) la condition de compatibilité (1.2.7c)

$$\mathbf{J}\mathbf{u} = 0, \quad (1.2.10)$$

et la condition d'équilibre (1.2.7d)

$${}^t\mathbf{L}\Phi = 0. \quad (1.2.11)$$

\mathbf{J} est une matrice booléenne (dans le cas d'un maillage conforme aux interfaces¹) qui indique quels points de la surface d'une sous-structure sont en contact avec les points de la

1. La définition d'un maillage conforme aux interfaces sera donnée un peu plus loin.

surface de la sous-structure qui y est couplée, et \mathbf{L} est la matrice qui localise les points de l'interface des sous-structures au sein du maillage global. Mathématiquement, on a $\mathbf{L} = \ker(\mathbf{J})$.

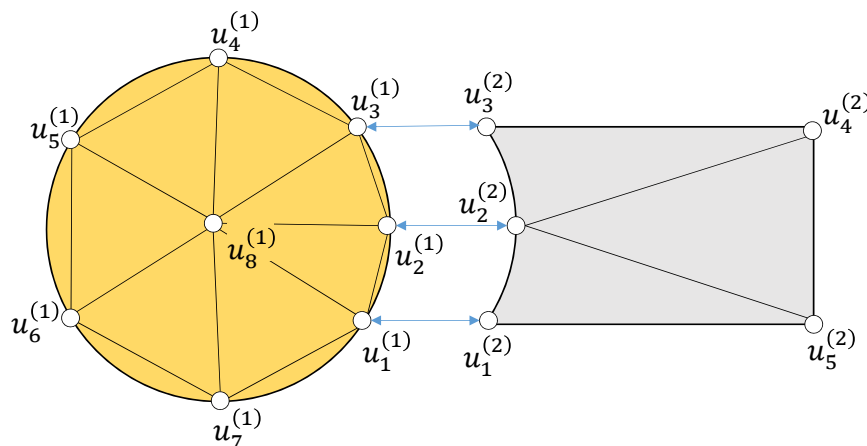


FIGURE 1.2.6 – Exemple de discrétisation de Ω .

La figure 1.2.6 présente un exemple de discrétisation de la figure 1.2.5. On a

$${}^t \mathbf{u} = \left[u_1^{(1)} \quad u_2^{(1)} \quad u_3^{(1)} \quad u_4^{(1)} \quad u_5^{(1)} \quad u_6^{(1)} \quad u_7^{(1)} \quad u_8^{(1)} \quad u_1^{(2)} \quad u_2^{(2)} \quad u_3^{(2)} \quad u_4^{(2)} \quad u_5^{(2)} \right], \quad (1.2.12)$$

et le couplage entre $\Omega^{(1)}$ et $\Omega^{(2)}$ est alors assuré par les points

$$\begin{cases} u_1^{(1)} = u_1^{(2)}, \\ u_2^{(1)} = u_2^{(2)}, \\ u_3^{(1)} = u_3^{(2)}. \end{cases} \quad (1.2.13)$$

La matrice \mathbf{J} s'écrit alors :

$$\mathbf{J} = \begin{array}{c} \begin{array}{ccccccccccccccc} u_1^{(1)} & u_2^{(1)} & u_3^{(1)} & u_4^{(1)} & u_5^{(1)} & u_6^{(1)} & u_7^{(1)} & u_8^{(1)} & u_1^{(2)} & u_2^{(2)} & u_3^{(2)} & u_4^{(2)} & u_5^{(2)} \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \end{array} \end{array},$$

et la matrice \mathbf{L} s'écrit :

$$\mathbf{L} = \begin{array}{c} \begin{array}{c} u_1^{(1)} \\ u_2^{(1)} \\ u_3^{(1)} \\ u_4^{(1)} \\ u_5^{(1)} \\ u_6^{(1)} \\ u_7^{(1)} \\ u_8^{(1)} \\ u_1^{(2)} \\ u_2^{(2)} \\ u_3^{(2)} \\ u_4^{(2)} \\ u_5^{(2)} \end{array} \end{array} \begin{array}{|cccccccccc|} \hline 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline \end{array}.$$

On obtient donc le système couplé suivant :

$$\begin{cases} \mathbf{K}\mathbf{u} = \mathbf{f} + \Phi, \\ \mathbf{J}\mathbf{u} = 0, \\ {}^t\mathbf{L}\Phi = 0. \end{cases} \quad (1.2.14)$$

Pour résoudre le système couplé (1.2.14), deux solutions s'offrent à nous : la formulation primale ou la formulation duale.

2.3.1 Formulation primale

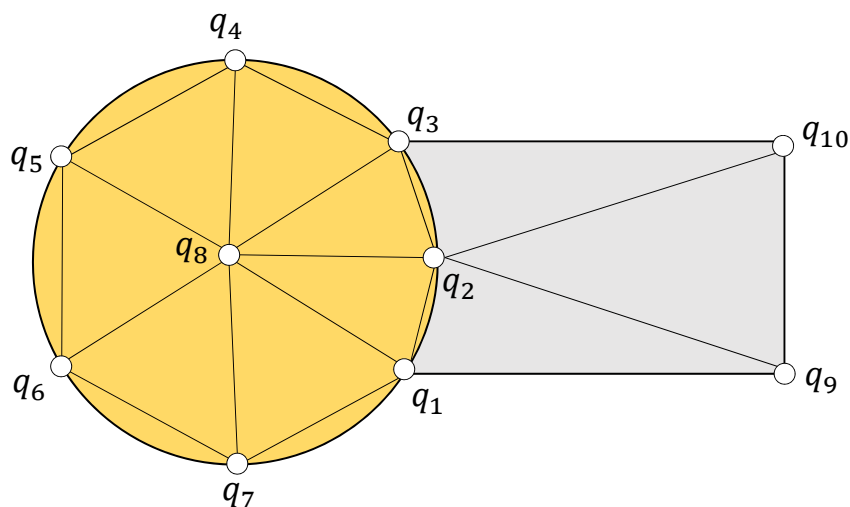
La formulation primale est la plus utilisée pour assembler les modèles éléments finis. Le principe de cette formulation est de définir un ensemble unique de degrés de liberté d'interface et d'éliminer les inconnues correspondantes aux flux d'interface à partir de la condition d'équilibre (1.2.11), *i.e.* en écrivant

$$\mathbf{u} = \mathbf{L}\mathbf{q}, \quad (1.2.15)$$

avec \mathbf{q} l'ensemble des degrés de libertés de la discrétisation de Ω (*i.e.* l'ensemble des $u_i^{(1)}$ et des $u_j^{(2)}$, sans répétition des points à l'interface, voir figure 1.2.7) :

$${}^t\mathbf{q} = \left[q_1 \quad q_2 \quad q_3 \quad q_4 \quad q_5 \quad q_6 \quad q_7 \quad q_8 \quad q_9 \quad q_{10} \right]. \quad (1.2.16)$$

Puisque $\mathbf{L} = \ker(\mathbf{J})$, $\mathbf{J}\mathbf{u} = \mathbf{J}\mathbf{L}\mathbf{q} = 0 \forall \mathbf{q}$, donc la condition de compatibilité (1.2.10) est

FIGURE 1.2.7 – Exemple de discrétisation de Ω .

implicitement respectée. Le système (1.2.14) devient donc

$$\begin{cases} \mathbf{KLq} = \mathbf{f} + \Phi, \\ {}^t\mathbf{L}\Phi = 0. \end{cases} \quad (1.2.17)$$

En multipliant les équations par \mathbf{L}^T , on obtient la forme compacte :

$$\tilde{\mathbf{K}}\mathbf{q} = \tilde{\mathbf{f}}, \quad (1.2.18)$$

avec

$$\begin{cases} \tilde{\mathbf{K}} = {}^t\mathbf{LKL}, \\ \tilde{\mathbf{f}} = {}^t\mathbf{L}\mathbf{f}. \end{cases}$$

2.3.2 Formulation duale

Cette fois-ci, tous les degrés de liberté sont conservés, et en particulier les degrés de liberté d'interface sont présents autant de fois qu'ils apparaissent dans les différents contacts entre sous-structures.

La forme duale est alors obtenue en satisfaisant la condition d'équilibre (1.2.11), en exprimant les flux Φ à l'interface sous la forme :

$$\Phi = -{}^t\mathbf{J}\lambda, \quad (1.2.19)$$

où λ est un ensemble de multiplicateurs de Lagrange.

On a donc bien $\mathbf{L}^T \Phi = -\mathbf{L}^T \mathbf{J}^T \lambda = 0$, et on obtient la formulation duale de (1.2.14) :

$$\begin{cases} \mathbf{K}\mathbf{u} + {}^t\mathbf{J}\lambda = \mathbf{f}, \\ \mathbf{J}\mathbf{u} = 0. \end{cases} \quad (1.2.20)$$

Sous forme matricielle, (1.2.20) s'écrit

$$\begin{pmatrix} \mathbf{K} & {}^t\mathbf{J} \\ \mathbf{J} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \lambda \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ 0 \end{pmatrix}. \quad (1.2.21)$$

Ce sont ces méthodes qui seront très développées dans les années 1990 avec l'avènement du calcul parallèle, notamment la méthode FETI.

Remarque 2.3.1 *L'équation obtenue en (1.2.21) est un problème de point selle, qui peut être résolu avec un algorithme d'Uzawa [21].*

2.4 La création de bibliothèques de composants : les maillages non-conformes

Comme dit en introduction, dans les exemples traités lors de cette thèse, à savoir les cartes électroniques, il est très intéressant de pouvoir faire un travail de réduction sur les composants électroniques avant de placer leurs modèles réduits sur une carte, afin de rentabiliser encore plus le temps nécessaire à l'obtention de ces modèles réduits et de pouvoir les réutiliser pour de nouvelles configurations.

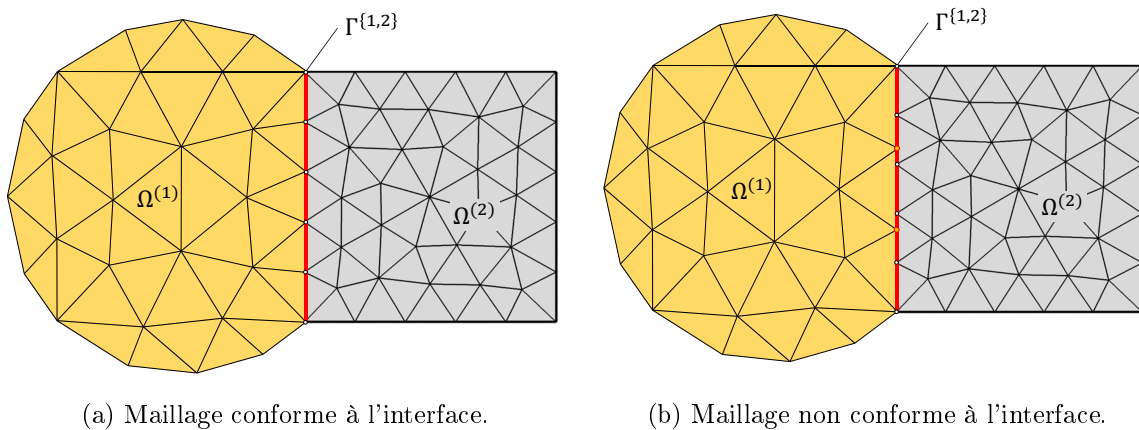


FIGURE 1.2.8 – Maillages conforme et non-conforme aux interfaces.

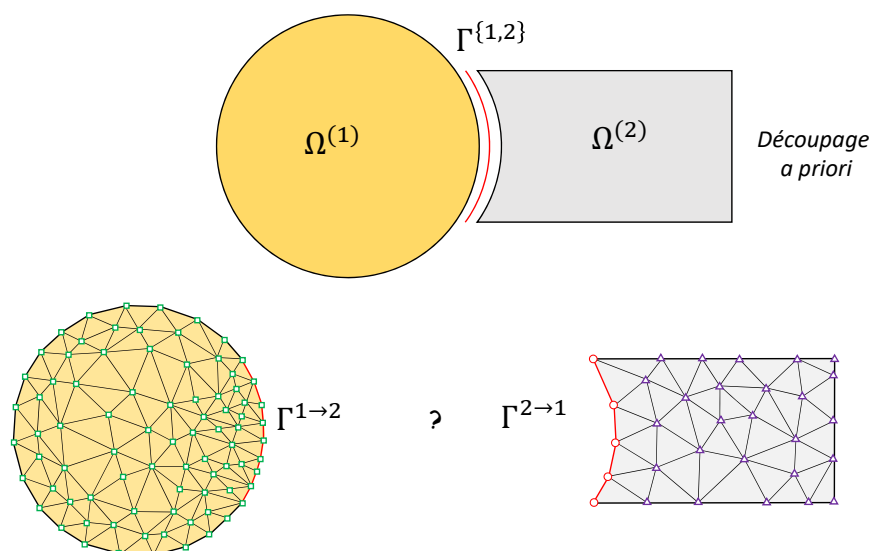


FIGURE 1.2.9 – Comment exprimer le couplage lorsque l'on veut raccorder deux sous-structures dans un découpage *a priori*, maillées indépendamment ?

L'objectif est donc de rendre nos méthodes compatibles avec des maillages non-conformes aux interfaces (voir figure 1.2.8), pour s'affranchir de cette limitation et ainsi réutiliser des modèles sous-structurés dans plusieurs configurations.

La difficulté réside ainsi dans l'expression de la matrice \mathbf{J} , qui sert à exprimer la condition de compatibilité (1.2.10) et qu'on retrouve notamment dans la formulation duale (1.2.21). Si cette matrice est une simple matrice booléenne dans le cas de maillages conformes aux interfaces, son expression est bien plus complexe dans le cas de maillages non-conformes.

On présente donc dans les sections suivantes les différentes méthodes de sous-structuration qui ont été développées pour répondre à ce problème particulier, avant de spécifier la méthode retenue.

Remarque 2.4.1 *Parler de conformité, ou de non-conformité, d'un maillage peut avoir plusieurs sens, notamment un sens géométrique. On peut parler de maillage géométriquement non-conforme lorsque celui-ci ne coïncide pas avec les bords du domaine à mailler, typiquement lorsqu'on souhaite mailler un domaine courbe (voir figure 1.2.10) : la frontière du maillage ne coïncide pas avec la frontière du domaine, en rouge.*

Ainsi, pour la suite, lorsque l'on parlera de maillage non-conforme, c'est uniquement pour désigner des maillages dont les nœuds ne coïncident pas sur une frontière commune.

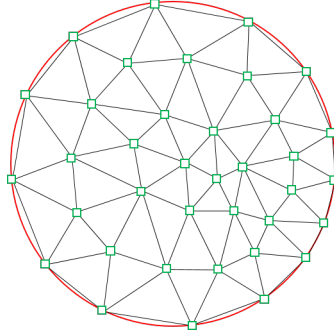


FIGURE 1.2.10 – Maillage géométriquement non-conforme.

2.5 La méthode des joints

La méthode des joints a été introduite pour la première fois par Bernardi *et al.* [22, 23] au début des années 90 pour la résolution d'équations aux dérivées partielles elliptiques en 2D, puis étudiée par de nombreux auteurs (liste non exhaustive : Abdoulaev [24], Belgacem [25, 26, 27], Maday [28], Lacour [29], Braess [30, 31, 32], Danek [33], Dryja [34], Dittmann [35], Hild [36], Hoppe [37], Rapetti [38, 39], Wohlmuth [40, 41, 42] et Krause[43]).

2.5.1 Partition de Ω

La méthode des joints s'appliquant sur des problèmes multi-domaines, on suppose pour cette explication que Ω est une union de $N_\Omega + 1$ sous-domaines telle que $\Omega = \bigcup_{i=0}^{N_\Omega} \overline{\Omega^{(i)}}$.

On prend les mêmes notations qu'en section 2.1 (page 15), et on note $\Gamma^{\{i,j\}} = \overline{\Omega^{(i)}} \cap \overline{\Omega^{(j)}}$ la frontière commune entre $\Omega^{(i)}$ et $\Omega^{(j)}$, $0 \leq i, j \leq N_\Omega$ (l'ordre des exposants est indifférent : $\Gamma^{\{i,j\}} \equiv \Gamma^{\{j,i\}}$). Γ désigne désormais le squelette interne de la partition de Ω , *i.e.* l'ensemble des interfaces entre les sous-domaines de Ω :

$$\Gamma = \bigcup_{i,j \in \{0, N_\Omega\}} \Gamma^{\{i,j\}}.$$

Attention, alors qu'en section 2.1, Γ désignait seulement $\Gamma^{\{1,2\}}$, puisqu'il n'y avait qu'une interface entre deux sous-domaines, ici Γ représente l'ensemble des interfaces, peu importe la décomposition de Ω , et ce pour tout le reste du manuscrit.

Cette décomposition peut être géométriquement conforme ou non (à ne pas confondre avec la notion de maillage non conforme évoquée dans la remarque 2.4.1). On parlera de décomposition géométriquement conforme lorsque l'intersection de deux sous-domaines $\Omega^{(i)}$

est soit vide, soit un point, soit une arête entière, soit une face entière pour les deux (voir figure 1.2.11a), et non conforme sinon (voir figure 1.2.11b).

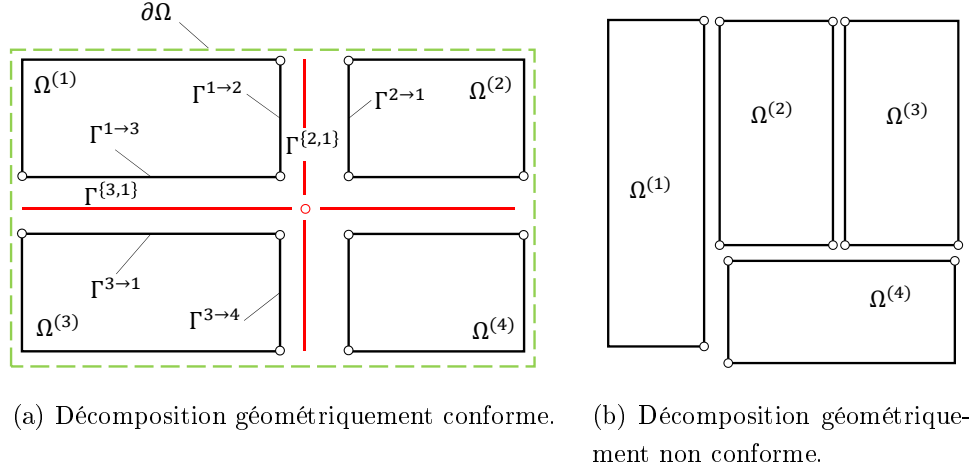


FIGURE 1.2.11 – Différents types de décompositions.

Le nom de la méthode des joints vient de la façon de partitionner le squelette, et plus particulièrement du choix de l'une des deux frontières situées de part et d'autre de chaque interface $\Gamma^{\{i,j\}}$ ($\Gamma^{i \rightarrow j}$ ou $\Gamma^{j \rightarrow i}$) en tant que *frontière mortier* (on peut aussi trouver les dénominations *frontière joint* ou *frontière maître*). Par abus de langage, on parlera parfois juste de *mortier* pour désigner cette frontière. L'autre frontière est alors appelée *frontière non-mortier* (ou *frontière non-joint*, ou *frontière esclave*). Pour chaque mortier, le sous-domaine associé est appelé *sous-domaine mortier* ou *sous-domaine maître* et pour chaque non-mortier, on appelle le sous-domaine associé *sous-domaine non-mortier* ou *sous-domaine esclave*.

L'ensemble des frontières mortiers forme alors une partition de Γ . De fait, l'ensemble des frontières non-mortiers forme aussi une partition de Γ . Un point important est que ces frontières mortiers doivent représenter des arêtes (en 2D) ou des faces (en 3D) entières de la frontière du sous-domaine considéré : on ne peut pas diviser une arête (ou une face) en une partie mortier et une autre non-mortier.

Pour illustrer cette méthode de partitionnement, on peut repartir des exemples donnés en figure 1.2.11.

Pour le cas conforme (figure 1.2.12), le partitionnement est aisé : pour chaque $\Gamma^{i \rightarrow j}$ choisie comme frontière mortier, $\Gamma^{j \rightarrow i}$ est de fait non-mortier.

Par exemple, on a pris $\Gamma^{1 \rightarrow 3}$, $\Gamma^{2 \rightarrow 1}$, $\Gamma^{4 \rightarrow 2}$ et $\Gamma^{3 \rightarrow 4}$ comme mortiers (en pointillés bleus). On vérifie qu'on a bien $\Gamma = \Gamma^{1 \rightarrow 3} \cup \Gamma^{2 \rightarrow 1} \cup \Gamma^{4 \rightarrow 2} \cup \Gamma^{3 \rightarrow 4}$. Les frontières non-mortiers (en vert) sont donc $\Gamma^{1 \rightarrow 2}$, $\Gamma^{3 \rightarrow 1}$, $\Gamma^{2 \rightarrow 4}$ et $\Gamma^{4 \rightarrow 3}$ (et on a de même $\Gamma = \Gamma^{1 \rightarrow 2} \cup \Gamma^{3 \rightarrow 1} \cup \Gamma^{2 \rightarrow 4} \cup \Gamma^{4 \rightarrow 3}$).

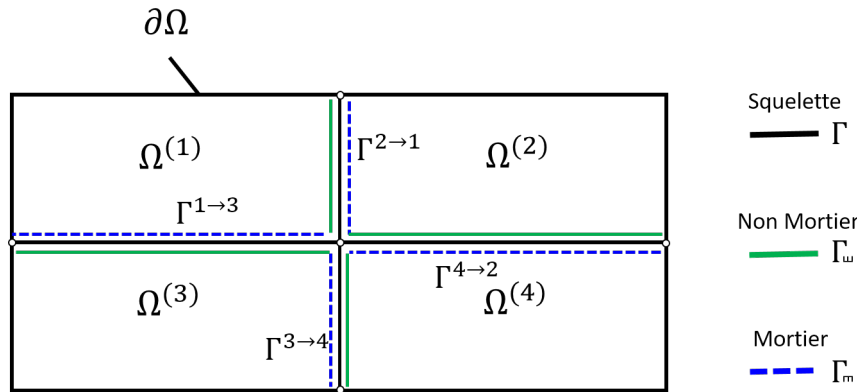


FIGURE 1.2.12 – Exemple de partition du squelette en mortiers (décomposition conforme).

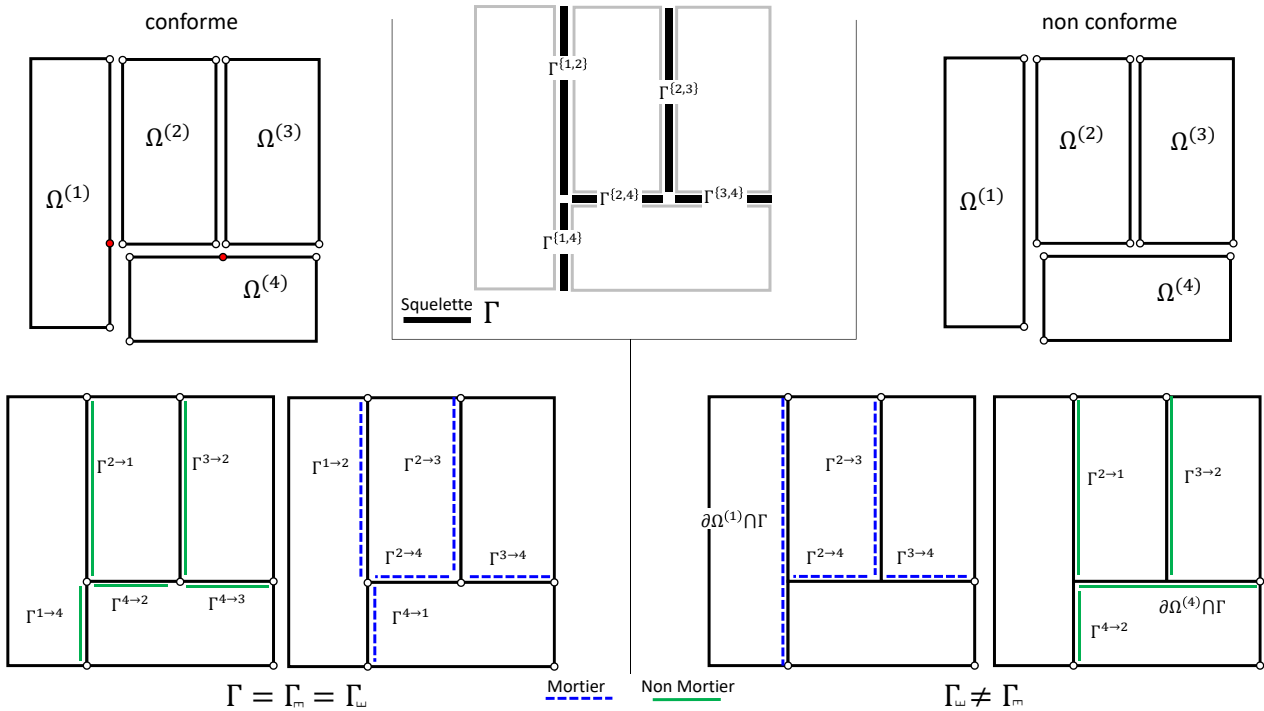


FIGURE 1.2.13 – Exemple de partition du squelette en mortiers (décomposition conforme à gauche et non conforme à droite).

Pour le cas d'une décomposition en domaines non-conforme (figure 1.2.13), il y a plus de contraintes lors du choix des frontières mortiers. On a choisi par exemple de prendre $\partial\Omega^{(1)} \cap \Gamma$ comme frontière mortier ; on doit donc obligatoirement prendre $\Gamma^{2 \rightarrow 1}$ et $\Gamma^{4 \rightarrow 1}$ comme frontières non-mortier. La difficulté du choix est due au fait que l'interface $\Gamma^{\{1,2\}}$ n'est pas une arête complète de $\partial\Omega^{(1)}$. L'autre cas possible est celui que l'on trouve aussi dans cette figure entre les sous-domaines $\Omega^{(2)}$, $\Omega^{(3)}$ et $\Omega^{(4)}$, les mortiers choisis sont les frontières $\Gamma^{2 \rightarrow 4}$ et $\Gamma^{3 \rightarrow 4}$,

l'arête entière du côté de $\Omega^{(4)}$ étant la frontière non-mortier correspondante. Un même côté mortier peut ainsi s'étendre sur les interfaces entre le sous-domaine mortier et plusieurs sous-domaines non-mortier. On pourrait penser contourner cette difficulté en coupant $\partial\Omega^{(1)} \cap \Gamma$ en deux parties, l'une en contact avec $\Omega^{(2)}$ et l'autre avec $\Omega^{(4)}$, comme illustré sur la partie gauche de la figure 1.2.13, et obtenir ainsi une décomposition conforme. Mais ceci est illusoire puisque le but est de pouvoir modéliser $\Omega^{(1)}$ sans préjuger s'il sera raccordé à un, deux ou plus de domaines par son interface.

2.5.2 Continuité de la solution à l'interface

L'intérêt de partitionner Γ ainsi est indissociable du deuxième point de la méthode : pour assurer la continuité de la solution du problème que l'on cherche à résoudre entre les interfaces des différents sous-domaines, on n'impose pas une continuité point à point sur chaque interface, mais une continuité faible. Pour décrire cette condition, on aura besoin de quelques notations.

On désigne par γ une frontière mortier et par δ une frontière non-mortier. Si on note Γ_{mortier} l'ensemble des frontières mortier du squelette de Ω et $\Gamma_{\text{non-mortier}}$ l'ensemble des non-mortier, et qu'on définit $N_{\text{mortier}} = \dim(\Gamma_{\text{mortier}})$ et $N_{\text{non-mortier}} = \dim(\Gamma_{\text{non-mortier}})$, on a $\bar{\Gamma} = \bigcup_{m=1}^{N_{\text{mortier}}} \bar{\gamma}_m = \bigcup_{m'=1}^{N_{\text{non-mortier}}} \bar{\delta}_{m'}$ ². On n'apportera pas ici plus de précision sur la façon d'indexer les mortiers et non-mortiers, puisqu'on ne va décrire la méthode que dans ses principes. Cependant, cette indexation reste cruciale au moment de l'implémentation dans le code de calcul qui sera utilisé.

Avec ces notations, on peut introduire la condition de continuité à l'interface proposée par la méthode originale, et qu'on appelle *condition mortier* ou *condition de joint*. Mais cette condition ne s'exprime que sur la solution cherchée sur un espace discrétisé. On reprend le problème couplé (1.2.7). On suppose donc que chaque $\Omega^{(k)}$ est discrétisé par un maillage éléments finis de paramètre h , et on note $V_h^{(k)} \subset H^1(\Omega^{(k)})$ l'espace de fonctions de dimension finie associé. On cherche alors la solution de (1.2.7) comme une fonction u_h telle que $u_h|_{\Omega^{(k)}} \in V_h^{(k)}$. La *condition de joint* qu'elle doit respecter sur Γ est :

pour chaque côté esclave $\delta \in \Gamma_{\text{non-mortier}}$:

$$\forall \phi_h \in M_h(\delta), \int_{\delta} (Tr_{\delta}(u_h|_{\Omega_{\delta}}) - Tr_{\gamma}(u_h|_{\Omega_{\gamma}})) \phi_h = 0, \quad (1.2.22)$$

où $M_h(\delta)$ est un espace de fonctions tests particulières définies sur δ , l'*espace des fonctions mortiers*. Tr_{δ} désigne l'opérateur trace de $H^1(\Omega_{\delta})$ vers $H^{1/2}(\delta)$, avec Ω_{δ} le sous-domaine associé au non-mortier δ (et similairement pour le mortier γ).

2. On rappelle que pour un espace métrisable X , \bar{X} représente l'adhérence de X , qui est l'ensemble des limites de suites convergentes à valeurs dans X .

Si la décomposition de Ω est conforme, alors γ est simplement le mortier “en face” du non-mortier δ . Dans le cas non-conforme en revanche, comme le montre l'exemple 1.2.13, il peut y avoir un mortier “plus long” ou plusieurs mortiers “plus courts” en face du non-mortier considéré. Cette notation est donc abusive, mais permet d'exprimer “simplement” la condition de joint.

L'intégrale présente dans (1.2.22) définit un produit scalaire sur $H^{1/2}(\delta)$. On en déduit que la condition (1.2.22) signifie que le saut de la fonction recherchée sur chaque bord non-joint doit être orthogonal à l'espace des fonctions tests défini sur δ . Elle revient à transformer la condition de raccordement (1.2.7c), où on impose $u^{(k)} = u^{(l)}$ sur $\Gamma^{\{k,l\}}$ en une condition faible. Cette formulation du raccordement permet ainsi d'avoir des maillages non-conformes aux interfaces, puisqu'on ne cherche plus à imposer la continuité de la solution point-à-point, mais seulement de manière intégrale.

2.5.3 Définition de $M_h(\delta)$ dans le cas 2D

Pour illustrer ce que peuvent être ces fonctions mortiers, en dimension 2, on peut choisir les fonctions de forme éléments finis ψ_i associées au maillage de δ , représentés en figure 1.2.14.

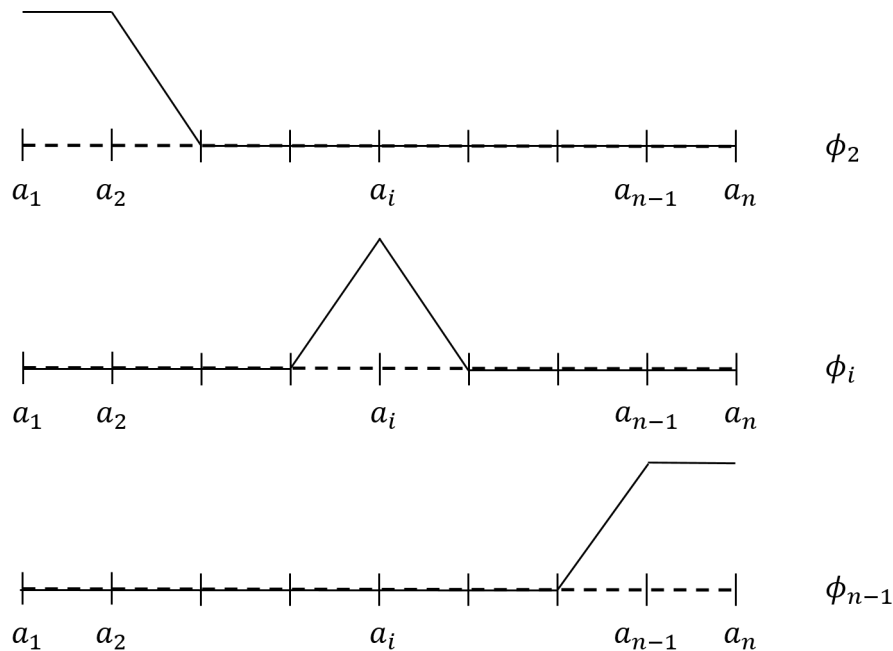


FIGURE 1.2.14 – Fonctions de base des multiplicateurs de Lagrange (2D).

Si le maillage de δ est composé de n points, les fonctions mortiers sont alors les $n - 2$ fonctions linéaires définies par

- si $j = 2$, $\phi_2 = \psi_1 + \psi_2$,
- si $j = 3, \dots, n - 2$, $\phi_i = \psi_i$,
- si $j = n - 1$, $\phi_{n-1} = \psi_{n-1} + \psi_n$.

2.5.4 Définition de $M_h(\delta)$ dans le cas 3D

Ben Belgacem [28] propose une définition de l'espace mortier en 3D lorsque le maillage à l'interface est constituée de triangles. Pour un maillage en quadrilatères, on pourra se référer à [44].

Pour l'implémentation, [45] donne une bonne description des fonctions tests sur l'interface 2D dans le cas 3D. Bien qu'exposée pour le cas de la magnétostatique, on pourra s'en inspirer pour notre application. Le choix de ces fonctions sera plus discuté dans la partie applications du manuscrit.

2.5.5 Autres méthodes de raccordement associées

Une autre méthode de raccordement, étudiée par Wohlmuth [40], est d'utiliser des bases duales pour les multiplicateurs de Lagrange ; dans la méthode des joints originale, si l'espace d'éléments finis défini sur l'interface est de dimension n , on ne peut utiliser comme mortier que des polynômes de degré au plus égal à $n - 1$. En utilisant la dualité, on peut définir un espace mortier *dual* de dimension n .

Pour les différents résultats numériques de la méthode des joints (stabilité, consistance convergence, estimation d'erreur *a priori*), on pourra à nouveau se référer à [25], [46] et [38].

Lacour et Maday [29] montrent que ce raccordement peut aussi se faire par une méthode FETI en utilisant des polynômes de Lagrange pour les ϕ_h dans (1.2.22), et ont mis en exergue la pertinence de choisir la méthode des joints plutôt que la méthode FETI. En effet, elle permet d'avoir de meilleures conditions de compatibilité entre les espaces discrets, de fournir une condition de convergence indépendante du paramètre de discrétisation h et d'avoir des matrices mieux conditionnées. Elle est aussi plus simple à mettre en œuvre car pour obtenir des résultats similaires à ceux obtenus avec les joints, il faut calculer des polynômes de Lagrange de degré assez élevé, au moins de degré 23 dans le cas étudié.

2.6 Méthodes de Galerkin discontinues (DG)

2.6.1 Origines des méthodes DG : les méthodes de pénalisation intérieure (IP)

2.6.1.1 Historique

Les méthodes de pénalisation intérieure découlent du problème étudié par Lions en 1968 [47], où il souhaite résoudre l'équation de Poisson (1.2.1), mais en remplaçant la condition de Dirichlet $u = 0$ par la condition approchée

$$\hat{u} + \mu^{-1} \frac{\partial \hat{u}}{\partial n} = 0, \quad (1.2.23)$$

avec μ un réel grand.

Lions a prouvé que pour tout $\mu \geq 0$, il existe une unique solution \hat{u} telle que \hat{u} converge vers u lorsque μ tend vers l'infini.

La formulation faible du problème généralisé est alors : trouver $\hat{u} \in H^1(\Omega)$ tel que

$$\int_{\Omega} \nabla \hat{u} \cdot \nabla v + \int_{\partial\Omega} \mu \hat{u} v = \int_{\Omega} f v, \quad \forall v \in H^1(\Omega). \quad (1.2.24)$$

On voit donc ici l'apparition de ce fameux paramètre de pénalisation, qui donnera son nom aux méthodes IP ("Interior Penalty"), et qui a été ajouté pour forcer le respect des conditions aux limites lorsque μ tend vers l'infini.

Cette approche a été réutilisée par Nitsche [48] dans le cadre d'une discrétisation du domaine Ω par une méthode éléments finis, qui a prouvé que si l'on prend $\mu = \eta/h$, où h est la taille de l'élément et η une constante suffisamment grande, alors la solution discrète converge vers la solution exacte avec un ordre optimal dans L^2 et H^1 .

2.6.1.2 Développement des méthodes IP

Les méthodes IP pour les problèmes elliptiques découlent de ces études, avec notamment la généralisation de la méthode de Nitsche aux problèmes elliptiques du second ordre [49].

Cette méthode a ensuite été étudiée dans la thèse d'Arnold en 1979, résumée dans l'article [50], pour les équations elliptiques linéaires et non-linéaires ainsi que pour les équations paraboliques.

Une autre approche très intéressante est celle de Becker *et al.* [51] où est développée la méthode de Nitsche dans le cadre des maillages non-conformes, répondant ainsi directement à notre problématique.

Pour l'illustrer, on se sert une fois de plus de l'exemple de partition d'un domaine Ω présentée en figure 1.2.5.

La méthode de Becker s'exprimant sur les discrétisations des sous-domaines, on suppose que $\Omega^{(1)}$ et $\Omega^{(2)}$ sont discrétisés selon des découpages éléments finis, notés respectivement $\mathcal{T}_h^{(1)}$ et $\mathcal{T}_h^{(2)}$. On note $h_{\Gamma^{\{i,j\}}}$ la taille du plus petit élément de $\Gamma^{\{i,j\}}$.

La méthode de Nitsche avec pénalisation de Becker est alors³ : trouver $u_h \in H^1(\Omega)$ tel que

$$\sum_{i=1}^2 \left[\int_{\Omega^{(i)}} \nabla u_i \cdot \nabla v_i + \gamma h_{\Gamma^{\{i,j\}}}^{-1} \int_{\Gamma^{\{i,j\}}} \llbracket u \rrbracket \llbracket v \rrbracket \right] - \int_{\Gamma^{\{i,j\}}} \frac{\partial u}{\partial n} \llbracket v \rrbracket - \int_{\Gamma^{\{i,j\}}} \frac{\partial v}{\partial n} \llbracket u \rrbracket = \sum_{i=1}^2 \int_{\Omega^{(i)}} f(v_i), \quad (1.2.25)$$

$\forall v \in H^1(\Omega)$, $i \neq j$, avec $\gamma > 0$ suffisamment grand.

On se référera à l'article de Becker *et al.* [51] pour les résultats de stabilité et de convergence de u_h .

2.6.1.3 Les méthodes IP aujourd'hui

Ces méthodes ont connu certes un développement important à la fin des années 1970, grâce au développement des méthodes DG et à leur application ensuite aux problèmes elliptiques, mais la tendance s'est inversée dès le début des années 1980. Cela peut être dû au manque de résultats convaincants face aux méthodes éléments finis conformes classiques, mais aussi au développement de nouvelles méthodes, comme la méthode des joints [22, 46].

2.6.2 Développement des méthodes DG

Inspirées des méthodes IP, les méthodes de Galerkin Discontinues (dites DG methods pour *Discontinuous Galerkin methods*) ont été introduites pour la première fois en 1973 par Reed et Hill [52] pour la résolution d'équations hyperboliques dans le cas du transport du neutron, alors que les méthodes IP étaient développées pour la résolution d'équations elliptiques et paraboliques.

Elles ont pour particularité, comme pour les méthodes IP, de ne pas imposer la continuité de la solution numérique sur les interfaces, ce qui permet de pouvoir résoudre les équations localement, et donc de paralléliser le calcul (ce qui constitue le plus grand intérêt), mais aussi en particulier de recourir à différents maillages de part et d'autre de ces interfaces, qui est le point qui nous intéresse.

3. On désigne le saut d'une fonction u à l'interface $\Gamma^{\{i,j\}}$ par $\llbracket u \rrbracket = u^{(j)}|_{\Gamma^{\{i,j\}}} - u^{(i)}|_{\Gamma^{\{i,j\}}}$

Les différentes méthodes DG pour la résolution d'équations hyperboliques non-linéaires vont en fait être définies par le choix des flux numériques aux interfaces. Arnold *et al.* [53, 54] en dressent une liste particulièrement fournie, en essayant d'unifier méthodes IP et méthodes DG. On pourra notamment citer, pour ce qui est des méthodes DG appliquées aux équations elliptiques (méthodes IP), Douglas [55], Rivière [56] (méthode NIPG), Babuška-Zláman [57] et Brezzi *et al.* [58]. Pour ce qui est des autres méthodes DG appliquées aux équations hyperboliques, on pourra citer Bassi-Rebay [59], Brezzi *et al.* [58], Cockburn [60] (méthode LDG) et Baumann-Oden [61].

On pourra aussi se référer à ce même article [53] pour les formes primales associées à ces choix et aux différents résultats de stabilité, convergence...

2.7 Choix de la méthode de raccordement

À la suite de cet état de l'art, il convient de choisir la méthode de sous-structuration que l'on va utiliser pour créer une bibliothèque de composants réduits. Les méthodes de Schwartz avec recouvrement sont inapplicables, puisqu'elles ne correspondent pas du tout à notre problématique. De même, la méthode de Schur ne convient pas non plus, car, comme on l'a expliqué plus tôt, ces méthodes reposent sur un découpage *a posteriori* du domaine d'étude, pour résoudre le problème localement avant d'en déduire la solution globale. Or, notre problème consiste à coupler plusieurs systèmes déjà définis, avec leurs paramètres distincts.

On va donc choisir une méthode de sous-structuration avec couplage aux interfaces. Toutefois, que l'on choisisse une méthode duale ou une méthode primale pour résoudre le système sous-structuré, celui-ci fait apparaître une matrice de couplage \mathbf{J} . Si celle-ci est simple à déterminer dans le cas de maillages conformes aux interfaces, la nécessité d'avoir des composants réduits réutilisables d'après le cahier des charges établi impose de travailler avec des maillages non-conformes. Le choix se réduit donc aux méthodes développées spécifiquement pour résoudre des problèmes de ce type : la méthode des joints ou les méthodes de Galerkin Discontinues.

On choisit d'utiliser la méthode des joints, par les bons résultats qu'elle permet d'obtenir [33, 39, 46], mais aussi, d'un point de vue plus pragmatique, car elle est la méthode la plus utilisée aujourd'hui, et celle qui est le plus référencée, notamment pour les études de contacts mécaniques qui doivent pouvoir utiliser des maillages évolutifs. Il y a donc un plus grand suivi par rapport à d'autres méthodes, ce qui est préférable puisqu'on dispose ainsi de plus de documentation et de supports pour l'implémentation informatique.

D'autres développements de la méthode des joints peuvent aussi être intéressants en terme de future optimisation de code, notamment l'implémentation parallèle ([24, 37, 62]).

Chapitre 3

État de l’art des méthodes de réduction

3.1 De la nécessité de réduire

Comme on l’a dit en introduction, de par la problématique de ce travail, *i.e.* la simulation thermique de cartes électroniques, la réduction des modèles est indispensable.

On peut envisager de réduire ces modèles de plusieurs façons. La première et la plus simple d’entre-elles est de simplifier la géométrie ou de réduire le nombre d’inconnues (ou d’états internes) en relâchant le maillage. Mais on perd alors finalement en résolution et en fidélité géométrique.

On peut aussi utiliser des “boîtes noires”, pour modéliser le comportement thermique des composants, comme des fonctions de transfert entrée/sortie ou des réseaux de neurones. Dans ce cas, on perd cependant toute l’information sur l’état interne du système

La méthode de réduction la plus utilisée actuellement dans le domaine de l’électronique est la méthode nodale utilisant des réseaux RC, que l’on va donc détailler juste après cette introduction.

L’approche par décomposition sur des bases réduites (POD, SVD, méthodes spectrales) présente l’avantage de ne pas perdre de résolution géométrique par rapport à une méthode type éléments ou volumes finis. Ces techniques sont au cœur de ce travail et on les présentera plus longuement dans la suite de ce manuscrit. On se focalisera en particulier sur la méthode modale, qui fait partie des méthodes spectrales.

Dans le but de comparer ces méthodes entre elles, on va les exprimer dans un formalisme associé à une discrétisation spatiale d’un domaine $\Omega \subset \mathbb{R}^d$. Cette discrétisation peut être faite à partir de différences finies, de volumes finis, d’éléments finis, de modes, ... On peut alors écrire l’équation de la chaleur sous la forme matricielle

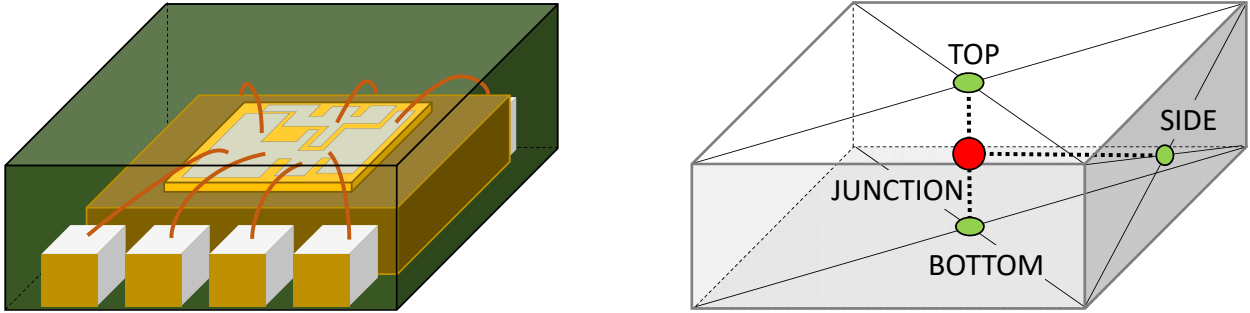


FIGURE 1.3.1 – Exemple de réseau nodal pour un composant dans un boîtier DFN8.

$$\mathbf{C}\dot{T} + (\mathbf{K} + \mathbf{H})T = \mathbf{B}U, \quad (1.3.1)$$

avec \mathbf{C} la matrice de capacité, \mathbf{K} la matrice de conductivité, \mathbf{H} la matrice d'échanges avec l'extérieur, T le vecteur des valeurs de la température aux points de la discrétisation choisie et $\mathbf{B}U$ le vecteur des sollicitations. Certaines méthodes peuvent aussi s'exprimer dans le cadre de l'analyse fonctionnelle, ce qui peut mettre en exergue des propriétés intéressantes qui vont permettre une simplification du système à résoudre, mais pas toutes.

3.2 Modèle nodal

La réduction nodale, ou la réduction par réseaux RC, est la méthode la plus fréquemment utilisée pour concevoir des modèles réduits de composants électroniques. On prend l'exemple d'un boîtier homogène dont on souhaite estimer la température interne (T_J). Le boîtier échange de la chaleur via ses surfaces avec le milieu extérieur à la température T_E , et l'intérieur du domaine est le siège d'un dégagement de chaleur Φ . Un modèle élémentaire peut être établi en considérant des zones isothermes, matérialisées par des nœuds. Ici, nous considérons un nœud interne pour le volume, et un nœud sur chaque face (par concision nous considérerons que la température est la même sur toutes les faces latérales), comme indiqué sur le schéma de droite de la figure 1.3.1. Le principe des réseaux nodaux repose sur l'analogie thermo-électrique : le flux thermique Φ échangé entre deux nœuds aux températures T et $T + \Delta T$ s'écrit $\Phi = \frac{\Delta T}{R}$, où R est la résistance thermique. Les flux thermiques entre les différents nœuds peuvent être modélisés par le réseau présenté sur la figure 1.3.2a. Les valeurs des températures s'obtiennent ensuite grâce à un bilan de flux : à l'équilibre thermique, la somme des flux thermiques reçus par chaque nœud doit être nulle. On obtient alors le système

matriciel suivant :

$$\begin{bmatrix} \frac{1}{R_{JT}} + \frac{1}{R_{JB}} + \frac{1}{R_{JS}} & -\frac{1}{R_{JT}} & -\frac{1}{R_{JB}} & -\frac{1}{R_{JS}} \\ -\frac{1}{R_{JT}} & \frac{1}{R_{JT}} + \frac{1}{R_{TE}} & 0 & 0 \\ -\frac{1}{R_{JB}} & 0 & \frac{1}{R_{JB}} + \frac{1}{R_{BE}} & 0 \\ -\frac{1}{R_{JS}} & 0 & 0 & \frac{1}{R_{JS}} + \frac{1}{R_{SE}} \end{bmatrix} \begin{bmatrix} T_J \\ T_T \\ T_B \\ T_S \end{bmatrix} = \begin{bmatrix} \Phi \\ \frac{T_E}{R_{TE}} \\ \frac{T_E}{R_{BE}} \\ \frac{T_E}{R_{SE}} \end{bmatrix}. \quad (1.3.2)$$

On retrouve la version stationnaire de la formulation matricielle (1.3.1), mais avec la matrice $(\mathbf{K} + \mathbf{H})$ remplacée par la matrice des résistances. Ce réseau peut être rendu plus réaliste si l'on considère qu'il y a de la conduction entre les surfaces latérales et les surfaces haute et basse. On pourrait même imaginer relier par une résistance les températures des parois haute et basse sans passer par le milieu, même si cette liaison n'est pas physique (voir figure 1.3.2c, mais sans capacité sur T_J). Dans ce cas, la détermination de la résistance thermique n'est plus évidente.

Lorsque l'on souhaite modéliser les phénomènes transitoires, une capacité est introduite (voir figure 1.3.2c) : la variation d'énergie du domaine ($\rho C_p V \frac{\partial T_J}{\partial t}$) est égale à la somme des flux thermiques reçus. On retrouve ainsi la formulation matricielle (1.3.1) :

$$\begin{bmatrix} \rho C_p V & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{T}_J \\ \dot{T}_T \\ \dot{T}_B \\ \dot{T}_S \end{bmatrix} + \begin{bmatrix} \frac{1}{R_{JT}} + \frac{1}{R_{JB}} + \frac{1}{R_{JS}} & -\frac{1}{R_{JT}} & -\frac{1}{R_{JB}} & -\frac{1}{R_{JS}} \\ -\frac{1}{R_{JT}} & \frac{1}{R_{JT}} + \frac{1}{R_{TE}} & 0 & 0 \\ -\frac{1}{R_{JB}} & 0 & \frac{1}{R_{JB}} + \frac{1}{R_{BE}} & 0 \\ -\frac{1}{R_{JS}} & 0 & 0 & \frac{1}{R_{JS}} + \frac{1}{R_{SE}} \end{bmatrix} \begin{bmatrix} T_J \\ T_T \\ T_B \\ T_S \end{bmatrix} = \begin{bmatrix} \Phi \\ \frac{T_E}{R_{TE}} \\ \frac{T_E}{R_{BE}} \\ \frac{T_E}{R_{SE}} \end{bmatrix}. \quad (1.3.3)$$

Pour des géométries simples (comme des barres, des murs, ...) dont les propriétés physiques sont connues, les valeurs des résistances peuvent être déterminées analytiquement ($R = \frac{L}{SK}$ ou L est la distance entre les deux nœuds, K la conductivité thermique du matériau et S la section de passage du flux). Lorsque le domaine à modéliser n'est plus homogène, ou que sa géométrie devient complexe, comme un composant électronique, il est illusoire de vouloir déterminer analytiquement un réseau physique. Il faut donc caractériser et la forme du réseau et les valeurs des résistances par d'autres moyens : cette opération peut passer par des mesures expérimentales, comme pour les réseaux de type Cauver-Foster [63] ou bien via une simulation numérique. C'est le principe de la méthodologie DELPHI, projet européen regroupant six partenaires entre 1993 et 1996 visant à créer une méthodologie de création de modèles réduits,

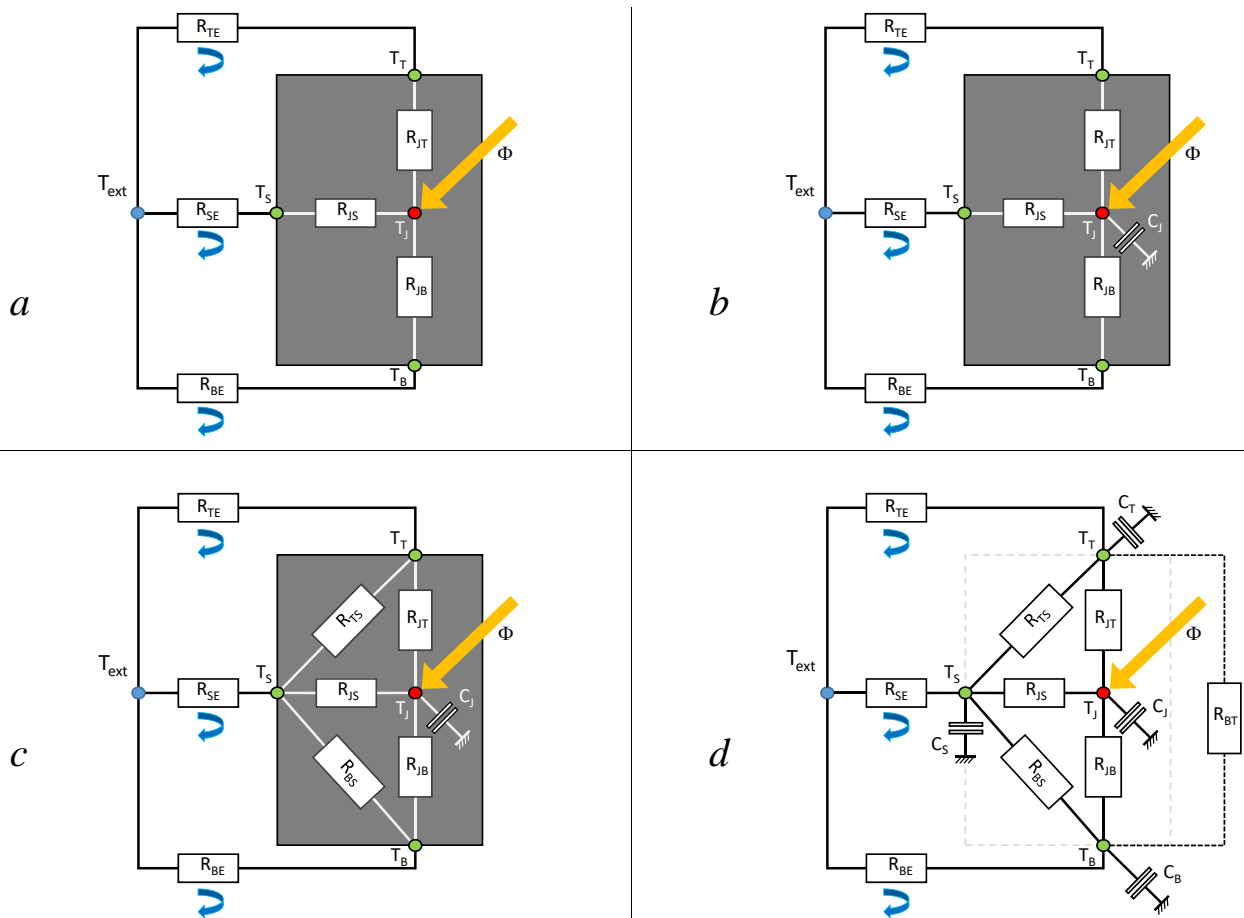


FIGURE 1.3.2 – Exemples de réseaux nodaux pour un même composant : (a) modèle stationnaire (sans inertie) issu d’une analyse physique par volumes finis ; (b) modèle dynamique (l’inertie est représentée par une capacité sur le nœud interne T_J) issu d’une analyse physique par volumes finis ; (c) modèle dynamique qui prend en compte la conduction entre les faces : la détermination des résistances thermique entre les faces est délicate car elle fait intervenir des phénomènes tridimensionnels ; (d) modèle dynamique dont les résistances thermiques et les capacités doivent être identifiées (par exemple à l’aide d’un algorithme génétique) : on s’éloigne d’une interprétation physique directe des résistances et des capacités, en particulier la résistance qui relie les nœuds “top” (T_T) et “bottom” (T_B) directement sans passer par le nœud central ne respecte pas les lois du transfert de chaleur par conduction thermique (diffusion). De même une surface n’a théoriquement pas de capacité thermique.

appelés “CTM” pour “Compact Thermal Model”, à partir de modèles détaillés, obtenus via des simulations de type éléments finis, volumes finis ou différences finies. Les modèles réduits transitoires sont appelés “DCTM”, ou “Dynamic Compact Thermal Model”.

Cette méthodologie a été reprise en 2008 par le JEDEC, organisme américain de standardisation des semi-conducteurs, pour fournir un guide de génération de modèles compacts [64]. Ce guide pose ainsi quatre critères qu'un modèle compact doit respecter pour être pertinent :

- la complexité doit être limitée : un réseau ne doit pas dépasser le nombre maximal de 10 nœuds ;
- le modèle doit être indépendant des conditions aux limites en garantissant une bonne précision de la température de jonction¹ pour 38 scénarios définis ;
- le partage des modèles doit pouvoir s'effectuer via un format de fichier neutre ;
- le modèle doit être documenté et appartenir au domaine public.

Le nombre de scénarios utilisés pour bien couvrir l'ensemble des conditions de 3ème espèce que le composant est susceptible de subir est variable selon les études, comme en attestent les 49 scénarios retenus par Thales à l'issue du programme DELPHI² pour réaliser leurs modèles compacts.

Des simulations reproduisant ces scénarios sont donc effectuées sur les composants à réduire, à partir de modèles détaillés. Les résistances (et capacités) sont identifiées pour que le CTM (DCTM) reproduise le mieux possible le comportement de ce composant lors de ces simulations, via l'optimisation d'une fonction objectif définie.

Dans la méthode développée par Thales, la forme du réseau est simple : tous les nœuds sont supposés être couplés entre eux. Pour identifier les résistances du réseau, une technique, introduite par [65] (mais seulement pour les composants mono-puce) et développée par Dia [66] est d'utiliser des algorithmes génétiques. Cette technique permet d'obtenir un réseau performant très rapidement. Dia a ainsi développé et amélioré cette technique, qui est utilisée au sein de Thales dans son processus industriel de réduction de composants. Il l'a notamment étendue aux composants multi-couches, tout en développant la possibilité de diviser les faces inférieures et supérieures du composant en deux surfaces, une interne ("inner") et une externe ("outer", voir figure 1.3.3).

Les CTM et DTCM obtenus par Dia présentent ainsi toujours moins de 10% de divergence au niveau des températures de jonction et des flux thermiques par rapport aux modèles détaillés, que ce soit en régime permanent ou en régime transitoire sur des scénarios d'activation/désactivation des puces, tout en permettant de gros gains de temps de calcul, puisque chaque composant est ainsi modélisé par une dizaine de nœuds.

Toutefois, le nombre de points d'observations proposés par cette méthode est très limité. Pour l'ingénieur chargé de l'étude thermique d'un équipement ces points névralgiques peuvent s'avérer tout à fait suffisants. Mais il peut arriver que la connaissance de l'intégralité du champ

1. La température de jonction est la température mesurée dans la zone active du composant (le silicium).

2. Thomson-CSF, une des sociétés fondatrice de Thales, était un des partenaires du programme DELPHI.

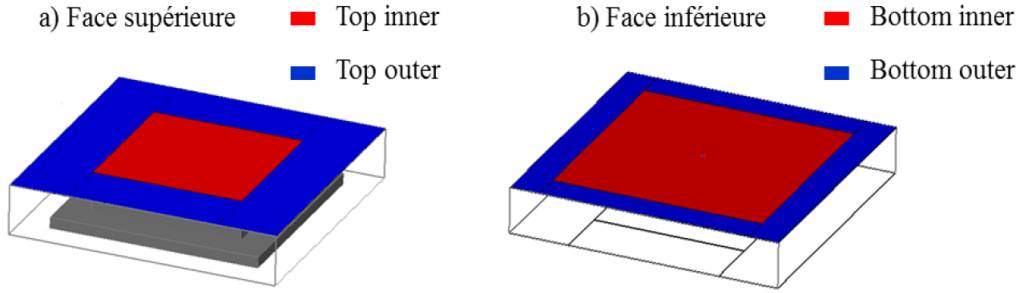


FIGURE 1.3.3 – Surfaces “inner” et “outer” des faces inférieure et supérieure.

de température soit nécessaire, comme par exemple pour déterminer des déformations thermomécaniques ou pour observer la température critique d'un composant dont la localisation n'est pas connue à l'avance.

3.3 Réduction modale

3.3.1 Principe

En 1822, Joseph Fourier publie *Théorie analytique de la chaleur* [67], où il propose une méthode de résolution de l'équation de la chaleur, en appliquant la méthode de séparation des variables de temps et d'espace pour exprimer la température sous la forme $T(M, t) = f(t)g(M)$, et en résolvant les équations vérifiées par f et g pour obtenir une décomposition dite en séries de Fourier de T .

On peut considérer la méthode modale comme une généralisation de la méthode de séparation des variables, où on cherche à obtenir une base de fonctions sur laquelle décomposer la température, *i.e.* on cherche T sous la forme

$$T(M, t) = \sum_{i=1}^{+\infty} x_i(t) V_i(M). \quad (1.3.4)$$

Si on connaît la base $\{V_i\}_{i \in \mathbb{N}^*}$, les inconnues deviennent alors les états d'excitations des modes $\{x_i\}_{i \in \mathbb{N}^*}$.

La réduction modale consiste ainsi à obtenir, non pas la base complète $\{V_i\}_{1 \leq i \leq N}$, mais une base réduite $\{\tilde{V}_i\}_{1 \leq i \leq \tilde{N}}$, avec $\tilde{N} \ll N$. On cherche donc une valeur approchée de la température

$$\tilde{T}(M, t) = \sum_{i=1}^{\tilde{N}} \tilde{x}_i(t) \tilde{V}_i(M), \quad (1.3.5)$$

et les inconnues sont les \tilde{N} états \tilde{x}_i .

En pratique, dans le cadre de simulations éléments finis, où on a une discrétisation \mathcal{T}_h du domaine d'étude, cette somme est finie et on cherche en réalité la température T_h associée à cette discrétisation.

Sous forme matricielle, si T est le vecteur des températures aux points de la discrétisation choisie, on l'exprime sous la forme

$$T = \tilde{\mathbf{V}}\tilde{X}, \quad (1.3.6)$$

avec $\tilde{\mathbf{V}}$ la matrice des modes de la base réduite $\{\tilde{V}_i\}_{1 \leq i \leq \tilde{N}}$.

En remplaçant T dans (1.3.1) par cette expression, on a

$$\mathbf{C}\tilde{\mathbf{V}}\dot{\tilde{X}} + (\mathbf{K} + \mathbf{H})\tilde{\mathbf{V}}\tilde{X} = \mathbf{B}U. \quad (1.3.7)$$

En remultipliant cette équation par ${}^t\tilde{\mathbf{V}}$, on obtient la forme réduite de (1.3.1) :

$${}^t\tilde{\mathbf{V}}\mathbf{C}\tilde{\mathbf{V}}\dot{\tilde{X}} + {}^t\tilde{\mathbf{V}}(\mathbf{K} + \mathbf{H})\tilde{\mathbf{V}}\tilde{X} = {}^t\tilde{\mathbf{V}}\mathbf{B}U. \quad (1.3.8)$$

Si on s'intéresse seulement à un certain nombre d'observables (valeurs moyennes sur une ou plusieurs surfaces, valeurs en des points particuliers), on définit une matrice d'observables $Y = \mathbf{G}X$. En posant les matrices réduites $\mathbf{C}_r = {}^t\tilde{\mathbf{V}}\mathbf{C}\tilde{\mathbf{V}}$, $\mathbf{A}_r = {}^t\tilde{\mathbf{V}}(\mathbf{K} + \mathbf{H})\tilde{\mathbf{V}}$ et $\mathbf{B}_r = {}^t\tilde{\mathbf{V}}\mathbf{B}$, le système modal réduit s'écrit :

$$\begin{cases} \mathbf{C}_r\dot{\tilde{X}} + \mathbf{A}_r\tilde{X} = \mathbf{B}_rU, \\ Y = \mathbf{G}X. \end{cases} \quad (1.3.9)$$

On a donc réduit le problème, en passant de N à \tilde{N} inconnues ($\dim(\mathbf{C}_r) = \dim(\mathbf{A}_r) = (\tilde{N} \times \tilde{N})$). Bien sûr, cette réduction implique qu'on a une perte en précision, il conviendra donc de la quantifier.

Pour obtenir ces bases modales, on dispose de plusieurs méthodes.

3.3.2 SVD, POD et PGD

Si on considère \mathbf{A} une réalisation d'une simulation thermique, la SVD ou "Singular Value Decomposition", ou encore décomposition en valeurs singulières est une méthode de décomposition où on factorise $\mathbf{A} \in M_{n,m}(\mathbb{R})$ sous la forme $\mathbf{A} = U\Sigma V^*$, avec $U \in M_{n,n}(\mathbb{R})$ et $V \in M_{m,m}(\mathbb{R})$ des matrices unitaires et Σ une matrice diagonale dont les coefficients sont les valeurs singulières de \mathbf{A} . Pour rappel, les valeurs singulières sont les λ_i tels que, si on note v_i les vecteurs propres de ${}^t\mathbf{A}\mathbf{A}$, on a ${}^t\mathbf{A}\mathbf{A}v_i = \lambda_i^2 v_i$.

3.3.2.1 POD

La POD, ou “Proper Orthogonal Decomposition”, est une application directe de la SVD [68], d’abord développée en mécanique des fluides puis dans les années 2000 en thermique [69, 70, 71].

Le principe de la POD est de chercher la base de modes V_i pour qu’elle approche au mieux en moyenne une réalisation $\mathbf{T}(M, t)$ du problème (\mathbf{T} étant définie comme dans le paragraphe précédent). Ces modes sont obtenus en résolvant l’équation aux valeurs propres

$$\langle (V_i | \mathbf{T}) \mathbf{T} \rangle = \lambda_i V_i, \quad (1.3.10)$$

où $(\cdot | \cdot)$ est le produit scalaire de l’espace de travail (généralement le produit scalaire dans $L^2(\Omega)$) et $\langle \cdot \rangle$ désigne ce qu’on appelle l’opérateur moyenne. Du choix de ce dernier va dépendre la méthode de POD utilisée.

Les deux grandes méthodes sont la méthode “classique”, introduite par Lumley [72] et la méthode “snapshot”. Dans la méthode classique, l’opérateur moyenne est défini comme une moyenne temporelle :

$$\langle u \rangle = \frac{1}{\tau} \int_0^\tau u dt. \quad (1.3.11)$$

La réalisation \mathbf{T} peut alors être décomposée sur ces modes :

$$\mathbf{T}(M, t) = \sum_{i \geq 0} a_i(t) V_i(M). \quad (1.3.12)$$

La méthode snapshot, introduite par Sirovich [73], considère que seulement K réalisations (que l’on appelle “snapshots”) $\{\mathbf{T}(M, t_i)\}_{1 \leq i \leq K}$ suffisent pour décrire le système étudié.

L’opérateur moyenne est cette fois-ci une moyenne spatiale :

$$\langle u \rangle = \int_{\Omega} u dv. \quad (1.3.13)$$

La POD snapshot consiste alors à chercher les coefficients $a(t_i)$, $1 \leq i \leq K$ tels que

$$V(M) = \sum_{i=1}^K a(t_i) \mathbf{T}(M, t_i). \quad (1.3.14)$$

On se ramène alors à un nouveau problème au valeurs propres, de dimension K à résoudre :

$$\frac{1}{K} \sum_{k=1}^K \mathcal{C}_{ik} a_k = \lambda a_i \quad \forall 1 \leq i \leq K, \quad (1.3.15)$$

avec $\mathcal{C}_{ik} = (\mathbf{T}(t_i) | \mathbf{T}(t_k))$.

3.3.2.2 PGD

La PGD, ou “Proper Generalized Decomposition”, est une généralisation à l’extrême du concept de séparation des variables, où on cherche T sous la forme

$$T(x, t, p^1, \dots, p^P) = \sum_i x_i(t) V_i(x) \prod_{j=1}^P F_i^j(p^j), \quad (1.3.16)$$

où les paramètres p^j peuvent être la capacité thermique massique, ou la conductivité. Les fonctions F_i^j sont calculées en enrichissant la base à chaque itération supplémentaire. Le calcul de chaque nouveau mode est effectué en résolvant un problème non linéaire issu de l’équation de la chaleur [74, 75].

Il est intéressant de noter que la POD et la PGD sont développées dès le départ pour des systèmes non-linéaires.

3.3.3 Modes de Krylov

Les modes de Krylov sont une autre manière de calculer des modes pour des systèmes paramétriques. Par exemple, si on considère le problème de réponse en fréquence linéaire [76]

$$\begin{aligned} (s\mathbf{E} - \mathbf{A})x &= bu, \\ y &= \mathbf{C}^T x, \end{aligned} \quad (1.3.17)$$

les modes de Krylov sont cherchés comme appartenant aux espaces de Krylov

$$\begin{aligned} K_q(\mathbf{A}^{-1}\mathbf{E}, \mathbf{A}^{-1}b) &= Vect(\mathbf{A}^{-1}b, \mathbf{A}^{-1}\mathbf{E}\mathbf{A}^{-1}b, (\mathbf{A}^{-1}\mathbf{E})^2\mathbf{A}^{-1}b, \dots, (\mathbf{A}^{-1}\mathbf{E})^{q-1}\mathbf{A}^{-1}b), \\ K_q(\mathbf{E}^T\mathbf{A}^{-T}, \mathbf{C}) &= Vect(\mathbf{C}, \mathbf{E}^T\mathbf{A}^{-T}\mathbf{C}, (\mathbf{E}^T\mathbf{A}^{-T})^2\mathbf{C}, \dots, (\mathbf{E}^T\mathbf{A}^{-T})^{q-1}\mathbf{C}). \end{aligned} \quad (1.3.18)$$

Le modèle réduit s’écrit alors

$$\begin{aligned} \mathbf{W}^T\mathbf{E}(s)\mathbf{V}\hat{x} &= \mathbf{W}^Tbu, \\ y &= \mathbf{C}^T\mathbf{V}\hat{x}, \end{aligned} \quad (1.3.19)$$

avec $\mathbf{W} \in K_q(\mathbf{A}^{-1}\mathbf{E}, \mathbf{A}^{-1}b)$ et $\mathbf{V} \in K_q(\mathbf{E}^T\mathbf{A}^{-T}, \mathbf{C})$.

Ces modes ont été utilisés avec succès par Codecasa *et al.* [77] pour créer des modèles réduits de composants électroniques. Les modèles réduits créés avec ces méthodes ont été utilisés pour simuler des scénarios de référence du standard JEDEC afin de créer des modèles nodaux.

Une autre application de la SVD qui en découle est la méthode hybride de SVD-Krylov [68, 78, 79], où on complète la base réduite de décomposition obtenue par la SVD avec des modes calculés par itérations de Krylov.

3.3.4 Méthode modale

Le principe de la méthode modale est de calculer la base de décomposition de T (1.3.6) en résolvant un problème aux valeurs propres adapté au problème physique à résoudre, ici l'équation de la chaleur.

Pour décomposer la température, on peut utiliser la base de Fourier, appelée aussi base "classique", puisqu'elle est caractérisée par des conditions aux limites similaires à celles du problème physique décrit en (2.4.1).

La base de Fourier est ainsi définie par le problème aux valeurs propres suivant :

$$\begin{cases} \Omega, & K_0 \Delta V_i = C_0 \lambda_i V_i, \\ \partial\Omega, & K_0 \nabla V_i \cdot \underline{n} = -h_0 V_i, \end{cases} \quad (1.3.20)$$

La température s'écrit alors $T = T_d + T_g = \sum_i x_i V_i + T_g$, où T_g est le champ glissant défini par

$$\begin{cases} \Omega, & K_0 \Delta T_g = \varpi, \\ \partial\Omega, & K_0 \nabla T_g \cdot \underline{n} = h_0 T_{ext} + \varphi. \end{cases} \quad (1.3.21)$$

Mais cette méthode est inadaptée à notre problématique ; en effet, les coefficients dans les conditions aux limites du problème aux valeurs propres (1.3.20) doivent être identiques à ceux du problème thermique (2.4.1). Si les coefficients du problème thermique viennent à évoluer (si h et K diffèrent de h_0 et K_0), alors les modes de Fourier ne forment plus une base sur laquelle on pourra décomposer T , ce qui limite son utilisation à des problèmes à conditions limites linéaires et stationnaires. Cette condition est totalement incompatible avec notre problématique, puisqu'on veut créer des modèles réduits pouvant être utilisés dans des conditions aux limites variables.

L'introduction d'une nouvelle base, dite de Branches a permis d'étendre la méthode à l'étude problème instationnaires, non-linéaires [80, 81] et avec ou sans transport [82, 83].

3.3.5 Méthode MIM

Le principe de la MIM est différent, puisque les modes ne sont pas calculés à partir d'un problème aux valeurs propres. Le modèle modal est directement identifié, c'est-à-dire les matrices \mathbf{A}_r , \mathbf{B}_r , \mathbf{C}_r et \mathbf{G} de (1.3.9) [84, 85]. Cette identification se fait à l'aide de données expérimentales ou numériques. L'intérêt principal de la MIM est qu'elle ne nécessite pas la connaissance d'un modèle (conditions aux limites, propriétés physiques, ...).

Pour le contrôle d'un modèle réel, c'est une méthode particulièrement efficace [86, 87, 88] ; par contre, si les conditions aux limites du système changent, le modèle ne pourra pas s'adapter.

3.4 Choix de la méthode de réduction

La réduction nodale avec des modèles RC permet d'avoir des modèles réduits très sobres en temps de calcul, puisque le nombre d'inconnues est limité à une dizaine d'observables. La résolution du système matriciel associé est donc très rapide. Le fait qu'ils soient construits à partir de dizaines de simulations de références permet de les rendre, dans une certaine mesure, indépendants des conditions aux limites. En revanche, on a uniquement une information sur les variables d'observation choisies au préalable (par définition en nombre limité), ce qui ne répond pas au cahier des charges. De plus, si ces méthodes ont prouvé leur efficacité pour des simulations stationnaires, les écarts obtenus pour des simulations dynamiques sont d'un ordre 10 fois plus élevé.

Les méthodes modales permettent elles aussi d'avoir des modèles réduits, en réduisant le nombre de modes utilisés pour décomposer T , tout en gardant une grande finesse géométrique, puisqu'on peut récupérer le champ de température sur l'intégralité du maillage utilisé pour la discrétisation, au contraire des méthodes nodales. Les méthodes POD ou PGD ainsi que l'utilisation de modes de Krylov ou encore de Branches et de Dirichlet-Steklov répondent ainsi au cahier des charges vis-à-vis des trois premiers points : la sobriété, la finesse géométrique et l'indépendance aux conditions aux limites possibles.

Pour le quatrième point, à savoir que les modèles réduits soient réutilisables, on a besoin de modes permettant de vérifier les conditions de compatibilité aux interfaces (*i.e.* la continuité des températures et des flux). Le travail de thèse étant limité dans le temps, il n'est pas possible de tester toutes les méthodes pour s'assurer qu'elles permettent d'assurer cette continuité. On décide donc de retenir une méthode dont on sait, via des travaux précédents celui-ci, qu'elle permet le couplage de modèles réduits puisque conçue dans ce but : la méthode modale.

Initialement, une base de Branches était utilisée. Nous montrerons qu'il est préférable d'utiliser une base de Dirichlet-Steklov qui généralise l'idée de la base de Branches. À l'aide de la base de Dirichlet-Steklov, on a accès à une formulation variationnelle dans un espace fonctionnel qui peut être décrit précisément. On verra qu'elle s'adapte assez naturellement à une formulation sous-structurée continue inspirée de la méthode des joints. Bien entendu, on peut obtenir à partir de la base de Dirichlet-Steklov une formulation matricielle, utile pour la mise en œuvre pratique.

Deuxième partie

Formulation sous-structurée réduite du problème thermique

Chapitre 4

Adaptation de la méthode des joints à un problème thermique sous-structuré

4.1 Formulation sous-structurée pour deux domaines

4.1.1 Équation de la chaleur

Dans un premier temps, on va exprimer le problème thermique sous-structuré pour un domaine Ω constitué de deux sous-domaines seulement (voir figure 2.4.1).

Avec les notations associées introduites au chapitre 2, on commence par exprimer l'équation de la chaleur sur chaque sous-domaine :

$$k, l \in \{1, 2\}, k \neq l, \Omega^{(k)}, \quad C^{(k)} \frac{\partial T}{\partial t} = \nabla(K^{(k)} \cdot \nabla T^{(k)}) + \varpi^{(k)}, \quad (2.4.1a)$$

$$\Gamma^{k \rightarrow l}, \quad K^{(k)} \cdot \nabla T^{(k)} \cdot \underline{n}_{k,l} = \varphi^{(k)}, \quad (2.4.1b)$$

$$\Sigma^{(k)}, \quad K^{(k)} \cdot \nabla T^{(k)} \cdot \underline{n}_{k,l} = h^{(k)}(T_{ext} - T^{(k)}), \quad (2.4.1c)$$

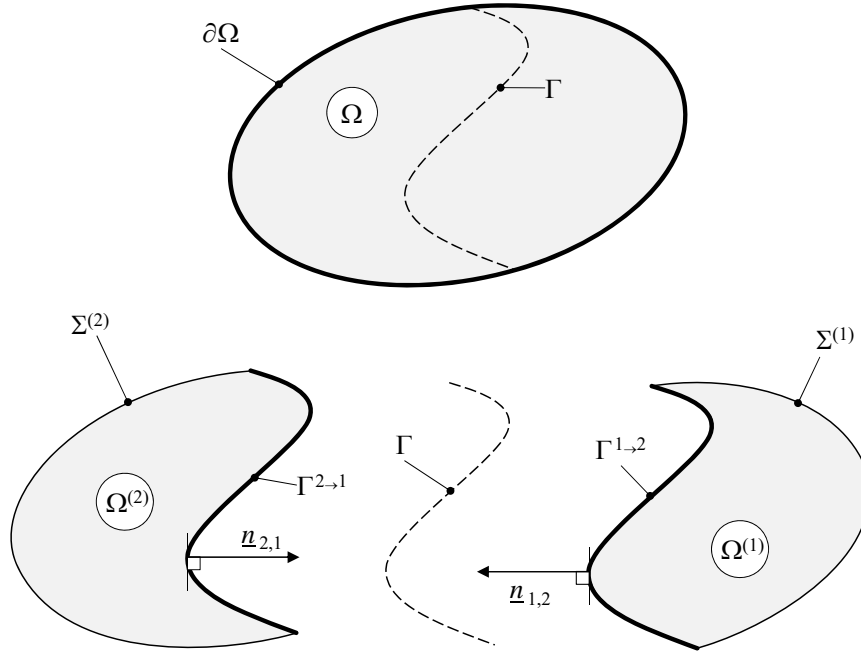
$$\Omega^{(k)}, \quad T^{(k)}(t=0) = T_0^{(k)}. \quad (2.4.1d)$$

Ici, $K^{(k)}$ est le tenseur des conductivités. L'équation (2.4.1b) définit $\varphi^{(k)}$ comme étant le flux sortant de $\Gamma^{k \rightarrow l}$.

4.1.1.1 Conditions de couplage

On doit ajouter à (2.4.1) des conditions de raccordement pour obtenir un système couplé. Tout d'abord, on impose la conservation du flux à l'interface, *i.e.*

$$\Gamma^{k \rightarrow l}, \quad \varphi^{(2)} = -\varphi^{(1)} = \varphi. \quad (2.4.2)$$

FIGURE 2.4.1 – Partition régulière du domaine Ω en deux sous domaines $\Omega^{(1)}$ et $\Omega^{(2)}$.

De plus, si on désigne par $R_c \geq 0$ une résistance de contact sur $\Gamma \equiv \Gamma^{\{1,2\}}$, on impose la condition

$$\Gamma^{k \rightarrow l}, \quad \llbracket T \rrbracket_{\Gamma^{2 \rightarrow 1}} = R_c \varphi, \quad (2.4.3)$$

où on désigne le saut d'une fonction u à l'interface Γ par $\llbracket u \rrbracket_{\Gamma^{p \rightarrow q}} = u^{(q)}|_{\Gamma} - u^{(p)}|_{\Gamma}$.

Le problème thermique sous-structuré est ainsi défini.

Définition 4.1.1 On cherche le champ de température solution du problème thermique :

$k, l \in \{1, 2\}, k \neq l,$

$$\left\{ \begin{array}{l} \Omega^{(k)}, \quad C^{(k)} \frac{\partial T}{\partial t} = \nabla(K^{(k)} \cdot \nabla T^{(k)}) + \varpi^{(k)}, \quad (2.4.4a) \\ \Gamma^{k \rightarrow l}, \quad K^{(k)} \cdot \nabla T^{(k)} \cdot \underline{n}_{k,l} = \varphi^{(k)}, \quad (2.4.4b) \\ \Sigma^{(k)}, \quad K^{(k)} \cdot \nabla T^{(k)} \cdot \underline{n}_{k,l} = h^{(k)}(T_{ext} - T^{(k)}), \quad (2.4.4c) \\ \Omega^{(k)}, \quad T^{(k)}(t=0) = T_0^{(k)}, \quad (2.4.4d) \\ \varphi^{(2)} = -\varphi^{(1)} = \varphi, \quad (2.4.4e) \\ \llbracket T \rrbracket_{\Gamma^{2 \rightarrow 1}} = R_c \varphi. \quad (2.4.4f) \end{array} \right.$$

4.1.2 Formulation variationnelle

On suppose que $T^{(k)} \in H^1(\Omega^{(k)})$ et on multiplie (2.4.4a) par une fonction test $v^{(k)} \in H^1(\Omega^{(k)})$ ($k \in \{1, 2\}$). En intégrant sur $\Omega^{(k)}$,

$$\int_{\Omega^{(k)}} C^{(k)} \frac{\partial T^{(k)}}{\partial t} v^{(k)} = \int_{\Omega^{(k)}} \nabla(K^{(k)} \cdot \nabla T^{(k)}) v^{(k)} + \int_{\Omega^{(k)}} \varpi^{(k)} v^{(k)}. \quad (2.4.5)$$

En utilisant la formule de Green on obtient

$$\int_{\Omega^{(k)}} \nabla(K^{(k)} \cdot \nabla T^{(k)}) v^{(k)} = - \int_{\Omega^{(k)}} \nabla T^{(k)} \cdot K^{(k)} \cdot \nabla v^{(k)} + \int_{\partial\Omega^{(k)}} K^{(k)} \cdot \nabla T^{(k)} \cdot \underline{n}_k v^{(k)}. \quad (2.4.6)$$

Or

$$\begin{aligned} \int_{\partial\Omega^{(k)}} K^{(k)} \cdot \nabla T^{(k)} \cdot \underline{n}_k v^{(k)} &= \int_{\Sigma^{(k)}} K^{(k)} \cdot \nabla T^{(k)} \cdot \underline{n}_k v^{(k)} + \int_{\Gamma} K^{(k)} \cdot \nabla T^{(k)} \cdot \underline{n}_{k,l} v^{(k)}, \quad k \neq l \\ &= \int_{\Sigma^{(k)}} h^{(k)} (T_{ext} - T^{(k)}) v^{(k)} + \int_{\Gamma} K^{(k)} \cdot \nabla T^{(k)} \cdot \underline{n}_{k,l} v^{(k)}. \end{aligned} \quad (2.4.7)$$

En injectant (2.4.6) et (2.4.7) dans (2.4.5) et en sommant sur les deux domaines,

$$\begin{aligned} &\sum_{k=1,2} \int_{\Omega^{(k)}} C^{(k)} \frac{\partial T^{(k)}}{\partial t} v^{(k)} + \sum_{k=1,2} \int_{\Omega^{(k)}} \nabla T^{(k)} \cdot K^{(k)} \cdot \nabla v^{(k)} \\ &= \sum_{k=1,2} \int_{\Omega^{(k)}} \varpi^{(k)} v^{(k)} + \sum_{k=1,2} \int_{\Sigma^{(k)}} h^{(k)} (T_{ext} - T^{(k)}) v^{(k)} + \int_{\Gamma} K^{(1)} \cdot \nabla T^{(1)} \cdot \underline{n}_{1,2} v^{(1)} + \int_{\Gamma} K^{(2)} \cdot \nabla T^{(2)} \cdot \underline{n}_{2,1} v^{(2)}. \end{aligned}$$

On reconnaît dans les deux dernières intégrales les flux de conduction $\varphi^{(1)}$ et $\varphi^{(2)}$:

$$\begin{aligned} &\sum_{k=1,2} \int_{\Omega^{(k)}} C^{(k)} \frac{\partial T^{(k)}}{\partial t} v^{(k)} + \sum_{k=1,2} \int_{\Omega^{(k)}} \nabla T^{(k)} \cdot K^{(k)} \cdot \nabla v^{(k)} \\ &= \sum_{k=1,2} \int_{\Omega^{(k)}} \varpi^{(k)} v^{(k)} + \sum_{k=1,2} \int_{\Sigma^{(k)}} h^{(k)} (T_{ext} - T^{(k)}) v^{(k)} + \int_{\Gamma} \varphi^{(1)} v^{(1)} + \int_{\Gamma} \varphi^{(2)} v^{(2)} \\ &= \sum_{k=1,2} \int_{\Omega^{(k)}} \varpi^{(k)} v^{(k)} + \sum_{k=1,2} \int_{\Sigma^{(k)}} h^{(k)} (T_{ext} - T^{(k)}) v^{(k)} - \int_{\Gamma} \varphi \llbracket v \rrbracket_{\Gamma^{2 \rightarrow 1}} \text{ d'après (2.4.4e)}. \end{aligned} \quad (2.4.8)$$

4.1.2.1 Formulation variationnelle utilisant une résistance de contact

On peut remplacer le terme de flux à l'interface par son expression tirée de (2.4.4f). On aurait alors une formulation du type :

Définition 4.1.2 Trouver $(T^{(1)}, T^{(2)}) \in H^1(\Omega^{(1)}) \times H^1(\Omega^{(2)})$ tels que

$$\begin{aligned} \sum_{k=1,2} \int_{\Omega^{(k)}} C^{(k)} \frac{\partial T^{(k)}}{\partial t} v^{(k)} + \sum_{k=1,2} \int_{\Omega^{(k)}} \nabla T^{(k)} \cdot K^{(k)} \cdot \nabla v^{(k)} + \sum_{k=1,2} \int_{\Sigma^{(k)}} h^{(k)} T^{(k)} v^{(k)} + \int_{\Gamma} \frac{[[v]]_{\Gamma^{2 \rightarrow 1}} [[T]]_{\Gamma^{2 \rightarrow 1}}}{R_c} \\ = \sum_{k=1,2} \int_{\Omega^{(k)}} \varpi^{(k)} v^{(k)} + \sum_{k=1,2} \int_{\Sigma^{(k)}} h^{(k)} T_{ext} v^{(k)}, \quad (2.4.9) \end{aligned}$$

quelque soit $(v^{(1)}, v^{(2)}) \in H^1(\Omega^{(1)}) \times H^1(\Omega^{(2)})$.

C'est cette formulation qui était utilisée pour les premiers travaux de sous-structuration modale développés par Laffay [89] et au début de cette thèse, mais elle présente plusieurs inconvénients. Tout d'abord, il est impossible de traiter une condition de contact parfait puisqu'il est nécessaire que $R_c \neq 0$. On ne peut que s'en approcher en choisissant une "petite" résistance de contact, dont la valeur s'obtient par une analyse physique (pour les contacts entre composants et PCB, les valeurs typiques sont entre 10^{-9} et $10^{-7} m^2 \cdot K \cdot W^{-1}$ [6]). Mais une valeur trop faible provoque des problèmes de convergence numérique, en donnant trop de poids au terme surfacique. Enfin, la prise en compte de maillages non conformes à l'interface est impossible avec cette formulation.

Au contraire, traiter φ comme une inconnue en suivant la méthode des joints permet d'avoir une condition de continuité imposée au sens faible, ce qui facilite le calcul à l'interface. Nous décrivons donc à présent cette méthode.

4.1.2.2 Formulation variationnelle inspirée de la méthode des joints

On conservant l'expression (2.4.8) sous la forme :

$$\begin{aligned} \sum_{k=1,2} \int_{\Omega^{(k)}} C^{(k)} \frac{\partial T^{(k)}}{\partial t} v^{(k)} + \sum_{k=1,2} \int_{\Omega^{(k)}} \nabla T^{(k)} \cdot K^{(k)} \cdot \nabla v^{(k)} \\ = \sum_{k=1,2} \int_{\Omega^{(k)}} \varpi^{(k)} v^{(k)} d\Omega^{(k)} + \sum_{k=1,2} \int_{\Sigma^{(k)}} h^{(k)} (T_{ext} - T^{(k)}) - \int_{\Gamma} \varphi [[v]]_{\Gamma^{2 \rightarrow 1}}, \quad (2.4.10) \end{aligned}$$

on constate qu'on a un "saut de fonction test".

Pour traiter la condition de continuité de flux, on va s'inspirer de la méthode des mortiers décrite dans la section 2.5 page 26 en conservant φ comme inconnue et en faisant apparaître

une condition de couplage faible à l'aide de (2.4.4f). On choisit $\Gamma^{1 \rightarrow 2}$ comme frontière mortier, la frontière non-mortier étant $\Gamma^{2 \rightarrow 1}$. On pose alors $\underline{n}_\Gamma = \underline{n}_{1,2}$.

Ainsi, on multiplie (2.4.4f) par une fonction test ϕ définie sur $\Gamma^{2 \rightarrow 1}$ et on intègre sur cette même interface :

$$\int_{\Gamma^{2 \rightarrow 1}} \llbracket T \rrbracket_{\Gamma^{2 \rightarrow 1}} \phi = R_c \int_{\Gamma^{2 \rightarrow 1}} \varphi \phi. \quad (2.4.11)$$

Remarque 4.1.1 Si $R_c = 0$, on retrouve une formulation variationnelle de la condition de raccordement (1.2.22), et qui traduit une condition de contact parfait entre les deux sous-domaines.

Finalement, en combinant (2.4.10) et (2.4.11), on obtient la formulation variationnelle du problème thermique.

Définition 4.1.3 On suppose que les sollicitations $\varpi^{(k)}$ sont dans $L^2(\Omega^{(k)})$. La formulation variationnelle associée au problème thermique introduit à la définition 2.4.4 est la suivante :

Trouver $(T^{(1)}, T^{(2)}, \varphi) \in H^1(\Omega^{(1)}) \times H^1(\Omega^{(2)}) \times W(\Gamma)$ tels que

$$\begin{aligned} \sum_{k=1,2} \int_{\Omega^{(k)}} C^{(k)} \frac{\partial T^{(k)}}{\partial t} v^{(k)} + \sum_{k=1,2} \int_{\Omega^{(k)}} \nabla T^{(k)} \cdot K^{(k)} \cdot \nabla v^{(k)} + \sum_{k=1,2} \int_{\Sigma^{(k)}} h^{(k)} T^{(k)} v^{(k)} + \int_{\Gamma^{(2,1)}} \varphi \llbracket v \rrbracket_{\Gamma^{2 \rightarrow 1}} \\ = \sum_{k=1,2} \int_{\Omega^{(k)}} \varpi^{(k)} v^{(k)} + \sum_{k=1,2} \int_{\Sigma^{(k)}} h^{(k)} T_{ext} v^{(k)}, \end{aligned} \quad (2.4.12a)$$

$$\int_{\Gamma^{2 \rightarrow 1}} \llbracket T \rrbracket_{\Gamma^{2 \rightarrow 1}} \phi = R_c \int_{\Gamma^{2 \rightarrow 1}} \varphi \phi, \quad (2.4.12b)$$

quelque soit $(v^{(1)}, v^{(2)}, \phi) \in H^1(\Omega^{(1)}) \times H^1(\Omega^{(2)}) \times W(\Gamma)$, où $W(\Gamma)$ est un espace que l'on va définir ultérieurement.

On a donc obtenu une formulation continue du problème thermique sous-structuré inspirée de la méthode des joints. On verra qu'on pourra conserver une formulation continue pour le problème réduit par la méthode modale, puisque celle-ci peut également s'exprimer dans le cadre de l'analyse fonctionnelle. La méthode modale faisant apparaître des bases infinies et dénombrables des espaces fonctionnels définis en 4.1.3, on va pouvoir donner une forme matricielle du problème indépendante de toute discrétisation spatiale, qui apparaîtra alors comme un complément pour le calcul effectif des bases.

4.1.3 Discrétisation du problème sous-structuré : principe

Soient $\{\psi_i^{(1)}\}_{i \in \mathbb{N}}$ et $\{\psi_i^{(2)}\}_{i \in \mathbb{N}}$ des bases des espaces de Hilbert $H^1(\Omega^{(1)})$ et $H^1(\Omega^{(2)})$ et soit $\{\phi_m\}_{1 \leq m \leq N_\phi}$ la famille des fonctions qui forment une base de l'espace mortier $W(\Gamma)$. Ces

bases étant infinies et dénombrables, on doit se limiter à un nombre fini de fonctions pour décomposer T et φ dans le cadre d'une formulation matricielle. On approche $T^{(1)}$, $T^{(2)}$ et φ par une décomposition sur les bases tronquées :

$$T^{(1)} = \sum_{i=1}^{N^{(1)}} X_i^{(1)} \psi_i^{(1)}, \quad T^{(2)} = \sum_{i=1}^{N^{(2)}} X_i^{(2)} \psi_i^{(2)}, \quad \varphi = \sum_{i=1}^{N_\phi} \eta_i \phi_i. \quad (2.4.13)$$

On pose $s_g^{k,l} = -\underline{n}_\Gamma \cdot \underline{n}_{k,l}$. Comme $\underline{n}_\Gamma = \underline{n}_{1,2}$, $s_g^{1,2} = -1$ et $s_g^{2,1} = 1$: ce coefficient vaut (-1) du côté mortier ($\Gamma^{1 \rightarrow 2}$) et 1 du côté non-mortier ($\Gamma^{2 \rightarrow 1}$), et va ainsi servir de marqueur pour différencier les deux côtés de Γ . Il permet aussi d'écrire le problème discret de façon plus compacte.

En injectant (2.4.13) dans (2.4.12) et en prenant les $\psi_j^{(k)}$ comme fonction test, on a $\forall k, l \in \{1, 2\}$, $k \neq l$, $\forall j \in \{1, \dots, N^{(k)}\}$,

$$\begin{aligned} \sum_{i=1}^{N^{(k)}} \frac{\partial X_i^{(k)}}{\partial t} \int_{\Omega^{(k)}} C^{(k)} \psi_i^{(k)} \psi_j^{(k)} + \sum_{i=1}^{N^{(k)}} X_i^{(k)} \int_{\Omega^{(k)}} \nabla \psi_i^{(k)} \cdot K^{(k)} \cdot \nabla \psi_j^{(k)} + \sum_{i=1}^{N^{(k)}} X_i^{(k)} \int_{\Sigma^{(k)}} h^{(k)} \psi_i^{(k)} \psi_j^{(k)} \\ + s_g^{k,l} \sum_{m=1}^{N_\phi} \eta_m \int_{\Gamma} \psi_j^{(k)} \phi_m = \int_{\Omega^{(k)}} \varpi^{(k)} \psi_j^{(k)} + \int_{\Sigma^{(k)}} h^{(k)} T_{ext} \psi_j^{(k)}, \end{aligned} \quad (2.4.14)$$

et

$$\forall m \in \{1, \dots, N_\phi\},$$

$$\sum_{k=1}^2 s_g^{k,l} \sum_{i=1}^{N^{(k)}} X_i^{(k)} \int_{\Gamma} \psi_i^{(k)} \phi_m - R_c \sum_{j=1}^{N_\phi} \eta_j \int_{\Gamma} \phi_m \phi_j = 0. \quad (2.4.15)$$

On introduit les matrices suivantes :

$$\begin{aligned} \mathbf{C}^{(k)} &= \left(\int_{\Omega^{(k)}} C^{(k)} \psi_i^{(k)} \psi_j^{(k)} \right)_{\substack{1 \leq i \leq N^{(k)} \\ 1 \leq j \leq N^{(k)}}}, \quad \mathbf{K}^{(k)} = \left(\int_{\Omega^{(k)}} \nabla \psi_i^{(k)} \cdot K^{(k)} \cdot \nabla \psi_j^{(k)} \right)_{\substack{1 \leq i \leq N^{(k)} \\ 1 \leq j \leq N^{(k)}}}, \\ \mathbf{H}^{(k)} &= \left(\int_{\Sigma^{(k)}} h^{(k)} \psi_i^{(k)} \psi_j^{(k)} \right)_{\substack{1 \leq i \leq N^{(k)} \\ 1 \leq j \leq N^{(k)}}}, \quad \mathbf{J}^{(k)} = s_g^{k,l} \left(\int_{\Gamma} \psi_i^{(k)} \phi_m \right)_{\substack{1 \leq i \leq N^{(k)} \\ 1 \leq m \leq N_\phi}}, \\ \mathbf{M}_c &= -R_c \left(\int_{\Gamma} \phi_i \phi_j d\Gamma \right)_{\substack{1 \leq i \leq N_\phi \\ 1 \leq j \leq N_\phi}}, \end{aligned} \quad (2.4.16)$$

et les vecteurs

$$\begin{aligned} \mathbf{\Pi}^{(k)} &= \left(\int_{\Omega^{(k)}} \varpi^{(k)} \psi_i^{(k)} \right)_{1 \leq i \leq N^{(k)}}, \quad \mathbf{\Theta}_{ext}^{(k)} = \left(\int_{\Sigma^{(k)}} h^{(k)} T_{ext} \psi_i^{(k)} \right)_{1 \leq i \leq N^{(k)}}, \\ \dot{\mathbf{X}}^{(k)} &= \left(\frac{\partial X_i^{(k)}}{\partial t} \right)_{1 \leq i \leq N^{(k)}}, \quad \mathbf{X}^{(k)} = \left(X_i^{(k)} \right)_{1 \leq i \leq N^{(k)}} \quad \text{et} \quad \mathbf{\Phi} = (\eta_i)_{1 \leq i \leq N_\phi}. \end{aligned} \quad (2.4.17)$$

La formulation matricielle du problème défini en 4.1.3 s'écrit alors :

$$\begin{bmatrix} \mathbf{C}^{(1)} & 0 & 0 \\ 0 & \mathbf{C}^{(2)} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} \dot{\mathbf{X}}^{(1)} \\ \dot{\mathbf{X}}^{(2)} \\ \dot{\Phi} \end{pmatrix} + \begin{bmatrix} \mathbf{K}^{(1)} + \mathbf{H}^{(1)} & 0 & \mathbf{J}^{(1)} \\ 0 & \mathbf{K}^{(2)} + \mathbf{H}^{(2)} & \mathbf{J}^{(2)} \\ {}^t\mathbf{J}^{(1)} & {}^t\mathbf{J}^{(2)} & \mathbf{M}_c \end{bmatrix} \begin{pmatrix} \mathbf{X}^{(1)} \\ \mathbf{X}^{(2)} \\ \Phi \end{pmatrix} = \begin{pmatrix} \mathbf{\Pi}^{(1)} + \Theta_{ext}^{(1)} \\ \mathbf{\Pi}^{(2)} + \Theta_{ext}^{(2)} \\ 0 \end{pmatrix}. \quad (2.4.18)$$

On retrouve ainsi la formulation duale de la méthode des joints, mais appliquée au problème thermique.

4.1.4 Exemple avec une base de fonctions spatiales discrétisées

Pour illustrer la formule (2.4.18), on va choisir une base de fonctions spatiales discrétisées par éléments finis. On suppose que les $\Omega^{(k)}$ sont des polygones ou des polyèdres et que l'on a pour chaque $\Omega^{(k)}$ une triangulation $\mathcal{T}_h^{(k)}$ constituée de triangles ou de tétraèdres \mathcal{K} , telle que

- $\overline{\Omega^{(k)}} = \bigcup_{\mathcal{K} \in \mathcal{T}_h^{(k)}} \mathcal{K}$,
- l'intersection entre deux éléments \mathcal{K}_1 et \mathcal{K}_2 de $\mathcal{T}_h^{(k)}$ est soit vide, soit un sommet soit une arête soit une face complète de \mathcal{K}_1 et \mathcal{K}_2 ,
- le rapport entre le diamètre d'un élément \mathcal{K} et le diamètre de son cercle inscrit ou de sa sphère inscrite est borné par une constante σ indépendante de \mathcal{K} et h .

4.1.4.1 Espaces fonctionnels

Pour $k \in \{1, 2\}$, on se donne un espace de fonctions de dimension finie associée à la discrétisation $\mathcal{T}_h^{(k)}$, noté $V_h^{(k)}$.

On suppose que $V_h^{(k)} \subset H^1(\Omega^{(k)})$.

Si une frontière de $\Omega^{(k)}$ est un côté esclave δ tel que défini en section 2.5, on note par W_h^δ l'espace des traces des fonctions de $V_h^{(k)}$ sur δ , défini par

$$W_h^\delta = \{v|_\delta, v \in V_h^{(k)}\}.$$

L'espace des fonctions mortiers \widetilde{W}_h^δ est ainsi un sous-espace de W_h^δ : $\widetilde{W}_h^\delta \subseteq W_h^\delta$.

4.1.4.2 Espaces fonctionnels en éléments finis

Dans notre cadre d'étude, $V_h^{(k)}$ est l'espace classique pour les éléments finis

$$V_h^{(k)} = \{v_h \in H^1(\Omega^{(k)}); \forall \mathcal{K} \in \mathcal{T}_h^{(k)}, v_h|_{\mathcal{K}} \in \mathcal{P}_l(\mathcal{K})\},$$

avec $\mathcal{P}_l(\mathcal{K})$ l'espace des polynômes à d variables et de degré total inférieur ou égal à l restreints à \mathcal{K} . Par conséquent, pour chaque δ , l'espace des traces est donné par

$$W_h^\delta = \{\phi_h \in H^1(\delta); \forall \mathcal{K} \in \mathcal{T}_h^\delta, \phi_h|_{\mathcal{K} \cap \delta} \in \mathcal{P}_l(\mathcal{K} \cap \delta)\},$$

avec \mathcal{T}_h^δ la restriction à δ de $\mathcal{T}_h^{(k)}$, où $\Omega^{(k)}$ est le domaine associé à δ .

Ici, $\delta = \Gamma^{2 \rightarrow 1}$, mais on pourra garder cette notation pour n'importe quelle frontière non-mortier δ dans le cas où Ω est décomposé en plus de deux sous-domaines.

On note $N^{(1)}$ le nombre de nœuds du maillage de $\Omega^{(1)}$ et $N^{(2)}$ le nombre de nœuds du maillage de $\Omega^{(2)}$.

On suppose que $\Omega^{(1)}$ et $\Omega^{(2)}$ sont maillés indépendamment, de sorte que l'on n'impose pas de conformité de maillage sur Γ , comme l'illustre par exemple la figure 2.4.2.

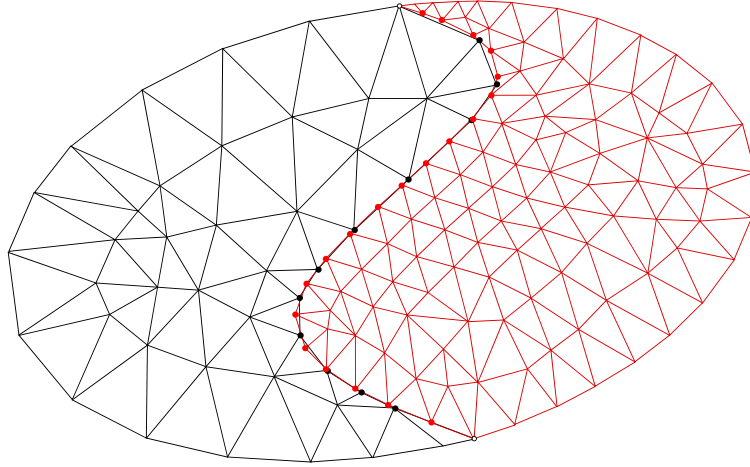


FIGURE 2.4.2 – Compatibilité des maillages. Exemple du raccordement de deux maillages incompatibles.

Ainsi, on choisit $\{\psi_i^{(1)}\}_{1 \leq i \leq N^{(1)}}$ comme étant la famille des fonctions de forme associées au maillage de $\Omega^{(1)}$ et $\{\psi_i^{(2)}\}_{1 \leq i \leq N^{(2)}}$ la famille des fonctions de forme associées au maillage de $\Omega^{(2)}$.

On choisit aussi $\{\phi_m\}_{1 \leq m \leq N_\phi}$ la famille des fonctions qui forment une base de l'espace mortier $\widetilde{W}_h^{\Gamma^{2 \rightarrow 1}}$. On discutera du choix de ces fonctions dans la partie consacrée à l'exemple d'application.

4.2 Généralisation à N domaines

4.2.1 Décomposition de Ω dans le cadre des cartes électroniques

Pour décrire le problème dans le cas d'un découpage en deux sous-domaines, on a pris un cas simple en 2D, mais la problématique de ce travail, à savoir la modélisation de cartes électroniques, est évidemment en 3D. Dans ce cadre, *i.e.* plusieurs composants sur un PCB, la sous-structuration naturelle est de choisir une sous-structure pour le PCB et une sous-structure par composant, comme le montre la figure 2.4.3, où six circuits intégrés (IC, ou "Integrated Circuit") sont posés sur un circuit imprimé. Deux choix s'offrent alors à nous pour

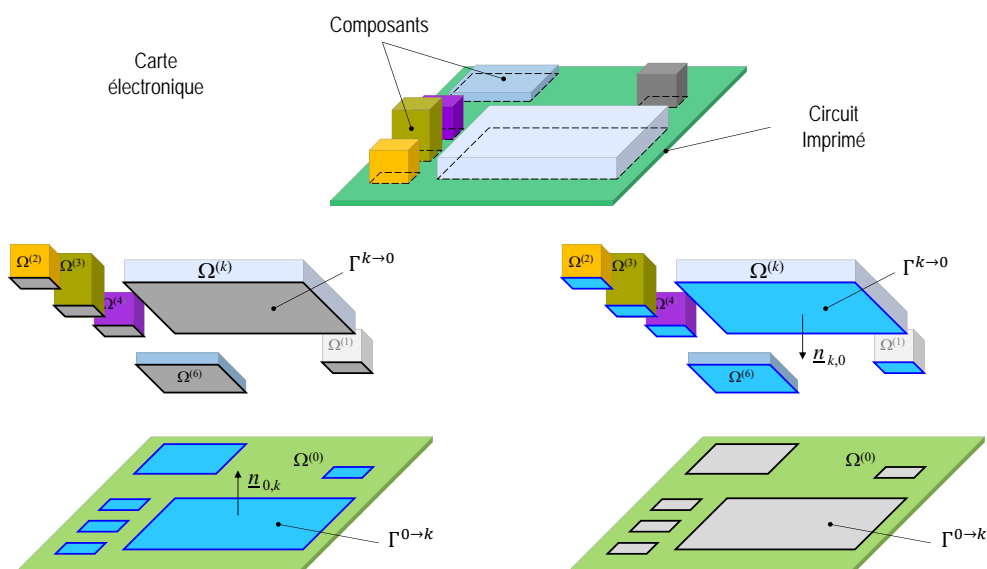


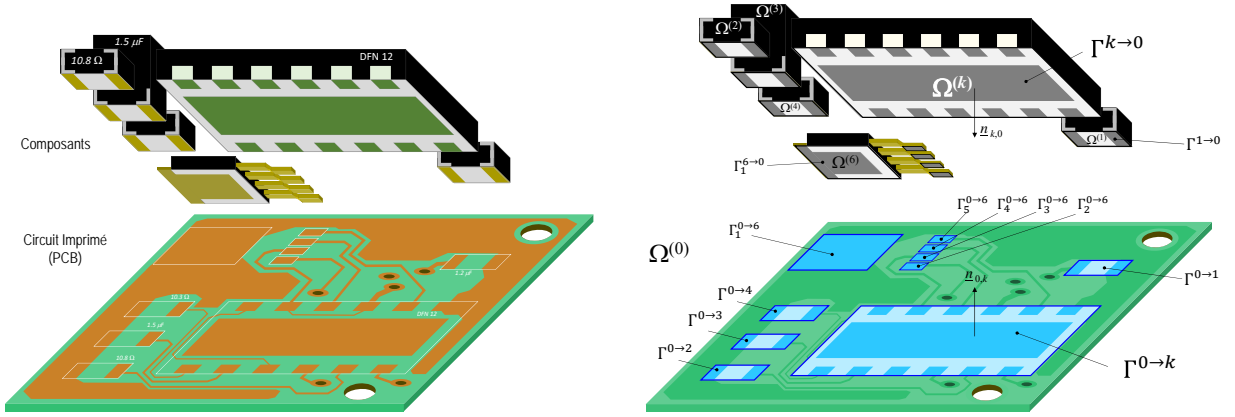
FIGURE 2.4.3 – Deux stratégies de choix des frontières maîtres d'une carte électronique.

l'affectation des côtés mortiers : côté PCB ou côté IC. Dans l'objectif de créer une bibliothèque de composants réduits, le choix logique semble être d'affecter le côté mortier du côté PCB :

- le masque d'implantation sur le PCB est supposé être toujours connu, puisqu'il est lié à la cartographie des pistes, et celle-ci est à refaire si la configuration de la carte change ;
- on choisit ainsi un découpage conforme, ce qui simplifie la méthodologie ;
- les ϕ_m sont ainsi calculés du côté du composant, et pourront être réutilisés pour modéliser l'implantation du composant sur un autre endroit de la carte, voire sur une autre carte. Le choix inverse pousserait à calculer les ϕ_m à chaque fois qu'on voudrait utiliser le modèle du composant.

Toutefois, ce modèle paraît encore un peu simpliste. La figure 2.4.4 représente un modèle

un peu plus réaliste de carte et permet de voir apparaître une nouvelle problématique : le couplage entre le PCB et un IC peut se faire via une ou plusieurs interfaces, comme par exemple pour le composant $\Omega^{(6)}$.



(a) Le masque d'implémentation est représenté en blanc sur le PCB.

(b) Les côtés mortiers (en bleu) sont choisis du côté du PCB.

FIGURE 2.4.4 – Schéma de carte “réaliste”.

Pour l'écriture d'un problème thermique à N domaines, on va donc considérer que pour chaque contact, on a plusieurs interfaces.

4.2.2 Formulation variationnelle

Pour la suite du manuscrit, on considère donc que Ω est décomposé en $N_\Omega + 1$ sous-domaines tels que $\Omega = \bigcup_{k=0}^{N_\Omega} \overline{\Omega^{(k)}}$ et $\Omega^{(k)} \cap \Omega^{(l)} = \emptyset$, comme dans la section 2.5. On reprend les mêmes notations que dans cette section et qu'en section 2.1 pour $\partial\Omega^{(k)}$, $\Sigma^{(k)}$ et $\Gamma^{\{k,l\}}$.

S'il y a plusieurs surfaces entre $\Omega^{(k)}$ et $\Omega^{(l)}$ (comme on l'a vu avec la figure 2.4.4), on note $\Gamma^{\{k,l\}} = \bigcup_{p=1}^{N_{\Gamma^{\{k,l\}}}} \Gamma_p^{\{k,l\}}$ avec $\Gamma_p^{\{k,l\}} \cup \Gamma_q^{\{k,l\}} = \emptyset$. $\overline{\Gamma}_p^{\{k,l\}} \cap \overline{\Gamma}_q^{\{k,l\}}$ peut alors être une arête, un point ou \emptyset . Pour généraliser, on supposera que l'on est dans ce cas pour la suite (le cas où l'interface entre $\Omega^{(k)}$ et $\Omega^{(l)}$ est réduite à une frontière commune se traduisant par $N_{\Gamma^{\{k,l\}}} = 1$). On peut donc garder $\Gamma = \bigcup_{k,l \in \{0, N_\Omega\}} \Gamma^{\{k,l\}}$ comme définition du squelette de la décomposition de Ω . On redéfinit $\Gamma^{k \rightarrow l} = \overline{\Omega^{(k)}} \cap \Gamma^{\{k,l\}}$ l'interface “du côté” de $\Omega^{(k)}$, et en particulier $\Gamma_p^{k \rightarrow l} = \overline{\Omega^{(k)}} \cap \Gamma_p^{\{k,l\}}$.

D'après la section 2.5.2, on rappelle qu'on note Γ_{mortier} l'ensemble des frontières mortier de Γ et $\Gamma_{\text{non-mortier}}$ l'ensemble des non-mortiers, et qu'on définit $N_{\text{mortier}} = \dim(\Gamma_{\text{mortier}})$ et $N_{\text{non-mortier}} = \dim(\Gamma_{\text{non-mortier}})$.

On veut maintenant, pour appliquer la méthode des joints, partitionner Γ en mortiers et en non-mortiers. Puisqu'on a choisi de se restreindre au cas d'une décomposition conforme, cela revient à choisir pour chaque interface $\Gamma_p^{\{k,l\}}$ un côté mortier et un côté non-mortier, indépendamment des autres (même si on l'a vu, il est plus pertinent de choisir systématiquement le côté PCB comme mortier pour le cas d'une carte électronique).

Ainsi, soit on affecte $\Gamma_p^{k \rightarrow l}$ à Γ_{\sqcap} , et alors $\Gamma_p^{l \rightarrow k} \in \Gamma_{\sqcup}$, et on a $s_{g_p}^{k,l} = -1$ et $s_{g_p}^{l,k} = +1$; soit on affecte $\Gamma_p^{k \rightarrow l}$ à Γ_{\sqcup} , et alors $\Gamma_p^{l \rightarrow k} \in \Gamma_{\sqcap}$, et on a $s_{g_p}^{k,l} = +1$ et $s_{g_p}^{l,k} = -1$.

Avec ces notations, on a alors $\Gamma_{\sqcap} = \bigcup_{s_{g_p}^{k,j} < 0} \bar{\Gamma}_p^{k \rightarrow j}$ et $\Gamma_{\sqcup} = \bigcup_{s_{g_q}^{k,j} > 0} \bar{\Gamma}_q^{k \rightarrow j}$.

La conformité de la décomposition implique aussi que $N_{\sqcap} = N_{\sqcup}$.

Définition 4.2.1 *On suppose que les sollicitations $\varpi^{(k)}$ sont dans $L^2(\Omega^{(k)})$. La formulation variationnelle associée au problème thermique introduit à la définition 2.4.4 est la suivante :*

Trouver $(T^{(k)}, \varphi_p^{\{k,l\}}) \in H^1(\Omega^{(k)}) \times W(\Gamma_p^{\{k,l\}})$ tels que

$$\begin{aligned} & \sum_{k=0}^{N_\Omega} \int_{\Omega^{(k)}} C^{(k)} \frac{\partial T^{(k)}}{\partial t} v^{(k)} + \sum_{k=0}^{N_\Omega} \int_{\Omega^{(k)}} \nabla T^{(k)} \cdot K^{(k)} \cdot \nabla v^{(k)} + \sum_{k=0}^{N_\Omega} \int_{\Sigma^{(k)}} h^{(k)} T^{(k)} v^{(k)} \\ & + \sum_{\Gamma_p^{\{k,l\}} \in \Gamma_{\sqcup}} \int_{\Gamma_p^{\{k,l\}}} \varphi_p^{k,l} \llbracket v \rrbracket_{\Gamma_p^{\{k,l\}}} = \sum_{k=0}^{N_\Omega} \int_{\Omega^{(k)}} \varpi^{(k)} v^{(k)} + \sum_{k=0}^{N_\Omega} \int_{\Sigma^{(k)}} h^{(k)} T_{ext} v^{(k)}, \end{aligned} \quad (2.4.19a)$$

$$\forall \Gamma_p^{\{k,l\}} \in \Gamma_{\sqcap}, \int_{\Gamma_p^{\{k,l\}}} \llbracket T \rrbracket_{\Gamma_p^{\{k,l\}}} \phi_p^{\{k,l\}} = R_{c_p}^{\{k,l\}} \int_{\Gamma_p^{\{k,l\}}} \varphi_p^{\{k,l\}} \phi_p^{\{k,l\}}, \quad (2.4.19b)$$

quelque soit $(v^{(k)}, \phi_p^{\{k,l\}}) \in H^1(\Omega^{(k)}) \times W(\Gamma_p^{\{k,l\}})$, $0 \leq k, l \leq N_\Omega$, $k \neq l$, $1 \leq p \leq N_{\Gamma}^{\{k,l\}}$.

4.2.3 Forme matricielle

De manière similaire au cas où l'on a 2 sous-domaines, on décompose une approximation de $T^{(k)}$ et de la densité de flux $\varphi_p^{k,j}$ associée au joint $\Gamma_p^{\{k,j\}}$ sur des bases tronquées de $H^1(\Omega^{(k)})$ et $W(\Gamma_p^{\{k,j\}})$:

$$T^{(k)} = \sum_{i=1}^{N^{(k)}} X_i^{(k)} \psi_i^{(k)}, \quad \varphi_p^{k,j} = \sum_{m=1}^{N_{\phi}^{\Gamma_p^{\{k,j\}}}} \eta_{p_m}^{\{k,j\}} \phi_{p_m}^{\{k,j\}}. \quad (2.4.20)$$

En injectant (2.4.20) dans (2.4.19) et en prenant $\psi_j^{(k)}$ comme fonction test, on a

$$\forall k \in \{0, \dots, N_\Omega\}, \forall j \in \{1, \dots, N^{(k)}\},$$

$$\begin{aligned} & \sum_{i=1}^{N^{(k)}} \frac{\partial X_i^{(k)}}{\partial t} \int_{\Omega^{(k)}} C^{(k)} \psi_i^{(k)} \psi_j^{(k)} + \sum_{i=1}^{N^{(k)}} X_i^{(k)} \int_{\Omega^{(k)}} \nabla \psi_i^{(k)} \cdot K^{(k)} \cdot \nabla \psi_j^{(k)} + \sum_{i=1}^{N^{(k)}} X_i^{(k)} \int_{\Sigma^{(k)}} h^{(k)} \psi_i^{(k)} \psi_j^{(k)} \\ & \quad + \sum_{l=0}^{N_\Omega} \sum_{p=1}^{N_\Gamma^{\{k,l\}}} s_{g_p}^{k,l} \sum_{m=1}^{N_\phi^{\{k,l\}}} \eta_{p_m}^{\{k,l\}} \int_{\Gamma_p^{\{k,l\}}} \psi_j^{(k)} \phi_{p_m}^{\{k,l\}} \\ & = \int_{\Omega^{(k)}} \varpi^{(k)} \psi_j^{(k)} + \int_{\Sigma^{(k)}} h^{(k)} T_{ext} \psi_j^{(k)}, \end{aligned} \tag{2.4.21}$$

et pour $0 \leq k, l \leq N_\Omega$, $k \neq l$, $1 \leq p \leq N_\Gamma^{\{k,l\}}$ tels que $\Gamma_p^{\{k,l\}} \in \Gamma_{\square}$, $\forall m \in \{1, \dots, N_\phi^{\{k,l\}}\}$,

$$s_g^{k,l} \sum_{i=1}^{N^{(k)}} X_i^{(k)} \int_{\Gamma_p^{\{k,l\}}} \psi_i^{(k)} \phi_{p_m}^{\{k,l\}} + s_g^{l,k} \sum_{i=1}^{N^{(l)}} X_i^{(l)} \int_{\Gamma_p^{\{k,l\}}} \psi_i^{(l)} \phi_{p_m}^{\{l,k\}} - \sum_{j=1}^{N_\phi^{\{k,l\}}} \eta_{p_j}^{\{k,l\}} \int_{\Gamma_p^{\{k,l\}}} R_{c_p}^{(k,l)} \phi_{p_m}^{\{k,l\}} \phi_{p_j}^{\{k,l\}} = 0. \tag{2.4.22}$$

Soient les matrices $\mathbf{C}^{(k)}$, $\mathbf{K}^{(k)}$, $\mathbf{H}^{(k)}$, $\mathbf{\Theta}_{ext}$, $\dot{\mathbf{X}}^{(k)}$ et $\mathbf{X}^{(k)}$, définies comme en (2.4.16) et (2.4.17), avec $0 \leq k \leq N_\Omega$.

On pose, pour $0 \leq k, l \leq N_\Omega$, $1 \leq p \leq N_\Gamma^{\{k,l\}}$,

$$\mathbf{J}_p^{(k,l)} = s_{g_p}^{k,l} \left(\int_{\Gamma_p^{\{k,l\}}} \psi_i^{(k)} \phi_{p_m}^{\{k,l\}} \right)_{\substack{1 \leq i \leq N^{(k)} \\ 1 \leq m \leq N_\phi^{\{k,l\}}}}, \quad \mathbf{M}_{c_p}^{(k,l)} = - \left(\int_{\Gamma_p^{\{k,l\}}} R_{c_p}^{(k,l)} \phi_{p_i}^{\{k,l\}} \phi_{p_m}^{\{k,l\}} \right)_{\substack{1 \leq i \leq N_\phi^{\{k,l\}} \\ 1 \leq m \leq N_\phi^{\{k,l\}}}}; \tag{2.4.23}$$

on peut ainsi former la matrice rectangulaire

$$\mathbf{J}^{(k,l)} = \left[\mathbf{J}_1^{(k,l)}, \dots, \mathbf{J}_{N_\Gamma^{\{k,l\}}}^{(k,l)} \right]. \tag{2.4.24}$$

On pose $\mathbf{\Phi}_p^{(k,l)} = \left(\eta_{p_m}^{\{k,l\}} \right)_{1 \leq m \leq N_\phi^{\{k,l\}}}$.

On pose enfin

$$\begin{aligned}
\mathbf{J} &= \begin{bmatrix} \mathbf{J}^{(0,0)} & \mathbf{J}^{(0,1)} & \dots & \mathbf{J}^{(0,N_{\perp})} \\ \mathbf{J}^{(1,0)} & \mathbf{J}^{(1,1)} & \dots & \mathbf{J}^{(1,N_{\perp})} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{J}^{(N_{\Omega},0)} & \mathbf{J}^{(N_{\Omega},1)} & \dots & \mathbf{J}^{(N_{\Omega},N_{\perp})} \end{bmatrix}, \quad \mathbf{M}_c = \left(\mathbf{M}_{c_p}^{(k,l)} \right)_{\Gamma_p^{\{k,l\}} \in \Gamma_{\perp}}, \quad \Phi = \left(\Phi_p^{(k,l)} \right)_{\Gamma_p^{\{k,l\}} \in \Gamma_{\perp}}, \\
\mathbf{C} &= \begin{bmatrix} \mathbf{C}^{(0)} & 0 & 0 & 0 \\ 0 & \mathbf{C}^{(1)} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \mathbf{C}^{(N_{\Omega})} \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} \mathbf{K}^{(0)} & 0 & 0 & 0 \\ 0 & \mathbf{K}^{(1)} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \mathbf{K}^{(N_{\Omega})} \end{bmatrix}, \\
\mathbf{H} &= \begin{bmatrix} \mathbf{H}^{(0)} & 0 & 0 & 0 \\ 0 & \mathbf{H}^{(1)} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \mathbf{H}^{(N_{\Omega})} \end{bmatrix}, \\
\dot{\mathbf{X}} &= \begin{bmatrix} \dot{\mathbf{X}}^{(0)} \\ \dot{\mathbf{X}}^{(1)} \\ \vdots \\ \dot{\mathbf{X}}^{(N_{\Omega})} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \mathbf{X}^{(0)} \\ \mathbf{X}^{(1)} \\ \vdots \\ \mathbf{X}^{(N_{\Omega})} \end{bmatrix}, \quad \mathbf{\Pi} = \begin{bmatrix} \mathbf{\Pi}^{(0)} \\ \mathbf{\Pi}^{(1)} \\ \vdots \\ \mathbf{\Pi}^{(N_{\Omega})} \end{bmatrix}, \quad \Theta_{ext} = \begin{bmatrix} \Theta_{ext}^{(0)} \\ \Theta_{ext}^{(1)} \\ \vdots \\ \Theta_{ext}^{(N_{\Omega})} \end{bmatrix}. \quad (2.4.25)
\end{aligned}$$

La formulation matricielle du problème défini en 4.2.1 s'écrit alors :

$$\begin{bmatrix} \mathbf{C} & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} \dot{\mathbf{X}} \\ \dot{\mathbf{X}}_{\varphi} \end{pmatrix} + \begin{bmatrix} \mathbf{K} + \mathbf{H} & \mathbf{J} \\ {}^t \mathbf{J} & \mathbf{M}_c \end{bmatrix} \begin{pmatrix} \mathbf{X} \\ \Phi \end{pmatrix} = \begin{pmatrix} \mathbf{\Pi} + \Theta_{ext} \\ 0 \end{pmatrix}. \quad (2.4.26)$$

Similairement à la section 4.1.4 page 55, on peut choisir pour la résolution d'utiliser comme fonctions $\{\psi_i^{(k)}\}_{1 \leq i \leq N^{(k)}}$ des fonctions de forme associées aux maillages des $\Omega^{(k)}$.

La famille $\{\phi_{p_m}^{\{k,j\}}\}_{1 \leq m \leq N_p^{\{k,j\}}}$ est alors la base de fonctions de l'espace mortier discrétisé associé à $\Gamma_p^{\{k,j\}}$ (calculées du côté non-mortier). Mais on verra dans le chapitre suivant qu'il est aussi possible d'utiliser comme fonctions de base $\{\psi_i^{(k)}\}_{1 \leq i \leq N^{(k)}}$ les modes de Dirichlet-Steklov; dans ce cas, il faudra spécifier une nouvelle famille de fonctions de type mortier.

4.2.3.1 Cas d'une carte électronique

On se place dans le cas d'une carte électronique, comme représentée sur la figure 2.4.4 et avec le découpage conforme décrit dans le paragraphe précédent. Les mortiers sont choisis

du côté du PCB et les non-mortiers du côté des composants. Par convention, on désignera le PCB par l'indice zéro ($\Omega^{(0)}$), et les IC seront désignés par $\{\Omega^{(k)}\}_{1 \leq k \leq N_\Omega}$. Alors les matrices \mathbf{J} et \mathbf{M}_c se simplifient. En effet, on définit dans ce cas pour $1 \leq k \leq N_\Omega$, $1 \leq p \leq N_{\Gamma^{0,k}}$ tels que $\Gamma_p^{\{0,k\}} \in \Gamma_{\square}$

$$\mathbf{J}_p^{(0,k)} = s_{g_p}^{k,0} \left(\int_{\Gamma_p^{\{k,0\}}} \psi_i^{(0)} \phi_{p_m}^{\{k,0\}} \right)_{\substack{1 \leq i \leq N^{(0)} \\ 1 \leq m \leq N_{\phi}^{\Gamma_p^{\{k,0\}}}}, \quad \mathbf{J}_p^{(k,k)} = s_{g_p}^{k,0} \left(\int_{\Gamma_p^{\{k,0\}}} \psi_i^{(k)} \phi_{p_m}^{\{k,0\}} \right)_{\substack{1 \leq i \leq N^{(k)} \\ 1 \leq m \leq N_{\phi}^{\Gamma_p^{\{k,0\}}}}, \quad (2.4.27)$$

$$\mathbf{M}_{c_p}^{(k)} = - \left(\int_{\Gamma_p^{\{k,0\}}} R_{c_p}^{(k,0)} \phi_{p_i}^{\{k,0\}} \phi_{p_m}^{\{k,0\}} \right)_{\substack{1 \leq i \leq N_{\phi}^{\Gamma_p^{\{k,0\}}} \\ 1 \leq m \leq N_{\phi}^{\Gamma_p^{\{k,0\}}}}. \quad (2.4.28)$$

On définit alors les matrices rectangulaires

$$\mathbf{J}^{(0,k)} = \left[\mathbf{J}_1^{(0,k)}, \dots, \mathbf{J}_{N_{\Gamma}^{\{0,k\}}}^{(0,k)} \right], \quad \mathbf{J}^{(k,k)} = \left[\mathbf{J}_1^{(k,k)}, \dots, \mathbf{J}_{N_{\Gamma}^{\{0,k\}}}^{(k,k)} \right] \quad (2.4.29)$$

et la matrice bloc diagonale

$$\mathbf{M}_c^{(k)} = \begin{bmatrix} \mathbf{M}_{c_1}^{(k)} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{M}_{c_{N_{\Gamma}^{\{0,k\}}}}^{(k)} \end{bmatrix} \quad (2.4.30)$$

pour $1 \leq k \leq N_\Omega$.

Les matrices \mathbf{J} et \mathbf{M}_c s'écrivent alors :

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}^{(0,1)} & \mathbf{J}^{(0,2)} & \dots & \mathbf{J}^{(0,N_\Omega)} \\ \mathbf{J}^{(1,1)} & 0 & 0 & 0 \\ 0 & \mathbf{J}^{(2,2)} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \mathbf{J}^{(N_\Omega,N_\Omega)} \end{bmatrix}, \quad \mathbf{M}_c = \begin{bmatrix} \mathbf{M}_c^{(1)} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{M}_c^{(N_\Omega)} \end{bmatrix}. \quad (2.4.31)$$

4.3 Premier bilan sur la méthode de sous-structuration avec joints

L'équation (2.4.26) nous donne une formulation du problème thermique utilisant la méthode des joints sur un maillage éléments finis. Bien qu'on n'ait pas encore la formulation

réduite qui nous intéresse, il est indispensable d'avoir une formulation complète, pour évaluer la réduction en comparant le champ réduit au champ détaillé. La méthode de réduction d'amalgame modal, décrite plus loin, nécessite aussi un champ détaillé. On implémente donc cette méthode sur SAMBA, le code du laboratoire destiné aux simulations thermiques réduites¹.

Pour évaluer le champ réduit, on pourrait cependant obtenir les champs détaillés via un logiciel tiers et se passer de cette première implémentation. Mais en procédant ainsi, il est plus aisé, et surtout plus légitime, d'estimer le gain en temps de calcul du modèle réduit. En effet, les modèles détaillé et réduit seront résolus avec le même ensemble d'outils de calcul et sur la même machine, la puissance de la machine n'interviendra donc pas dans le gain en temps, qui sera ainsi « absolu ».

La difficulté de l'implémentation de la méthode des joints réside dans l'évaluation de la matrice \mathbf{J} . D'une part, par le choix des fonctions mortiers $\{\phi_m^{\{k,l\}}\}_m$, qui, on le rappelle, sont définies sur chaque non-mortier $\Gamma^{\{k,l\}} \in \Gamma_{\square}$. De manière similaire à l'exemple 2D donné en section 2.5.3 page 30, on peut utiliser les fonctions de forme définies sur l'interface. D'autres choix sont possibles, on les détaillera dans la partie de ce manuscrit dédiée à la mise en œuvre de la méthode et aux applications. D'autre part, pour chaque interface, on a deux projections à calculer : du côté mortier et du côté non-mortier. Du côté non-mortier, cette projection est simple puisque les fonctions mortiers sont définies sur le maillage de ce côté. De l'autre côté en revanche, puisqu'on n'impose aucune condition sur la conformité des maillages à l'interface, la projection doit se faire sur des fonctions définies sur un maillage différents de celui des fonctions mortier. Une manière simple de contourner ce problème et de calculer ces projections à l'aide d'une formule de quadrature sur un troisième maillage, sans épaisseur, constitué de quadrilatères sur lequel on peut appliquer la formule des trapèzes, comme illustré sur la figure 2.4.5.

1. S'il n'a fallu que quelques secondes pour écrire cette phrase ... l'implémentation concrète des méthodes utilisées pour ces recherches a représenté une part significative du travail de thèse (et qu'il est difficile de retranscrire dans le manuscrit).

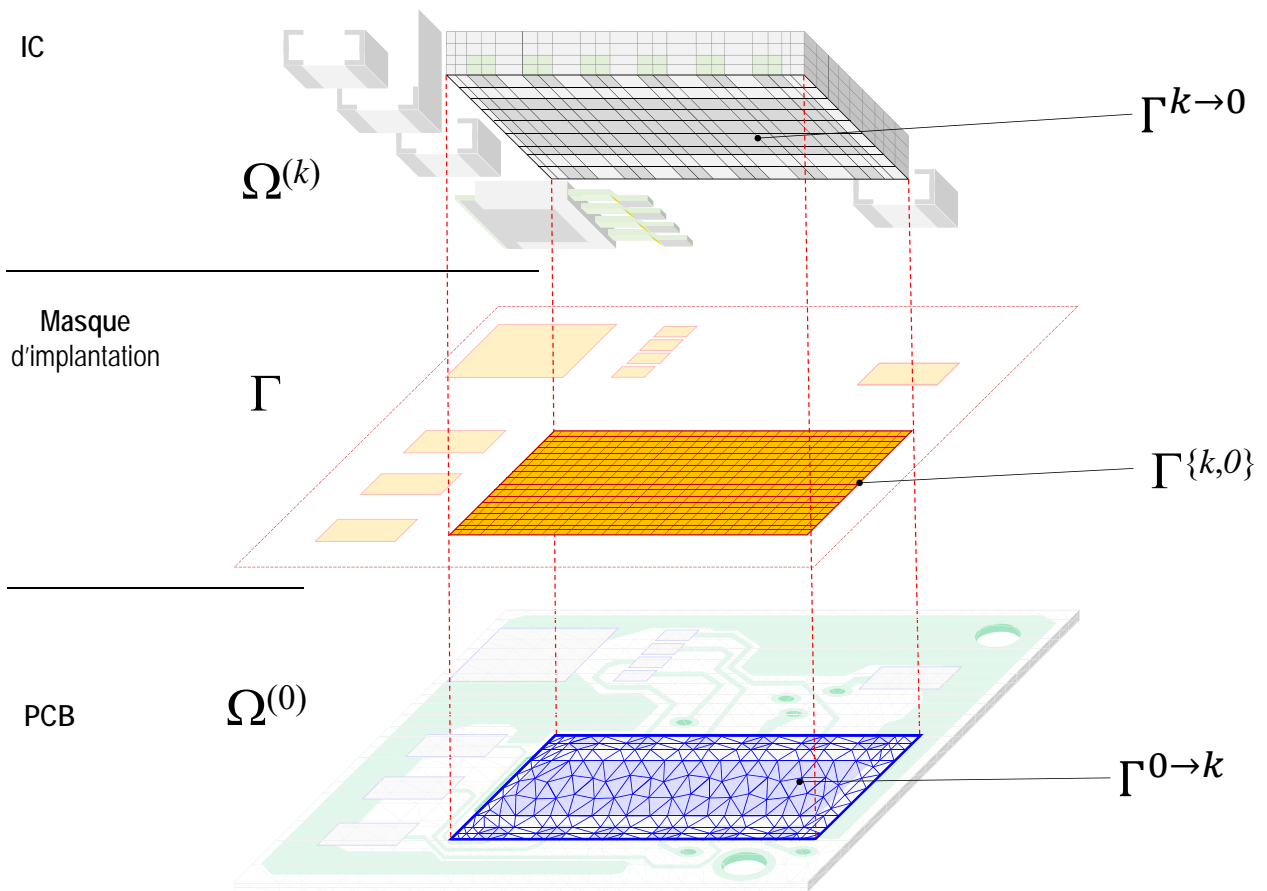


FIGURE 2.4.5 – Maillage du contact entre le composant $\Omega^{(k)}$ et le circuit imprimé $\Omega^{(0)}$.

Chapitre 5

Adaptation de la méthode de réduction modale à la sous-structuration par joints

On va dans un premier temps définir les bases de Dirichlet et de Steklov. Nous montrerons ensuite que leur réunion forme une base de l'espace fonctionnel $H^1(\Omega)$, qui est l'espace de travail approprié pour le problème thermique. Nous construirons un produit scalaire adapté à cette base. La norme induite va permettre de quantifier la distance entre deux champs de température. Ceci ouvrira la voie à la réduction de la base par amalgame modal .

5.1 Définition de la base utilisée

5.1.1 Base de Dirichlet

5.1.1.1 Définition

Définition 5.1.1 *La base de Dirichlet est définie par l'ensemble des couples $(\lambda^D, V^D) \in \mathbb{R}^+ \times H_0^1(\Omega)$ solutions du problème aux valeurs propres suivant :*

$$\begin{cases} \Omega & -\nabla \cdot (K \cdot \nabla V^D) = \lambda^D C V^D, \\ \partial\Omega & V^D = 0. \end{cases} \quad (2.5.1)$$

On appelle modes de Dirichlet les vecteurs propres V^D solutions de ce problème.

Les valeurs propres λ^D , associées à chaque vecteur propre V^D , sont homogènes à des fréquences, et leurs inverses sont les temps caractéristiques des modes propres associés, tandis que les modes sont des champs de température correspondant au problème thermique stationnaire, homogène sur la frontière.

5.1.1.2 Formulation variationnelle

Définition 5.1.2 *La formulation faible du problème aux valeurs propres (2.5.1) est :*

Trouver $(\lambda^D, V^D) \in \mathbb{R}^+ \times H_0^1(\Omega)$ tels que

$$\int_{\Omega} \nabla V^D \cdot K \cdot \nabla u = \lambda^D \int_{\Omega} C V^D u, \quad (2.5.2)$$

quelque soit $u \in H_0^1(\Omega)$.

5.1.1.3 Opérateurs bilinéaires associés

À partir de la formulation variationnelle (2.5.2), on peut définir, pour u et $v \in H_0^1(\Omega)$, les opérateurs bilinéaires suivants :

$$\forall u, v \in H_0^1(\Omega), \quad \mathcal{A}_{\Omega}(u, v) = \int_{\Omega} \nabla u \cdot K \cdot \nabla v, \quad \mathcal{C}_{\Omega}(u, v) = \int_{\Omega} C u v. \quad (2.5.3)$$

Proposition 5.1.1 *L'ensemble des modes de Dirichlet $\{V_i^D\}_{i \in \mathbb{N}}$ forme une base infinie mais dénombrable de l'espace de Hilbert $H_0^1(\Omega)$.*

Preuve 5.1.1 *C'est un résultat classique d'analyse fonctionnelle [90],[91], [92]. On rappelle que l'opérateur $\mathcal{A}_{\Omega}(\cdot, \cdot)$ est coercif dans $H_0^1(\Omega)$.*

□

On appellera cette base la “base de Dirichlet”. Ces modes permettent de bien reconstruire la température dans le domaine Ω quand il est excité par une puissance interne, mais pas sur sa frontière $\partial\Omega$.

On a représenté un exemple de champ de température reconstruit à l'aide des modes de Dirichlet sur la figure 2.5.1. On considère une plaque, en 2D, et on applique une source de chaleur; on impose une condition de Dirichlet aux frontières (voir figure 2.5.1a). On peut alors reconstruire le champ de température considéré (figure 2.5.1b) en combinant les modes successifs de Dirichlet, pondérés par les états adéquats (figure 2.5.1c). Si on prend assez de modes, on obtient ainsi un champ très proche du champ de référence. Mais l'utilisation des modes de Dirichlet nous empêche de reconstruire un champ aux frontières, puisque, par définition, ils sont nuls sur ces frontières. On a donc besoin de modes supplémentaires pour avoir une décomposition complète d'un champ T quelconque : on choisit d'utiliser les modes de Steklov.

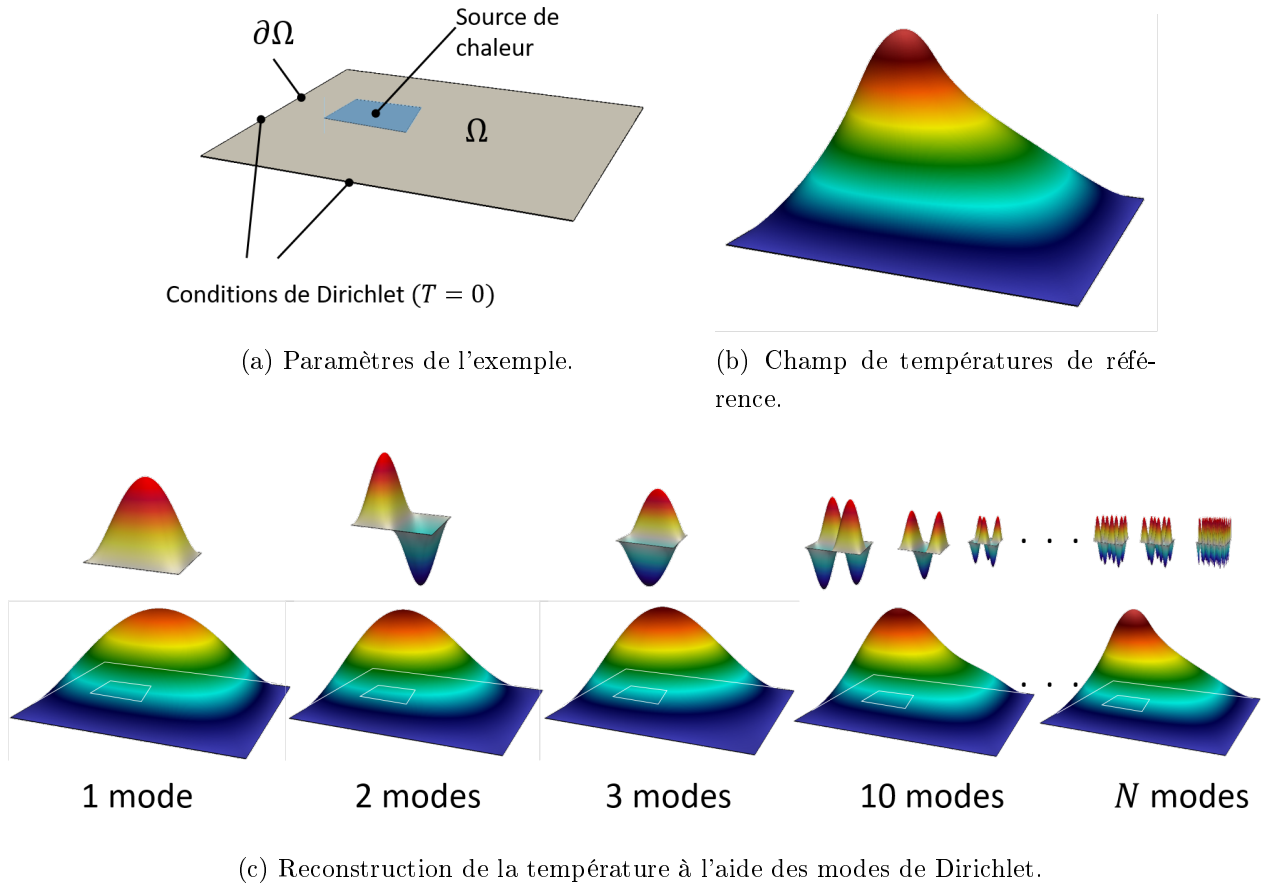


FIGURE 2.5.1 – Exemple de champ de température reconstruit à l'aide de modes de Dirichlet.

5.1.2 Base de Steklov

5.1.2.1 Définition

Définition 5.1.3 La base de Steklov est définie par l'ensemble des couples $(\mu^S, S) \in \mathbb{R}^+ \times H^{1/2}(\partial\Omega)$ solutions du problème aux valeurs propres suivant [93] :

$$\begin{cases} \Omega & -\nabla \cdot (K \cdot \nabla V^S) = 0, \\ \partial\Omega & K \cdot \nabla V^S \cdot \underline{n} = \mu^S h_S \zeta(x) S, \quad V^S|_{\partial\Omega} = S. \end{cases} \quad (2.5.4)$$

S est le mode propre de Steklov, mais par abus de langage, on appellera modes de Steklov les fonctions $V^S \in H^1(\Omega)$, qui sont les relèvements harmoniques de S dans Ω .

La valeur propre μ_S est sans dimension et h_S est un paramètre homogène à un coefficient d'échange ($[h_S] = W.m^{-2}.K^{-1}$). On rajoute une fonction pouvant dépendre de l'espace $\zeta(x)$.¹

1. Pour donner une expression de la valeur propre de Steklov de dimension analogue à la valeur propre de

Les modes de Steklov sont des champs stationnaires pour un flux imposé à la frontière, dont la valeur est proportionnelle à la température sur cette frontière.

5.1.2.2 Formulation variationnelle

La formulation faible du problème aux valeurs propres (2.5.4) est :

Trouver $(\mu^S, S) \in \mathbb{R}^+ \times H^{1/2}(\Omega)$ tels que

$$\begin{cases} V^S|_{\partial\Omega} = S, \\ \int_{\Omega} \nabla V^S \cdot K \cdot \nabla u = \mu^S h_s \int_{\partial\Omega} \zeta(x) V^S|_{\partial\Omega} u|_{\partial\Omega}. \end{cases} \quad (2.5.5)$$

5.1.2.3 Opérateur bilinéaire associé

À partir de la formulation variationnelle (2.5.5), on peut définir, pour u et $v \in H^1(\Omega)$, l'opérateur bilinéaire suivant :

$$\forall u, v \in H^1(\Omega), c_{\partial\Omega}(u, v) = \int_{\partial\Omega} \zeta(x) u|_{\partial\Omega} v|_{\partial\Omega}. \quad (2.5.6)$$

Définition 5.1.4 *Le problème variationnel aux valeurs propres de Steklov est défini par :*

$$\text{Trouver } (\mu^S, V^S) \in \mathbb{R}^+ \times H^1(\Omega) \text{ tels que quelque soit } v \in H^1(\Omega) \quad (2.5.7)$$

$$\mathcal{A}_{\Omega}(V^S, v) = \mu^S h_s c_{\partial\Omega}(V^S, v).$$

Proposition 5.1.2 *L'ensemble des modes de Steklov $\{S_i\}_{i \in \mathbb{N}}$ forme une base infinie mais dénombrable de l'espace de Hilbert $H^{1/2}(\partial\Omega)$.*

Dirichlet, on peut définir comme nouvelle valeur propre de Steklov $\lambda_S = \frac{\mu^S}{\tau_0}$, avec $\tau_0 = \frac{C_0 L_0^2}{K_0}$; λ^S est ainsi homogène à une fréquence. Les quantités C_0 et K_0 sont des valeurs de référence choisies parmi les paramètres du système. La longueur de référence L_0 est définie par $L_0 = \frac{\mathcal{Vol}(\Omega)}{\mathcal{S}(\Omega)}$, avec $\mathcal{Vol}(\Omega)$ le volume du domaine Ω et

$\mathcal{S}(\Omega)$ l'aire totale de $\partial\Omega$. On pose alors $h_s = \frac{K_0}{L_0}$. Cela revient à choisir un nombre de Biot $Bi = \frac{h_s L_0}{K_0} = 1$.

La condition aux limites de l'équation (2.5.4) se réécrit alors :

$$K \cdot \nabla V^S \cdot \underline{n} = \lambda^S \zeta_0(x) S,$$

avec $\zeta_0 = C_0 L_0 \zeta(x)$. On retrouve ainsi la condition aux limites des modes de Branches (voir par exemple [82]).

Preuve 5.1.2 La démonstration de cette proposition repose sur la définition d'un opérateur auxiliaire $b_\varepsilon(\cdot, \cdot)$ à $\mathcal{A}_\Omega(\cdot, \cdot)$, que l'on démontre être bilinéaire, symétrique, continu et coercif. L'opérateur auxiliaire possède les mêmes modes propres que ceux de l'opérateur $\mathcal{A}_\Omega(\cdot, \cdot)$ et ses valeurs propres μ_ε sont décalées vers la droite par rapport à celles de $\mathcal{A}_\Omega(\cdot, \cdot)$ notées μ (concrètement on a $\mu_k = \mu_{\varepsilon k} - \varepsilon$). On donne la preuve complète de cette proposition en annexe.

□

On appellera cette base la “base de Steklov”.

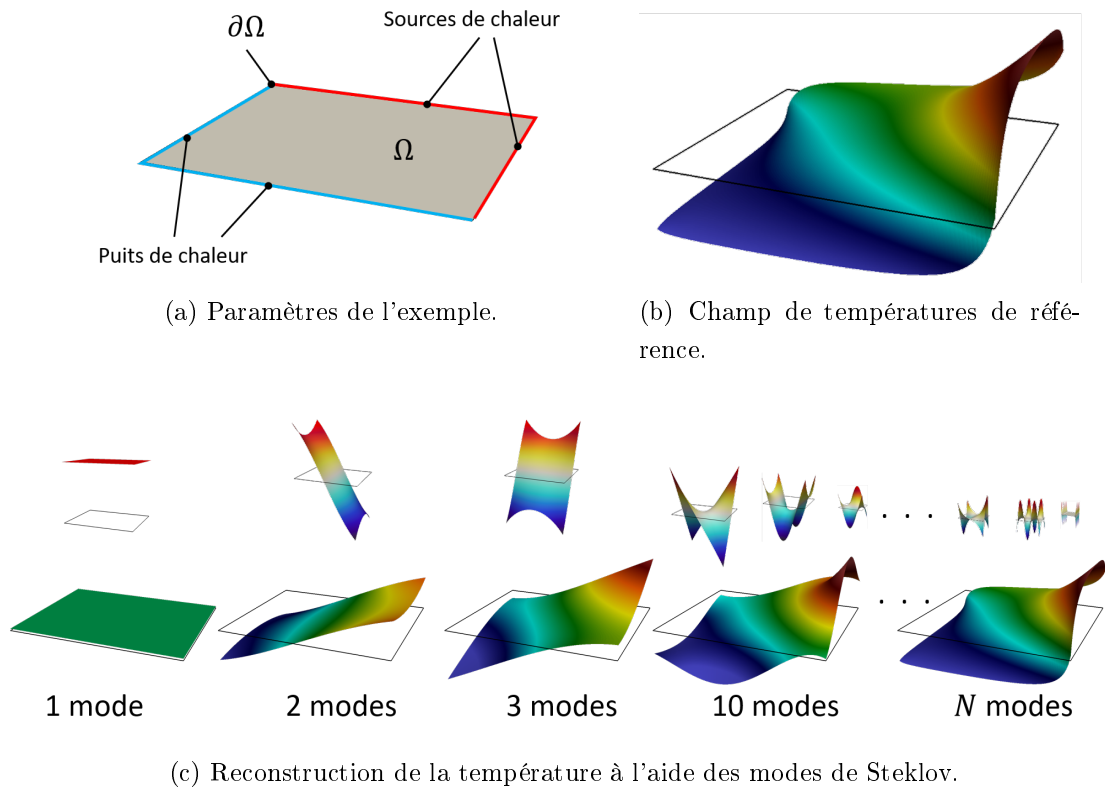


FIGURE 2.5.2 – Exemple de champ de température reconstruit à l'aide de modes de Steklov.

On a représenté un exemple de champ de température reconstruit à l'aide des modes de Steklov sur la figure 2.5.2. On considère toujours une plaque, en 2D, mais cette fois on n'applique pas de source de chaleur dans le domaine. En revanche, on applique une source de chaleur sur deux arêtes de la plaque et un puits de chaleur sur les deux autres (voir figure 2.5.2a). On obtient ainsi un champ asymétrique (figure 2.5.2b), que l'on reconstruit en combinant les modes successifs de Steklov, pondérés par les états adéquats (figure 2.5.1c). Si

on prend assez de modes, on obtient un champ très proche du champ de référence. L'idée est donc d'utiliser ces deux bases conjointement pour décomposer la température.

5.1.3 Base de Dirichlet-Steklov

Définition 5.1.5 Soit $\{V_i^{\mathcal{X}}\}_{i \in \mathbb{N}}$ la réunion des modes propres de Dirichlet $\{V_i^D\}_{i \in \mathbb{N}}$ et des modes de Steklov $\{V_i^S\}_{i \in \mathbb{N}}$. On appelle cette réunion "base de Dirichlet-Steklov".

5.2 Propriétés de la base de Dirichlet-Steklov

5.2.1 Définition d'un nouveau produit scalaire sur $H^1(\Omega)$

Définition 5.2.1 Soient $u, v \in H^1(\Omega)$. On définit le produit scalaire $(u|v)_{H(\Omega)}$ par

$$\begin{aligned} (u|v)_{H(\Omega)} &= \int_{\Omega} \nabla u \cdot K \cdot \nabla v + h_0 \int_{\partial\Omega} \zeta(x) uv \\ &= \mathcal{A}_{\Omega}(u, v) + h_0 c_{\partial\Omega}(u, v), \end{aligned} \quad (2.5.8)$$

avec $h_0 > 0$ un facteur de pondération homogène à un coefficient d'échange.

La norme induite par ce produit scalaire s'écrit

$$\forall u \in H^1(\Omega), \quad \|u\|_{H(\Omega)} = \sqrt{\int_{\Omega} \nabla u \cdot K \cdot \nabla u + h_0 \int_{\partial\Omega} \zeta(x) (u)^2}. \quad (2.5.9)$$

Proposition 5.2.1 La norme $\|\cdot\|_{H(\Omega)}$ est équivalente à la norme usuelle dans $H^1(\Omega)$

On rappelle que la norme usuelle dans $H^1(\Omega)$ s'écrit :

$$\forall u \in H^1(\Omega), \quad \|u\|_{H^1(\Omega)} = \sqrt{\|\nabla u\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\Omega)}^2}. \quad (2.5.10)$$

Preuve 5.2.1 Pour démontrer l'équivalence de ces deux normes, la preuve repose sur l'utilisation de l'inégalité de Poincaré, de l'inégalité de la trace et de l'identité remarquable $(a+b)^2 \leq 2(a^2 + b^2)$. Le développement complet de cette preuve est donné en annexe.

□

On peut ainsi normaliser les bases de Dirichlet et Steklov à l'aide de $\|\cdot\|_{H(\Omega)}$.

5.2.2 Propriétés d'orthogonalité

Les propriétés d'orthogonalité sont essentielles dans les méthodes modales, puisque ce sont grâce à celles-ci que l'on peut démontrer que la décomposition de T sur ces bases existe et est unique.

Ce nouveau produit scalaire sur $H^1(\Omega)$ permet ainsi d'établir des relations d'orthogonalité non seulement pour les modes de Steklov, mais aussi pour les modes de Dirichlet. Ces relations entre les bases sont fondamentales, et c'est seulement quand on établira une relation d'orthogonalité entre modes de Dirichlet et modes de Steklov qu'on pourra parler de base de Dirichlet-Steklov.

Proposition 5.2.2 *Si $\|V_i^D\|_{H(\Omega)} = 1$, alors les modes de Dirichlet vérifient les relations d'orthogonalité :*

$$\begin{aligned} \forall i, j \in \mathbb{N}, \quad (V_i^D | V_j^D)_{H(\Omega)} &= \delta_i^j, \\ \mathcal{A}_\Omega(V_i^D, V_j^D) &= \delta_i^j, \\ \mathcal{C}_\Omega(V_i^D, V_j^D) &= \frac{\delta_i^j}{\lambda_i^D}. \end{aligned} \tag{2.5.11}$$

Preuve 5.2.2 *Ces propriétés se démontrent classiquement en soustrayant les équations que vérifient deux modes propres V_i^D et V_j^D . On développera en annexe.*

Proposition 5.2.3 *Si $\|V_i^S\|_{H(\Omega)} = 1$, alors les modes de Steklov vérifient les relations d'orthogonalité :*

$$\begin{aligned} \forall i, j \in \mathbb{N}, \quad (V_i^S | V_j^S)_{H(\Omega)} &= \delta_i^j, \\ \mathcal{A}_\Omega(V_i^S, V_j^S) &= \delta_i^j \frac{\mu_i^S}{\frac{h_0}{h_s} + \mu_i^S}, \\ c_{\partial\Omega}(V_i^S, V_j^S) &= \frac{\delta_i^j}{\frac{h_0}{h_s} + \mu_i^S}. \end{aligned} \tag{2.5.12}$$

Preuve 5.2.3 *Idem que précédemment, voir en annexe.*

On a énoncé les propriétés d'orthogonalité pour les modes de Dirichlet et les modes de Steklov. Il reste maintenant à établir les propriétés d'orthogonalité entre ces deux bases.

Proposition 5.2.4 *Les modes de Dirichlet $\{V_i^D\}_{i \in \mathbb{N}}$ et de Steklov $\{V_i^S\}_{i \in \mathbb{N}}$ vérifient les deux relations d'orthogonalité suivantes :*

$$\begin{aligned} \forall i, j \in \mathbb{N}, \quad (V_i^D | V_j^S)_{H(\Omega)} &= 0, \\ \mathcal{A}_\Omega(V_i^D, V_j^S) &= 0. \end{aligned} \tag{2.5.13}$$

Preuve 5.2.4 Cette preuve est assez directe, puisqu'elle découle directement des définitions de $(u|v)_{H(\Omega)}$ et des modes de Dirichlet et de Steklov. En effet, soit V_j^S un mode de Steklov. On a, $\forall u \in H^1(\Omega)$,

$$(V_j^S|u)_{H(\Omega)} = \mathcal{A}_\Omega(V_j^S, u) + h_0 c_{\partial\Omega}(V_j^S, u) = (h_s \mu_j^S + h_0) c_{\partial\Omega}(V_j^S, u)$$

par définition des modes de Steklov (2.5.7). De plus, on a

$$\forall i, j \in \mathbb{N} \quad c_{\partial\Omega}(V_i^D, V_j^S) = \int_{\partial\Omega} \zeta(x) V_i^D|_{\partial\Omega} S_j = 0$$

puisque $V_i^D \in H_0^1(\Omega)$. En remplaçant u par V_i^D et en remarquant que $(\cdot|\cdot)_{H(\Omega)}$ est symétrique, on a

$$\begin{aligned} \forall i, j \in \mathbb{N}, \quad (V_i^D|V_j^S)_{H(\Omega)} &= (h_s \mu_j^S + h_0) c_{\partial\Omega}(V_i^D, V_j^S) = 0, \\ \mathcal{A}_\Omega(V_i^D, V_j^S) &= h_s \mu_j^S c_{\partial\Omega}(V_i^D, V_j^S) = 0. \end{aligned}$$

□

En revanche, il n'y a pas d'orthogonalité L^2 entre ces deux familles, que l'on considère Ω (ce qui est vrai pour le cas des modes de Dirichlet seuls) ou bien encore $\bar{\Omega}$. Il existe une orthogonalité pour le produit scalaire sur $\partial\Omega$ puisque les traces des modes de Dirichlet sont nulles, ce que l'on vient de voir dans cette preuve. Cette orthogonalité peut être éventuellement invoquée pour déterminer l'état du mode "plat" par projection, car ce dernier étant associé à une valeur propre nulle il présente une singularité.

Des propositions 5.2.2, 5.2.3 et 5.2.4, on déduit directement la propriété d'orthogonalité des modes de la base de Dirichlet-Steklov :

Proposition 5.2.5 Les bases de Dirichlet et Steklov sont orthogonales vis à vis du produit scalaire défini par (2.5.8). On peut normaliser ces bases à l'aide de la norme induite par (2.5.9).

$$\begin{aligned} \forall \mathcal{X}, \mathcal{Y} \in \{D, S\}, \forall i, j \in \mathbb{N} \text{ tels que } \|V_i^{\mathcal{X}}\|_{H(\Omega)} &= 1, \\ (V_i^{\mathcal{X}}|V_j^{\mathcal{Y}})_{H(\Omega)} &= \delta_{\mathcal{X}}^{\mathcal{Y}} \delta_i^j. \end{aligned} \tag{2.5.14}$$

Ces relations d'orthogonalité, et en particulier cette dernière, vont ainsi permettre d'affirmer que la réunion des modes propres de Dirichlet et de Steklov ($\{V_i^D\}_{i \in \mathbb{N}}, \{S_j\}_{j \in \mathbb{N}}$) forme une base hilbertienne de $H^1(\Omega)$. Grâce à cela, on va pouvoir projeter un champ de température sur cette base pour en trouver les coefficients de décomposition.

On énonce finalement le résultat fondamental suivant :

Proposition 5.2.6 *La réunion des bases de Dirichlet et de Steklov $(\{V_i^D\}, \{S_j\})_{i,j \in \mathbb{N}}$ (respectivement $\{V_i^D\}_{i \in \mathbb{N}} \oplus \{V_j^S\}_{j \in \mathbb{N}}$) forme une base hilbertienne de $H_0^1(\Omega) \times H^{1/2}(\partial\Omega)$ (respectivement de $H_0^1(\Omega) \oplus \mathbb{E}(\Omega) = H^1(\Omega)$, où $\mathbb{E}(\Omega) \subset H^1(\Omega)$ représente l'espace des relèvements harmoniques). En particulier :*

$$\exists x_i^D, x_j^S \text{ tels que } \forall u \in H^1(\Omega), \quad u = \sum_{\mathcal{X} \in \{D,S\}} \sum_{i=1}^{\infty} x_i^{\mathcal{X}} V_i^{\mathcal{X}}, \quad (2.5.15)$$

avec $x_i^{\mathcal{X}} = (u|V_i^{\mathcal{X}})_{H(\Omega)}$.

Preuve 5.2.5 *Pour démontrer cette proposition, on aura d'abord besoin de démontrer que les relèvements harmoniques des modes de Steklov forment une base hilbertienne de $\mathbb{E}(\Omega)$ pour le produit scalaire $(\cdot|\cdot)_{H(\Omega)}$. On utilise ensuite les propositions 5.1.1 et 5.1.2 pour en déduire la proposition. On renvoie une fois de plus à la démonstration complète en annexe.*

Ceci nous assure qu'on peut décomposer T sur la base de Dirichlet-Steklov, et ainsi avoir une formulation modale réduite de (2.4.26).

En reprenant et en combinant les exemples donnés pour les modes de Dirichlet (figure 2.5.1) et de Steklov (figure 2.5.2), on obtient un champ de température qui s'exprime à la fois sur Ω et sur $\partial\Omega$. On peut alors le reconstruire en combinant les mêmes modes de Dirichlet et Steklov utilisés dans ces exemples, comme le montre la figure 2.5.3.

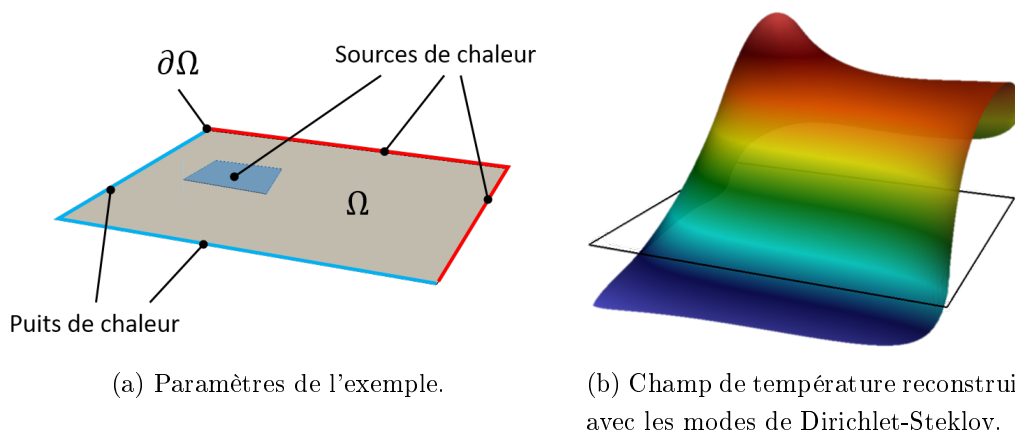


FIGURE 2.5.3 – Exemple de champ de température reconstruit à l'aide de modes de Dirichlet-Steklov.

5.3 Formulation matricielle des bases de Dirichlet et Steklov

La résolution des problèmes aux valeurs propres n'est pas envisageable par voie analytique, sauf dans de rares cas académiques. Pour des géométries complexes, les modes de la base de Dirichlet-Steklov peuvent être calculés à partir d'une discrétisation par éléments finis comme celle introduite au chapitre 4. Il est ainsi possible d'accéder à une expression approchée des modes sur chaque sous-domaine, indépendamment des autres.

Dans la suite, on ne décrit donc le processus de discrétisation des problèmes aux valeurs propres que sur un domaine Ω . On considère une base de fonctions éléments finis $\{\psi_i\}_{1 \leq i \leq N}$, où N représente le nombre de nœuds du maillage. L'espace d'approximation est noté $V_h(\Omega) \subset H^1(\Omega)$, le paramètre h caractérisant la finesse du maillage.

5.3.1 Modes de Dirichlet.

On décompose les modes de Dirichlet sur la base des fonctions éléments finis :

$$V^D = \sum_{i=1}^N v_i^D \psi_i, \quad (2.5.16)$$

où $v_i^D = V^D(x_i)$ est la valeur du mode de Dirichlet au nœud m_i du maillage. En substituant cette décomposition dans (2.5.2) et en choisissant comme fonctions d'essai v les fonctions de forme ψ_i on a :

$$\sum_{i=1}^N v_i^D \int_{\Omega} \nabla \psi_i \cdot K \cdot \nabla \psi_j = \lambda^D \sum_{i=1}^N v_i^D \int_{\Omega} c \psi_i \psi_j. \quad (2.5.17)$$

Les valeurs aux nœuds situés sur la frontière étant nulles, il faut restreindre les sommes aux nœuds internes. On note N_0 le nombre de nœuds internes. On introduit donc les matrices :

$$\mathbf{K}_{\Omega} = \left(\int_{\Omega} \nabla \psi_i \cdot K \cdot \nabla \psi_j \right)_{\substack{1 \leq i, j \leq N_0 \\ m(x_i) \in \Omega}}, \quad \mathbf{C}_{\Omega} = \left(\int_{\Omega} c \psi_i \psi_j \right)_{\substack{1 \leq i, j \leq N_0 \\ m(x_i) \in \Omega}},$$

$$V_0^D = (v_i^D)_{1 \leq i \leq N_0}.$$

Le problème aux valeurs propres de Dirichlet discrétisé s'écrit sous la forme matricielle :

$$\mathbf{K}_{\Omega} V_0^D = \lambda^D \mathbf{C}_{\Omega} V_0^D. \quad (2.5.18)$$

Les matrices \mathbf{K}_Ω et \mathbf{C}_Ω sont symétriques et définies positives. On peut réécrire les relations d'orthogonalité, en complétant les vecteurs propres par des zéros aux nœuds situés sur le bord :

$$\forall i, j \in \mathbb{N}, i \neq j, \quad {}^t\mathbf{V}_i^D \mathbf{K} \mathbf{V}_j^D = {}^t\mathbf{V}_i^D \mathbf{C} \mathbf{V}_j^D = 0. \quad (2.5.19)$$

5.3.2 Modes de Steklov.

On décompose les modes de Steklov sur la base des fonctions éléments finis :

$$V^S = \sum_{i=1}^N v_i^S \psi_i, \quad (2.5.20)$$

où $v_i^S = V^S(x_i)$ est la valeur du mode de Steklov au nœud m_i du maillage. En substituant cette décomposition dans (2.5.5) et en choisissant comme fonctions d'essais v les fonctions de forme ψ_i , on trouve :

$$\sum_{i=1}^N v_i^S \int_{\Omega} \nabla \psi_i \cdot \mathbf{K} \cdot \nabla \psi_j = h_s \mu^S \sum_{i=1}^N v_i^S \int_{\partial\Omega} \zeta(x) \psi_i \psi_j. \quad (2.5.21)$$

Ici les nœuds situés sur la frontière doivent être conservés. On retrouve dans l'expression de gauche la matrice \mathbf{K} . On introduit une nouvelle matrice :

$$\mathbf{C}_{\partial\Omega} = \left(\int_{\partial\Omega} \zeta(x) \psi_i \psi_j \right)_{1 \leq i, j \leq N}, \quad \mathbf{V}^S = (v_i^S)_{1 \leq i, j \leq N}.$$

Cette matrice possède de nombreux termes nuls puisque l'intégrale ne porte que sur le bord du domaine (elle est singulière).

Le problème aux valeurs propres de Steklov pour l'opérateur de la chaleur discrétisé s'écrit sous la forme matricielle :

$$\mathbf{K} \mathbf{V}^S = h_s \mu^S \mathbf{C}_{\partial\Omega} \mathbf{V}^S. \quad (2.5.22)$$

Les matrices \mathbf{K} et $\mathbf{C}_{\partial\Omega}$ sont symétriques. On peut réécrire les relations d'orthogonalité :

$$\forall i, j \in \mathbb{N}, i \neq j, \quad {}^t\mathbf{V}_i^S \mathbf{K} \mathbf{V}_j^S = {}^t\mathbf{V}_i^S \mathbf{C}_{\partial\Omega} \mathbf{V}_j^S = 0. \quad (2.5.23)$$

On a de plus les relations d'orthogonalité :

$$\left. \begin{array}{l} \forall \mathcal{X}, \mathcal{Y} \in \{D, S\} \\ \forall i, j \in \mathbb{N}, i \neq j \end{array} \right\} \begin{array}{l} {}^t V_i^{\mathcal{Y}} \mathbf{K} V_j^{\mathcal{X}} = 0, \\ {}^t V_i^{\mathcal{Y}} \mathbf{C}_{\partial\Omega} V_j^{\mathcal{X}} = 0, \\ {}^t V_i^{\mathcal{Y}} \mathbf{K} V_j^{\mathcal{X}} + {}^t V_i^{\mathcal{Y}} \mathbf{C}_{\partial\Omega} V_j^{\mathcal{X}} = 0, \end{array} \quad (2.5.24)$$

la dernière relation étant l'expression du produit scalaire $(\cdot|\cdot)_{H(\Omega)}$ déclinée sous forme matricielle.

5.3.3 Exemples de modes

Les modes de Dirichlet et de Steklov ont des localisations spatiales très différentes. En effet, lorsque l'ordre augmente, les modes de Steklov se concentrent sur le bord du domaine. Ils ne sont pas excités par une sollicitation de puissance interne. À l'inverse, les modes de Dirichlet dont la trace est nulle sur le bord du domaine ne sont pas excités par un flux imposé en surface.

À titre d'exemple, on a calculé les modes de Dirichlet-Steklov pour un pavé droit de dimensions $2.4 \times 4 \times 2 \text{ mm}^3$. La figure 2.5.4 présente des modes de Dirichlet de différents ordres. Si les premiers sont très réguliers et symétriques, les approximations numériques des modes de Dirichlet perdent en qualité pour les ordres très élevés : si la discrétisation comporte N nœuds, on estime en général que seuls les $N/10$ premiers modes de Dirichlet sont retrouvés avec une excellente précision. Au delà, des artefacts numériques peuvent apparaître.

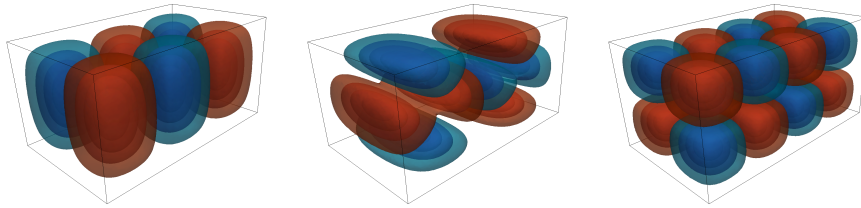


FIGURE 2.5.4 – Exemples de modes de Dirichlet calculés sur un pavé droit : les modes sont normalisés entre -1 et $+1$ pour la visualisation, et on a représenté l'isovaleur -0.2 en bleu et $+0.2$ en rouge. On note que pour les modes de Dirichlet, les isovaleurs sont fermées.

Les premiers modes de Steklov sont définis dans le volume et sur la frontière (figure 2.5.5). Toutefois, dès qu'on atteint l'ordre 10 (mode de droite sur la figure), on constate qu'ils se concentrent sur la frontière du pavé.

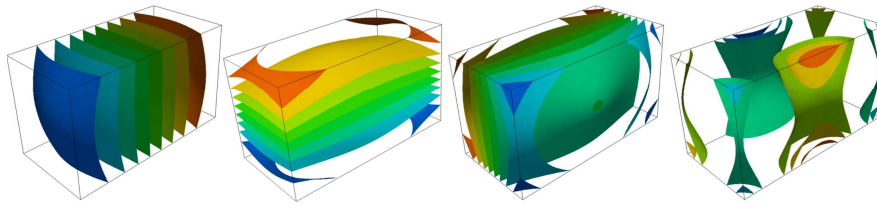


FIGURE 2.5.5 – Exemples des premiers modes de Steklov calculés sur un pavé droit : les modes sont aussi normalisés entre -1 et $+1$ pour la visualisation, et on représente dix isovaleurs.

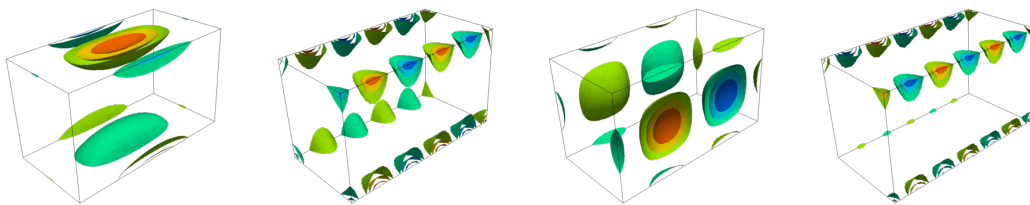


FIGURE 2.5.6 – Exemples de modes de Steklov d'ordre plus élevé calculés sur un pavé droit.

Si on calcule les modes de Steklov suivants et qu'on s'intéresse à des ordres plus élevés (figure 2.5.6), cette concentration sur la frontière est très visible. On commence aussi à remarquer que les modes de Steklov s'expriment soit principalement sur des faces opposées, soit sur des arêtes, montrant de fait l'existence de deux sous-familles de modes de Steklov en 3D : les modes de face et les modes d'arêtes.

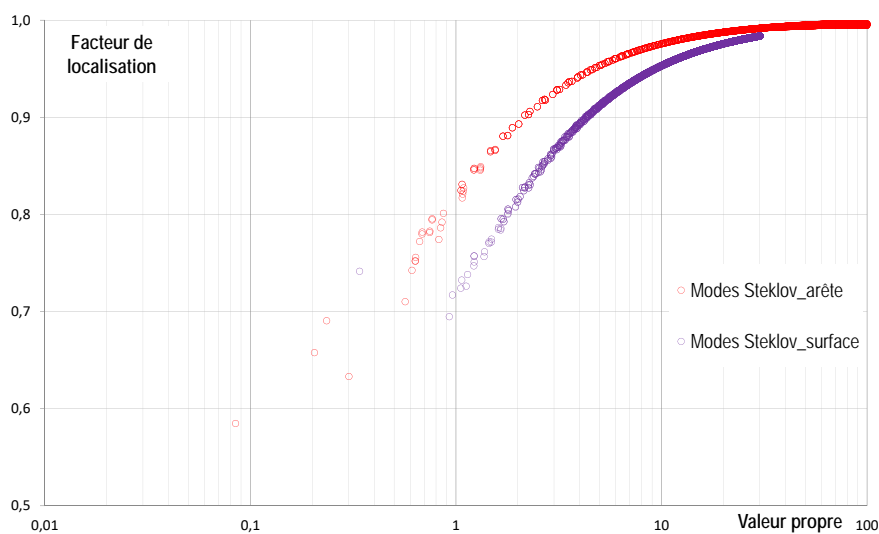


FIGURE 2.5.7 – Localisation des modes de Steklov pour le pavé droit.

L'existence² des deux sous-familles de modes de Steklov peut-être mise en évidence en calculant leur facteur de localisation, *i.e.* la proportion d'un mode qui s'exprime sur la surface.

Le facteur de localisation d'une fonction u , $Loc(u)$, est donné par la formule

$$Loc(u) = \frac{\int_{\partial\Omega} L_r u^2}{\int_{\partial\Omega} L_r u^2 + \int_{\Omega} u^2}. \quad (2.5.25)$$

La figure 2.5.7 présente ainsi les facteurs de localisation des modes de Steklov calculés pour ce même pavé droit en fonction de leur valeur propre, plus le facteur étant proche de 1, plus le mode s'exprimant à la frontière. Les deux catégories de modes sont alors clairement distinctes, avec en violet les modes de Steklov s'exprimant sur la surface et en rouge s'exprimant sur les arêtes. Ces deux familles se rapprochent asymptotiquement, et tendent à s'exprimer exclusivement sur la frontière du domaine et plus du tout dans le domaine.

5.4 Réduction de la base

On veut réduire le nombre de modes propres à conserver dans la décomposition modale du champ de température. En effet, la décomposition modale n'est utilisable que si nous limitons les sommes infinies à un nombre "faible" de termes. En pratique, on ne calcule que des approximations des modes propres, par exemple via une discrétisation en éléments finis en résolvant les problèmes matriciels définis dans la section précédente. Le nombre de modes est donc lié à la finesse du maillage. Si ce maillage comporte N_0^S noeuds sur la surface $\partial\Omega$ et N_0^D noeuds interne dans le domaine Ω seuls sont accessibles N_0^S modes de Steklov et N_0^D modes de Dirichlet respectivement.

Une première idée pour réduire la base initiale consiste à sélectionner les modes les plus lents, c'est à dire ceux ayant les valeurs propres les plus faibles. Compte tenu du classement des modes selon leur constante de temps obtenus lors de la résolution du problème aux valeurs propres par la méthode de Lanczos, cette technique a l'avantage d'être immédiate. Mais il est alors difficile de choisir la valeur de la constante de temps qui va servir de "coupure" dans le nombre de modes retenu. De plus, comment faire pour tronquer la base de Steklov, puisque leurs valeurs propres ne sont pas des constantes de temps ? Une deuxième idée consiste ainsi à faire une troncature énergétique en retenant les modes les plus dominants. C'est sur ce principe que repose la méthode utilisée : la réduction par amalgame modal.

2. Elle peut être montrée rigoureusement dans ce cas simple d'un pavé droit homogène où il existe une solution analytique [94].

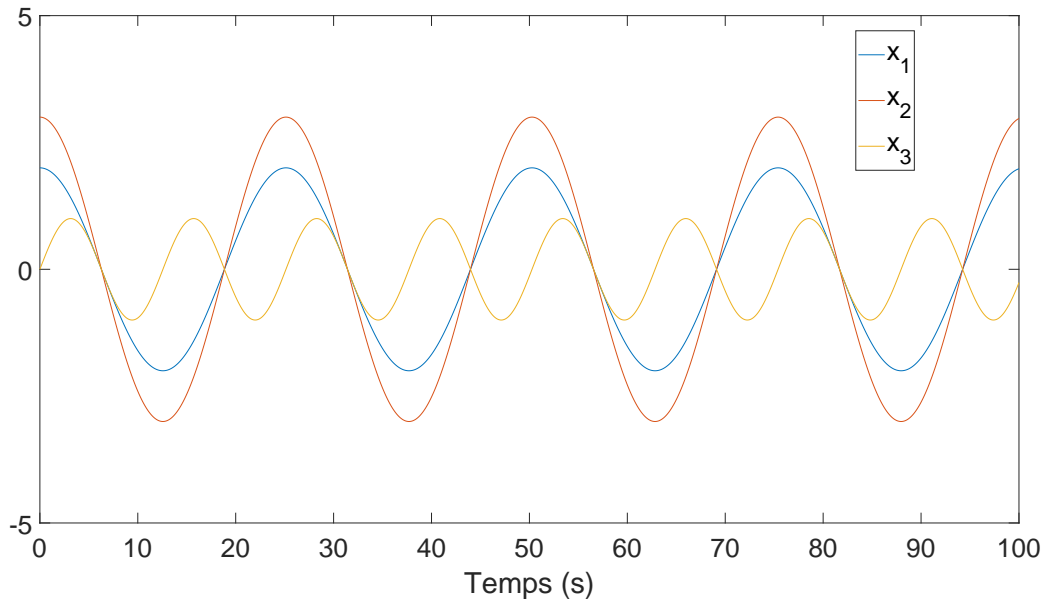


FIGURE 2.5.8 – Exemple d'états d'excitation d'une base constituée de trois modes.

5.4.1 Principe de la méthode de réduction par amalgame modal

La méthode d'amalgame modal a été initialement développée par Oulefki et Neveu [95] pour les bases de Fourier, puis adaptée par la suite aux modes de Branches [89, 96]. On va alors directement l'appliquer aux modes de Dirichlet-Steklov. Comme son nom l'indique, cette méthode consiste à amalgamer à des modes principaux, ou dominants, ou encore maîtres, un certain nombre de modes dits mineurs ou esclaves, de façon à conserver l'information spatiale de chacun des modes.

Pour l'illustrer, on va prendre un exemple simple, où on a une base initiale de trois modes quelconques, que l'on souhaite réduire en gardant seulement deux modes. On dispose pour cela d'états associés à ces modes, que l'on appelle états de référence, et qui correspondent à une réalisation $T(M, t) = \sum_{i=1}^3 x_i(t)V_i(M)$. Prenons par exemple les états x_i représentés sur la figure 2.5.8. On voit alors qu'on peut écrire

$$\begin{aligned} T &= x_1V_1 + x_2V_2 + x_3V_3 \\ &= x_1(V_1 + \alpha V_2) + x_3V_3 \\ &= x_1\tilde{V}_1 + x_3V_3. \end{aligned}$$

On appelle alors $\tilde{V}_1 = V_1 + \alpha V_2$ un *mode amalgamé*, V_1 étant le mode dominant et V_2 le mode esclave.

C'est sur ce principe que l'on va réduire une base initiale $\{V_i^{\mathcal{X}}\}_{\mathcal{X} \in \{D, S\}}$ de $N^D + N^S$ modes de Dirichlet-Steklov en une base réduite $\{\tilde{V}_i^{\mathcal{X}}\}_{\mathcal{X} \in \{D, S\}}$ de $\tilde{N} = \tilde{N}^D + \tilde{N}^S$ modes amalgamés.

5.4.2 Écart de réduction et dominance

Le cas pris en exemple est simpliste, et on n'a eu besoin que d'amalgamer V_1 et V_2 pour réduire la base. En pratique, un mode amalgamé peut-être constitué de plusieurs dizaines de modes. Chaque mode amalgamé s'écrit :

$$\forall \mathcal{X} \in \{D, S\}, 1 \leq i \leq \tilde{N}^{\mathcal{X}}, \quad \tilde{V}_i^{\mathcal{X}} = \sum_{p=1}^{N_i^{\mathcal{X}}} \alpha_{i,p}^{\mathcal{X}} V_{i,p}^{\mathcal{X}}. \quad (2.5.26)$$

Le champ de température sera alors approché par la décomposition réduite

$$\tilde{T} = \sum_{\mathcal{X} \in \{D, S\}} \sum_{I=1}^{\tilde{N}^{\mathcal{X}}} \tilde{x}_I^{\mathcal{X}} \tilde{V}_I^{\mathcal{X}}. \quad (2.5.27)$$

Afin de conserver les propriétés d'orthogonalité de la base, tous les modes de la base initiale ne sont utilisés qu'une seule et unique fois, soit en tant que mode dominant, soit en tant que mode esclave.

On détermine la répartition entre modes dominants et esclaves, ainsi que les différents coefficients de pondération $\alpha_{i,p}^{\mathcal{X}}$, de telle sorte que l'écart de réduction de la réponse en température $\Delta T = |T - \tilde{T}|$ entre le modèle d'origine et le modèle réduit soit le plus faible possible. Pour cela, on minimise l'écart quadratique de réduction de la réponse en température entre les champs T et \tilde{T} dans l'espace d'énergie.

Définition 5.4.1 *L'écart quadratique de réduction de la réponse en température entre les champs T et \tilde{T} dans l'espace d'énergie est défini par*

$$\|\Delta T\|_{L^2([0, \Delta\tau]; H^1(\Omega))}^2 = \int_0^{\Delta\tau} \|\Delta T\|_{H(\Omega)}^2. \quad (2.5.28)$$

On parle en fait d'espace d'énergie par abus de langage, puisque cette quantité n'a pas la dimension physique d'une quantité de chaleur ou d'un travail, mais est mesurée en "joule kelvin" (à ne pas confondre avec l'entropie du système, exprimée en "joule par kelvin").

Le champ T est le champ de référence à partir duquel on va réduire la base. Celui-ci nous permet d'obtenir des *états de référence* en projetant les modes de la base initiale sur T selon le produit scalaire $(\cdot, \cdot)_{H(\Omega)}$. On a alors la décomposition

$$T = \sum_{\mathcal{X} \in \{D, S\}} \sum_{i=1}^{N^{\mathcal{X}}} x_i^{\mathcal{X}} V_i^{\mathcal{X}}. \quad (2.5.29)$$

Les modes dominants sont ensuite déterminés en calculant les *dominances modales* de chaque mode de la base.

Définition 5.4.2 *La dominance modale d'un mode de Dirichlet-Steklov est définie par*

$$\mathcal{X} \in \{D, S\}, \quad \mathcal{D}_i^{\mathcal{X}} = \int_0^{\Delta t} (x_i^{\mathcal{X}}(s))^2. \quad (2.5.30)$$

5.4.3 Algorithme d'amalgame

Dans les travaux précédents de l'équipe ThE du LMEE, le nombre de modes réduits obtenus par amalgame est arbitraire : on fixe un nombre souhaité de modes réduits et on laisse l'algorithme d'amalgame construire les modes amalgamés jusqu'à ce qu'il atteigne le nombre requis.

5.4.3.1 Initialisation

Le mode dont la dominance est la plus élevée, *i.e.* telle que $\mathcal{D}_i^{\mathcal{X}} = \max$ est choisi comme mode dominant. Puisqu'on réduit une base de Dirichlet-Steklov, on amalgame les modes de Dirichlet qu'avec des modes de Dirichlet, idem pour les modes de Steklov. On choisit donc un mode de Dirichlet dominant et un mode de Steklov dominant. Le premier mode de Steklov (le "mode plat") est quant à lui écarté de la procédure, et est gardé tel quel.

On va ensuite amalgamer les autres modes, de fait esclaves, aux modes dominants. On forme ainsi des sous-espaces d'amalgames, constitués chacun d'un mode maître et de modes esclaves associés.

Définition 5.4.3 *La dominance modale de l'écart, notée $\Delta \mathcal{D}_{I,q}^{\mathcal{X}}$, pour le $q^{\text{ème}}$ mode du $I^{\text{ème}}$ sous espace d'amalgame est définie par*

$$\Delta \mathcal{D}_{I,q}^{\mathcal{X}} = \int_0^{\Delta t} (x_{I,q}^{\mathcal{X}} - \tilde{x}_I^{\mathcal{X}} \alpha_{I,q}^{\mathcal{X}})^2, \quad (2.5.31)$$

$V_{I,1}^{\mathcal{X}}$ étant le mode maître de ce sous espace.

Les états réduits $\tilde{x}_I^{\mathcal{X}}$ étant inaccessibles, on impose $\tilde{x}_I^{\mathcal{X}} = x_{I,1}^{\mathcal{X}}$, puisque c'est le mode maître qui impose sa dynamique au mode amalgamé.

On peut alors déterminer les facteurs optimaux d'amalgame $\alpha_{I,q}^{\mathcal{X}}$, autrement dit ceux qui minimisent l'écart entre modèle complet et réduit.

Proposition 5.4.1 *Pour une partition donnée de la base de Dirichlet-Steklov le facteur d'amalgame $\alpha_{I,q}^{opt \mathcal{X}}$ qui minimise l'énergie de l'écart de la réponse en température entre le modèle complet et le modèle réduit est donné par la formule :*

$$\alpha_{I,q}^{opt \mathcal{X}} = \frac{\int_0^{\Delta t} x_{I,q}^{\mathcal{X}}(t) x_{I,1}^{\mathcal{X}}(t)}{\int_0^{\Delta t} (x_{I,1}^{\mathcal{X}}(t))^2}. \quad (2.5.32)$$

La dominance modale de l'écart est alors minimale et s'écrit :

$$\Delta \mathcal{D}_{I,q}^{\min \mathcal{X}} = \mathcal{D}_{I,q}^{\mathcal{X}} - \left(\alpha_{I,q}^{opt \mathcal{X}} \right)^2 \mathcal{D}_{I,1}^{\mathcal{X}}, \quad (2.5.33)$$

avec $0 \leq \Delta \mathcal{D}_{I,q}^{\min \mathcal{X}} \leq \mathcal{D}_{I,q}^{\mathcal{X}}$.

À la fin de cette initialisation, on a donc deux modes amalgamés contenant tous les modes de chaque famille, combinés avec les coefficients α calculés grâce à (2.5.32).

5.4.3.2 Étape n

On choisit, toutes familles confondues, comme n -ème nouveau mode maître le mode esclave apportant la plus grande contribution à l'énergie de l'écart de réduction, *i.e.* tel que

$$\Delta \mathcal{D}_{K,m}^{\min \mathcal{Y}} = \max_{\substack{\mathcal{X} \in \{D,S\} \\ I \in \{1, \dots, \tilde{n}^{\mathcal{X}}\} \\ k \in \{1, \dots, N_I^{\mathcal{X}}\}}} \left(\Delta \mathcal{D}_{I,k}^{\min \mathcal{X}} \right),$$

où $\tilde{n}^{\mathcal{X}}$ est le nombre de modes amalgamés de la famille \mathcal{X} à cette étape.

On donne à ce nouveau mode l'ordre $J = \tilde{n}^{\mathcal{Y}}$ dans la base amalgamée. Pour chaque mode esclave de la famille \mathcal{Y} , on calcule son facteur d'amalgame optimal $\alpha_{J,q}^{opt \mathcal{Y}}$ avec le nouveau mode maître et on détermine le minimum de la dominance modale de l'écart $\Delta \mathcal{D}_{J,q}^{\min \mathcal{Y}}$.

Alors, si $\Delta \mathcal{D}_{J,q}^{\min \mathcal{Y}} < \Delta \mathcal{D}_{I,q}^{\min \mathcal{Y}}$, ou en d'autres termes, si l'énergie de l'écart de réduction de la réponse en température $\Delta \mathcal{D}_{J,q}^{\min \mathcal{Y}}$ est plus faible lorsque le mode esclave considéré est amalgamé au nouveau mode maître J qu'avec son mode maître actuel d'ordre $I < J$, alors il change de maître et on réorganise les partitions : le sous-espace d'amalgame d'ordre J contient un mode en plus ($N_J^{\mathcal{Y}} = N_I^{\mathcal{Y}} + 1$), et le sous-espace d'amalgame d'ordre I contient un mode en moins ($N_I^{\mathcal{Y}} = N_I^{\mathcal{Y}} - 1$). Sinon, le mode esclave considéré reste attaché à son mode maître actuel.

On répète ainsi cette opération pour chaque mode esclave de la famille \mathcal{Y} , puis, si on a atteint le nombre de modes réduits voulu \tilde{N} , on s'arrête là. Sinon, on passe à l'itération $n + 1$.

On finalise l'algorithme en normant les modes amalgamés par rapport à la norme $\|\cdot\|_{H(\Omega)}$:

$$\mathcal{X} \in \{D, S\}, 1 \leq i \leq \tilde{N}^{\mathcal{X}}, \quad \tilde{V}_i^{\mathcal{X}} = \frac{\tilde{V}_i^{\mathcal{X}}}{\sqrt{\sum_{p=1}^{\tilde{N}_i^{\mathcal{X}}} (\alpha_{i,p}^{\mathcal{X}})^2}}. \quad (2.5.34)$$

Pour avoir une estimation de l'écart de réduction après simulation avec les modes réduits, on calcule *l'écart de projection* ΔT^{proj} :

$$\Delta T^{proj} = |T - T^{proj}|, \quad (2.5.35)$$

avec

$$T^{proj} = \sum_{\mathcal{X} \in \{D, S\}} \sum_{I=1}^{\tilde{N}^{\mathcal{X}}} \tilde{x}_I^{proj, \mathcal{X}} \tilde{V}_I^{\mathcal{X}}, \quad (2.5.36)$$

le *champ projeté*. Les *états projetés* $\tilde{x}_I^{proj, \mathcal{X}}$ sont obtenus en "projetant" T sur la base réduite selon $(\cdot | \cdot)_{H(\Omega)}$: $\tilde{x}_I^{proj, \mathcal{X}} = (T | \tilde{V}_I^{\mathcal{X}})_{H(\Omega)}$.

Remarque 5.4.1 *Pour construire la base réduite amalgamée, on a eu besoin d'états de référence, issus d'un champ de référence T . Notre base est donc conçue pour réduire l'écart de réduction avec ce champ particulier, ce qui peut sembler être en contradiction avec l'objectif de créer des bibliothèques de composants réduits réutilisables. Le choix de ce champ sera donc crucial pour que la base amalgamée puisse être utilisée pour reconstruire plusieurs conditions aux limites et plusieurs configurations, mais on verra dans la partie dédiée aux applications à quel point on peut "détourner" une base réduite et obtenir des bons résultats avec des simulations très différentes du champ de référence.*

5.4.4 Vers un nouveau critère d'arrêt

La valeur exacte de l'erreur de réduction, correspondant à l'évaluation de $\Delta T = |T - \tilde{T}|$ s'obtient donc tardivement, puisqu'on obtient les états réduits de l'expression \tilde{T} après résolution du problème réduit utilisant la base amalgamée (décrit dans la section ci-après), et si on veut la réduire, il faut refaire un amalgame en indiquant un nombre plus élevé de modes réduits.

Ce procédé peut être fastidieux si on cherche à obtenir une erreur de réduction inférieure à un seuil spécifique, et reste très empirique. C'est pourquoi, au cours de cette thèse, la

faisabilité d'un amalgame dont le critère d'arrêt serait une estimation *a priori* de l'écart en température entre simulation réduite et simulation complète a été étudiée.

Dans le processus d'amalgame, l'écart est minimisé dans l'espace d'énergie lié à la norme $\|\cdot\|_{H(\Omega)}$. Pour estimer l'erreur de réduction en terme d'écart quadratique sur les températures, on doit relier $\|\cdot\|_{H(\Omega)}$ à la norme $\|\cdot\|_{L^2(\Omega)}$ usuelle. C'est le sens de la proposition suivante qui montre que la convergence dans $H^1(\Omega)$ entraîne la convergence dans $L^2(\Omega)$ (mais aussi dans $L^1(\Omega)$).

Proposition 5.4.2 *Soit la fonction $u \in H^1(\Omega)$ où Ω est un ouvert régulier de \mathbb{R}^d . Alors :*

$$\forall u \in H^1(\Omega), \exists C_* > 0 \text{ telle que } \|u\|_{L^2(\Omega)} \leq C_* \|u\|_{H(\Omega)}. \quad (2.5.37)$$

À partir de cette minimisation, on peut ainsi se fixer une valeur maximale de l'écart souhaité en terme d'écart de température et se ramener à un écart dans l'espace d'énergie pour obtenir le nombre suffisant de modes amalgamés.

La difficulté réside dans l'estimation de C_* , ou plutôt dans une majoration suffisamment fine de C_* . En effet, si on donne une valeur trop grande pour C_* , on peut être amené à construire des centaines de modes amalgamés pour reconstruire le champ de température alors que quelques dizaines suffisent, amoindrissant ainsi le facteur de réduction. La démonstration de cette proposition et de valeurs de majorants de C_* pour des géométries particulières (boule, pavé droit de \mathbb{R}^d, \dots) seront donnés en annexe. Pour des géométries complexes telles que la carte modélisée dans la partie suivante, la recherche d'une valeur assez fine d'un majorant de C_* n'a malheureusement pas encore abouti.

5.5 Écriture modale du problème sous-structuré

La base de Dirichlet-Steklov apparaît comme un candidat idéal pour décomposer la température, puisque elle sépare les contributions purement volumiques et les contributions surfaciques, ce qui est très avantageux pour le couplage et l'obtention d'une formulation modale du problème thermique sous-structuré défini en 4.2.1.

Pour chaque $k \in \{0, \dots, N_\Omega\}$, et d'après la proposition 5.2.6, on décompose les champs de température $T^{(k)}$ sur une base de Dirichlet-Steklov dans l'espace $H_0^1(\Omega^{(k)}) \times H^{1/2}(\partial\Omega^{(k)})$. On prolonge ces bases par zéro dans le reste du domaine $\Omega \setminus \Omega^{(k)}$. Chaque $T^{(k)}$ s'exprime alors sous la forme :

$$T^{(k)} = \sum_{\mathcal{X} \in \{D, S\}} \sum_{j=1}^{N_{\mathcal{X}}^{(k)}} x_j^{\mathcal{X}, (k)} V_j^{\mathcal{X}, (k)}. \quad (2.5.38)$$

On note $N_D^{(k)}$ et $N_S^{(k)}$ le nombre de modes de Dirichlet et de Steklov respectivement. Ces bases peuvent être réduites ou non, selon qu'on utilise une base réduite via l'amalgame modal présenté en section précédente ou qu'on utilise une base complète³, mais cela n'influence pas l'écriture modale du problème.

Comme en section 4.2.3 page 59, sur chaque frontière non-mortier $\Gamma_p^{\{k,j\}}$, on décompose les densités de flux de $\partial\Omega^{(k)}$ sur la base $\{\phi_{p_m}^{\{k,j\}}\}_{1 \leq m \leq N_\phi^{\Gamma_p^{\{k,j\}}}}$ de fonctions de l'espace mortier discrétisé associé :

$$\varphi_p^{k,j} = \sum_{m=1}^{N_\phi^{\Gamma_p^{\{k,j\}}}} \eta_{p_m}^{\{k,j\}} \phi_{p_m}^{\{k,j\}}. \quad (2.5.39)$$

On rappelle que, par définition, $V_i^{D,(k)}|_{\partial\Omega^{(k)}} = 0$ et $V_i^{S,(k)}|_{\partial\Omega^{(k)}} = S_i^{(k)}$.

5.5.1 Projection des équations sur la base modale.

En remplaçant la température et le flux par leur décomposition sur les modes propres de Dirichlet-Steklov et sur les fonctions tests ϕ_m dans les équations (2.4.19a) et (2.4.19b), et en utilisant comme fonctions test ces mêmes modes propres de Dirichlet-Steklov ainsi que les fonctions ϕ_m , on obtient le jeu d'équations couplées suivant :

$$\begin{aligned} 0 \leq k \leq N_\Omega, \quad 1 \leq i \leq N_D^{(k)}, \quad & \sum_{\mathcal{X} \in \{D,S\}} \sum_{j=1}^{N_{\mathcal{X}}^{(k)}} x_j^{\mathcal{X},(k)} \int_{\Omega^{(k)}} \nabla V_i^{D,(k)} \cdot K^{(k)} \cdot \nabla V_j^{\mathcal{X},(k)} + \\ & \sum_{\mathcal{X} \in \{D,S\}} \sum_{j=1}^{N_{\mathcal{X}}^{(k)}} \frac{\partial}{\partial t} x_j^{\mathcal{X},(k)} \int_{\Omega^{(k)}} C^{(k)} V_i^{D,(k)} V_j^{\mathcal{X},(k)} = \int_{\Omega^{(k)}} \varpi^{(k)} V_i^{D,(k)} \end{aligned} \quad (2.5.40a)$$

$$\begin{aligned} 0 \leq k \leq N_\Omega, \quad 1 \leq i \leq N_S^{(k)}, \quad & \sum_{\mathcal{X} \in \{D,S\}} \sum_{j=1}^{N_{\mathcal{X}}^{(k)}} x_j^{\mathcal{X},(k)} \int_{\Omega^{(k)}} \nabla V_i^{S,(k)} \cdot K^{(k)} \cdot \nabla V_j^{\mathcal{X},(k)} + \\ & \sum_{\mathcal{X} \in \{D,S\}} \sum_{j=1}^{N_{\mathcal{X}}^{(k)}} \frac{\partial}{\partial t} x_j^{\mathcal{X},(k)} \int_{\Omega^{(k)}} C^{(k)} V_i^{S,(k)} V_j^{\mathcal{X},(k)} + \sum_{j=1}^{N_{\mathcal{X}}^{(k)}} x_j^{\mathcal{X},(k)} \int_{\Sigma^{(k)}} h^{(k)} S_i^{(k)} S_j^{(k)} + \\ & \sum_{l=0}^{N_\Omega} \sum_{p=1}^{N_\Gamma^{\{k,l\}}} S_g^{k,l} \sum_{m=1}^{N_\phi^{\Gamma_p^{\{k,l\}}}} \eta_{p_m}^{\{k,l\}} \int_{\Gamma_p^{\{k,l\}}} S_i^{(k)}|_{\Gamma_p^{\{k,l\}}} \phi_{p_m}^{\{k,l\}} = \int_{\Omega^{(k)}} \varpi^{(k)} V_i^{S,(k)} + \int_{\Sigma^{(k)}} h^{(k)} T_{ext} S_i^{(k)} \end{aligned} \quad (2.5.40b)$$

3. *i.e.* dont le nombre de modes est égal au nombre de points de la discrétisation de $\Omega^{(k)}$.

$$\begin{aligned}
0 \leq k, l \leq N_\Omega, k \neq l, 1 \leq p \leq N_\Gamma^{\{k,l\}} \mid \Gamma_p^{k,l} \subset \Gamma_{\square}^{(k)}, 1 \leq n \leq N_\phi^{\Gamma_p^{\{k,l\}}} \\
s_{g_p}^{k,l} \sum_{j=1}^{N_S^{(k)}} x_j^{S,(k)} \int_{\Gamma_p^{\{k,l\}}} \phi_{p_n}^{\{k,l\}} S_j^{(k)} \Big|_{\Gamma_p^{\{k,l\}}} + s_{g_p}^{l,k} \sum_{j=1}^{N_S^{(l)}} x_j^{S,(l)} \int_{\Gamma_p^{\{k,l\}}} \phi_{p_n}^{\{k,l\}} S_j^{(l)} \Big|_{\Gamma_p^{\{k,l\}}} \\
- \sum_{m=1}^{N_\phi^{\Gamma_p^{\{k,l\}}}} \eta_{p_m}^{\{k,l\}} \int_{\Gamma_p^{\{k,l\}}} R_{c_p}^{\{k,l\}} \phi_{p_n}^{\{k,l\}} \phi_{p_m}^{\{k,l\}} = 0. \quad (2.5.40c)
\end{aligned}$$

5.5.2 Forme matricielle

On va donner une formulation matricielle des équations d'état (2.5.40). Pour cela, on introduit des nouvelles matrices de conductivité, de capacité, de joints, . . . , mais modales :

$$0 \leq k \leq N_\Omega,$$

$$\begin{aligned}
\tilde{\mathbf{C}}_{\mathcal{X}\mathcal{Y}}^{(k)} &= \left(\int_{\Omega^{(k)}} C^{(k)} V_i^{\mathcal{X},(k)} V_j^{\mathcal{Y},(k)} \right)_{\substack{1 \leq i \leq N_{\mathcal{X}}^{(k)} \\ 1 \leq j \leq N_{\mathcal{Y}}^{(k)}}}, \quad \tilde{\mathbf{K}}_{\mathcal{X}\mathcal{Y}}^{(k)} = \left(\int_{\Omega^{(k)}} \nabla V_i^{\mathcal{X},(k)} \cdot K^{(k)} \cdot \nabla V_j^{\mathcal{Y},(k)} \right)_{\substack{1 \leq i \leq N_{\mathcal{X}}^{(k)} \\ 1 \leq j \leq N_{\mathcal{Y}}^{(k)}}} \\
\tilde{\mathbf{H}}^{(k)} &= \left(\int_{\Sigma^{(k)}} h^{(k)} S_i^{(k)} \Big|_{\Sigma^{(k)}} S_j^{(k)} \Big|_{\Sigma^{(k)}} \right)_{\substack{1 \leq i \leq N_S^{(k)} \\ 1 \leq j \leq N_S^{(k)}}}.
\end{aligned}$$

Dans le cas général, les matrices $\mathbf{K}_{\mathcal{X}\mathcal{X}}^{(k)}$ sont symétriques mais pleines. Si le tenseur de conductivité $K^{(k)}$ est identique à la fonction utilisée pour déterminer les modes propres alors ces matrices sont diagonales du fait des relations d'orthogonalité dans $H^1(\Omega^{(k)})$. En particulier, la matrice $\mathbf{K}_{DS}^{(k)}$ qui couple les modes de Dirichlet et de Steklov d'un domaine est rectangulaire. Si le tenseur de conductivité $K^{(k)}$ est identique à la fonction utilisée pour déterminer les modes propres alors la matrice est nulle du fait de ces mêmes relations d'orthogonalité.

On pose, pour $0 \leq k, l \leq N_\Omega$, $1 \leq p \leq N_\Gamma^{\{k,l\}}$ tels que $\Gamma_p^{\{k,l\}} \in \Gamma_{\square}$,

$$\tilde{\mathbf{J}}_p^{(k,l)} = s_{g_p}^{k,l} \left(\int_{\Gamma_p^{\{k,l\}}} S_i^{(k)} \Big|_{\Gamma_p^{\{k,l\}}} \phi_{p_m}^{\{k,l\}} \right)_{\substack{1 \leq i \leq N_S^{(k)} \\ 1 \leq m \leq N_\phi^{\Gamma_p^{\{k,l\}}}},$$

on peut ainsi former la matrice rectangulaire

$$\tilde{\mathbf{J}}^{(k,l)} = \left[\tilde{\mathbf{J}}_1^{(k,l)}, \dots, \tilde{\mathbf{J}}_{N_\Gamma^{\{k,l\}}}^{(k,l)} \right],$$

et la matrice

$$\tilde{\mathbf{J}} = \begin{bmatrix} \tilde{\mathbf{J}}^{(0,0)} & \tilde{\mathbf{J}}^{(0,1)} & \dots & \tilde{\mathbf{J}}^{(0,N_{\sqcup})} \\ \tilde{\mathbf{J}}^{(1,0)} & \tilde{\mathbf{J}}^{(1,1)} & \dots & \tilde{\mathbf{J}}^{(1,N_{\sqcup})} \\ \vdots & & \ddots & \\ \tilde{\mathbf{J}}^{(N_{\Omega},0)} & \tilde{\mathbf{J}}^{(N_{\Omega},1)} & \dots & \tilde{\mathbf{J}}^{(N_{\Omega},N_{\sqcup})} \end{bmatrix}.$$

On garde les mêmes définitions qu'en section 4.2.3 page 59 pour les matrices $\mathbf{M}_{c_p}^{(k,l)}$ et \mathbf{M}_c et le vecteur Φ , puisqu'ils dépendent uniquement des fonctions mortiers.

On pose enfin les vecteurs :

$$0 \leq k \leq N_{\Omega},$$

$$\tilde{\Pi}_{\mathcal{X}}^{(k)} = \left(\int_{\Omega^{(k)}} \varpi^{(k)} V_i^{\mathcal{X},(k)} \right)_{1 \leq i \leq N_{\mathcal{X}}^{(k)}}, \quad \tilde{\Theta}_{ext}^{(k)} = \left(\int_{\Sigma^{(k)}} h^{(k)} T_{ext} S_i^{(k)} |_{\Sigma^{(k)}} \right)_{1 \leq i \leq N_{\mathcal{X}}^{(k)}},$$

$$\mathbf{X}_{\mathcal{X}}^{(k)} = \left(x_j^{\mathcal{X},(k)} \right)_{1 \leq j \leq N_{\mathcal{X}}^{(k)}}.$$

En assemblant toutes les matrices de chaque sous-domaine $\Omega^{(k)}$ et de chaque interface mortier $\Gamma_p^{(k,l)}$, on peut définir les “super-matrices” :

$$\tilde{\mathbf{K}}_{\mathcal{X}\mathcal{X}} = \begin{bmatrix} \tilde{\mathbf{K}}_{\mathcal{X}\mathcal{X}}^{(0)} & 0 & 0 & 0 \\ 0 & \tilde{\mathbf{K}}_{\mathcal{X}\mathcal{X}}^{(1)} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \tilde{\mathbf{K}}_{\mathcal{X}\mathcal{X}}^{(N_{\Omega})} \end{bmatrix}, \quad \tilde{\mathbf{K}}_{DS} = \begin{bmatrix} \tilde{\mathbf{K}}_{DS}^{(0)} & 0 & 0 & 0 \\ 0 & \tilde{\mathbf{K}}_{DS}^{(1)} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \tilde{\mathbf{K}}_{DS}^{(N_{\Omega})} \end{bmatrix},$$

$$\tilde{\mathbf{C}}_{\mathcal{X}\mathcal{Y}} = \begin{bmatrix} \tilde{\mathbf{C}}_{\mathcal{X}\mathcal{Y}}^{(0)} & 0 & 0 & 0 \\ 0 & \tilde{\mathbf{C}}_{\mathcal{X}\mathcal{Y}}^{(1)} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \tilde{\mathbf{C}}_{\mathcal{X}\mathcal{Y}}^{(N_{\Omega})} \end{bmatrix}, \quad \tilde{\mathbf{H}} = \begin{bmatrix} \tilde{\mathbf{H}}^{(0)} & 0 & 0 & 0 \\ 0 & \tilde{\mathbf{H}}^{(1)} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \tilde{\mathbf{H}}^{(N_{\Omega})} \end{bmatrix},$$

$$\mathbf{X}_{\mathcal{X}} = \begin{pmatrix} \mathbf{X}_{\mathcal{X}}^{(0)} \\ \mathbf{X}_{\mathcal{X}}^{(1)} \\ \vdots \\ \mathbf{X}_{\mathcal{X}}^{(N_{\Omega})} \end{pmatrix}, \quad \tilde{\Pi}_{\mathcal{X}} = \begin{pmatrix} \tilde{\Pi}_{\mathcal{X}}^{(0)} \\ \tilde{\Pi}_{\mathcal{X}}^{(1)} \\ \vdots \\ \tilde{\Pi}_{\mathcal{X}}^{(N_{\Omega})} \end{pmatrix}, \quad \tilde{\Theta}_{ext} = \begin{pmatrix} \tilde{\Theta}_{ext}^{(0)} \\ \tilde{\Theta}_{ext}^{(1)} \\ \vdots \\ \tilde{\Theta}_{ext}^{(N_{\Omega})} \end{pmatrix}.$$

Finalement, le système d'équations modales (2.5.40) s'écrit sous forme matricielle condensée :

$$\begin{bmatrix} \tilde{\mathbf{K}}_{DD} & \tilde{\mathbf{K}}_{DS} & 0 \\ {}^t\tilde{\mathbf{K}}_{DS} & \tilde{\mathbf{K}}_{SS} + \tilde{\mathbf{H}} & \tilde{\mathbf{J}} \\ 0 & {}^t\tilde{\mathbf{J}} & \mathbf{M}_c \end{bmatrix} \begin{pmatrix} \mathbf{X}_D \\ \mathbf{X}_S \\ \Phi \end{pmatrix} + \begin{bmatrix} \tilde{\mathbf{C}}_{DD} & \tilde{\mathbf{C}}_{DS} & 0 \\ {}^t\tilde{\mathbf{C}}_{DS} & \tilde{\mathbf{C}}_{SS} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} \dot{\mathbf{X}}_D \\ \dot{\mathbf{X}}_S \\ \dot{\Phi} \end{pmatrix} = \begin{pmatrix} \tilde{\Pi}_D \\ \tilde{\Pi}_S + \tilde{\Theta}_{ext} \\ 0 \end{pmatrix}. \quad (2.5.41)$$

Une fois ce système matriciel résolu, on peut reconstituer tout simplement le champ de température par la formule

$$\mathbf{T}^r = \mathbf{V}^D \mathbf{X}_D + \mathbf{V}^S \mathbf{X}_S, \quad (2.5.42)$$

où \mathbf{V}^D et \mathbf{V}^S sont les matrices où sont stockés les modes de Dirichlet et de Steklov (*i.e.* les matrices de passage de l'espace modal à l'espace des températures). Le champ obtenu ainsi est le *champ réduit*.

5.5.2.1 Forme matricielle pour une carte électronique.

Dans le cas particulier où on considère une carte électronique, on obtient les équations modales suivantes, où on distingue les modes de Steklov du PCB ($\Omega^{(0)}$) de ceux des composants ($\Omega^{(k)}$, $1 \leq k \leq N_\Omega$), soit :

- Pour les modes de Dirichlet, $0 \leq k \leq N_\Omega$, $1 \leq i \leq N_D^{(k)}$:

$$\begin{aligned} \sum_{\mathcal{X} \in \{D, S\}} \sum_{j=1}^{N_{\mathcal{X}}^{(k)}} x_j^{\mathcal{X}, (k)} \int_{\Omega^{(k)}} \nabla V_i^{D, (k)} \cdot K^{(k)} \cdot \nabla V_j^{\mathcal{X}, (k)} \\ + \sum_{\mathcal{X} \in \{D, S\}} \sum_{j=1}^{N_{\mathcal{X}}^{(k)}} \frac{\partial}{\partial t} x_j^{\mathcal{X}, (k)} \int_{\Omega^{(k)}} C^{(k)} V_i^{D, (k)} V_j^{\mathcal{X}, (k)} = \int_{\Omega^{(k)}} \varpi^{(k)} V_i^{D, (k)}. \end{aligned} \quad (2.5.43)$$

- Pour les modes de Steklov du PCB, $k = 0$, $1 \leq i \leq N_S^{(0)}$:

$$\begin{aligned} \sum_{\mathcal{X} \in \{D, S\}} \sum_{j=1}^{N_{\mathcal{X}}^{(0)}} x_j^{\mathcal{X}, (0)} \int_{\Omega^{(0)}} \nabla V_i^{S, (0)} \cdot K^{(0)} \cdot \nabla V_j^{\mathcal{X}, (0)} \\ + \sum_{\mathcal{X} \in \{D, S\}} \sum_{j=1}^{N_{\mathcal{X}}^{(0)}} \frac{\partial}{\partial t} x_j^{\mathcal{X}, (0)} \int_{\Omega^{(0)}} C^{(0)} V_i^{S, (0)} V_j^{\mathcal{X}, (0)} + \sum_{j=1}^{N_{\mathcal{X}}^{(0)}} x_j^{\mathcal{X}, (0)} \int_{\Sigma^{(0)}} h^{(0)} S_i^{(0)} S_j^{(0)} \\ + \sum_{l=0}^{N_\Omega} \sum_{p=1}^{N_\Gamma^{\{l, 0\}}} s_g^{0, l} \sum_{m=1}^{N_\phi^{\Gamma_p^{\{l, 0\}}}} \eta_{p_m}^{\{l, 0\}} \int_{\Gamma_p^{\{l, 0\}}} S_i^{(0)}|_{\Gamma_p^{\{l, 0\}}} \phi_{p_m}^{\{l, 0\}} = \int_{\Omega^{(0)}} \varpi^{(0)} V_i^{S, (0)} + \int_{\Sigma^{(0)}} h^{(0)} T_{ext} S_i^{(0)}. \end{aligned} \quad (2.5.44)$$

- Pour les modes de Steklov des composants, $1 \leq k \leq N_\Omega$, $1 \leq i \leq N_S^{(k)}$:

$$\begin{aligned}
& \sum_{\mathcal{X} \in \{D, S\}} \sum_{j=1}^{N_{\mathcal{X}}^{(0)}} x_j^{\mathcal{X},(k)} \int_{\Omega^{(k)}} \nabla V_i^{S,(k)} \cdot K^{(k)} \cdot \nabla V_j^{\mathcal{X},(k)} \\
& + \sum_{\mathcal{X} \in \{D, S\}} \sum_{j=1}^{N_{\mathcal{X}}^{(k)}} \frac{\partial}{\partial t} x_j^{\mathcal{X},(k)} \int_{\Omega^{(k)}} C^{(k)} V_i^{S,(k)} V_j^{\mathcal{X},(k)} + \sum_{j=1}^{N_{\mathcal{X}}^{(k)}} x_j^{\mathcal{X},(k)} \int_{\Sigma^{(k)}} h^{(k)} S_i^{(k)} S_j^{(k)} \\
& + \sum_{p=1}^{N_{\Gamma}^{\{k,0\}}} s_{g_p}^{k,0} \sum_{m=1}^{N_{\phi}^{\{k,0\}}} \eta_{p_m}^{\{k,0\}} \int_{\Gamma_p^{\{k,0\}}} S_i^{(k)}|_{\Gamma_p^{\{k,0\}}} \phi_{p_m}^{\{k,0\}} = \int_{\Omega^{(k)}} \varpi^{(k)} V_i^{S,(k)} + \int_{\Sigma^{(k)}} h^{(k)} T_{ext} S_i^{(k)}. \quad (2.5.45)
\end{aligned}$$

• Pour les fonctions mortiers, $1 \leq l \leq N_{\Omega}$, $1 \leq p \leq N_{\Gamma}^{\{l,0\}}$ tels que $\Gamma_p^{0,l} \subset \Gamma_{\cup}$, $1 \leq n \leq N_{\phi}^{\{l,0\}}$:

$$\begin{aligned}
& s_{g_p}^{0,l} \sum_{j=1}^{N_S^{(0)}} x_j^{S,(0)} \int_{\Gamma_p^{\{l,0\}}} \phi_{p_n}^{\{l,0\}} S_j^{(0)}|_{\Gamma_p^{\{l,0\}}} + s_{g_p}^{l,0} \sum_{j=1}^{N_S^{(l)}} x_j^{S,(l)} \int_{\Gamma_p^{\{l,0\}}} \phi_{p_n}^{\{l,0\}} S_j^{(l)}|_{\Gamma_p^{\{l,0\}}} \\
& - \sum_{m=1}^{N_{\phi}^{\{l,0\}}} \eta_{p_m}^{\{l,0\}} \int_{\Gamma_p^{\{l,0\}}} r_{c_p}^{\{l,0\}} \phi_{p_n}^{\{l,0\}} \phi_{p_m}^{\{l,0\}} = 0.
\end{aligned}$$

Les matrices modales de conduction, de capacité, ..., sont les mêmes que dans le cas général, en revanche on a une simplification des matrices $\tilde{\mathbf{J}}$ et \mathbf{M}_c . On définit : $1 \leq k \leq N_{\Omega}$, $1 \leq p \leq N_{\Gamma}^{\{0,k\}}$

$$\begin{aligned}
\tilde{\mathbf{J}}_p^{(0,k)} &= s_{g_p}^{0,k} \left(\int_{\Gamma_p^{\{k,0\}}} S_i^{(0)}|_{\Gamma_p^{\{k,0\}}} \phi_{p_m}^{\{k,0\}} \right)_{\substack{1 \leq i \leq N_S^{(0)} \\ 1 \leq m \leq N_{\phi}^{\Gamma_p^{\{k,0\}}}}, \\
\tilde{\mathbf{J}}_p^{(k,k)} &= s_{g_p}^{k,0} \left(\int_{\Gamma_p^{\{k,0\}}} S_i^{(k)}|_{\Gamma_p^{\{k,0\}}} \phi_{p_m}^{\{k,0\}} \right)_{\substack{1 \leq i \leq N_S^{(k)} \\ 1 \leq m \leq N_{\phi}^{\Gamma_p^{\{k,0\}}}}.
\end{aligned}$$

$$\mathbf{M}_{c_p}^{(k)} = - \left(\int_{\Gamma_p^{\{k,0\}}} R_{c_p}^{\{k,0\}} \phi_{p_i}^{\{k,0\}} \phi_{p_m}^{\{k,0\}} \right)_{\substack{1 \leq i \leq N_{\phi}^{\Gamma_p^{\{k,0\}}} \\ 1 \leq m \leq N_{\phi}^{\Gamma_p^{\{k,0\}}}} .$$

On peut ainsi former les matrices bloc rectangulaires $\tilde{\mathbf{J}}^{(l,k)}$ et bloc diagonale $\mathbf{M}_c^{(k)}$ pour $0 \leq l \leq N_{\Omega}$ et $1 \leq k \leq N_{\Omega}$:

$$\begin{aligned} \tilde{\mathbf{J}}^{(0,k)} &= \begin{bmatrix} \tilde{\mathbf{J}}_1^{(0,k)} & \tilde{\mathbf{J}}_2^{(0,k)} & \dots & \tilde{\mathbf{J}}_{N_{\Gamma}^{\{0,k\}}}^{(0,k)} \end{bmatrix}, & \mathbf{M}_c^{(k)} &= \begin{bmatrix} \mathbf{M}_{c_1}^{(k)} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{M}_{c_{N_{\Gamma}^{\{0,k\}}}}^{(k)} \end{bmatrix}. \\ \tilde{\mathbf{J}}^{(k,k)} &= \begin{bmatrix} \tilde{\mathbf{J}}_1^{(k,k)} & \tilde{\mathbf{J}}_2^{(k,k)} & \dots & \tilde{\mathbf{J}}_{N_{\Gamma}^{\{0,k\}}}^{(k,k)} \end{bmatrix}, \end{aligned}$$

De même, les vecteurs des coefficients de décomposition des fonctions mortier s'écrivent :

$$1 \leq k \leq N_{\Omega}, \quad \Phi_p^{(k)} = \left(\eta_{p_m}^{\{k,0\}} \right)_{1 \leq m \leq N_{\phi}^{\Gamma_p^{\{k,0\}}}}, \quad \Phi^{(k)} = \begin{pmatrix} \Phi_1^{(k)} \\ \vdots \\ \Phi_{N_{\phi}^{\Gamma_p^{\{k,0\}}}}^{(k)} \end{pmatrix}.$$

Au niveau des domaines, on a ainsi l'écriture simplifiée de $\tilde{\mathbf{J}}$, \mathbf{M}_c et Φ :

$$\begin{aligned} \tilde{\mathbf{J}} &= \begin{bmatrix} \tilde{\mathbf{J}}^{(0,1)} & \tilde{\mathbf{J}}^{(0,2)} & \dots & \tilde{\mathbf{J}}^{(0,N_{\Omega})} \\ \tilde{\mathbf{J}}^{(1,1)} & 0 & 0 & 0 \\ 0 & \tilde{\mathbf{J}}^{(2,2)} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \tilde{\mathbf{J}}^{(N_{\Omega},N_{\Omega})} \end{bmatrix}, & \mathbf{M}_c &= \begin{bmatrix} \mathbf{M}_c^{(1)} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{M}_c^{(N_{\Omega})} \end{bmatrix}, \\ \Phi &= \begin{pmatrix} \Phi^{(1)} \\ \vdots \\ \Phi^{(N_{\Omega})} \end{pmatrix}. \end{aligned}$$

5.6 Bilan sur la formulation modale sous-structurée

On a donc réussi à construire une formulation modale du problème sous-structuré dans le cadre de la méthode des joints. On a pu développer cette formulation tout en restant dans le cadre de l'analyse fonctionnelle. La formulation matricielle qui en découle est indépendante de la méthode de discrétisation spatiale utilisée : on a ainsi pu obtenir une formulation avec des éléments finis, mais aussi avec des modes de Dirichlet-Steklov, qui vaut pour des bases réduites ou non. Pour des géométries simples, les modes peuvent être calculés analytiquement,

et la base peut ainsi être aussi riche qu'on le souhaite. Le calcul numérique des modes n'est nécessaire que pour pallier à la complexité géométrique du domaine à modéliser.

Cette formulation ne fait pas non plus explicitement référence à des maillages, conformes ou non, des surfaces de contact, comme c'est le cas pour la méthode des joints "classique". C'est donc, comme on l'a évoqué dans le chapitre précédent, au moment du calcul des matrices de contact \mathbf{J} que repose la difficulté, qui nécessite deux projections sur deux maillages différents si on décide (et on le fera) de travailler avec des maillages non conformes aux interfaces.

Pour valider ce formalisme il reste à le tester numériquement : c'est l'objet de la partie suivante du manuscrit.

Troisième partie

Applications et exemples numériques

Chapitre 6

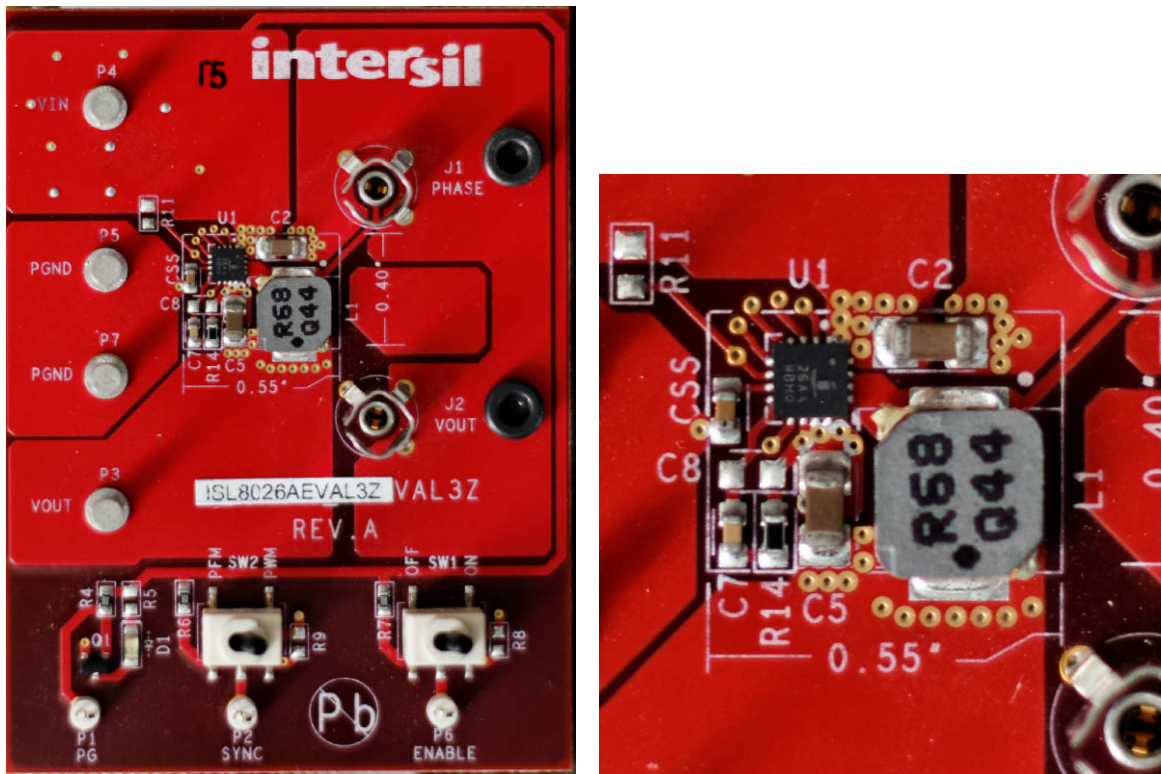
Modélisation par sous-structuration modale conforme : exemple de la carte d'évaluation d'un convertisseur de tension à découpage ISL8026

Dans ce premier exemple d'application un seul aspect de la méthode sera présenté : la réduction de modèle utilisant des bases de Dirichlet-Steklov réduites. Le modèle est tout de même sous-structuré, mais pas via la méthode des joints ; les sous-structures seront couplées en remplaçant le terme de flux à l'interface par son expression en fonction de la résistance de contact R_c . On rappelle que la formulation variationnelle du problème sous-structuré couplé par une résistance de contact est donnée en (2.4.9) (page 52). En conséquence, on devra se limiter à des maillages conformes aux interfaces entre les sous-structures.

En dépit de cette limitation, qui fait que les modèles réduits de composants créés ne pourront pas être réutilisés dans d'autres configurations, l'étude que l'on va présenter dans ce chapitre permettra tout de même de juger de la robustesse de la base de Dirichlet-Steklov pour reconstruire une large gamme de conditions aux limites. Si on se réfère au cahier des charges établi dans le chapitre 1 (section 1.4), on va ainsi mesurer la capacité des modèles réduits à répondre aux trois premières conditions.

6.1 Description de la carte de test

La carte test est celle déjà présentée comme exemple dans le chapitre 1 : la carte INTERSIL ISL8026EVAL3Z (voir figure 3.6.1), qui est un démonstrateur pour le composant ISL8026 [5].



(a) Vue de la surface supérieure.

(b) Zoom sur la surface supérieure.

FIGURE 3.6.1 – Vues de la carte Intersil ISL8026EVAL3Z.

Le choix de cette carte est motivé par plusieurs raisons :

- c’est une “vraie” carte électronique, qui permet de tester la méthode sur un cas concret plutôt que sur des cas plus “académiques” constitués de simples blocs homogènes apposés sur une plaque tout aussi homogène ;
- bien qu’elle soit constituée d’une vingtaine de composants, elle reste assez simple à reproduire via un outil de modélisation en 3 dimensions ;
- c’est une carte abordable (moins de 100€), ce qui rend possible des mesures expérimentales ;
- cette carte est très bien documentée, et elle a notamment déjà été étudiée par Thales par l’intermédiaire de Cheikh Dia [66], pour l’évaluation de leur propres méthodes de réduction de modèles.

Le composant ISL8026 est un convertisseur DC-DC qui peut délivrer un courant de 3 A pour une tension d’entrée pouvant varier de 2.7 à 5.5 V. Pour minimiser sa surface d’implantation sur le circuit imprimé, celui-ci est encapsulé dans un boîtier QFN16 de dimensions $3 \times 3 \times 0,75 \text{ mm}^3$.

Ce convertisseur est associé à 7 résistances CMS (pour “Composants Montés en surface”) de dimension normalisée 0603, 5 capacités de dimension 0603 et 4 de dimension 1206. Ces composants sont répartis sur la face supérieure (figure 3.6.1) ainsi que sur la face inférieure de la carte. Les dimensions 0603 et 1206 correspondent à des standards utilisés par les fabricants de composants. Les composants 0603 sont de dimensions $1,55 \times 0,85 \times 0,45 \text{ mm}^3$, tandis que les composants 1206 sont de dimensions $3,2 \times 1,6 \times 0,55 \text{ mm}^3$.

Une inductance, de dimensions $7,3 \times 6,6 \times 2,4 \text{ mm}^3$ est aussi placée en série, et peut être traversée par un courant intense, et donc dégager une puissance thermique importante. Ces différents composants sont placés sur un circuit imprimé de dimensions $70,12 \times 51,59 \times 0,73 \text{ mm}^3$. Le contact entre les composants et la carte fait apparaître une résistance de contact.

Lors de la mise en fonctionnement, et du fait de rendements imparfaits, les différents composants dissipent une puissance thermique.

6.2 Fonctionnement du convertisseur DC-DC. Estimation de la puissance thermique dissipée

Nous présentons ici le principe simplifié d’un hacheur de courant pour expliquer la provenance de la dissipation thermique. Une explication plus détaillée est fournie en annexe.

6.2.1 Principe de fonctionnement d’un convertisseur DC-DC.

Les convertisseurs de tension continu-continu ont pour fonction de fournir une tension continue variable à partir d’une tension continue fixe. La figure 3.6.2 donne une schématisation “naïve” d’un hacheur de tension (on parle aussi de régulateur à découpage). C’est un simple interrupteur (cellule de commutation) qui s’ouvre et se ferme périodiquement. Si la période de hachage est T l’interrupteur est fermé pendant le temps αT . Le facteur α est appelé facteur cyclique ou facteur de modulation d’impulsion.

Nous pouvons calculer la tension de sortie moyenne U_R :

$$U_R = \langle u_R \rangle = \frac{1}{T} \int_0^T u_R(t) dt = \frac{1}{T} \int_0^{\alpha T} U_e dt + \frac{1}{T} \int_{\alpha T}^T 0 dt.$$

Soit

$$\alpha = \frac{U_R}{U_e}. \quad (3.6.1)$$

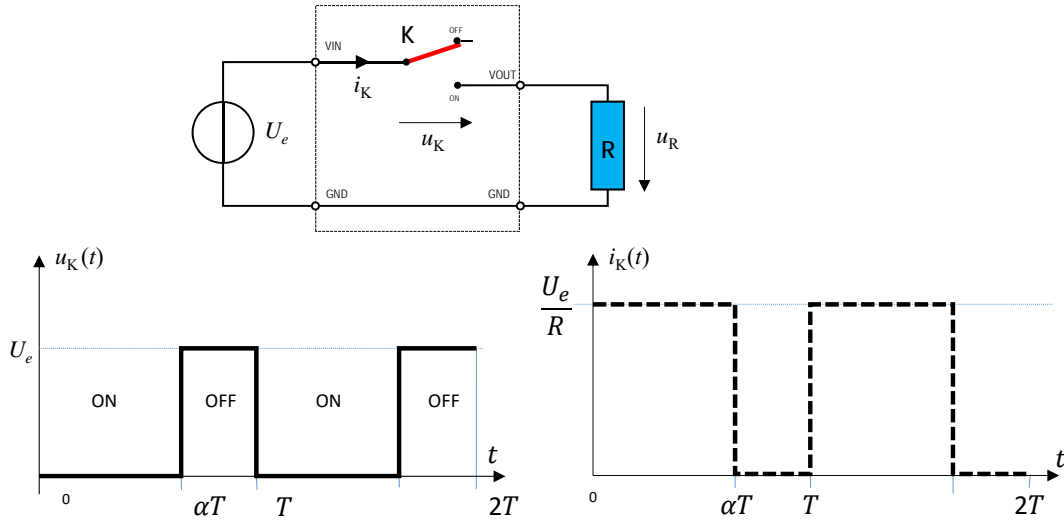


FIGURE 3.6.2 – Schéma de principe d'un hacheur de courant élémentaire. Digramme temporel de tension et d'intensité.

Pour un convertisseur parfait la puissance d'entrée est égale à la puissance de sortie, ce qui donnerait pour les intensités moyennes :

$$\alpha = \frac{U_R}{U_e} = \frac{I_e}{I_R}.$$

Bien entendu si l'idée est juste ce schéma est beaucoup trop simple. En particulier tension et courant sont totalement discontinus ce qui n'est pas acceptable. A elle seule la cellule de commutation est insuffisante pour réaliser un hacheur de courant. Il faut lui adjoindre un dispositif de stockage/dé-stockage rapide d'énergie. La figure 3.6.3 montre un tel dispositif. Le hacheur est alimenté par un générateur de courant délivrant une tension U_e . La charge du hacheur est constituée ici d'une simple résistance R . L'inductance L et la capacité C forment un filtre passe-bas dont le but est de limiter l'ondulation résultant du découpage sur la tension et le courant de sortie. Si ces éléments sont correctement dimensionnés, on peut supposer que $i_s(t)$ et $u_r(t)$ sont continus. La figure 3.6.4 montre les deux phases du cycle de fonctionnement d'un hacheur.

Phase $0 \leq t < \alpha T$. Les interrupteurs K_1 et K_2 sont respectivement fermé et ouvert. La tension aux bornes de K_1 est nulle, et le courant qui traverse K_1 vaut $i_{K1} = i_L$. À l'inverse, il n'y a pas de courant traversant K_2 , et la tension aux bornes de cet interrupteur vaut U_e (voir figure 3.6.5).

Phase $\alpha T \leq t < T$. Cette fois les interrupteurs K_1 et K_2 sont respectivement ouvert et fermé. La tension aux bornes de K_2 est nulle, et le courant qui le traverse vaut $i_{K2} = i_L$. Le

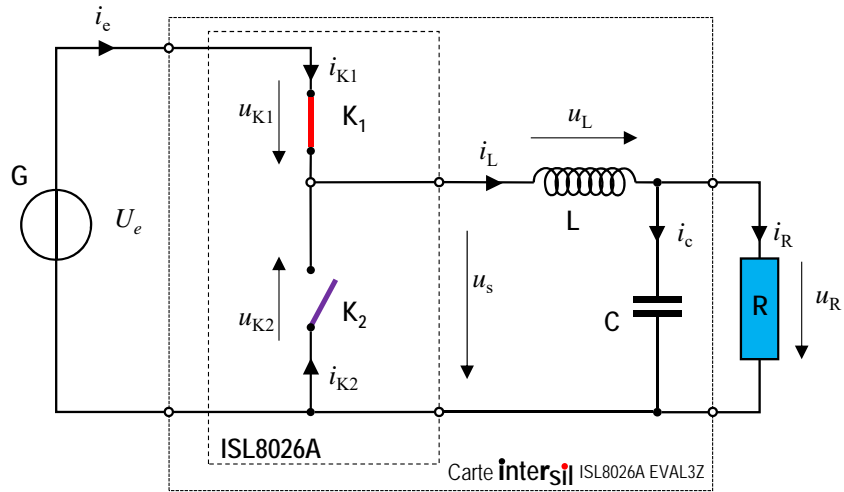


FIGURE 3.6.3 – Schéma de principe d'un hacheur de courant abaisseur de tension.

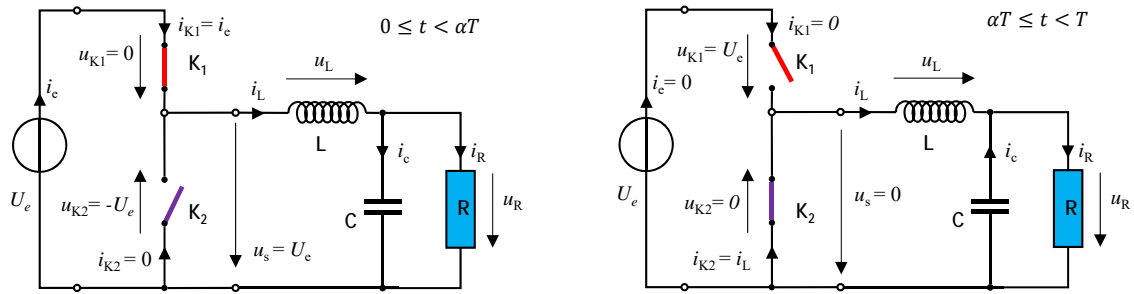


FIGURE 3.6.4 – Phases de fonctionnement d'un hacheur de courant.

courant traversant K_1 est nul, et la tension aux bornes de cet interrupteur vaut U_e (voir figure 3.6.5).

Un modèle de circuit RLC (les calculs sont détaillés en annexe) permet d'avoir une estimation des grandeurs I_M et I_m , ainsi que l'expression de I_L . En particulier, on démontre que :

$$I_e = \alpha I_L \quad \text{avec} \quad I_L = \langle i_L(t) \rangle . \quad (3.6.2)$$

De plus, sous l'approximation $\tau = R/L \gg T$ (qui correspond au critère de dimensionnement de l'inductance par rapport à la fréquence de hachage et à la résistance de charge), le bilan de puissance donne

$$\Pi_R = \Pi_e \sim \alpha^2 \frac{U_e^2}{R} . \quad (3.6.3)$$

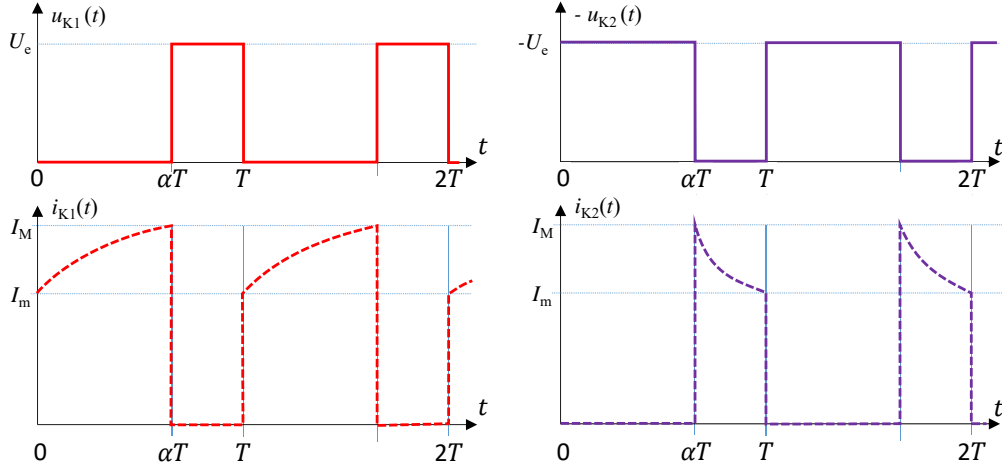


FIGURE 3.6.5 – Diagramme temporel de l'évolution de la tension aux bornes des interrupteurs K_1 et K_2 et du courant les traversant.

Puisque $U_R = \alpha U_e$ on peut en déduire que l'intensité efficace dans la résistance est $I_R = \frac{1}{\alpha} I_e$.

6.2.2 Origine de la dissipation thermique dans le hacheur.

En réalité le hacheur n'est pas parfait et une partie de la puissance fournie par le générateur de tension est dissipée en chaleur. C'est cette puissance qui nous intéresse particulièrement pour nos modèles thermiques. On définit le rendement η du hacheur par :

$$\Pi_e = \Pi_R + \Pi_{th}, \quad \eta = \frac{\Pi_R}{\Pi_e}. \quad (3.6.4)$$

Expression où Π_{th} est la puissance thermique dissipée dans le hacheur. On peut la scinder en deux parties :

- Dissipation dans la cellule de commutation [97] :
 - Les transistors (les interrupteurs) possèdent une résistance interne. Dans le cas de l'ISL8026A le constructeur indique les valeurs suivantes : $R_p = 36 m\Omega$ et $R_n = 13 m\Omega$. Il y a donc des pertes par effet Joule.
 - Les fronts montants et descendants ne sont pas des échelons parfaits mais plutôt des rampes. Pendant la durée de commutation, notée δ , le produit de la tension par l'intensité n'est pas nul. Il y a donc une perte par commutation. Si on assimile la variation à une rampe alors la puissance perdue à chaque commutation du transistor est $UI/6$ (on a $U_e I_e = U_R I_R$ et $\Delta I_R \ll I_R$). Il y a 4 commutations par cycle. Les mesures donnent approximativement $\delta \sim 10 ns$.

On peut tenter de construire un modèle élémentaire de dissipation dans le QFN pour voir l'influence des paramètres :

$$\Pi_{th}^{QFN} = \alpha R_p I_e^2 + (1 - \alpha) R_n I_R^2 + \frac{2}{3} \delta f U_e I_e. \quad (3.6.5)$$

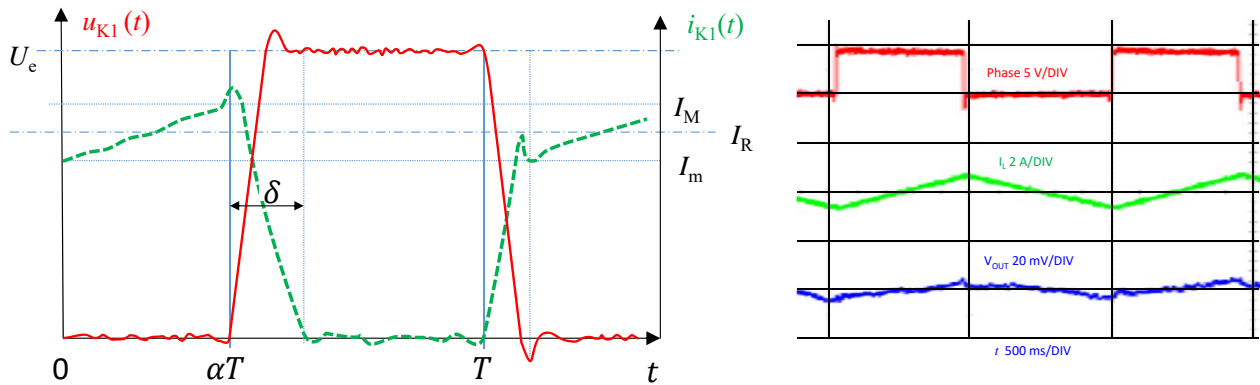


FIGURE 3.6.6 – Diagramme temporel montrant le défaut de commutation. À gauche schéma de principe (défaut amplifiés), à droite enregistrement des signaux sur un oscilloscope numérique.

- Dissipation dans l'inductance [98] :
 - Le fil de la bobine de l'inductance dissipe de l'énergie par effet Joule. Si on note r_L la résistance de la bobine, la puissance dégagée sera $r_L I_L^2$. La résistance de la bobine serait faible en régime continu. Cependant au delà de quelques centaines de MHz il faut considérer l'effet de peau qui augmente considérablement la perte par effet Joule. C'est donc le cas pour l'ISL8026A puisque qu'il fonctionne à une fréquence de $2 MHz$. À cette fréquence l'épaisseur de peau pour un fil de cuivre est de $47 \mu m$, que l'on peut rapporter au diamètre du fil qui est de $400 \mu m$. Logiquement la résistance sera une fonction de la fréquence et la puissance sera dissipée en "surface" du fil.
 - Il y a aussi des pertes par hystérésis (le champ magnétique s'inverse périodiquement) et par courants de Foucault. Elles sont proportionnelles respectivement à la fréquence et au carré de la fréquence du courant qui traverse la bobine. Dans notre cas la bobine de l'inductance est enrobée dans un matériau magnétique métallique ("integrated molded coil"). La puissance sera dissipée dans la totalité du volume.
 - Les pertes par effet condensateur entre les spires de la bobine ne sont significatives que pour des très hautes fréquences (VHF, UHF, c'est à dire au delà de $30 MHz$) ce qui n'est pas notre cas.

Le fabricant de l'inductance montée sur la carte (TDK) donne une formule de dissipation thermique simple :

$$\Pi_{th}^L = R_{ac} I_R^2. \quad (3.6.6)$$

La résistance R_{ac} globalise l'ensemble des phénomènes générant la dissipation thermique. C'est une fonction de la fréquence. Elle est donnée sur la figure 3.6.7.

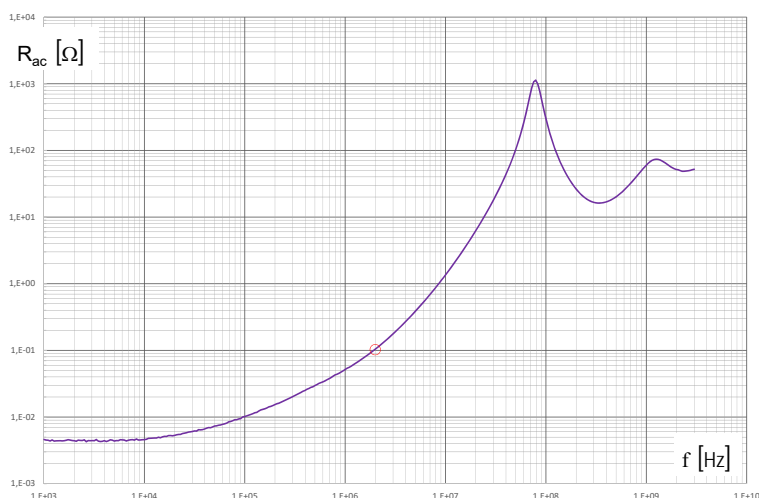
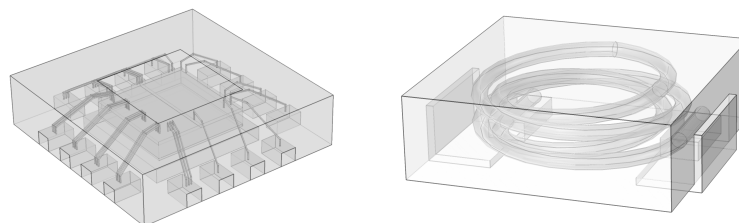


FIGURE 3.6.7 – Variation de la résistance d'une inductance. D'après document TDK.

Notons que la dissipation dans les condensateurs est marginale. Il en est de même dans les résistances utilisées pour la partie contrôle.

6.3 Modélisation de la carte de test

On fixe la puissance dissipée par la carte et les composants à 0.5W. Une partie de cette puissance est absorbée par la carte qui chauffe à son tour. Ainsi, tous les composants sont couplés par la conduction dans la carte. Pour prédire le comportement thermique d'un composant, il est donc nécessaire de simuler la carte dans son intégralité. La nature même du problème incite à le traiter par sous-structuration. Une division naturelle est de suivre le découpage fonctionnel de la carte : chaque composant sera considéré comme une sous-structure, et l'ensemble des couches du PCB est vu comme une seule sous-structure. Le modèle est donc un couplage de 19 sous-structures.



(a) Boitier QFN16 du composant ISL8026. (b) Boitier de l'inductance.

FIGURE 3.6.8 – Boitiers QFN16 et de l'inductance.

6.3.1 Modélisation des composants

Le convertisseur de tension est composé d'un support en OLIN C194 (un alliage cuivre-fer), de 16 broches du même matériau, et d'une puce en silicium collée sur le support à l'aide d'une colle ABLESTIK 71-1. Les fils de connexion entre la puce et les broches, qui entraînent un fort drain thermique, ont été pris en compte. Ils sont aussi en OLIN C194, mais certains composants peuvent avoir des fils de connexion en or (comme les BGA présentés dans le chapitre 1). Le tout est coulé dans une boîte en résine SUMITOMO-EME, formant ainsi un boitier de type QFN16 (figure 3.6.8a). La fonction n'est appliquée que sur une partie de la puce, ce qui entraîne une dissymétrie du champ thermique.

L'inductance (figure 3.6.8b) est constituée de deux broches en cuivre reliées par un solénoïde, en cuivre aussi. Pour cette modélisation, on suppose que le matériau englobant le solénoïde a les mêmes propriétés que la résine du QFN16, car on manque d'informations sur les propriétés thermiques et magnétiques de cette résine ferrique. Toutefois, on s'accorde cette imprécision, puisque comme on le verra plus tard, l'inductance, qui est un composant passif, n'est pas celui présentant la température critique atteinte par le système.

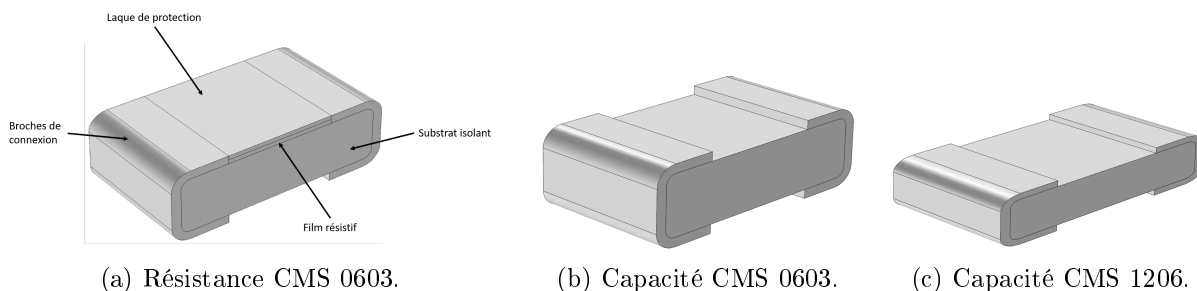


FIGURE 3.6.9 – Composants Montés en Surface de la carte.

Matériau	Conductivité thermique $K (W.m^{-1}.K^{-1})$	Masse volumique $\rho (kg.m^{-3})$	Capacité thermique massique $c_p (J.kg^{-1}.K^{-1})$
Cuivre	401	8960	385
OLIN C194	260	8900	385
Silicium	144	2330	708
Colle ABLESTIK 71-1	2.1	1900	920
Résine SUMITOMO-EME	0.67	1900	920
Nickel	90.7	8902	440
Alumine	23.5	3970	880
Ruthenium	50	6970	182
Époxy	2.1	1900	920

TABLE 3.6.1 – Propriétés physiques des matériaux utilisés pour les composants

Les résistances CMS (figure 3.6.9a) sont constituées d'un substrat isolant, généralement une alumine, et de broches de connexion en nickel reliées sur le dessus du composant par un film résistif métallique (ici du ruthenium), protégé par une laque de protection isolante en époxy. La longueur et l'épaisseur de ce film résistif déterminent la résistance ohmique. Ici, toutes les résistances sont identiques.

Les capacités (figures 3.6.9b et 3.6.9c) ne sont quant à elles constituées de deux broches en nickel reliées par un substrat en alumine.

Les propriétés thermiques de l'ensemble des matériaux utilisés pour modéliser les composants sont présentées dans la table 3.6.1

Après modélisation, ces composants sont discrétisés par un maillage tétraédrique P1 (voir figure 1.2.8). Le maillage du QFN16 est constitué de 163 000 nœuds (figure 3.6.10a), celui de l'inductance est constitué de 47 000 nœuds (figure 3.6.10b), chaque résistance et capacité 0603 est maillée avec environ 1 500 nœuds (3.6.10c, 3.6.10d) et les capacités de dimension 1206 avec environ 3 500 nœuds (3.6.10e).

6.3.2 Modélisation du PCB

Le PCB de la carte est constitué de 7 couches (voir figure 3.6.11), numérotées de 1 à 7 en commençant par la couche inférieure : les couches 1, 3, 5 et 7 (figure 3.6.12) sont des feuillards de cuivre de $35 \mu m$ d'épaisseur et les couches 2, 4 et 6 sont des couches diélectriques en résine

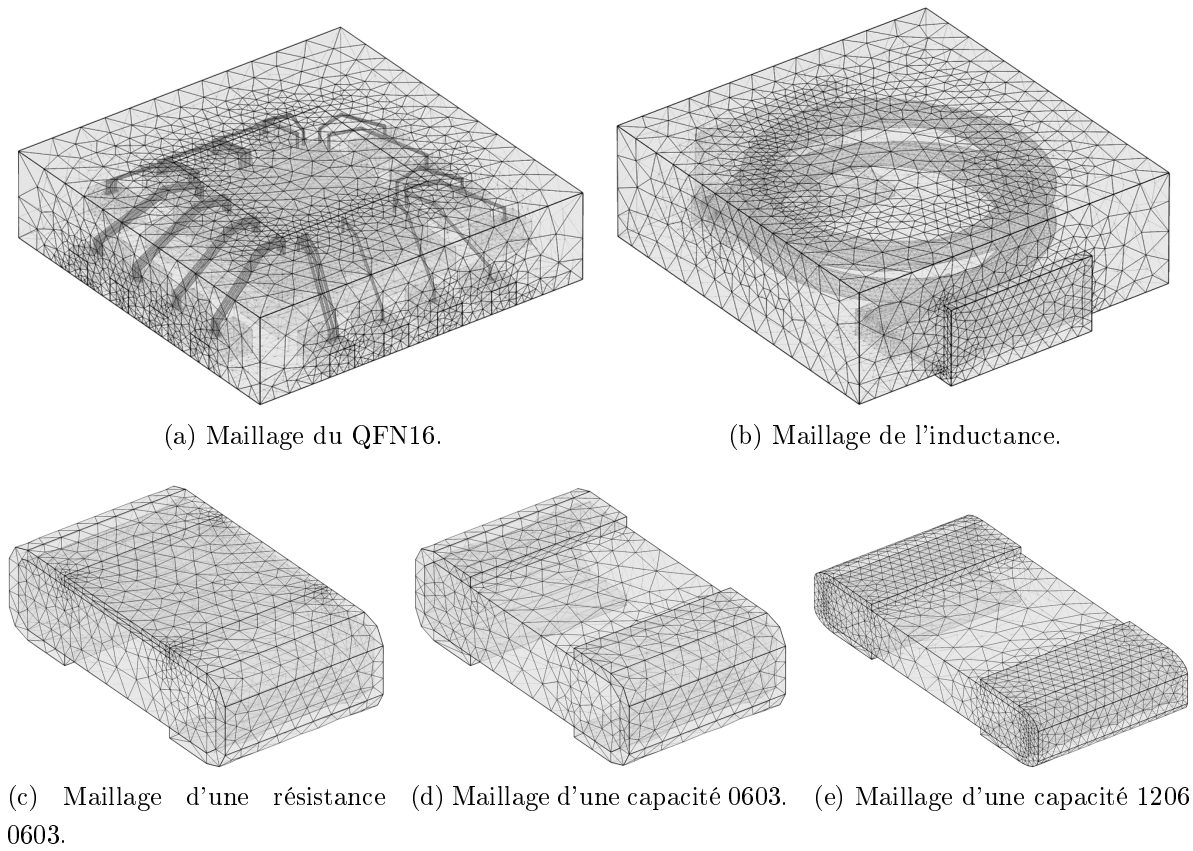


FIGURE 3.6.10 – Maillages des composants.

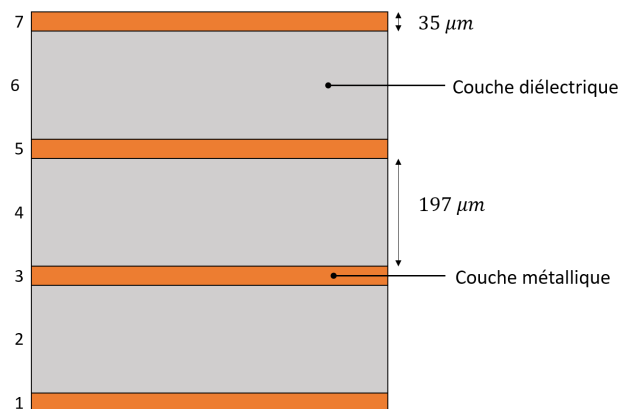


FIGURE 3.6.11 – Couches de la carte de test ISL8026EVAL3Z.

époxy, de $197 \mu m$ d'épaisseur.

Pour le modéliser, on peut décider de prendre en compte tous les détails du routage de la carte, ce qui donnerait le modèle 3D présenté en figure 3.6.13. Mais comme on l'a dit dans le

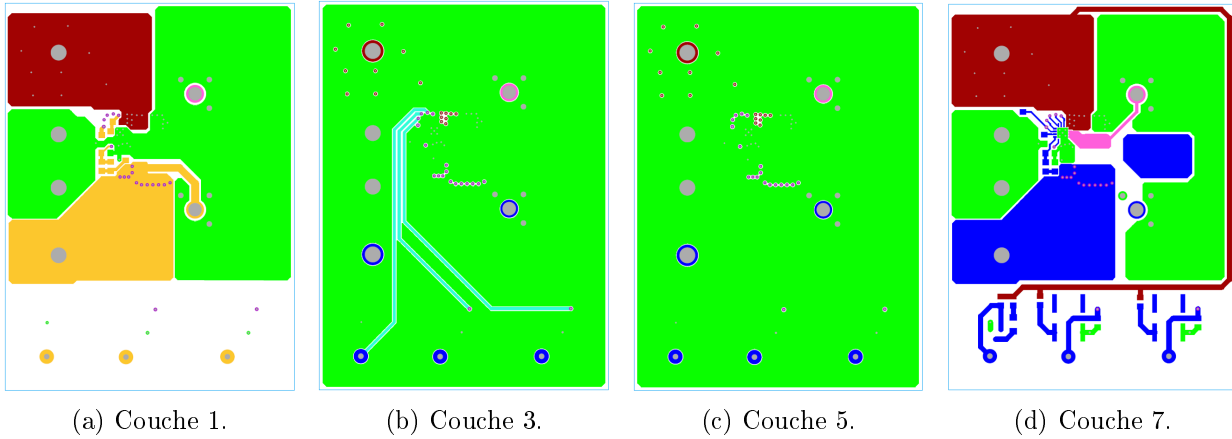


FIGURE 3.6.12 – Couches métalliques de la carte de test.

chapitre 1, le maillage dépasserait aisément le million de nœuds. L’objectif est bien de réduire le modèle de PCB, mais un obstacle d’ordre matériel nous empêche de modéliser finement le PCB : la machine que j’utilise pour implémenter et tester la méthode de réduction est une station de travail portable munie d’un processeur Xeon cadencé à 2.80 GHz et de 32Gb de mémoire RAM, assez puissant pour la plupart des calculs, mais pas pour des modèles de plusieurs millions de nœuds. Et bien qu’on puisse calculer et réduire les bases de Dirichlet-Steklov du PCB pour obtenir un modèle réduit qui puisse être traité, on a tout de même besoin d’une simulation du modèle complet pour pouvoir quantifier l’erreur due à la réduction.

Pour ce faire, on pourrait décider de modéliser une couche homogène munie d’une conductivité anisotrope équivalente au modèle à 7 couches, mais des tests précédents ont montré que cette simplification était bien trop grossière, avec notamment une très large sous-estimation de la conductivité suivant l’axe vertical.

La méthode retenue est celle de l’approche “Metal fraction” [99], utilisée par Thales, entreprise partenaire de cette thèse. Le principe est le suivant : chaque couche est modélisée finement et maillée par des quadrilatères réguliers. La fraction métallique présente dans chaque cellule (qui fait toute l’épaisseur de la couche considérée) est ensuite calculée, pour donner une valeur variant de 0 à 1 : une cellule avec une fraction métallique de 1 est ainsi entièrement constituée de métal (ici du cuivre).

On peut donc calculer une conductivité équivalente pour chaque cellule, donnée par la formule $K_{equiv} = m_f * K_{cu} + (1 - m_f) * K_{ex}$, avec m_f la valeur de la fraction métallique dans la cellule, et K_{cu} et K_{ex} les conductivités du cuivre et de la résine époxy respectivement (données dans la table 3.6.2¹). On obtient la masse volumique et la capacité thermique massique de la

1. Attention, ce sont des valeurs différentes de celles données dans la table 3.6.1, car ce ne sont pas

Matériau	Conductivité thermique K ($W.m^{-1}.K^{-1}$)	Masse volumique ρ ($kg.m^{-3}$)	Capacité thermique massique c_p ($J.kg^{-1}.K^{-1}$)
Cuivre	385	8933	397
Résine époxy	0.3	1250	1300

TABLE 3.6.2 – Propriétés physiques des matériaux utilisés pour le PCB

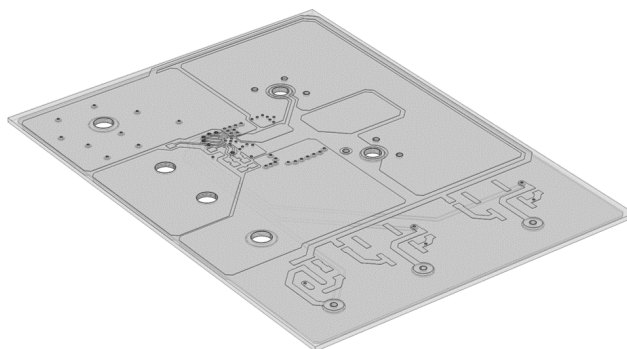


FIGURE 3.6.13 – Modélisation détaillée du PCB.

même façon. On obtient un champ de conductivité (respectivement de masse volumique et de capacité thermique massique) que l'on va pouvoir projeter sur n'importe quel autre maillage.

On modélise donc 7 couches homogènes, de même épaisseurs respectives que les couches décrites sur la figure 3.6.11. Ces couches sont ensuite discrétisées par un maillage tétraédrique éléments finis sur lequel est projeté le champ de conductivité de chaque couche, obtenu grâce à la “Metal Fraction” fournie par Thales (figure 3.6.14). Pour rendre le champ de conductivité projeté sur la couche supérieure (la couche 7, voir figure 3.6.14b) plus représentatif du routage de la carte, on a resserré le maillage du PCB dans la zone où sont implantés les composants. Le maillage du PCB est donc constitué de seulement 183 000 nœuds, au lieu des millions nécessaires dans le cas d'une modélisation fine des circuits de chaque couche.

Le maillage total de la carte est finalement constitué de 422 000 nœuds.

6.3.3 Résistance de contact

Comme dit en introduction de ce chapitre, on établit le couplage entre les 19 sous-structures constitutives de cette carte en remplaçant le flux aux interfaces par son expression en fonction du saut de température et de la résistance de contact R_c , suivant la formulation variationnelle donnée en (2.4.9). Pour choisir la valeur de cette résistance, on s'inspire de

exactement les mêmes matériaux (différente pureté du cuivre et une résine époxy de composition différente).

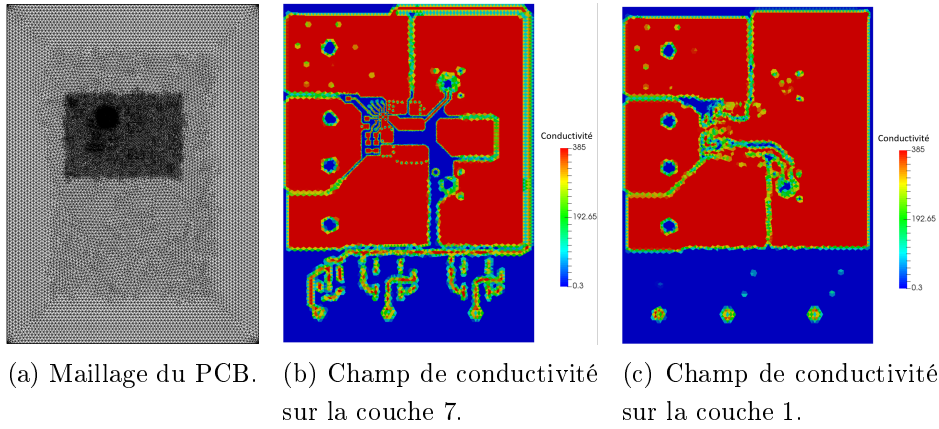


FIGURE 3.6.14 – Projection de la “Metal Fraction” sur le maillage du PCB.

Configuration de contact	R_c ($m^2.K.W^{-1}$)
Micro-électronique (brasage)	$\in [10^{-9}, 10^{-7}]$
Collage	$\in [10^{-6}, 10^{-4}]$
Assemblage par serrage	$\in [10^{-4}, 10^{-2}]$

TABLE 3.6.3 – Différentes valeurs de R_c en fonction du contact [6]

Battaglia *et al.* [6] qui fournissent des valeurs de résistances de contact mesurées dans des configurations réelles en fonction du contact appliqué.

D’après la table 3.6.3, on choisit d’appliquer un coefficient R_c de $10^{-7} m^2.K.W^{-1}$ entre chaque composant et le PCB. On pourrait prendre une valeur encore plus faible, mais le modèle utilisant un couplage des sous-structures via une résistance de contact atteint alors sa limite, et le solveur numérique utilisé pour résoudre le système matriciel correspondant converge beaucoup plus lentement, la contrainte imposée aux interfaces étant trop forte.

6.4 Réduction du modèle

6.4.1 Formulation du problème réduit modal

Le modèle réduit est obtenu en exprimant dans l’équation (2.4.9) la température par sa formulation modale réduite, et en remplaçant les fonctions test par la base des vecteurs propres. En effet, à condition de partager le même espace fonctionnel, toute fonction f est décomposable sur la base de vecteurs propres. On trouve alors l’équation que doivent satisfaire les états $x_i^{\mathcal{X},(k)}$. Pour faciliter la lecture, nous nous limitons au cas où seuls deux domaines

sont en contact via une seule frontière. On pourra se référer pour plus de détails à la thèse P.O. Laffay [89] ou celle de B. Gaume [100].

Définition 6.4.1

$$\begin{aligned}
& q, k \in \{1, 2\} , \quad q \neq k , \quad \mathcal{X} \in \{D, S\} , \quad \forall j \in \{1, N^{(k)}\} , \forall l \in \{1, N^{(1)}\} , \forall m \in \{1, N^{(2)}\} \\
& \sum_{k=1,2} \sum_{i=1}^{N^{(k)}} \frac{\partial x_i^{\mathcal{X},(k)}}{\partial t} \int_{\Omega^{(k)}} C^{(k)} V_i^{\mathcal{X},(k)} V_j^{\mathcal{Y},(k)} + \sum_{k=1,2} \sum_{i=1}^{N^{(k)}} x_i^{\mathcal{X},(k)} \int_{\Omega^{(k)}} \nabla V_i^{\mathcal{X},(k)} \cdot K^{(k)} \cdot \nabla V_j^{\mathcal{Y},(k)} \\
& \quad + \sum_{k=1,2} \sum_{i=1}^{N^{\mathcal{S},(k)}} x_i^{\mathcal{S},(k)} \int_{\Sigma^{(k)}} h^{(k)} V_i^{\mathcal{S},(k)} V_j^{\mathcal{S},(k)} \\
& \quad + \frac{1}{R_c} \int_{\Gamma} (V_m^{\mathcal{S},(2)} - V_l^{\mathcal{S},(1)}) \left(\sum_{i=1}^{N^{\mathcal{S},(2)}} x_i^{\mathcal{S},(2)} V_i^{\mathcal{S},(2)} - \sum_{i=1}^{N^{\mathcal{S},(1)}} x_i^{\mathcal{S},(1)} V_i^{\mathcal{S},(1)} \right) \\
& \quad = \sum_{k=1,2} \int_{\Omega^{(k)}} \varpi^{\mathcal{Y},(k)} V_j^{(k)} + \sum_{k=1,2} \int_{\Sigma^{(k)}} h^{(k)} T_{ext} V_j^{\mathcal{S},(k)} . \quad (3.6.7)
\end{aligned}$$

Pour obtenir le modèle réduit de la carte, la prochaine étape consiste à calculer et réduire les bases de chaque sous-structure. On résout pour cela les systèmes matriciels issus des problèmes aux valeurs propres de Dirichlet (2.5.18) et de Steklov (2.5.22), à l'aide d'une méthode de Lanczos [101] implémentée dans la librairie Arpack [102].

6.4.2 Choix du paramètre de Steklov

Pour comprendre la problématique liée au choix du paramètre de Steklov, prenons l'exemple suivant : deux pavés droits de dimensions respectives $4 \times 1 \times 0,5 \text{ mm}^3$ et $4 \times 2,2 \times 0,5 \text{ mm}^3$ sont couplés par une résistance de contact. Le pavé le plus petit (en bleu sur la figure 3.6.15) dissipe une puissance thermique de 10^9 W.m^{-3} et a comme propriétés physiques $\kappa = 260 \text{ W.m}^{-1}.\text{K}^{-1}$ et $C = 3,43 \times 10^6 \text{ J.m}^{-3}.\text{K}^{-1}$. L'autre sous-structure est caractérisée par $\kappa = 0,67 \text{ W.m}^{-1}.\text{K}^{-1}$ et $C = 1,75 \times 10^6 \text{ J.m}^{-3}.\text{K}^{-1}$. Bien que très simples géométriquement, ces deux domaines ont les ordres de grandeurs caractéristiques de l'électronique. Ces deux solides échangent de la chaleur avec le milieu extérieur à une température de 0°C via un coefficient d'échange $h = 10 \text{ W.m}^{-2}.\text{K}^{-1}$.

Une simulation éléments finis est effectuée (le champ en régime permanent est donné à titre d'exemple sur la figure 3.6.15). Une base modale est calculée, puis réduite grâce à la méthode d'amalgame : 10 modes sont conservés pour chacune des structures. Le champ de température est alors projeté sur la base réduite. On rappelle les propriétés d'orthogonalité :

$$(V_i^{\mathcal{X}} | V_j^{\mathcal{Y}})_H = \delta_i^j \delta_{\mathcal{X}}^{\mathcal{Y}}.$$

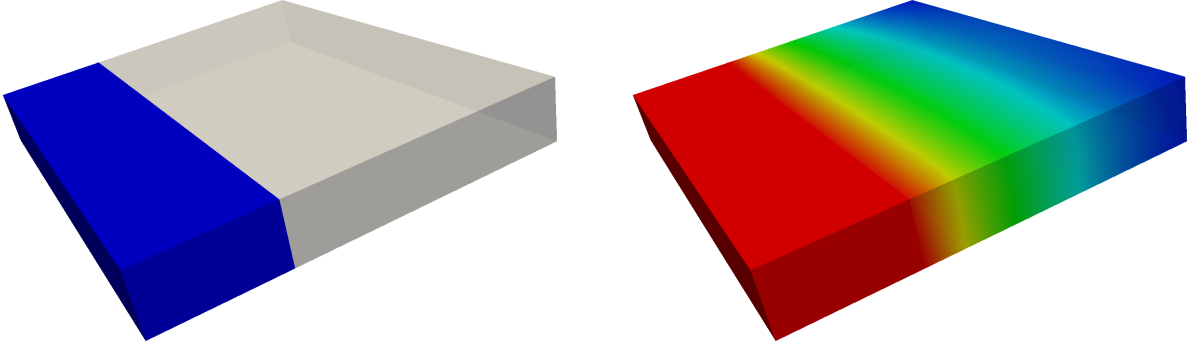


FIGURE 3.6.15 – Géométrie traitée (à gauche) et résultat de la simulation (à droite).

Comme $T = \sum_{\mathcal{X} \in \{D, S\}} \sum_{I=1}^{N^{\mathcal{X}}} x_I^{\mathcal{X}} V_I^{\mathcal{X}}$, on a

$$\left(T | \tilde{V}_j^{\mathcal{X}} \right)_H = \tilde{x}_j^{\mathcal{X}}.$$

Il est alors possible de reconstituer la température selon :

$$\tilde{T}_{proj} = \sum_{\mathcal{X} \in \{D, S\}} \sum_{I=1}^{\tilde{N}^{\mathcal{X}}} \tilde{x}_I^{\mathcal{X}} \tilde{V}_I^{\mathcal{X}}.$$

Évidemment, cette opération n'a aucun aspect pratique, puisqu'il faut connaître le champ de température pour déterminer \tilde{T}_{proj} . Cependant, cette projection est utile pour déterminer la "qualité" de la base réduite. En effet, à cause de la réduction, $\tilde{T}_{proj} \neq T$. On peut alors définir un écart entre ces deux champs qui caractérisera la qualité de la base. Cet écart est l'horizon à atteindre pour l'écart entre le champ de température de référence et le champ obtenu à l'issue de la simulation réduite. Finalement, une simulation modale réduite est exécutée selon l'équation (2.4.9). Un champ de température réduit est ainsi obtenu, noté \tilde{T} .

Une analyse de sensibilité à la résistance de contact est conduite. Les résultats sont présentés sur la figure 3.6.16. Les écarts quadratique entre la simulation éléments finis et d'une part la température réduite projetée et d'autre part la température modale simulée sont représentés. Pour des valeurs modérément faibles de la résistance de contact, les écarts sont très faibles (on parle de 0.2 K d'écart sur une plage de température montant jusqu'à 6600 K), et les écarts projetés et simulés sont du même ordre de grandeur. Lorsque la résistance de contact passe en dessous de $10^{-6} K.m^2.W^{-1}$ l'écart entre la température issue du modèle éléments finis et la température projetée reste constante. Par contre, la température modale

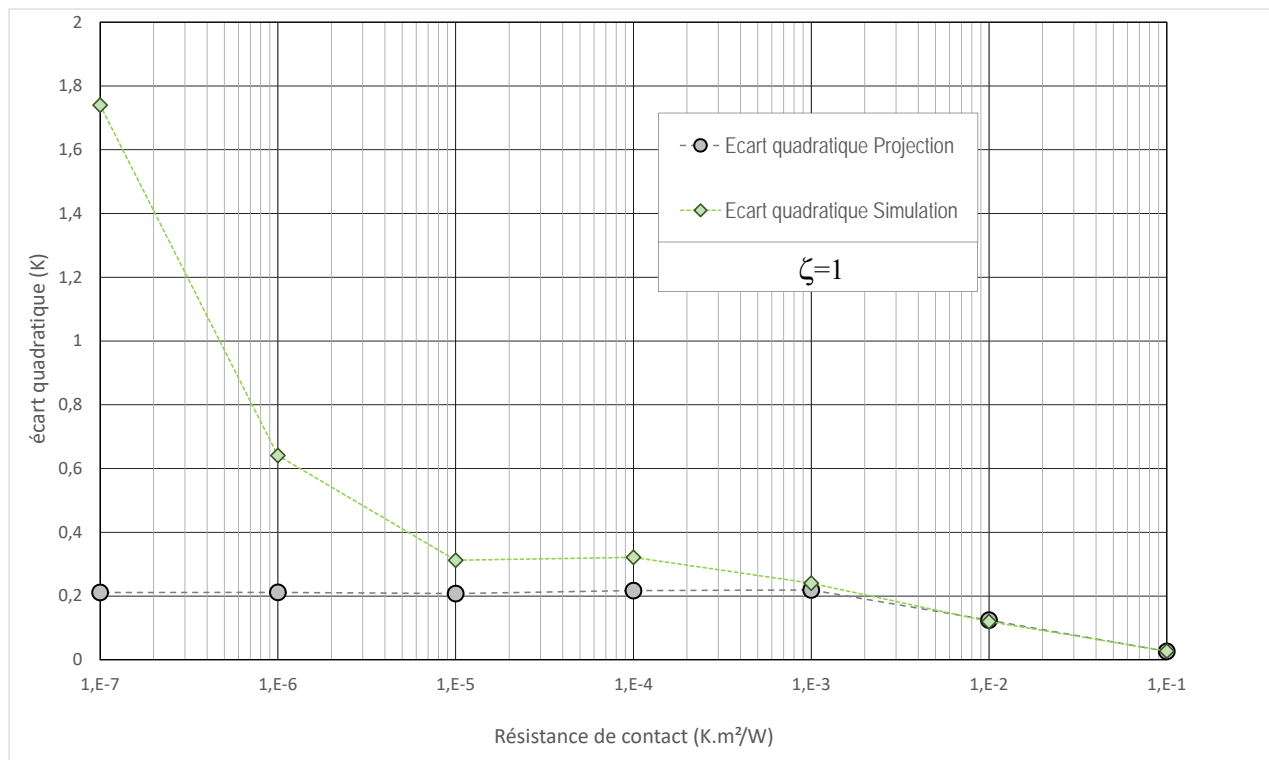


FIGURE 3.6.16 – Écart entre la simulation éléments finis et deux températures modales obtenues par projection et par simulation, en fonction de la résistance de contact.

simulée s'éloigne de la simulation de référence, et l'écart devient près de dix fois supérieur à l'écart de projection. En d'autres termes, l'écart réellement obtenu est dix fois plus grand que ce que l'on est en droit d'espérer. Même si cet écart reste raisonnable, il est gênant puisqu'il apparaît dans la gamme de résistance de contact qui nous intéresse.

Pour comprendre ce phénomène, essayons d'analyser le champ de température donné en figure 3.6.15. Comme la conductivité de la sous-structure 1 est élevée, le champ de température dans cette partie est quasiment uniforme. Ainsi, si on isole la sous-structure 2, elle est d'un côté refroidie par le milieu extérieur avec un coefficient d'échange de $h = 10 \text{ W.m}^{-2}.\text{K}^{-1}$, et de l'autre elle est chauffée par la sous-structure 1 à température quasi-uniforme mais avec un coefficient d'échange équivalent de $h = 10^7 \text{ W.m}^{-2}.\text{K}^{-1}$, ce qui entraîne une forte dissymétrie du champ de température. Or les modes avec lesquels nous avons fait la simulation sont calculés avec des conditions aux limites symétriques.

Afin d'améliorer la capacité des modes à reconstruire un champ de température dissymétrique, il est possible de les calculer en utilisant des conditions aux limites de Steklov non symétriques. Pour cela, on peut jouer sur la valeur de la fonction de Steklov $\zeta(x)$ présente dans l'équation (2.5.4). Cette dernière est choisie à 1 pour les frontières en contact avec le milieu extérieur, et une étude paramétrique est menée en fonction de sa valeur sur la frontière de contact. Pour cette étude, la valeur de la résistance de contact est égale à $10^{-7} \text{ K.m}^2.\text{W}^{-1}$. La figure 3.6.17 présente les résultats de cette étude en terme d'écart quadratique entre la solution éléments finis et les résultats issus de l'analyse modale. On remarque que l'écart de projection reste globalement constant lorsque ζ augmente. Ainsi, dissymétriser les modes n'influe pas sur la capacité de la base réduite à reconstruire la solution. Par contre, on remarque que l'écart de simulation (celui qui nous intéresse) décroît sensiblement (il y a près d'un facteur 10 sur l'écart quadratique) lorsque ζ passe de 1 à 100. Pour des valeurs plus grandes de ζ , la précision se maintient sur au moins 4 décades. Il n'y a pas de valeur optimale précise de ζ nécessitant une étude de sensibilité préalable : il suffira de prendre une valeur assez grande de ζ sur la frontière en contact (supérieure à 10^3 pour fixer les idées).

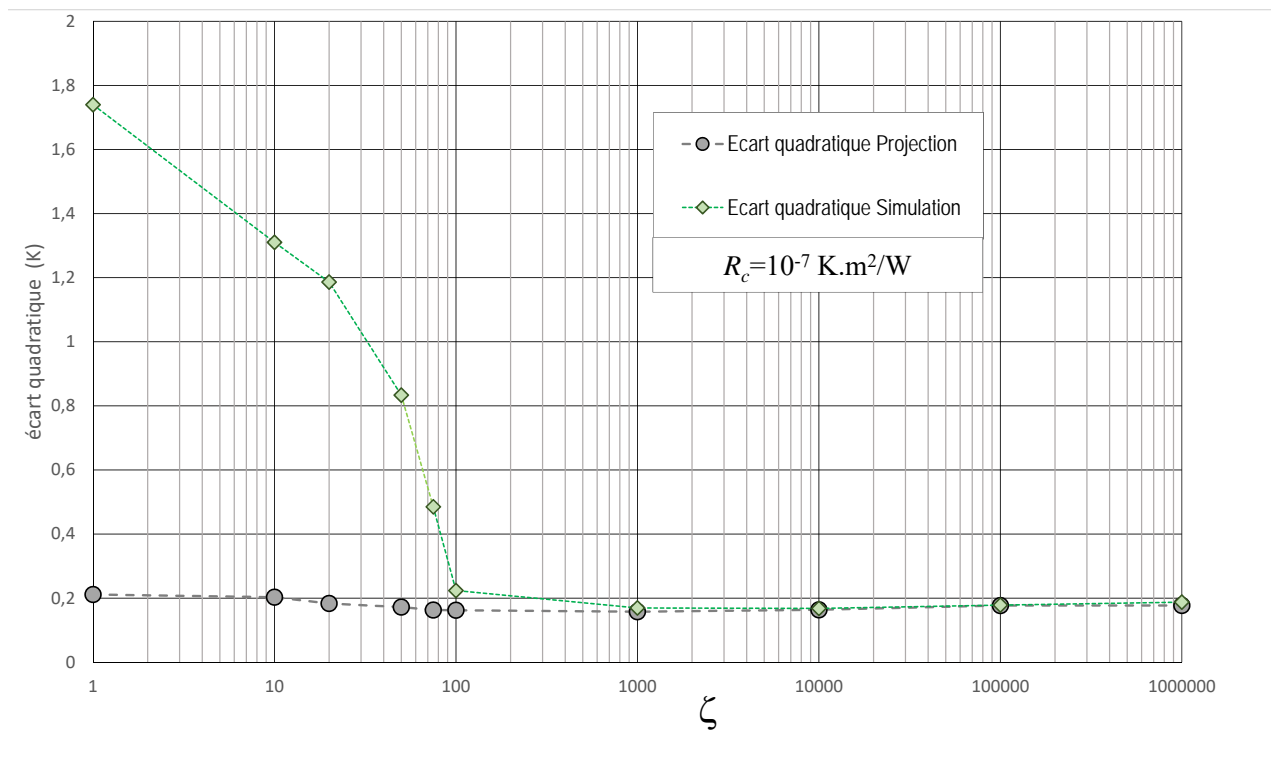


FIGURE 3.6.17 – Écart entre la simulation éléments finis et deux températures modales en fonction de la valeur du paramètre ζ sur la frontière de contact.

6.4.3 Calcul des modes

Puisque les systèmes matriciels à résoudre pour calculer les modes (équations (2.5.18) et (2.5.22)) sont issus des discrétisations en éléments finis des sous-structures, on peut calculer autant de modes de Dirichlet qu'il y a de nœuds dans le volume, et autant de modes de Steklov que le nombre de nœuds aux frontières. Par exemple, pour le QFN16, il y a seulement 3 585 nœuds en surface sur un total de 163 000 nœuds. On a donc calculé l'intégralité des modes de Steklov accessibles, mais seulement 10% des modes de Dirichlet, soit une base d'environ 20 000 modes avant réduction. La figure 3.6.18 présente des exemples de modes du QFN16. On observe ainsi que les modes de Dirichlet (figures 3.6.18a et 3.6.18b) sont localisés dans la résine, et on distingue l'empreinte des plots de connexions, du support et mêmes des fils. Les modes de Steklov (figures 3.6.18c et 3.6.18d), sont quant à eux localisés sur les faces et les arêtes, comme on avait pu l'observer pour les exemples de modes du pavé droit.

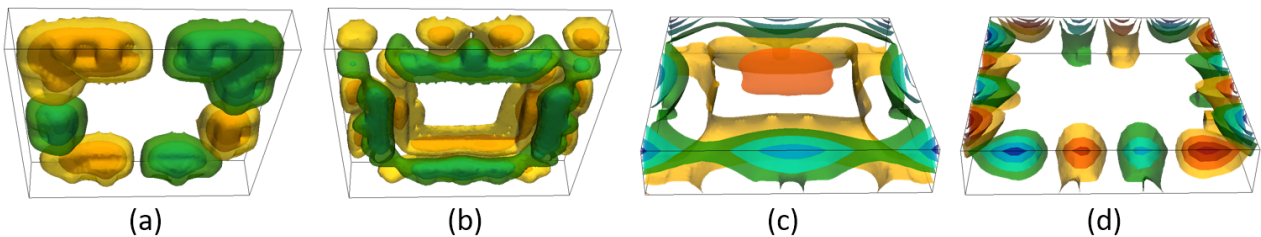


FIGURE 3.6.18 – QFN16 : (a) & (b) exemples de modes de Dirichlet (vue du dessous), (c) & (d) modes de Steklov (vue du dessus).

Pour l'inductance, on calcule le même ratio de modes de Dirichlet (10%) et de Steklov (100%), pour un total de 7 000 modes. On présente quelques exemples en figure 3.6.19. Le mode de Dirichlet en 3.6.19b est particulièrement reconnaissable puisqu'il fait apparaître le solénoïde de l'inductance.

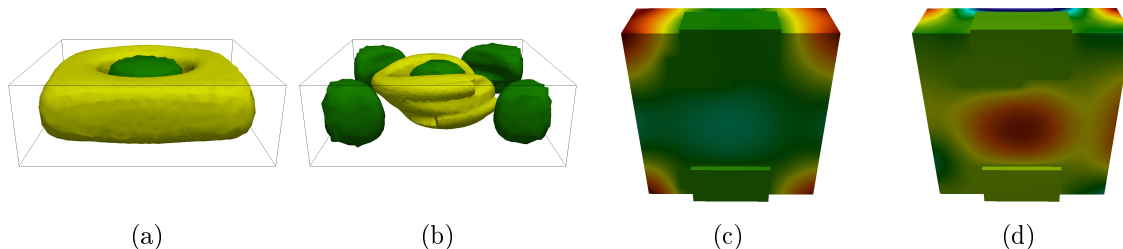


FIGURE 3.6.19 – Inductance : (a) & (b) exemples de modes de Dirichlet (vue du dessus), (c) & (d) modes de Steklov (vue du dessous).

Le maillage du PCB étant constitué de 36 000 nœuds de surface et de 146 000 nœuds, on ne calcule que 50% des modes de Steklov et 5% des modes de Dirichlet. Sur la figure 3.6.20,

on a représenté trois modes de Steklov du PCB, chacun étant très localisé aux différentes interfaces des composants, notamment sur les masques d'implémentation de l'inductance et du QFN16.

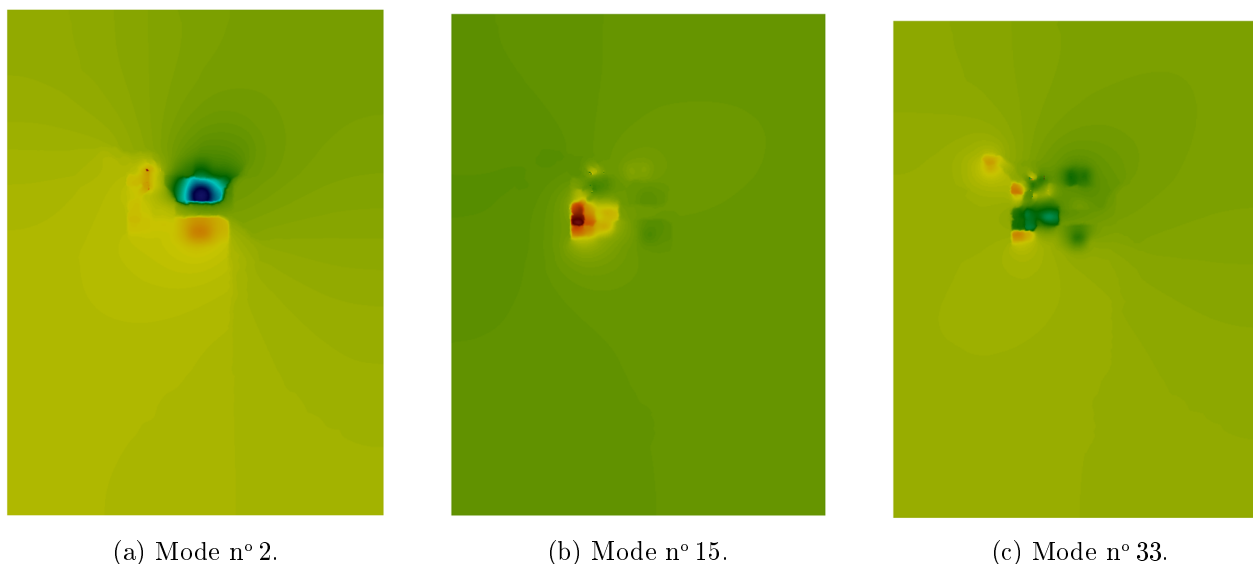


FIGURE 3.6.20 – PCB : Exemples de modes de Steklov (vue du dessus).

Pour ces trois composants, on a en quelque sorte déjà opéré une réduction, puisqu'on ne calcule qu'une partie des modes accessibles, afin de gagner en temps de calcul. Pour les autres composants, qui ne sont constitués que de quelques milliers de nœuds, les bases complètes ont été calculées.

Au total, 8 heures ont été nécessaires pour la détermination des bases modales des 19 sous-structures. Un temps non-négligeable qu'il faut prendre en compte pour juger de la "rentabilité" de la construction des modèles réduits.

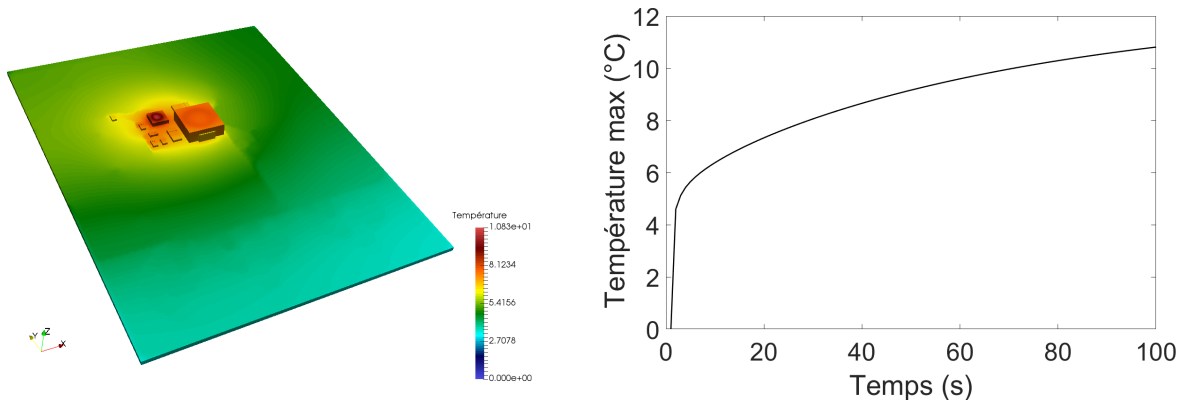
6.4.4 Réduction des bases

La technique d'amalgame modal, présentée en section 5.4.1, est utilisée pour réduire les bases des différentes sous-structures. Les états de référence, nécessaires pour la procédure, sont obtenus en projetant les modes sur un champ de température dit de référence. Ce champ est obtenu grâce à une simulation utilisant la méthode des éléments finis, où on considère un coefficient d'échange global de $10 \text{ W.m}^{-2}.\text{K}^{-1}$ sur toutes les surfaces externes, avec une température externe de 0°C et une puissance interne constante. Les puissances dissipées par les composants sont choisis selon les résultats de l'étude de [66] : $0,18 \text{ W}$ dans la partie

supérieure de la puce du boîtier QFN16, $0,09 W$ dans le solénoïde de cuivre de l'inductance, $0,01 W$ pour chaque capacité et résistance de dimension 0603 et $0,02 W$ pour les capacités de dimension 1206, portant le total de la puissance dissipée par les composants à $0,47 W$.

Le champ de température de référence est obtenu au terme d'une simulation transitoire de $100 s$ (figure 3.6.21a). On montre sur la figure 3.6.21b l'évolution de la température maximale au sein du QFN16. On observe que la température augmente d'abord très rapidement (jusqu'à $6^{\circ}C$) avant d'augmenter plus progressivement jusqu'à ce que le QFN16 atteigne $10,83^{\circ}C$ à $100 s$. On note qu'au terme des $100 s$, le régime permanent n'est pas atteint.

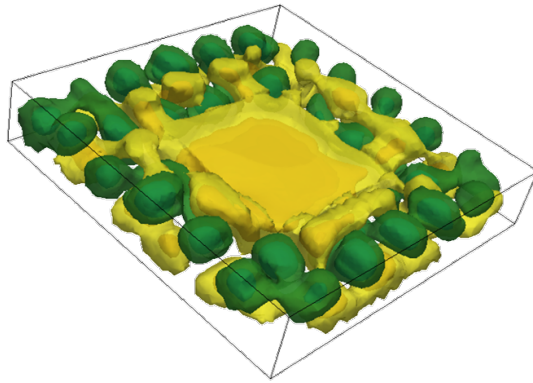
On crée ainsi un premier problème réduit de dimension $370 : 100$ modes réduits sont retenus pour le QFN16, 100 pour le PCB et 10 modes pour chacun des autres composants. Cette procédure de réduction (obtention d'un champ de référence et amalgame des modes) dure environ une heure, à ajouter au temps de calcul des bases initiales.



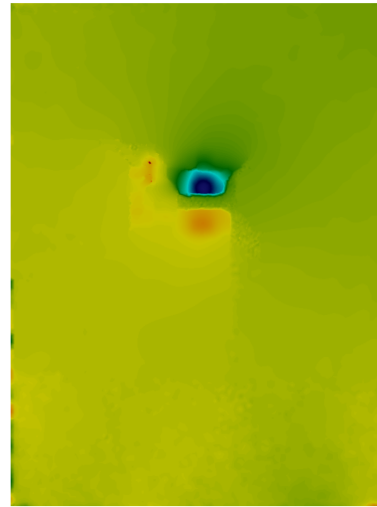
(a) Champ de température de référence pour l'amalgame. (b) Température maximale du QFN16 au cours du temps.

FIGURE 3.6.21 – Simulation de référence pour l'amalgame.

La figure 3.6.22 présente des exemples de modes amalgamés : le mode de Dirichlet du QFN16 (3.6.22a) présente ainsi beaucoup plus d'information que les modes de Dirichlet calculés, représentés sur la figure 3.6.18. Le mode de Steklov amalgamé représenté en 3.6.22b est très semblable au mode 3.6.20a ; le mode dominant porte ici très bien son nom, puisque les autres modes qui lui sont assignés sont très peu exprimés. On note toutefois des différences au niveau du PCB par rapport au mode non amalgamé, dues aux modes esclaves.



(a) Mode amalgamé de Dirichlet pour le QFN16.



(b) Mode amalgamé de Steklov n° 3 pour le PCB.



(c) Mode amalgamé de Steklov n° 95 pour le PCB.

FIGURE 3.6.22 – Exemples de modes amalgamés pour le QFN16 et le PCB.

6.5 Résultats

6.5.1 Scénario de test

Les bases réduites issues de la procédure d'amalgame sont donc construites pour reconstruire au mieux la simulation de référence. Pour tester la polyvalence des modes de Dirichlet-Steklov, un scénario avec des paramètres différents est simulé (voir figure 3.6.23) : cette fois-ci,

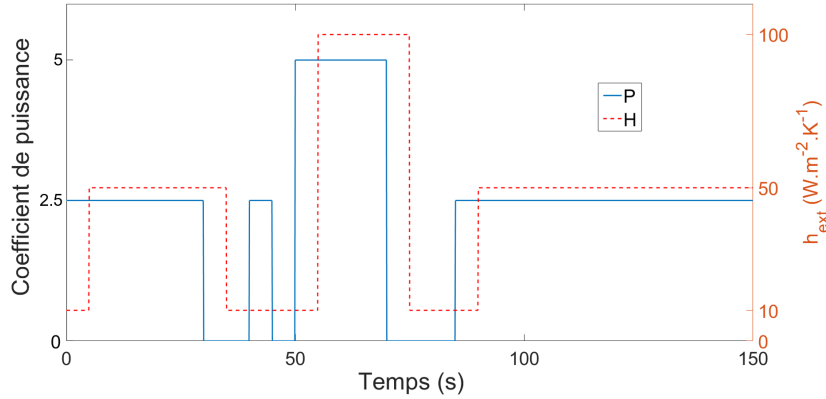


FIGURE 3.6.23 – Scénario de test.

la puissance varie tout au long d’une simulation de 150 s. Pour cela, un coefficient global est appliqué aux puissances dissipées par les composants, multipliant celles-ci par 0, 2,5 ou 5, simulant ainsi une série d’activations et de désactivations de la carte. La puissance dissipée est plus importante, entraînant une élévation plus haute de la température. On souhaite ainsi montrer l’indépendance de la base réduite vis-à-vis de ce paramètre.

Le coefficient d’échange global h_{ext} varie aussi, de concert avec la variation de puissance mais avec un décalage de 5 s : quand la puissance dissipée augmente, h_{ext} augmente aussi, passant de $10 W.m^{-2}.K^{-1}$ à 50 ou $100 W.m^{-2}.K^{-1}$, simulant ainsi l’action d’un ventilateur qui s’active automatiquement après la “mise en route” de la carte, avec un petit délai. Le détail du scénario est alors le suivant : la carte est activée pendant 30 s, au cours desquelles est dissipé un total de $2,5 \times 0,47 = 1,175 W$. Le ventilateur s’active 5 s après la mise en route : h_{ext} passe de $10 W.m^{-2}.K^{-1}$ (ordre de grandeur pour la convection naturelle) à $50 W.m^{-2}.K^{-1}$, puis se désactive 5 s après la mise hors-tension de la carte. Celle-ci est réactivée 5 s plus tard pendant seulement 5 s, mais cette fois h_{ext} reste à $10 W.m^{-2}.K^{-1}$. Encore 5 s plus tard, on remet en route la carte, mais cette fois-ci la puissance totale dissipée est de $2,35 W$. Le coefficient d’échange passe alors de 10 à $100 W.m^{-2}.K^{-1}$, avec toujours un délai de 5 s. La carte est désactivée après 20 s, puis réactivée 15 s plus tard mais ne dissipe cette fois plus que $1,175 W$ au total. Le coefficient d’échange global h_{ext} varie ainsi de 100 à 10 puis $50 W.m^{-2}.K^{-1}$ avec ce même délai de 5 s sur la variation de puissance. Ces paramètres sont ensuite maintenus jusqu’à la fin de la simulation.

On fait une première simulation de ce scénario avec une méthode d’éléments finis classique, qui va nous servir de point de comparaison pour évaluer l’écart de réduction. Le temps de calcul pour l’obtenir est de 915 s.

6.5.2 Résultats : comparaison avec une simulation détaillée

Le scénario de test est enfin simulé avec le modèle réduit. Le temps de calcul est cette fois de 2,8 s : le modèle modal réduit est 332 fois plus rapide que le modèle éléments finis. On répond donc au premier point du cahier des charges, à savoir être sobre en puissance de calcul. Il reste maintenant à voir si on a bien reconstitué les différentes conditions aux limites, *i.e.* si le modèle réduit ne s'éloigne pas trop du modèle complet.

Pour quantifier la qualité du modèle réduit, on introduit plusieurs quantités : soient T^{EF} le champ de température obtenu avec la méthode des éléments finis et T^{mod} le champ de température obtenu avec le modèle modal réduit, alors les écarts quadratiques $\langle \varepsilon \rangle_{L^2}$ et maximum ε_{∞} entre les deux champs sont définis par

$$\langle \varepsilon \rangle_{L^2} = \left[\frac{1}{V} \frac{1}{\tau} \int_0^{\tau} \int_{\Omega} (T^{EF} - T^{mod})^2 dt dv \right]^{\frac{1}{2}}, \quad \varepsilon_{\infty} = \max_{\tau} \max_{\Omega} |T^{EF} - T^{mod}|, \quad (3.6.8)$$

On définit de plus

$$\varepsilon_{\max}(t) = \max_{\Omega} T^{EF} - \max_{\Omega} T^{mod}, \quad \max(\varepsilon_{\max}) = \max_{\tau} (\varepsilon_{\max}(t)), \quad (3.6.9)$$

car la température maximale atteinte par la puce à l'intérieur du QFN16 est un paramètre critique : dans les procédés industriels, un écart maximum sur toute la durée de la simulation réduite $\max(\varepsilon_{\max}) \leq 1 K$ est requis.

Sur la figure 3.6.24, on a représenté l'évolution de la température maximale $\max_{\Omega} T^{mod}$ atteinte par le QFN16 (car c'est le composant qui chauffe le plus) au cours du temps (ligne pleine bleue); les élévations brutales de températures correspondent ainsi à l'activation du composant, ensuite limitées par l'activation du "ventilateur" (on le voit très bien entre 55 s et 70 s).

On a aussi représenté sur cette figure $\varepsilon_{\max}(t)$ pour le même composant, en ligne mauve pointillée : les pics correspondent ainsi aux différentes variations de paramètre (puissance ou coefficient d'échange), mais on constate que cet écart critique ne dépasse jamais 1 K, et même jamais 1% de la plage de température du QFN16.

Puisque la méthode modale permet de récupérer l'ensemble du champ de température, on va pouvoir localiser graphiquement cet écart. La figure 3.6.25 représente le champ de température de la carte à $t = 51 s$ (à gauche) et le champ écart $|T^{mod} - T^{FE}|$ (à droite). Le champ de température est bien représenté dans sa complexité, puisqu'on peut observer la diffusion de la puissance dissipée par la puce dans le QFN16 et qui suit les pistes du PCB, avant de chauffer les capacités et résistances. On constate alors que l'écart le plus grand est localisé dans les coins du PCB, le nœud correspondant à $\varepsilon = 3,61 K$ se trouvant dans le coin

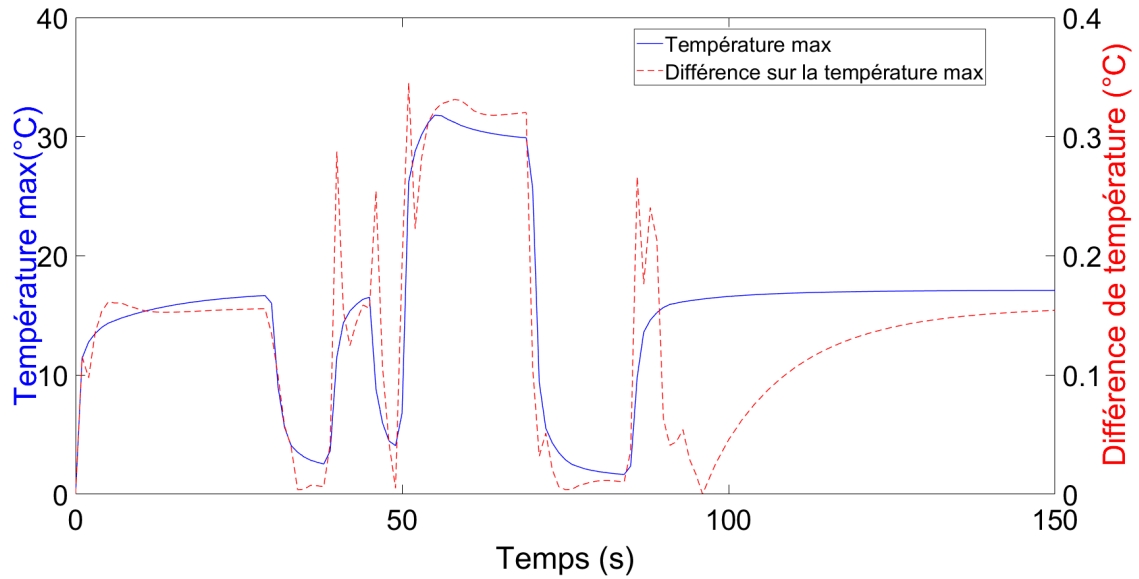


FIGURE 3.6.24 – Évolution de $\max_{\Omega} T^{mod}(t)$ et $\varepsilon_{\max}(t)$ pour le QFN16 au cours du temps.

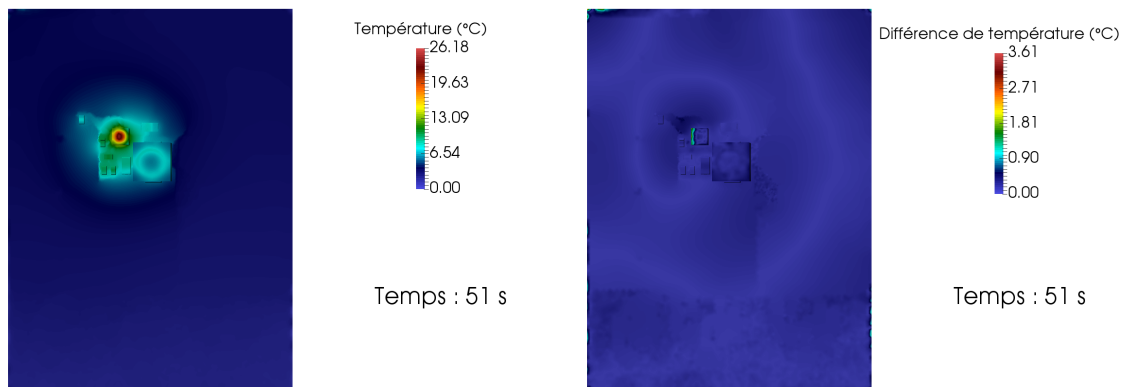


FIGURE 3.6.25 – T^{mod} (gauche) et $|T^{mod} - T^{EF}|$ (droite) à $t = 51s$, quand ε_{∞} est maximal.

inférieur gauche, qui n'est pas du tout un endroit critique, et ce qui n'a que peu d'incidence sur les composants. On constate aussi un écart au sein du QFN16 : celui-ci correspond au pic observé sur le graphe en figure 3.6.24. On va donc s'intéresser à cette partie de la carte plus en détail (figure 3.6.26) : alors qu'on a modélisé le QFN16 avec seulement 100 degrés de libertés (les modes réduits), on arrive tout de même à voir les détails de l'intérieur du boîtier, comme les fils de connexion. L'écart avec la simulation éléments finis est ainsi localisé dans la résine, où la température est environ inférieure de dix degrés à la température de la puce, ce qui montre que le modèle est très fiable au point critique. Cet écart est dû à l'activation soudaine

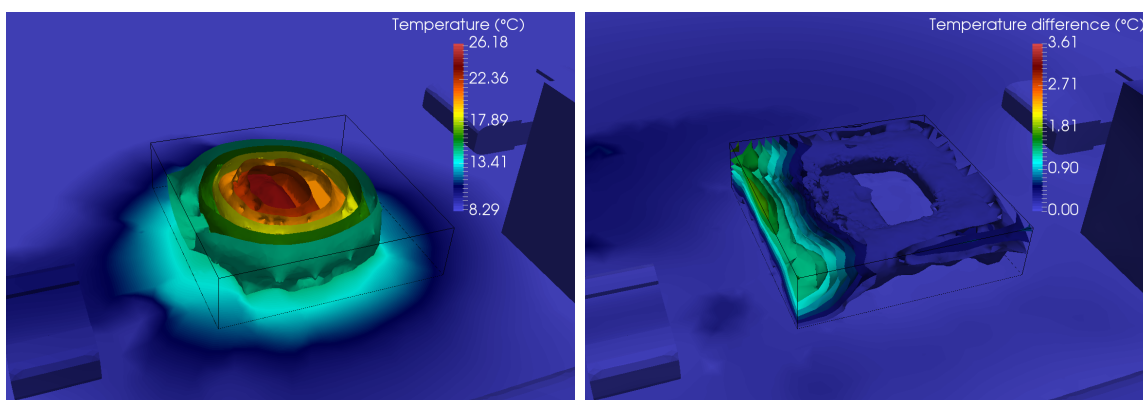


FIGURE 3.6.26 – T^{mod} (gauche) et $|T^{mod} - T^{EF}|$ (droite) à $t = 51s$, quand ϵ_∞ est maximal.

du composant, est resté très éphémère, comme le montre la figure 3.6.27, où on représente les mêmes champs mais une seconde plus tard ; l'écart avec la simulation en éléments finis s'est presque totalement estompé.

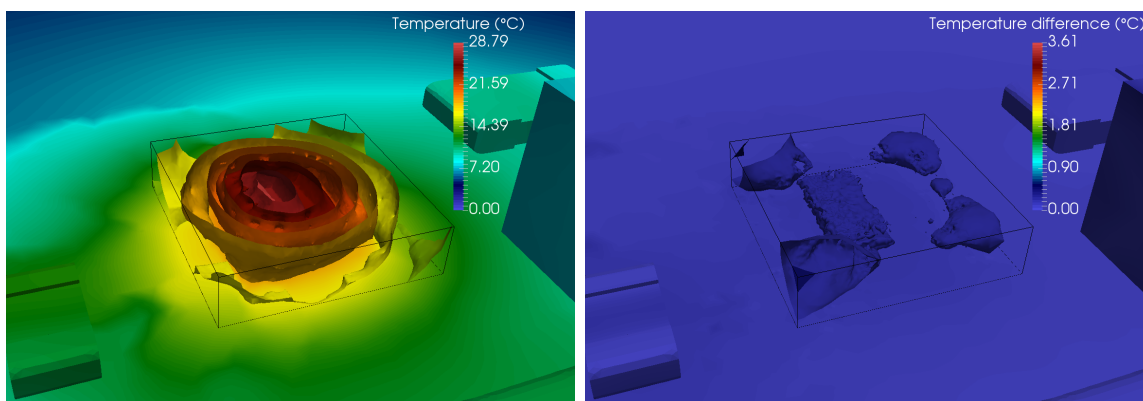


FIGURE 3.6.27 – T^{mod} (gauche) et $|T^{mod} - T^{EF}|$ (droite) à $t = 52s$.

Puisque l'algorithme d'amalgame utilisé ne permet que d'avoir le nombre de modes réduits voulu pour chaque sous-structure, on a dû attendre de faire la simulation réduite pour savoir si on respectait le critère $\max(\epsilon_{\max}) \leq 1K$. On peut se demander si ce critère est respecté avec un modèle réduit encore plus petit. On conduit alors une étude de sensibilité pour étudier l'influence de l'ordre de réduction sur la précision du modèle réduit en variant conjointement le nombre de modes réduits retenus pour le PCB et le QFN16, de 20 pour chacune des deux sous-structures à 1000 (on garde toutefois 10 modes pour les autres composants, puisqu'aucun écart notable n'y a été détecté). Les résultats de cette étude sont présentés dans la table 3.6.4. On note qu'avec seulement 30 modes pour le PCB et le QFN16, on réalise la condition

$\max(\epsilon_{\max}) \leq 1 K$. Le modèle réduit correspondant est alors de dimension totale 230, soit une réduction d'un facteur proche de 2000 par rapport au 422 000 degrés de libertés du modèle complet. Le modèle réduit est résolu en 1.6 s, c'est-à-dire en 300 fois moins de temps.

On remarque toutefois qu'à mesure que le nombre de modes augmente, $\max(\epsilon_{\max})$ ne décroît pas de manière monotone. En effet, $\max(\epsilon_{\max})$ représente un écart en un point du maillage à un point dans le temps, alors que l'amalgame est conçu pour minimiser un écart quadratique global, qui lui décroît effectivement, comme le montre la colonne $\langle \epsilon \rangle_{L^2}$.

Enfin, on remarque que ϵ_{∞} peut atteindre jusqu'à $k K$, mais cet écart est très localisé en temps et en espace, comme on on vient de le voir avec les figures 3.6.26 et 3.6.27.

TABLE 3.6.4 – Différences sur le QFN16 et le PCB, et temps de calcul en fonction de l'ordre de réduction

Nombre de modes	Structure	$\langle \epsilon \rangle_{L^2}$	ϵ_{∞}	$\max(\epsilon_{\max})$	Temps de calcul (s)	Gain
20	PCB	0,122	5,405	1,076	1,3	715
20	QFN16	0,247	2,915	1,310		
30	PCB	0,097	2,480	0,479	1,6	585
30	QFN16	0,149	2,395	0,999		
100	PCB	0,045	3,610	0,213	2,8	332
100	QFN16	0,119	1,879	0,345		
200	PCB	0,042	2,756	0,220	4,8	190
200	QFN16	0,084	1,000	0,417		
300	PCB	0,040	2,021	0,219	7,2	128
300	QFN16	0,077	0,723	0,535		
500	PCB	0,037	2,564	0,214	13,7	67
500	QFN16	0,066	0,734	0,444		
1000	PCB	0,036	0,643	0,215	39,8	23
1000	QFN16	0,062	0,730	0,339		

6.6 Bilan de cette première étude

L'objectif de cette étude était de réduire un modèle réaliste de carte électronique, pour éprouver la réduction d'un modèle représentatif des propriétés physiques d'une telle carte. Les composants, et notamment le QFN16, ont ainsi été modélisés avec un très grand niveau de détails (jusqu'au fils de connexion entre la puce et les pattes du boîtier), et la modélisation du PCB en utilisant une approche "Metal Fraction" permet d'avoir une représentation précise

des propriétés physiques des différentes couches du circuit imprimé. Avec une réduction du modèle d'un facteur 2000, on arrive à maintenir un écart inférieur à 1% sur le paramètre critique entre modèle réduit et modèle détaillé, tout au long de la simulation de test. La capacité de la base de Dirichlet-Steklov à reconstruire des conditions aux limites variables a ainsi été mise en évidence.

Cette réduction permet ainsi de réduire le temps de résolution d'un facteur 300 : le coût initial de la réduction (les 8 heures de temps de calcul des bases plus l'heure requise pour les amalgamer par rapport à une simulation de référence) peut donc être rapidement rentabilisé en faisant des études paramétriques de la carte, ce qui représente des dizaines de simulations.

Bien sûr, le temps de réduction pourrait être encore plus rentabilisé si les composants réduits pouvaient être réutilisés pour d'autres configurations. Ce n'est malheureusement pas possible dans le cas présent, puisqu'on s'est limité à des maillages conformes aux interfaces. C'est pourquoi, dans le chapitre suivant, on va présenter un exemple de sous-structuration modale utilisant la méthode des joints.

Il faut toutefois noter que la sous-structuration du maillage conforme présente quand même un intérêt : le calcul des modes des différentes sous-structures est beaucoup plus rapide que si on avait dû calculer les modes de l'ensemble de la carte d'un coup.

Chapitre 7

Exemple d'application de la sous-structuration modale non-conforme avec joints

Pour vérifier le bien fondé de la méthode modale non conforme avec joints, nous avons commencé par traiter un modèle bidimensionnel simple. Cette étude a fait l'objet d'une communication au Congrès de la Société Française de Thermique 2018 à Pau. On trouvera le texte de cette communication en annexe.

Ce travail a montré de très bons résultats en terme d'écart de réduction et confirme la pertinence de notre approche. Cependant, la géométrie reste assez éloignée du cas tridimensionnel, qui reste notre objectif. Nous proposons de mettre en œuvre la méthode de sous-structuration modale avec joints à une carte électronique simplifiée tridimensionnelle.

Cependant, par rapport à la méthode conforme vue au chapitre précédent, deux nouveaux ingrédients rentrent dans la formulation du modèle : le choix et le nombre de fonctions mortiers utilisées sur les interfaces de contact. Nous allons donc commencer par analyser ce problème sur une géométrie élémentaire afin de bien dégager les paramètres essentiels, avec dans un premier temps une étude de l'influence de la méthode des joints dans une formulation éléments finis classique et dans un second temps l'influence sur la formulation modale. C'est seulement ensuite que sera traité le cas de la carte simplifiée.

7.1 Validation du modèle éléments finis sous-structuré avec joints

La première étape, avant de considérer toute réduction modale, est de s'assurer que le modèle détaillé utilisé comme référence est validé. On compare dans cette section les résultats obtenus avec une simulation sous-structurée selon la méthode des joints utilisant des éléments finis (2.4.26) aux résultats obtenus via une simulation non sous-structurée classique (on parlera de modèle "monobloc"). Cette simulation monobloc est obtenue à l'aide du même code de calcul SAMBA, elle même validée avec des logiciels commerciaux.

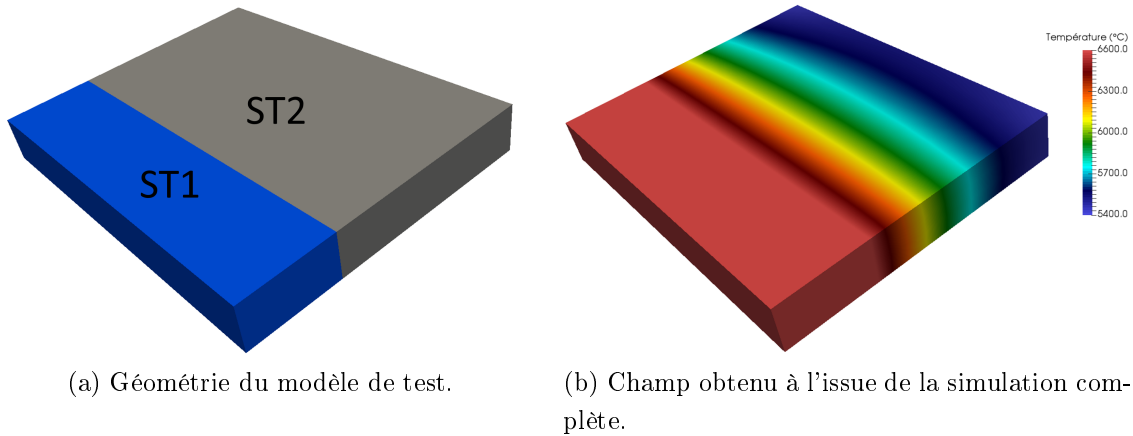


FIGURE 3.7.1 – Description de la géométrie de test et champ de température obtenue à l'issue d'une simulation dynamique de 1000 s.

Pour cette étude, on reprend l'exemple sous-structuré utilisé en section 6.4.2 pour illustrer l'importance du paramètre de Steklov, à savoir deux pavés droits de dimensions respectives $4 \times 1 \times 0,5 \text{ mm}^3$ et $4 \times 2,2 \times 0,5 \text{ mm}^3$. On rappelle que le pavé le plus petit (en bleu sur 3.7.1a) dissipe une puissance thermique de 10^9 W.m^{-3} et a comme propriétés physiques $\kappa = 260 \text{ W.m}^{-1}.\text{K}^{-1}$ et $C = 3,43 \times 10^6 \text{ J.m}^{-3}.\text{K}^{-1}$, l'autre pavé étant caractérisé par $\kappa = 0,67 \text{ W.m}^{-1}.\text{K}^{-1}$ et $C = 1,75 \times 10^6 \text{ J.m}^{-3}.\text{K}^{-1}$, les deux solides échangeant de la chaleur avec le milieu extérieur à une température de 0°C via un coefficient d'échange $h = 10 \text{ W.m}^{-2}.\text{K}^{-1}$. Le petit pavé, en bleu, est choisi comme étant la sous-structure 1 et le gros pavé, en gris, étant la sous-structure 2. On suppose qu'on a un contact parfait entre les deux sous-structures (*i.e.* $R_c = 0 \text{ m}^2.\text{K.W}^{-1}$). On représente sur la figure 3.7.1b le champ de température obtenu à l'issue des 1000 s de la simulation dynamique monobloc. Le régime stationnaire est atteint, après une élévation de 6600°C .

7.1.1 Choix des fonctions mortiers

En reprenant les notations du chapitre 4, le couplage entre les sous-structures s'effectue à travers les matrices $\mathbf{J}_p^{(0,k)}$ et $\mathbf{J}_p^{(k,k)}$ définies en (2.4.27). Pour rappel, ces matrices sont définies par

$$\mathbf{J}_p^{(0,k)} = s_{g_p}^{k,0} \left(\int_{\Gamma_p^{\{k,0\}}} \psi_i^{(0)} \phi_{p_m}^{\{k,0\}} \right)_{\substack{1 \leq i \leq N^{(0)} \\ 1 \leq m \leq N_\phi^{\Gamma_p^{\{k,0\}}}}} , \quad \mathbf{J}_p^{(k,k)} = s_{g_p}^{k,0} \left(\int_{\Gamma_p^{\{k,0\}}} \psi_i^{(k)} \phi_{p_m}^{\{k,0\}} \right)_{\substack{1 \leq i \leq N^{(k)} \\ 1 \leq m \leq N_\phi^{\Gamma_p^{\{k,0\}}}}} .$$

Rappelons que de manière complètement contre intuitive, les fonctions mortiers $\phi_{p_m}^{\{k,0\}}$ sont définies sur le maillage du coté non-mortier. Dans la littérature, ces fonctions ont des formes bien particulières. En deux dimensions, on a donné leur allure dans la section 2.5.3. D'après [46], si on a n points sur la frontière non mortier, on a $n - 2$ fonctions mortiers, qui peuvent être les fonctions définies en 2.5.3 (page 30), et qu'on a choisi d'utiliser dans l'exemple en 2D. En trois dimensions, leur définition est plus délicate [28]. Les raisons poussant à ne pas prendre les fonctions de forme éléments finis à l'interface, mais des fonctions formant un ensemble de co-dimension 2 ne sont pas très explicites. Il semble que prendre l'intégralité des fonctions de forme à l'interface pose un problème lorsqu'un point est en commun entre trois sous-structures. Dans le cas de carte électronique, ce cas de figure ne devrait pas se produire. Nous avons alors choisi d'utiliser toutes les fonctions de forme du maillage de la frontière non mortier : $\phi_{p_m}^{\{k,0\}} = \psi_i^{(k)}$, $1 \leq i \leq N^{(k)}$

Le calcul de la matrice $\mathbf{J}_p^{(k,k)}$ est alors immédiat. En revanche, ce n'est pas le cas pour $\mathbf{J}_p^{(0,k)}$ puisque les fonctions $\psi_i^{(0)}$ et $\phi_{p_m}^{\{k,0\}}$ ne sont pas définies sur le même maillage. Un maillage fictif intermédiaire constitué de quadrangles est créé entre les deux surfaces en contact. Les fonctions $\psi_i^{(0)}$ et $\phi_{p_m}^{\{k,0\}}$ sont projetées sur ce maillage ; l'intégrale $\int_{\Gamma_p^{\{k,0\}}} \psi_i^{(0)} \phi_{p_m}^{\{k,0\}}$ est ensuite calculée en utilisant une formule des trapèzes. On peut alors assembler le système (2.4.26) et le résoudre numériquement. Puisque nous n'avons pas choisi les fonctions mortiers "réglementaires", il convient de valider cette entorse à l'orthodoxie.

7.1.2 Avec un maillage conforme à l'interface

Dans un premier temps, on compare les deux modèles (monobloc et sous-structuré avec joints) à partir d'un même maillage, nécessairement conforme à l'interface pour effectuer la simulation monobloc (voir figure 3.7.2a). Ce maillage est constitué de 738 nœuds sur la sous-structure 1 et 1 452 sur la sous-structure 2. Il y a 84 nœuds sur l'interface de contact. La surface choisie comme étant la surface non-mortier (et où seront calculées les fonctions mortiers) est la surface de contact du côté de la sous-structure 2.

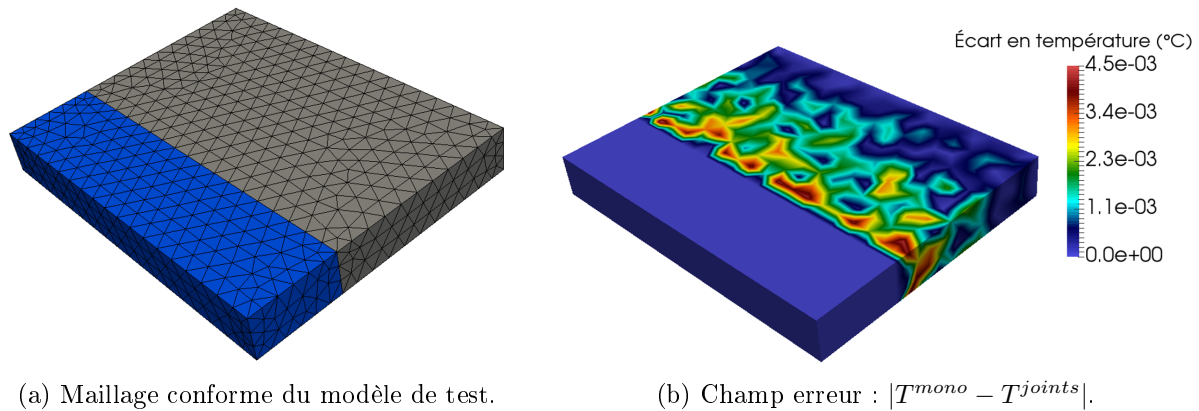


FIGURE 3.7.2 – Maillage conforme de la géométrie de test et champ écart obtenu à l'issue d'une simulation dynamique de 1000 s.

La figure 3.7.2b présente le champ de l'écart de température entre les deux modèles lorsque le régime permanent est atteint. Certes, l'écart est supérieur au zéro machine, mais les $4,5 \times 10^{-3} K$ d'écart maximum doivent être comparés aux $6600^\circ C$ de la plage de température, soit un écart relatif inférieur à 10^{-6} . L'écart moyen est lui de l'ordre de $10^{-6} K$, soit un écart relatif de l'ordre de 10^{-10} c'est-à-dire le critère d'arrêt du solveur matriciel. Le modèle mortier est donc validé lorsque le maillage est conforme à l'interface.

7.1.3 Avec un maillage non-conforme à l'interface

Évidemment, l'objectif de la méthode des joints est de pouvoir sous-structurer des maillages non-conformes aux interfaces. On applique alors cette méthode sur un maillage constitué de 2 408 nœuds pour la sous-structure 1 (le petit pavé, en bleu) et 1 452 nœuds pour la sous-structure 2 (le gros pavé, en gris) (voir la figure 3.7.4a)). Cependant, une comparaison entre un modèle non sous-structuré conforme aux interfaces et un modèle sous-structuré non-conforme est plus délicate, puisque par construction les maillages ne coïncident pas. Un calcul d'erreur point à point est impossible. La validation se fait alors en deux temps. Dans un premier temps, la température maximale atteinte au cours du temps par les deux modèles est comparée sur la figure 3.7.3. La comparaison est concluante, puisque l'écart ne dépasse jamais $3 \times 10^{-3} K$, soit là encore un écart relatif inférieur à 10^{-6} . Nous représentons sur la figure 3.7.4b un montage des isovaleurs des champs de température obtenus par les deux modèles. La partie supérieure (maillage noir) est calculée sans sous-structuration, alors que la partie inférieure (maillage rouge) est le résultat d'une simulation avec la méthode des joints. Les iso-valeurs se superposent parfaitement.

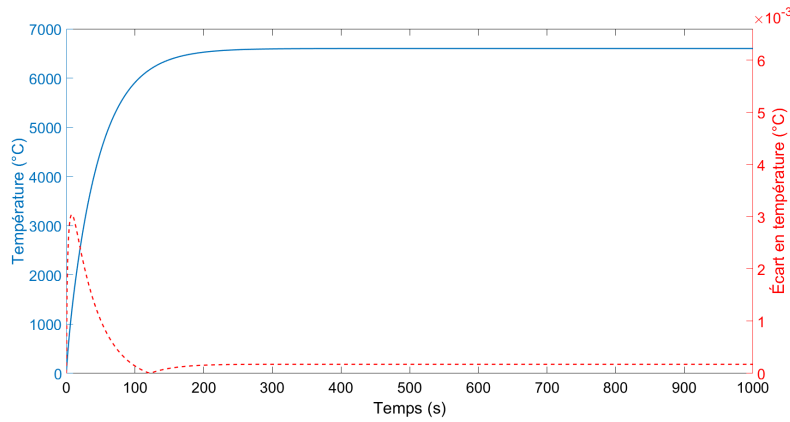
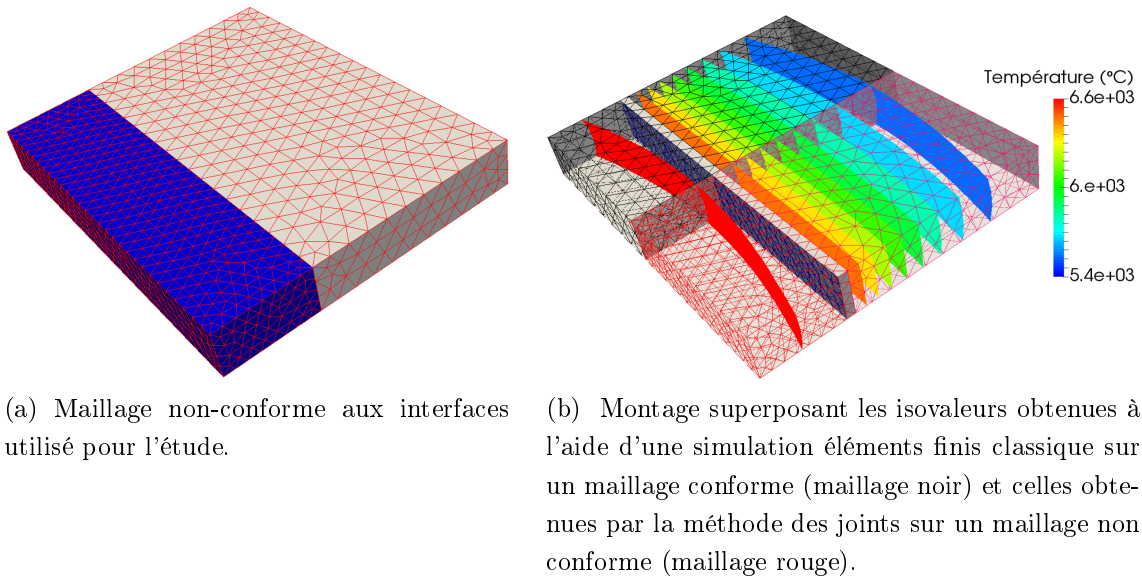


FIGURE 3.7.3 – Évolution de la température maximale dans la sous-structure 2. En rouge (échelle de droite) l'écart sur cette quantité entre le modèle éléments finis classique sur un maillage conforme et la méthode des joints sur un maillage non-conforme aux interfaces.



(a) Maillage non-conforme aux interfaces utilisé pour l'étude.

(b) Montage superposant les isovaleurs obtenues à l'aide d'une simulation éléments finis classique sur un maillage conforme (maillage noir) et celles obtenues par la méthode des joints sur un maillage non conforme (maillage rouge).

FIGURE 3.7.4 – Simulation éléments finis avec la méthode des joints.

7.2 Réduction modale

7.2.1 Réduction modale des fonctions mortiers

Les fonctions mortiers définies dans la littérature étant liées aux fonctions de forme éléments finis, elles ne sont pas réductibles. Il est dès lors difficile de concevoir l'utilité d'une réduction de modèle avec quelques dizaines de modes par sous-structure, couplés par quelques

centaines voire milliers de fonctions mortiers. Tout l'intérêt de la réduction de modèle serait perdu. Puisque les fonctions utilisées doivent former une base de l'espace mortier défini sur $\Gamma_p^{\{k,0\}}$, on peut envisager d'utiliser des modes pour réduire ces bases, selon la même méthodologie utilisée pour les sous-structures. On construit alors une base de fonctions propres de l'opérateur Laplacien sur $\Gamma_p^{\{k,0\}}$, avec des conditions limites sur $\partial\Gamma_p^{\{k,0\}}$ à préciser :

- Une condition de Dirichlet ne convient pas puisque les fonctions propres formeront une base de $H_0^1(\Gamma_p^{\{k,0\}})$ mais il n'y a pas de raison que le flux s'annule sur la frontière $\partial\Gamma_p^{\{k,0\}}$;
- Une condition de Neumann donnera une dérivée normale du flux nulle sur le bord $\nabla\varphi \cdot \underline{n}_{\partial\Gamma_p^{\{k,0\}}} = 0$, alors que le flux φ lui-même ne s'annulera pas. Les modes de Neumann $\{\phi_i^N\}_{i \in \mathbb{N}}$ forment une base d'un sous espace de $H^1(\Gamma_p^{\{k,0\}})$ et le premier mode propre est constant, ce qui présente un avantage puisqu'il permet de reconstituer à lui seul un flux uniforme. Or, compte tenu de la petite dimension des interfaces et de la grande quantité de chaleur dégagée par les composants via ces interfaces, il n'est pas irraisonné de considérer que le flux est uniforme. Si on note λ_i^N la valeur propre d'ordre i associée au mode ϕ_i^N , le problème aux valeurs propres de Neumann local s'écrit :

$$\begin{aligned} \Gamma_p^{\{k,0\}} \quad \Delta\phi^N &= \lambda^N \phi^N, \\ \partial\Gamma_p^{\{k,0\}} \quad \nabla\phi^N \cdot \underline{n}_{\partial\Gamma_p^{\{k,0\}}} &= 0. \end{aligned} \quad (3.7.1)$$

- Une condition limite de Steklov sur $\partial\Gamma_p^{\{k,0\}}$ est aussi envisageable, et on obtiendrait alors une base de Branches. Le problème de Branches local s'écrit :

$$\begin{aligned} \Gamma_p^{\{k,0\}} \quad \Delta\phi^B &= \lambda^B \phi^B, \\ \partial\Gamma_p^{\{k,0\}} \quad \nabla\phi^B \cdot \underline{n}_{\partial\Gamma_p^{\{k,0\}}} &= \lambda^B \phi^B. \end{aligned} \quad (3.7.2)$$

Les modes de Branches ainsi définis forment des bases hilbertiennes de $H_0^1(\Gamma_p^{\{k,0\}}) \times H^{1/2}(\partial\Gamma_p^{\{k,0\}})$ et $L^2(\Gamma_p^{\{k,0\}}) \times L^2(\partial\Gamma_p^{\{k,0\}})$.

La piste des modes de Branches est très intéressante, d'autant plus que le problème local de Branches est posé en 2D, et les travaux précédents de l'équipe ThE ont montré que le problème est d'autant plus régulier que la dimension est faible, et donne d'excellents résultats dans cette dimension. Cependant, pour être pleinement efficace, la réduction de la base de branche demande une certaine expertise. Pour des raisons de simplicité et de rapidité de mise en œuvre, les modes de Neumann ont été donc été choisis. Des exemples de ces modes calculés sur l'interface entre les deux sous-structures sont représentés sur la figure 3.7.5.

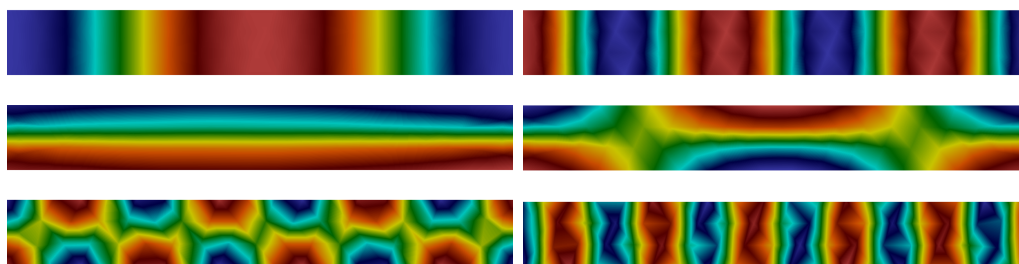


FIGURE 3.7.5 – Exemples de modes de Neumann pour l'interface de contact.

L'influence sur la précision du nombre de modes de Neumann conservés pour raccorder les deux sous-structures est résumée sur la figure 3.7.6. Ici, la température n'a pas été réduite dans les sous-structures : les deux modèles éléments finis sont couplés par des modes de Neumann calculés à l'interface. Ainsi, seule l'influence de cette réduction est caractérisée. Les résultats sont mitigés. Notons tout d'abord que la méthode converge : l'écart relatif est inférieur à 10^{-10} lorsque la base complète est conservée. Au titre des déceptions (relatives), on aurait souhaité une convergence plus régulière. En effet, l'écart ne varie pas lorsque l'on passe de 10 modes gardés à 30 modes, ou encore entre 50 et 75 modes. De plus, l'ajout du seul dernier mode fait chuter l'erreur de cinq ordres de grandeur. Malgré cela, on note tout de même que l'écart engendré par la réduction du flux à l'interface est très faible. Avec un seul mode gardé (donc une hypothèse de flux homogène), l'écart est de $5^{\circ}C$, soit un écart relatif inférieur à 0.1%. Avec 6 modes, il diminue encore d'un ordre de grandeur. Par ailleurs, il faut garder à l'esprit que l'objectif est de réduire aussi le champ de température, ce qui engendre également un écart.

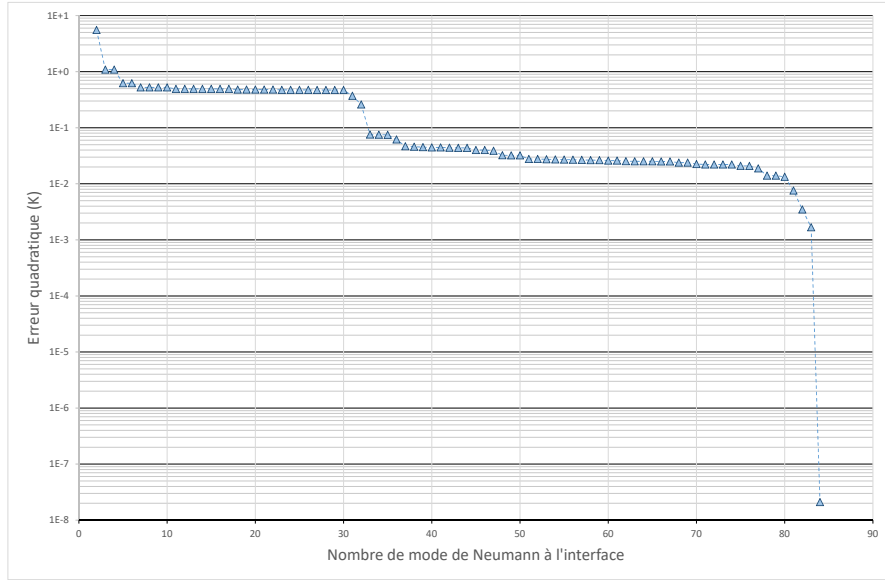


FIGURE 3.7.6 – Écart quadratique sur la température entre d’une part le modèle éléments finis sous-structuré avec les fonctions de formes éléments finis pour les fonctions mortiers et d’autre part le modèle éléments finis avec des modes de Neumann pour les fonctions mortiers. Le nombre de modes de Neumann conservé est en abscisse.

7.2.2 Réduction modale en flux et en température

On étudie dans cette section, sur un cas simple, l’interaction entre les réductions modales en température et la réduction modale en flux.

Les bases de Dirichlet-Steklov complètes des sous-structures 1 et 2 sont calculées dans un premier temps avec une valeur de ζ de 1. Elles sont ensuite réduites par la méthode d’amalgame modal avec la simulation détaillée comme référence.

On crée ainsi quatre modèles réduits, correspondant à autant de bases réduites : un modèle avec 10 modes amalgamés pour chaque sous-structure, un deuxième avec 20 modes amalgamés par sous-structure, un troisième avec 50 modes et enfin un quatrième modèle constitué de 100 modes amalgamés par sous-structure. On calcule la base de modes de Neumann sur la frontière de contact de la sous-structure 2. Il y a 84 nœuds sur cette frontière, on a donc calculé 84 modes de Neumann.

On rappelle qu’on peut reconstituer la température selon

$$\tilde{T}_{proj} = \sum_{\mathcal{X} \in \{D, S\}} \sum_{I=1}^{\tilde{N}^{\mathcal{X}}} \tilde{x}_I^{\mathcal{X}} \tilde{V}_I^{\mathcal{X}},$$

en calculant les états réduits à partir du champ de température de référence T selon

$(T|\tilde{V}_j^{\mathcal{X}})_H = \tilde{x}_j^{\mathcal{X}}$, ce qui va nous permettre d'estimer la "qualité" de la base réduite à partir de l'écart entre \tilde{T}_{proj} et T .

Les écarts entre champ de référence et champ projeté $|T - \tilde{T}_{proj}|$ pour chaque sous-structure sont présentés dans la table 3.7.1.

Nombre de modes	$ T^{(1)} - \tilde{T}_{proj}^{(1)} $	$ T^{(2)} - \tilde{T}_{proj}^{(2)} $
10	1,1E-04	0,08
20	3,04E-05	0,02
50	5,92E-06	0,008
100	2,37E-06	0,003

TABLE 3.7.1 – Écarts entre champ projeté et champ de référence en fonction du nombre de modes des bases réduites pour les sous-structures 1 et 2

Ces écarts représentent donc l'objectif à atteindre en terme d'écart entre la simulation réduite et la simulation de référence.

Pour étudier l'influence de la réduction des fonctions mortiers sur l'écart global de réduction, une étude de sensibilité est conduite pour chaque modèle réduit. Le champ obtenu à l'issu de la simulation du modèle réduit (c'est à dire en résolvant l'équation d'état (2.5.41)) sera noté \tilde{T} .

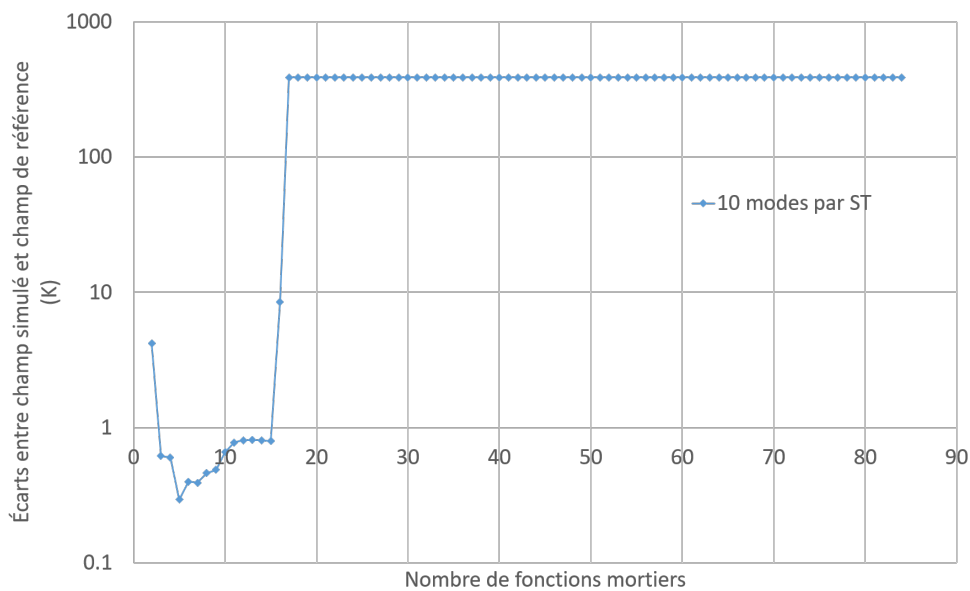


FIGURE 3.7.7 – Écart moyen entre le champ issu de la simulation réduite avec 10 modes par sous-structure et champ de référence en fonction du nombre de mortiers utilisés à l'interface.

La figure 3.7.7 présente l'écart moyen entre la simulation réduite et la simulation de référence en fonction du nombre de modes de Neumann choisis à l'interface, de 2 à l'ensemble des 84 modes de la base, pour le modèle réduit avec 10 modes par sous-structure. Pour 2 modes de Neumann à l'interface, l'écart moyen est aux alentours de $4 K$, ce qui reste très acceptable dans le cadre d'une simulation dont la plage de température est de $6600 K$, mais qui est néanmoins très loin de l'écart de projection. En comparant avec la courbe 3.7.6, il est clair que c'est la réduction du flux qui bride la qualité de la réduction. L'écart diminue ensuite à mesure que l'on ajoute des modes de Neumann. Cependant, au delà de 5 modes, l'écart augmente. Cette remontée est d'abord très progressive, pour se stabiliser à environ $0,8 K$ d'écart, avant de complètement exploser à partir de 16 modes de Neumann, pour se maintenir à un écart de $387 K$. Cette divergence s'explique. Sous forme condensée, le système à résoudre lorsque la résistance de contact est nulle est de la forme :

$$\begin{bmatrix} \mathbf{A} & \mathbf{J} \\ {}^t\mathbf{J} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \Phi \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ 0 \end{bmatrix},$$

où \mathbf{A} est une matrice regroupant les matrices réduite $\tilde{\mathbf{K}}_{xy}$ et $\tilde{\mathbf{H}}$ de l'équation (2.5.41). Pour déterminer la solution, il faut alors résoudre le problème suivant :

$${}^t\mathbf{J}\mathbf{A}^{-1}\mathbf{J}\Phi = {}^t\mathbf{J}\mathbf{A}^{-1}\mathbf{f}.$$

Or au delà d'un certain nombre de modes gardés pour le flux, le rang de la matrice ${}^t\mathbf{J}\mathbf{A}^{-1}\mathbf{J}$ devient strictement inférieur à sa dimension : La matrice n'est donc plus inversible. Il est dès lors compréhensible que le résultat soit incohérent.

Le nombre de mortiers utilisés sur l'interface apparaît donc clairement comme un paramètre critique qu'il est très important de prendre en compte, avec un optimum pour avoir l'écart le plus faible possible, et une valeur limite à partir de laquelle la solution diverge. Pour ce cas précis, l'optimum est à 5 fonctions mortiers, ce qui semble indiquer qu'il faut prendre moins de mortiers que de modes de Dirichlet-Steklov amalgamés.

Pour s'en assurer, on trace sur la figure 3.7.8 la même courbe pour chaque modèle réduit. Sur ce graphe, la courbe bleue représente l'évolution de l'écart pour le modèle réduit avec 10 modes amalgamés par sous-structure, la courbe orange pour le modèle avec 20 modes par sous-structure, la courbe grise pour le modèles à 50 modes et la courbe jaune pour le modèle à 100 modes. Si on s'intéresse d'abord à la courbe jaune, on observe le même comportement que la courbe bleue, mais "décalé" vers la droite du graphique, avec un optimum à 11 mortiers et une divergence à partir de 35 mortiers. Pour le modèle avec 50 modes réduits par sous-structure, on observe aussi le même comportement, l'optimum étant atteint à 33 mortiers. L'écart augmente ensuite progressivement à mesure qu'on ajoute des modes de Neumann pour

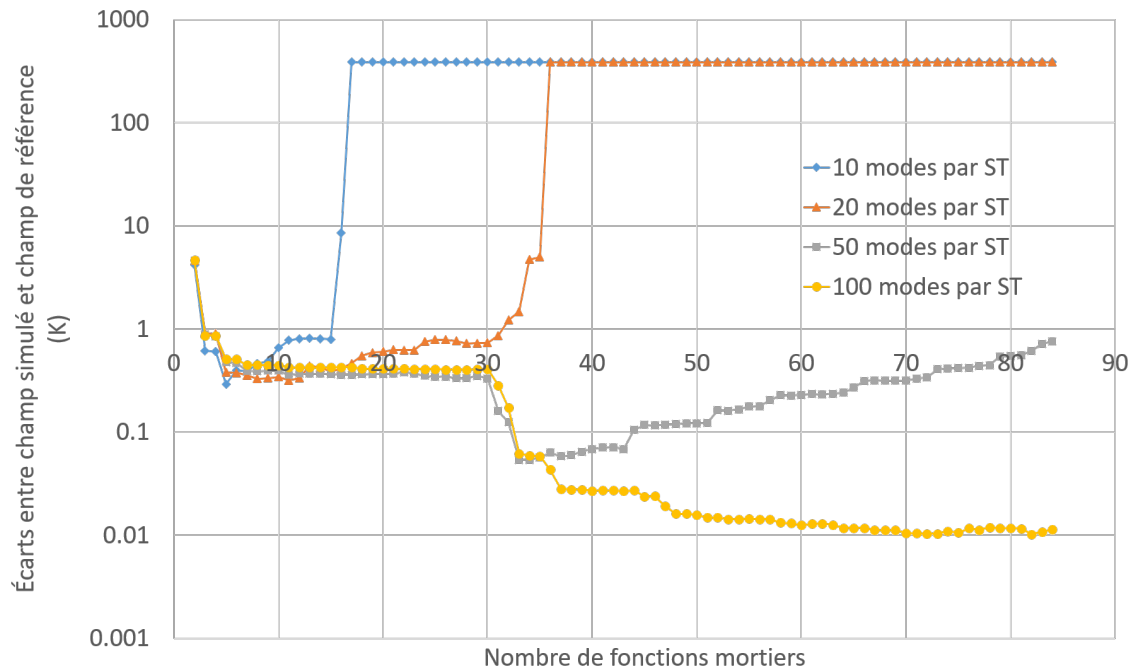


FIGURE 3.7.8 – Écart moyen entre champ issu de la simulation réduite avec 10, 20, 50 et 100 modes par sous-structures et champ de référence en fonction du nombre de mortiers utilisés à l’interface.

l’interface, mais le solveur ne diverge pas : on atteint en effet la limite de mortiers à ajouter, mais on peut supposer qu’on retrouverait cette augmentation soudaine de l’écart si le maillage à l’interface était plus fin, et donc le nombre de modes de Neumann calculables plus élevé.

Il est toutefois intéressant de noter que, et ce pour le modèle à 100 modes également, l’écart plonge de $5 K$ à $0,5 K$ quand on passe de 2 à 5 mortiers, mais reste stable pendant un long moment, avant de “plonger” vers l’optimum. Aux alentours de 30 modes de Neumann, l’écart chute là encore très rapidement (à l’instar de ce que l’on observait sur la figure 3.7.6), avant de croître de manière monotone avec l’augmentation du nombre de mortiers utilisés. En revanche, l’écart pour le modèle à 100 modes continue de décroître, jusqu’à se stabiliser autour de $0,01 K$. Cependant, il convient de nuancer ce propos, puisqu’on remarque qu’après 70 mortiers utilisés, il n’y a plus de décroissance strictement monotone, et que l’écart semble fluctuer de manière très sensible autour de $0,011 K$, et est en réalité en augmentation sur les deux derniers points de la courbe. Il ressort de cette analyse, que la réduction du flux à l’interface bride la précision du modèle réduit. On a ainsi mis en évidence l’importance et l’influence des fonctions mortiers sur l’écart de réduction, et il ressort de cette analyse que le nombre de fonctions utilisées doit être inférieur au nombre de modes, entre 2 fois et 1,25 fois

moins pour ce cas précis.

Toutefois, un autre problème apparaît, et ce pour tous les modèles réduits; si on prend le cas avec 50 modes réduits, l'écart optimal est à $0,05 K$. Cet écart semble insignifiant, d'autant plus lorsque on le compare à la plage de température de $6600 K$ de la simulation, mais d'après la table 3.7.1, on était en droit d'atteindre un écart moyen de $0,008 K$, *i.e.* l'écart de projection pour la sous-structure 2, qui est le plus important. On retrouve ainsi le même phénomène que celui étudié dans la section 6.4.2. On va donc augmenter la valeur de ζ aux frontières en contact pour le calcul des bases.

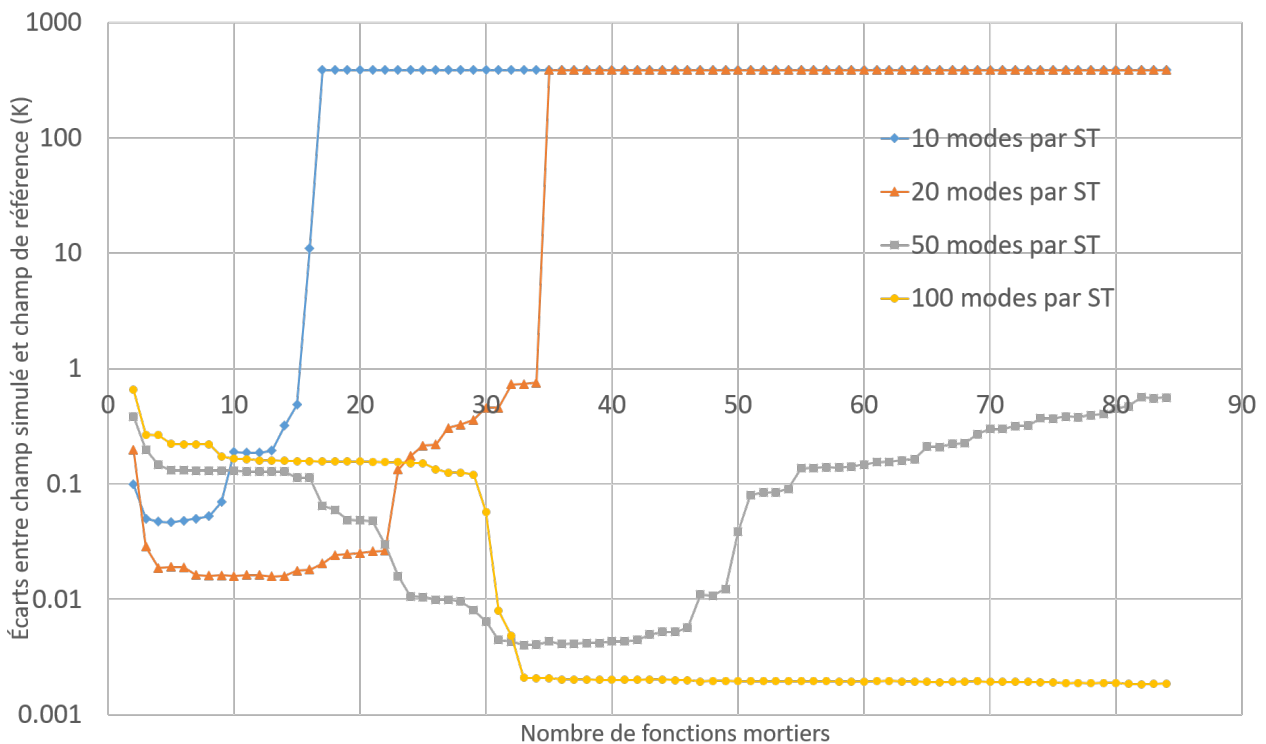


FIGURE 3.7.9 – Écart moyen entre le champ issu de la simulation réduite avec 10, 20, 50 et 100 modes par sous-structure et champ de référence, en fonction du nombre de mortiers utilisés à l'interface, les bases étant calculées avec $\zeta = 10\,000$ sur la frontière de contact.

La figure 3.7.9 présente ainsi les mêmes courbes que la figure 3.7.8, mais cette fois-ci avec des bases calculées avec une valeur $\zeta = 10\,000$ imposée sur la frontière de contact. Si les allures sont similaires d'un graphe à l'autre, l'échelle est différente, avec un écart bien plus faible (presque 10 fois moins élevé) lorsque ζ est plus élevé.

Dans le détail, si on compare l'écart obtenu après simulation avec 100 modes réduits par sous-structure, on est aux alentours de $0,002 K$, ce qui est en accord avec l'écart calculé à partir du champ projeté pour la sous-structure 2 (voir table 3.7.2). Par ailleurs, en comparant

Nombre de modes	$ T^{(1)} - \tilde{T}_{proj}^{(1)} $	$ T^{(2)} - \tilde{T}_{proj}^{(2)} $
10	1,8E-04	0,057
20	5,15E-05	0,021
50	1,78E-05	0,005
100	9,02E-06	0,002

TABLE 3.7.2 – Écart entre champ projeté et champ de référence en fonction du nombre de modes des bases réduites pour les sous-structures 1 et 2 avec $\zeta = 10\,000$.

cette courbe à celle illustrée par la figure 3.7.5, on se rend compte que la précision du modèle est déterminée par la réduction du flux à l'interface jusqu'à environ 33 modes gardés, mais qu'elle est ensuite limitée par la réduction modale du champ de température.

On a ainsi confirmation des résultats obtenus par l'étude pour le cas de simulations sous-structurées avec un maillage conforme, à savoir que l'augmentation de la valeur du nombre de Steklov sur les interfaces de contact permet d'aligner les écarts de projection et de simulation. Il ressort également de cette étude que, pour que le modèle réduit converge vers le modèle complet, il convient de choisir le bon nombre de mortiers, suffisant pour permettre la convergence mais sans dépasser une valeur critique qui ferait complètement diverger le problème. Heureusement, il n'existe pas "une" valeur optimale du nombre de mode de Neumann à conserver, mais toute une gamme sur laquelle l'écart varie très peu.

On va donc devoir étudier l'influence du nombre de mortiers sur le modèle de carte simplifiée étudié dans ce chapitre pour obtenir un modèle réduit performant.

7.3 Modélisation d'un ensemble complexe

Nous traitons finalement un modèle élémentaire de carte électronique. Il est constitué d'un circuit imprimé de conductivité orthotrope sur lequel sont placés huit composants (voir figure 3.7.10). Deux composants (le 3 et le 7) sont actifs et dégagent une puissance thermique interne. Le composant 3 dégage 9 W de manière homogène dans tout son volume. Le composant 7 représente une esquisse simplifiée de puce en silicium encapsulée dans une résine, et la puissance ($4,32\text{ W}$) est dégagée sur la partie supérieure de la puce. Un dissipateur thermique en cuivre permet d'évacuer cette puissance vers le circuit imprimé. Les sous-structures 3 et 7 seront les composants critiques dont la température maximale sera surveillée.

Les autres composants sont passifs, et sont des blocs homogènes de différentes taille et forme. Les dimensions de ces différents éléments sont les suivantes : $12,3 \times 13,6 \times 0,1\text{ cm}$ pour le PCB, $2 \times 1 \times 0,75\text{ cm}$ pour les composants 1 et 2, $2 \times 3,2 \times 0,75\text{ cm}$ pour le composant 3,

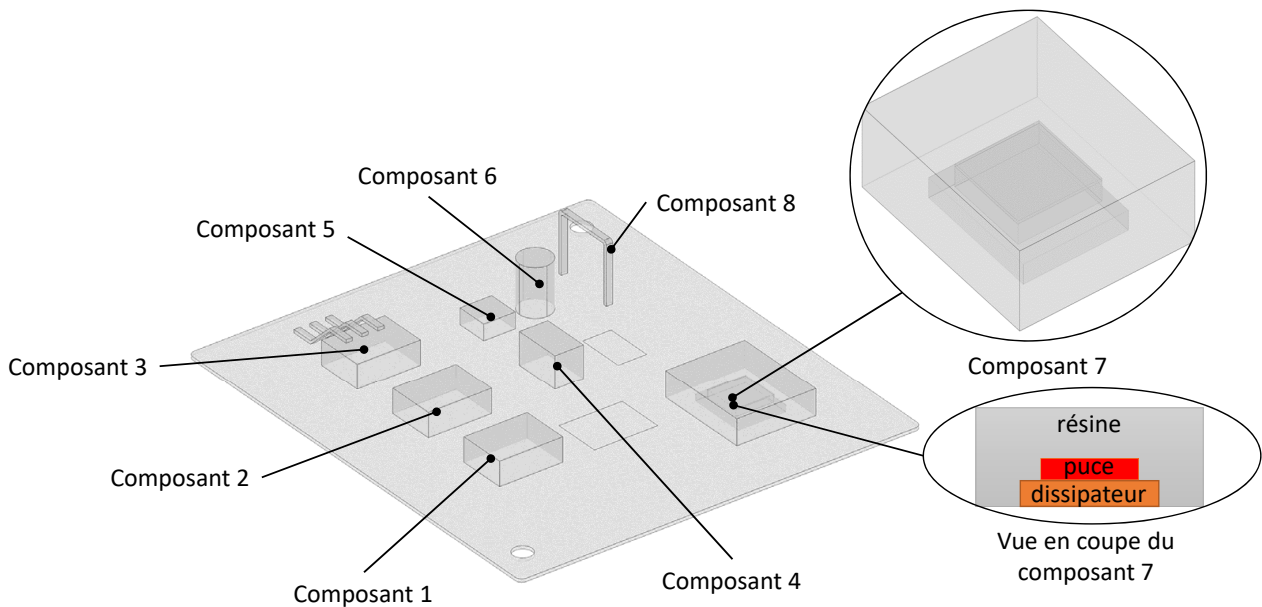


FIGURE 3.7.10 – Modèle de carte élémentaire avec les différents composants numérotés.

$0,9 \times 1,5 \times 1 \text{ cm}$ pour le composant 4, $1 \times 1 \times 0,5 \text{ cm}$ pour le composant 5, $1 \times 1 \times 1,5 \text{ cm}$ pour le composant 6, $2,5 \times 3 \times 0,8 \text{ cm}$ pour le composant 7 et $0,2 \times 2 \times 1,9 \text{ cm}$ pour le composant 8.

Les propriétés physiques de tous ces éléments sont résumés dans la table 3.7.3 et sont représentatives de celles rencontrées dans l'électronique. On considère enfin qu'on a un coefficient d'échange global $h_{ext} = 10 \text{ W.m}^{-2}.\text{K}^{-1}$ sur toutes les surfaces externes, avec une température externe de 0°C .

La décomposition en sous-structures sera la même que précédemment : le PCB et chaque composant constituent un ensemble de neuf sous-structures. Cet exemple simple est composé de plusieurs structures en contact. On note que les composants 3 et 8 échangent de la chaleur avec le PCB via plusieurs frontières géométriquement distinctes. La frontière entre le PCB et le composant 7 est quant à elle "hétérogène" : la partie centrale relie le PCB au dissipateur en cuivre, alors que la partie externe couple le PCB avec une résine isolante. Cet exemple expose ainsi les différentes configurations rencontrées dans un cadre industriel.

7.4 Validation du modèle sous-structuré détaillé

Avant de tester la réduction du modèle de test, on va dans un premier temps comparer les résultats obtenus avec une simulation sous-structurée selon la méthode des joints utilisant

Composant	$\kappa (W.m^{-1}.K^{-1})$	$C = \rho.c_p$ ($MJ.m^{-3}.K^{-1}$)
PCB	20,7 selon x	1,75
	20,7 selon y	
	0,34 selon z	
Composant 1	90,7	3,92
Composant 2	90,7	3,92
Composant 3	puce : 144	1,65
	pattes : 401	3,45
Composant 4	23,5	3,49
Composant 5	90,7	3,92
Composant 6	50	1,27
Composant 7	résine : 0,67	1,75
	support : 260	3,43
	puce : 144	1,65
Composant 8	2,1	1,75

TABLE 3.7.3 – Propriétés physiques des matériaux utilisés pour la carte de test (la couche 2 du PCB est la couche en contact avec les composants)

des éléments finis (2.4.26) aux résultats obtenus via une simulation détaillée complète telle qu'utilisée classiquement lorsqu'on ne souhaite pas faire de décomposition de domaines.

Le maillage non-conforme utilisé est celui présenté en figure 3.7.11. Il est constitué d'un total de 9 079 nœuds, très inégalement répartis : on a 7 279 nœuds pour le PCB, mais 1120 pour le composant 3, 2264 pour le composant 7 et entre 200 et 400 nœuds pour les autres composants, pour un total de 12 415 nœuds.

L'agrandissement sur la figure 3.7.11 montre que les nœuds ne coïncident pas à l'interface. Le champ de température au terme de 1000s de simulation est présenté sur la figure 3.7.12. La température maximale est comparée à celle obtenue avec une simulation éléments finis classique. On note ainsi une différence de $0,02^{\circ}C$ sur la température maximale, puisque en non conforme on n'atteint que $101,11^{\circ}C$. Si on compare l'évolution de la température maximale avec les autres simulations (figure 3.7.13), elle est très semblable en conforme et non conforme, avec seulement un très faible écart en stationnaire. Cet écart peut être dû à la plus grande finesse du maillage des composants dans le cas non-conforme.

Ce très faible écart en température nous amène à considérer que les résultats obtenus avec un maillage non-conforme sont validés par rapport à une simulation monobloc. On va donc

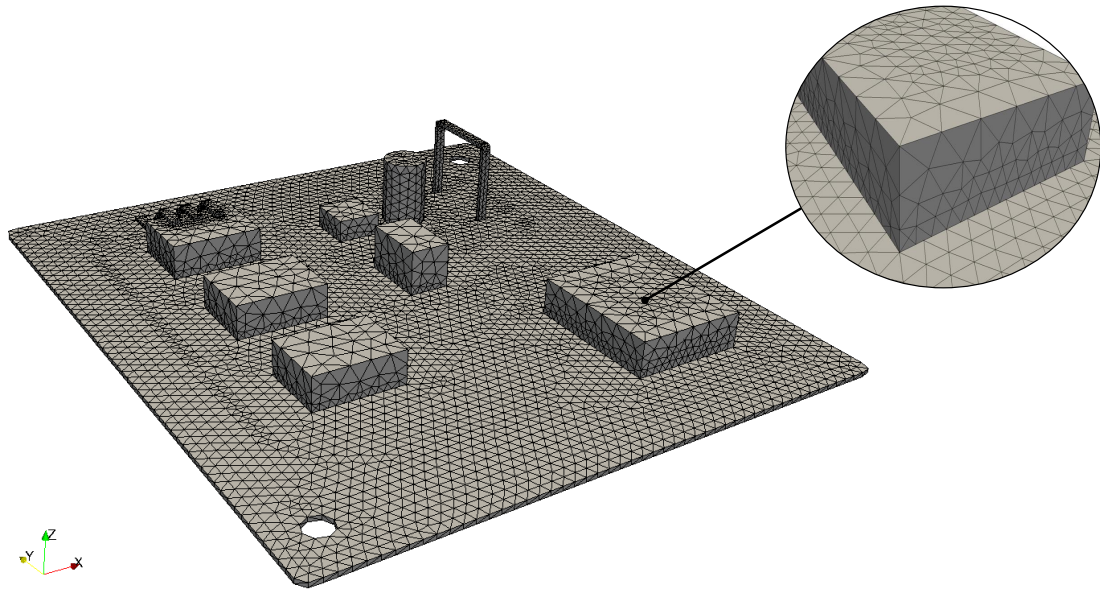


FIGURE 3.7.11 – Maillage non conforme pour le modèle de test (avec zoom sur le composant 2).

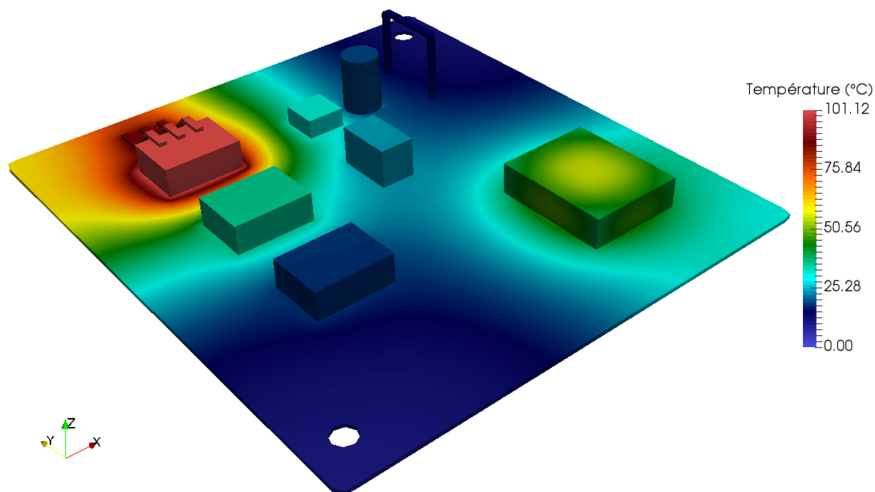


FIGURE 3.7.12 – Champ de température à $t = 1000s$ de la simulation avec méthode des joints utilisant un maillage non conforme.

pouvoir prendre ces résultats comme référence pour estimer l'écart de réduction à la suite d'une simulation utilisant des modèles réduits.

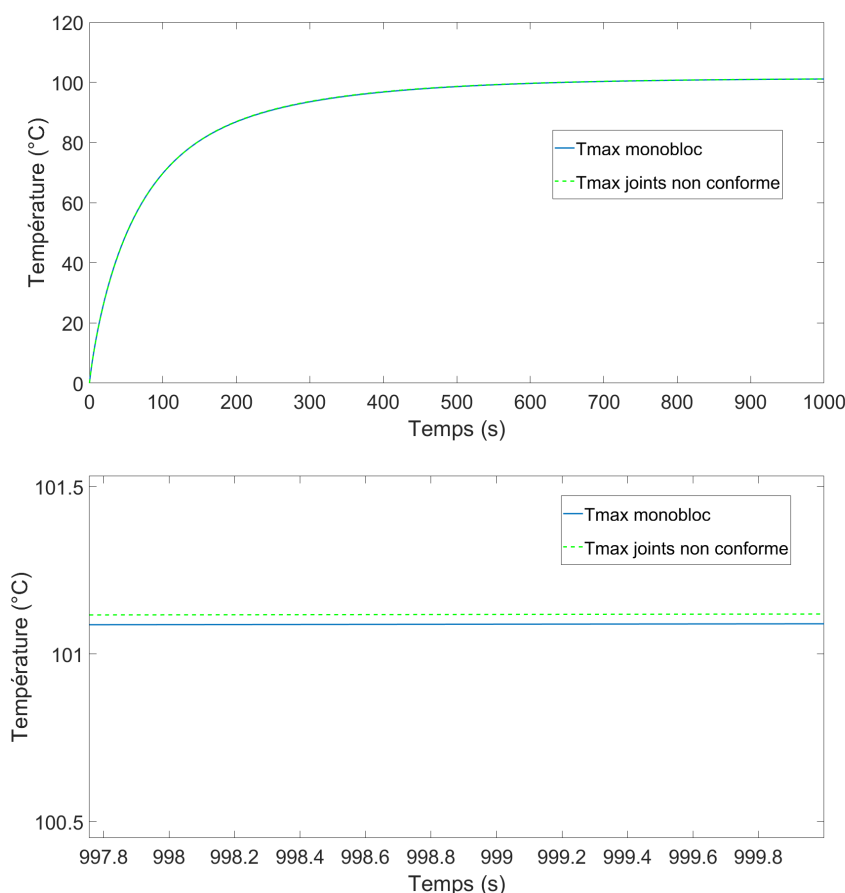


FIGURE 3.7.13 – Évolution de la température maximale au cours du temps avec : en bleu, la température maximale au cours de la simulation monobloc, et en vert (ligne pointillée) la température maximale au cours de la simulation sous-structurée avec méthode des joints pour le maillage non conforme.

7.5 Réduction de modèle

7.5.1 Création des modèles réduits

Du fait du faible nombre de nœuds, les bases de Dirichlet-Steklov de toutes les sous-structures sont calculées en seulement trois minutes. Les modes sont calculés avec un paramètre de Steklov $\zeta = 10^4$ sur les interfaces de contact.

La figure 3.7.14 présente ainsi des exemples de modes de Steklov pour chaque sous-structure. On observe à quel point ceux-ci sont “spécialisés” aux interfaces de contact, où l’on a appliqué un grand coefficient ζ , ce qui permet un meilleur couplage des sous-structures.

Les bases sont ensuite réduites par amalgame modal avec comme référence le champ de

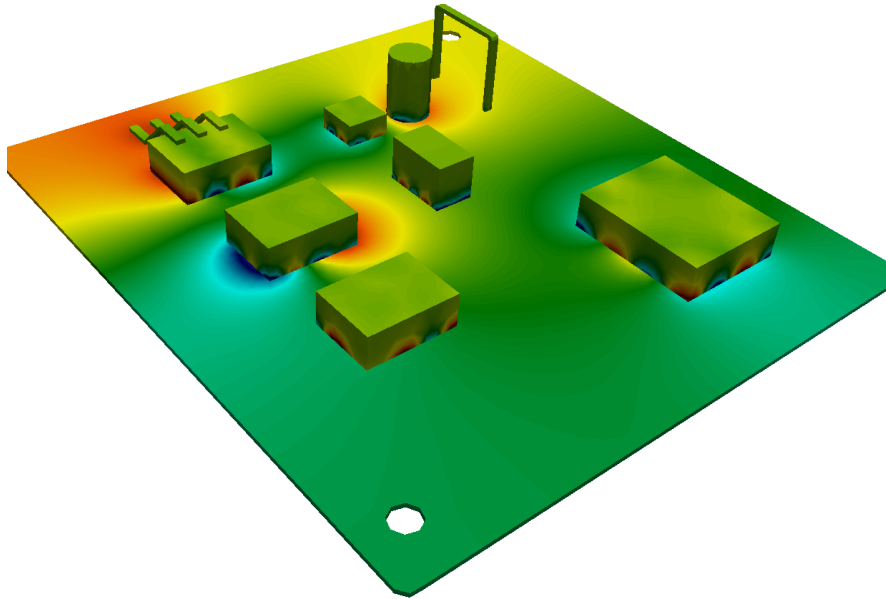


FIGURE 3.7.14 – Assemblage de modes de Steklov des différentes sous-structures du modèle.

température obtenu à l'issue de la simulation dynamique de 1 000s effectuée dans la section précédente. À la suite de l'analyse des bases réduites après projection sur le champ de référence, on retient 30 modes pour le PCB, 30 modes pour le composant 7 et 10 modes pour chacun des autres composants. En effet, le PCB et le composant 7 sont plus complexes que les autres sous-structures, et nécessitent des bases plus riches.

Le modèle réduit est donc de dimension 150, soit un facteur de réduction de 86.

On calcule ensuite pour chaque interface une dizaine de modes de Neumann seulement (quelques exemples sont exposés sur la figure 3.7.15). La jonction entre le composant 7 et le circuit imprimé est “inhomogène” puisque la partie centrale relie le PCB au dissipateur en cuivre, alors que la partie externe couple le PCB avec une résine isolante. Deux choix s'offrent à nous : soit prendre des fonctions mortiers englobant toute la frontière, soit définir deux famille de fonctions mortiers : l'une pour la frontière avec la résine, et l'autre pour la frontière avec le dissipateur. C'est ce dernier choix qui a été fait, sa pertinence pourra être discutée : en effet, comme illustré par la figure 3.7.15d, la continuité du flux dans le plan de l'interface n'est pas garantie.

On retient entre 2 et 10 modes de Neumann sur chaque frontière de contact : le modèle total comprenant les matrices réduites de chaque sous-structure et les matrices de joints \mathbf{J} est alors de dimension 195.

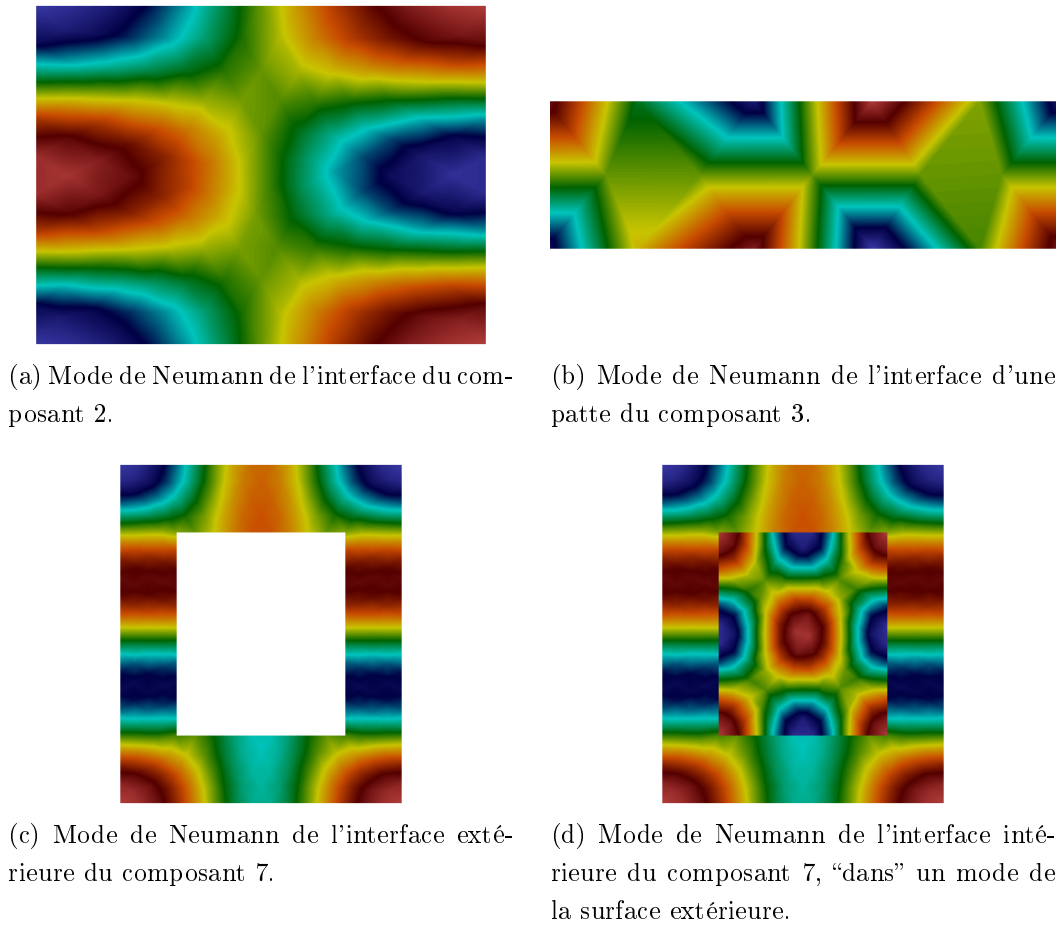


FIGURE 3.7.15 – Exemples de modes de Neumann pour les interfaces de contact.

7.5.2 Simulations réduites

7.5.2.1 Simulation de référence

On commence par faire une simulation réduite dynamique de 1000 s avec les mêmes paramètres que précédemment, à savoir un coefficient d'échange global $h_{ext} = 10 W.m^{-2}.K^{-1}$ sur toutes les surfaces externes, avec une température externe de $0^{\circ}C$, et une puissance dégagée constante. Le champ de température obtenu à l'issue de la simulation réduite est présenté en figure 3.7.16. La température maximale atteinte est $101,94^{\circ}C$, soit un écart de $0,67^{\circ}C$ avec le modèle complet. Si on regarde l'évolution de la température maximale du composant 3 et de l'écart par rapport à la température maximale obtenue avec le modèle complet (figure 3.7.17a), qui était le paramètre critique de l'exemple du chapitre 6, on constate que cet écart dépasse légèrement le seuil de 1% au début de la simulation.

Ce critère est presque respecté pour le composant 7 (3.7.17b), avec seulement $0,61^{\circ}C$

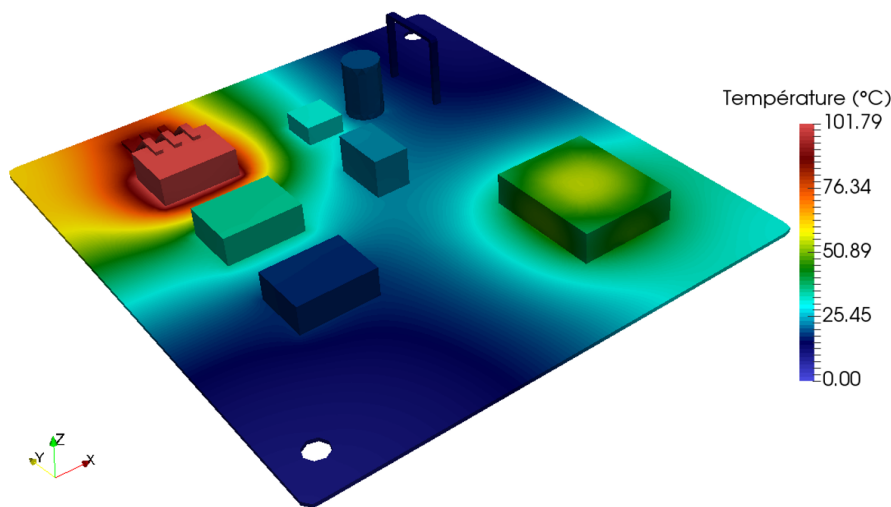
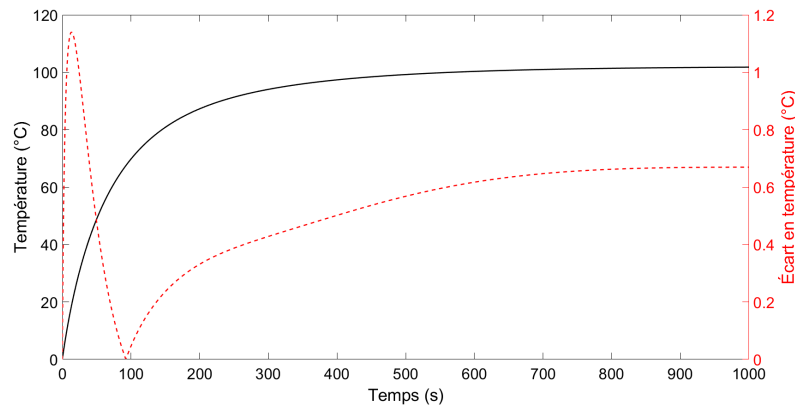


FIGURE 3.7.16 – Champ de température à $t = 1000s$ de la simulation avec méthode des joints réduite.

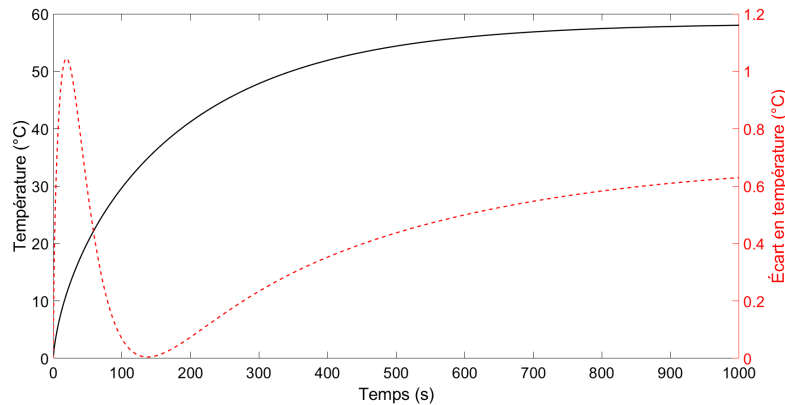
d'écart une fois le régime permanent atteint. On remarque sur ces deux graphiques un minima à $0^{\circ}C$ d'écart : celui-ci correspond à l'inversion de signe de l'écart (dont on donne ici la valeur absolue) : avant ce minima, la température de la simulation réduite est inférieure à celle de la simulation complète, et supérieure ensuite. On a aussi un maxima dans les premiers pas de temps, qui se résorbe tout de suite après : en appliquant une puissance dissipée immédiate, on est assez sévère pour le modèle réduit, qui n'est pas très performant à cet instant. En réduisant la base, on écarte les modes à dynamiques rapides. Cela est d'autant plus vrai dans une base de Dirichlet-Steklov où seuls les modes de Dirichlet ont une dynamique propre. Or, dans cet exemple, la procédure d'amalgame n'a fait ressortir qu'un seul mode maître de Dirichlet pour chaque sous-structure. Il n'est dès lors pas étonnant d'avoir un écart significatif aux premiers instants de la simulation. Mais cela reste très limité dans le temps, et à ce moment de la simulation la température reste basse, il n'y a donc pas trop d'impact sur l'évaluation de la température critique atteinte par les composants au cours du temps.

Quoiqu'il en soit, l'écart est très faible : sur toute la durée de la simulation ($1\ 000s$), l'écart moyen est de $0,41^{\circ}C$. Si on regarde les écarts un peu plus en détail, on s'aperçoit que cet écart moyen vient surtout du PCB : l'écart est de $0,43^{\circ}C$ pour cette sous-structure. Le circuit imprimé couplant toutes les autres sous-structures est critique. Il "absorbe" les erreurs provenant des autres sous-structures, et les propage. La réduction du PCB doit donc faire l'objet d'un soin particulier.

Concernant les temps de calcul, le modèle réduit a été résolu en moins de $1,1s$ contre $330s$ pour le modèle complet, soit un facteur de réduction de 300. Toutefois, ces temps de calcul



(a) Évolution du composant 3.



(b) Évolution du composant 7.

FIGURE 3.7.17 – Évolution de la température maximale atteinte par les composant 3 et 7 au cours de la simulation réduite (courbe noire), et écart (absolu) par rapport à la température maximale au cours de la simulation sous-structurée complète (courbe rouge).

sont à relativiser : le maillage utilisé est très petit, avec seulement 12 000 nœuds, il convient donc de tester la méthode avec des modèles plus imposants, comme la carte ISL8026EVAL3Z vu dans le chapitre 6, dont les temps de résolution sont plus importants, et avec les mêmes solveurs que ceux utilisés pour résoudre le problème avec un maillage conforme. Ces derniers sont en effet plus optimisés mais leur implémentation pour la méthode des joints représente encore une importante quantité de travail.

7.5.2.2 Simulation avec un coefficient d'échange $h_{ext} = 10 \text{ W.m}^{-2}.\text{K}^{-1}$ et puissance dissipée doublée

Afin de tester la polyvalence des composants réduits que l'on vient de créer, on va les utiliser pour effectuer une simulation en changeant les paramètres : dans un premier temps, on double la puissance dissipée par les deux composants 3 et 7. Le champ de température

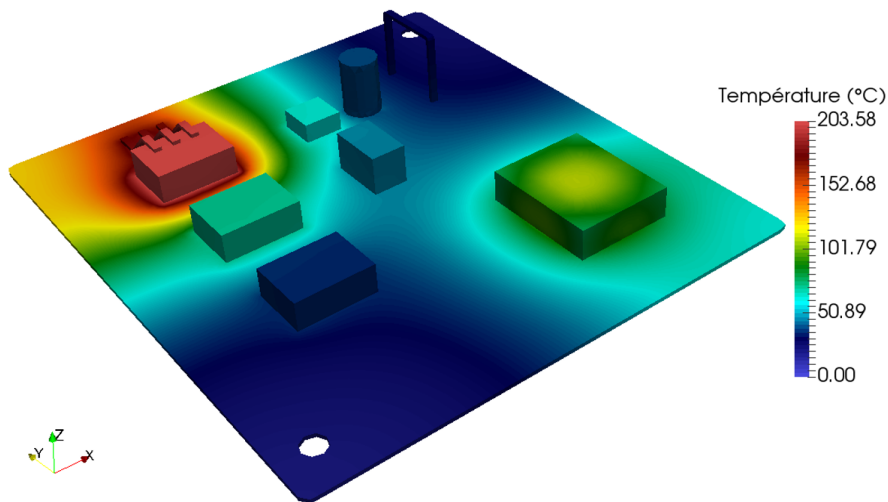
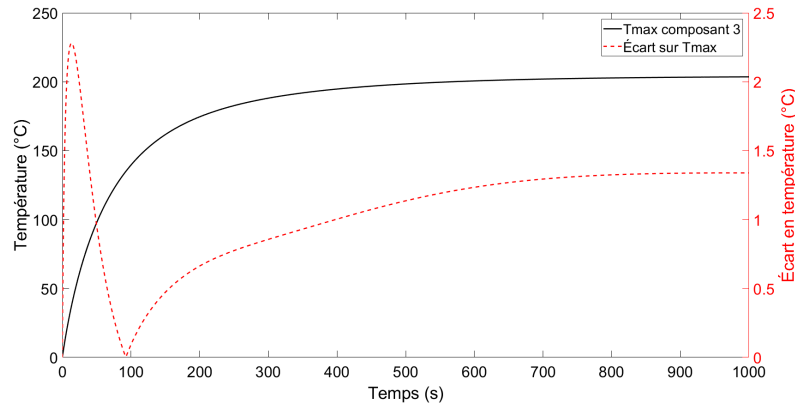


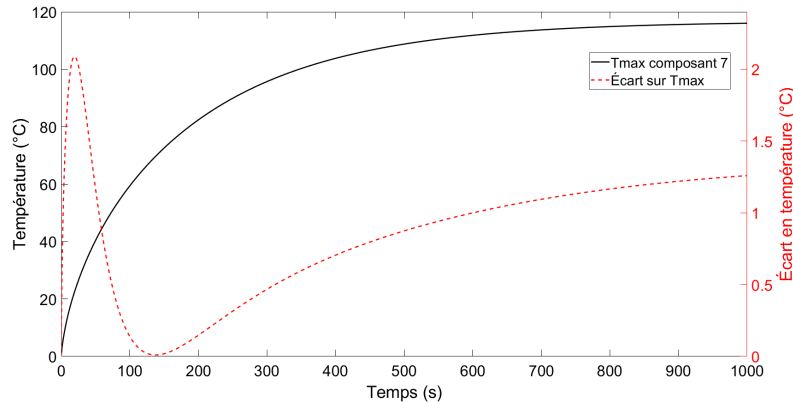
FIGURE 3.7.18 – Champ de température à $t = 1000\text{s}$ de la simulation avec méthode des joints réduite, avec une puissance dissipée par les composants doublée.

obtenu à la suite d'une simulation réduite avec ce nouveau paramètre est présenté sur la figure 3.7.18. La température maximale atteinte par la carte est alors de $203,58^\circ\text{C}$. Si on compare les champs de température issus de cette simulation réduite à ceux issus d'une simulation détaillée, on obtient un écart moyen de $0,83^\circ\text{C}$ sur les 1000 s de simulation. Dans le détail, l'écart sur le composant 3 est de $1,14^\circ\text{C}$, et de $0,75^\circ\text{C}$ sur le composant 7. On reste ainsi en dessous de 1% d'écart pour le composant 3 (figure 3.7.19a) et légèrement au dessus pour le composant 7 (3.7.19b).

Concernant les temps de calcul, le modèle réduit a été résolu en $1,12\text{ s}$, contre 375 s pour le modèle détaillé correspondant, soit un facteur de réduction de 335. Les écarts observés sont donc tout à fait acceptables par rapport aux gains en temps de calcul.



(a) Évolution du composant 3.



(b) Évolution du composant 7.

FIGURE 3.7.19 – Évolution de la température maximale atteinte par les composant 3 et 7 au cours de la simulation réduite (courbe noire), et écart (absolu) par rapport à la température maximale au cours de la simulation sous-structurée complète (courbe rouge) lorsque ces deux composants dégagent une puissance doublée.

7.5.2.3 Simulation avec un coefficient d'échange $h_{ext} = 50 W.m^{-2}.K^{-1}$

Dans un second temps, on teste le modèle réduit avec un coefficient d'échange $h_{ext} = 50 W.m^{-2}.K^{-1}$, *i.e.* pour modéliser une convection forcée, comme si la carte était refroidie par un ventilateur.

Après comparaison avec une simulation détaillée correspondante, on obtient un écart moyen de $1,75^{\circ}C$ entre modèles réduit et détaillé. C'est un écart assez élevé, surtout que la température maximale atteinte est de $44^{\circ}C$. Cet écart est encore plus élevé pour le composant 7, qui présente un écart moyen de $4,12^{\circ}C$ avec la simulation détaillée, soit presque 10% d'écart. On augmente la taille des bases réduites pour réduire cet écart : on retient cette fois

70 modes amalgamés pour le PCB et également 70 modes amalgamés pour le composant 7. Le champ réduit obtenu est représenté sur la figure 3.7.20. L'écart moyen est alors de seulement $0,38^{\circ}\text{C}$, et plus remarquable encore, de seulement $0,34^{\circ}\text{C}$ pour le composant 7. On a plus que divisé l'écart par 10. Malgré l'augmentation de la taille des modèles réduits, le temps de calcul reste très faible, avec 2,9 s pour résoudre le problème réduit, contre 373 pour le problème détaillé. Si on regarde l'évolution de la température maximale pour le composant 3,

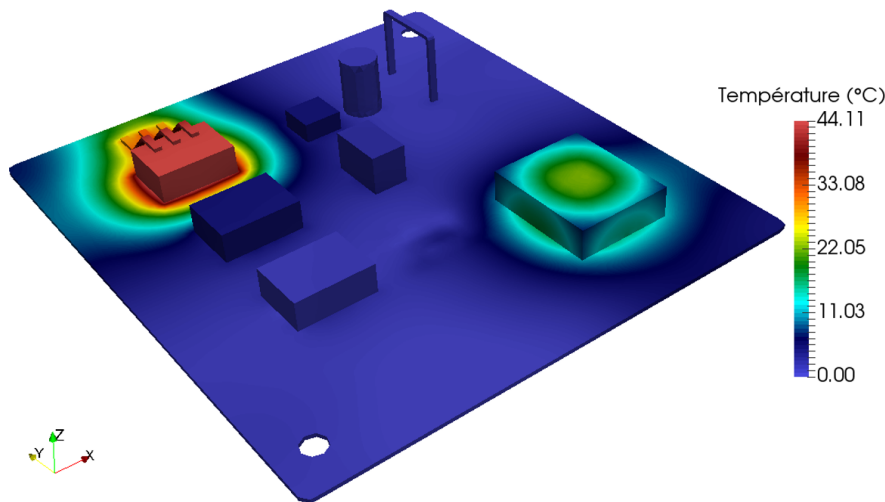
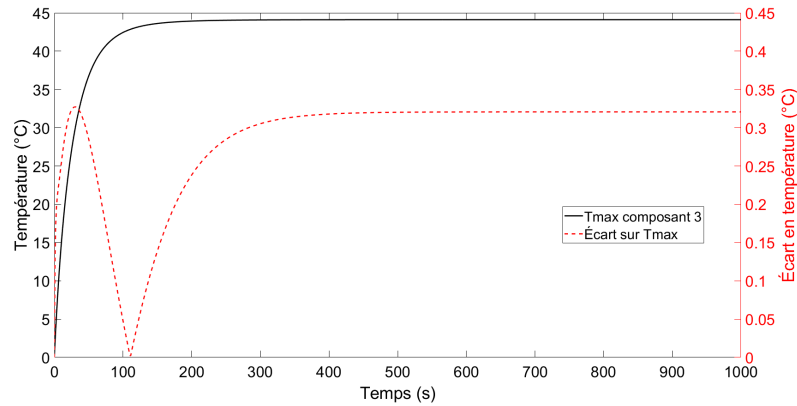


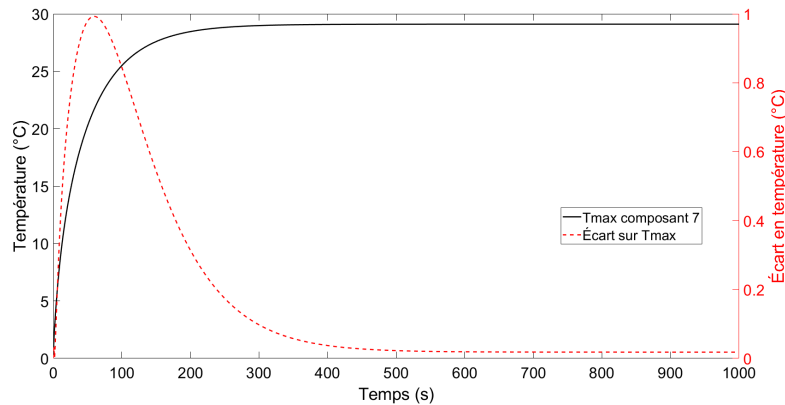
FIGURE 3.7.20 – Champ de température à $t = 1000\text{s}$ de la simulation avec méthode des joints réduite, avec $h_{ext} = 50\text{ W}\cdot\text{m}^{-2}\cdot\text{K}^{-1}$.

qui est celui qui chauffe le plus (figure 3.7.21a), le régime stationnaire est évidemment atteint plus rapidement. L'écart entre les modèles détaillé et réduit ne dépasse jamais 1% sur cette température critique.

Pour le composant 7 (figure 3.7.21b), on a de très bon résultats, avec un écart en deçà de $0,02^{\circ}\text{C}$ en régime permanent, mais on a comme précédemment un pic à 1°C d'écart pendant la montée en température.



(a) Évolution du composant 3.



(b) Évolution du composant 7.

FIGURE 3.7.21 – Évolution de la température maximale atteinte par les composant 3 et 7 au cours de la simulation réduite (courbe noire), et écart (absolu) par rapport à la température maximale au cours de la simulation sous-structurée complète (courbe rouge) avec $h_{ext} = 50 \text{ W.m}^{-2}.K^{-1}$.

7.5.3 Simulation d'une nouvelle configuration

Pour ce test, on change la configuration de la carte : le composant 1 est cette fois-ci déplacé sur une autre surface d'implémentation prévue pour le recevoir, à côté du composant 7 (voir figure 3.7.22). On ne remaille pas l'ensemble du modèle, on conserve tous les maillages et modèles réduits associés créés lors de la réduction du modèle précédent.

On conserve les paramètres de simulation du modèle de référence utilisé pour la réduction, à savoir une puissance constante dans le temps et un coefficient d'échange global $h_{ext} = 10 \text{ W.m}^{-2}.K^{-1}$. Le champ de température obtenu au terme d'une simulation dynamique de 1000 s est présenté en figure 3.7.23. La température maximale atteinte est de 101,06 °C.

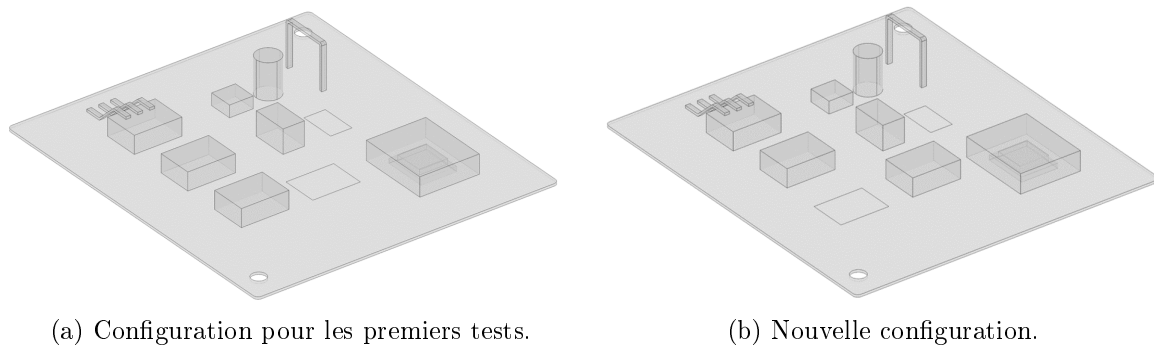
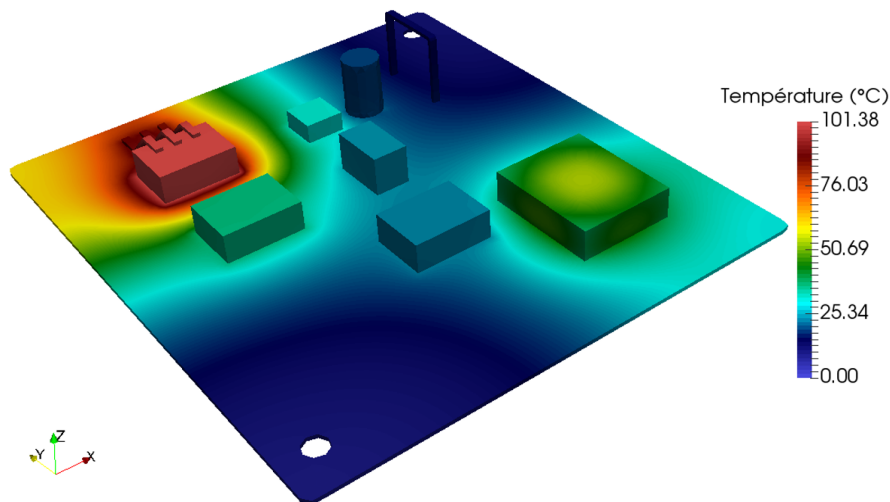


FIGURE 3.7.22 – Le composant 1 change de surface d'implémentation.

FIGURE 3.7.23 – Champ de température à $t = 1000s$ de la simulation éléments finis avec méthode des joints de la nouvelle configuration.

On effectue ensuite une simulation réduite de cette nouvelle configuration, utilisant les modèles réduits créés initialement (*i.e.* avec 30 modes pour le PCB, 30 modes pour le composant 7 et 10 modes pour chacun des autres composants). Le résultat de cette simulation est présenté sur la figure 3.7.24.

On constate tout de suite un écart sur la température maximale atteinte par le modèle réduit par rapport au modèle complet, avec environ $2.3^{\circ}C$ d'écart, soit un peu plus de 2% de différence. On va regarder en détail l'évolution de la température maximale sur le composant 1, qui a été déplacé, et les composants 3 et 7. Les températures maximales atteintes par les modèles complet et réduit sur ces composants sont représentés sur le graphe en figure 3.7.25.

Ainsi, pour les composants 3 et 7, les écarts sont assez faibles, avec respectivement $1,7^{\circ}C$ et $0,7^{\circ}C$ d'écart. En revanche, pour le composant 1, les simulations complètes et réduites

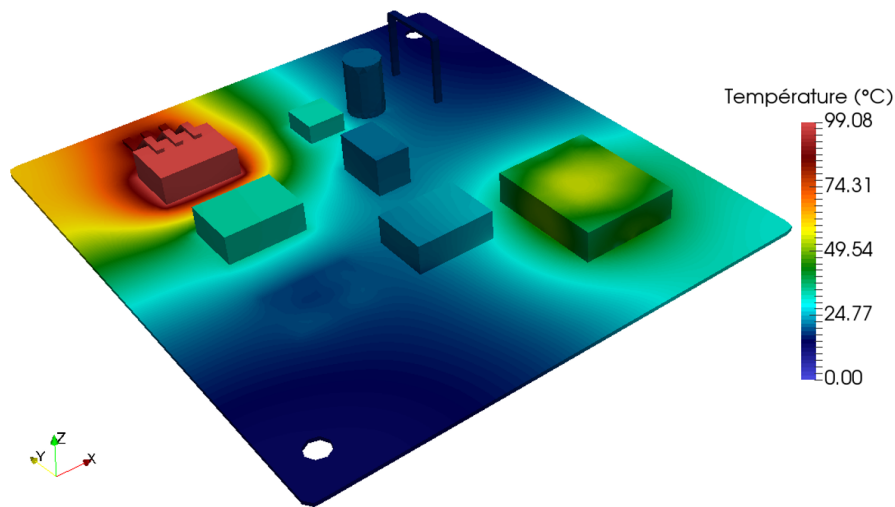


FIGURE 3.7.24 – Champ de température à $t = 1000s$ de la simulation réduite avec méthode des joints de la nouvelle configuration.

évoluent de manière similaire, mais avec un décalage en température, qui est plus élevée dans le modèle réduit. En stationnaire, la température maximale est de $28^{\circ}C$ pour le modèle réduit contre $23^{\circ}C$ pour le modèle détaillé. On a donc un écart assez important sur ce composant déplacé. On voit donc sur cet exemple simple que même avec une réduction naïve, les résultats sont assez bons, puisque l'écart sur ce composant non critique est de $5^{\circ}C$, à comparer aux $100^{\circ}C$ d'augmentation de l'ensemble de la carte. De plus, le comportement des composants actifs est fidèlement reproduit. La précision pourrait encore être améliorée en choisissant avec plus de soin le champ de référence utilisé pour réduire les bases des composants. En particulier, comme l'augmentation de température des composants passifs est provoquée uniquement par le circuit imprimé, sa réduction doit être bien réfléchie pour pouvoir être encore plus polyvalente. Il faut notamment éviter des modes très (trop?) spécialisés pour reconstruire le champ de référence avec des paramètres bien précis.

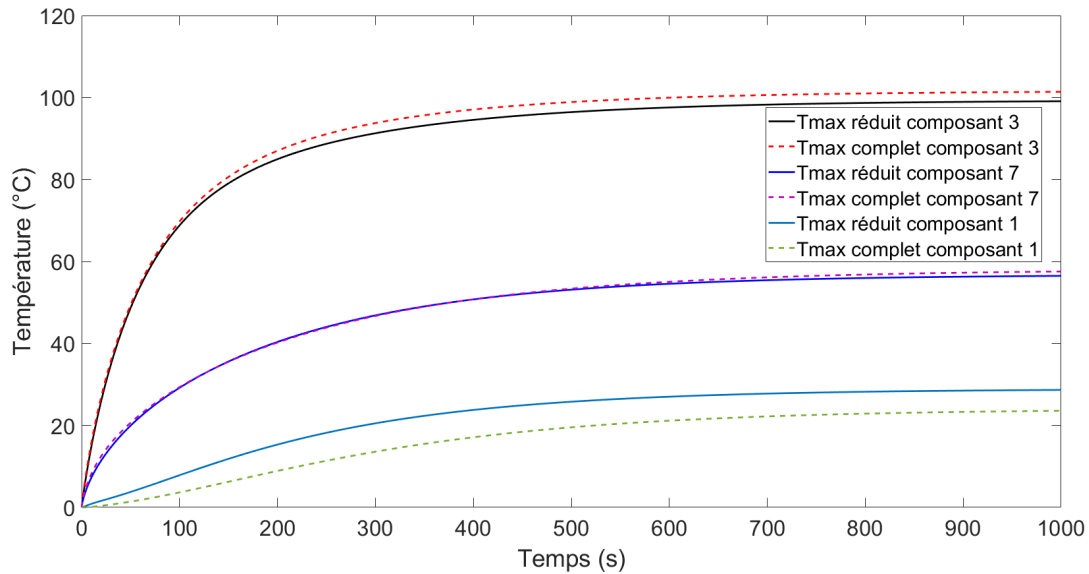


FIGURE 3.7.25 – Évolution de la température maximale au cours du temps pendant les simulation complète (lignes pleines) et réduite (lignes pointillées) pour, de haut en bas, les composants 3, 7 et 1 dans la nouvelle configuration.

7.6 Conclusions et perspectives de cette deuxième étude

L'objectif de cette étude était d'évaluer l'efficacité de la méthode de sous-structuration modale sur des maillages non-conformes utilisant la méthode des joints. L'étude sur un modèle simple avec un seul contact a montré de très bons résultats, puisqu'on parvient à retrouver l'écart de projection après une simulation réduite. On a pu observer qu'il existait ainsi une plage de valeurs optimales pour le nombre de fonctions mortiers utilisées à l'interface pour minimiser l'écart.

Le cas suivant, présentant un modèle de carte élémentaire, a permis de faire une première évaluation dans le cas de contacts multiples. Les résultats présentés sont là aussi très satisfaisants puisque un seul modèle réduit a permis de simuler, avec une bonne précision, plusieurs configurations. En particulier, pour la première fois dans notre équipe de recherche, la géométrie du modèle a été modifiée après réduction. C'est une véritable avancée dans le domaine de la réduction modale et cela ouvre d'intéressantes perspectives industrielles.

Il ressort de cette étude que le nombre de mortiers est un point névralgique, puisque les différents tests menés pour présenter ces résultats ont montré une influence réciproque des différents mortiers. En effet, l'écart entre modèles réduit et détaillés peut varier pour un composant si le nombre de mortiers à l'interface d'un autre composant varie. Cela est peut-

être dû au fait de la configuration particulière d'une carte, à savoir que tous les composants sont reliés au PCB. Si un écart augmente sur ce dernier, il va se répercuter sur les autres composants, en particulier sur les composants passifs. Il reste donc un travail à faire sur le choix des fonctions mortiers, qui apparaît comme crucial, mais aussi sur la référence utilisée pour l'amalgame. En effet, on voit un effet de la spécialisation des modes pour reconstruire la simulation de référence pour le cas où on a déplacé le composant, avec un petit artefact qui apparaît à son ancien emplacement pendant la simulation réduite (figure 3.7.24).

La question se pose ainsi sur le nombre de fonctions mortiers choisis pour chaque interface, mais aussi sur la façon de les choisir. Pour cet exemple, on a choisi les modes de Neumann en les prenant dans l'ordre de leur valeur propre. On pourrait envisager de les choisir eux aussi via une méthode d'amalgame, mais là encore la référence utilisée aura son importance. Plus généralement, se pose aussi la question d'utiliser une ou plusieurs familles de modes comme fonctions mortiers pour des surfaces de contact hétérogènes (comme pour le composant 7 par exemple).

Enfin, on a choisi d'utiliser des modes de Neumann de manière arbitraire. On pourrait aussi tester d'autres modes, comme les modes de Branches évoqués au début du chapitre.

Chapitre 8

Conclusion générale et perspectives

L'objectif de cette thèse est de répondre à une problématique industrielle précise : la simulation de cartes électroniques en un temps raisonnable. Notre partenaire industriel ayant développé depuis des années ses propres procédés, il a fallu proposer une méthode permettant de s'intégrer à terme dans la chaîne de calcul, tout en apportant une véritable plus-value. En particulier, le cahier des charges était strict : il s'agissait de faire une bibliothèque de modèles thermiques réduits de composants indépendants des conditions aux limites, et de pouvoir les coupler à un circuit imprimé. Les modèles de composants doivent

- Être sobres en puissance de calcul,
- Accepter toutes conditions aux limites tout en respectant une certaine finesse géométrique,
- Être réutilisables.

Parmi les différentes options envisageables brièvement rappelées dans l'état de l'art, une approche de réduction modale a été choisie pour les composants et pour le circuit imprimé. Le premier écueil fut de choisir une famille de fonction possédant les propriétés mathématiques adéquates. Deux bases ont été utilisées : une base dite de Dirichlet et une base dite de Steklov. Un important travail théorique a été mené pour démontrer que la réunion de ces bases forme une base pour l'espace de Hilbert $H^1(\Omega)$. Ce résultat essentiel n'est pas classique bien qu'il ne repose que sur des connaissances standard de l'analyse fonctionnelle. De même nous avons mis au jour une norme adaptée à cette base et établi le lien entre $\|u\|_{H(\Omega)}$ et les normes dans $H^1(\Omega)$, $L^2(\Omega)$ et $L^1(\Omega)$. On est alors bien loin de la problématique industrielle ! Il est cependant nécessaire de s'assurer mathématiquement qu'une méthode fonctionne avant de la mettre en œuvre numériquement. Je ne peux ici que souligner l'importance d'avoir effectué ma thèse au sein d'une entreprise qui a une grande culture de la recherche. Malgré une véritable attente de résultats, Thales a laissé une part importante mais nécessaire à des recherches plus

fondamentales.

Cette nouvelle base, associée à son nouveau produit scalaire a amené à modifier l'algorithme d'amalgame modal, socle des méthodes de réduction utilisées au laboratoire.

L'autre aspect du cahier des charges concerne le couplage entre le modèle réduit d'un composant et le circuit imprimé. Là aussi, plusieurs approches étaient possibles. Des approches précédentes (voir par exemple les thèses de P.O. Laffay [89] ou B. Gaume [100]) ont consisté à coupler des modèles modaux entre eux via une résistance de contact. Cette approche a été utilisée dans le cas d'une carte électronique réelle modélisée finement. Dans le cadre de cette thèse, une analyse de sensibilité a montré sur un cas simple que le choix initial de la base de Steklov est prépondérant pour la qualité de l'approximation modale. En particulier, il est nécessaire de pondérer différemment la condition aux limites de Steklov selon qu'il s'agit d'une interface de contact ou d'une frontière externe. Grâce à cette pondération, l'approche de couplage par une résistance de contact donne d'excellents résultats dans le cas de maillages conformes aux interfaces. Cependant elle ne résiste pas à un découpage à priori, c'est-à-dire dans le cas de maillage non-conforme aux interfaces.

Il a donc fallu trouver une nouvelle voie. Notre choix s'est porté sur la méthode des mortiers. Là encore, un travail théorique a été nécessaire pour adapter cette méthode (initialement créée pour la méthode des éléments finis) au formalisme modal.

Si la réduction modale des champs de température était un aspect bien maîtrisé par l'équipe de recherche, la réduction du flux à l'interface était un saut dans l'inconnu. Nous avons dû trouver une base modale permettant de reconstruire le flux. Nous avons choisi une base de Neumann. Un des regrets de cette thèse est de ne pas avoir pu analyser cette réduction avec le même niveau de détail mathématique que la réduction des champs de température. À notre décharge, les raisons qui ont motivé les inventeurs de la méthode mortier à choisir un espace de co-dimension 2 restent obscures... Nous avons cependant conduit sur un cas simple une étude numérique qui justifie *a posteriori* le choix de cette base. L'utilisation des modes de Branches comme base mortier demanderait à être explorée.

La méthode de sous-structuration modale par la méthode des joints a d'abord été validée et étudiée sur des cas simples. Un premier exemple en 2D a été traité pour faire la preuve du concept (Voir annexe F). Un second exemple très simple, mais en 3D a été ensuite considéré. En effet, étant donné le nombre important de variables entrant dans la méthode (nombre de modes pour la réduction de la température, nombre de modes pour la réduction du flux à l'interface, paramètres des bases modales, . . .), une étude paramétrique sur une carte complète aurait été inexploitable. Les résultats principaux de cette étude sont les suivants :

- La méthode converge avec le nombre de modes pour la réduction en température ;
- La méthode converge jusqu'à un certain point avec le nombre de modes retenus pour

- le flux. Mais au-delà d'une certaine valeur du nombre de fonction de flux, la précision des résultats se dégrade, jusqu'au moment où le système linéaire n'est plus inversible ;
- Le nombre de fonctions mortiers nécessaire à une bonne précision est fonction du nombre de mode retenus pour la réduction de la température. Heureusement, il n'existe pas "une" valeur optimale, mais bien une plage et une heuristique reliant le nombre de mode et le nombre de fonctions mortiers peut être trouvée ;
 - Le dernier enseignement tiré de cet exemple concerne la base de Steklov. Les paramètres utilisés pour son calcul ne sont pas neutres. En particulier, il est nécessaire de dissymétriser la base de Steklov en imposant des conditions aux limites différentes sur l'interface de contact et sur les frontières externes. Là encore le choix de la valeur du facteur de Steklov ζ sur la frontière de contact n'est pas critique, la plage admissible s'étendant sur plus de trois décades.

Enfin, un dernier exemple de carte simplifiée a été étudié. Il confirme les résultats obtenus à l'issue de l'étude précédente, avec pour la première fois dans notre équipe de recherche, la modification de la géométrie du modèle après réduction. C'est un résultat majeur de cette thèse qui ouvre d'intéressantes perspectives industrielles. En revanche, il est plus compliqué de déterminer la plage de mortiers nécessaires pour obtenir la convergence lorsque l'on a plusieurs interfaces, notamment dans le cas d'une carte électronique où tous les composants sont liés au PCB. Il conviendra donc de pousser plus en avant cette étude pour mieux comprendre le comportement des modèles réduits et des influences mutuelles avec les mortiers.

Étant donnés les enjeux économiques autour de l'électronique, il va de soi que la thermique des composants électroniques est un sujet qui occupe depuis des années plusieurs équipes de recherche de par le monde. Il serait donc présomptueux d'affirmer que cette thèse clôt le sujet. Elle apporte néanmoins une ouverture vers une nouvelle approche. Ces 3 ans de thèse auront permis d'asseoir un outil théorique et un début d'outil numérique qui pourra à court terme déterminer si cette approche tient ses promesses. Au-delà des nécessaires optimisations informatiques, un travail de thermicien sera nécessaire pour répondre pleinement au problème posé. Le premier aspect, et à mon sens le plus important, sera de travailler sur l'amalgame. En effet, celui-ci nécessite des états de référence. Si dans un découpage *a posteriori* ceux-ci sont faciles à obtenir, leur obtention s'avère délicate dans le cas d'un découpage *a priori*. En effet, que prendre comme condition aux limites pour l'interface de couplage, quand on ne connaît pas l'autre structure à coupler ? Dans le cas de cartes électroniques le masque d'implantation du composant formant une "mini carte" pourrait être utilisé comme structure générique . Un autre travail essentiel sera de valider nos expérimentations numériques par des expériences physiques, notamment en environnement contrôlé afin de maîtriser les conditions aux limites.

Des compléments plus théoriques pourront aussi être apportés. Comme évoqué plus haut, la réduction du flux à l'interface a été faite de façon brute, et pourrait probablement être améliorée. En particulier, nous nous sommes contentés de choisir les modes de Neumann par ordre des valeurs propres. Une méthode d'amalgame pourrait-elle être envisagée ? Il faudrait alors coupler l'amalgame modal sur les températures avec cette nouvelle procédure. La clé de ce nouvel amalgame sera un produit scalaire (et la norme induite) qui rassemble ces deux aspects. Par ailleurs, le circuit imprimé est un empilement de structures essentiellement bi-dimensionnelles. Une suite logique de ce travail serait de travailler sur le couplage modal de coques.

Annexes

Annexe A

Décomposition d'une fonction de $H^1(\Omega)$ sur la base de Dirichlet-Steklov

À partir des formulations variationnelles nous avons introduit, pour des fonctions u et v de $H^1(\mathcal{D})$, les formes bilinéaires suivantes :

$$\mathcal{A}_{\mathcal{D}}(u, v) = \int_{\mathcal{D}} \overrightarrow{\text{grad}} u \cdot K \cdot \overrightarrow{\text{grad}} v, \quad \mathcal{C}_{\mathcal{D}}(u, v) = \int_{\mathcal{D}} \rho C_p u v, \quad c_{\partial\mathcal{D}}(u, v) = \int_{\partial\mathcal{D}} \zeta u|_{\partial\mathcal{D}} v|_{\partial\mathcal{D}}, \quad (3.A.1)$$

où \mathcal{D} est le domaine dimensionné. Nous pouvons associer à chacune de ces formes une expression adimensionnée pour les fonctions de $H^1(\Omega)$, où contrairement au corps du texte, Ω représente à présent le domaine adimensionné tel quel $x = \frac{X}{L_0}$:

$$a_{\Omega}(u, v) = \int_{\Omega} \nabla u \cdot \kappa \cdot \nabla v, \quad c_{\Omega}(u, v) = \int_{\Omega} c u v, \quad c_{\partial\Omega}(u, v) = \int_{\partial\Omega} \zeta u|_{\partial\Omega} v|_{\partial\Omega}. \quad (3.A.2)$$

On a les relations :

$$K_o L_o = \frac{\mathcal{A}_{\mathcal{D}}(u, v)}{a_{\Omega}(u, v)} = \frac{\mathcal{C}_{\mathcal{D}}(u, v)}{c_{\Omega}(u, v)} = \frac{c_{\partial\mathcal{D}}(u, v)}{c_{\partial\Omega}(u, v)}. \quad (3.A.3)$$

Dans cette expression le coefficient d'échelle $K_o L_o$ est mesuré en $W.K$. Dans la suite nous travaillerons essentiellement avec les formes sans dimensions des opérateurs.

On note alors , $c = \frac{\rho C_p}{C_0}$, $\kappa = \frac{K}{K_0}$, $\frac{h_s L_0}{K_0} = 1$.

Si on pose $Bi = \frac{hL_0}{K_0}$ et $\pi_0 = \frac{\varpi L_0^2}{K_0}$, on peut exprimer le problème thermique (2.4.1) sous

forme locale :

$$\begin{cases} \Omega, & c \frac{\partial T}{\partial t} - \nabla \cdot (\kappa \cdot \nabla T) = \pi_0, \\ \partial\Omega, & \kappa \cdot \nabla T \cdot \underline{n} = \text{Bi}(T_{ext} - T), \\ \Omega, t = 0, & T(x, 0) = T_0(x). \end{cases} \quad (3.A.4)$$

A.1 Continuité et coercivité de l'opérateur de la chaleur.

Notons que si l'opérateur bilinéaire est à la fois continu et coercif nous avons :

$$C' \|u\|_{F(\Omega)}^2 \leq a(u, u) \leq C \|u\|_{F(\Omega)}^2 \implies a(u, u) \equiv \|u\|_{F(\Omega)}.$$

Autrement dit l'opérateur $a(., .)$ induit une norme équivalente à la norme sur $F(\Omega)$.

Proposition A.1.1 *Les opérateurs bilinéaires $c_\Omega(., .)$ et $c_{\partial\Omega}(., .)$ sont symétriques, continus et coercifs dans $L^2(\Omega)$ et $L^2(\partial\Omega)$ respectivement.*

▲

Preuve A.1.1 *La symétrie est évidente. Examinons la continuité. En remarquant la majoration $c^{\max} = \max_{\mathbf{x} \in \Omega} c(\mathbf{x}) > 0$ nous obtenons immédiatement d'après l'inégalité de Schwarz :*

$$|c_\Omega(u, v)| = \left| \int_{\Omega} c u v \right| \leq c^{\max} \left| \int_{\Omega} u v \right| = c^{\max} |(u | v)_{L^2(\Omega)}| \leq c^{\max} \|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}.$$

Ce qui prouve que la forme bilinéaire $c_\Omega(., .)$ est continue dans $L^2(\Omega)$.

De même nous avons par définition la minoration $c^{\min} = \max_{\mathbf{x} \in \Omega} c(\mathbf{x}) > 0$, donc :

$$|c_\Omega(u, u)| = \left| \int_{\Omega} c u^2 \right| \geq c^{\min} \int_{\Omega} u^2 = c^{\min} \|u\|_{L^2(\Omega)}^2.$$

Ce qui prouve que la forme bilinéaire $c_\Omega(., .)$ est coercive dans $L^2(\Omega)$ et qu'elle induit une norme équivalente à la norme usuelle dans $L^2(\Omega)$.

En partant cette fois de l'encadrement $0 < \zeta^{\min} = \min_{\mathbf{x} \in \partial\Omega} \zeta(\mathbf{x}) \leq \zeta \leq \max_{\mathbf{x} \in \partial\Omega} \zeta(\mathbf{x}) = \zeta^{\max} < \infty$ nous obtenons avec la même démarche mais cette fois pour des fonctions de $L^2(\partial\Omega)$:

$$|c_{\partial\Omega}(u, v)| = \left| \int_{\partial\Omega} \zeta u v \right| \leq \zeta^{\max} \left| \int_{\partial\Omega} u v \right| = \zeta^{\max} |(u | v)_{L^2(\partial\Omega)}| \leq \zeta^{\max} \|u\|_{L^2(\partial\Omega)} \|v\|_{L^2(\partial\Omega)},$$

$$c_{\partial\Omega}(u, u) = \int_{\partial\Omega} \zeta u^2 \geq \zeta^{\min} \int_{\partial\Omega} u^2 = \zeta^{\min} \|u\|_{L^2(\partial\Omega)}^2.$$

Ce qui prouve que la forme bilinéaire $c_{\partial\Omega}(\cdot, \cdot)$ est continue et coercive dans $L^2(\partial\Omega)$ et qu'elle induit une norme équivalente à la norme usuelle dans ce même espace. ■

Si nous considérons des fonctions de l'espace $L^2(\bar{\Omega}) = L^2(\Omega) \cap C^0(\bar{\Omega})$ nous pouvons lui associer un produit scalaire défini par :

$$\forall u, v \in L^2(\bar{\Omega}) \quad (u | v)_{L^2(\bar{\Omega})} = \int_{\Omega} u v + \int_{\partial\Omega} u v, \quad (3.A.5)$$

En conséquence de la proposition A.1.1 la forme bilinéaire symétrique définie par $c_{\bar{\Omega}}(\cdot, \cdot) = c_{\Omega}(\cdot, \cdot) + c_{\partial\Omega}(\cdot, \cdot)$ est continue et coercive dans $L^2(\bar{\Omega})$ et induit une norme équivalente à la norme 3.A.5.

Définition A.1.1 Soit $c(\cdot, \cdot)$ la forme bilinéaire symétrique continue et positive définie par¹ :

$$\gamma, \gamma' \geq 0; \gamma + \gamma' \neq 0 \quad c(u, v) = \gamma c_{\partial\Omega}(u, v) + \gamma' c_{\Omega}(u, v).$$

On définit un nouvel opérateur auxiliaire, bilinéaire et symétrique, noté $b_{\varepsilon}(\cdot, \cdot)$:

$$\forall \varepsilon > 0 \quad b_{\varepsilon}(u, v) = a_{\Omega}(u, v) + \varepsilon c(u, v), \quad \forall u, v \in H^1(\Omega).$$

Expression où $a_{\Omega}(\cdot, \cdot)$ est défini par (3.A.2) et ε désigne un réel positif borné quelconque². △

Proposition A.1.2 Sous les conditions de la définition A.1.1 l'opérateur bilinéaire symétrique $b_{\varepsilon}(\cdot, \cdot)$ est continu et coercif dans $H^1(\Omega)$. ▲

L'inégalité de Poincaré joue un rôle majeur dans la preuve de la proposition A.1.2 que nous donnons à présent.

1. Ici γ' et γ sont sans dimensions. En pratique ces facteurs prendront les valeurs 0 ou 1.

2. Du point de vue physique si nous revenions à des opérateurs dimensionnés ce paramètre aurait la dimension d'une fréquence et serait mesuré en Hertz .

Preuve A.1.2 Les propriétés de symétrie, linéarité et coercivité de la forme bilinéaire $c(.,.)$ découlent directement de la proposition A.1.1. La symétrie et la bilinéarité de l'opérateur auxiliaire sont évidentes. Examinons la continuité.

$$\begin{aligned} |b_\varepsilon(u, v)| &= |a_\Omega(u, v) + \varepsilon c(u, v)| \leq |a_\Omega(u, v)| + \varepsilon |c(u, v)| \\ &= |a_\Omega(u, v)| + \varepsilon (|\gamma c_{\partial\Omega}(u, v) + \gamma' c_\Omega(u, v)|) \\ &\leq |a_\Omega(u, v)| + \varepsilon \gamma |c_{\partial\Omega}(u, v)| + \varepsilon \gamma' |c_\Omega(u, v)|. \end{aligned}$$

Nous utilisons l'inégalité de Schwarz :

$$|b_\varepsilon(u, v)| \leq \sqrt{a_\Omega(u, u) a_\Omega(v, v)} + \varepsilon \gamma \sqrt{c_{\partial\Omega}(u, u) c_{\partial\Omega}(v, v)} + \varepsilon \gamma' \sqrt{c_\Omega(u, u) c_\Omega(v, v)}.$$

Montrons que l'opérateur $a_\Omega(.,.)$ est continu dans $H^1(\Omega)$. Pour des fonctions $u, v \in H^1(\Omega)$ on a d'après l'inégalité de Schwarz :

$$|a_\Omega(u, v)| \leq \left| \int_\Omega \nabla u \cdot \kappa \cdot \nabla v \right|^{\frac{1}{2}} \left| \int_\Omega \nabla v \cdot \kappa \cdot \nabla v \right|^{\frac{1}{2}}.$$

Mais la matrice de conductivité est définie positive (voir annexe B.9) et nous avons démontré que si $K^{\max} = \max_{1 \leq i \leq d, \mathbf{x} \in \Omega} K_{oi}(\mathbf{x})$, expression où les coefficients K_{oi} sont les conductivités principales, alors en introduisant $\kappa^{\max} = K^{\max}/K_o$:

$$\left| \int_\Omega \nabla u \cdot \kappa \cdot \nabla v \right| \leq \kappa^{\max} \|\nabla u\|_{L^2(\Omega)}^2 \leq \kappa^{\max} (\|\nabla u\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\Omega)}^2) = \kappa^{\max} \|u\|_{H^1(\Omega)}^2.$$

Ainsi nous obtenons l'inégalité :

$$|a_\Omega(u, v)| \leq \kappa^{\max} \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}.$$

Ce qui prouve que l'opérateur bilinéaire symétrique $a_\Omega(.,.)$ est continu dans $H^1(\Omega)$ (et aussi dans $H_0^1(\Omega)$ puisque $H_0^1(\Omega) \subset H^1(\Omega)$). En particulier

$$a_\Omega(u, u) \leq \kappa^{\max} \|\nabla u\|_{L^2(\Omega)}^2.$$

En utilisant les propriétés de l'opérateur $c(.,.)$ nous écrivons :

$$c_{\partial\Omega}(u, u) \leq \zeta^{\max} \|u\|_{L^2(\partial\Omega)}^2 \leq l_\gamma \zeta^{\max} \|u\|_{H^1(\Omega)}^2 \quad \text{et} \quad c_\Omega(u, u) \leq c^{\max} \|u\|_{L^2(\Omega)}^2 \leq c^{\max} \|u\|_{H^1(\Omega)}^2,$$

où nous avons utilisé le théorème de la trace $\exists l_\gamma > 0$ telle que $\|u\|_{L^2(\partial\Omega)} \leq \sqrt{l_\gamma} \|u\|_{H^1(\Omega)}$.

Rassemblons ces différentes inégalités en posant :

$$C = \max(\kappa^{\max}, \varepsilon \gamma' c^{\max}) + \varepsilon \gamma l_\gamma \zeta^{\max} > 0.$$

Donc

$$|b_\varepsilon(u, v)| \leq C \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}.$$

Ce qui prouve que l'opérateur auxiliaire $b_\varepsilon(\cdot, \cdot)$ est continu sur $H^1(\Omega)$.

Examinons à présent la coercivité. Avec nos hypothèses :

$$\begin{aligned} |b_\varepsilon(u, u)| &= |a_\Omega(u, u) + \varepsilon c(u, u)| = a_\Omega(u, u) + \varepsilon c(u, u) \\ &\geq \kappa^{\min} \|\nabla u\|_{L^2(\Omega)}^2 + \varepsilon(\gamma \zeta^{\min} \|u\|_{L^2(\partial\Omega)}^2 + \gamma' c^{\min} \|u\|_{L^2(\Omega)}^2) \\ &\geq \kappa^{\min} \|\nabla u\|_{L^2(\Omega)}^2 + \varepsilon \gamma \zeta^{\min} \|u\|_{L^2(\partial\Omega)}^2 \\ &\geq \min\left(\kappa^{\min}, \varepsilon \gamma \zeta^{\min}\right) \left(\|\nabla u\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\partial\Omega)}^2\right). \end{aligned}$$

Mais d'après l'inégalité de Poincaré (voir annexe B) :

$$\|u\|_{L^2(\Omega)}^2 \leq l_\square^2 \left(\|u\|_{L^2(\partial\Omega)} + \|\nabla u\|_{L^2(\Omega)}\right)^2 \leq 2l_\square^2 \left(\|u\|_{L^2(\partial\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2\right).$$

Soit en ajoutant aux deux membres le terme $\|\nabla u\|_{L^2(\Omega)}^2$ pour faire apparaître la norme H^1 :

$$\begin{aligned} \|u\|_{H^1(\Omega)}^2 = \|\nabla u\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\Omega)}^2 &\leq \|\nabla u\|_{L^2(\Omega)}^2 + 2l_\square^2 \left(\|u\|_{L^2(\partial\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2\right) \\ &\leq (2l_\square^2 + 1) \|\nabla u\|_{L^2(\Omega)}^2 + 2l_\square^2 \|u\|_{L^2(\partial\Omega)}^2 \\ &\leq (2l_\square^2 + 1) \left(\|\nabla u\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\partial\Omega)}^2\right). \end{aligned}$$

Il suffit alors de poser

$$C' = \frac{\min(\kappa^{\min}, \varepsilon \gamma \zeta^{\min})}{2l_\square^2 + 1}$$

pour obtenir la minoration :

$$|b_\varepsilon(u, u)| \geq C' \|u\|_{H^1(\Omega)}^2.$$

Ce qui prouve que la forme bilinéaire $b_\varepsilon(\cdot, \cdot)$ est coercive dans $H^1(\Omega)$ et qu'elle induit une norme équivalente à la norme usuelle dans $H^1(\Omega)$. ■

Notons qu'il n'est pas possible de se passer des hypothèses $\varepsilon > 0$ et $\gamma + \gamma' \neq 0$ dans cette démonstration. Ceci prouve à contrario que l'opérateur de la chaleur $a_\Omega(\cdot, \cdot)$ associé à une condition limite de Neumann ou de Steklov ne peut pas être coercif. Nous allons voir comment il est possible de contourner cette difficulté.

Définition A.1.2 Nous reprenons les notations et les hypothèses de la proposition A.1.2. Nous définissons deux problèmes aux valeurs propres :

$$\text{Trouver } (u, \mu) \in H^1(\Omega) \times \mathbb{R}^+ \text{ tels que } a_\Omega(u, v) = \mu c(u, v) \quad \forall v \in H^1(\Omega); \quad (3.A.6)$$

$$\text{Trouver } (u_\varepsilon, \mu(\varepsilon)) \in H^1(\Omega) \times \mathbb{R}^+ \text{ tels que } b_\varepsilon(u_\varepsilon, v) = \mu(\varepsilon) c(u_\varepsilon, v) \quad \forall v \in H^1(\Omega). \quad (3.A.7)$$

△

Proposition A.1.3 *L'opérateur de la chaleur $a_\Omega(\cdot, \cdot)$ pour le problème aux valeurs propres (3.A.6) (resp. $b_\varepsilon(\cdot, \cdot)$ pour le problème auxiliaire (3.A.7)) admet une suite $(\mu_k)_{k \in \mathbb{N}}$ infinie dénombrable de valeurs propres réelles positives (resp. $(\mu_k(\varepsilon))_{k \in \mathbb{N}}$ de valeurs propres réelles strictement positives). Les valeurs propres des deux opérateurs vérifient :*

$$\forall k \in \mathbb{N} \quad \mu_k = \mu_k(\varepsilon) - \varepsilon.$$

En particulier $\mu_1 = 0$ et $\mu_1(\varepsilon) = \varepsilon$ sont les premières valeurs propres. Elles sont simples et associées à un mode propre constant dans le domaine.

De plus les modes propres des deux opérateurs vérifient (après normalisation) :

$$\forall k \in \mathbb{N}; \forall(\mathbf{x}) \in \bar{\Omega}, \quad u_k(\mathbf{x}) = u_\varepsilon(\mathbf{x}),$$

$$\forall i, j \in \mathbb{N}, \quad c(u_i, u_j) = \delta_i^j, \quad a_\Omega(u_i, u_j) = \mu_i \delta_i^j, \quad b_\varepsilon(u_i, u_j) = (\mu_i + \varepsilon) \delta_i^j.$$

▲

Preuve A.1.3 .

L'opérateur $b_\varepsilon(\cdot, \cdot)$ est bilinéaire, symétrique, continu et coercif dans $H^1(\Omega)$.

De plus $H^1(\Omega) \subset L^2(\Omega)$ avec injection compacte³ (théorème de Rellich) et $H^1(\Omega)$ est dense dans $L^2(\Omega)$. On peut donc lui appliquer la théorie spectrale des opérateurs auto-adjoints. En particulier il admet un spectre infini dénombrable de valeurs propres réelles strictement positives. La première valeur propre est simple (opérateur "Laplacien"). Les modes propres sont orthogonaux par rapport au produit scalaire formé par $c(\cdot, \cdot)$ dans L^2 soit

$$\forall i, j \in \mathbb{N}, i \neq j, \quad c(u_i(\varepsilon), u_j(\varepsilon)) = 0.$$

Si de plus les modes sont normalisés par rapport à $c(\cdot, \cdot)$ (càd. $c(u_i(\varepsilon), u_i(\varepsilon)) = 1$ alors

$$\forall i, j \in \mathbb{N}, \quad b_\varepsilon(u_i(\varepsilon), u_j(\varepsilon)) = \mu_i(\varepsilon) \delta_i^j.$$

$b_\varepsilon(\cdot, \cdot)$ est un produit scalaire pour les éléments de $H^1(\Omega)$ et ses modes sont aussi orthogonaux pour ce produit scalaire.

L'équation (3.A.7) étant vérifiée pour tout $v \in H^1(\Omega)$ nous pouvons choisir pour v une solution u de l'équation (3.A.6). De même cette dernière équation est en particulier satisfaite si nous choisissons $v = u_\varepsilon$. On a donc les deux égalités :

$$\begin{aligned} a_\Omega(u_\varepsilon, u) + \varepsilon c(u_\varepsilon, u) &= \mu(\varepsilon) c(u_\varepsilon, u), \\ a_\Omega(u, u_\varepsilon) &= \mu c(u, u_\varepsilon). \end{aligned}$$

3. Le terme "injection compacte" signifie que l'opérateur \mathcal{I} qui à une fonction $u \in H^1(\Omega)$ associe la fonction $\mathcal{I}(u) = u \in L^2(\Omega)$ est continu et relativement compact. Les hypothèses sur les deux espaces reviennent à considérer que de toute suite bornée de $H^1(\Omega)$ on peut extraire une sous-suite convergente dans $L^2(\Omega)$.

Mais les formes bilinéaires sont symétriques. Nous obtenons donc par soustraction :

$$\varepsilon = \mu(\varepsilon) - \mu.$$

Ainsi un réel μ est valeur propre du problème (3.A.6) si et seulement si $\mu + \varepsilon$ est valeur propre du problème auxiliaire (3.A.7). Nous en déduisons que les valeurs propres du problème (3.A.6) sont donc exactement les réels :

$$\forall k \in \mathbb{N} \quad \mu_k = \mu_k(\varepsilon) - \varepsilon. \quad (3.A.8)$$

Le spectre de $a_\Omega(\cdot, \cdot)$ associé au problème (3.A.6) est infini dénombrable. Il est en quelque sorte "décalé vers la gauche" de $-\varepsilon$ par rapport au spectre de $b_\varepsilon(\cdot, \cdot)$ associé au problème (3.A.7). De plus $\mu_k(\varepsilon) > 0$ par coercivité de $b_\varepsilon(\cdot, \cdot)$. L'égalité (3.A.8) montre que $\mu_k > -\varepsilon$ et ceci pour tout $\varepsilon > 0$. En faisant tendre ε vers 0 ceci prouve que $\mu_k \geq 0$. Dans le cas présent la première valeur propre de $b_\varepsilon(\cdot, \cdot)$ est simple et sa valeur est $\mu_1(\varepsilon) = \varepsilon$. Ainsi la première valeur propre de (3.A.6) est aussi simple et $\mu_1 = 0$.

Les modes propres $\{u_{\varepsilon k}\}_{k \in \mathbb{N}}$ de l'opérateur $b_\varepsilon(\cdot, \cdot)$ forment une base de $H^1(\Omega)$. A ce titre un mode propre u_k solution de (3.A.6) est décomposable sur cette base :

$$u_k = \sum_{i=1}^{\infty} x_i u_{\varepsilon i}.$$

Donc :

$$a_\Omega(u_k, v) = \mu_k c(u_k, v) = (\mu_k(\varepsilon) - \varepsilon) c(u_k, v),$$

$$a_\Omega(u_k, v) + \varepsilon c(u_k, v) = \mu_k(\varepsilon) c(u_k, v),$$

$$b_\varepsilon(u_k, v) = \mu_k(\varepsilon) c(u_k, v), \quad \forall v \in H^1(\Omega),$$

L'égalité est donc vraie pour $v = u_{\varepsilon j} \forall j$,

Soit en remplaçant u_k par sa décomposition sur les $u_{\varepsilon j}$

$$b_\varepsilon\left(\sum_{i=1}^{\infty} x_i u_{\varepsilon i}, u_{\varepsilon j}\right) = \mu_k(\varepsilon) c\left(\sum_{i=1}^{\infty} x_i u_{\varepsilon i}, u_{\varepsilon j}\right) \text{ donc par linéarité}$$

$$\sum_{i=1}^{\infty} x_i b_\varepsilon(u_{\varepsilon i}, u_{\varepsilon j}) = \mu_k(\varepsilon) \sum_{i=1}^{\infty} x_i c(u_{\varepsilon i}, u_{\varepsilon j}).$$

Soit en utilisant l'orthonormalité des modes propres $u_{\varepsilon i}$

$$\sum_{i=1}^{\infty} x_i \mu_j(\varepsilon) \delta_i^j = \mu_k(\varepsilon) \sum_{i=1}^{\infty} x_i \delta_i^j. \text{ Autrement dit } x_j (\mu_j(\varepsilon) - \mu_k(\varepsilon)) = 0$$

Donc si $\mu_j(\varepsilon) \neq \mu_k(\varepsilon)$ alors $x_j = 0$.

Nous en déduisons $u_k = x_k u_{\varepsilon k}$. Nous pouvons normaliser les modes u_k pour avoir $c(u_k, u_k) = 1$. Finalement les deux familles de modes propres sont identiques : $\forall k \in \mathbb{N}^* \quad u_k = u_{\varepsilon k}$.



Nous pouvons choisir pour $c(.,.)$ l'un des opérateurs qui suivent :

$$- c(u, v) = \int_{\Omega} c u v = c_{\Omega}(u, v) \quad \text{avec} \quad \gamma' = 1 \quad ; \quad \gamma = 0.$$

Ce choix correspond à l'opérateur auxiliaire pour le problème de Neumann. C'est aussi l'opérateur auxiliaire du problème de Dirichlet si nous travaillons dans $H_0^1(\Omega)$.

$$- c(u, v) = \int_{\partial\Omega} \zeta u v = c_{\partial\Omega}(u, v) \quad \text{avec} \quad \gamma' = 0 \quad ; \quad \gamma = 1.$$

Ce choix correspond à l'opérateur auxiliaire du problème de Steklov. Attention toutefois ici c'est l'injection de $H^{1/2}(\partial\Omega)$ dans $L^2(\partial\Omega)$ qui est continue et compacte et $H^{1/2}(\partial\Omega)$ est dense dans $L^2(\partial\Omega)$. Par conséquent les modes de Steklov forment bien une base hilbertienne de $H^{1/2}(\partial\Omega)$ mais les relèvements harmoniques de ces modes ne forment qu'une base d'un sous espace $E(\Omega) \subsetneq H^1(\Omega)$ (inclusion stricte, voir définition A.4.1). Pour obtenir une base de $H^1(\Omega)$ il faudra lui adjoindre la base des modes de Dirichlet de $H_0^1(\Omega)$.

$$- c(u, v) = c_{\bar{\Omega}}(u, v) = c_{\Omega}(u, v) + c_{\partial\Omega}(u, v) \quad \text{avec} \quad \gamma' = 1 \quad ; \quad \gamma = 1.$$

Ce choix correspond à l'opérateur auxiliaire du problème de Branche. L'injection de $H_0^1(\Omega) \times H^{1/2}(\partial\Omega)$ dans $L^2(\Omega) \times L^2(\partial\Omega)$ est compacte et $H_0^1(\Omega) \times H^{1/2}(\partial\Omega)$ est dense dans $L^2(\Omega) \times L^2(\partial\Omega)$.

$$- c(u, v) = \int_{\partial\Omega} \zeta u v \quad \text{avec} \quad \gamma' = 0 \quad ; \quad \gamma = 1 \quad \zeta = \text{Bi}/\varepsilon \quad 0 < \text{Bi}^{\min} \leq \text{Bi} \leq \text{Bi}^{\max}.$$

On retrouve dans ce cas l'opérateur de la chaleur avec des conditions limites de Fourier

$$a_{\Omega}^{\text{Bi}}(u, v) = \int_{\Omega} \nabla u \cdot \kappa \cdot \nabla v + \int_{\partial\Omega} \text{Bi} u v.$$

$$- c(u, v) = \int_{\partial\Omega} \zeta u v + c_{\Omega}(u, v) \quad \text{avec} \quad \gamma' = 1 \quad ; \quad \gamma = 1 \quad ; \quad \zeta = \text{Bi}/\varepsilon.$$

Nous retrouvons cette fois l'opérateur de la chaleur auxiliaire pour le problème aux valeurs propres avec des conditions limites de Fourier $a_{\Omega}^{\text{Bi}}(u, v) + \varepsilon c_{\Omega}(u, v)$.

A.2 Propriété de la base de Dirichlet.

A.2.1 Base de Dirichlet.

Proposition A.2.1 *Les modes de Dirichlet $\{V_i^D\}_{i \in \mathbb{N}}$ forment une base infinie mais dénombrable de l'espace de Hilbert $H_0^1(\Omega)$.*

▲

Preuve A.2.1 *C'est un résultat classique d'analyse fonctionnelle [90],[91], [92]. Il a été justifié dans la proposition A.1.3. Rappelons que l'opérateur $a_\Omega(.,.)$ est coercif dans $H_0^1(\Omega)$.*

■

Remarque A.2.1 *Les valeurs propres sont strictement positives et la première valeur propre est simple. Elle est associée à un mode qui ne change pas de signe dans le domaine. Les valeurs propres tendent vers l'infini. Une valeur propre multiple engendre un sous espace propre de dimension égale à l'ordre de multiplicité.*

A.2.2 Orthogonalité des modes de Dirichlet.

Définition A.2.1 *Soient $u, v \in H^1(\Omega)$. On donne la forme adimensionnée du produit scalaire défini en (2.5.8) $(u|v)_{H(\Omega)}$:*

$$\begin{aligned} (u|v)_{H(\Omega)} &= \int_{\Omega} \nabla u \cdot \kappa \cdot \nabla v + \mu_0 \int_{\partial\Omega} \zeta(x) uv \\ &= a_\Omega(u, v) + \mu_0 c_{\partial\Omega}(u, v), \end{aligned} \quad (3.A.9)$$

avec $\mu_0 > 0$ un facteur de pondération sans dimension. La norme induite est définie par

$$\forall u \in H^1(\Omega), \quad \|u\|_{H(\Omega)} = \sqrt{\int_{\Omega} \nabla u \cdot \kappa \cdot \nabla u + \mu_0 \int_{\partial\Omega} \zeta(x) (u)^2}. \quad (3.A.10)$$

Proposition A.2.2 *Si $\|V_i^D\|_{H(\Omega)} = 1$ alors les modes de Dirichlet vérifient les relations d'orthogonalité :*

$$\begin{aligned} \forall i, j \in \mathbb{N}, \quad (V_i^D | V_j^D)_{H(\Omega)} &= \delta_i^j, \\ a_\Omega(V_i^D, V_j^D) &= \delta_i^j, \quad c_\Omega(V_i^D, V_j^D) = \frac{\delta_i^j}{\mu_i^D}. \end{aligned} \quad (3.A.11)$$

▲

Bien entendu on peut donner des expressions dimensionnées équivalentes⁴.

Preuve A.2.2 *Considérons deux modes propres V_i^D et V_j^D . Ils vérifient d'après (2.5.2) :*

$$\forall u \in H_0^1(\Omega), \quad a_\Omega(V_i^D, u) = \mu_i^D c_\Omega(V_i^D, u), \quad a_\Omega(V_j^D, u) = \mu_j^D c_\Omega(V_j^D, u).$$

Choisissons tour à tour $u = V_j^D$ et $u = V_i^D$, soit :

$$a_\Omega(V_i^D, V_j^D) = \mu_i^D c_\Omega(V_i^D, V_j^D), \quad a_\Omega(V_j^D, V_i^D) = \mu_j^D c_\Omega(V_j^D, V_i^D).$$

Les formes bilinéaires sont symétriques. Nous trouvons par différence :

$$i \neq j \quad (\mu_i^D - \mu_j^D) c_\Omega(V_i^D, V_j^D) = 0 \Rightarrow c_\Omega(V_i^D, V_j^D) = 0,$$

d'où les relations

$$\forall i, j \in \mathbb{N}; i \neq j, \quad a_\Omega(V_i^D, V_j^D) = 0,$$

$$\forall i \neq j, \quad (V_i^D | V_j^D)_{H(\Omega)} = a_\Omega(V_i^D, V_j^D) + \mu_0 c_{\partial\Omega}(V_i^D, V_j^D) = a_\Omega(V_i^D, V_j^D) = 0.$$

Si $i = j$ alors : $a_\Omega(V_i^D, V_i^D) = \mu_i^D c_\Omega(V_i^D, V_i^D)$. Si de plus nous imposons $\|V_i^D\|_{H^1(\Omega)} = 1$, en remarquant que la trace d'un mode de Dirichlet est nulle sur le bord du domaine, nous en déduisons :

$$(V_i^D | V_i^D)_{H(\Omega)} = \|V_i^D\|_{H(\Omega)}^2 = a_\Omega(V_i^D, V_i^D) + \mu_0 c_{\partial\Omega}(V_i^D, V_i^D) = a_\Omega(V_i^D, V_i^D) = 1,$$

donc

$$c_\Omega(V_i^D, V_i^D) = \frac{1}{\mu_i^D}.$$

Cette dernière expression est bien définie puisque la valeur propre de Dirichlet est strictement positive.

■

A.3 Propriété de la base de Steklov.

A.3.1 Définition de la Base de Steklov [1] [2] [3] [4]

La base de Steklov est constituée d'un ensemble de fonctions qui sont définies par leur trace sur $\partial\Omega$. On peut associer de manière univoque à chaque mode de Steklov son relèvement

4. $(V_i^D | V_j^D)_{H(\mathcal{D})} = K_o L_o \delta_i^j, \quad \mathcal{A}_{\mathcal{D}}(V_i^D, V_j^D) = K_o L_o \delta_i^j, \quad c_{\mathcal{D}}(V_i^D, V_j^D) = K_o L_o \frac{\delta_i^j}{\lambda_i^D}.$

harmonique dans le domaine Ω . Par abus de langage nous désignerons aussi par "mode de Steklov" le relèvement harmonique. Les fonctions de la base de Steklov vérifient elles aussi un problème aux valeurs propres. Avant de le décrire donnons une définition. .

Définition A.3.1 Une fonction $u \in H^1(\Omega)$ est dite harmonique dans le domaine Ω si en tout point du domaine on a :

$$\Omega \quad \nabla \cdot (\kappa \cdot \nabla u) = 0.$$

△

Si on se donne une fonction $f \in L^2(\partial\Omega)$, alors son relèvement harmonique sur Ω noté f^Δ est la fonction de $H^1(\Omega)$ dont la trace sur $\partial\Omega$ est égale à la fonction imposée $f^\Delta|_{\partial\Omega} = f$. La base de Steklov que nous allons introduire engendre l'espace des fonctions harmoniques dont la trace ne s'annule pas sur le bord $\partial\Omega$ du domaine. Il est noté $H^{1/2}(\partial\Omega)$. C'est un sous espace dense de $L^2(\partial\Omega)$. Les notions de relèvement harmonique et de trace et leurs propriétés sont décrites dans la section B.5. Nous pouvons à présent introduire le problème aux valeurs propres de Steklov :

Définition A.3.2 Le problème variationnel aux valeurs propres de Steklov est défini par :

$$\begin{aligned} \text{Trouver } (\mu^S, V^S) \in \mathbb{R}^+ \times H^1(\Omega) \text{ tels que quelque soit } v \in H^1(\Omega) \\ a_\Omega(V^S, v) = \mu^S c_{\partial\Omega}(V^S, v). \end{aligned} \tag{3.A.12}$$

△

On reconnaît dans le terme de droite de cette formulation la forme bilinéaire symétrique positive notée $c_{\partial\Omega}(.,.)$

$$c_{\partial\Omega}(u, v) = \int_{\partial\Omega} \zeta u|_{\partial\Omega} v|_{\partial\Omega}. \tag{3.A.13}$$

Cette forme bilinéaire est un produit scalaire de $L^2(\partial\Omega)$. Ce produit scalaire induit une norme équivalente à la norme usuelle dans $L^2(\partial\Omega)$ tel que :

$$\forall w \in L^2(\partial\Omega), \quad c_{\partial\Omega}(w, w) \equiv \|w\|_{L^2(\partial\Omega)}^2 = \int_{\partial\Omega} w^2. \tag{3.A.14}$$

Proposition A.3.1 Les modes de Steklov $\{S_i\}_{i \in \mathbb{N}}$ ($S_i = V_i^S|_{\partial\Omega}$) forment une base infinie mais dénombrable de $H^{1/2}(\partial\Omega)$.

▲

Preuve A.3.1 Ce résultat est une conséquence de la proposition A.1.3. Rappelons que l'opérateur $a_\Omega(.,.)$ n'est pas coercif, mais que nous avons utilisé pour établir la propriété un opérateur auxiliaire $b_\varepsilon(.,.)$ bilinéaire, symétrique, continu et coercif. L'opérateur $b_\varepsilon(.,.)$ est caractérisé par un paramètre numérique ε de valeur réelle strictement positive (mais arbitraire à part cela). L'opérateur auxiliaire possède les mêmes modes propres que ceux de l'opérateur $a_\Omega(.,.)$ et ses valeurs propres μ_ε sont décalées vers la droite par rapport à celles de $a_\Omega(.,.)$ notées μ (concrètement on a $\mu_k = \mu_\varepsilon - \varepsilon$).

■

Remarque A.3.1 La fonction $V_1^S = 1$ constante dans le domaine Ω est une solution du problème aux valeurs propres de Steklov (3.A.12). Ce mode "plat" est associé à la valeur propre nulle $\mu_1^S = 0$. Ici le mode "plat" est normé pour le problème adimensionné. Ce mode joue un rôle particulier dans les modélisations de la température. Il permet en particulier d'ajuster la température moyenne du système. Nous lui réserverons un traitement particulier dans le processus de réduction modale par amalgame.

A.3.2 Orthogonalité

Proposition A.3.2 Si $\|V_i^S\|_{H(\Omega)} = 1$ alors les modes de Steklov vérifient les relations d'orthogonalité :

$$\begin{aligned} \forall i, j \in \mathbb{N}, \quad (V_i^S | V_j^S)_{H(\Omega)} &= \delta_i^j, \\ a_\Omega(V_i^S, V_j^S) &= \delta_i^j \frac{\mu_i^S}{\mu_o + \mu_i^S}, \quad c_{\partial\Omega}(V_i^S, V_j^S) = \delta_i^j \frac{1}{\mu_o + \mu_i^S}. \end{aligned} \quad (3.A.15)$$

▲

Preuve A.3.2 Considérons deux modes propres V_i^S et V_j^S . Ils vérifient d'après (3.A.12) :

$$\forall u \in H^1(\Omega), \quad a_\Omega(V_i^S, u) = \mu_i^S c_{\partial\Omega}(V_i^S, u), \quad a_\Omega(V_j^S, u) = \mu_j^S c_{\partial\Omega}(V_j^S, u).$$

Choisissons tour à tour $u = V_j^S$ et $u = V_i^S$, soit :

$$a_\Omega(V_i^S, V_j^S) = \mu_i^S c_{\partial\Omega}(V_i^S, V_j^S), \quad a_\Omega(V_j^S, V_i^S) = \mu_j^S c_{\partial\Omega}(V_j^S, V_i^S).$$

Les formes bilinéaires sont symétriques. Nous trouvons par différence :

$$i \neq j \quad (\mu_i^S - \mu_j^S) c_{\partial\Omega}(V_i^S, V_j^S) = 0 \Rightarrow c_{\partial\Omega}(V_i^S, V_j^S) = 0,$$

d'où les relations

$$\forall i, j \in \mathbb{N}; i \neq j, \quad a_{\Omega}(V_i^S, V_j^S) = 0,$$

$$(V_i^S | V_j^S)_{H(\Omega)} = a_{\Omega}(V_i^S, V_j^S) + \mu_o c_{\partial\Omega}(V_i^S, V_j^S) = 0.$$

Si $i = j$ alors : $a_{\Omega}(V_i^S, V_i^S) = \mu_i^S c_{\partial\Omega}(V_i^S, V_i^S)$. Si de plus nous imposons $\|V_i^S\|_{H(\Omega)} = 1$ nous en déduisons :

$$(V_i^S | V_i^S)_{H(\Omega)} = \|V_i^S\|_{H(\Omega)}^2 = a_{\Omega}(V_i^S, V_i^S) + \mu_o c_{\partial\Omega}(V_i^S, V_i^S) = (\mu_o + \mu_i^S) c_{\partial\Omega}(V_i^S, V_i^S) = 1,$$

donc

$$c_{\partial\Omega}(V_i^S, V_i^S) = \frac{1}{\mu_o + \mu_i^S} \quad ; \quad a_{\Omega}(V_i^S, V_i^S) = \frac{\mu_i^S}{\mu_o + \mu_i^S}.$$

Ces expressions sont bien définies pour $\mu_i^S \geq 0$. La valeur propre nulle correspond au mode de Steklov constant dans le domaine. ■

A.4 Propriétés de la base de Dirichlet-Steklov.

A.4.1 Orthogonalité entre les modes de Dirichlet et de Steklov.

Proposition A.4.1 Les modes de Dirichlet $(V_i^D)_{i \in \mathbb{N}}$ et de Steklov $(V_j^S)_{j \in \mathbb{N}}$ vérifient les deux relations d'orthogonalité :

$$\forall i, j \in \mathbb{N}, \quad (V_i^D | V_j^S)_{H(\Omega)} = 0, \tag{3.A.16}$$

$$a_{\Omega}(V_i^D, V_j^S) = 0.$$

▲

Notons en revanche qu'il n'y a pas d'orthogonalité L^2 entre ces deux familles, que l'on considère Ω (ce qui est vrai pour le cas des modes de Dirichlet seuls) ou bien encore $\bar{\Omega}$ (comme ce serait le cas pour les modes de Branche). Il existe une orthogonalité pour le produit scalaire sur $\partial\Omega$ puisque les traces des modes de Dirichlet sont nulles. Cette orthogonalité n'a que très peu d'intérêt.

Preuve A.4.1 Elle est évidente si nous choisissons comme fonction d'essai v dans la formulation variationnelle (3.A.12) un mode de Dirichlet $V_i^D \in H_0^1(\Omega) \subset H^1(\Omega)$ dont la trace sur $\partial\Omega$ est nulle. donc

$$\forall u \in H^1(\Omega), \quad (V_j^S | u)_{H(\Omega)} = a_\Omega(V_j^S, u) + \mu_o c_{\partial\Omega}(V_j^S, u) = (\mu_j^S + \mu_o) c_{\partial\Omega}(V_j^S, u).$$

Nous remarquons de manière évidente l'égalité :

$$\forall i, j \in \mathbb{N}, \quad c_{\partial\Omega}(V_i^D, V_j^S) = \int_{\partial\Omega} \zeta V_i^D |_{\partial\Omega} S_j = 0.$$

Par conséquent en substituant V_i^D à u ⁵ :

$$\forall i, j \in \mathbb{N} \left| \begin{array}{l} (V_j^S | V_i^D)_{H(\Omega)} = (\mu_j^S + \mu_o) c_{\partial\Omega}(V_j^S, V_i^D) = 0, \\ a_\Omega(V_j^S, V_i^D) = 0. \end{array} \right.$$

D'où le résultat annoncé pour les deux relations d'orthogonalité. En associant cette propriété avec les deux relations d'orthogonalité (3.A.11) et (3.A.15), vis-à-vis du produit scalaire $(\cdot | \cdot)_{H(\Omega)}$ dans $H^1(\Omega)$, des modes de Dirichlet entre eux d'une part et des modes de Steklov entre eux d'autre part, on peut conclure que la réunion des bases propres de Dirichlet et de Steklov forme bien une base orthonormée de l'espace $H(\Omega) = H_0^1(\Omega) \times H^{1/2}(\partial\Omega)$.

Notons au passage que pour $j \in \mathbb{N}$:

$$\int_{\Omega} (\nabla \cdot \kappa \cdot \nabla V_i^D) V_j^S = - \int_{\Omega} \nabla V_j^S \cdot \kappa \cdot \nabla V_i^D + \int_{\partial\Omega} (\kappa \cdot \nabla V_i^D \cdot n) S_j.$$

Utilisons le problème aux valeurs propres de Dirichlet sous sa forme locale adimensionnée issue de (2.5.1) pour transformer le membre de gauche. Le premier terme du membre de droite s'annule. Nous utilisons enfin le problème aux valeurs propres de Steklov sous sa forme locale adimensionnée issue de (2.5.4) pour transformer le second terme du membre de droite, ce qui donne finalement la relation :

$$\int_{\Omega} c V_i^D V_j^S = \frac{1}{\mu_i^D \mu_j^S} \int_{\partial\Omega} \zeta (\kappa \cdot \nabla V_i^D \cdot n) (\kappa \cdot \nabla V_j^S \cdot n) \neq 0.$$

■

Finalement si nous choisissons de normaliser les modes propres à l'aide de la norme $\| \cdot \|_{H(\Omega)}$ induite par le produit scalaire défini par (3.A.9). Nous pouvons résumer toutes ces relations par :

5. On prendra garde à ce que $a_\Omega(V_i^D, V_j^S) \neq \mu_i^D c_\Omega(V_i^D, V_j^S)$ car $V_j^S \notin H_0^1(\Omega)$.

Proposition A.4.2 *Les bases de Dirichlet et de Steklov sont orthogonales vis à vis du produit scalaire défini par (3.A.9). On peut normaliser ces bases à l'aide de la norme induite (3.A.10).*

$$\forall i, j \in \mathbb{N}, \forall \mathcal{X}, \mathcal{Y} \in \{D, S\} \left| \begin{array}{l} (V_i^{\mathcal{X}} | V_j^{\mathcal{Y}})_{H(\Omega)} = \delta_i^j \delta_{\mathcal{X}}^{\mathcal{Y}}, \\ \|V_i^{\mathcal{X}}\|_{H(\Omega)} = 1. \end{array} \right. \quad (3.A.17)$$

▲

Preuve A.4.2 *Ce résultat est la conséquence directe des propositions (3.A.11) et (3.A.15). ■*

La relation d'orthogonalité définie par (3.A.17) va permettre d'affirmer que la réunion des modes propres de Dirichlet et de Steklov $(\{V_i^D\}_{i \in \mathbb{N}}, \{V_j^S\}_{j \in \mathbb{N}})$ forme une base hilbertienne de $H^1(\Omega) = H_0^1(\Omega) \oplus E(\Omega)$ ⁶. Grâce à cela nous pourrons facilement projeter un champ de température sur cette base pour en trouver les coefficients de décomposition. Dans cette décomposition nous ne garderons que les termes principaux ou dominants dans un sens qu'il faudra préciser. Il sera ainsi possible d'établir une démarche complète de réduction d'un problème thermique.

A.4.2 Décomposition d'une fonction de $H^1(\Omega)$.

A.4.2.1 Espace de relèvement harmonique.

Considérons une fonction $u \in H^1(\Omega)$. Le théorème de la trace assure que la trace de toute fonction u sur le bord du domaine Ω (notée $u|_{\partial\Omega} = \gamma_o(u)$) existe et est unique. La trace $u|_{\partial\Omega}$ appartient à l'image $H^{1/2}(\partial\Omega)$ de l'application trace qui est un espace de Hilbert.

Définition A.4.1 *L'espace des relèvements harmoniques $E(\Omega) \subset H^1(\Omega)$ est défini par :*

$$E(\Omega) = \{\hat{u} \in H^1(\Omega) \text{ t.q. } \exists u \in H^1(\Omega), \hat{u}|_{\partial\Omega} = u|_{\partial\Omega} \text{ sur } \partial\Omega \text{ et } \hat{u} = \mathcal{R}(u|_{\partial\Omega}) \text{ sur } \Omega\}, \quad (3.A.18)$$

où \mathcal{R} désigne le relèvement harmonique dans le domaine d'une fonction définie sur son bord.

△

Proposition A.4.3 *Considérons les fonctions $u \in H^1(\Omega)$. A chaque fonction u on peut associer un unique couple de fonctions $(\overset{\circ}{u}, \hat{u}) \in H_0^1(\Omega) \times E(\Omega)$ tel que :*

$$\overset{\circ}{u} = u - \hat{u}. \quad (3.A.19)$$

6. $E(\Omega) \underset{(\cdot, \cdot)_{H(\Omega)}}{\perp} H_0^1(\Omega)$ et $(H_0^1(\Omega))^\perp = E(\Omega)$

▲

Preuve A.4.3 Puisque u est une fonction de $H^1(\Omega)$ sa trace $u|_{\partial\Omega}$ est bien définie. C'est un élément de $H^{1/2}(\partial\Omega)$. Le théorème de Lax-Milgram assure que le relèvement harmonique \hat{u} de $u|_{\partial\Omega}$ existe et est unique ($\hat{u} \in E(\Omega)$). Formons la différence $\hat{u} = u - \hat{u}$. La fonction \hat{u} est une fonction de $H^1(\Omega)$ (comme combinaison linéaire de deux fonctions de $H^1(\Omega)$). Elle est unique. Nous pouvons calculer sa trace qui par construction vérifie :

$$\hat{u}|_{\partial\Omega} = \gamma_0(u - \hat{u}) = u|_{\partial\Omega} - \hat{u}|_{\partial\Omega} = 0.$$

Autrement dit $\hat{u} \in H_0^1(\Omega)$.

■

Proposition A.4.4 Les relèvements harmoniques des modes de Steklov forment une base hilbertienne de $E(\Omega)$ pour le produit scalaire $(\cdot|\cdot)_{H(\Omega)}$.

▲

Preuve A.4.4 Les modes de Steklov $\{S_j\}_{j \in \mathbb{N}}$ forment une base de l'espace des traces $H^{1/2}(\partial\Omega)$. À chaque mode de Steklov est associée une fonction V_j^S de $H^1(\Omega)$ qui est, d'après la forme locale adimensionnée issue de (2.5.4), le relèvement harmonique du mode de Steklov :

$$V_j^S|_{\partial\Omega} = S_j, \quad V_j^S = \mathcal{R}(S_j).$$

Autrement dit $V_j^S \in E(\Omega)$.

Soit $w \in E(\Omega)$. Sa trace est décomposable sur la base de Steklov :

$$\exists \alpha_j \in \mathbb{R} \text{ telles que } w|_{\partial\Omega} = \sum_{j=1}^{\infty} \alpha_j S_j.$$

Soit en prenant le relèvement de cette expression :

$$w = \mathcal{R}(w|_{\partial\Omega}) = \mathcal{R}\left(\sum_{j=1}^{\infty} \alpha_j S_j\right) = \sum_{j=1}^{\infty} \alpha_j \mathcal{R}(S_j) = \sum_{j=1}^{\infty} \alpha_j V_j^S.$$

Utilisons la relation d'orthogonalité (3.A.15) $((V_i^S|V_j^S)_{H(\Omega)} = \delta_i^j)$ nous obtenons l'expression des coefficients de décomposition α :

$$(w|V_i^S)_{H(\Omega)} = \sum_{j=1}^{\infty} \alpha_j (V_j^S|V_i^S)_{H(\Omega)} = \sum_{j=1}^{\infty} \alpha_j \delta_i^j = \alpha_i.$$

Cette décomposition est unique.

■

A.4.2.2 Base de Dirichlet-Steklov.

Proposition A.4.5 *La réunion des bases de Dirichlet et de Steklov $(\{V_i^D\}, \{S_j\})_{i,j \in \mathbb{N}}$ (resp. $\{V_i^D\}_{i \in \mathbb{N}} \oplus \{V_j^S\}_{j \in \mathbb{N}}$) forme une base hilbertienne de $H_0^1(\Omega) \times H^{1/2}(\partial\Omega)$ (resp. $H_0^1(\Omega) \oplus \mathbf{E}(\Omega) = H^1(\Omega)$).*

En particulier :

$$\exists x_i^D, x_j^S \text{ telles que } u = \sum_{\mathcal{X} \in \{D,S\}} \sum_{i=1}^{\infty} x_i^{\mathcal{X}} V_i^{\mathcal{X}} \quad \forall u \in H^1(\Omega), \quad (3.A.20)$$

avec $x_i^{\mathcal{X}} = (u|V_i^{\mathcal{X}})_{H(\Omega)}$

▲

Preuve A.4.5 *D'après la proposition A.4.3 toute fonction u de $H^1(\Omega)$ se décompose de manière unique comme la somme d'une fonction $\overset{\circ}{u}$ de l'ensemble des fonctions de trace nulle $H_0^1(\Omega)$ et d'une fonction $\overset{\Delta}{u}$ de l'espace des relèvements harmoniques $\mathbf{E}(\Omega)$.*

Mais toute fonction $\overset{\circ}{u}$ de $H_0^1(\Omega)$ est décomposable sur la base des modes de Dirichlet $\{V_i^D\}_{i \in \mathbb{N}}$. De plus nous avons montré à la proposition A.4.4 que toute fonction $\overset{\Delta}{u}$ de $\mathbf{E}(\Omega)$ est décomposable sur la base des relèvements harmoniques des modes de Steklov $\{V_j^S\}_{j \in \mathbb{N}}$. Bien entendu la trace de $\overset{\Delta}{u}$ est décomposable sur la base de Steklov de $H^{1/2}(\partial\Omega)$. Enfin les coefficients de décomposition sont déterminés en faisant le produit scalaire de la décomposition avec un des modes propres et en utilisant la propriété d'orthogonalité par rapport au produit scalaire $(\cdot|\cdot)_{H(\Omega)}$ en supposant que les modes ont été normalisés par rapport à la norme induite par ce produit scalaire (base hilbertienne).

■

A.4.3 Équivalence entre la norme $\|u\|_{H(\Omega)}$ et la norme usuelle dans $H^1(\Omega)$.

Nous avons introduit par (3.A.10) la norme $\|u\|_{H(\Omega)}$. Pour rappel :

$$\|u\|_{H(\Omega)} = \sqrt{\int_{\Omega} \nabla u \cdot \kappa \cdot \nabla u + \mu_o \int_{\partial\Omega} \zeta u^2}.$$

Si les fonctions u sont exprimées en Kelvin il en est de même de cette norme (les autres paramètres sont sans dimensions ici).

Proposition A.4.6 *La norme $\|u\|_{H(\Omega)}$ est équivalente à la norme usuelle dans $H^1(\Omega)$.*



Preuve A.4.6 Rappelons les deux inégalités suivantes :

Inégalité de Poincaré :

$$\forall u \in H^1(\Omega), \exists l_{\square} > 0 \text{ telle que } \|u\|_{L^2(\Omega)} \leq l_{\square} \left(\|\nabla u\|_{L^2(\Omega)} + \|u|_{\partial\Omega}\|_{L^2(\partial\Omega)} \right). \quad (3.A.21)$$

Cette inégalité généralise l'inégalité classique dans $H_0^1(\Omega)$.

En remarquant que $\forall a, b, (a + b)^2 \leq 2(a^2 + b^2)$ nous pouvons aussi écrire :

$$\|u\|_{L^2(\Omega)}^2 \leq 2l_{\square}^2 \left(\|\nabla u\|_{L^2(\Omega)}^2 + \|u|_{\partial\Omega}\|_{L^2(\partial\Omega)}^2 \right). \quad (3.A.22)$$

Inégalité de Trace :

$$\forall u \in H^1(\Omega), \exists l_{\gamma} > 0 \text{ telle que } \|u|_{\partial\Omega}\|_{L^2(\partial\Omega)} \leq \sqrt{l_{\gamma}} \|u\|_{H^1(\Omega)}, \quad (3.A.23)$$

où la norme usuelle dans $H^1(\Omega)$ est définie par

$$\|u\|_{H^1(\Omega)}^2 = \|\nabla u\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\Omega)}^2. \quad (3.A.24)$$

Commençons par rechercher une majoration. Le tenseur de conductivité est défini positif.

Il existe donc deux constantes réelles strictement positives et bornées, telles que :

$$\kappa^{\min} = \frac{K^{\min}}{K_o} \leq \kappa_{i,j} \leq \kappa^{\max} = \frac{K^{\max}}{K_o}, \quad \kappa^{\min} \|\nabla u\|_{L^2(\Omega)}^2 \leq a_{\Omega}(u, u) \leq \kappa^{\max} \|\nabla u\|_{L^2(\Omega)}^2.$$

De même il existe deux constantes réelles strictement positives et bornées, telles que :

$$\zeta^{\min} \leq \zeta \leq \zeta^{\max}, \quad \zeta^{\min} \|u\|_{L^2(\partial\Omega)}^2 \leq c_{\partial\Omega}(u, u) \leq \zeta^{\max} \|u\|_{L^2(\partial\Omega)}^2.$$

Nous en déduisons l'inégalité

$$\|u\|_{H(\Omega)}^2 = a_{\Omega}(u, u) + \mu_o c_{\partial\Omega}(u, u) \leq \kappa^{\max} \|\nabla u\|_{L^2(\Omega)}^2 + \mu_o \zeta^{\max} \|u|_{\partial\Omega}\|_{L^2(\partial\Omega)}^2.$$

Soit en utilisant l'inégalité de trace (3.A.23) et la définition (3.A.24) de la norme usuelle dans $H^1(\Omega)$

$$\|u\|_{H(\Omega)}^2 \leq \kappa^{\max} \|\nabla u\|_{L^2(\Omega)}^2 + \mu_o \zeta^{\max} l_{\gamma} \|u\|_{H^1(\Omega)}^2 = (\kappa^{\max} + \mu_o \zeta^{\max} l_{\gamma}) \|\nabla u\|_{L^2(\Omega)}^2 + \mu_o \zeta^{\max} l_{\gamma} \|u\|_{L^2(\Omega)}^2.$$

Donc

$$\|u\|_{H(\Omega)}^2 \leq (\kappa^{\max} + \mu_o \zeta^{\max} l_{\gamma}) \left(\|\nabla u\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\Omega)}^2 \right).$$

Soit en posant $C_1 = \sqrt{\kappa^{\max} + \mu_o \zeta^{\max} l_\gamma}$, la majoration

$$\|u\|_{H(\Omega)} \leq C_1 \|u\|_{H^1(\Omega)} .$$

Recherchons à présent une minoration. A nouveau en utilisant le caractère défini positif de la matrice de conductivité nous écrivons

$$\|u\|_{H(\Omega)}^2 \geq \min(\kappa, \mu_o \zeta) \left(\|\nabla u\|_{L^2(\Omega)}^2 + \|u|_{\partial\Omega}\|_{L^2(\partial\Omega)}^2 \right) .$$

Ajoutons à chaque membre de (3.A.22) (déduite de l'inégalité de Poincaré) le carré de la norme du gradient. Nous trouvons

$$\|u\|_{H^1(\Omega)}^2 \leq (2l_\square^2 + 1) \|\nabla u\|_{L^2(\Omega)}^2 + 2l_\square^2 \|u|_{\partial\Omega}\|_{L^2(\partial\Omega)}^2 \leq (2l_\square^2 + 1) \left(\|\nabla u\|_{L^2(\Omega)}^2 + \|u|_{\partial\Omega}\|_{L^2(\partial\Omega)}^2 \right) .$$

Soit en posant $C_2 = \sqrt{\min(\kappa, \mu_o \zeta) / (2l_\square^2 + 1)}$, la minoration

$$\|u\|_{H(\Omega)} \geq C_2 \|u\|_{H^1(\Omega)} .$$

Nous avons donc montré

$$\forall u \in H^1(\Omega) , \exists C_1, C_2 \in \mathbb{R}^+ \setminus \{0\} \text{ telles que } C_2 \|u\|_{H^1(\Omega)} \leq \|u\|_{H(\Omega)} \leq C_1 \|u\|_{H^1(\Omega)} .$$

Ce qui prouve que $\|u\|_{H(\Omega)} \equiv \|u\|_{H^1(\Omega)}$.

■

A.4.4 Complétude de la base de Dirichlet-Steklov

Pour démontrer la complétude de la base de Dirichlet-Steklov, on aura besoin d'établir quelques résultats préliminaires sur les espaces fonctionnels impliqués.

On considère un domaine borné de \mathbb{R}^d qu'on suppose suffisamment régulier (pour que la formule de Stokes ait un sens dans $H^1(\Omega)$ et pour que les fonctions régulières soient denses dans $H^1(\Omega)$). On note γ_o l'opérateur trace défini de $H^1(\Omega)$ vers $L^2(\partial\Omega)$, $H_0^1(\Omega)$ le noyau de γ_o et $H^{1/2}(\partial\Omega)$ l'image de γ_o . On norme $H^1(\Omega)$ par

$$\|u\|_{H^1(\Omega)} = \sqrt{\|u\|_{L^2(\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2}, \quad \|u\|_{L^2(\Omega)}^p = \left(\int_{\Omega} |u(x)|^p \right)^{1/p}, \quad (3.A.25)$$

et $H^{1/2}(\partial\Omega)$ par

$$\|g\|_{H^{1/2}(\partial\Omega)} = \inf_{\gamma_o(u)=g} \|u\|_{H^1(\Omega)} . \quad (3.A.26)$$

γ_o est continue de $H^1(\Omega)$ dans $L^2(\partial\Omega)$ (Th. de la trace)

$$\|\gamma_o(u)\|_{L^2(\partial\Omega)}^2 \leq l_\gamma \|u\|_{H^1(\Omega)}^2, \quad \|\gamma_o(u)\|_{L^2(\partial\Omega)}^p = \left(\int_{\partial\Omega} |\gamma_o(u(\sigma))|^p \right)^{1/p}. \quad (3.A.27)$$

En particulier, $H_0^1(\Omega)$ est un sous-espace fermé⁷ de $H^1(\Omega)$. En décomposant $H^1(\Omega)$ en somme directe orthogonale

$$H^1(\Omega) = H_0^1(\Omega) \oplus (H_0^1(\Omega))^{\perp_{H^1}},$$

on a :

- γ_o est une bijection de $(H_0^1(\Omega))^{\perp_{H^1}}$ sur $H^{1/2}(\partial\Omega)$,
- γ_o est isométrique de $(H_0^1(\Omega))^{\perp_{H^1}}$ sur $H^{1/2}(\partial\Omega)$.

En particulier, $H^{1/2}(\partial\Omega)$ est un espace de Hilbert pour la norme $\|\cdot\|_{H^{1/2}(\partial\Omega)}$. Enfin il est évident⁸ que $H^1(\Omega)$ s'injecte continument dans $L^2(\Omega)$ et $H^{1/2}(\partial\Omega)$ s'injecte continument dans $L^2(\partial\Omega)$:

$$\|u\|_{L^2(\Omega)} \leq \|u\|_{H^1(\Omega)}, \quad \|g\|_{L^2(\partial\Omega)} \leq \sqrt{l_\gamma} \|g\|_{H^{1/2}(\partial\Omega)}. \quad (3.A.28)$$

Rappelons l'inégalité de Poincaré (théorème B.4.1, dont on donne la démonstration dans l'annexe B) :

$$\|u\|_{L^2(\Omega)} \leq l_\square \left(\|\nabla u\|_{L^2(\Omega)} + \|\gamma_o(u)\|_{L^2(\partial\Omega)} \right), \quad (3.A.29)$$

qui peut encore s'écrire après quelques transformations⁹ :

$$\|u\|_{H^1(\Omega)} \leq \sqrt{2l_\square^2 + 1} \left(\|\nabla u\|_{L^2(\Omega)}^2 + \|\gamma_o(u)\|_{L^2(\partial\Omega)}^2 \right)^{1/2}. \quad (3.A.30)$$

Enfin on a les résultats de compacité et de densité suivants :

- L'injection de $H^1(\Omega) \subset L^2(\Omega)$ est compacte.
- L'injection de $H^{1/2}(\partial\Omega) \subset L^2(\partial\Omega)$ est compacte.
- $H_0^1(\Omega)$ est dense dans $L^2(\Omega)$
- $H^1(\Omega)$ est dense dans $L^2(\Omega)$
- $H^{1/2}(\partial\Omega)$ est dense dans $L^2(\partial\Omega)$

7. Voir théorème de projection .

8.

— $\|u\|_{H^1(\Omega)}^2 = \|\nabla u\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\Omega)}^2 \geq \|u\|_{L^2(\Omega)}^2$
 — $\|\gamma_o(u)\|_{L^2(\partial\Omega)} \leq \sqrt{l_\gamma} \|u\|_{H^1(\Omega)}$ mais cette inégalité est vraie $\forall u_1$ telle que $\gamma_o(u_1) = \gamma_o(u)$. On peut donc choisir $\|\gamma_o(u)\|_{L^2(\partial\Omega)} \leq \sqrt{l_\gamma} \|u_1\|_{H^1(\Omega)} = \sqrt{l_\gamma} \inf \|u\|_{H^1(\Omega)} = \sqrt{l_\gamma} \|\gamma_o(u)\|_{H^{1/2}(\partial\Omega)}$

9. En élevant au carré puis en ajoutant aux deux membres le terme $\|\nabla u\|_{L^2(\Omega)}^2$ et en remarquant que $(a+b)^2 \leq 2(a^2+b^2)$, puis enfin en prenant la racine.

- $H_0^1(\Omega) \times H^{1/2}(\partial\Omega)$ est dense dans $L^2(\Omega) \times L^2(\partial\Omega)$
- $H_0^1(\Omega) \times H^{1/2}(\partial\Omega) \subset L^2(\Omega) \times L^2(\partial\Omega)$ avec injection continue et compacte.

A.4.4.1 Produit scalaire adapté aux bases de Dirichlet Steklov dans $H^1(\Omega)$.

On définit un nouveau produit scalaire dans $H^1(\Omega)$. Ce produit scalaire se révélera utile pour l'utilisation conjointe des bases propres de Dirichlet et de Steklov. En effet les modes de Dirichlet et de Steklov seront orthogonaux par rapport à ce nouveau produit scalaire. On considère

- Le tenseur symétrique de conductivité thermique défini positif $\kappa : \Omega \mapsto \mathcal{M}_d(\mathbb{R})$ tel que :

$$\forall \mathbf{x} \in \Omega, \forall U \in (H^1(\Omega))^d, \quad \kappa^{\min} |U|^2 \leq {}^T U \cdot \kappa(x) \cdot U \leq \kappa^{\max} |U|^2, \quad (3.A.31)$$

où si $0 < \kappa_{oj}(\mathbf{x}) < \infty$ sont les valeurs propres de κ au point \mathbf{x} , les deux bornes κ^{\min} et κ^{\max} sont définies par :

$$\kappa^{\min} = \min_{\substack{\mathbf{x} \in \Omega \\ 1 \leq j \leq d}} \kappa_{oj}(\mathbf{x}), \quad \kappa^{\max} = \max_{\substack{\mathbf{x} \in \Omega \\ 1 \leq j \leq d}} \kappa_{oj}(\mathbf{x}). \quad (3.A.32)$$

- La capacité calorifique notée c et qui est une fonction régulière sur Ω telle que :

$$\forall \mathbf{x} \in \Omega, \quad 0 < c^{\min} \leq c(\mathbf{x}) \leq c^{\max} < \infty. \quad (3.A.33)$$

- Le facteur de Steklov noté ζ et qui est une fonction régulière sur $\partial\Omega$ telle que :

$$\forall \sigma \in \partial\Omega, \quad 0 < \zeta^{\min} \leq \zeta(\sigma) \leq \zeta^{\max} < \infty. \quad (3.A.34)$$

- Une constante $\mu_o > 0$ (facteur d'équilibrage volume-surface).

On définit la forme bilinéaire sur $H^1(\Omega)$

$$\langle u|v \rangle = \int_{\Omega} \nabla u \cdot \kappa \cdot \nabla v + \mu_o \int_{\partial\Omega} \zeta \gamma_o(u) \gamma_o(v). \quad (3.A.35)$$

On a d'après le théorème de la trace et $\forall u, v \in H^1(\Omega)$:

$$\begin{aligned} |\langle u|v \rangle_{H(\Omega)}| &\leq \kappa^{\max} \|\nabla u\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} + \mu_o \zeta^{\max} \|\gamma_o(u)\|_{L^2(\partial\Omega)} \|\gamma_o(v)\|_{L^2(\partial\Omega)} \\ &\leq (\kappa^{\max} + \mu_o l_{\gamma} \zeta^{\max}) \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}. \end{aligned}$$

En particulier l'inégalité est vraie en choisissant $v = u$.

De plus, on a $\forall u \in H^1(\Omega)$, et en utilisant l'inégalité de Poincaré :

$$\begin{aligned} \|u\|_{H(\Omega)}^2 &= |\langle u|u \rangle_{H(\Omega)}| \geq \kappa^{\min} \|\nabla u\|_{L^2(\Omega)}^2 + \mu_o \zeta^{\min} \|\gamma_o(u)\|_{L^2(\partial\Omega)}^2 \\ &\geq \frac{\min(\kappa^{\min}, \mu_o \zeta^{\min})}{2l_{\square}^2 + 1} \|u\|_{H^1(\Omega)}^2. \end{aligned}$$

Ce qui prouve que la forme bilinéaire $\langle \cdot | \cdot \rangle_{H(\Omega)}$ est continue et définie positive sur $H^1(\Omega)$. Elle définit ainsi un produit scalaire sur cet espace. Sa norme induite $\|u\|_{H(\Omega)} = \sqrt{\langle u|u \rangle_{H(\Omega)}}$ est équivalente à la norme usuelle sur $H^1(\Omega)$ (*i.e.* $\|u\|_{H(\Omega)} \equiv \|u\|_{H^1(\Omega)}$).

On définit de plus deux formes bilinéaires sur $L^2(\Omega)$ et $L^2(\partial\Omega)$ respectivement

$$c_{\Omega}(u, v) = (cu|v)_{L^2(\Omega)} = \int_{\Omega} c u v, \quad c_{\partial\Omega}(u, v) = (cu|v)_{L^2(\partial\Omega)} = \int_{\partial\Omega} \zeta u v. \quad (3.A.36)$$

Ces deux formes bilinéaires sont symétriques, continues et définies positives (coercives) respectivement sur $L^2(\Omega)$ et $L^2(\partial\Omega)$. Elles induisent des normes et on a les équivalences $\sqrt{c_{\Omega}(u, u)} \equiv \|u\|_{L^2(\Omega)}$ et $\sqrt{c_{\partial\Omega}(u, u)} \equiv \|u\|_{L^2(\partial\Omega)}$.

On note $E(\Omega) = (H_0^1(\Omega))^{\perp_H}$ l'orthogonal de $H_0^1(\Omega)$ pour le produit scalaire $\langle \cdot | \cdot \rangle_{H(\Omega)}$.

Nous pouvons munir $H^{1/2}(\partial\Omega)$ de la norme $\|g\|_{H,1/2} = \inf_{\gamma_o(u)=g} \|u\|_{H(\Omega)}$. Dans ces conditions on a alors :

- γ_o est une bijection de $E(\Omega)$ sur $H^{1/2}(\partial\Omega)$
- γ_o est isométrique de $(E, \|\cdot\|_{H(\Omega)})$ sur $(H^{1/2}(\partial\Omega), \|\cdot\|_{H,1/2})$
- Sur $H^{1/2}(\partial\Omega)$ on a l'équivalence des normes $\|\cdot\|_{H^{1/2}(\partial\Omega)} \equiv \|\cdot\|_{H,1/2}$

On notera \mathcal{R} l'opérateur de relèvement :

$$\begin{aligned} \mathcal{R} : H^{1/2}(\partial\Omega) &\rightarrow E(\Omega) \\ w &\mapsto \gamma_o(\mathcal{R}(w)) = w. \end{aligned} \quad (3.A.37)$$

On a les résultats fondamentaux suivants :

- $u \in E(\Omega) \Leftrightarrow u \in H^1(\Omega)$ et $\nabla \cdot \kappa \cdot \nabla u = 0$,
- $u = \mathcal{R}(g) \Leftrightarrow u \in H^1(\Omega), \gamma_o(u) = g$ et $\nabla \cdot \kappa \cdot \nabla u = 0$.

A.4.4.2 Modes de Dirichlet.

Pour $f \in L^2(\Omega)$, la forme linéaire $v \in H_0^1(\Omega) \mapsto \int_{\Omega} c f v dx$ est continue sur $H_0^1(\Omega)$. D'après le théorème de Lax-Milgram, il existe un unique $u \in H_0^1(\Omega)$ tel que, pour tout

$v \in H_0^1(\Omega)$:

$$\int_{\Omega} \nabla u \cdot \kappa \cdot \nabla v = \int_{\Omega} c f v, \quad (3.A.38)$$

ou encore, puisque les fonction de $H_0^1(\Omega)$ sont de trace nulle $\langle u|v \rangle_{H(\Omega)} = c_{\Omega}(f, v)$.

On note \mathcal{D}_o l'opérateur avec condition limite de Dirichlet homogène :

$$\begin{aligned} \mathcal{D}_o : L^2(\Omega) &\rightarrow H_0^1(\Omega) \\ f &\mapsto u = \mathcal{D}_o(f) \Leftrightarrow \begin{cases} \Omega & -\nabla \cdot \kappa \cdot \nabla u = c f, \\ \partial\Omega & u = 0. \end{cases} \end{aligned} \quad (3.A.39)$$

En utilisant l'inégalité de Cauchy-Schwarz puis celle de Poincaré (dans H_0^1) on peut écrire :

$$\left| \int_{\Omega} c f u \right| \leq c^{\max} |(f|u)_{L^2(\Omega)}| \leq c^{\max} \|f\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} \leq c^{\max} l_{\square} \|f\|_{L^2(\Omega)} \|\nabla u\|_{L^2(\Omega)}. \quad (3.A.40)$$

D'autre part en choisissant $v = u$ dans (3.A.38) et en utilisant le caractère défini positif du tenseur κ nous obtenons :

$$\left| \int_{\Omega} c f u \right| = \int_{\Omega} \nabla u \cdot \kappa \cdot \nabla u \geq \kappa^{\min} \|\nabla u\|_{L^2(\Omega)}^2. \quad (3.A.41)$$

En combinant (3.A.40) et (3.A.41) :

$$\|\nabla u\|_{L^2(\Omega)} \leq \frac{c^{\max} l_{\square}}{\kappa^{\min}} \|f\|_{L^2(\Omega)}. \quad (3.A.42)$$

En utilisant à nouveau l'inégalité de Poincaré on a pour les fonctions de $H_0^1(\Omega)$:

$$\|u\|_{H^1(\Omega)} \leq \sqrt{1 + l_{\square}^2} \|\nabla u\|_{L^2(\Omega)} \leq \frac{c^{\max} l_{\square} \sqrt{1 + l_{\square}^2}}{\kappa^{\min}} \|f\|_{L^2(\Omega)}. \quad (3.A.43)$$

Enfin puisque $u = \mathcal{D}_o(f)$:

$$\|\mathcal{D}_o(f)\|_{H^1(\Omega)} \leq C \|f\|_{L^2(\Omega)}, \quad C = \frac{c^{\max} l_{\square} \sqrt{1 + l_{\square}^2}}{\kappa^{\min}}. \quad (3.A.44)$$

Cette inégalité prouve que l'opérateur \mathcal{D}_o est continu de $L^2(\Omega)$ vers $H_0^1(\Omega)$.

On appelle maintenant \mathcal{D} le même opérateur mais considéré comme allant de $L^2(\Omega)$ vers $L^2(\Omega)$: $\mathcal{D} = Id_{H^1 \rightarrow L^2} \circ \mathcal{D}_o$. L'opérateur \mathcal{D} est compact de $L^2(\Omega)$ vers $L^2(\Omega)$. De plus, \mathcal{D} est autoadjoint pour le produit scalaire de $L^2(\Omega)$ puisque :

$$c_{\Omega}(f, \mathcal{D}(v)) = \int_{\Omega} c f \mathcal{D}(v) = \langle \mathcal{D}(f) | \mathcal{D}(v) \rangle_{H(\Omega)}. \quad (3.A.45)$$

Par densité de $H_0^1(\Omega)$ dans $L^2(\Omega)$, l'opérateur \mathcal{D} est injectif. De plus, si f est un mode propre de \mathcal{D} , de valeur propre ϑ , on a par définition :

$$\vartheta c_{\Omega}(f, f) = \int_{\Omega} c f \mathcal{D}(f) = \|\mathcal{D}(f)\|_{H(\Omega)}^2. \quad (3.A.46)$$

Donc $\vartheta > 0$.

On sait alors qu'il existe une base orthonormée $\{V_i^D\}_{i \in \mathbb{N}}$ de $H_0^1(\Omega)$ pour la norme $\|\cdot\|_{H(\Omega)}$ composée de vecteurs propres de \mathcal{D} et telle que :

$$\mathcal{D}(V_i^D) = \vartheta_i^D c V_i^D \quad \text{avec} \quad \vartheta_1^D \geq \vartheta_2^D \geq \dots \geq \vartheta_i^D \geq \vartheta_{i+1}^D \geq \dots \quad \text{et} \quad \lim_{i \rightarrow \infty} \vartheta_i^D = 0.$$

La condition $\mathcal{D}(V_i^D) = \vartheta_i^D c V_i^D$ se réécrit en posant $\mu_i^D = 1/\vartheta_i^D$:

$$\begin{cases} \Omega & -\nabla \cdot \kappa \cdot \nabla V_i^D = \mu_i^D c V_i^D, \\ \partial\Omega & V_i^D = 0. \end{cases} \quad (3.A.47)$$

Sous forme variationnelle faible le problème de recherche des modes propres de Dirichlet s'écrit¹⁰ :

Trouver $(V_i^D, \mu_i^D) \in H_0^1(\Omega) \times \mathbb{R}^+$ telles que

$$\langle V_i^D | u \rangle_{H(\Omega)} = \mu_i^D c_{\Omega}(V_i^D, u), \quad \forall u \in H_0^1(\Omega). \quad (3.A.48)$$

A.4.4.3 Modes de Steklov.

Pour $\varphi \in L^2(\partial\Omega)$, la forme linéaire $g \in H^{1/2}(\partial\Omega) \mapsto \int_{\partial\Omega} \zeta \varphi g \, d\sigma$ est continue sur $H^{1/2}(\partial\Omega)$. D'après le théorème de Lax-Milgram, il existe un unique $\phi \in H^{1/2}(\partial\Omega)$ tel que, pour tout $v \in H^{1/2}(\partial\Omega)$:

$$\int_{\Omega} \nabla \mathcal{R}(\phi) \cdot \kappa \cdot \nabla \mathcal{R}(v) + \mu_o \int_{\partial\Omega} \zeta \phi v = \int_{\partial\Omega} \zeta \varphi v, \quad (3.A.49)$$

ou encore $\langle \mathcal{R}(\phi) | \mathcal{R}(v) \rangle_{H(\Omega)} = c_{\partial\Omega}(\varphi, v)$.

10. Où plus explicitement $\int_{\Omega} \nabla V_i^D \cdot \kappa \cdot \nabla u = \mu_i^D \int_{\Omega} c V_i^D u$.

On note \mathcal{S}_o l'opérateur tel que :

$$\begin{aligned} \mathcal{S}_o : L^2(\partial\Omega) &\rightarrow H^{1/2}(\partial\Omega) \\ \varphi &\mapsto \phi = \mathcal{S}_o(\varphi) \\ u = \mathcal{R}(\mathcal{S}_o(\varphi)) &\Leftrightarrow \begin{cases} \Omega & -\nabla \cdot \kappa \cdot \nabla u = 0, \\ \partial\Omega & \kappa \cdot \nabla u \cdot \underline{n} + \mu_o \zeta \gamma_o(u) = \zeta \varphi. \end{cases} \end{aligned} \quad (3.A.50)$$

En utilisant l'inégalité de Cauchy-Schwarz puis celle de la trace on peut écrire :

$$\left| \int_{\partial\Omega} \zeta \varphi \phi \right| \leq \zeta^{\max} |(\varphi|\phi)_{L^2(\partial\Omega)}| \leq \zeta^{\max} \|\varphi\|_{L^2(\partial\Omega)} \|\phi\|_{L^2(\partial\Omega)} \leq \zeta^{\max} \sqrt{l_\gamma} \|\varphi\|_{L^2(\partial\Omega)} \|\mathcal{R}(\phi)\|_{H^1(\Omega)}. \quad (3.A.51)$$

D'autre part en choisissant $v = \varphi$ dans (3.A.49) et en utilisant l'inégalité de Poincaré (3.A.30) nous obtenons :

$$\left| \int_{\partial\Omega} \zeta \varphi \phi \right| = \int_{\Omega} \nabla \mathcal{R}(\phi) \cdot \kappa \cdot \nabla \mathcal{R}(\phi) + \mu_o \int_{\partial\Omega} \zeta \phi^2 \geq \kappa^{\min} \|\nabla \mathcal{R}(\phi)\|_{L^2(\Omega)}^2 + \mu_o \zeta^{\min} \|\phi\|_{L^2(\partial\Omega)}^2 \quad (3.A.52)$$

$$\geq \frac{\min(\kappa^{\min}, \mu_o \zeta^{\min})}{2l_\square^2 + 1} \|\mathcal{R}(\phi)\|_{H^1(\Omega)}^2. \quad (3.A.53)$$

En combinant les deux relations (3.A.51) - (3.A.52), en utilisant la définition de la norme (3.A.26) et puisque $\phi = \mathcal{S}_o(\varphi)$:

$$\|\mathcal{S}_o(\varphi)\|_{H^{1/2}(\partial\Omega)} = \inf_{\gamma_o(\mathcal{R}(\phi))=\phi} \|\mathcal{R}(\phi)\|_{H^1(\Omega)} \leq \|\mathcal{R}(\phi)\|_{H^1(\Omega)} \leq C \|\varphi\|_{L^2(\partial\Omega)}, \quad (3.A.54)$$

où nous avons posé :

$$C = \frac{(2l_\square^2 + 1) \zeta^{\max} \sqrt{l_\gamma}}{\min(\kappa^{\min}, \mu_o \zeta^{\min})}.$$

On a donc $\|\mathcal{S}_o(\varphi)\|_{H^{1/2}(\partial\Omega)} \leq C \|\varphi\|_{L^2(\partial\Omega)}$, de sorte que l'opérateur \mathcal{S}_o est continu de $L^2(\partial\Omega)$ vers $H^{1/2}(\partial\Omega)$.

On note à présent \mathcal{S} le même opérateur mais considéré comme allant de $L^2(\partial\Omega)$ vers $L^2(\partial\Omega)$: $\mathcal{S} = Id_{H^{1/2} \rightarrow L^2(\partial\Omega)} \circ \mathcal{S}_o$. L'opérateur \mathcal{S} est compact de $L^2(\partial\Omega)$ vers $L^2(\partial\Omega)$. De plus, \mathcal{S} est autoadjoint pour le produit scalaire de $L^2(\partial\Omega)$ puisque :

$$c_{\partial\Omega}(\varphi, \mathcal{S}(\phi)) = \int_{\partial\Omega} \zeta \varphi \mathcal{S}(\phi) = \langle \mathcal{R}(\mathcal{S}(\varphi)) | \mathcal{R}(\mathcal{S}(\phi)) \rangle_{H(\Omega)}. \quad (3.A.55)$$

Par densité de $H^{1/2}(\partial\Omega)$ dans $L^2(\partial\Omega)$, l'opérateur \mathcal{S} est injectif. De plus, si φ est un mode propre de \mathcal{S} , de valeur propre ϑ , on a :

$$\vartheta c_{\partial\Omega}(\varphi, \varphi) = \int_{\partial\Omega} \zeta \varphi \mathcal{S}(\varphi) dv = \|\mathcal{R}(\mathcal{S}(\varphi))\|_{H(\Omega)}^2. \quad (3.A.56)$$

La valeur propre vérifie donc $\vartheta > 0$.

On sait alors qu'il existe une base orthonormée $\{V_i^S\}_{i \in \mathbb{N}}$ de $E(\Omega)$ pour la norme $\|\cdot\|_{H(\Omega)}$ composée de vecteurs propres de $\mathcal{R} \circ \mathcal{S} \circ \gamma_o$ et telle que :

$$\mathcal{R}(\mathcal{S}(\gamma_o(V_i^S))) = \vartheta_i^S \zeta V_i^S \quad \text{avec} \quad \vartheta_1^S \geq \vartheta_2^S \geq \dots \vartheta_i^S \geq \vartheta_{i+1}^S \geq \dots \quad \text{et} \quad \lim_{i \rightarrow \infty} \vartheta_i^S = 0.$$

La condition $\mathcal{R}(\mathcal{S}(\gamma_o(V_i^S))) = \vartheta_i^S \zeta V_i^S$ se réécrit en posant $\mu_i^S = 1/\vartheta_i^S$:

$$\begin{cases} \Omega & -\nabla \cdot \kappa \cdot \nabla V_i^S = 0, \\ \partial\Omega & \kappa \cdot \nabla V_i^S \cdot \underline{n} + \mu_o \zeta V_i^S = \mu_i^S \zeta V_i^S. \end{cases} \quad (3.A.57)$$

Notons ici que la première valeur propre est simple et est associée à un mode uniforme dans tout le domaine (y compris sur le bord). On a $\mu_1^S = \mu_o$. Sous forme variationnelle faible le problème de recherche des modes propres de Steklov s'écrit :

Trouver $(V_i^S, \mu_i^S) \in E(\Omega) \times \mathbb{R}^+$ telles que

$$\langle V_i^S | u \rangle_{H(\Omega)} = \mu_i^S c_{\partial\Omega}(V_i^S, u), \quad (3.A.58)$$

où plus explicitement

$$\int_{\Omega} \nabla V_i^S \cdot \kappa \cdot \nabla u dv + \mu_o \int_{\partial\Omega} \zeta V_i^S u da = \mu_i^S \int_{\partial\Omega} \zeta V_i^S u da.$$

A.4.4.4 Orthogonalité de la base de Dirichlet-Steklov.

On définit la formes bilinéaire¹¹ sur $H^1(\Omega)$

$$a_{\Omega}(u, v) = \int_{\Omega} \nabla u \cdot \kappa \cdot \nabla v. \quad (3.A.59)$$

11. Cette forme est symétrique continue mais pas coercive. Elle définit un semi-produit scalaire dans $H^1(\Omega)$.

On a alors les relations d'orthogonalité :

$$\forall i, j \in \mathbb{N} \quad \left\{ \begin{array}{l} \langle V_i^{\mathcal{X}} | V_j^{\mathcal{Y}} \rangle = \delta_i^j \delta_{\mathcal{X}}^{\mathcal{Y}}, \quad \forall \mathcal{X}, \mathcal{Y} \in \{D, S\}, \\ c_{\partial\Omega}(V_i^{\mathcal{X}}, V_j^{\mathcal{Y}}) = \frac{\delta_i^j}{\mu_i^S} \delta_{\mathcal{X}}^S \delta_{\mathcal{Y}}^S, \quad \forall \mathcal{X}, \mathcal{Y} \in \{D, S\}, \\ a_{\Omega}(V_i^{\mathcal{X}}, V_j^{\mathcal{Y}}) = \begin{cases} \left(\frac{\mu_i^S - \mu_o}{\mu_i^S} \right) \delta_i^j & si \quad \mathcal{X} = \mathcal{Y} = S, \\ \delta_i^j & si \quad \mathcal{X} = \mathcal{Y} = D, \\ 0 & si \quad \mathcal{X} \neq \mathcal{Y}. \end{cases} \end{array} \right. \quad (3.A.60)$$

On a aussi les relations additionnelles

$$\forall i, j \in \mathbb{N} \quad c_{\Omega}(V_i^{\mathcal{X}}, V_j^{\mathcal{Y}}) = \begin{cases} \frac{\delta_i^j}{\mu_i^D} & \mathcal{X} = \mathcal{Y} = D, \\ \neq 0 & si \quad \mathcal{X} = S; \mathcal{Y} \in \{D, S\}. \end{cases} \quad (3.A.61)$$

Attention l'orthogonalité par rapport à $c_{\Omega}(\cdot, \cdot)$ n'est que partielle. Elle concerne uniquement les modes de Dirichlet. C'est pour cette raison que les dynamiques des modes de Dirichlet et de Steklov se retrouvent couplées dans le problème d'évolution thermique instationnaire.

A.4.4.5 Décomposition d'une fonction de $H^1(\Omega)$ sur la base de Dirichlet-Steklov.

Définition A.4.2 Soit F un sous-espace vectoriel de H . On définit l'orthogonal de F par :

$$F^{\perp} = \{u \in H, \text{ tel que } \forall v \in F (u|v)_H = 0\}.$$

△

F^{\perp} est un sous-espace fermé de H et $F^{\perp} \cap F = \emptyset$, $H = F \oplus F^{\perp}$.
En particulier on a $H^{\perp} = \{0\}$.

Théorème: A.4.1 THÉORÈME DE PROJECTION [103] P.102

Soit H un espace de Hilbert, muni d'un produit scalaire $(\cdot|\cdot)_H$ et de la norme associée $\|\cdot\|_H$. Soit F un sous-espace vectoriel fermé de H . Alors tout vecteur u de H se décompose de manière unique sous la forme :

$$u = \overset{\circ}{u} + \overset{\Delta}{u}, \quad \overset{\circ}{u} \in F \quad \overset{\Delta}{u} \in F^\perp. \quad (3.A.62)$$

De plus, on a

$$\|u - \overset{\circ}{u}\|_H = \inf\{\|u - v\|_H, v \in F\},$$

et $\overset{\circ}{u}$ est caractérisé par le fait que pour tout $v \in F$, $(u - \overset{\circ}{u}|v)_H = 0$. On dit que $\overset{\circ}{u} = P_F(u)$ est le projeté orthogonal de u sur F .

▲

Nous avons montré que $H_0^1 = Vect\{V_i^D\}_{i \in \mathbb{N}}$, $E = Vect\{V_i^S\}_{i \in \mathbb{N}}$ et $H^1(\Omega) = H_0^1(\Omega) \oplus E(\Omega)$ (avec $H_0^1(\Omega) \perp_H E(\Omega)$ et $E \cap H_0^1(\Omega) = \emptyset$).

On peut donc décomposer toutes les fonctions de $H^1(\Omega)$ sur la base de Dirichlet-Steklov. Plus exactement :

$$\begin{aligned} \exists (x_i^D, x_i^S)_{i \in \mathbb{N}} \quad \text{tels que } \forall u \in H^1(\Omega), \\ u = \overset{\circ}{u} + \overset{\Delta}{u}, \\ \overset{\circ}{u} = \sum_{i=1}^{\infty} x_i^D V_i^D, \quad \overset{\Delta}{u} = \sum_{i=1}^{\infty} x_i^S V_i^S. \end{aligned} \quad (3.A.63)$$

On écrira donc formellement :

$$u = \sum_{\mathcal{X} \in \{D,S\}} \sum_{i=1}^{\infty} x_i^{\mathcal{X}} V_i^{\mathcal{X}}, \quad x_i^{\mathcal{X}} = \langle u | V_i^{\mathcal{X}} \rangle_{H(\Omega)}. \quad (3.A.64)$$

En pratique on fera une décomposition réduite avec $N = N_D + N_S$ modes propres, à partir d'une base de référence comportant $N_o = N_{oD} + N_{oS}$ modes avec $N \lll N_o$. Nous écrirons :

— Approximation par troncature d'ordre très élevé ("référence numérique")

$$u \cong u_{N_o} = \sum_{\mathcal{X} \in \{D,S\}} \sum_{i=1}^{N_{o\mathcal{X}}} x_i^{\mathcal{X}} V_i^{\mathcal{X}}. \quad (3.A.65)$$

— Réduction par troncature énergétique

$$u_N = \sum_{\mathcal{X} \in \{D,S\}} \sum_{i=1}^{N_{\mathcal{X}}} x_i^{\mathcal{X}} V_i^{\mathcal{X}}. \quad (3.A.66)$$

— Réduction par amalgame modal

$$\tilde{u}_N = \sum_{\mathcal{X} \in \{D, S\}} \sum_{i=1}^{N_{\mathcal{X}}} \tilde{x}_i^{\mathcal{X}} \tilde{V}_i^{\mathcal{X}}. \quad (3.A.67)$$

La méthode de troncature énergétique converge dans $L^2(]0, \Delta\tau[; H^1(\Omega))$ et les solutions amalgamées vérifient :

$$\|u_{N_o} - \tilde{u}_N\|_{H(\Omega)} \leq \|u_{N_o} - u_N\|_{H(\Omega)}. \quad (3.A.68)$$

On a de plus :

$$\frac{1}{\sqrt{\mathcal{V}_{ol}(\Omega)}} \|u_{N_o} - \tilde{u}_N\|_{L^1(\Omega)} \leq \|u_{N_o} - \tilde{u}_N\|_{L^2(\Omega)} \leq C_* \|u_{N_o} - \tilde{u}_N\|_{H(\Omega)}. \quad (3.A.69)$$

Annexe B

Résultats mathématiques généraux

La démonstration de la complétude de la base de Dirichlet Steklov donnée en annexe A repose sur l'utilisation d'un certain nombre de théorèmes classiques de l'analyse fonctionnelle. Le lecteur intéressé pourra consulter avec profit les ouvrages mathématiques cités dans la bibliographie et que nous avons utilisé (mais il en existe bien d'autres !). Toutefois pour faciliter la lecture du manuscrit nous avons regroupé dans la présente annexe les points essentiels utiles à notre démarche. Certaines démonstrations sont mêmes montrées car les idées qu'elles mettent en oeuvre sont reprises dans l'annexe B et les suivantes. Enfin par souci de cohérence nous avons harmonisé les notations avec celles de notre manuscrit.

B.1 Produit scalaire et norme.

Définition B.1.1 (Produit scalaire.) *Si F est un espace vectoriel (sur le corps de nombres réels \mathbb{R} pour fixer les idées) le produit scalaire de deux éléments $u, v \in F$, que nous noterons $(\cdot|\cdot)_F$, sera défini par :*

$$\begin{array}{l|l} \text{symétrie} & (u|v)_F = (v|u)_F \quad \forall u, v \in F, \\ \text{linéarité} & ((u+v)|w)_F = (u|w)_F + (v|w)_F \quad \forall u, v, w \in F, \\ \text{homogénéité} & (\alpha u|v)_F = \alpha(u|v)_F \quad \forall \alpha \in \mathbb{R} \quad \forall u, v \in F, \\ \text{défini positif} & (u|u)_F = 0 \implies u = 0(p.p.) \quad \forall u \in F. \end{array} \quad (3.B.1)$$

Si la dernière propriété n'est pas vérifiée on parle de semi-produit scalaire.

Δ

Exemples :

— Si $\underline{U} = \sum_{i=1}^d u_i \underline{e}_i$ alors $U \cdot V = \sum_{i=1}^d u_i v_i$ est un produit scalaire sur \mathbb{R}^d ,

— $\int_{\Omega} uv$ est un produit scalaire sur l'espace $L^2(\Omega)$ des fonctions de carré sommable¹.

On dit que deux éléments de F sont orthogonaux si leur produit scalaire est nul.

Définition B.1.2 (Norme.) Si F est un espace vectoriel muni du produit scalaire $(\cdot|\cdot)_F$, La norme induite par ce produit scalaire sera notée $\|\cdot\|_F$ et définie par :

$$\forall u \in F, \quad \|u\|_F = \sqrt{(u|u)_F}. \quad (3.B.2)$$

On a l'inégalité de Cauchy-Schwarz²

$$\forall u, v \in F, \quad |(u|v)_F| \leq \|u\|_F \|v\|_F. \quad (3.B.3)$$

Si $(\cdot|\cdot)_F$ est un semi produit scalaire on parle de semi-norme.

△

Définition B.1.3 (Continuité et coercivité.) Soit $a(\cdot, \cdot)$ un opérateur bilinéaire à coefficients réels et à valeurs réelles sur un espace de Hilbert $F(\Omega)$. On dit que $a(\cdot, \cdot)$ est :

- Symétrique si : $\forall u, v \in F(\Omega), \quad a(u, v) = a(v, u),$
- Continu si : $\forall u, v \in F(\Omega), \quad \exists C > 0$ telle que $|a(u, v)| \leq C \|u\|_{F(\Omega)} \|v\|_{F(\Omega)},$
- Coercif si : $\forall u \in F(\Omega), \quad \exists C' > 0$ telle que $a(u, u) \geq C' \|u\|_{F(\Omega)}^2.$

△

B.2 Espace de Hilbert.

B.2.1 Suites de Cauchy.

Définition B.2.1 Une suite $(u_n)_{n \in \mathbb{N}}$ dans un espace métrique (E, d_{ist}) où d_{ist} est une distance est dite de Cauchy si :

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ tel que } \forall p, q \in \mathbb{N}, \quad d_{ist}(u_p, u_q) \leq \varepsilon. \quad (3.B.4)$$

△

1. $L^2(\Omega)$ est bien un espace vectoriel car $\int (u+v)^2 \leq \int u^2 + \int v^2 < \infty$.
 2. Il est évident que $(u|v)_F \leq |(u|v)_F|$.

Toute suite convergente est de Cauchy mais la réciproque est fautive en général. Les suites de Cauchy sont bornées. L'image d'une suite de Cauchy par une application uniformément continue est une suite de Cauchy. Enfin une suite de Cauchy a au plus une valeur d'adhérence et si cette valeur existe alors la suite de Cauchy converge.

Définition B.2.2 *Un espace métrique complet est un espace métrique dans lequel toute suite de Cauchy converge.*

△

La propriété de complétude dépend de la distance (ou de la norme). Il est donc important de toujours préciser la distance que l'on prend quand on parle d'espace complet. Intuitivement un espace est complet s'il " n'a aucun point manquant " ³ .

Définition B.2.3 *Un espace de Hilbert est un espace vectoriel muni d'un produit scalaire (espace préhilbertien) et qui est complet pour la norme induite par ce produit scalaire. Un espace est complet si toute suite de Cauchy converge pour la norme considérée.*

△

B.3 Fonction à support borné.

Dire qu'une fonction régulière u a son support borné dans le fermé $\bar{\Omega}$ signifie qu'elle s'annule à l'infini si le fermé n'est pas borné (on dit aussi que la fonction a un support compact dans $\bar{\Omega}$). On prendra garde que la fonction u ne s'annule pas nécessairement sur le bord $\partial\Omega$ du domaine Ω .

On définit l'espace $C_c^\infty(\bar{\Omega})$ comme l'espace des fonctions régulières de classe C^∞ à support borné (ou compact) dans le fermé $\bar{\Omega}$. Si Ω est borné toutes les fonctions de $C^\infty(\bar{\Omega})$ ont nécessairement un support borné ce qui implique $C^\infty(\bar{\Omega}) = C_c^\infty(\bar{\Omega})$. Les fonctions de $C_c^\infty(\bar{\Omega})$ ne s'annulent pas nécessairement ⁴ sur $\partial\Omega$, contrairement aux fonctions de $C_c^\infty(\Omega)$. L'espace $C_c^\infty(\bar{\Omega})$ est dense dans $H^1(\Omega)$.

3. L'exemple type est l'ensemble \mathbb{Q} des nombres rationnels. Il n'est pas complet car il existe des suites de Cauchy de nombres rationnels qui convergent vers des irrationnels. Par exemple la suite $x_1 = 1, x_{n+1} = \frac{x_n}{2} + \frac{1}{x_n}$ est de Cauchy et $(1, \frac{3}{2}, \frac{17}{12}, \frac{577}{408}, \frac{665857}{470832}, \dots)$ converge (très rapidement) vers $\sqrt{2} \notin \mathbb{Q}$. En fait \mathbb{Q} est dense dans \mathbb{R} et on démontre par ailleurs que \mathbb{R} est complet.

4. Si Ω n'est pas borné les fonctions de $C_c^\infty(\bar{\Omega})$ s'annulent à l'infini.

On note $C_c^\infty(\Omega)$ l'espace des fonctions de classe C^∞ à support compact dans l'ouvert Ω . Les fonctions de l'espace $C_c^\infty(\Omega)$ et toutes leurs dérivées s'annulent sur le bord du domaine Ω . L'espace $C_c^\infty(\Omega)$ est dense à la fois dans $H_0^1(\Omega)$ (par définition⁵) et dans $L^2(\Omega)$. Autrement dit

$$\forall u \in L^2(\Omega), \text{ il existe une suite } u_n \in C_c^\infty \text{ telle que } \lim_{n \rightarrow \infty} \|u_n - u\|_{L^2(\Omega)} = 0. \quad (3.B.5)$$

B.4 Inégalité de Poincaré.

Un très grand nombre de preuves repose sur un résultat technique de minoration de la norme L^2 connu sous le nom d'inégalité de Poincaré.

Théorème: B.4.1 INÉGALITÉ DE POINCARÉ *On considère un domaine borné Ω de \mathbb{R}^d . On note u les fonctions de l'espace $H^1(\Omega)$. On désigne par $l_\square(\Omega)$ (ou l_\square) une constante réelle strictement positive⁶.*

$$\forall u \in H^1(\Omega), \exists l_\square > 0 \text{ telle que } \|u\|_{L^2(\Omega)} \leq l_\square (\|u|_{\partial\Omega}\|_{L^2(\partial\Omega)} + \|\nabla u\|_{L^2(\Omega)}). \quad (3.B.6)$$

▲

Démonstration: B.4.1 *On trouvera les éléments de démonstration des inégalités de Poincaré par exemple dans [90] [92].*

Malheureusement, ce sont des démonstrations par contradiction qui ne disent rien de la valeur de la constante l_\square dans l'inégalité de Poincaré. Pour déterminer une expression explicite de cette constante, il faut utiliser une démarche "constructive". Elle est classique dans $H_0^1(\Omega)$. Nous en donnons une démonstration moins habituelle dans $H^1(\Omega)$. Nous faisons l'hypothèse que le domaine Ω est compris dans la bande $|x_1| < l_1/2$ (quitte à faire un changement de repère avec une translation pour recentrer le domaine et une rotation pour se ramener à x_1). On note $\underline{n} = n_i \underline{e}_i$ la normale unitaire en un point σ de la frontière $\partial\Omega$ orientée vers l'extérieur du domaine Ω . Nous remarquons en particulier que $\forall i, n_i \leq 1$.

$$\|u\|_{L^2(\Omega)}^2 = \int_{\Omega} u^2(x) dx = \int_{\Omega} u^2(x) \frac{\partial x_1}{\partial x_1} dx \quad \text{car } \frac{\partial x_1}{\partial x_1} = 1$$

5. L'espace de H_0^1 est l'adhérence de $C_c^\infty(\Omega)$ dans $H^1(\Omega)$. C'est aussi le sous espace de $H^1(\Omega)$ constitué des fonctions qui s'annulent sur le bord $\partial\Omega$ puisque c'est le cas des fonctions de $C_c^\infty(\Omega)$. Comme $C_c^\infty(\Omega)$ est un sous espace strict de $C_c^\infty(\bar{\Omega})$, l'espace $H_0^1(\Omega)$ est strictement plus petit que $H^1(\Omega)$ (Ω borné). On écrit $H_0^1(\Omega) \subsetneq H^1(\Omega)$ (inclusion stricte).

6. Cette constante a la dimension d'une longueur. Elle dépend de la taille du domaine mais elle est indépendante des fonctions u .

$$\begin{aligned}
&= \int_{\partial\Omega} \sigma_1 n_1 u^2(\sigma) d\sigma - \int_{\Omega} x_1 \frac{\partial u^2(x)}{\partial x_1} dx && \text{par la formule de Stokes} \\
&= \int_{\partial\Omega} \sigma_1 n_1 u^2(\sigma) d\sigma - 2 \int_{\Omega} x_1 u(x) \frac{\partial u(x)}{\partial x_1} dx && \text{en dérivant } u^2(x) \\
&\leq \frac{l_1}{2} \int_{\partial\Omega} u^2(\sigma) d\sigma + l_1 \int_{\Omega} |u(x) \frac{\partial u(x)}{\partial x_1}| dx \\
&\leq \frac{l_1}{2} \int_{\partial\Omega} u^2(\sigma) d\sigma + \frac{1}{2} \int_{\Omega} |u(x)|^2 dx + \frac{l_1^2}{2} \int_{\Omega} \left| \frac{\partial u(x)}{\partial x_1} \right|^2 dx,
\end{aligned}$$

car $|ab| \leq \frac{1}{2}(\epsilon a^2 + \frac{b^2}{\epsilon}) \forall \epsilon > 0$ et en choisissant $\epsilon = l_1$. Donc en réorganisant :

$$\|u\|_{L^2(\Omega)}^2 \leq l_1 \left(\int_{\partial\Omega} u^2(\sigma) d\sigma + l_1 \int_{\Omega} \left| \frac{\partial u(x)}{\partial x_1} \right|^2 dx \right).$$

$$\text{Enfin } \left| \frac{\partial u(x)}{\partial x_1} \right|^2 \leq \sum_{i=1}^d \left| \frac{\partial u(x)}{\partial x_i} \right|^2 = |\nabla u|^2 :$$

$$\|u\|_{L^2(\Omega)}^2 \leq l_1 \left(\int_{\partial\Omega} u^2 da + l_1 \int_{\Omega} |\nabla u|^2 dv \right).$$

Donc en posant :

$$l_{\square} = \min_{1 \leq i \leq d} (l_i \max(1, l_i)), \quad (3.B.7)$$

nous retrouvons la minimisation (3.B.6)

■

La valeur $l_{\square} = \min_{1 \leq i \leq d} (l_i \max(1, l_i))$ n'est pas optimale. Elle est assez pertinente pour une sphère, un parallélépipède rectangle. Elle l'est beaucoup moins pour des corps creux par exemple. Dans le cas de l'espace $H_0^1(\Omega)$ des fonctions de trace nulle sur le bord on peut déterminer la constante optimale. Elle est fonction de la première valeur propre. Nous avons :

Proposition B.4.1 Pour les fonctions de traces nulles notons μ_1^D la plus petite valeur propre du problème de Dirichlet :

$$\int_{\Omega} \nabla V_i^D \cdot \kappa \cdot \nabla v = \mu_i^D \int_{\Omega} c V_i^D v \quad \forall v \in H_0^1(\Omega).$$

Alors pour toutes les fonctions de $H_0^1(\Omega)$ on a l'inégalité de Poincaré suivante :

$$\|u\|_{L^2(\Omega)} \leq \sqrt{\frac{1}{c^{\min} \mu_1^D}} \|u\|_{H(\Omega)}. \quad (3.B.8)$$

Dans cette expression $c^{\min} = \min_{\Omega} c(x)$

▲

Preuve B.4.1 Pour montrer 3.B.8 commençons par montrer une inégalité pour les fonctions propres de Dirichlet. Les modes de Dirichlet forment une base infinie mais dénombrable de $H_0^1(\Omega)$. La première valeur propre est simple. On a :

$$0 < \mu_1^D < \mu_2^D \leq \mu_3^D \leq \dots$$

Les modes vérifient les relations d'orthogonalité :

$$\int_{\Omega} \nabla V_i^D \cdot \kappa \cdot \nabla V_j^D = \int_{\Omega} c V_i^D V_j^D = 0, \quad \forall i, j \in \mathbb{N}^*, i \neq j.$$

La normalisation de ces modes impose $\|V_i^D\|_{H(\Omega)} = 1$, et puisque la trace des modes est nulle sur la frontière :

$$(V_i^D | V_j^D)_{H(\Omega)} = \int_{\Omega} \nabla V_i^D \cdot \kappa \cdot \nabla V_j^D = \delta_i^j.$$

En utilisant dans la définition du problème aux valeurs propres de Dirichlet le mode propre V_i^D comme fonction test nous trouvons :

$$\int_{\Omega} \nabla V_i^D \cdot \kappa \cdot \nabla V_i^D = \mu_i^D \int_{\Omega} c V_i^D V_i^D,$$

d'où en tenant compte de la normalisation :

$$\int_{\Omega} c V_i^D V_i^D = \frac{1}{\mu_i^D} \leq \frac{1}{\mu_1^D}.$$

Ce résultat préliminaire étant acquis considérons une fonction $u \in H_0^1(\Omega)$. Elle est décomposable sur la base de Dirichlet. Autrement dit en utilisant l'orthogonalité :

$$u = \sum_{i=1}^{\infty} \alpha_i V_i^D \quad \text{avec } \forall i \in \mathbb{N}^* \quad \alpha_i = (u | V_i^D)_{H(\Omega)}.$$

Calculons la norme de u :

$$\|u\|_{H(\Omega)}^2 = \left(\sum_{i=1}^{\infty} \alpha_i V_i^D \middle| \sum_{j=1}^{\infty} \alpha_j V_j^D \right)_{H(\Omega)} = \sum_{i=1}^{\infty} \alpha_i^2.$$

D'autre part :

$$\int_{\Omega} c u^2 = \int_{\Omega} c \left(\sum_{i=1}^{\infty} \alpha_i V_i^D \right) \left(\sum_{j=1}^{\infty} \alpha_j V_j^D \right) = \sum_{i=1}^{\infty} \alpha_i^2 \int_{\Omega} c V_i^D V_i^D = \sum_{i=1}^{\infty} \frac{\alpha_i^2}{\mu_i^D},$$

soit en utilisant le résultat préliminaire :

$$\int_{\Omega} c u^2 \leq \frac{1}{\mu_1^D} \sum_{i=1}^{\infty} \alpha_i^2 = \frac{1}{\mu_1^D} \|u\|_{H(\Omega)}^2.$$

Enfin la capacité étant une fonction positive bornée nous pouvons écrire :

$$\int_{\Omega} c u^2 \geq \min_{\Omega} c(x) \int_{\Omega} u^2 = c^{\min} \|u\|_{L^2(\Omega)}^2.$$

Finalement en combinant ces résultats

$$c^{\min} \|u\|_{L^2(\Omega)}^2 \leq \frac{1}{\mu_1^D} \|u\|_{H(\Omega)}^2.$$

En prenant la racine de cette expression nous obtenons la proposition B.4.1

■

Notons que pour un pavé droit de \mathbb{R}^d les expressions analytiques des modes de Dirichlet sont connues. On a en particulier :

$$\mu_1^D = \pi^2 \left(\sum_{i=1}^d \frac{1}{l_i^2} \right).$$

B.5 Notion de trace sur le bord d'un domaine.

Dans nos démonstrations nous avons besoin de manipuler les fonctions u sur le bord du domaine $\partial\Omega$ (et aussi la dérivée normale $\nabla u \cdot \underline{n}$).

Définition B.5.1 On considère un ouvert Ω borné et régulier de classe C^1 . On définit l'application trace notée γ_0 :

$$\begin{aligned} \gamma_0 : H^1(\Omega) \cap C(\bar{\Omega}) &\longrightarrow L^2(\partial\Omega) \cap C(\bar{\partial\Omega}) \\ u &\longmapsto \gamma_0(u) = u|_{\partial\Omega}. \end{aligned} \tag{3.B.9}$$

△

Théorème: B.5.1 THÉORÈME DE LA TRACE.

L'application trace γ_0 se prolonge par continuité en une application linéaire continue de $H^1(\Omega)$ dans $L^2(\partial\Omega)$, notée encore γ_0 . En particulier, il existe une longueur constante notée $l_\alpha > 0$ telle que :

$$\|u\|_{L^2(\partial\Omega)} \leq \sqrt{l_\alpha} \|u\|_{H^1(\Omega)}, \quad \forall u \in H^1(\Omega). \quad (3.B.10)$$

▲

Le théorème de la trace donne un sens à la valeur d'une fonction de $H^1(\Omega)$ sur le bord $\partial\Omega$. Ce résultat n'est pas vrai pour les fonctions de $L^2(\Omega)$, autrement dit il n'y a pas de notion de trace pour des fonctions qui sont seulement de carré sommable. Le noyau de l'application trace est $\text{Ker}(\gamma_0) = H_0^1(\Omega)$. L'espace $H_0^1(\Omega)$ apparaît comme le sous espace de $H^1(\Omega)$ des fonctions qui s'annulent sur le bord. On note $\text{Im}(\gamma_0) = H^{1/2}(\partial\Omega)$ l'image de l'application trace. C'est un sous espace strict, dense dans $L^2(\partial\Omega)$. L'application trace joue un rôle fondamental.

Démonstration: B.5.1 La démonstration du théorème de la trace est relativement simple dans un domaine "semi infini". Soit Ω un ouvert régulier de \mathbb{R}^d semi infini dans une direction de l'espace. Pour fixer les idées nous choisissons la direction \underline{e}_1 et nous noterons $\Omega = \{x \in \mathbb{R}^d, x_1 > 0\}$. On considère une fonction $u \in C_c^\infty(\bar{\Omega})$ ($C_c^\infty(\bar{\Omega})$ est dense dans $H^1(\Omega)$).

$$\int_0^\infty \frac{\partial u^2}{\partial x_1} dx_1 = [u^2(x_1, \dots, x_d)]_{x_1=0}^\infty = u^2(\infty, \dots, x_d) - u^2(0, \dots, x_d) = -u^2(0, \dots, x_d),$$

puisque u est à support compact et s'annule à l'infini. De plus $\frac{\partial u^2}{\partial x_1} = 2u \frac{\partial u}{\partial x_1}$. Donc

$$|u(0, x_2, \dots, x_d)|^2 = -2 \int_0^\infty u(x_1, x_2, \dots, x_d) \frac{\partial u(x_1, x_2, \dots, x_d)}{\partial x_1} dx_1,$$

et en remarquant l'inégalité $a^2 + b^2 \geq 2ab$ nous obtenons

$$|u(0, x_2, \dots, x_d)|^2 \leq \int_0^\infty \left(|u(x_1, x_2, \dots, x_d)|^2 + \left| \frac{\partial}{\partial x_1} u(x_1, x_2, \dots, x_d) \right|^2 \right) dx_1.$$

Intégrons selon x_2, \dots, x_d

$$\int_{\mathbb{R}^{d-1}} |u(0, x_2, \dots, x_d)|^2 dx_2 \dots dx_d \leq \int_{\Omega} \left(|u(x)|^2 + \left| \frac{\partial}{\partial x_1} u(x) \right|^2 \right) dx$$

$$\leq \int_{\Omega} \left(|u(x)|^2 + \sum_{i=1}^d \left| \frac{\partial}{\partial x_i} u(x) \right|^2 \right) dx = \|u\|_{H^1(\Omega)}^2.$$

Mais ici $\mathbb{R}^{d-1} = \partial\Omega$ donc $\|u\|_{L^2(\partial\Omega)}^2 \leq \|u\|_{H^1(\Omega)}^2$. Enfin $C_c^\infty(\bar{\Omega})$ est dense dans $H^1(\Omega)$ d'où le résultat (avec $l_\gamma = 1$).

Pour généraliser le résultat à un ouvert régulier de classe C^1 il faut faire appel à un argument de cartes locales du bord. Cela permet de se ramener localement pour chaque carte au cas semi infini que nous venons de démontrer.

■

Notons qu'à partir du théorème de la trace il est possible de définir d'autres opérateurs de trace. En particulier on désigne par γ_1 l'opérateur de trace normale qui permet de parler de la valeur sur le bord $\partial\Omega$ de la dérivée normale $\nabla u \cdot \underline{n}$ du gradient d'une fonction u définie dans le domaine Ω .

Définition B.5.2 On considère encore un ouvert Ω borné régulier de classe C^1 . On définit l'application trace normale notée γ_1 :

$$\begin{aligned} \gamma_1 : H^2(\Omega) \cap C^1(\bar{\Omega}) &\longrightarrow L^2(\bar{\Omega}) \cap C(\partial\bar{\Omega}) \\ u &\longmapsto \gamma_1(u) = (\nabla u \cdot \underline{n})|_{\partial\Omega}. \end{aligned} \tag{3.B.11}$$

△

Théorème: B.5.2 TRACE NORMALE.

L'application trace normale γ_1 se prolonge par continuité en une application linéaire continue de $H^2(\Omega)$ dans $L^2(\partial\Omega)$, notée encore γ_1 . En particulier, il existe une constante positive telle que⁷ :

$$\|\nabla u \cdot \underline{n}\|_{L^2(\partial\Omega)} \leq C \|u\|_{H^2(\Omega)}, \quad \forall u \in H^2(\Omega). \tag{3.B.12}$$

▲

Démonstration: B.5.2 Voir par exemple [90]

■

7. $\|u\|_{H^2(\Omega)} = \|u\|_{L^2(\Omega)} + \|\nabla u\|_{L^2(\Omega)} + \|\nabla^2 u\|_{L^2(\Omega)}$

Les théorèmes de traces permettent de généraliser différentes formules de Green pour les fonctions de $H^1(\Omega)$, $H^2(\Omega)$.

Il est évident que l'application trace n'est pas bijective (deux fonction u et u' de $H^1(\Omega)$ peuvent avoir la même trace sans être égale presque partout). On peut chercher à "prolonger" une fonction définie sur le bord à l'intérieur du domaine. C'est par exemple le rôle de l'opérateur de relèvement harmonique que nous définissons à présent.

Définition B.5.3 Relèvement harmonique \mathcal{R}

$$\begin{aligned} \mathcal{R} : H^{1/2}(\partial\Omega) &\longrightarrow H^1(\Omega) \\ w &\longmapsto \overset{\Delta}{w} = \mathcal{R}(w), \\ &\left| \begin{array}{l} \Omega \quad -\nabla \cdot (\kappa \cdot \nabla \overset{\Delta}{w}) + \varepsilon c \overset{\Delta}{w} = 0, \\ \partial\Omega \quad \overset{\Delta}{w}|_{\partial\Omega} = w. \end{array} \right. \end{aligned} \quad (3.B.13)$$

△

Remarquons la présence du terme d'ordre zéro $\varepsilon c \overset{\Delta}{w}$ qui assure la coercivité de l'opérateur de relèvement harmonique.

Théorème: B.5.3 *L'opérateur $\mathcal{R} : L^2(\partial\Omega) \rightarrow L^2(\Omega)$ est continu pour toutes les solutions u du problème du Laplacien-Dirichlet défini par :*

$$\begin{aligned} \Omega \quad &-\nabla \cdot (\kappa \cdot \nabla u) + \varepsilon c u = \varpi, \\ \partial\Omega \quad &u|_{\partial\Omega} = 0, \end{aligned} \quad (3.B.14)$$

où u et ϖ satisfont respectivement⁸ : $u \in H^{3/2+\xi}(\Omega)$ et $\forall \varpi \in L^2(\Omega)$.

▲

Notons que du point de vue physique la sollicitation ϖ est une puissance volumique (adimensionnée ici).

8. **Théorème** : Si Ω est un ouvert borné régulier de classe C^1 , et si $m > d/2$, alors l'espace de Hilbert $H^m(\Omega)$ est un sous espace de l'ensemble $C(\bar{\Omega})$ des fonctions continues sur $\bar{\Omega}$. De plus par application réitérée de cette propriété à une fonction et à ses dérivées on a la proposition : si il existe un entier $k \geq 0$ tel que $m - d/2 > k$, alors $H^m(\Omega)$ est un sous-espace de l'ensemble $C^k(\bar{\Omega})$ des fonctions k fois différentiables sur $\bar{\Omega}$. ▲

Concrètement si nous posons $m = d/2 + \xi$ avec $\xi > 0$ alors nous pouvons choisir $k = 0$ et donc $H^{d/2+\xi}(\Omega) \subset C(\bar{\Omega})$. En géométrie bidimensionnelle on peut choisir $k = 1$ et dans ce cas $H^{1+\xi}(\Omega) \subset C^1(\bar{\Omega})$. Nous constatons que le problème bidimensionnel est plus régulier que le problème tridimensionnel.

Démonstration: B.5.3 Soit $v \in H^{1/2}(\partial\Omega)$. Son relèvement harmonique $\hat{v} = \mathcal{R}(v)$ est défini dans le sens usuel comme une fonction de l'espace de Hilbert $H^1(\Omega)$.

Introduisons la forme linéaire \mathcal{L} :

$$\mathcal{L}(\varpi) = \int_{\Omega} \hat{v} \varpi \quad \forall \varpi \in L^2(\Omega). \quad (3.B.15)$$

Cependant une simple intégration par partie (formule de Green) donne :

$$\int_{\Omega} \left(\hat{v} \nabla \cdot \kappa \cdot \nabla u - u \nabla \cdot \kappa \cdot \nabla \hat{v} \right) = \int_{\partial\Omega} \left(\hat{v} |_{\partial\Omega} \kappa \cdot \nabla u - u |_{\partial\Omega} \kappa \cdot \nabla \hat{v} \right) \cdot \underline{n}.$$

Mais u s'annule sur le bord d'après (3.B.14) et $\hat{v}|_{\partial\Omega} = v$.

De plus dans le domaine $\nabla \cdot \kappa \cdot \nabla u = \varepsilon c u - \varpi$ d'après (3.B.14) et $\nabla \cdot \kappa \cdot \nabla \hat{v} = \varepsilon c \hat{v}$ d'après (3.B.13) d'où

$$\int_{\Omega} \hat{v} \varpi = - \int_{\partial\Omega} v \kappa \cdot \nabla u \cdot \underline{n}. \quad (3.B.16)$$

Donc en utilisant (3.B.16) et l'inégalité de Schwarz nous avons :

$$|\mathcal{L}(\varpi)| \leq \kappa^{\max} \|v\|_{L^2(\partial\Omega)} \|\nabla u \cdot \underline{n}\|_{L^2(\partial\Omega)}.$$

Mais d'après le théorème de la trace normale : $\exists C' > 0$ telle que $\|\nabla u \cdot \underline{n}\|_{L^2(\partial\Omega)} \leq C' \|u\|_{H^{3/2+\xi}(\Omega)}$.

De plus puisque v dépend continument de ϖ , et grâce à l'hypothèse de régularité, le théorème du graphe fermé prouve que l'application $\mathcal{G}(\varpi)$:

$$\begin{aligned} \mathcal{G} : L^2(\Omega) &\longrightarrow H^{3/2+\xi}(\Omega) \\ \varpi &\longmapsto u. \end{aligned} \quad (3.B.17)$$

est continue. Donc $\exists C'' > 0$ telle que $\|u\|_{H^{3/2+\xi}(\Omega)} \leq C'' \|\varpi\|_{L^2(\Omega)}$. En résumé :

$$\exists C > 0 \text{ telle que } |\mathcal{L}(\varpi)| \leq C \|v\|_{L^2(\partial\Omega)} \|\varpi\|_{L^2(\Omega)}.$$

En utilisant la norme du relèvement :

$$\begin{aligned} \|\hat{v}\|_{L^2(\Omega)} \equiv \|\mathcal{R}(v)\|_{L^2(\Omega)} &= \sup_{\varpi \in L^2(\Omega)} \int_{\Omega} \hat{v} \varpi \leq \sup_{\varpi \in L^2(\Omega)} |\mathcal{L}(\varpi)| \leq \kappa^{\max} C' C'' \|v\|_{L^2(\partial\Omega)}. \\ \|\varpi\|_{L^2(\Omega)} = 1 & \qquad \qquad \qquad \|\varpi\|_{L^2(\Omega)} = 1 \end{aligned} \quad (3.B.18)$$

nous en déduisons l'existence d'une constante C sans dimension telle que :

$$\forall v \in H^{1/2}(\partial\Omega), \quad \|\hat{v}\|_{L^2(\Omega)} \leq C \|v\|_{L^2(\partial\Omega)}. \quad (3.B.19)$$

Enfin puisque $H^{1/2}(\partial\Omega) \subset L^2(\partial\Omega)$ et que $H^{1/2}(\partial\Omega)$ est dense dans $L^2(\partial\Omega)$, l'opérateur de relèvement harmonique \mathcal{R} peut être étendu par continuité en un opérateur de $L^2(\partial\Omega)$ dans $L^2(\Omega)$ ce qui achève la preuve. ■

B.6 Formules de Courant-Fischer.

Dans cette partie H et L désignent deux espaces de Hilbert réels de dimension infinie. On suppose que $H \subset L$ avec injection compacte et que H est dense dans L .

$a(\cdot, \cdot)$ et $c(\cdot, \cdot)$ sont deux formes bilinéaires symétriques. La forme $a(\cdot, \cdot)$ est continue et coercive sur H . La forme $c(\cdot, \cdot)$ est continue et positive sur L . Le problème aux valeurs propres sous forme variationnelle associé à cette paire d'opérateurs s'écrit :

$$\text{Trouver } (\mu, u) \in \mathbb{R} \times H \quad \text{tel que} \quad a(u, v) = \mu c(u, v) \quad \forall v \in H. \quad (3.B.20)$$

Le problème variationnel aux valeurs propres (3.B.20) admet une infinité dénombrable de solutions. On note $(\mu_k)_{k \geq 1}$ la suite croissante de ses valeurs propres. Dans cette suite chaque valeur propre est répétée avec son ordre de multiplicité :

$$\mu_1 \leq \mu_2 \leq \dots < \underbrace{\mu_{k-m_q} = \mu_{k-m_q+1} = \dots = \mu_k}_{m_q+1 \text{ fois}} < \dots$$

Les valeurs propres sont réelles et positives. Chaque valeur propre est associée à une fonction propre notée V_k . L'ensemble des fonctions propres forme une base orthogonale de l'espace H (resp. L) par rapport au produit scalaire défini par $a(\cdot, \cdot)$ (resp. $c(\cdot, \cdot)$).

Définition B.6.1 *Sous les hypothèses précédentes on définit le quotient de Rayleigh d'une fonction $u \in H \setminus \{0\}$ par :*

$$R(u) = \frac{a(u, u)}{c(u, u)}.$$

△

Le quotient de Rayleigh présente un grand intérêt pour l'étude théorique des valeurs propres du problème (3.B.20). En effet on peut voir le problème de recherche des valeurs propres comme une suite de minimisations du quotient de Rayleigh. A chaque nouvelle minimisation le nouveau mode propre doit être orthogonal aux modes trouvés aux étapes antérieures. Plus précisément nous avons le théorème :

Théorème: B.6.1 FORMULES DE COURANT-FISCHER

Pour tout $p \geq 1$ on note \mathcal{F}_p l'ensemble des sous espaces vectoriels de dimension p de H . Par

convention $\mathcal{F}_0 = \{0\}$. Pour tout $k \geq 1$ la k -ème valeur propre de (3.B.20) est donnée par les formules :

$$\mu_k = \min_{F \in \mathcal{F}_k} \left(\max_{u \in F \setminus \{0\}} R(u) \right) = \max_{F \in \mathcal{F}_{k-1}} \left(\min_{u \in F^\perp \setminus \{0\}} R(u) \right).$$

En particulier la première valeur propre vérifie :

$$\mu_1 = \min_{u \in H \setminus \{0\}} R(u).$$

▲

Démonstration: B.6.1 *Considérons une fonction $u \in F$. Elle est décomposable sur la base des fonctions propres de (3.B.20). Notons α_j ses coefficients de décomposition et V_j les modes propres⁹ :*

$$\forall u \in F, \quad u = \sum_{j=1}^{\infty} \alpha_j V_j.$$

Il découle des propriétés d'orthogonalité de la base propre les deux identités :

$$\forall u \in F, \quad a(u, u) = \sum_{j=1}^{\infty} \mu_j \alpha_j^2 c(V_j, V_j), \quad c(u, u) = \sum_{j=1}^{\infty} \alpha_j^2 c(V_j, V_j).$$

Nous poserons $\beta_j^2 = \alpha_j^2 c(V_j, V_j)$.

— Si $F_k = \text{Vect}(V_j)_{1 \leq j \leq k}$ désigne le sous espace de dimension k engendré par les k premiers modes propres on a les deux identités :

$$\forall u \in F_k, \quad a(u, u) = \sum_{j=1}^k \mu_j \beta_j^2, \quad c(u, u) = \sum_{j=1}^k \beta_j^2.$$

En tenant compte du fait que $\forall j \leq k$ on a $\frac{\mu_j}{\mu_k} \leq 1$ le quotient de Rayleigh d'un tel vecteur vérifie :

$$R(u) = \frac{\sum_{j=1}^k \mu_j \beta_j^2}{\sum_{j=1}^k \beta_j^2} = \mu_k \frac{\sum_{j=1}^k \frac{\mu_j}{\mu_k} \beta_j^2}{\sum_{j=1}^k \beta_j^2} \leq \mu_k.$$

Soit

$$\mu_k = \max_{u \in F_k \setminus \{0\}} R(u).$$

9. Pour plus de généralité nous ne supposons pas les modes normés par rapport à $c_\Omega(\cdot, \cdot)$ comme il est souvent d'usage lors de l'étude des valeurs propres du Laplacien-Dirichlet. En effet lors du processus de réduction des modèles par amalgame modal d'autres normes "énergétiques" se révèlent plus avantageuses.

- Si $F_{k-1} \in \mathcal{F}_{k-1}$ désigne le sous espace engendré par les $k-1$ premiers vecteurs propres $F_{k-1} = \text{Vect}(V_j)_{1 \leq j \leq k-1}$. La décomposition sur la base des vecteurs propres d'une fonction u de son espace orthogonal F_{k-1}^\perp s'écrit :

$$u = \sum_{j=k}^{\infty} \alpha_j V_j.$$

En remarquant qu'ici $\forall j \geq k$ on a $\frac{\mu_j}{\mu_k} \geq 1$ ceci permet d'écrire son quotient de Rayleigh :

$$R(u) = \frac{\sum_{j=k}^{\infty} \mu_j \beta_j^2}{\sum_{j=k}^{\infty} \beta_j^2} = \mu_k \frac{\sum_{j=k}^{\infty} \frac{\mu_j}{\mu_k} \beta_j^2}{\sum_{j=k}^{\infty} \beta_j^2} \geq \mu_k.$$

Soit

$$\mu_k = \min_{u \in F_{k-1}^\perp \setminus \{0\}} R(u).$$

- Considérons à présent un sous espace F quelconque de \mathcal{F}_k . Puisque $\dim F = k$ et $\dim F_{k-1} = k-1$ l'intersection $F \cap F_{k-1}^\perp$ n'est pas réduite à $\{0\}$. Par conséquent¹⁰ :

$$\max_{u \in F \setminus \{0\}} R(u) \geq \max_{u \in F \cap F_{k-1}^\perp \setminus \{0\}} R(u) \geq \min_{u \in F \cap F_{k-1}^\perp \setminus \{0\}} R(u) \geq \min_{u \in F_{k-1}^\perp \setminus \{0\}} R(u) = \mu_k.$$

Donc $\forall F \in \mathcal{F}_k$ $\mu_k \leq \max_{u \in F \setminus \{0\}} R(u)$ et puisque $\mu_k = \max_{u \in F_k \setminus \{0\}} R(u)$ nous en déduisons la première égalité de Courant-Fischer.

$$\mu_k = \min_{F \in \mathcal{F}_k} \left(\max_{u \in F \setminus \{0\}} R(u) \right).$$

- La démonstration de la seconde égalité est très semblable. Nous considérons cette fois un sous espace F quelconque de \mathcal{F}_{k-1} . A nouveau l'intersection $F^\perp \cap F_k$ n'est pas réduite à $\{0\}$. Par conséquent en utilisant des arguments d'inclusions semblables à ceux du cas précédent nous en déduisons :

$$\min_{u \in F^\perp \setminus \{0\}} R(u) \leq \min_{u \in F^\perp \cap F_k \setminus \{0\}} R(u) \leq \max_{u \in F^\perp \cap F_k \setminus \{0\}} R(u) \leq \max_{u \in F_k \setminus \{0\}} R(u) = \mu_k.$$

Donc $\forall F^\perp$ $\mu_k \geq \min_{u \in F^\perp \setminus \{0\}} R(u)$ et puisque $\mu_k = \min_{u \in F_{k-1}^\perp \setminus \{0\}} R(u)$ nous en déduisons la deuxième égalité de Courant-Fischer :

$$\mu_k = \max_{F \in \mathcal{F}_{k-1}} \left(\min_{u \in F^\perp \setminus \{0\}} R(u) \right).$$

■

10. Nous utilisons ici les deux inclusions $F \cap F_{k-1}^\perp \subset F$, $F \cap F_{k-1}^\perp \subset F_{k-1}^\perp$ et l'inégalité "triviale" $\max_{u \in F} \geq \min_{u \in F}$!

Comme conséquence des formules de Courant-Fischer examinons par exemple la relation entre les valeurs propres de deux domaines dont l'un est inclus dans l'autre.

Proposition B.6.1 *Soient Ω et Ω' deux ouverts réguliers de \mathbb{R}^d tels que $\Omega \subset \Omega'$. Soient $b_\omega(u, v)$ et $c_\omega(u, v)$ un couple d'opérateurs bilinéaires symétriques continu dans $H^1(\omega)$. L'opérateur différentiel $b_\omega(\cdot, \cdot)$ est supposé coercif sur $H^1(\omega)$ et $c_\omega(\cdot, \cdot)$ est positif sur $L^2(\omega)$. Nous supposons que :*

$$\text{si } \omega_1 \cap \omega_2 = \emptyset, \quad b_{\omega_1 \cup \omega_2}(\cdot, \cdot) = b_{\omega_1}(\cdot, \cdot) + b_{\omega_2}(\cdot, \cdot), \quad c_{\omega_1 \cup \omega_2}(\cdot, \cdot) = c_{\omega_1}(\cdot, \cdot) + c_{\omega_2}(\cdot, \cdot).$$

On associe à ce couple d'opérateurs les problèmes aux valeurs propres :

$$\text{Trouver } (u^D, \mu^D) \in H_0^1(\omega) \times \mathbb{R}^+ \quad b_\omega(u^D, v) = \mu^D(\omega) c_\omega(u^D, v) \quad \forall v \in H_0^1(\omega); \quad (3.B.21)$$

$$\text{Trouver } (u, \mu) \in H^1(\omega) \times \mathbb{R}^+ \quad b_\omega(u, v) = \mu(\omega) c_\omega(u, v) \quad \forall v \in H^1(\omega). \quad (3.B.22)$$

Sous ces conditions on a l'inégalité :

$$\forall n \in \mathbb{N}^*, \quad \mu_n(\Omega) \geq \mu_n^D(\Omega') \geq \mu_n(\Omega'). \quad (3.B.23)$$

▲

Preuve B.6.1 *Les hypothèses faites sur les opérateurs et les espaces de travail assurent pour (3.B.21) l'existence d'une infinité dénombrable de solutions propres réelles strictement positives.*

D'après la première formule de Courant-Fischer du théorème B.6.1, les modes propres du problème auxiliaire défini dans $H^1(\Omega)$ vérifient :

$$\mu_n(\Omega) = \min_{F \in \mathcal{F}_n(\Omega)} \left(\max_{u \in F \setminus \{0\}} \frac{b_\Omega(u, u)}{c_\Omega(u, u)} \right).$$

Ici $\mathcal{F}_n(\Omega)$ ($\forall n \geq 1$) désigne l'ensemble des sous espaces vectoriels de dimension n de $H^1(\Omega)$. Désignons par $\widetilde{\mathcal{F}_n(\Omega')}$ l'ensemble des sous espaces vectoriels de dimension n de $H_0^1(\Omega')$ formés des prolongements \tilde{u} des fonctions $u \in H^1(\Omega)$. Puisque dans $\Omega' \setminus \Omega$ le prolongement \tilde{u} est nul il est immédiat d'écrire :

$$\frac{b_\Omega(u, u)}{c_\Omega(u, u)} = \frac{b_\Omega(\tilde{u}, \tilde{u}) + b_{\Omega' \setminus \Omega}(\tilde{u}, \tilde{u})}{c_\Omega(\tilde{u}, \tilde{u}) + c_{\Omega' \setminus \Omega}(\tilde{u}, \tilde{u})} = \frac{b_{\Omega'}(\tilde{u}, \tilde{u})}{c_{\Omega'}(\tilde{u}, \tilde{u})}.$$

Désignons par $\mathcal{F}_n(\Omega')$ l'ensemble des sous espaces vectoriels de dimension n de $H_0^1(\Omega')$. L'ensemble $\widetilde{\mathcal{F}_n(\Omega')}$ ne contient que des sous espaces formés par des prolongements par zéro dans

$\Omega' \setminus \Omega$ des fonctions $u \in H^1(\Omega)$. Ainsi par construction ces fonctions sont de trace nulle sur $\partial\Omega'$. Donc $\widetilde{\mathcal{F}_n(\Omega')} \subset \mathcal{F}_n(\Omega') \subset H_0^1(\Omega') \subset H^1(\Omega')$. Nous en déduisons :

$$\begin{aligned} \mu_n(\Omega) &= \min_{\mathbf{F} \in \widetilde{\mathcal{F}_n(\Omega')}} \left(\max_{\tilde{u} \in \mathbf{F} \setminus \{0\}} \frac{b_{\Omega'}(\tilde{u}, \tilde{u})}{c_{\Omega'}(\tilde{u}, \tilde{u})} \right) \\ &\geq \min_{\mathbf{F} \in \mathcal{F}_n(\Omega') \subset H_0^1(\Omega')} \left(\max_{u \in \mathbf{F} \setminus \{0\}} \frac{b_{\Omega'}(u, u)}{c_{\Omega'}(u, u)} \right) = \mu_n^D(\Omega') \\ &\geq \min_{\mathbf{F} \in \mathcal{F}_n(\Omega') \subset H^1(\Omega')} \left(\max_{u \in \mathbf{F} \setminus \{0\}} \frac{b_{\Omega'}(u, u)}{c_{\Omega'}(u, u)} \right) = \mu_n(\Omega'). \end{aligned}$$

ce qui achève la preuve ■

Remarque B.6.1 Considérons des opérateurs de type $b_\Omega(u, v) = a_\Omega(u, v) + \varepsilon c_\Omega(u, v)$ où la forme bilinéaire $a_\Omega(u, v)$ n'est pas nécessairement coercive dans $H^1(\Omega)$. Introduisons les problèmes aux valeurs propres :

$$\text{Trouver } (u_\varepsilon, \mu(\varepsilon)) \in H^1(\Omega) \times \mathbb{R}^+ \quad b_\Omega(u_\varepsilon, v) = \begin{cases} \mu^D(\varepsilon, \Omega) c_\Omega(u_\varepsilon, v) & \forall v \in H_0^1(\Omega), \\ \mu(\varepsilon, \Omega) c_\Omega(u_\varepsilon, v) & \forall v \in H^1(\Omega). \end{cases}$$

$$\text{Trouver } (u, \mu(\Omega)) \in H^1(\Omega) \times \mathbb{R}^+ \quad a_\Omega(u, v) = \begin{cases} \mu^D(\Omega) c_\Omega(u, v) & \forall v \in H_0^1(\Omega), \\ \mu(\Omega) c_\Omega(u, v) & \forall v \in H^1(\Omega). \end{cases}$$

D'après ce qui précède nous avons l'inégalité $\mu(\varepsilon, \Omega) \geq \mu^D(\varepsilon, \Omega') \geq \mu(\varepsilon, \Omega')$, mais d'après la proposition A.1.3 nous savons que :

$$\mu_n(\Omega) = \mu_n(\varepsilon, \Omega) - \varepsilon.$$

Ces égalités étant vraies $\forall \varepsilon > 0$. Nous avons alors l'inégalité :

$$\mu_n(\Omega) + \varepsilon \geq \mu_n^D(\Omega') + \varepsilon \geq \mu_n(\Omega') + \varepsilon.$$

En soustrayant ε des deux membres de l'inégalité précédente nous trouvons finalement la relation générale pour les valeurs propres du couple d'opérateurs $a_\Omega(.,.)$ et $c_\Omega(.,.)$:

$$\mu_n(\Omega) \geq \mu_n^D(\Omega') \geq \mu_n(\Omega').$$

Ce résultat général est applicable aux valeurs propres de :

— *Dirichlet.* L'espace de travail est $H_0^1(\Omega)$. Ici $a_\Omega(u, v) = \int_{\Omega} \nabla u \cdot \kappa \cdot \nabla v$ est directement coercif dans $H_0^1(\Omega)$. L'opérateur de capacité s'écrit $c_\Omega(u, v) = \int_{\Omega} c u v$ positif dans $L^2(\Omega)$.

- *Fourier.* L'espace de travail est $H^1(\Omega)$. L'opérateur $a_\Omega(u, v) = \int_\Omega \nabla u \cdot \kappa \cdot \nabla v + \int_{\partial\Omega} h u v$ est directement coercif dans $H^1(\Omega)$. On a $c_\Omega(u, v) = \int_\Omega c u v$ positif dans $L^2(\Omega)$
- *Neumann.* L'espace de travail est $H^1(\Omega)$. Ici l'opérateur $a_\Omega(u, v) = \int_\Omega \nabla u \cdot \kappa \cdot \nabla v$ n'est plus coercif dans $H^1(\Omega)$. L'opérateur de capacité s'écrit encore $c_\Omega(u, v) = \int_\Omega c u v$ positif dans $L^2(\Omega)$.

En revanche le résultat ne s'applique pas aux modes de Branche. En effet l'opérateur de capacité à considérer est défini par $c_{\bar{\Omega}}(u, v) = \int_\Omega c u v + \int_{\partial\Omega} \zeta u v$. Cet opérateur comporte un terme surfacique et il est défini sur $L^2(\bar{\Omega})$ (et pas sur $L^2(\Omega)$). Lors du prolongement des fonctions de Ω dans Ω' on aura :

$$c_{\bar{\Omega}}(u, u) = c_\Omega(u, u) + c_{\partial\Omega}(u, u) = c_\Omega(\tilde{u}, \tilde{u}) + c_{\partial\Omega}(\tilde{u}, \tilde{u}) + c_{\Omega' \setminus \Omega}(\tilde{u}, \tilde{u}) + c_{\partial\Omega'}(\tilde{u}, \tilde{u}) \\ = c_{\bar{\Omega}'}(\tilde{u}, \tilde{u}) + c_{\partial\Omega}(\tilde{u}, \tilde{u}) \neq c_{\bar{\Omega}'}(\tilde{u}, \tilde{u}).$$

Ainsi le terme surfacique ne s'élimine pas (les fonctions u sont dans $H^1(\Omega)$ et pas dans $H_0^1(\Omega)$). Il en est d'ailleurs de même pour l'opérateur b avec :

$$b_\Omega(u, u) = b_{\Omega'}(\tilde{u}, \tilde{u}) + \varepsilon c_{\partial\Omega}(\tilde{u}, \tilde{u}) \neq b_{\Omega'}(\tilde{u}, \tilde{u}).$$

On se retrouve alors avec le rapport :

$$\frac{b_\Omega(u, u)}{c_{\bar{\Omega}}(u, u)} = \frac{b_{\Omega'}(\tilde{u}, \tilde{u}) + \varepsilon c_{\partial\Omega}(\tilde{u}, \tilde{u})}{c_{\bar{\Omega}'}(\tilde{u}, \tilde{u}) + c_{\partial\Omega}(\tilde{u}, \tilde{u})} \neq \frac{b_{\Omega'}(\tilde{u}, \tilde{u})}{c_{\bar{\Omega}'}(\tilde{u}, \tilde{u})}.$$

Et nous ne savons pas relier les deux derniers termes par une inégalité (dans un sens ou dans un autre !).

Les mêmes remarques s'appliquent pour les modes de Steklov puisque l'opérateur de capacité est défini cette fois uniquement sur la frontière donc dans $L^2(\partial\Omega)$.

B.7 Densité et Compacité.

Théorème: B.7.1 DE RELICH

On considère dans \mathbb{R}^d un ouvert Ω régulier et de classe C^1 . De toute suite bornée de $H^1(\Omega)$ on peut extraire une sous-suite convergente dans $L^2(\Omega)$. On dit que l'injection canonique $u \mapsto u$ de $H^1(\Omega)$ dans $L^2(\Omega)$ est un opérateur compact.

▲

Démonstration: B.7.1 [91] , [92]



On peut remplacer $H^1(\Omega)$ par $H_0^1(\Omega)$ dans le théorème de Rellich. Dans ce cas le théorème est vrai même si l'ouvert n'est pas régulier. Attention le théorème peut être faux si l'ouvert n'est pas borné.

On a aussi le résultat de compacité $H_0^1(\Omega) \times H^{1/2}(\partial\Omega) \subset L^2(\Omega) \times L^2(\partial\Omega)$ avec injection continue et compacte.

Enfin indiquons les résultats de densité suivants :

- $H^1(\Omega)$ est dense dans $L^2(\Omega)$,
- $H^{1/2}(\partial\Omega)$ est dense dans $L^2(\partial\Omega)$,
- $H_0^1(\Omega) \times H^{1/2}(\partial\Omega)$ est dense dans $L^2(\Omega) \times L^2(\partial\Omega)$.

B.8 Convergence.

Définition B.8.1

Soit $(E, \|\cdot\|_E)$ un espace pré-hilbertien. Soit $(u_n)_{n \in \mathbb{N}}$ une suite d'éléments de E . On dit que u_n converge

- fortement vers u dans E et on note $u_n \rightarrow u$ si et seulement si

$$\lim_{n \rightarrow \infty} \|u_n - u\|_E = 0,$$

- faiblement vers u dans E et on note $u_n \rightharpoonup u$ si et seulement si

$$\forall v \in E \quad \lim_{n \rightarrow \infty} (u_n - u | v)_E = 0.$$



Par l'inégalité de Cauchy-Schwarz on peut écrire : $|(u_n - u | v)_E| \leq \|u_n - u\|_E \|v\|_E$.
Donc la convergence forte entraîne la convergence faible.

B.9 Propriétés du tenseur de conductivité thermique.

Commençons par rappeler ce qu'est une matrice définie positive.

Définition B.9.1 Soit $\mathbb{A} \in \mathcal{M}_d(\mathbb{R})$ ¹¹ (resp. $\underline{v} \in \mathcal{E}_d(\mathbb{R})$ ¹²) une matrice carrée (resp. un vecteur) de dimension d , à coefficients réels. La matrice \mathbb{A} est dite symétrique définie positive si elle est symétrique et si :

$$\forall \underline{v} \in \mathcal{E}_d(\mathbb{R}) \setminus \{0\}, \quad {}^t \underline{v} \cdot \mathbb{A} \cdot \underline{v} > 0.$$

△

Proposition B.9.1 On considère un domaine \mathcal{D} de \mathbb{R}^d et une base orthonormée $\{\underline{e}_i\}_{1 \leq i \leq d}$ de $\mathcal{E}_d(\mathbb{R})$. On désigne par $K \in \mathcal{M}_d(\mathbb{R})$ la matrice de conductivité thermique dont les coefficients sont notés K_{ij} . On suppose que $K_{ij} \in L^\infty(\mathcal{D})$.

Dans ces conditions la matrice de conductivité thermique K :

- est symétrique et définie positive.
- ses coefficients vérifient $K_{ii} > 0$; $K_{ij} < \sqrt{K_{ii} K_{jj}}$ $j \neq i$; $\det K > 0$.
- admet des valeurs propres $K_{oi} > 0$
- $\exists K^{\min}, K^{\max} \in \mathbb{R}^+ \setminus \{0\}$ tels que :

$$K^{\min} \|\underline{v}\|_{L^2(\mathcal{D})}^2 \leq \int_{\mathcal{D}} {}^t \underline{v} \cdot K \cdot \underline{v} \leq K^{\max} \|\underline{v}\|_{L^2(\mathcal{D})}^2 \quad \forall \underline{v} \in \mathcal{E}_d(\mathbb{R}).$$

▲

Preuve B.9.1 La démonstration de la proposition B.9.1 repose à la fois sur des lois physiques (thermodynamique, physique statistique) et sur des propriétés mathématiques. Désignons par T le champ de la température à valeur réelle strictement positive dans un domaine $\mathcal{D} \in \mathbb{R}$. Nous supposons $T \in H^1(\mathcal{D})$. La génération interne d'entropie en fonction du temps dans le domaine Ω soumis à un gradient de température noté ∇T vérifie :

$$\frac{dS_i}{dt} = \int_{\mathcal{D}} \underline{q} \cdot \nabla \left(\frac{1}{T} \right). \quad (3.B.24)$$

Dans cette expression \underline{q} désigne la densité surfacique du flux de chaleur¹³. La loi de Fourier, bien vérifiée par l'expérience dans le domaine qui nous intéresse ici, suppose que le

11. $\mathcal{M}_d(\mathbb{R})$ désigne l'espace vectoriel des matrices carrées de dimension d à coefficients réels

12. $\mathcal{E}_d(\mathbb{R})$ désigne l'espace vectoriel des vecteurs de dimension d à coefficients réels

13. En thermodynamique des processus irréversibles on représente la variation d'entropie comme la somme de produits de forces (qui produisent un processus irréversible) par des flux qui décrivent la réponse du système à ces forces [104](p.74). Par exemple l'expérience commune montre qu'un déséquilibre de la température engendre un flux de chaleur. C'est un effet directe. Il existe aussi des effets croisés. Par exemple le flux de chaleur peut être engendré par une différence de potentiel électrique (effet thermoélectrique Peletier). Si \underline{q} est un "flux" alors $\nabla \frac{1}{T}$ est la "force généralisée" qui a engendré ce flux.

vecteur densité de flux de chaleur \underline{q} est relié au gradient de la température par la formule :

$$\underline{q} = -K \cdot \nabla T. \quad (3.B.25)$$

Dans l'expression (3.B.25) K désigne un tenseur du second ordre de dimension d à coefficients réels. Ses coefficients, que nous supposons bornés, dépendent de l'espace (si le domaine Ω n'est pas homogène) et éventuellement de la température (conduction non linéaire). Par abus de notation nous noterons aussi $K = (K_{ij})_{1 \leq i, j \leq d}$ la matrice des conductivités thermiques ($K \in \mathcal{M}_d(\mathbb{R})$). Les lois de réciprocité d'Onsager [105] déduites de la physique statistique imposent de plus la symétrie¹⁴ du tenseur des conductivités thermiques $K_{ij} = K_{ji}$. Enfin la cristallographie, grâce à l'étude de l'organisation spatiale des atomes, permet de montrer que certains termes extra-diagonaux peuvent être nuls car ils doivent satisfaire en plus une relation de symétrie de type $K_{ij} = -K_{ji}$.

En remarquant que $\nabla \left(\frac{1}{T}\right) = -\frac{1}{T^2} \nabla T$ nous obtenons une nouvelle expression de la génération interne d'entropie :

$$\frac{dS_i}{dt} = \int_{\mathcal{D}} \frac{1}{T^2} \nabla T \cdot K \cdot \nabla T. \quad (3.B.26)$$

Mais le second principe de la thermodynamique affirme que la génération interne d'entropie liée au déséquilibre thermique (traduit par un gradient de température non nul) est strictement positive [104].

Ainsi $\frac{dS_i}{dt} > 0$ entraîne $\nabla T \cdot K \cdot \nabla T > 0 \quad \forall \nabla T \neq \underline{0}$.

La matrice K est donc définie positive d'après la définition B.9.1. Ce type de matrice vérifie de nombreuses propriétés. En particulier tous ses mineurs principaux sont strictement positifs et notamment (si $d \leq 3$) :

- mineurs d'ordre un : $1 \leq i \leq d, \quad K_{ii} > 0,$
- mineurs d'ordre deux : $1 \leq i, j \leq d \quad j \neq i, \quad K_{ii} K_{jj} - K_{ij}^2 > 0,$
- mineur d'ordre trois : $\det K > 0.$

Le théorème spectral pour les matrices prouve que toute matrice réelle symétrique est diagonalisable. Si de plus la matrice est définie positive alors toutes ses valeurs propres sont strictement positives. Ainsi si nous notons $(K_{oi})_{1 \leq i \leq d}$ la suite des valeurs propres¹⁵ de la matrice de conductivité thermique alors $K_{oi} > 0$. Il existe une matrice orthogonale de changement

14. On suppose qu'il n'y a pas d'effets électromagnétiques. Dans les matériaux purement conducteurs où n'existe que l'effet Joule cette hypothèse reste valable [104] (p.178,p.181).

15. On parle aussi pour la valeur propre K_{oi} de conductivités principales dans la direction \underline{e}_i .

de base $\mathbb{P} \in \mathcal{M}_d(\mathbb{R})$ telle que :

$${}^t\mathbb{P}\mathbb{P} = \mathbb{I}_d, \quad {}^t\mathbb{P}K\mathbb{P} = (K_{oi} \delta_i^j)_{1 \leq i, j \leq d}.$$

En particulier $\forall \underline{v}, \exists \underline{w}$ tel que $\underline{v} = \mathbb{P} \cdot \underline{w}$ et $|\underline{v}| = |\underline{w}|$.

En effet $|\underline{v}|^2 = {}^t\underline{v} \cdot \underline{v} = {}^t\underline{w} \cdot {}^t\mathbb{P}\mathbb{P} \cdot \underline{w} = {}^t\underline{w} \cdot \underline{w} = |\underline{w}|^2$.

De plus :

$$\int_{\mathcal{D}} {}^t\underline{v} \cdot K \cdot \underline{v} = \int_{\mathcal{D}} {}^t\underline{w} \cdot {}^t\mathbb{P}K\mathbb{P} \cdot \underline{w} = \int_{\mathcal{D}} \sum_{i=1}^d K_{oi} w_i^2.$$

Les composantes v_i d'un vecteur $\underline{v} = \sum_{i=1}^d v_i \underline{e}_i$ sont de carré sommable : $v_i(\mathbf{x}) \in L^2(\mathcal{D})$.

Nous pouvons écrire $|\underline{w}|^2 = \sum_{i=1}^d w_i^2$ et $\|\underline{w}\|_{L^2(\mathcal{D})}^2 = \int_{\mathcal{D}} |\underline{w}|^2$. Donc

$$\exists C > 0 \text{ telle que } \int_{\mathcal{D}} \sum_{i=1}^d K_{oi} w_i^2 \leq C \int_{\mathcal{D}} \sum_{i=1}^d w_i^2 = C \|\underline{w}\|_{L^2(\mathcal{D})}^2 = C \|\underline{v}\|_{L^2(\mathcal{D})}^2.$$

Nous pouvons choisir pour C la plus grande conductivité principale dans le domaine, autrement dit :

$$K^{\max} = \max_{1 \leq i \leq d; \mathbf{x} \in \mathcal{D}} K_{oi}(\mathbf{x}).$$

Nous pouvons procéder de la même manière pour trouver un minorant.

Cette fois nous poserons $K^{\min} = \min_{1 \leq i \leq d; \mathbf{x} \in \Omega} K_{oi}(\mathbf{x})$. Finalement nous trouvons l'encadrement :

$$K^{\min} \|\underline{v}\|_{L^2(\mathcal{D})}^2 \leq \int_{\mathcal{D}} {}^t\underline{v} \cdot K \cdot \underline{v} \leq K^{\max} \|\underline{v}\|_{L^2(\mathcal{D})}^2.$$

ce qui achève la preuve de la proposition B.9.1. Cette proposition est essentielle pour démontrer la coercivité de l'opérateur de la chaleur.

Dans la pratique nous utiliserons cet encadrement pour $u \in H^1(\mathcal{D})$ en posant $\underline{v} = \nabla u$ et avec $|\nabla u| \in L^2(\mathcal{D})$.

Notons pour finir que si le domaine \mathcal{D} est homogène et isotrope alors on a :

$$\forall \mathbf{x} \in \mathcal{D} \quad K(\mathbf{x}) = K_o \left(\delta_i^j \right)_{1 \leq i, j \leq d}.$$

Le tenseur de conductivité est sphérique et on peut le réduire à un scalaire K_o . Pour ce cas particulier :

$$K^{\min} = K^{\max} = K_o \quad \text{et} \quad \int_{\mathcal{D}} {}^t \underline{v} \cdot K \cdot \underline{v} = K_o \|\underline{v}\|_{L^2(\mathcal{D})}^2.$$

Dans la pratique l'hypothèse d'isotropie est couramment admise pour de nombreux matériaux car elle permet de bien retranscrire les phénomènes thermiques observés.

■

Annexe C

Proposition d'un algorithme d'amalgame pour la base de Dirichlet-Steklov

Principe :

Nous proposons un algorithme qui construit progressivement la partition de l'espace des modes de Dirichlet-Steklov. A chaque étape les modes esclaves sont affectés à l'espace d'amalgame de leur famille respective où ils contribuent le moins à l'énergie de l'écart de réduction de la réponse en température (entre le modèle complet et le modèle réduit). A l'issue d'une étape le mode esclave (toutes familles confondues) qui contribue le plus cette fois à l'énergie de l'écart de réduction de la réponse en température est choisi comme nouveau mode maître. Il est alors seul dans son sous espace d'amalgame avant une nouvelle répartition de modes esclaves.

— *Etape 0. Initialisation :*

On fixe un seuil de tolérance ϵ pour l'énergie de l'écart de la réponse en température¹. Pour chaque famille Dirichlet (D) et Steklov (S), le mode "plat" étant exclu de la procédure : On fixe le nombre maximum de modes amalgamés $\tilde{N}_{max}^{\mathcal{X}}$

— *0.1 Calcul des états de référence $x_i^{\mathcal{X}}(t)$ avec $i \in \{1, \dots, N^{\mathcal{X}}\}$. Projection directe ou résolution de l'équation d'état complète linéarisée.*

— *0.2 Calcul des dominances modales $\mathcal{D}_i^{\mathcal{X}}$.*

— *0.3 Choix comme mode maître du mode le plus énergétique, autrement dit celui*

1. Ce choix est certainement le point le plus délicat de l'approche proposée. La norme de l'écart de réduction est exprimée dans l'espace d'énergie. Mais il ne s'agit pas d'une énergie au sens thermodynamique. On fixera un écart quadratique en température $\varepsilon_{\tilde{N}}$ qui peut être traduit en une norme de l'écart dans $L^2(\Omega)$ défini par

$$\varepsilon_{\tilde{N}} = \sqrt{\frac{\|\Delta T\|_{L^2(\Delta t \times \mathcal{D})}^2}{\mathcal{V}_{oi}(\mathcal{D}) \Delta t}}.$$

dont la contribution $\mathcal{D}_i^{\mathcal{X}}$ à l'énergie de la réponse en température est la plus forte. Il vérifie :

$$\mathcal{D}_i^{\mathcal{X}} = \max_{k \in \mathcal{P}_1^{\mathcal{X}}} (\mathcal{D}_k^{\min \mathcal{X}}).$$

Affectation du premier indice $k(1, 1) = i$ dans la partition $\mathcal{P}_1^{\mathcal{X}}$.

Les autres modes sont considérés comme esclaves et leurs indices sont ajoutés à la suite dans la partition $\mathcal{P}_1^{\mathcal{X}} = \{k(1, 1), k(1, 2), \dots, k(1, N^{\mathcal{X}})\}$.

Pour chaque mode esclave.

— 0.31 Calcul du minimum de la dominance modale de l'écart de réduction $\Delta \mathcal{D}_k^{\min \mathcal{X}}$ des modes esclaves affectés au mode maître $k \in \mathcal{P}_1^{\mathcal{X}}$.

— 0.32 Mise à jour de la dimension des partitions :

$$\tilde{N}^{\mathcal{X}} = 1$$

$$N_1^{\mathcal{X}} = N^{\mathcal{X}}$$

— 0.4 Choix, toutes familles confondues, comme nouveau mode maître du mode esclave apportant la plus grande contribution à l'énergie de l'écart de la réponse en température. Ce mode vérifie :

$$\Delta \mathcal{D}_{1,q}^{\min \mathcal{Y}} = \max_{\substack{\mathcal{X} \in \{\text{D}, \text{S}\} \\ k \in \mathcal{P}_1^{\mathcal{X}}}} (\Delta \mathcal{D}_k^{\min \mathcal{X}}).$$

Il est de type \mathcal{Y} (D ou S) et a le rang q dans la partition $\mathcal{P}_1^{\mathcal{Y}}$

Affectation de son indice $k(1, q)$ dans la nouvelle partition $\mathcal{P}_2^{\mathcal{Y}} = \{k(2, 1)\}$ tel que $k(2, 1) = k(1, q)$.

Mise à jour de la dimension des partitions :

$$\tilde{N}^{\mathcal{Y}} = 2$$

$$N_1^{\mathcal{Y}} = N^{\mathcal{Y}} - 1 \quad \text{et} \quad N_2^{\mathcal{Y}} = 1$$

Passage à l'étape $n = 1$.

— Etape n .

Nous ajoutons un nouvel espace d'amalgame ne comportant pour l'instant que le mode maître sélectionné à l'étape $n - 1$. Nous supposons qu'il est de type \mathcal{Y} et que son ordre dans la base amalgamée est $J = \tilde{N}^{\mathcal{Y}}$. En dehors de ce nouveau sous espace d'amalgame qui ne comporte que son mode maître, chaque mode esclave est affecté dans l'un des sous espace d'amalgame de Dirichlet ou de Steklov où sa contribution à l'énergie de l'écart de réduction à la réponse en température $\|\Delta T\|_{L^2(\Delta t; H^1(\Omega))}^2$ est minimum.

— 1.1 Pour chaque mode esclave de type \mathcal{Y} de chaque sous espace d'amalgame d'ordre

$I < \tilde{N}^{\mathcal{Y}}$ et de rang $p > 1$:

- 1.11 Calcul du facteur d'amalgame $\alpha_{J,q}^{opt\mathcal{Y}}$ si ce mode était amalgamé au nouveau mode maître d'ordre J (au dernier rang noté $q > 1$).
- 1.12 Calcul du minimum de la dominance modale de l'écart de réduction pour le nouveau sous espace d'amalgame

$$\Delta \mathcal{D}_{J,q}^{\min \mathcal{Y}}$$

- 1.13 Comparaison de la contribution à l'énergie de l'écart de réduction de la réponse en température $\Delta \mathcal{D}_{J,q}^{\min \mathcal{Y}}$ avec celle que le mode génère quand il est affecté à l'espace d'amalgame actuel $\Delta \mathcal{D}_{I,p}^{\min \mathcal{Y}}$

$$SI \quad \Delta \mathcal{D}_{J,q}^{\min \mathcal{Y}} < \Delta \mathcal{D}_{I,p}^{\min \mathcal{Y}}$$

- ALORS La nouvelle contribution est inférieure à la contribution actuelle

* Le mode esclave change de maître.

* Réorganisation des partitions.

On ajoute l'indice $k(J, q) = k(I, p)$ à la partition $\mathcal{P}_J^{\mathcal{Y}}$.

La partition $\mathcal{P}_J^{\mathcal{Y}}$ comporte à présent q modes. $N_J^{\mathcal{Y}} = N_J^{\mathcal{Y}} + 1 = q$

On élimine cet indice de la partition $\mathcal{P}_I^{\mathcal{Y}}$. $N_I^{\mathcal{Y}} = N_I^{\mathcal{Y}} - 1$

- SINON le mode esclave reste attaché à son maître actuel.

- 1.14 Passage au mode esclave suivant.

- 1.2 Une fois l'ensemble des modes esclaves de type \mathcal{Y} affectés à leur maître :

$$SI \quad \|\Delta T\|_{L^2([0, \Delta t]; H^1(\Omega))} \leq \epsilon, \text{ et (ou) } \tilde{N}^{\mathcal{X}} = \tilde{N}_{max}^{\mathcal{X}}.$$

- ALORS.

* Ordre du modèle $\tilde{N} = \tilde{N}^D + \tilde{N}^S$

* Construction des combinaisons linéaires des modes

* Normalisation des modes amalgamés

* Détermination de l'écart en température $\varepsilon_{\tilde{N}}$ (basée sur la norme $L^2(\Omega)$)

Ce calcul est facultatif. Par construction on est assuré que cet écart sera nécessairement inférieur à l'objectif fixé initialement $\varepsilon_{\tilde{N}}^{objc} \leq \varepsilon^{objc}$.

* Fin.

- SINON.

* Choix comme nouveau mode maître du mode esclave apportant la plus grande

contribution à l'énergie de l'écart de réduction de la réponse en température, toutes familles confondues. Ce mode vérifie :

$$\Delta \mathcal{D}_{K,m}^{\min \mathcal{Z}} = \max_{\substack{\mathcal{X} \in \{D,S\} \\ I \in \{1, \tilde{N}^{\mathcal{X}}\} \\ k \in \{1, \dots, N_I^{\mathcal{X}}\}}} (\Delta \mathcal{D}_{I,k}^{\min \mathcal{X}}).$$

Il est de type \mathcal{Z} (D ou S) et a le rang m dans la partition $\mathcal{P}_K^{\mathcal{Z}}$

* Ajout d'un sous espace d'amalgame $\tilde{N}^{\mathcal{Z}} = \tilde{N}^{\mathcal{Z}} + 1$

* Mise à jour de la dimensions des partitions :

Affectation de son indice $k(K, m)$ dans la nouvelle partition $\mathcal{P}_{\tilde{N}^{\mathcal{Z}}}^{\mathcal{Z}} = \{k(\tilde{N}^{\mathcal{Z}}, 1)\}$

tel que $k(\tilde{N}^{\mathcal{Z}}, 1) = k(K, m)$.

$$N_{\tilde{N}^{\mathcal{Z}}}^{\mathcal{Z}} = 1$$

Élimination du mode dans la partition $\mathcal{P}_K^{\mathcal{Z}}$. $N_K^{\mathcal{Z}} = N_K^{\mathcal{Z}} - 1$

* Passage à l'étape $n + 1$.

Remarque C.0.1 La procédure d'amalgame pour la base de Dirichlet-Steklov présente de grandes similitudes avec celle déjà utilisée pour la base de branche. Toutefois la dominance a été construite à l'aide d'un produit scalaire différent. Rappelons que pour les modes de branche le produit scalaire utilisé est :

$$(u | v)_{L^2(\bar{\Omega})} = c_{\bar{\Omega}}(u, v) = \int_{\Omega} c u v + \int_{\partial\Omega} \zeta u|_{\partial\Omega} v|_{\partial\Omega}. \quad (3.C.1)$$

Les modes propres de branche sont orthogonaux pour ce produit scalaire et ils sont normalisés par rapport à la norme induite. L'espace d'énergie associé aux modes de branche est $L^2(]0, t[\times \bar{\Omega})$.

Dans notre approche de réduction de la base de Dirichlet-Steklov nous avons réalisé un amalgame par famille, Dirichlet d'un côté et Steklov de l'autre côté. En réalité il est possible de s'affranchir de cette contrainte puisque elle ne remet pas en cause la relation d'orthogonalité des modes ainsi amalgamés. Toutefois en procédant ainsi on va mélanger des modes de frontière et des modes de domaine avec pour conséquence une perte de richesse de la base ainsi obtenue. Bien d'autres stratégies pour construire la base amalgamée, et plus particulièrement la partition, sont envisageables. Par exemple nous pouvons commencer par un amalgame des modes de Steklov en utilisant cette fois pour "l'espace d'énergie" $L^2(\partial\Omega)$. On contrôle alors la précision sur le bord. En retranchant le relèvement harmonique de cette contribution "glissante" de la réponse en température nous obtenons la réponse dynamique qui est décomposable sur la base de Dirichlet. On applique alors la procédure d'amalgame mais avec $H = L^2(\Omega)$. L'intérêt de procéder ainsi réside dans l'interprétation physique des énergies des écarts si la

capacité et le coefficient de Steklov sont constants sur le domaine (l'énergie de l'écart est alors proportionnelle à l'erreur quadratique sur la température). Il reste à écrire cette procédure pour vérifier sa faisabilité. Soulignons que ces différentes approches de l'amalgame ne conduisent pas a priori aux mêmes partitions de l'espace et donc pas au même minimum local. C'est seulement au prix de nombreux tests numériques qu'il serait possible de départager ces méthodes en ne perdant pas de vue au delà de la précision l'objectif de la réduction du temps de calcul.

Annexe D

Lien entre la norme $\|\cdot\|_{H(\Omega)}$ et la norme dans $L^2(\Omega)$

Dans le processus d'amalgame, l'écart (noté $|\tilde{u}_N - u|$) entre le modèle de référence et le modèle réduit à l'ordre N est minimisé dans l'espace d'énergie lié à la norme $\|\cdot\|_{H(\Omega)}$. Dans cet espace la méthode converge. Pour pouvoir estimer a priori l'erreur de réduction en terme d'écart quadratique sur les températures il est utile de relier $\|\cdot\|_{H(\Omega)}$ à la norme $\|\cdot\|_{L^2(\Omega)}$ usuelle. C'est le sens de la proposition suivante qui montre que la convergence dans $H^1(\Omega)$ entraîne la convergence dans $L^2(\Omega)$ (mais aussi dans $L^1(\Omega)$).

D.1 Majoration de la norme dans $L^2(\Omega)$.

Proposition D.1.1 *Soit la fonction $u \in H^1(\Omega)$ où Ω est un ouvert régulier de \mathbb{R}^d . Alors :*

$$\forall u \in H^1(\Omega), \exists C_* > 0 \text{ telle que } \frac{1}{\sqrt{\mathcal{V}_{ol}(\Omega)}} \|u\|_{L^1(\Omega)} \leq \|u\|_{L^2(\Omega)} \leq C_* \|u\|_{H(\Omega)}, \quad (3.D.1)$$

$$\text{avec } \mathcal{V}_{ol}(\Omega) = \int_{\Omega} dv^1 \quad ; \quad C_* \leq C_o = \frac{l^{\min}}{\sqrt{\kappa^{\min}}} \quad ; \quad \mu_o = \frac{\kappa^{\min}}{l^{\min} \zeta^{\min}}. \quad (3.D.2)$$

▲

Cette minimisation fixe le facteur de pondération μ_o entre le terme surfacique et le terme volumique dans le produit scalaire $(\cdot|\cdot)_{H(\Omega)}$. Notons que nous n'avons trouvé qu'un majorant

1. Dans cette annexe (à l'exception de la dernière section D.4), la mesure d'intégration est précisée $\left(\int_{\Omega} f \equiv \int_{\Omega} f dv \text{ et } \int_{\partial\Omega} f \equiv \int_{\partial\Omega} f da\right)$

de la constante optimale C_* . Pour des corps “pleins” cette majoration donne une bonne approximation de C_* . En revanche pour des corps “creux” (comme un tube, un sphère creuse, ...) la valeur C_o peut être loin de la valeur optimale C_* . Ici il faut reprendre la démonstration dans chaque cas afin de trouver la “bonne” longueur l^{\min} . Par exemple pour des coques minces $l^{\min} \sim 2e$ où e est l'épaisseur de la coque.

Preuve D.1.1 La relation entre les normes $\|\cdot\|_{L^1}$ et $\|\cdot\|_{L^2}$ résulte de l'application de l'inégalité de Schwarz.

$$\int_{\Omega} |u w| dv \leq \sqrt{\int_{\Omega} |u|^2 dv} \sqrt{\int_{\Omega} |w|^2 dv}.$$

Soit en prenant $w = 1$

$$\|u\|_{L^1(\Omega)} = \int_{\Omega} |u| dv = \int_{\Omega} |u| 1 dv \leq \sqrt{\int_{\Omega} |u|^2 dv} \sqrt{\int_{\Omega} 1 dv} = \|u\|_{L^2(\Omega)} \sqrt{\mathcal{V}_{ol}(\Omega)}$$

La preuve du résultat pour le lien entre les normes $\|\cdot\|_{L^2}$ et $\|\cdot\|_H$ utilise la même démarche que celle utilisée dans la démonstration constructive de l'inégalité de Poincaré dans $H^1(\Omega)$ donnée au paragraphe B.4 (les constantes physiques n'apparaissent pas). Comme précédemment pour le cas $H_0^1(\Omega)$ nous faisons l'hypothèse que le domaine Ω est compris dans la bande $|x_1| < l_1/2$ (quitte à faire un changement de repère avec une translation pour recentrer le domaine et une rotation pour se ramener à x_1). On note $\underline{n} = n_i \underline{e}_i$ la normale unitaire en un point σ de la frontière $\partial\Omega$ orientée vers l'extérieur du domaine Ω .

$$\begin{aligned} \|u\|_{L^2(\Omega)}^2 &= \int_{\Omega} u^2(x) dv = \int_{\Omega} u^2(x) \frac{\partial x_1}{\partial x_1} dv && \text{car } \frac{\partial x_1}{\partial x_1} = 1 \\ &= \int_{\partial\Omega} \sigma_1 n_1 u^2(\sigma) da - \int_{\Omega} x_1 \frac{\partial u^2(x)}{\partial x_1} dv && \text{par la formule de Stokes} \\ &= \int_{\partial\Omega} \sigma_1 n_1 u^2(\sigma) da - 2 \int_{\Omega} x_1 u(x) \frac{\partial u(x)}{\partial x_1} dv && \text{en dérivant } u^2(x) \\ &\leq \int_{\partial\Omega} |\sigma_1| |n_1| u^2(\sigma) da + 2 \int_{\Omega} |x_1| |u(x)| \frac{\partial u(x)}{\partial x_1} |dv. \end{aligned}$$

Examinons successivement les deux termes du membre de droite de cette inégalité. Nous remarquons que les composantes de la normale vérifient $\forall i, |n_i| \leq 1$. De plus $|x_1| < l_1/2$ et $\zeta^{\min} \leq \zeta \leq \zeta^{\max}$. On a donc pour le premier terme :

$$\int_{\partial\Omega} |\sigma_1| |n_1| u^2(\sigma) da \leq \frac{l_1}{2} \int_{\partial\Omega} u^2(\sigma) da \leq \frac{l_1}{2\zeta^{\min}} \int_{\partial\Omega} \zeta u^2 da.$$

Pour le deuxième terme en remarquant l'inégalité $|ab| \leq \frac{1}{2}(\epsilon a^2 + \frac{b^2}{\epsilon})$ qui est vraie $\forall \epsilon > 0$ nous obtenons :

$$2 \int_{\Omega} |x_1| |u(x)| \frac{\partial u(x)}{\partial x_1} dv \leq l_1 \int_{\Omega} |u(x)| \frac{\partial u(x)}{\partial x_1} dv \leq \frac{l_1}{2\epsilon} \int_{\Omega} |u(x)|^2 dv + \epsilon \frac{l_1}{2} \int_{\Omega} \left| \frac{\partial u(x)}{\partial x_1} \right|^2 dv.$$

Donc en combinant ces différentes inégalités et en posant $\epsilon = \alpha l_1$ avec $\alpha > 1$:

$$\frac{2\alpha - 1}{2\alpha} \|u\|_{L^2(\Omega)}^2 \leq \frac{l_1}{2\zeta^{\min}} \int_{\partial\Omega} \zeta u^2 da + \frac{\alpha l_1^2}{2} \int_{\Omega} \left| \frac{\partial u(x)}{\partial x_1} \right|^2 dv.$$

De plus $\left| \frac{\partial u(x)}{\partial x_1} \right|^2 \leq \sum_{i=1}^d \left| \frac{\partial u(x)}{\partial x_i} \right|^2 = |\nabla u|^2$. Enfin, le tenseur de conductivité étant défini positif, $\int_{\Omega} |\nabla u|^2 dv \leq \frac{1}{\kappa^{\min}} \int_{\Omega} \nabla u \cdot \kappa \cdot \nabla u da$. Donc :

$$\begin{aligned} \|u\|_{L^2(\Omega)}^2 &\leq \frac{l_1}{2} \frac{2\alpha}{2\alpha - 1} \left(\frac{1}{\zeta^{\min}} \int_{\partial\Omega} \zeta u^2 da + \frac{\alpha l_1}{\kappa^{\min}} \int_{\Omega} \nabla u \cdot \kappa \cdot \nabla u dv \right) \\ &\leq l_1 \max \left(\frac{\alpha}{2\alpha - 1} \frac{1}{\mu_o \zeta^{\min}}, \frac{\alpha^2}{2\alpha - 1} \frac{l_1}{\kappa^{\min}} \right) \left(\int_{\Omega} \nabla u \cdot \kappa \cdot \nabla u dv + \mu_o \int_{\partial\Omega} \zeta u^2 da \right). \end{aligned}$$

Notons que $\frac{\alpha^2}{2\alpha - 1}$ est minimum pour $\alpha = 1$.

$$\|u\|_{L^2(\Omega)}^2 \leq l_1 \max \left(\frac{1}{\mu_o \zeta^{\min}}, \frac{l_1}{\kappa^{\min}} \right) \|u\|_{H(\Omega)}^2.$$

Pour minimiser le max il faut choisir $\mu_o \geq \frac{\kappa^{\min}}{l_1 \zeta^{\min}}$. Bien entendu on va choisir la direction où l_i est minimum, soit $l^{\min} = \min_{1 \leq i \leq d} (l_i)$. Nous trouvons donc

$$\|u\|_{L^2(\Omega)}^2 \leq \frac{l^{\min 2}}{\kappa^{\min}} \|u\|_{H(\Omega)}^2.$$

Autrement dit la constante optimale existe et vérifie :

$$C_* \leq C_o = \frac{l^{\min}}{\sqrt{\kappa^{\min}}}.$$

Mais la norme $\|u\|_{H(\Omega)}^2$ sera minimum à son tour si nous choisissons $\mu_o = \frac{\kappa^{\min}}{l^{\min} \zeta^{\min}}$. Nous retrouvons la minimisation (3.D.1)

■

D.2 Encadrement de la constante $C_*(\Omega)$.

On peut chercher un majorant de la constante $C_*(\Omega)$ dans des cas particuliers. Pour cela énonçons un résultat préliminaire.

Proposition D.2.1 *Soit $l > 0$ une longueur arbitraire (sans dimension ici). Si dans la norme $\|\cdot\|_{H(\Omega)}$ nous choisissons un facteur de pondération tel que $\mu_o = \frac{\kappa^{\min}}{l \zeta^{\min}}$ alors :*

$$\kappa^{\min} \left(\int_{\Omega} |\nabla u|^2 dv + \frac{1}{l} \int_{\partial\Omega} u^2 da \right) \leq \|u\|_{H(\Omega)}^2 \leq \kappa^{\min} \alpha_o^2 \left(\int_{\Omega} |\nabla u|^2 dv + \frac{1}{l} \int_{\partial\Omega} u^2 da \right). \quad (3.D.3)$$

Expression où nous avons posé : $\alpha_o = \sqrt{\max\left(\frac{\kappa^{\max}}{\kappa^{\min}}, \frac{\zeta^{\max}}{\zeta^{\min}}\right)} \geq 1$.

▲

Preuve D.2.1 *La démonstration est immédiate. Le tenseur de conductivité est défini positif. Il suffit de choisir respectivement pour κ^{\min} et κ^{\max} la plus petite et la plus grande valeur propre du tenseur de conductivité dans tout le domaine. Les principes de la thermodynamique assurent que $\kappa^{\min} > 0$ et $\kappa^{\max} < \infty$.*

D'autre part le facteur de Steklov vérifie les conditions $0 < \zeta^{\min} \leq \zeta \leq \zeta^{\max} < \infty$. D'où le résultat.

■

D.2.1 Cas d'une boule de \mathbb{R}^d .

Si $\Omega = \{x \in \mathbb{R}^d \text{ tel que } |x|^2 = \sum_{i=1}^d x_i^2 \leq r^2\}$ alors $l^{\min} = 2r$ d'où $C_*(\Omega) \leq \frac{2r}{\sqrt{\kappa^{\min}}}$. Nous allons tester des fonctions particulières (et simples à calculer) pour tenter d'obtenir un encadrement pertinent. Donnons trois exemples :

- Considérons la fonction particulière $u_o(x) = x_1$. Notons s_d l'aire de la boule unité de \mathbb{R}^{d+2} . Rappelons que le volume de la boule unité se déduit de son aire et est donné par $v_d = \frac{s_d}{d}$. Un calcul explicite donne :

$$\|u_o\|_{L^2(\Omega)}^2 = \int_{\Omega} u^2(x) dv = \frac{s_d}{d(d+2)} r^{d+2}.$$

2. D'après [106] on a le résultat classique $s_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}$.

En particulier $s_1 = 2$, $s_2 = 2\pi$, $s_3 = 4\pi$, $s_4 = 2\pi^2$, $s_5 = \frac{8}{3}\pi^2$, $s_6 = \pi^3$

En effet si $\rho = \sqrt{\sum_{i=1}^d x_i^2}$ alors

$$\forall i \quad I_i = \int_{\Omega} x_i^2 dv \quad \text{donc} \quad I = \sum_{i=1}^d I_i = \int_{\Omega} \rho^2 dv = \frac{s_d}{d+2} r^{d+2} \quad \text{d'où} \quad I_1 = \frac{I}{d}.$$

Avec le même type de démarche nous obtenons :

$$\begin{aligned} \int_{\partial\Omega} u_o^2(\sigma) da &= \frac{s_d}{d} r^{d+1}, \\ \int_{\Omega} |\nabla u_o(x)|^2 dv &= v_d = \frac{s_d}{d} r^d \quad \text{car} \quad |\nabla u_o(x)| = 1, \end{aligned}$$

soit :

$$\begin{aligned} \int_{\Omega} |\nabla u_o|^2 dv + \frac{1}{2r} \int_{\partial\Omega} u_o^2 da &= \frac{s_d}{d} r^d + \frac{1}{2r} \frac{s_d}{d} r^{d+1} = \frac{3}{2} \frac{s_d}{d} r^d \\ &= \frac{3(d+2)}{2r^2} \left(\frac{s_d r^{d+2}}{d(d+2)} \right) = \frac{3(d+2)}{2r^2} \|u_o\|_{L^2(\Omega)}^2. \end{aligned}$$

Donc en revenant à la norme $\|u\|_{H(\Omega)}$ à l'aide de (3.D.3) (en prenant $l = 2r$) nous écrirons :

$$\begin{aligned} \|u\|_{H(\Omega)}^2 &\leq \kappa^{\min} \alpha_o^2 \left(\int_{\Omega} |\nabla u|^2 dv + \frac{1}{2r} \int_{\partial\Omega} u^2 da \right) = \kappa^{\min} \alpha_o^2 \frac{3(d+2)}{2r^2} \|u_o\|_{L^2(\Omega)}^2 \\ &\leq \kappa^{\min} \alpha_o^2 \frac{3(d+2)}{2r^2} C_*^2 \|u_o\|_{H(\Omega)}^2. \end{aligned}$$

En utilisant la proposition D.1.1 nous obtenons l'encadrement :

$$\sqrt{\frac{1}{6(d+2)}} \frac{1}{\alpha_o} \frac{2r}{\sqrt{\kappa^{\min}}} \leq C_*(\Omega) \leq \frac{2r}{\sqrt{\kappa^{\min}}}. \quad (3.D.4)$$

— Choisissons la fonction radiale $u_o(x) = \rho$. Nous trouvons puisque $|\nabla \rho|^2 = 1$:

$$\int_{\Omega} |\nabla u_o|^2 dv + \frac{1}{2r} \int_{\partial\Omega} u_o^2 da = \frac{s_d}{d} r^d + \frac{1}{2r} s_d r^{d+1} = \frac{d+2}{2d} s_d r^d = \frac{(d+2)^2}{2dr^2} \|u_o\|_{L^2(\Omega)}^2.$$

Donc en revenant à la norme $\|u\|_{H(\Omega)}$ nous écrirons :

$$\|u\|_{H(\Omega)}^2 \leq \kappa^{\min} \alpha_o^2 \frac{(d+2)^2}{2dr^2} \|u_o\|_{L^2(\Omega)}^2 \leq \kappa^{\min} \alpha_o^2 \frac{(d+2)^2}{2dr^2} C_*^2 \|u_o\|_{H(\Omega)}^2.$$

Nous obtenons le nouvel encadrement :

$$\frac{\sqrt{d}}{\sqrt{2}(d+2)} \frac{1}{\alpha_o} \frac{2r}{\sqrt{\kappa^{\min}}} \leq C_*(\Omega) \leq \frac{2r}{\sqrt{\kappa^{\min}}}. \quad (3.D.5)$$

- Avec une fonction constante $u_o(x) = 1$ le terme de diffusion s'élimine puisque $|\nabla u_o|^2 = 0$, soit :

$$\int_{\Omega} |\nabla u_o|^2 dv + \frac{1}{2r} \int_{\partial\Omega} u_o^2 da = \frac{1}{2r} s_d r^{d-1} = \frac{d}{2r^2} \frac{s_d}{d} r^d = \frac{d}{2r^2} v_d = \frac{d}{2r^2} \|u_o\|_{L^2(\Omega)}^2.$$

En revenant à la norme $\|u\|_{H(\Omega)}$ nous écrivons :

$$\|u\|_{H(\Omega)}^2 \leq \kappa^{\min} \alpha_o^2 \frac{d}{2r^2} \|u_o\|_{L^2(\Omega)}^2 \leq \kappa^{\min} \alpha_o^2 \frac{d}{2r^2} C_*^2 \|u_o\|_{H(\Omega)}^2.$$

Cette expression aboutit à l'encadrement :

$$\sqrt{\frac{1}{2d}} \frac{1}{\alpha_o} \frac{2r}{\sqrt{\kappa^{\min}}} \leq C_*(\Omega) \leq \frac{2r}{\sqrt{\kappa^{\min}}}. \quad (3.D.6)$$

Numériquement nous résumons ces résultats dans le tableau 3.D.1. Ils montrent que parmi les trois fonctions que nous avons testées, la meilleure estimation est obtenue avec la fonction constante. En ce sens nous pouvons dire que (3.D.6) donne une "bonne" approximation de $C_*(\Omega)$ pour la sphère pleine.

D.2.2 Cas d'un pavé droit de \mathbb{R}^d .

Soit le pavé droit $\Omega = \prod_{i=1}^d]-l_i/2, l_i/2[$. Supposons pour fixer les idées que $l^{\min} = l_1$ d'où avec le choix $\mu_o = \frac{\kappa^{\min}}{l^{\min} \zeta^{\min}}$ la borne supérieure $C_*(\Omega) \leq \frac{l_1}{\sqrt{\kappa^{\min}}}$.

Pour trouver une borne inférieure considérons à nouveau des fonctions particulières :

- Examinons le cas où $u_o(x) = x_1$. Un calcul explicite donne :

$$\begin{aligned} \|u_o\|_{L^2(\Omega)}^2 &= \int_{\Omega} u^2(x) dv = \int_{-\frac{l_1}{2}}^{\frac{l_1}{2}} x_1^2 dx_1 \prod_{i=2}^d \int_{-\frac{l_i}{2}}^{\frac{l_i}{2}} dx_i \\ &= \frac{2}{3} \left(\frac{l_1}{2}\right)^3 \prod_{i=2}^d l_i = \frac{1}{3} \left(\frac{l_1}{2}\right)^2 \prod_{i=1}^d l_i = \frac{l_1^2}{12} \prod_{i=1}^d l_i. \end{aligned}$$

Remarquons qu'avec nos hypothèses $l_1/l_i \leq 1$. Donc

$$\begin{aligned} \int_{\partial\Omega} u_o^2(\sigma) da &= 2 \left(\frac{l_1}{2}\right)^2 \prod_{i=2}^d \int_{-\frac{l_1}{2}}^{\frac{l_1}{2}} dx_i + 2 \sum_{j=2}^d \int_{-\frac{l_1}{2}}^{\frac{l_1}{2}} x_1^2 dx_1 \prod_{i=2, i \neq j}^d \int_{-\frac{l_i}{2}}^{\frac{l_i}{2}} dx_i \\ &= 2 \left(\frac{l_1}{2}\right)^2 \prod_{i=2}^d l_i + \frac{4}{3} \sum_{j=2}^d \left(\frac{l_1}{2}\right)^3 \prod_{i=2, i \neq j}^d l_i = \left(1 + \frac{1}{3} \sum_{j=2}^d \frac{l_1}{l_j}\right) \frac{l_1}{2} \prod_{i=1}^d l_i \\ &\leq \frac{d+2}{3} \left(\frac{l_1}{2}\right) \prod_{i=1}^d l_i. \end{aligned}$$

Enfin puisque $|\nabla u_o(x)| = 1$

$$\int_{\Omega} |\nabla u_o(x)|^2 dv = \prod_{i=1}^d l_i.$$

En combinant ces résultats :

$$\int_{\Omega} |\nabla u_o|^2 dv + \frac{1}{l_1} \int_{\partial\Omega} u_o^2 da \leq \prod_{i=1}^d l_i + \frac{d+2}{6} \prod_{i=1}^d l_i = \frac{2(d+8)}{l_1^2} \left(\frac{l_1^2}{12} \prod_{i=1}^d l_i\right) = \frac{2(d+8)}{l_1^2} \|u_o\|_{L^2(\Omega)}^2.$$

En revenant à la norme $\|u_o\|_{H(\Omega)}$ à l'aide de (3.D.3) nous obtenons :

$$\begin{aligned} \|u_o\|_{H(\Omega)}^2 &\leq \kappa^{\min} \max\left(\frac{\kappa^{\max}}{\kappa^{\min}}, \frac{\zeta^{\max}}{\zeta^{\min}}\right) \left(\int_{\Omega} |\nabla u_o|^2 dv + \frac{1}{l_1} \int_{\partial\Omega} u_o^2 da\right) \\ &\leq \kappa^{\min} \alpha_o^2 \frac{2(d+8)}{l_1^2} \|u_o\|_{L^2(\Omega)}^2 \leq \kappa^{\min} \alpha_o^2 \frac{2(d+8)}{l_1^2} C_*^2 \|u_o\|_{H(\Omega)}^2. \end{aligned}$$

Donc pour ce choix particulier de fonction ($u(x) = x_1$) et en utilisant la proposition D.1.1 nous trouvons l'encadrement :

$$\sqrt{\frac{1}{2(d+8)}} \frac{1}{\alpha_o} \frac{l^{\min}}{\sqrt{\kappa^{\min}}} \leq C_*(\Omega) \leq \frac{l^{\min}}{\sqrt{\kappa^{\min}}}. \quad (3.D.7)$$

— Considérons une fonction constante. Pour fixer les idées choisissons $u_o(x) = 1$.

Un calcul direct donne $\|u_o\|_{L^2(\Omega)}^2 = \prod_{i=1}^d l_i$. Avec nos hypothèses $l_1/l_i \leq 1$. Donc

$$\begin{aligned} \int_{\partial\Omega} u_o^2(\sigma) da &= 2 \prod_{i=2}^d \int_{-\frac{l_1}{2}}^{\frac{l_1}{2}} dx_i + 2 \sum_{j=2}^d \int_{-\frac{l_1}{2}}^{\frac{l_1}{2}} dx_1 \prod_{i=2, i \neq j}^d \int_{-\frac{l_i}{2}}^{\frac{l_i}{2}} dx_i = 2 \prod_{i=2}^d l_i + 2 \sum_{j=2}^d l_1 \prod_{i=2, i \neq j}^d l_i \\ &= \frac{2}{l_1} \prod_{i=1}^d l_i + \sum_{j=2}^d \frac{2}{l_j} \prod_{i=1}^d l_i = \left(\frac{2}{l_1} \prod_{i=1}^d l_i \right) \sum_{j=1}^d \frac{l_1}{l_j} \leq \frac{2d}{l_1} \prod_{i=1}^d l_i. \end{aligned}$$

En combinant ces résultats, et puisqu'ici $|\nabla u_o(x)| = 0$:

$$\int_{\Omega} |\nabla u_o|^2 dv + \frac{1}{l_1} \int_{\partial\Omega} u_o^2 da \leq \frac{2d}{l_1^2} \prod_{i=1}^d l_i = \frac{2d}{l_1^2} \|u_o\|_{L^2(\Omega)}^2.$$

En revenant à la norme $\|u_o\|_{H(\Omega)}$ à l'aide de (3.D.3) nous obtenons :

$$\|u_o\|_{H(\Omega)}^2 \leq \kappa^{\min} \alpha_o^2 \frac{2d}{l_1^2} \|u_o\|_{L^2(\Omega)}^2 \leq \kappa^{\min} \alpha_o^2 \frac{2d}{l_1^2} C_*^2 \|u_o\|_{H(\Omega)}^2.$$

Donc pour le choix particulier d'une fonction constante nous trouvons l'encadrement :

$$\sqrt{\frac{1}{2d}} \frac{1}{\alpha_o} \frac{l^{\min}}{\sqrt{\kappa^{\min}}} \leq C_*(\Omega) \leq \frac{l^{\min}}{\sqrt{\kappa^{\min}}} \quad (3.D.8)$$

Notons que nous trouvons la même borne inférieure que dans le cas de la sphère (avec la même fonction test constante). Cette borne est meilleur que la précédente.

Numériquement nous résumons ces résultats dans le tableau 3.D.1. Ils montrent que parmi les deux fonctions que nous avons testées, la meilleur estimation est obtenue avec la fonction constante. En ce sens nous pouvons dire que (3.D.8) constitue une "bonne" approximation de $C_*(\Omega)$ pour le pavé droit.

D.2.3 Cas d'un sphère creuse de faible épaisseur de \mathbb{R}^d .

Considérons le cas d'une sphère creuse $\Omega = \{x \in \mathbb{R}^d \text{ tel que } r_o^2 \leq \sum_{i=1}^d x_i^2 \leq r^2\}$. On notera $e = r - r_o$ l'épaisseur de la sphère. Nos définitions nous conduisent, comme dans le cas de la sphère pleines, à l'estimation $C_*(\Omega) \leq \frac{2r}{\sqrt{\kappa^{\min}}}$. Mais cette estimation n'est pas du tout optimale comme nous allons le voir. Pour cela il faut reprendre la démarche de la preuve D.1.1.

Remarquons que si $\rho = \sqrt{\sum_{i=1}^d x_i^2}$, on a $\sum_{i=1}^d \frac{x_i}{\rho} \frac{\partial(\rho-r_o)}{\partial x_i} = 1$

$$\begin{aligned} \|u\|_{L^2(\Omega)}^2 &= \int_{\Omega} u^2(x) dv = \int_{\Omega} u^2(x) \sum_{i=1}^d \frac{x_i}{\rho} \frac{\partial(\rho-r_o)}{\partial x_i} \\ &= \sum_{i=1}^d \left(\int_{\partial\Omega} (\rho-r_o) \frac{\sigma_i}{\rho} n_i u^2(\sigma) da - \int_{\Omega} (\rho-r_o) \frac{\partial}{\partial x_i} \left(u^2(x) \frac{x_i}{\rho} \right) dv \right). \end{aligned}$$

Mais

$$\frac{\partial}{\partial x_i} \left(u^2(x) \frac{x_i}{\rho} \right) = \frac{x_i}{\rho} \frac{\partial u^2(x)}{\partial x_i} + u^2(x) \frac{\partial}{\partial x_i} \left(\frac{x_i}{\rho} \right) = 2u(x) \frac{x_i}{\rho} \frac{\partial u(x)}{\partial x_i} + u^2(x) \left(\frac{1}{\rho} - \frac{x_i^2}{\rho^3} \right),$$

et

$$\sum_{i=1}^d \left(\frac{1}{\rho} - \frac{x_i^2}{\rho^3} \right) = \frac{d-1}{\rho},$$

donc

$$\begin{aligned} \|u\|_{L^2(\Omega)}^2 &= \sum_{i=1}^d \left(\int_{\partial\Omega} (\rho-r_o) \frac{\sigma_i}{\rho} n_i u^2(\sigma) da - 2 \int_{\Omega} (\rho-r_o) \frac{x_i}{\rho} u(x) \frac{\partial u(x)}{\partial x_i} dv \right) \\ &\quad - \int_{\Omega} (\rho-r_o) \frac{d-1}{\rho} u^2(x) dv. \end{aligned}$$

Nous remarquons :

- Sur $\partial\Omega$ on a $\rho = r_o$ ou r . Donc $(\rho - r_o) \leq e$.
- Puisque $n_i \leq 1$ on a aussi $\sum_{i=1}^d n_i \frac{\sigma_i}{\rho} \leq \frac{1}{\rho} \sum_{i=1}^d \sigma_i \leq \frac{1}{\rho} (\sum_{i=1}^d \sigma_i^2)^{1/2} = 1$.
- Dans Ω on a aussi $|\sum_{i=1}^d (\rho - r_o) \frac{x_i}{\rho} \frac{\partial u(x)}{\partial x_i}| \leq e |\nabla u(x)|$.

Donc :

$$\|u\|_{L^2(\Omega)}^2 \leq e \int_{\partial\Omega} u^2(\sigma) da + 2e \int_{\Omega} |u(x) \nabla u(x)| dv.$$

Examinons successivement les deux termes du membre de droite de cette inégalité :

$$e \int_{\partial\Omega} u^2(\sigma) da \leq \frac{e}{\zeta_{\min}} \int_{\partial\Omega} \zeta u^2 da.$$

3.

$$\frac{\partial}{\partial x_k} \left(\left(\sum_{i=1}^d x_i^2 \right)^{1/2} - r_o \right) = \frac{\partial}{\partial x_k} \left(\sum_{i=1}^d x_i^2 \right)^{1/2} = \frac{1}{2} \frac{\frac{\partial}{\partial x_k} \sum_{i=1}^d x_i^2}{\sqrt{\sum_{i=1}^d x_i^2}} = \frac{x_k}{\rho} \quad \text{soit} \quad \sum_{i=1}^d \frac{x_i}{\rho} \frac{\partial(\rho-r_o)}{\partial x_i} = \sum_{i=1}^d \frac{x_i^2}{\rho^2} = 1.$$

Pour le deuxième terme en remarquant l'inégalité $|ab| \leq \frac{1}{2}(\epsilon a^2 + \frac{b^2}{\epsilon})$ qui est vraie $\forall \epsilon > 0$ nous obtenons :

$$2e \int_{\Omega} |u(x) \nabla u(x)| dv \leq \frac{e}{\epsilon} \int_{\Omega} |u(x)|^2 dv + \epsilon e \int_{\Omega} |\nabla u(x)|^2 dv.$$

Donc en combinant ces différentes inégalités et en posant $\epsilon = 2\alpha e$ avec $\alpha > 1/2$:

$$\frac{2\alpha - 1}{2\alpha} \|u\|_{L^2(\Omega)}^2 \leq \frac{e}{\zeta^{\min}} \int_{\partial\Omega} \zeta u^2 da + 2\alpha e^2 \int_{\Omega} \left| \frac{\partial u(x)}{\partial x_1} \right|^2 dv.$$

Enfin le tenseur de conductivité étant défini positif $\int_{\Omega} |\nabla u|^2 dv \leq \frac{1}{\kappa^{\min}} \int_{\Omega} \nabla u \cdot \kappa \cdot \nabla u dv$. Donc :

$$\begin{aligned} \|u\|_{L^2(\Omega)}^2 &\leq 2e \frac{\alpha}{2\alpha - 1} \left(\frac{1}{\zeta^{\min}} \int_{\partial\Omega} \zeta u^2 da + \frac{2\alpha e}{\kappa^{\min}} \int_{\Omega} \nabla u \cdot \kappa \cdot \nabla u dv \right) \\ &\leq 2e \max \left(\frac{\alpha}{2\alpha - 1} \frac{1}{\mu_o \zeta^{\min}}, \frac{\alpha^2}{2\alpha - 1} \frac{2e}{\kappa^{\min}} \right) \left(\int_{\Omega} \nabla u \cdot \kappa \cdot \nabla u dv + \mu_o \int_{\partial\Omega} \zeta u^2 da \right). \end{aligned}$$

Notons que $\frac{\alpha^2}{2\alpha - 1}$ est minimum pour $\alpha = 1$.

$$\|u\|_{L^2(\Omega)}^2 \leq 2e \max \left(\frac{1}{\mu_o \zeta^{\min}}, \frac{2e}{\kappa^{\min}} \right) \|u\|_{H(\Omega)}^2.$$

Pour minimiser le max il faut choisir $\mu_o \geq \frac{\kappa^{\min}}{2e\zeta^{\min}}$. Mais la norme $\|u\|_{H(\Omega)}^2$ sera minimum à son tour si nous choisissons $\mu_o = \frac{\kappa^{\min}}{2e\zeta^{\min}}$. Nous trouvons donc

$$\|u\|_{L^2(\Omega)}^2 \leq \frac{4e^2}{\kappa^{\min}} \|u\|_{H(\Omega)}^2.$$

Autrement dit la constante optimale vérifie la majoration :

$$C_* \leq \frac{2e}{\sqrt{\kappa^{\min}}}. \quad (3.D.9)$$

Notons que pour établir ce résultat nous n'avons pas fait d'hypothèse sur l'épaisseur. Il s'applique donc aussi bien aux coques sphériques épaisses ou minces.

Nous pouvons chercher une borne inférieure pour $C_*(\Omega)$. Nous nous placerons cette fois dans le cas où $e \ll r$, autrement dit $u_o(x) \approx 2r(\rho - r_o)$.

— Considérons en premier lieu la fonction particulière $u_o(x) = \rho^2 - r_o^2$.

$$\|u_o\|_{L^2(\Omega)}^2 = \int_{\Omega} u_o^2(x) dv \approx s_d r^{d-1} 4 r^2 \int_{r_o}^r (\rho - r_o)^2 d\rho = \frac{4}{3} s_d r^{d+1} e^3$$

et

$$\int_{\partial\Omega} u^2(\sigma) da = s_d r^{d-1} (r^2 - r_o^2)^2 \approx 4 s_d r^{d+1} e^2.$$

Enfin puisque $|\nabla u_o(x)|^2 = 4\rho^2$ nous en déduisons :

$$\int_{\Omega} |\nabla u_o(x)|^2 dv = \int_{r_o}^r 4\rho^2 s_d \rho^{d-1} d\rho = \frac{4 s_d}{d+2} (r^{d+2} - r_o^{d+2}) = \frac{4 s_d}{d+2} r^{d+2} \left(1 - \left(1 - \frac{e}{r}\right)^{d+2}\right)$$

$$\approx 4 s_d r^{d+1} e.$$

Donc en utilisant les résultats du calcul explicite nous obtenons la relation :

$$\begin{aligned} \int_{\Omega} |\nabla u_o|^2 dv + \frac{1}{2e} \int_{\partial\Omega} u_o^2 da &= 4 s_d r^{d+1} e + \frac{1}{2e} 4 s_d r^{d+1} e^2 \\ &\approx 6 s_d r^{d+1} e \\ &\approx \frac{9}{2e^2} \frac{4}{3} s_d r^{d+1} e^3 \\ &= \frac{18}{4e^2} \|u_o\|_{L^2(\Omega)}^2. \end{aligned}$$

En revenant à la norme $\|u\|_{H(\Omega)}$ à l'aide de (3.D.3) (en prenant $l = 2e$) nous obtenons :

$$\begin{aligned} \|u_o\|_{H(\Omega)}^2 &\leq \kappa^{\min} \max\left(\frac{\kappa^{\max}}{\kappa^{\min}}, \frac{\zeta^{\max}}{\zeta^{\min}}\right) \left(\int_{\Omega} |\nabla u_o|^2 dv + \frac{1}{2e} \int_{\partial\Omega} u_o^2 da \right) \\ &\lesssim \kappa^{\min} \alpha_o^2 \frac{18}{4e^2} \|u_o\|_{L^2(\Omega)}^2 \leq \kappa^{\min} \alpha_o^2 \frac{18}{4e^2} C_*^2 \|u_o\|_{H(\Omega)}^2. \end{aligned}$$

Dans le cas d'une coque sphérique mince la longueur l^{\min} qui doit être retenue est deux fois l'épaisseur e . Nous obtenons finalement l'encadrement :

$$\frac{1}{3\sqrt{2}} \frac{1}{\alpha_o} \frac{2e}{\sqrt{\kappa^{\min}}} \lesssim C_*(\Omega) \leq \frac{2e}{\sqrt{\kappa^{\min}}}. \quad (3.D.10)$$

4. Par exemple si $d = 3$ le gradient en coordonnées sphériques s'écrit $\nabla \cdot = \frac{\partial}{\partial \rho} \underline{e}_r + \frac{1}{\rho} \frac{\partial}{\partial \theta} \underline{e}_\theta + \frac{1}{\rho \sin \theta} \frac{\partial}{\partial \phi} \underline{e}_\phi$.

On a donc $\nabla u_o = 2\rho \underline{e}_r$ et $|\nabla u_o|^2 = 4\rho^2$.

— Dans un deuxième temps nous utilisons la fonction constante $u_o(x) = 1$.

$$\|u_o\|_{L^2(\Omega)}^2 = \int_{\Omega} dv \approx s_d r^{d-1} e$$

et
$$\int_{\partial\Omega} da \approx 2 s_d r^{d-1} \quad \text{il y a deux faces!}$$

Enfin puisque $|\nabla u_o(x)| = 0$ nous en déduisons la relation :

$$\int_{\Omega} |\nabla u_o|^2 dv + \frac{1}{2e} \int_{\partial\Omega} u_o^2 da \approx \frac{1}{e} s_d r^{d-1} e = \frac{1}{e^2} s_d r^{d-1} e \approx \frac{1}{e^2} \|u_o\|_{L^2(\Omega)}^2.$$

En revenant à la norme $\|u\|_{H(\Omega)}$ à l'aide de (3.D.3) (en prenant $l = 2e$) nous obtenons :

$$\|u_o\|_{H(\Omega)}^2 \lesssim \kappa^{\min} \alpha_o^2 \frac{1}{e^2} \|u_o\|_{L^2(\Omega)}^2 \leq \kappa^{\min} \alpha_o^2 \frac{4}{4e^2} C_*^2 \|u_o\|_{H(\Omega)}^2.$$

Nous obtenons un autre encadrement :

$$\frac{1}{2} \frac{1}{\alpha_o} \frac{2e}{\sqrt{\kappa^{\min}}} \lesssim C_*(\Omega) \leq \frac{2e}{\sqrt{\kappa^{\min}}}. \quad (3.D.11)$$

Numériquement les deux encadrements obtenus sont, dans le cadre de l'approximation des coques minces, indépendants de la dimension de l'espace. Celui obtenu avec la fonction test constante est légèrement meilleur que celui construit avec une fonction radiale ($\frac{1}{3\sqrt{2}} \approx 0,2357 < \frac{1}{2}$). Nous retrouvons le même ordre de grandeur que pour la sphère pleine. Cependant la sphère pleine et la sphère creuse ne partagent pas la même dimension minimale l^{\min} qui est respectivement le rayon extérieur r et deux fois l'épaisseur $2e$. Pour la coque mince on a $e \ll r$. En ce sens nous pouvons dire que (3.D.9) donne une majoration efficace pour la sphère creuse (pas nécessairement mince) et (3.D.11) fournit un encadrement utile dans le cas d'une coque sphérique mince. A contrario (3.D.2) n'est pas une "bonne" approximation de $C_*(\Omega)$ pour la coque sphérique mince.

D.2.4 Cas d'un cylindre droit de \mathbb{R}^d .

Soit le cylindre droit $\Omega = \{x \in \mathbb{R}^d, d \geq 3 \text{ tel que } \sqrt{x_1^2 + x_2^2} \leq r, \forall i > 2 \quad |x_i| \leq li/2\}$.

Pour trouver une borne inférieure considérons à nouveau des fonctions particulières. En nous basant sur les résultats des géométries déjà traitées nous considérons directement une fonction constante. Pour fixer les idées choisissons $u_o(x) = 1$.

Un calcul direct donne $\|u_o\|_{L^2(\Omega)}^2 = \pi r^2 \prod_{i=3}^d l_i$. Supposons de plus que ce cylindre est long, autrement dit $\forall i > 2 \quad 2r \leq l_i$ et $l^{\min} = 2r$.

$$\begin{aligned} \int_{\partial\Omega} u_o^2(\sigma) da &= 2\pi r \prod_{i=3}^d l_i + 2\pi r^2 \sum_{j=3}^d \prod_{i=3, i \neq j}^d l_i \\ &= \frac{2}{r} \left(1 + \frac{1}{2} \sum_{j=3}^d \frac{2r}{l_j} \right) \pi r^2 \prod_{i=3}^d l_i \leq \frac{d}{r} \pi r^2 \prod_{i=3}^d l_i. \end{aligned}$$

En combinant ces résultats, et puisqu'ici $|\nabla u_o(x)| = 0$:

$$\int_{\Omega} |\nabla u_o|^2 dv + \frac{1}{2r} \int_{\partial\Omega} u_o^2 da \leq \frac{d}{2r^2} \|u_o\|_{L^2(\Omega)}^2.$$

En revenant à la norme $\|u_o\|_{H(\Omega)}$ à l'aide de (3.D.3) nous obtenons :

$$\|u_o\|_{H(\Omega)}^2 \leq \kappa^{\min} \alpha_o^2 \frac{d}{2r^2} \|u_o\|_{L^2(\Omega)}^2 \leq \kappa^{\min} \alpha_o^2 \frac{2d}{4r^2} C_*^2 \|u_o\|_{H(\Omega)}^2.$$

Donc pour la fonction particulière constante nous trouvons l'encadrement de la constante optimale d'un long cylindre droit de \mathbb{R}^d avec le choix $\mu_o = \frac{\kappa^{\min}}{2r \zeta^{\min}}$:

$$\frac{1}{\sqrt{2d}} \frac{1}{\alpha_o} \frac{2r}{\sqrt{\kappa^{\min}}} \leq C_*(\Omega) \leq \frac{2r}{\sqrt{\kappa^{\min}}}. \quad (3.D.12)$$

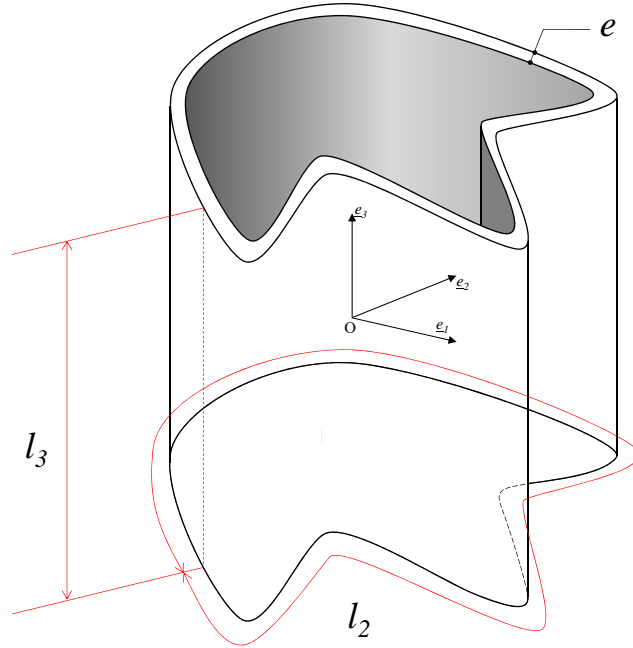
Examinons cette fois le cas du cylindre court. Pour fixer les idées nous supposons $l^{\min} = l_3$, autrement dit $\forall i > 3 \quad l_3 \leq l_i$ et $l_3 \leq 2r$. Cette fois il faut choisir $\mu_o = \frac{\kappa^{\min}}{l_3 \zeta^{\min}}$ ce qui abouti à la borne supérieure $C_*(\Omega) \leq \frac{l_3}{\sqrt{\kappa^{\min}}}$. En suivant la même démarche que pour le cylindre long nous obtenons :

$$\begin{aligned} \int_{\partial\Omega} u_o^2(\sigma) da &= 2\pi r \prod_{i=3}^d l_i + 2\pi r^2 \prod_{i=4}^d l_i + 2\pi r^2 \sum_{j=4}^d l_3 \prod_{i=4, i \neq j}^d l_i \\ &= \frac{2}{l_3} \left(2 \frac{l_3}{2r} + 1 + \sum_{j=4}^d \frac{l_3}{l_j} \right) \pi r^2 \prod_{i=3}^d l_i \leq \frac{2d}{l_3} \pi r^2 \prod_{i=3}^d l_i. \end{aligned}$$

et finalement l'encadrement :

$$\frac{1}{\sqrt{2d}} \frac{1}{\alpha_o} \frac{l_3}{\sqrt{\kappa^{\min}}} \leq C_*(\Omega) \leq \frac{l_3}{\sqrt{\kappa^{\min}}}. \quad (3.D.13)$$

Voir le tableau 3.D.1 pour les approximations numériques.

FIGURE 3.D.1 – Tube mince d'épaisseur e

D.2.5 Cas d'un tube cylindrique droit de \mathbb{R}^d .

Soit le tube cylindrique droit à base circulaire défini par :

$$\Omega = \{x \in \mathbb{R}^d, d \geq 3, \text{ tel que } r_o \leq \sqrt{x_1^2 + x_2^2} \leq r, \forall i > 2 \quad |x_i| \leq l_i/2\}.$$

Pour trouver une borne inférieure nous considérons une fonction constante. Nous nous placerons dans le cas d'un tube mince d'épaisseur e telle que $e/l_i \ll 1$ (figure 3.D.1).

On a les approximations :

$$\begin{aligned} \|u_o\|_{L^2(\Omega)}^2 &= \int_{\Omega} u_o^2(x) dv = \int_{\Omega} dv \approx e \prod_{i=2}^d l_i. \\ \int_{\partial\Omega} u_o^2(\sigma) da &= \int_{\partial\Omega} da \approx l_2 \prod_{i=3}^d l_i + 2e l_2 \sum_{j=3}^d \prod_{i=3, i \neq j}^d l_i \\ &= \left(1 + 2 \sum_{j=3}^d \frac{e}{l_j}\right) \prod_{i=2}^d l_i \lesssim \frac{2d-3}{e} e \prod_{i=2}^d l_i. \end{aligned}$$

En combinant ces résultats et en choisissant $l^{\min} = 2e$:

$$\int_{\Omega} |\nabla u_o|^2 dv + \frac{1}{2e} \int_{\partial\Omega} u_o^2 da \lesssim \frac{2d-3}{2e^2} \|u_o\|_{L^2(\Omega)}^2.$$

En revenant à la norme $\|u_o\|_{H(\Omega)}$ à l'aide de (3.D.3) nous obtenons :

$$\|u_o\|_{H(\Omega)}^2 \leq \kappa^{\min} \alpha_o^2 \frac{2d-3}{2e^2} \|u_o\|_{L^2(\Omega)}^2 \leq \kappa^{\min} \alpha_o^2 \frac{2(2d-3)}{4e^2} C_*^2 \|u_o\|_{H(\Omega)}^2.$$

Donc pour la fonction particulière constante nous trouvons une borne inférieure⁵ de la constante optimale d'un tube droit mince de \mathbb{R}^d (avec le choix $\mu_o = \frac{\kappa^{\min}}{2e\zeta^{\min}}$) :

$$\frac{1}{\sqrt{2(2d-3)}} \frac{1}{\alpha_o} \frac{2e}{\sqrt{\kappa^{\min}}} \leq C_*(\Omega) \leq \frac{\min_{1 \leq i \leq d} l_i}{\sqrt{\kappa^{\min}}}. \quad (3.D.14)$$

Les approximations numériques sont reportées dans le tableau 3.D.1

D.3 Estimation d'erreur dans l'espace d'énergie.

Définition D.3.1 Si T est la solution exacte du problème (3.A.4) et \tilde{T} la solution approchée par un modèle réduit d'ordre \tilde{N} nous pouvons former en chaque point du domaine Ω et à chaque instant de l'intervalle de simulation $]0, \Delta t[$ la différence :

$$\Delta T = |T - \tilde{T}| \in H^1(\Omega).$$

Les erreurs quadratiques de réduction dans l'espace des températures $L^2(]0, \Delta t[\times \Omega)$ et d'énergie $L^2(]0, \Delta t[; H^1(\Omega))$ notées respectivement $\varepsilon_{\tilde{N}}$ et $\mathcal{E}_{\tilde{N}}$ sont mesurées en Kelvin et définies par :

$$\varepsilon_{\tilde{N}} = \frac{\|\Delta T\|_{L^2(\Delta t \times \Omega)}}{\sqrt{\mathcal{V}_{ol} \Delta t}} \quad \text{avec} \quad \|\Delta T\|_{L^2(\Delta t \times \Omega)} = \left(\int_0^{\Delta t} \|\Delta T\|_{L^2(\Omega)}^2 dt \right)^{1/2}, \quad (3.D.15)$$

$$\mathcal{E}_{\tilde{N}} = \frac{\|\Delta T\|_{L^2(\Delta t; H^1(\Omega))}}{\sqrt{\mathcal{V}_{ol} \Delta t}} \quad \text{avec} \quad \|\Delta T\|_{L^2(\Delta t; H^1(\Omega))} = \left(\int_0^{\Delta t} \|\Delta T\|_{H^1(\Omega)}^2 dt \right)^{1/2}. \quad (3.D.16)$$

△

Dans cette définition nous voyons apparaître la norme usuelle dans $L^2(\Omega)$ ainsi que le volume adimensionné du domaine :

$$\|u\|_{L^2(\Omega)} = \int_{\Omega} u^2 dv, \quad \mathcal{V}_{ol} = \int_{\Omega} dv.$$

5. Pour la borne supérieure, il doit être possible, comme dans le cas de la sphère, d'établir une majoration de type : $C_*(\Omega) \leq \frac{2e}{\sqrt{\kappa^{\min}}}$.

u_o	d	1	2	3	4	5	6
Boule							
$x_1 \longleftrightarrow$	$\sqrt{\frac{1}{6(d+2)}}$	0,2357	0,2041	0,1826	0,1667	0,1543	0,1443
$\rho \longleftrightarrow$	$\frac{\sqrt{d}}{\sqrt{2}(d+2)}$	0,2357	0,2500	0,2449	0,2357	0,2259	0,2165
$C_{st} \longleftrightarrow$	$\sqrt{\frac{1}{2d}}$	0,7071	0,5000	0,4082	0,3536	0,3162	0,2887
Pavé droit							
$x_1 \longleftrightarrow$	$\sqrt{\frac{1}{2(d+8)}}$	0,2357	0,2236	0,2132	0,2041	0,1961	0,1890
$C_{st} \longleftrightarrow$	$\sqrt{\frac{1}{2d}}$	0,7071	0,5000	0,4082	0,3536	0,3162	0,2887
Cylindre droit							
$C_{st} \longleftrightarrow$	$\frac{1}{\sqrt{2d}}$			0,4082	0,3536	0,3162	0,2887
Tube mince							
$C_{st} \longleftrightarrow$	$\frac{1}{\sqrt{2(2d-3)}}$			0,4082	0,3162	0,2673	0,2357
Sphère mince							
$\rho \longleftrightarrow$	$\frac{1}{3\sqrt{2}}$	0,2357					
$C_{st} \longleftrightarrow$	$\frac{1}{2}$	0,5					

TABLE 3.D.1 – Valeur numérique pour la borne inférieure de la constante C_* en fonction de la dimension de l'espace \mathbb{R}^d . Cas d'un matériau homogène et isotrope et facteur de Steklov uniforme ($\alpha_o = 1$).

Proposition D.3.1 *Les erreurs de réduction dans l'espace des températures et dans l'espace d'énergie sont liées par l'inégalité*

$$\varepsilon_{\tilde{N}} \leq C_* \mathcal{E}_{\tilde{N}}. \quad (3.D.17)$$

▲

Preuve D.3.1 *Les champs de température T et \tilde{T} sont des fonctions de $H^1(\Omega)$. Nous pouvons*

utiliser l'inégalité (3.D.1) pour les normes spatiales, ce qui donne le résultat. Notons que les deux erreurs de réduction s'expriment ici en Kelvin puisque nous travaillons sur des opérateurs adimensionnés (mis à part la température). ■

D.4 Estimation d'énergie.

Nous considérons le problème (3.A.4). Soit T la solution de la formulation variationnelle associée :

$$\int_{\Omega} \nabla T \cdot \kappa \cdot \nabla v + \int_{\partial\Omega} \text{Bi } T |_{\partial\Omega} v |_{\partial\Omega} = - \int_{\Omega} c \frac{\partial T}{\partial t} v + \int_{\partial\Omega} \text{Bi } T_{ex} v |_{\partial\Omega} + \int_{\Omega} \varpi v. \quad (3.D.18)$$

Avec la condition initiale $T(t=0) = T_o$. En régime permanent l'équation précédente se simplifie :

$$\int_{\Omega} \nabla T \cdot \kappa \cdot \nabla v + \int_{\partial\Omega} \text{Bi } T |_{\partial\Omega} v |_{\partial\Omega} = \int_{\partial\Omega} \text{Bi } T_{ex} v |_{\partial\Omega} + \int_{\Omega} \varpi v. \quad (3.D.19)$$

Proposition D.4.1 *En régime permanent la solution de (3.D.19) vérifie l'estimation d'énergie*

$$\exists C > 0 \text{ telle que } \|T\|_{H(\Omega)} \leq C (\|\varpi\|_{L^2(\Omega)} + \|T_{ex}\|_{L^2(\partial\Omega)}) . \quad (3.D.20)$$

▲

Preuve D.4.1 *Remplaçons dans (3.D.19) la fonction v par la solution T . Pour le membre de gauche de cette nouvelle égalité nous avons, puisque le tenseur κ est défini positif :*

$$\begin{aligned} \int_{\Omega} \nabla T \cdot \kappa \cdot \nabla T + \int_{\partial\Omega} \text{Bi } T |_{\partial\Omega}^2 &\geq \kappa^{\min} \|\nabla T\|_{L^2(\Omega)}^2 + \frac{h^{\min} L_o}{K_o} \|T |_{\partial\Omega}\|_{L^2(\partial\Omega)}^2 \\ &\geq \min(\kappa, \text{Bi}/\mu_o) \left(\|\nabla T\|_{L^2(\Omega)}^2 + \mu_o \|T |_{\partial\Omega}\|_{L^2(\partial\Omega)}^2 \right) \\ &\geq \frac{\min(\kappa, \text{Bi}/\mu_o)}{2} \left(\|\nabla T\|_{L^2(\Omega)} + \sqrt{\mu_o} \|T |_{\partial\Omega}\|_{L^2(\partial\Omega)} \right)^2 . \end{aligned}$$

Pour le membre de droite en utilisant l'inégalité de Schwarz :

$$\int_{\partial\Omega} \text{Bi } T_{ex} T |_{\partial\Omega} + \int_{\Omega} \varpi T \leq \text{Bi}^{\max} \|T_{ex}\|_{L^2(\partial\Omega)} \|T |_{\partial\Omega}\|_{L^2(\partial\Omega)} + \|\varpi\|_{L^2(\Omega)} \|T\|_{L^2(\Omega)} ,$$

et l'inégalité de Poincaré

$$\begin{aligned}
&\leq \text{Bi}^{max} \|T_{ex}\|_{L^2(\partial\Omega)} \|T|_{\partial\Omega}\|_{L^2(\partial\Omega)} + l_{\square} \|\varpi\|_{L^2(\Omega)} (\|\nabla T\|_{L^2(\Omega)} + \|T|_{\partial\Omega}\|_{L^2(\partial\Omega)}) \\
&\leq (\text{Bi}^{max} \|T_{ex}\|_{L^2(\partial\Omega)} + l_{\square} \|\varpi\|_{L^2(\Omega)}) \|T|_{\partial\Omega}\|_{L^2(\partial\Omega)} + l_{\square} \|\varpi\|_{L^2(\Omega)} \|\nabla T\|_{L^2(\Omega)} \\
&\leq \max(\text{Bi}, l_{\square}) (\|T_{ex}\|_{L^2(\partial\Omega)} + \|\varpi\|_{L^2(\Omega)}) \|T|_{\partial\Omega}\|_{L^2(\partial\Omega)} + l_{\square} \|\varpi\|_{L^2(\Omega)} \|\nabla T\|_{L^2(\Omega)} \\
&\leq (\|T_{ex}\|_{L^2(\partial\Omega)} + \|\varpi\|_{L^2(\Omega)}) (\max(\text{Bi}, l_{\square}) \|T|_{\partial\Omega}\|_{L^2(\partial\Omega)} + l_{\square} \|\nabla T\|_{L^2(\Omega)}) \\
&\leq (\|T_{ex}\|_{L^2(\partial\Omega)} + \|\varpi\|_{L^2(\Omega)}) \max\left(\frac{\text{Bi}}{\mu_o}, \frac{l_{\square}}{\mu_o}, l_{\square}\right) (\sqrt{\mu_o} \|T|_{\partial\Omega}\|_{L^2(\partial\Omega)} + \|\nabla T\|_{L^2(\Omega)}).
\end{aligned}$$

Donc en simplifiant dans les deux membres par $\sqrt{\mu_o} \|T|_{\partial\Omega}\|_{L^2(\partial\Omega)} + \|\nabla T\|_{L^2(\Omega)}$

$$\sqrt{\mu_o} \|T|_{\partial\Omega}\|_{L^2(\partial\Omega)} + \|\nabla T\|_{L^2(\Omega)} \leq 2 \frac{\max\left(\frac{\text{Bi}}{\mu_o}, \frac{l_{\square}}{\mu_o}, l_{\square}\right)}{\min\left(\frac{\text{Bi}}{\mu_o}, \kappa\right)} (\|T_{ex}\|_{L^2(\partial\Omega)} + \|\varpi\|_{L^2(\Omega)}).$$

Enfin

$$\|T\|_{H(\Omega)} = \sqrt{\mu_o \|T|_{\partial\Omega}\|_{L^2(\partial\Omega)}^2 + \|\nabla T\|_{L^2(\Omega)}^2} \leq \sqrt{\mu_o} \|T|_{\partial\Omega}\|_{L^2(\partial\Omega)} + \|\nabla T\|_{L^2(\Omega)}.$$

Nous obtenons l'inégalité proposée si nous posons

$$C = 2 \frac{\max\left(\frac{\text{Bi}}{\mu_o}, \frac{l_{\square}}{\mu_o}, l_{\square}\right)}{\min\left(\frac{\text{Bi}}{\mu_o}, \kappa\right)}.$$

■

La réponse thermique du système est bornée par les sollicitations extérieures appliquées sur ce système. Autrement dit la solution dépend continument des sollicitations⁶. Notons qu'il est possible d'obtenir une estimation d'énergie pour le régime variable en temps. Elle fait apparaître la norme sur $L^2(\Omega)$ des températures initiale $T(t=0)$ et finale $T(\Delta t)$.

Proposition D.4.2 *En régime variable la solution de (3.D.18) vérifie l'estimation d'énergie*

$$\begin{aligned}
&\exists C > 0 \text{ telle que} \\
&\frac{1}{2} \|T(t)\|_{L^2(\Omega)}^2 + \|T\|_{L^2(\Delta t; H(\Omega))}^2 \leq C \left(\frac{1}{2} \|T(0)\|_{L^2(\Omega)}^2 + \|\varpi\|_{L^2(\Delta t \times \Omega)}^2 + \|T_{ex}\|_{L^2(\Delta t \times \partial\Omega)}^2 \right).
\end{aligned} \tag{3.D.21}$$

▲

6. Le problème est bien posé au sens d'Hadamard.

Preuve D.4.2 *La démonstration utilise une intégration sur le temps de l'équation (3.D.18) où nous choisissons comme fonction d'essai le champ de température lui même. Par rapport aux formules du régime permanent nous obtenons un terme supplémentaire :*

$$\int_0^{\Delta t} \int_{\Omega} c \frac{\partial T(\mathbf{x}, t)}{\partial t} T(\mathbf{x}, t) \, dx \, dt .$$

On remarque que $\frac{\partial T(\mathbf{x}, t)}{\partial t} T(\mathbf{x}, t) = \frac{1}{2} \frac{\partial T^2(\mathbf{x}, t)}{\partial t}$. On permute alors l'ordre des dérivations, soit :

$$\int_0^{\Delta t} \int_{\Omega} c \frac{1}{2} \frac{\partial T^2(\mathbf{x}, t)}{\partial t} \, dt \, dx = \frac{1}{2} \int_{\Omega} c \int_0^{\Delta t} \frac{\partial T^2(\mathbf{x}, t)}{\partial t} \, dt \, dx = \frac{1}{2} \int_{\Omega} c (T^2(\mathbf{x}, t) - T^2(\mathbf{x}, 0)) \, dx .$$

On utilise ensuite les mêmes arguments que pour le cas stationnaire.

■

Annexe E

Fonctionnement de la carte ISL EVAL8026A

E.1 Fonctionnement du convertisseur DC-DC. Estimation de la puissance thermique dissipée.

E.1.1 Principe de fonctionnement d'un convertisseur DC-DC.

E.1.1.1 Constituants principaux d'un convertisseur hacheur de tension.

Les convertisseurs de tension continu-continu ont pour fonction de fournir une tension continue variable à partir d'une tension continue fixe. La figure 3.E.1 donne une schématisation "naïve" d'un hacheur de tension (on parle aussi de régulateur à découpage). C'est un simple interrupteur (cellule de commutation) qui s'ouvre et se ferme périodiquement. Si la période de hachage est T l'interrupteur est fermé pendant le temps αT . Le facteur α est appelé facteur cyclique ou facteur de modulation d'impulsion.

Nous pouvons calculer la tension de sortie moyenne U_R :

$$U_R = \langle u_R \rangle = \frac{1}{T} \int_0^T u_R(t) dt = \frac{1}{T} \int_0^{\alpha T} U_e dt + \frac{1}{T} \int_{\alpha T}^T 0 dt,$$

soit

$$\alpha = \frac{U_R}{U_e}. \quad (3.E.1)$$

Pour un convertisseur parfait la puissance d'entrée est égale à la puissance de sortie, ce qui donnerait pour les intensités moyennes :

$$\alpha = \frac{U_R}{U_e} = \frac{I_e}{I_R}.$$

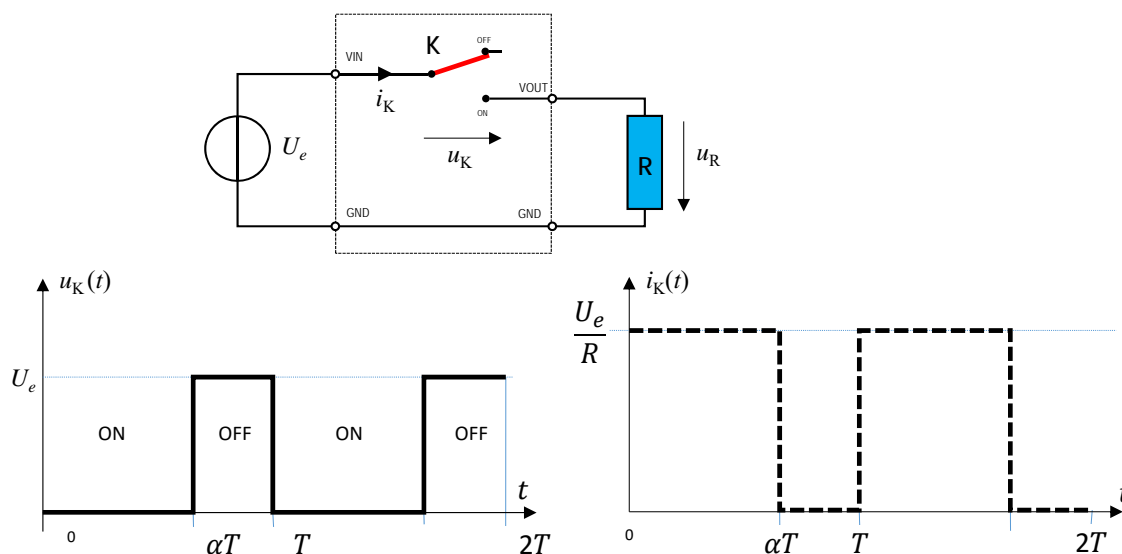


FIGURE 3.E.1 – Schéma de principe d'un hacheur de courant élémentaire. Digramme temporel de tension et d'intensité.

Bien entendu si l'idée est juste ce schéma est beaucoup trop simple. En particulier tension et courant sont totalement discontinus ce qui n'est pas acceptable.

Examinons comment est réalisée la cellule de commutation dans un composant industriel utilisé typiquement pour alimenter des dispositifs électroniques à partir d'une batterie. Le circuit ISL8026A¹ est un hacheur abaisseur de tension² synchrone³. Le schéma bloc est donné sur la figure 3.E.2. Il est beaucoup plus complexe que notre simple interrupteur. En réalité il possède de nombreuses fonctions annexes (limitation de courant, de température, aide à la mise en et hors tension, compensation, etc). Isolons la fonction principale de hachage qu'est la cellule de commutation. Pour cela simplifions de manière drastique le schéma bloc. Nous ne retenons que les éléments essentiels à savoir l'interrupteur à amorçage et à blocage commandés (ici un transistor⁴ type P) et un interrupteur à blocage et amorçage spontanés (MOSFET type N) et l'horloge de commande. Ici les transistors sont utilisés comme des interrupteurs (très

1. Il existe une multitude de circuits intégrés de ce type (8026) et de nombreux fabricants. Ils sont tous basés sur le même principe. Le modèle ISL8026AFRTAJZ-T7A a été retenu comme exemple car il a été étudié par Thales TGS à des fins de validation de modèle et qu'il possède une carte d'évaluation que nous avons testé. Ce composant coûte moins d'un euro pièce.

2. Car la tension moyenne de sortie est inférieure à celle de l'entrée. On utilise aussi le qualificatif "dévolteur" ou "buck"

3. Dans un hacheur synchrone la diode D du hacheur standard est remplacée par l'interrupteur K_2 .

4. C'est un transistor à effet de champ à structure métal-oxyde-semi-conducteur. MOSFET est l'acronyme anglais pour Metal Oxide Semiconductor Field Effect Transistor.

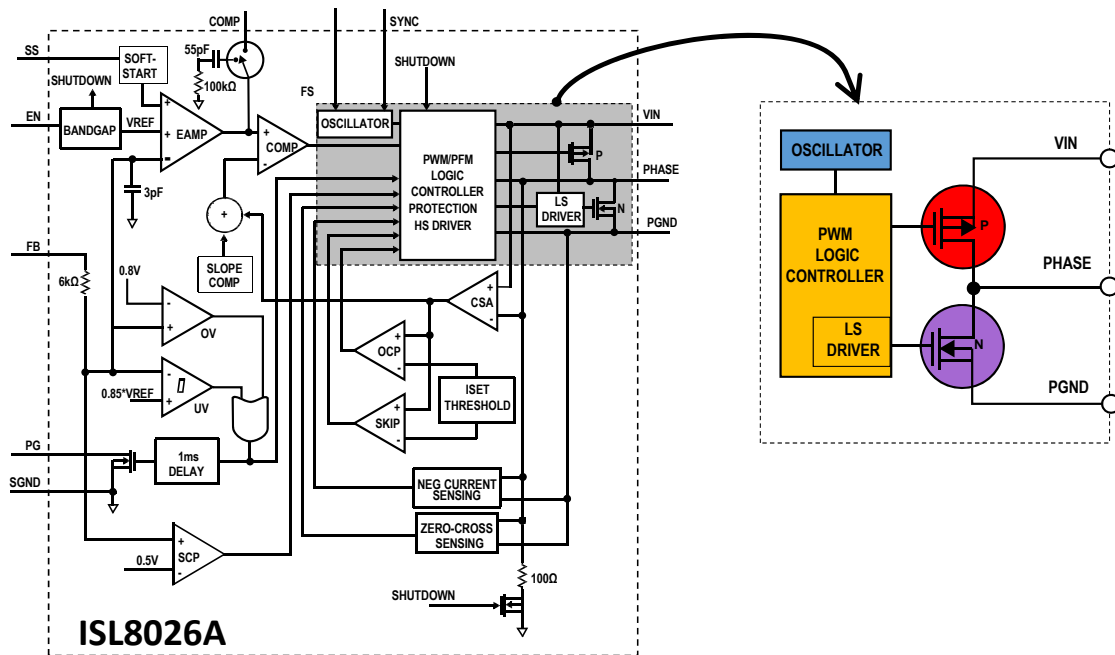


FIGURE 3.E.2 – Bloc diagramme du hacheur ISL8026A d’après [5]. Sur la figure de droite on a extrait du bloc diagramme la cellule de commutation.

rapide !). Sur la figure 3.E.3 on a représenté à gauche deux réalisations classiques de la cellule de commutation d’un hacheur mono-cadrant⁵. On adoptera finalement le schéma représenté à droite de la figure 3.E.3 où les MOSFET ont été remplacés par un couple d’interrupteurs K_1 et K_2 à commande complémentaire ($\{K_1$ ouvert, K_2 fermé $\}$ ou $\{K_1$ fermé, K_2 ouvert $\}$).

A elle seule la cellule de commutation est insuffisante pour réaliser un hacheur de courant. Il faut lui adjoindre un dispositif de stockage dé-stockage rapide d’énergie. La figure 3.E.4 montre un tel dispositif. Le hacheur est alimenté par un générateur de courant délivrant une tension U_e . La charge du hacheur est constituée ici d’une simple résistance R . L’inductance L et la capacité C forment un filtre passe-bas dont le but est de limiter l’ondulation résultant du découpage sur la tension et le courant de sortie. Si ces éléments sont correctement dimensionnés, on peut supposer que $i_s(t)$ et $u_R(t)$ sont continus.

5. Un hacheur mono cadrant ne fonctionne que dans le sens générateur de courant vers récepteur de tension. Dans les hacheurs multicable le récepteur peut renvoyer du courant vers le générateur. C’est par exemple le cas d’un variateur de vitesse d’un moteur à courant continu lorsque la variateur est utilisé en frein.

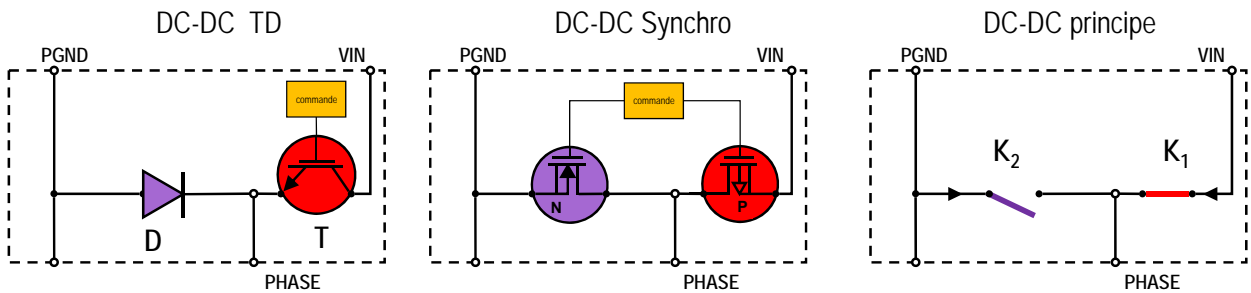


FIGURE 3.E.3 – Schémas de principe de la cellule de commutation d’un hacheur. Les deux premiers schémas proposent des réalisations pratiques. Sur la représentation de gauche l’interrupteur commandé à l’amorçage et au blocage est un IGBT (noté T). Ses commutations commandées entraîneront les commutations spontanées de la diode D. La figure centrale montre une réalisation utilisant deux transistors MOSFET, l’un de type P l’autre de type N, commandés de manière synchronisée. La figure de droite est un schéma général que nous utiliserons pour décrire le fonctionnement. Il comporte deux interrupteurs K_1 et K_2 à commande complémentaire (l’horloge qui cadence le découpage et la commande complémentaire ne sont pas représentées sur ce schéma générique.).

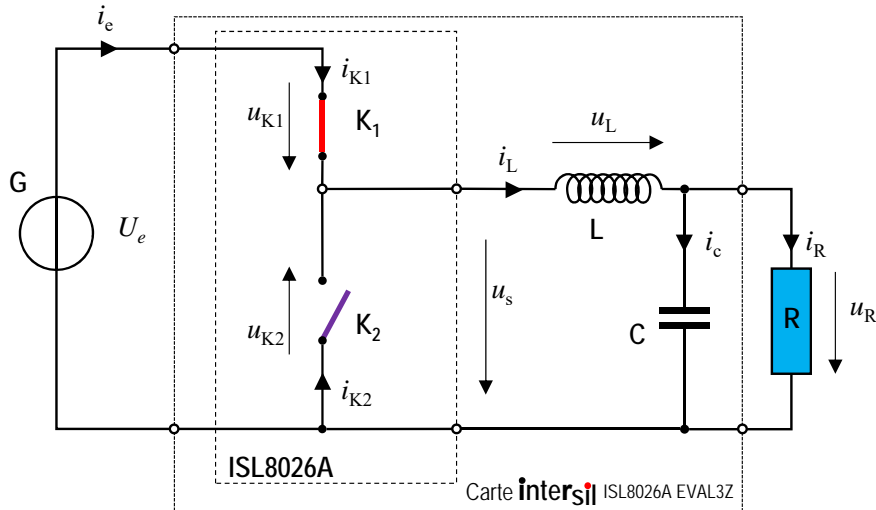


FIGURE 3.E.4 – Schéma de principe d’un hacheur de courant abaisseur de tension.

E.1.1.2 Phases de fonctionnement du hacheur.

La figure 3.E.5 montre les deux phases du cycle de fonctionnement d’un hacheur. On peut écrire la relation :

$$u_s(t) = u_L(t) + u_R(t) = L \frac{d}{dt} i_L + R i_L(t). \quad (3.E.2)$$

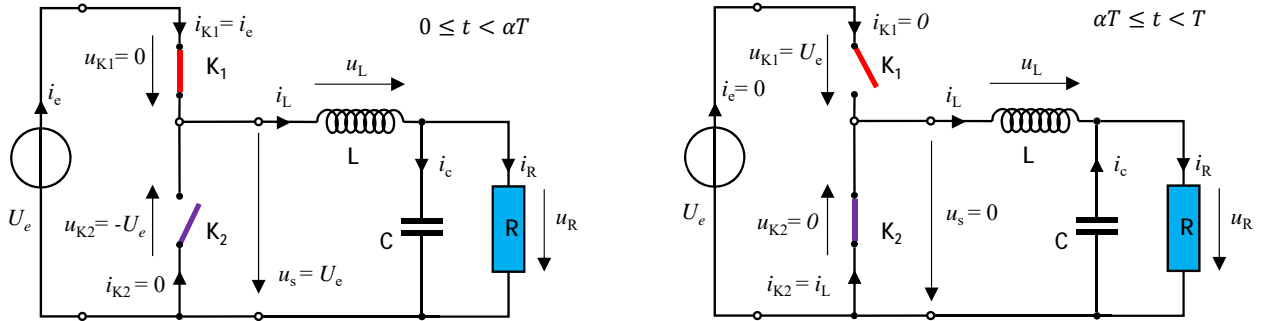


FIGURE 3.E.5 – Phases de fonctionnement d'un hacheur de courant.

E.1.1.2.1 Phase $0 \leq t < \alpha T$. Les interrupteurs K_1 et K_2 sont respectivement fermé et ouvert. Le courant qui traverse K_1 vaut $i_L(t)$. Dans cette phase du cycle l'inductance est soumise à une tension $U_e - u_R > 0$ d'où $\frac{d}{dt}i_L(t) > 0$. Négligeons dans un premier temps la présence de la capacité. Donc puisque dans cette phase $u_s(t) = U_e$ et en notant $\tau = R/L$ la constante de temps nous obtenons après intégration :

$$i_L(t) = I_m + \frac{U_e}{R}(1 - \exp -\frac{t}{\tau}). \quad (3.E.3)$$

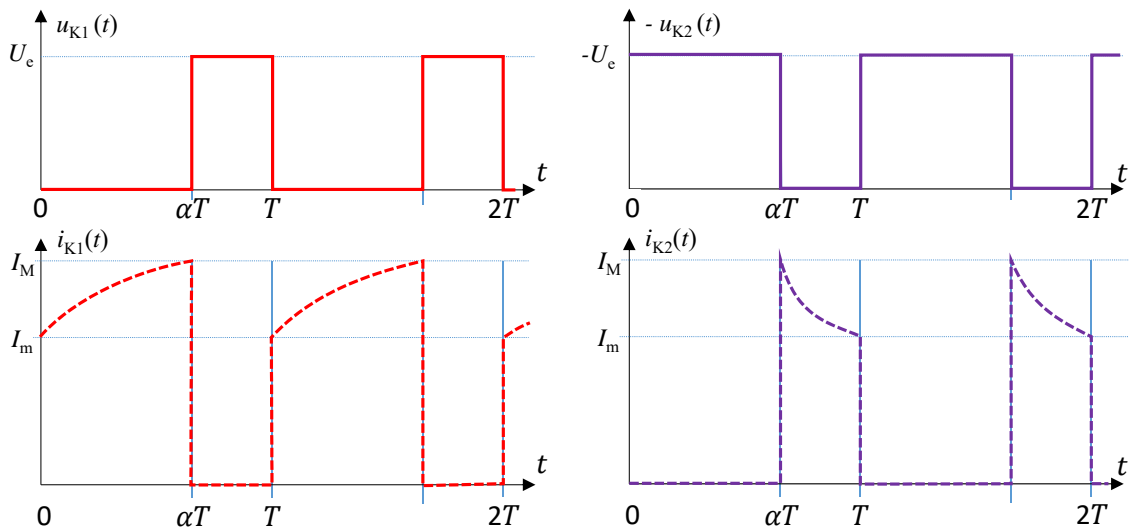


FIGURE 3.E.6 – Digramme temporel de l'évolution de la tension aux bornes des interrupteurs K_1 et K_2 et du courant les traversant.

E.1.1.2.2 Phase $\alpha T \leq t < T$. Cette fois les interrupteurs K_1 et K_2 sont respectivement ouvert et fermé. Dans cette phase du cycle l'inductance est soumise à une tension $-U_e < 0$ (car $u_s = 0$) d'où $\frac{d}{dt}i_L(t) < 0$. L'intégration de (3.E.2) donne :

$$i_L(t) = I_M + \frac{U_e}{R} \exp -\frac{t}{\tau}. \quad (3.E.4)$$

E.1.1.2.3 Raccordement des deux phases du cycle de hachage. Déterminons les constantes I_M et I_m en utilisant la continuité de l'intensité. On montre immédiatement :

$$I_m = \frac{U_e}{R} \frac{1 - \exp \alpha T / \tau}{1 - \exp T / \tau}, \quad I_M = \frac{U_e}{R} \frac{1 - \exp -\alpha T / \tau}{1 - \exp -T / \tau}. \quad (3.E.5)$$

Mais $T \gg \tau$ pour ce type de système. En développant au second ordre on trouve en posant $f = 1/T$:

$$\Delta I_L = I_M - I_m \cong \frac{U_e}{L f} \alpha (1 - \alpha). \quad (3.E.6)$$

On peut remarquer que l'oscillation maximum sera atteinte pour un rapport cyclique $\alpha = 0,5$, autrement dit $\Delta I_{Lmax} = \frac{U_e}{4L f}$. L'oscillation sera d'autant plus faible que la fréquence de hachage sera élevée (mais dans ce cas la puissance dissipée augmente aussi).

Pendant la phase de conduction de K_1 on a $i_e(t) = i_L(t)$ alors que pendant la phase de blocage (K_1 ouvert) $i_e = 0$. On a alors si $\tau \ll T$:

$$I_e = \langle i_e(t) \rangle = \frac{1}{T} \int_0^T i_e(t) dt = \frac{1}{T} \int_0^{\alpha T} i_L(t) dt = \frac{1}{T} \int_0^{\alpha T} \left(\frac{\Delta I_L}{\alpha T} t + \left(I_L - \frac{\Delta I_L}{2} \right) \right) dt,$$

d'où

$$I_e = \alpha I_L. \quad (3.E.7)$$

E.1.1.2.4 Prise en compte simplifiée de la capacité de lissage. La capacité absorbe l'oscillation de courant de l'inductance. En notant $\tilde{a}(t)$ l'oscillation d'une grandeur a telle que $a(t) = A + \tilde{a}(t)$ avec $A = \langle a(t) \rangle$ on peut écrire :

$$i_c(t) = C \frac{d}{dt} \tilde{u}_R(t) \sim \tilde{i}_L(t).$$

L'oscillation de tension aux bornes de la résistance est donc proportionnelle à l'intégrale de l'oscillation de l'intensité traversant l'inductance. On trouve après calcul compte tenu des expressions (3.E.3) et (3.E.4) :

$$\Delta U_R = \frac{U_e}{8 L C f^2} \alpha (1 - \alpha). \quad (3.E.8)$$

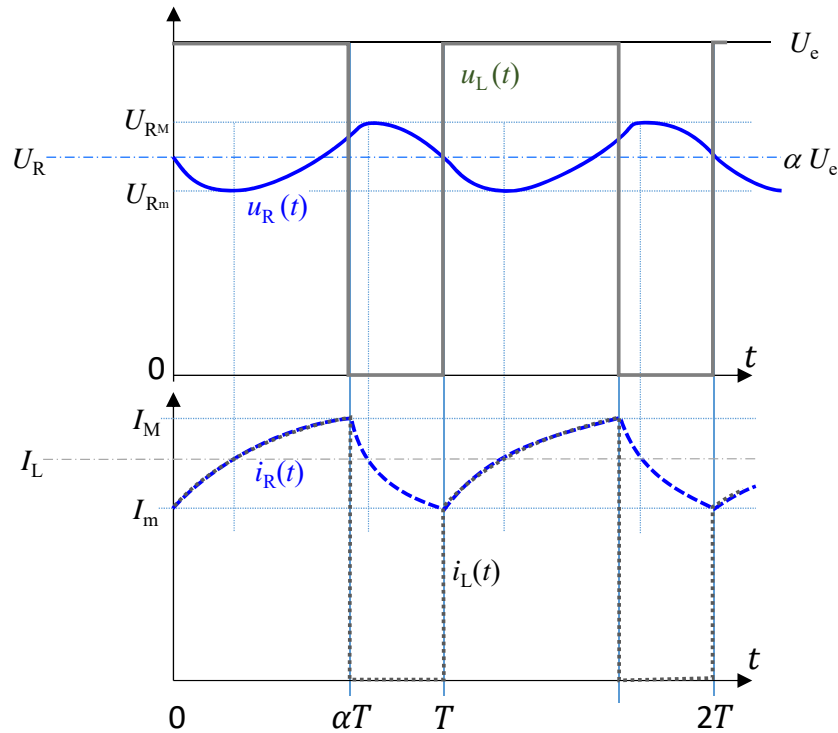


FIGURE 3.E.7 – Diagramme temporel des signaux de sortie du hacheur et de l'inductance. Ici les ondulations ont été très amplifiées par rapport à la réalité. Ces signaux peuvent être enregistrés sur la carte d'évaluation.

L'oscillation de la tension de sortie est donc beaucoup plus faible que celle de l'intensité (dépendance quadratique à la fréquence de hachage). En pratique on peut pratiquement considérer que la tension et l'intensité dans la charge vérifie respectivement $u_R(t) = \alpha U_e$ et $i_R(t) = I_e/\alpha$.

E.1.1.2.5 Bilan de puissance pour le hacheur idéal. La puissance fournie par le générateur de tension est $\Pi_e = U_e \langle i_e(t) \rangle$ sachant que $i_e(t) = i_{K_1}(t)$ avec $i_{K_1}(t) = i_L(t)$ pour $t \in [0, \alpha T]$ et $i_{K_1}(t) = 0$ pour $t \in [\alpha T, T]$. Après intégration on trouve :

$$\Pi_e = \alpha^2 \frac{U_e^2}{R} + U_e \frac{\tau}{T} \left(I_m - \frac{U_e}{R} \right) (1 - \exp -\alpha \frac{T}{\tau}). \quad (3.E.9)$$

Cette puissance est égale à la puissance dissipée dans la résistance de charge. Avec l'approximation $\tau \ll T$ qui correspond au critère de dimensionnement de l'inductance par rapport à la fréquence de hachage et à la résistance de charge on peut écrire :

$$\Pi_R = \Pi_e \sim \alpha^2 \frac{U_e^2}{R}. \quad (3.E.10)$$

Puisque $U_R = \alpha U_e$ on peut en déduire que l'intensité efficace dans la résistance est $I_R = \frac{1}{\alpha} I_e$.

E.1.2 Origine de la dissipation thermique dans le hacheur.

En réalité le hacheur n'est pas parfait et une partie de la puissance fournie par le générateur de tension est dissipée en chaleur. C'est cette puissance qui nous intéresse particulièrement pour nos modèles thermiques. On définit le rendement η du hacheur par :

$$\Pi_e = \Pi_R + \Pi_{th}, \quad \eta = \frac{\Pi_R}{\Pi_e}. \quad (3.E.11)$$

Expression où Π_{th} est la puissance thermique dissipée dans le hacheur. On peut la scinder en deux parties :

Dissipation dans la cellule de commutation [97] :

- Les transistors possèdent une résistance interne (plus forte pour le type P que pour le type N). Dans le cas de l'ISL8026A le constructeur indique les valeurs suivantes : $R_p = 36 \text{ m}\Omega$ et $R_n = 13 \text{ m}\Omega$. Il y a donc des pertes par effet Joule.
- Les fronts montants et les descendants ne sont pas des échelons parfaits mais plutôt des rampes. Pendant la durée de commutation, notée δ le produit de la tension par l'intensité n'est pas nul. Il y a donc une perte par commutations. Si on assimile la variation à une rampe alors la puissance perdue à chaque commutation du transistor est $UI/6$ (on a $U_e I_e = U_R I_R$ et $\Delta I_R \ll I_R$). Il y a 4 commutations par cycle. Les mesures donnent approximativement $\delta \sim 10 \text{ ns}$.

On peut tenter de construire un modèle élémentaire de dissipation dans le QFN pour voir l'influence des paramètres :

$$\Pi_{th}^{QFN} = \alpha R_p I_e^2 + (1 - \alpha) R_n I_R^2 + \frac{2}{3} \delta f U_e I_e. \quad (3.E.12)$$

Dissipation dans l'inductance [98] :

- Le fil de la bobine de l'inductance dissipe de l'énergie par effet Joule. Si on note r_L la résistance de la bobine, la puissance dégagée sera $r_L I_L^2$. La résistance de la bobine serait faible en régime continu. Cependant au delà de quelques centaines de kHz il faut considérer l'effet de peau qui augmente considérablement la perte par effet Joule. C'est donc le cas pour l'ISL8026A puisque qu'il fonctionne à une fréquence de 2 MHz . A cette fréquence l'épaisseur de peau pour un fil de cuivre est de $47 \mu\text{m}$, que l'on peut rapporter au diamètre du fil qui est de $400 \mu\text{m}$. Logiquement la résistance sera une fonction de la fréquence et la puissance sera dissipée en "surface" du fil.
- Il y a aussi des pertes par hystérésis (le champ magnétique s'inverse périodiquement) et par courants de Foucault. Elles sont proportionnelles respectivement à la fréquence et

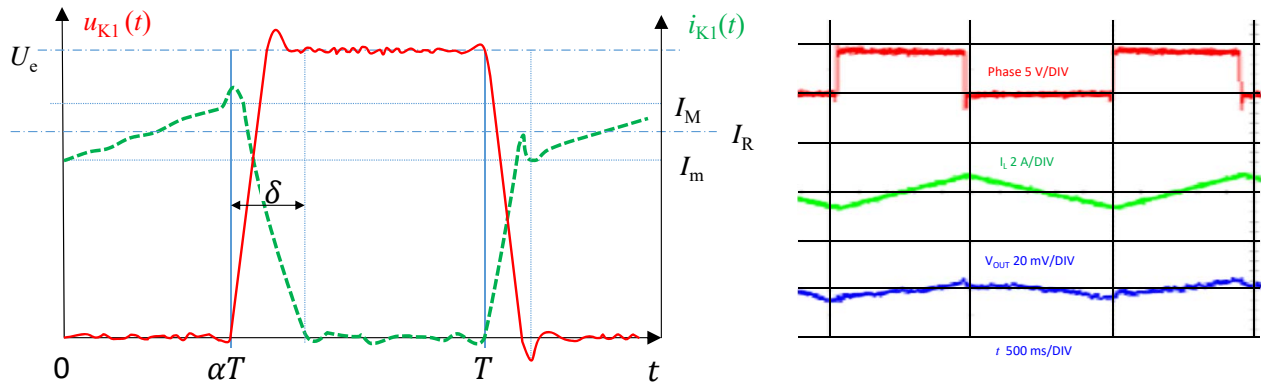


FIGURE 3.E.8 – Diagramme temporel montrant le défaut de commutation. À gauche schéma de principe (défaut amplifiés), à droite enregistrement des signaux sur un oscilloscope numérique.

au carré de la fréquence du courant qui traverse la bobine. Dans notre cas la bobine de l'inductance est enrobée dans un matériau magnétique métallique (integrated molded coil). La puissance sera dissipée dans la totalité du volume.

- Les pertes par effet condensateur entre les spires de la bobine ne sont significatives que pour des très hautes fréquences (VHF, UHF) c'est à dire au delà de $30MHz$ ce qui n'est pas notre cas.

Le fabricant de l'inductance montée sur la carte (TDK) donne une formule de dissipation thermique simple :

$$\Pi_{th}^L = R_{ac} I_R^2. \quad (3.E.13)$$

La résistance R_{ac} globalise l'ensemble des phénomènes générant la dissipation thermique. C'est une fonction de la fréquence. Elle est donnée sur la figure 3.E.9.

Notons que la dissipation dans les condensateurs est marginale. Il en est de même dans les résistances utilisées pour la partie contrôle.

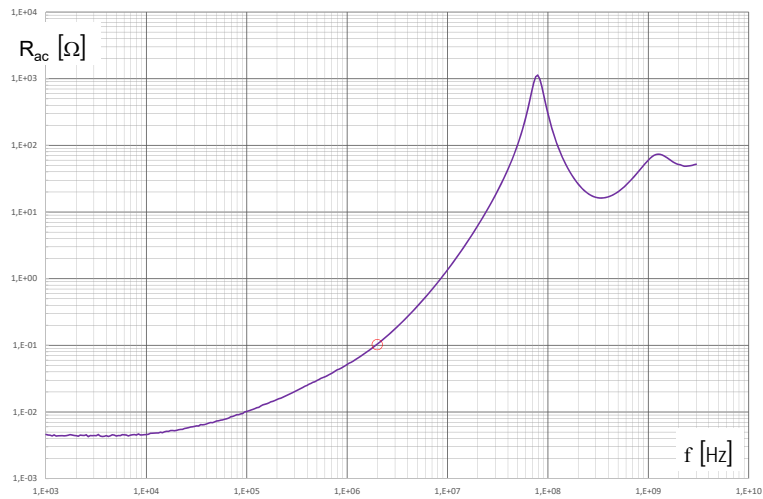


FIGURE 3.E.9 – Variation de la résistance d'une inductance. D'après document TDK.

Annexe F

Réduction modale sous-structurée avec des maillages non-conformes en 2D

L'étude présentée dans cette annexe a fait l'objet d'une publication au Congrès de la Société Française de Thermique 2018, qui s'est tenu à Pau du 29 mai au 1er juin 2018 à Pau. L'ensemble de la publication, qui a fait l'objet d'une reproduction dans le Bulletin de Liaison 2018 N° 3 de la SFT est reproduit ci-après.

Réduction d'un problème thermique par sous-structuration modale sur des maillages non-conformes

Sébastien GROSJEAN^{1,2*}, Frédéric JOLY¹, Karine VERA², Alain NEVEU¹

¹ Université Paris-Saclay, Laboratoire de Mécanique et d'Energétique d'Evry, Antenne de Brétigny sur Orge, IUT d'Evry Val d'Essonne, Département GTE, Chemin de la Tuilerie, 91731 Brétigny sur Orge

² Thales Communications & Security, 4 Avenue des Louvresses, 92230 Gennevilliers

* (auteur correspondant : s.grosjean@iut.univ-evry.fr)

Résumé - Cette communication traite du raccord de modèles réduits modaux calculés indépendamment les uns des autres (sous-structuration non conforme). Cette approche offre plus de souplesse par rapport à la technique de sous-structuration conforme développée antérieurement par notre équipe. Elle se rapproche ainsi de la pratique de description des systèmes par les industriels. La méthode des joints écrite initialement pour les éléments finis a été adaptée au formalisme modal. Pour cela, des nouvelles fonctions de couplage aux interfaces sont introduites. Un exemple en dimension 2 illustre la méthode.

1. Introduction

De nombreuses industries (électronique, aérospatiale, ...) se servent de composants utilisés pour de multiples applications. Il est alors intéressant de créer des modèles réduits caractérisant le comportement thermique de ces briques élémentaires, et de les archiver dans des bibliothèques de composants. Puisque certains composants peuvent être employés des dizaines de fois, le temps passé à créer les modèles réduits sera rentabilisé. Dans les industries citées plus haut, la méthode de réduction utilisée est la méthode nodale [1]. Si ces modèles ont fait leurs preuves et donnent des résultats tout à fait satisfaisants pour les applications actuelles, ils pourraient atteindre leurs limites si la résolution spatiale demandée augmente. En effet, dans ces méthodes le nombre d'observables est directement lié au nombre d'inconnues. L'augmentation de la résolution spatiale s'accompagne donc mécaniquement d'une augmentation du nombre d'inconnues. Par ailleurs, la détermination du réseau de résistances thermiques couplant les différents nœuds deviendra aussi problématique. Enfin, les modèles nodaux rencontrent une limite dans la représentation d'entités complexes comme des circuits imprimés.

Les méthodes modales découplent complètement la notion d'inconnue et celle d'observable. Elles consistent à rechercher la solution sous la forme d'une somme pondérée de champs élémentaires, appelés modes. Elles ont été étendues à des ensembles multi-composants par l'introduction de la méthode de sous-structuration modale [2]. Il s'agit de décomposer le système en sous-structures, de calculer les modes de chacune de ces entités, puis de les rassembler pour résoudre le problème d'origine. Cette méthode répond à la problématique exposée ci-dessus, puisqu'elle permet en théorie le couplage de modèles modaux créés indépendamment. À l'heure actuelle, le couplage utilisé nécessite d'avoir un maillage conforme aux interfaces : les deux composants devant être couplés entre eux doivent être maillés en même temps pour que les nœuds coïncident. Cette limitation empêche le développement de bibliothèques de modèles réduits modaux, puisque un modèle modal doit être recalculé pour chaque nouveau maillage.

Nous proposons dans cette communication une méthode originale de réduction modale inspirée de la méthode des joints [3] permettant de coupler entre eux des modèles modaux dont les maillages ne sont pas conformes aux interfaces. Nous décrivons dans une première partie le principe de la méthode dans un cadre général, puis nous l'illustrons par un exemple bi-dimensionnel.

2. Modèle mathématique

2.1. La méthode des joints

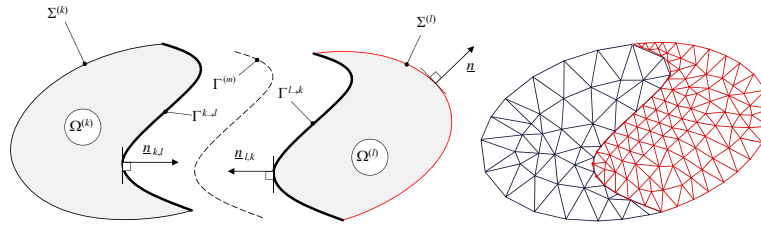


Figure 1 - Décomposition d'un domaine, définition des frontières, maillage non-conforme

La méthode des joints, ou “mortar method” a été développée initialement au laboratoire Jacques-Louis Lions dans les années 90 pour résoudre le problème de deux domaines en contact et dont les maillages ne se raccordent pas [3]. On considère un domaine $\Omega = \Omega^{(1)} \cup \Omega^{(2)}$, divisé par une interface Γ . On note $\partial\Omega = \cup_k \Sigma^{(k)}$ la frontière séparant Ω du milieu extérieur (voir Fig. 1). Sur ce domaine, l'évolution temporelle de la température est modélisée par l'équation de la chaleur :

$$c \frac{\partial T}{\partial t} = \underline{\nabla} \cdot \underline{\kappa} \cdot \underline{\nabla} T + \varpi \quad \text{sur } \Omega \quad (1)$$

$$\underline{\kappa} \cdot \underline{\nabla} T \cdot \underline{n} = h(T_f - T) \quad \text{sur } \partial\Omega \quad (2)$$

À ces équations, il faut rajouter les deux conditions de raccord de part et d'autres de l'interface Γ , la continuité du flux de chaleur et des températures. Ici, on suppose un contact parfait¹. La conformité géométrique impose $\underline{n}_{1,2} = -\underline{n}_{2,1}$, et nous posons arbitrairement $\underline{n} = \underline{n}_{2,1}$.

$$\underline{\kappa}^{(1)} \cdot \underline{\nabla} T^{(1)} \cdot \underline{n} = \underline{\kappa}^{(2)} \cdot \underline{\nabla} T^{(2)} \cdot \underline{n} = \varphi \quad (3)$$

$$T^{(2)}|_{\Gamma} - T^{(1)}|_{\Gamma} = 0 \quad (4)$$

La formulation variationnelle faible de l'équation (1) s'écrit en intégrant sur les domaines $\Omega^{(k)}$ et en multipliant par une fonction test $g^{(k)} \in H^1(\Omega^{(k)})$, qui est ici l'espace fonctionnel adapté dans lequel nous recherchons la solution.

$$\begin{aligned} \sum_{k=1,2} \int_{\Omega^{(k)}} c^{(k)} \frac{\partial T^{(k)}}{\partial t} g^{(k)} = & - \sum_{k=1,2} \int_{\Omega^{(k)}} \underline{\nabla} T^{(k)} \cdot \underline{\kappa}^{(k)} \cdot \underline{\nabla} g^{(k)} + \sum_{k=1,2} \int_{\Omega} \varpi^{(k)} g^{(k)} \\ & + \sum_{k=1,2} \int_{\Sigma^{(k)}} \underline{\kappa}^{(k)} \cdot \underline{\nabla} T^{(k)} \cdot \underline{n}^{(k)} g^{(k)} d\sigma + \sum_{k=1,2} \int_{\Gamma} \underline{\kappa}^{(k)} \cdot \underline{\nabla} T^{(k)} \cdot \underline{n}^{(k)} g^{(k)} d\sigma \end{aligned} \quad (5)$$

1. Rien ne s'oppose à la prise en compte d'une résistance de contact

Nous reconnaissons dans les termes intégraux de surface la densité de flux de conduction $\underline{\kappa} \cdot \underline{\nabla} T \cdot \underline{n}$. Le terme intégral sur la frontière extérieure $\partial\Omega$ s'exprime classiquement à l'aide de l'Éq. (2). En revanche, à l'interface de couplage, le flux n'est pas exprimé en fonction de températures, mais devient une nouvelle variable. En utilisant la contrainte de continuité du flux de chaleur à l'interface de raccordement (Eq. (3)), l'Eq. (5) se réécrit :

$$\int_{\Omega} c \frac{\partial T}{\partial t} g = - \int_{\Omega} \underline{\nabla} T \cdot \underline{\kappa} \cdot \underline{\nabla} g + \int_{\Omega} \varpi g + \int_{\partial\Omega} h(T_f - T) g d\sigma + \int_{\Gamma} \varphi \llbracket g \rrbracket d\sigma \quad (6)$$

où l'on a défini le saut de fonction test $\llbracket g \rrbracket = g^{(2)}|_{\Gamma} - g^{(1)}|_{\Gamma}$. Le flux de conduction, en tant qu'inconnue supplémentaire, est exprimé sur la base de nouvelles fonctions :

$$\varphi = \sum_i \varphi_i \phi_i(\underline{x}) \quad (7)$$

où les $\phi_i(\underline{x})$, que nous appellerons "fonctions mortier", sont des fonctions spatiales à préciser ultérieurement. La continuité de la température (4) est imposée au sens faible, le saut de température à l'interface devant être orthogonal aux fonctions mortiers :

$$\int_{\Gamma} \llbracket T \rrbracket_{\Gamma} \phi = 0 \quad (8)$$

L'équation (8) a la particularité de ne pas forcer la continuité de la solution numérique sur les interfaces, ce qui permet de pouvoir résoudre les équations localement, et donc de recourir à différents maillages de part et d'autre de ces interfaces. La résolution des équations (6)-(8) en éléments finis a fait l'objet de très nombreuses publications (voir par exemple [4]). En revanche, à notre connaissance, il n'existe pas de mise en œuvre modale de cette méthode. Pourtant, le formalisme des méthodes modales se prête bien à la méthode des joints comme nous le montrons à présent.

2.2. Applications aux méthodes modales

Nous cherchons la température comme une somme pondérée de fonctions élémentaires. Afin d'assurer la continuité aux interfaces, nous calculons deux bases. La première est issue du problème aux valeurs propres de Dirichlet [5] (température imposée nulle aux frontières). La seconde est issue du problème aux valeurs propres de Steklov [6]. On désigne respectivement par (V^D, λ^D) et (V^S, λ^S) les couples formés par le mode propre et sa valeur propre associée pour les problèmes de Dirichlet et de Steklov :

$$\begin{array}{l|l} \Omega & -\underline{\nabla} \cdot \underline{\kappa} \cdot \underline{\nabla} V^D = \lambda^D c V^D \\ \partial\Omega & V^D = 0 \end{array} \quad \left| \quad \begin{array}{l} \underline{\nabla} \cdot \underline{\kappa} \cdot \underline{\nabla} V^S = 0 \\ \underline{\kappa} \cdot \underline{\nabla} V^S \cdot \underline{n} = \lambda^S \zeta S, \quad S = V^S|_{\partial\Omega} \end{array} \right. \quad (9)$$

Les modes de Dirichlet étant insuffisants pour reconstruire la solution sur les bords, on leur adjoint les modes de Steklov qui permettent de reconstituer l'hétérogénéité des conditions aux limites. Le coefficient de Steklov ζ , dont la valeur est arbitraire, est choisi de façon à compenser les variations de conductivité dans le cas de matériaux hétérogènes. La température est alors cherchée comme

$$T(\underline{x}, t) = \sum_i V_i^D(\underline{x}) x_i^D(t) + \sum_j V_j^S(\underline{x}) x_j^S(t) \quad (10)$$

Avec cette formulation, les inconnues du système deviennent les coefficients de pondération x_i^D et x_j^S , appelés états d'excitation.

Les flux sont aussi cherchés comme une décomposition sur une base de fonctions (voir Eq. (7)). Plusieurs options sont envisageables pour ces fonctions. En nous inspirant de l'allure des fonctions de forme présentées dans de nombreuses publications ([7] par exemple), mais en nous affranchissant à ce stade de la notion de maillage, nous avons choisi d'exprimer les flux sur une base de Neumann :

$$-\nabla^2 \phi = \lambda \phi \quad \text{sur } \Gamma \quad | \quad \underline{\nabla} \phi \cdot \underline{n} = 0 \quad \text{sur } \partial\Gamma \quad (11)$$

où $\partial\Gamma$ est la frontière de l'interface de contact.

Pour trouver les équations des états des modes propres de Dirichlet et de Steklov et des flux, il suffit de remplacer dans (6) et (8) la température et le flux respectivement par leur expression (7) et (10) et d'utiliser comme fonctions tests les modes propres de Dirichlet-Steklov ainsi que les fonctions ϕ . Nous introduisons quatre matrices modales de conduction, de capacité, de convection et de joint, définies dans le cadre de N_Ω domaines couplés via N_c interfaces :

$$\begin{aligned} \mathbb{K}_{\mathcal{X}\mathcal{Y}}^{(k)} &= \left(\int_{\Omega^{(k)}} \underline{\nabla} V_i^{\mathcal{X},(k)} \cdot \underline{\underline{\kappa}}^{(k)} \cdot \underline{\nabla} V_j^{\mathcal{Y},(k)} \right)_{\substack{1 \leq i \leq N_{\mathcal{X}}^{(k)} \\ 1 \leq j \leq N_{\mathcal{Y}}^{(k)}}} & \mathbb{C}_{\mathcal{X}\mathcal{Y}}^{(k)} &= \left(\int_{\Omega^{(k)}} c^{(k)} V_i^{\mathcal{X},(k)} V_j^{\mathcal{Y},(k)} \right)_{\substack{1 \leq i \leq N_{\mathcal{X}}^{(k)} \\ 1 \leq j \leq N_{\mathcal{Y}}^{(k)}}} \\ \mathbb{H}_{cv}^{(k)} &= \left(\int_{\Sigma^{(k)}} h^{(k)} S_i^{(k)}|_{\Sigma^{(k)}} S_j^{(k)}|_{\Sigma^{(k)}} \right)_{\substack{1 \leq i \leq N_S^{(k)} \\ 1 \leq j \leq N_S^{(k)}}} & \mathbb{J}_m^{(k)} &= s_m^k \left(\int_{\Gamma^{(m)}} S_i^{(k)}|_{\Gamma^{(m)}} \phi_p^{(m)} \right)_{\substack{1 \leq i \leq N_S^{(k)} \\ 1 \leq p \leq N_\phi^{(m)}}} \end{aligned}$$

Ici, $\mathcal{X}, \mathcal{Y} \in \{\text{D}, \text{S}\}$, $N_{\mathcal{X}}^{(k)}$ est le nombre de modes de type \mathcal{X} du sous-domaine (k) , et $N_\phi^{(m)}$ le nombre de fonctions mortier associés à la frontière $\Gamma^{(m)}$. La matrice $\mathbb{J}_m^{(k)}$ couple $\Omega^{(k)}$ et $\Omega^{(l)}$ via la frontière $\Gamma^{(m)}$. Le facteur s_m^k vaut (-1) si le mortier est défini du côté de la structure k , (1) si il est défini du côté de l (donc $s_m^k = -s_m^l$) et, bien sûr, 0 si k et l ne sont pas en contact.

De même, introduisons les vecteurs des états modaux, des coefficients du flux et des sollicitations définis respectivement par :

$$\begin{aligned} X_{\mathcal{X}}^{(k)} &= \left(x_j^{\mathcal{X},(k)} \right)_{1 \leq j \leq N_{\mathcal{X}}^{(k)}} & X_\varphi^{(m)} &= (\varphi_j)_{1 \leq j \leq N_\phi^{(m)}} \\ \Pi_{\mathcal{X}}^{(k)} &= \left(\int_{\Omega^{(k)}} \varpi^{(k)} V_i^{\mathcal{X},(k)} \right)_{1 \leq i \leq N_{\mathcal{X}}^{(k)}} & \Theta_{ex}^{(k)} &= \left(\int_{\Sigma^{(k)}} h^{(k)} T_{ex}^{(k)} S_i^{(k)}|_{\Sigma^{(k)}} \right)_{1 \leq i \leq N_S^{(k)}} \end{aligned}$$

Nous obtenons, après assemblage des sous-structures, les matrices bloc associées où \mathbb{A} représente \mathbb{C} , \mathbb{K} , ou \mathbb{H} , et où la matrice de couplage \mathbb{J} est de dimension $N_\Omega \times N_c$:

$$\mathbb{A}_{\mathcal{X}\mathcal{Y}} = \begin{bmatrix} \mathbb{A}_{\mathcal{X}\mathcal{Y}}^{(0)} & 0 & 0 & 0 \\ 0 & \mathbb{A}_{\mathcal{X}\mathcal{Y}}^{(1)} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \mathbb{A}_{\mathcal{X}\mathcal{Y}}^{(N_\Omega)} \end{bmatrix} \quad \mathbb{J} = \begin{bmatrix} \mathbb{J}^{(0,0)} & \mathbb{J}^{(0,1)} & \dots & \mathbb{J}^{(0,N_c)} \\ \mathbb{J}^{(1,0)} & \mathbb{J}^{(1,1)} & \dots & \mathbb{J}^{(1,N_c)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{J}^{(N_\Omega,0)} & \mathbb{J}^{(N_\Omega,1)} & \dots & \mathbb{J}^{(N_\Omega,N_c)} \end{bmatrix}$$

Finalement, le système d'équations modales instationnaires sous-structuré s'écrit sous forme matricielle condensée :

$$\begin{bmatrix} \mathbb{C}_{DD} & \mathbb{C}_{DS} & 0 \\ {}^T\mathbb{C}_{DS} & \mathbb{C}_{SS} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} \dot{X}_D \\ \dot{X}_S \\ \dot{X}_\varphi \end{pmatrix} + \begin{bmatrix} \mathbb{K}_{DD} & \mathbb{K}_{DS} & 0 \\ {}^T\mathbb{K}_{DS} & \mathbb{K}_{SS} + \mathbb{H}_{cv} & \mathbb{J} \\ 0 & {}^T\mathbb{J} & 0 \end{bmatrix} \begin{pmatrix} X_D \\ X_S \\ X_\varphi \end{pmatrix} = \begin{pmatrix} \Pi_D \\ \Pi_S + \Theta_{ex} \\ 0 \end{pmatrix} \quad (12)$$

	Plots de connexion et dissipateur thermique	Puce	Fil	Piste en cuivre	Résine	Diélectrique	Radiateur
c ($MJ.m^{-3}.K^{-1}$)	3.43	1.75	3.42	3.55	1.65	1.63	2.42
κ ($W.m^{-1}.K^{-1}$)	260	150	317	385	0.66	0.3	237

Tableau 1 - Propriétés physiques des matériaux utilisés dans cette étude

Sous cette forme, et même en dehors de toute considération sur les maillages, la méthode est non conforme car les bases sont réduites. C'est une non conformité "fonctionnelle". Bien entendu si en plus le problème est résolu par une approximation aux éléments finis cette non conformité fonctionnelle peut se doubler d'une non conformité "du maillage". La structure de ce système matriciel est très semblable à celle obtenue par la méthode des joints en éléments finis. La différence essentielle est la dimension des matrices qui est plus faible pour le modèle modal avec des bases réduites. Le problème matriciel obtenu est un problème dit de "point-selle", et est résolu par l'algorithme d'Uzawa. Un schéma d'Euler implicite au premier ordre est utilisé pour la discrétisation temporelle.

3. Illustration de la méthode

Le cas test étudié est donné à titre d'illustration. Il représente schématiquement un composant électronique (Integrated Circuit, IC) "posé" sur un circuit imprimé (Printed Circuit Board, PCB). Si pour des raisons pédagogiques les dimensions ne sont délibérément pas réalistes, la conductivité thermique s'étend sur une large gamme représentative des composants électroniques (les propriétés sont données dans le tableau 1). Au sein du composant, une puce en silicium dégage une puissance thermique de $5 \times 10^6 W.m^{-3}$, qu'un dissipateur thermique permet de drainer en majorité vers le milieu extérieur. Un radiateur peut éventuellement être placé entre le dissipateur thermique et l'air ambiant. Deux radiateurs seront testés (composants 3 et 4, voir Fig. 3). L'information électrique générée par la puce est transmise à un circuit imprimé via des fils en or et des plots de connexion. Une partie de la chaleur dégagée par la puce se propage via ces fils, chauffant les plots de connexion, ce qui engendre des contraintes thermo-mécaniques indésirables. Le circuit imprimé est quant à lui modélisé comme un matériau diélectrique traversé par des pistes en cuivre.

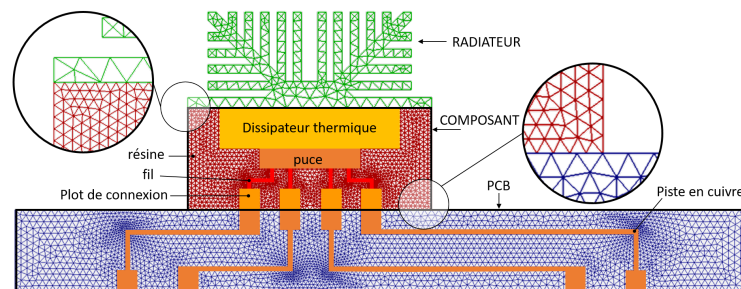


Figure 2 - Géométrie du cas traité

Le problème a été séparé en différentes sous-structures (radiateur, IC, PCB) et chacune a été maillée séparément. Nous présentons sur la figure 2 un agrandissement des maillages. On

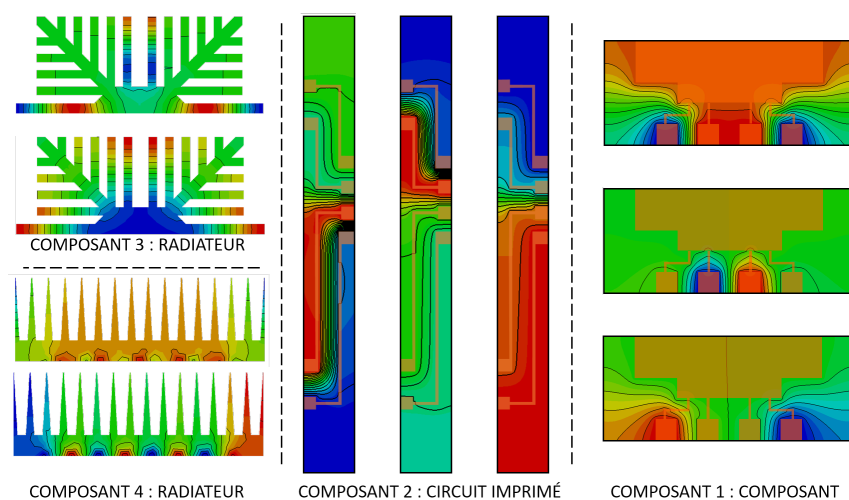


Figure 3 - Bibliothèque modale de composant

remarque que les nœuds ne coïncident pas aux interfaces, sauf aux points définissant la géométrie. De plus, si la taille des mailles est du même ordre de grandeur entre le circuit imprimé et le composant, ce n'est pas le cas entre le composant et le radiateur. Le nombre de nœuds des maillages du IC, du PCB et des deux radiateurs est respectivement de 2241, 3251, 766 et 321.

Les modes de Dirichlet-Steklov de chaque sous-structure sont alors calculés indépendamment. La figure 3 présente des modes de Steklov disposés dans une bibliothèque de composant. Les différences de propriétés sculptent les modes. On voit ainsi les isothermes des modes contourner les parties les plus conductrices pour se concentrer dans les parties isolantes où l'on s'attend à obtenir les gradients les plus importants. On remarque aussi des modes concentrés autour des plots de connexion, et qui permettent le couplage entre les sous-structures.

Toutefois, dans cette méthode, les modes de Steklov ne sont plus les seuls à assurer le raccordement. Au lieu d'imposer les deux conditions de continuité à l'aide de la seule température (*i.e.* par les seuls modes de Steklov), l'ajout des fonctions mortiers donne un degré de liberté supplémentaire pour assurer séparément la compatibilité des températures et des flux sur l'interface. De façon imagée et enfantine, on assemble des briques Lego® et Duplo® (les modes de Steklov de chaque sous-structure) en interposant une couche de pâte à modeler (les fonctions mortiers). Bien entendu, mathématiquement les fonctions mortier n'ont pas d'épaisseur.

Pour illustrer les potentialités qu'offre cette méthode, trois cas transitoires sont simulés à l'aide de la même bibliothèque de composants. Suivant [4], le nombre de fonctions mortiers par frontière est égal au nombre de nœuds sur celle-ci moins 2. Notons qu'avec notre méthode, le nombre de fonctions mortiers peut être réduit.

Dans le cas (a), l'ensemble IC et PCB est refroidi par convection naturelle ($h = 10 \text{ W.m}^{-2}.K^{-1}$). La figure 4 montre pour un même maillage le champ de température à l'équilibre obtenu après 10^4 s par sous-structuration avec la méthode des éléments finis (à gauche) et la méthode modale (à droite). L'ordre de réduction choisi est de 50 modes par sous-structure. On constate que

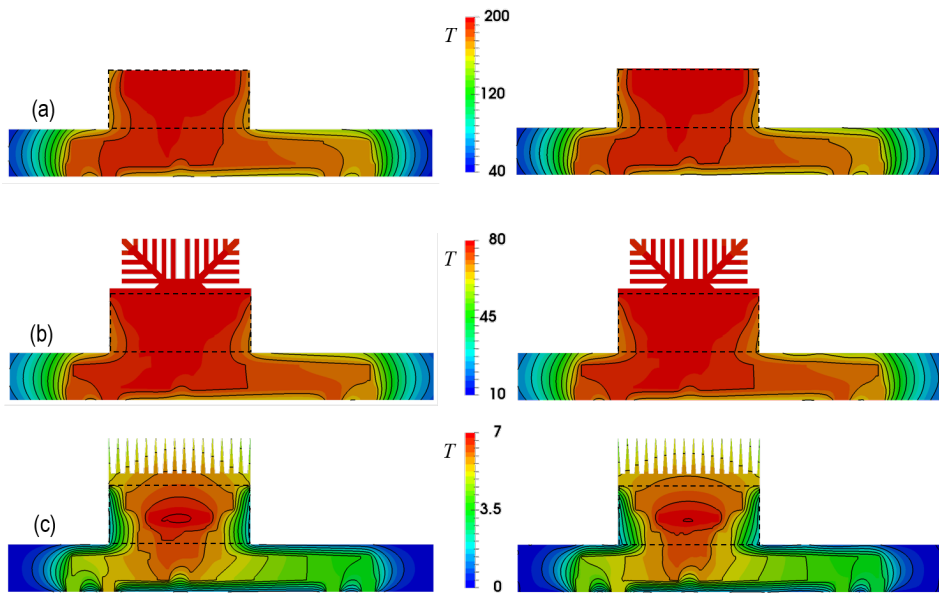


Figure 4 - Champs de température en régime permanent calculés avec la méthode des éléments finis et par modèles modaux sous-structurés pour trois cas tests

le modèle réduit reproduit fidèlement le modèle détaillé. En particulier, les isothermes restent continues aux interfaces malgré la réduction modale et la non-conformité des maillages, ce qui montre l'efficacité de la méthode des joints.

Nous rajoutons pour le cas (b) un radiateur, le modèle réduit du PCB et du composant restant les mêmes que pour le cas (a). L'ensemble est là encore refroidi par convection naturelle. La visualisation des champs de température montre la très bonne adaptabilité des modèles réduits. Outre le fait que la diminution de température causée par l'ajout du radiateur est parfaitement prédite, on remarque que même l'allure des isothermes au niveau des parois latérales du composant sont bien reproduites.

Finalement, nous changeons de radiateur et de conditions aux limites pour le cas (c) ($h = 200 \text{ W.m}^{-2}.\text{K}^{-1}$, convection forcée), mais nous gardons la même bibliothèque modale pour le composant et le circuit imprimé. Les performances du modèle réduit sont là aussi très satisfaisantes. Par ailleurs, on remarque que l'allure du champ de température dans le composant et dans la carte est significativement différente entre le cas (c) et les deux autres cas. Cependant, un même modèle réduit permet de reproduire ces deux allures avec une très bonne précision.

Afin de quantifier plus précisément les performances du modèle réduit, le tableau 2 présente la norme L^2 de l'écart relatif entre les modèles éléments finis et le modèle réduit modal sous-structuré en fonction du nombre de modes par sous-structure. L'écart relatif diminue linéairement avec le nombre de modes. Nous présentons aussi l'ordre de grandeur du gain de temps CPU (ce gain ne tient pas compte du temps de calcul off-line des modes). Cette table confirme les bonnes performances du modèle réduit, puisqu'avec uniquement 50 modes par sous-structure, l'écart moyen entre le modèle détaillé et le modèle réduit est inférieur au

	5	10	20	50	100
ε (%) cas (a)	2,97	2,03	0,55	0,18	0,07
ε (%) cas (b)	3,13	2,2	0,37	0,19	0,08
ε (%) cas (c)	10,82	8,59	2,85	0,99	0,74
Gain CPU cas (c)	2700	1350	670	180	55

Tableau 2 - Évolution en fonction du nombre de modes par sous-structure de l'écart relatif (en norme L^2 et en %) entre le modèle éléments finis et le modèle modal sous-structuré

pourcent dans les trois cas étudiés, et cela pour un gain de temps de 180. Cette précision de 1% est obtenue dès 20 modes pour les deux premiers cas. La relative contre-performance en pourcentage du modèle réduit pour le cas (c) doit être contre-balançée par la faible gamme de température (7 K) due au fort coefficient de convection : en variable dimensionnée, l'écart moyen est de l'ordre de 0.2 K avec 50 modes, c'est à dire un écart tout à fait admissible.

4. Conclusion et perspectives

Nous avons décrit dans cette communication une méthode originale permettant de coupler des modèles réduits créés séparément. Pour cela, nous avons adapté la méthode des joints à la méthode de réduction modale sur une base de Dirichlet-Steklov. Une illustration a été donnée pour prouver le concept de la méthode.

Cette étude préliminaire amène des questions, comme le choix des fonctions mortiers. D'autres tests non présentés ici montrent qu'une réduction supplémentaire peut être conduite au niveau de ces fonctions. L'influence de cette réduction pourra faire l'objet d'une étude plus complète. Le choix du découpage des frontières est une question ouverte : dans l'exemple donné ici, le composant est séparé du circuit imprimé par 9 frontières correspondant au découpage physique. À chacune de ces interfaces est associée une famille de fonctions mortiers. Quelles seraient les performances du modèle réduit si ces frontières étaient réunies en une seule ?

Cette méthode ouvre d'intéressantes potentialités, puisqu'elle allie la modularité des méthodes nodales et la résolution spatiale des méthodes modales. Elle permet surtout la création de bibliothèques de modèles réduits de composants réutilisables d'une application à l'autre.

Références

- [1] E. Monier-Vinard, B. Rogié, V. Bissuel, N. Laraqi, O. Daniel, and M-C. Kotelon, State of the art of numerical thermal characterization of electronic component, *17th International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE)* (Montpellier, France, 2016).
- [2] P. O. Laffay, O. Quemener, A. Neveu, Developing a method for coupling branch modal models, *IJTS*, 48, (2009), 1060–1067
- [3] C. Bernardi, Y. Maday, and A. T. Patera, A new conforming approach to domain decomposition : The mortar element method, *Non Linear Partial Differential Equations and Applications, College de France Seminar*, Pitman edition (1990), 13–51.
- [4] C. Bernardi, Y. Maday, and F. Rapetti, Basics and some applications of the mortar element method, *GAMM-Mitteilungen*, 28-2 (2005), 97–123.
- [5] G. Allaire, *Analyse numérique et optimisation*, Edition de l'École Polytechnique-Ellipse (2006)
- [6] F. Bourquin, F. d'Hennezel, Numerical study of an intrinsic component mode synthesis method, *Comput. Meth. Appl. Mech. Eng.*, 97 (1992), 49–76
- [7] C. Lacour and Y. Maday, Two different approaches for matching nonconforming grids : The mortar element method and the FETI method, *BIT Numerical Mathematics*, 37-3 (1997), 720–738.

Bibliographie

- [1] W. STEKLOFF [V.A. STEKLOV] : Sur les problèmes fondamentaux de la physique mathématique. *Annales sci. ENS, Sér. 3, 19*, pages 191–259 and 455–490, 1902.
- [2] W. STEKLOFF [V.A. STEKLOV] : Sur la condition de fermeture des systèmes de fonctions orthogonales. *Comp. rend. Acad. sci. Paris, 141*, pages 1116–1119, 1910.
- [3] W. STEKLOFF [V.A. STEKLOV] : Sur la théorie de fermeture des systèmes de fonctions orthogonales dépendant d'un nombre quelconque de variables. *Mém. Acad. Sci.St. Pétersbourg, Cl. Phys. Math., Sér. 8, 30, no. 4*, pages 1–87, 1911.
- [4] N. KUZNETSOV et AL : The Legacy of Vladimir Andreevich Steklov. *Notice of the AMS, Vol.61, nb.1*, 2014.
- [5] INTERSIL : Application Note 1759, 3A/4A Low Quiescent Current High Efficiency Synchronous Buck Regulator.
- [6] Jean-Luc BATTAGLIA, Andrzej KUSIAK et Jean-Rodolphe PUIGGALI : *Introduction aux transferts thermiques - 2e édition : Cours et exercices corrigés*. Dunod, Paris, 2e édition édition, mars 2014.
- [7] J. R. CULHAM et M. M. YOVANOVICH : Factors affecting the calculation of effective conductivity in printed circuit boards [thermal analysis]. In *ITherm'98. Sixth Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (Cat. No.98CH36208)*, pages 460–467, mai 1998.
- [8] Y. SHABANY : Component size and effective thermal conductivity of printed circuit boards. In *ITherm 2002. Eighth Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (Cat. No.02CH37258)*, pages 489–494, mai 2002.
- [9] Y. S. MUZYCHKA, J. R. CULHAM et M. M. YOVANOVICH : Thermal Spreading Resistance of Eccentric Heat Sources on Rectangular Flux Channels. *Journal of Electronic Packaging*, 125(2):178–185, juin 2003.

- [10] N. RINALDI : Generalized image method with application to the thermal modeling of power devices and circuits. *IEEE Transactions on Electron Devices*, 49(4):679–686, avril 2002.
- [11] Y. S. MUZYCHKA, K. R. BAGNALL et E. N. WANG : Thermal Spreading Resistance and Heat Source Temperature in Compound Orthotropic Systems With Interfacial Resistance. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 3(11):1826–1841, novembre 2013.
- [12] Sébastien VINTROU, Najib LARAQI et Abderrahmane BAÏRI : Thermal impedance of multi-finger microelectronic structures : Exact analytical model. *Journal of Physics D : Applied Physics*, 42(24):245501, 2009.
- [13] D. De KLERK, D. J. RIXEN et S. N. VOORMEEREN : General Framework for Dynamic Substructuring : History, Review and Classification of Techniques. *AIAA Journal*, 46(5): 1169–1181, mai 2008.
- [14] H. A. SCHWARZ : *II. Ueber einen Grenzübergang durch alternirendes Verfahren*. 1870. Published : Wolf J. XV. 272-286. 1870 (1870). MSC2010 : 31A05 = Harmonic, subharmonic, superharmonic functions (two-dimensional).
- [15] Leo SARIO : Alternating method on arbitrary Riemann surfaces. *Pacific Journal of Mathematics*, 3(3):631–645, septembre 1953.
- [16] Harvey COHN : *Conformal Mapping on Riemann Surfaces*. Dover Publ, New York, unabridged and corr. republ édition, 1980. OCLC : 251692335.
- [17] Henri Paul de SAINT-GERVAIS : *Uniformisation des surfaces de Riemann : retour sur un théorème centenaire*. ENS éditions, Lyon, 2010.
- [18] Barry SMITH, Petter E. BJØRSTAD et William D. GROPP : *Domain Decomposition : Parallel Multilevel Methods for Elliptic Partial Differential Equations = XE-AU*. Cambridge Univ. Press, Cambridge, 1. pbk. ed édition, 2004. OCLC : 254691168.
- [19] Andrea TOSELLI et Olof B. WIDLUND : *Domain Decomposition Methods — Algorithms and Theory*, volume 34 de *Springer Series in Computational Mathematics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.
- [20] J. S. PRZEMIENIECKI : Matrix Structural Analysis of Substructures. *AIAA Journal*, 1(1):138–147, 1963.
- [21] Hirofumi UZAWA et Kenneth J ARROW : *Preference, Production and Capital : Selected Papers of Hirofumi Uzawa*. Cambridge University Press, Cambridge, 1989. OCLC : 776972500.

- [22] Christine BERNARDI, Yvon MADAY et Anthony T. PATERA : Domain Decomposition by the Mortar Element Method. *In Asymptotic and Numerical Methods for Partial Differential Equations with Critical Parameters*, pages 269–286. Springer, 1993.
- [23] Christine BERNARDI, Yvon MADAY et Anthony T. PATERA : A New Conforming Approach to Domain Decomposition : The Mortar Element Method. *In Collège de France Seminar XI*, College de France Seminar, pages 13–51, 1994.
- [24] Gassav S. ABDOULAEV, Yves ACHDOU, Yuri A. KUZNETSOV et Christophe PRUD'HOMME : On a Parallel Implementation of the Mortar Element Method. *ESAIM : Mathematical Modelling and Numerical Analysis*, 33(2):245–259, mars 1999.
- [25] F. B. BELGACEM, P. HILD et P. LABORDE : The Mortar Finite Element Method for Contact Problems. *Mathematical and Computer Modelling*, 28(4):263–271, août 1998.
- [26] Faker Ben BELGACEM : The Mortar Finite Element Method with Lagrange Multipliers. *Numerische Mathematik*, 84(2):173–197, 1999.
- [27] F. Ben BELGACEM, L. K. CHILTON et P. SESHAIYER : The Hp-Mortar Finite-Element Method for the Mixed Elasticity and Stokes Problems. *Computers & Mathematics with Applications*, 46(1):35–55, juillet 2003.
- [28] F. Ben BELGACEM et Yvon MADAY : The Mortar Element Method For Three Dimensional Finite Elements. *RAIRO - Modélisation mathématique et analyse numérique*, 31(2):289–302, 1997.
- [29] Catherine LACOUR et Yvon MADAY : Two Different Approaches for Matching Nonconforming Grids : The Mortar Element Method and the FETI Method. *BIT Numerical Mathematics*, 37(3):720–738, 1997.
- [30] Dietrich BRAESS, Wolfgang DAHMEN et Christian WIENERS : A Multigrid Algorithm for the Mortar Finite Element Method. *SIAM Journal on Numerical Analysis*, 37(1):48–69, 1999.
- [31] D. BRAESS, M. DRYJA et W. HACKBUSH : A Multigrid Method for Nonconforming FE-Discretisations with Application to Non-Matching Grids. *Computing*, 63(1):1–25, juillet 1999.
- [32] Dietrich BRAESS et Rüdiger VERFÜRTH : Multigrid Methods for Nonconforming Finite Element Methods. *SIAM Journal on Numerical Analysis*, 27(4):979–986, 1990.
- [33] J. DANEK et H. KUTAKOVA : The Mortar Finite Element Method in 2D : Implementation in Matlab. 2008.

- [34] Maksymilian DRYJA : An Iterative Substructuring Method for Elliptic Mortar Finite Element Problems with Discontinuous Coefficients. *Contemporary Mathematics*, 218:94–103, 1998.
- [35] M. DITTMANN, M. FRANKE, I. TEMIZER et C. HESCH : Isogeometric Analysis and Thermomechanical Mortar Contact Problems. *Computer Methods in Applied Mechanics and Engineering*, 274:192–212, juin 2014.
- [36] Patrick HILD : Numerical Implementation of Two Nonconforming Finite Element Methods for Unilateral Contact. *Computer Methods in Applied Mechanics and Engineering*, 184(1):99–123, mars 2000.
- [37] R. HOPPE, Yu ILIASH, Yu KUZNETSOV, Yu VASSILEVSKI et B. WOHLMUTH : Analysis and Parallel Implementation of Adaptive Mortar Element Methods. *EAST WEST JOURNAL OF NUMERICAL MATHEMATICS*, 6:223–248, 1998.
- [38] Francesca RAPETTI : *Approximation Des Équations de La Magnétodynamique En Domaine Tournant Par La Méthode Des Éléments Avec Joints*. Thèse de doctorat, Paris 6, 2000.
- [39] F. RAPETTI, E. BOUILLAULT, L. SANTANDREA, A. BUFFA, Y. MADAY et A. RAZEK : Calculation of Eddy Currents with Edge Elements on Non-Matching Grids in Moving Structures. *IEEE Transactions on Magnetics*, 36(4):1351–1355, juillet 2000.
- [40] Barbara I. WOHLMUTH : A Mortar Finite Element Method Using Dual Spaces for the Lagrange Multiplier. *SIAM Journal on Numerical Analysis*, 38(3):989–1012, janvier 2000.
- [41] Barbara I. WOHLMUTH : Iterative Solvers Based on Domain Decomposition. *In Discretization Methods and Iterative Solvers Based on Domain Decomposition*, Lecture Notes in Computational Science and Engineering Free Preview. Springer-Verlag Berlin Heidelberg, springer édition, 2001.
- [42] Barbara I. WOHLMUTH : Discretization Techniques Based on Domain Decomposition. *In Discretization Methods and Iterative Solvers Based on Domain Decomposition*, Lecture Notes in Computational Science and Engineering Free Preview. Springer-Verlag Berlin Heidelberg, springer édition, 2001.
- [43] B. WOHLMUTH et R. KRAUSE : Monotone Multigrid Methods on Nonmatching Grids for Nonlinear Multibody Contact Problems. *SIAM Journal on Scientific Computing*, 25(1):324–347, janvier 2003.
- [44] Faker BEN BELGACEM : *Discrétisations 3D Non Conformes Par La Méthode de Décomposition de Domaine Des Éléments Avec Joints : Analyse Mathématique et Mise En Oeuvre Pour Le Problème de Poisson*. Paris 6, janvier 1993.

- [45] Frédéric BOUILLAULT, Annalisa BUFFA, Yvon MADAY et Francesca RAPETTI : The Mortar Edge Element Method in Three Dimensions : Application to Magnetostatics. *SIAM Journal on Scientific Computing*, 24(4):1303–1327, 2003.
- [46] Christine BERNARDI, Yvon MADAY et Francesca RAPETTI : Basics and Some Applications of the Mortar Element Method. *GAMM-Mitteilungen*, 28(2):97–123, 2005.
- [47] J. L. LIONS : Problèmes aux Limites Non-homogènes à Données Irrégulières. In *Numerical Analysis of Partial Differential Equations*, C.I.M.E. Summer Schools, pages 283–292. Springer, Berlin, Heidelberg, 2010.
- [48] Joachim NITSCHKE : Uber Ein Variationsprinzip Zur Losung von Dirichlet-Problemen Bei Verwendung von Teilraumen, Die Keinen Randbedingungen Unterworfen Sind. In *Abhandlungen Aus Dem Mathematischen Seminar Der Universitaet Hamburg*, volume 36, pages 9–15. Springer, 1971.
- [49] Mary Fanett WHEELER : An Elliptic Collocation-Finite Element Method with Interior Penalties. *SIAM Journal on Numerical Analysis*, 15(1):152–161, 1978.
- [50] D. ARNOLD : An Interior Penalty Finite Element Method with Discontinuous Elements. *SIAM Journal on Numerical Analysis*, 19(4):742–760, août 1982.
- [51] Roland BECKER, Peter HANSBO et Rolf STENBERG : A Finite Element Method for Domain Decomposition with Non-Matching Grids. *ESAIM : Mathematical Modelling and Numerical Analysis*, 37(2):209–225, mars 2003.
- [52] W.H REED et T.R HILL : Triangular Mesh Methods For The Neutron Transport Equation. *Proceedings of the American Nuclear Society*, 1973.
- [53] D. ARNOLD, F. BREZZI, B. COCKBURN et L. MARINI : Unified Analysis of Discontinuous Galerkin Methods for Elliptic Problems. *SIAM Journal on Numerical Analysis*, 39(5):1749–1779, janvier 2002.
- [54] Douglas N ARNOLD, F. BREZZI, B. COCKBURN et Donatella MARINI : Discontinuous Galerkin Methods for Elliptic Problems. *Lecture Notes in Computational Science and Engineering*, 11, février 2000.
- [55] Jim DOUGLAS et Todd DUPONT : Interior Penalty Procedures for Elliptic and Parabolic Galerkin Methods. In *Computing Methods in Applied Sciences*, Lecture Notes in Physics, pages 207–216. Springer, Berlin, Heidelberg, 1976.
- [56] Béatrice RIVIÈRE, Mary F. WHEELER et Vivette GIRAULT : Improved Energy Estimates for Interior Penalty, Constrained and Discontinuous Galerkin Methods for Elliptic Problems. Part I. *Computational Geosciences*, 3(3-4):337–360, décembre 1999.

- [57] I. BABUŠKA et M. ZLÁMAL : Nonconforming Elements in the Finite Element Method with Penalty. *SIAM Journal on Numerical Analysis*, 10(5):863–875, octobre 1973.
- [58] F. BREZZI, G. MANZINI, D. MARINI, P. PIETRA et A. RUSSO : Discontinuous Galerkin Approximations for Elliptic Problems. *Numerical Methods for Partial Differential Equations*, 16(4):365–378, juillet 2000.
- [59] F. BASSI et S. REBAY : A High-Order Accurate Discontinuous Finite Element Method for the Numerical Solution of the Compressible Navier-Stokes Equations. *Journal of Computational Physics*, 131(2):267–279, mars 1997.
- [60] B. COCKBURN et C. SHU : The Local Discontinuous Galerkin Method for Time-Dependent Convection-Diffusion Systems. *SIAM Journal on Numerical Analysis*, 35(6):2440–2463, décembre 1998.
- [61] Carlos Erik BAUMANN et J. Tinsley ODEN : A Discontinuous Hp Finite Element Method for Convection-Diffusion Problems. *Computer Methods in Applied Mechanics and Engineering*, 175(3):311–341, juillet 1999.
- [62] A. SAMAKE, S. BERTOLUZZA, M. PENNACCHIO, Christophe PRUD’HOMME et C. ZAZA : A Parallel Implementation of the Mortar Element Method in 2D and 3D. *ESAIM : Proceedings*, 43:213–224, décembre 2013.
- [63] J. W. SOFIA : Analysis of thermal transient data with synthesized dynamic models for semiconductor devices. *IEEE Transactions on Components, Packaging, and Manufacturing Technology : Part A*, 18(1):39–47, mars 1995.
- [64] JEDEC SOLID STATE TECHNOLOGY ASSOCIATION : DELPHI Compact Thermal Model Guideline, 2008.
- [65] P. ARUNASALAM, K. N. SEETHARAMU et I. A. AZID : Determination of thermal compact model via evolutionary genetic optimization method. *IEEE Transactions on Components and Packaging Technologies*, 28(2):345–352, juin 2005.
- [66] Cheikh Tidiane DIA : *Génération de Modeles Compacts Thermiques Dynamiques de Composants Electroniques via Les Algorithmes Genetiques*. Thèse de doctorat, Paris 10, décembre 2015.
- [67] Jean-Baptiste-Joseph FOURIER : *Théorie analytique de la chaleur*. 1822.
- [68] A.C. ANTOULAS : An Overview of Approximation Methods for Large-Scale Dynamical Systems. *Annual Reviews in Control*, 29(2):181–190, janvier 2005.
- [69] Adam FIC, Ryszard A. BIAŁECKI et Alain J. KASSAB : Solving Transient Nonlinear Heat Conduction Problems by Proper Orthogonal Decomposition and the Finite-

- Element Method. *Numerical Heat Transfer, Part B : Fundamentals*, 48(2):103–124, août 2005.
- [70] H. M. PARK et W. S. JUNG : Recursive Solution of an Inverse Heat Transfer Problem in Rapid Thermal Processing Systems. *International Journal of Heat and Mass Transfer*, 44(11):2053–2065, 2001.
- [71] Jeanne A. ATWELL et Belinda B. KING : Proper Orthogonal Decomposition for Reduced Basis Feedback Controllers for Parabolic Equations. *Mathematical and computer modelling*, 33(1-3):1–19, 2001.
- [72] JL LUMLEY : The Structure of Inhomogeneous Turbulent Flows. In AM YAGLOM et VI TATARSKI, éditeurs : *Atmospheric Turbulence and Radio Propagation*, pages 166–178. Nauka, 1967.
- [73] Lawrence SIROVICH : Turbulence and the Dynamics of Coherent Structures Part I : Coherent Structures. *Quarterly of Applied Mathematics*, 45(3):561–571, 1987.
- [74] Francisco CHINESTA, Roland KEUNINGS et Adrien LEYGUE : *The Proper Generalized Decomposition for Advanced Numerical Simulations : A Primer*. SpringerBriefs in Applied Sciences and Technology. Springer International Publishing, 2014.
- [75] Julien BERGER, Sihem GUERNOUTI, Monika WOLOSZYN et Francisco CHINESTA : Proper Generalised Decomposition for Heat and Moisture Multizone Modelling. *Energy and Buildings*, 105:334–351, octobre 2015.
- [76] Ulrike BAUR : Multipoint Padé approximation for parametric model order reduction. page 6.
- [77] Brice ROGIE, Lorenzo CODECASA, Eric MONIER-VINARD, Valentin BISSUEL, Najib LARAQI, Olivier DANIEL, Dario D’AMORE, Alessandro MAGNANI, Vincenzo D’ALESSANDRO et Niccolò RINALDI : Delphi-like dynamical compact thermal models using model order reduction. In *2017 23rd International Workshop on Thermal Investigations of ICs and Systems (THERMINIC)*, pages 1–6, Amsterdam, septembre 2017. IEEE.
- [78] Serkan GUGERCIN : An Iterative SVD-Krylov Based Method for Model Reduction of Large-Scale Dynamical Systems. *Linear Algebra and its Applications*, 428(8-9):1964–1986, avril 2008.
- [79] David BINION et Xiaolin CHEN : A Krylov Enhanced Proper Orthogonal Decomposition Method for Efficient Nonlinear Model Reduction. *Finite Elements in Analysis and Design*, 47(7):728–738, juillet 2011.

- [80] Olivier QUÉMÉNER, Alain NEVEU et Etienne VIDECOQ : A Specific Reduction Method for the Branch Modal Formulation : Application to a Highly Non-Linear Configuration. *International Journal of Thermal Sciences*, 46(9):890–907, septembre 2007.
- [81] Etienne VIDECOQ, Myriam LAZARD, Olivier QUEMENER et Alain NEVEU : Online Temperature Prediction Using a Branch Eigenmode Reduced Model Applied to Cutting Process. *Numerical Heat Transfer, Part A : Applications*, 55(7):683–705, mars 2009.
- [82] 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, mars 2008.
- [83] Olivier QUÉMÉNER, Frédéric JOLY et Alain NEVEU : On-Line Heat Flux Identification from a Rotating Disk at Variable Speed. *International Journal of Heat and Mass Transfer*, 53(7-8):1529–1541, mars 2010.
- [84] D. PETIT et R. PASQUETTI : Réduction de modèle par identification de modes dominants : Application à un modèle bidimensionnel de diffusion thermique. *Revue de Physique Appliquée*, 25(8):831–842, 1990.
- [85] 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(21):3177–3189, 1998.
- [86] M. GIRAULT, E. VIDECOQ et D. PETIT : Estimation of Time-Varying Heat Sources through Inversion of a Low Order Model Built with the Modal Identification Method from in-Situ Temperature Measurements. *International Journal of Heat and Mass Transfer*, 53(1-3):206–219, janvier 2010.
- [87] Y. ROUIZI, M. GIRAULT, Y. FAVENNEC et D. PETIT : Model Reduction by the Modal Identification Method in Forced Convection : Application to a Heated Flow over a Backward-Facing Step. *International Journal of Thermal Sciences*, 49(8):1354–1368, août 2010.
- [88] Etienne VIDECOQ, Alain NEVEU, Olivier QUEMENER, Manuel GIRAULT et Daniel PETIT : Comparison of Two Nonlinear Model Reduction Techniques : The Modal Identification Method and the Branch Eigenmodes Reduction Method. *Numerical Heat Transfer, Part B : Fundamentals*, 49(6):537–558, décembre 2006.
- [89] Pierre-Olivier LAFFAY : *Sous-Structuration de Systemes Thermiques Par Modes de Branche*. Thèse de doctorat, Université D’Evry Val-d’Essonne, 2008.
- [90] G. ALLAIRE : *Analyse Numérique et Optimisation*. Ellipse, 2006.

- [91] Francis HIRSCH et Gilles LACOMBE : *Éléments d'analyse Fonctionnelle : Cours et Exercices*. 1997.
- [92] H. BREZIS : *Analyse Fonctionnelle*. 1991.
- [93] Frédéric BOURQUIN et Frédéric D'HENNEZEL : Numerical Study of an Intrinsic Component Mode Synthesis Method. *Computer Methods in Applied Mechanics and Engineering*, 97(1):49–76, mai 1992.
- [94] Alain NEVEU : étude d'un pavé droit, 2018.
- [95] Abdelhakim OULEFKI : *Reduction de Modeles Thermiques Par Amalgame Modal*. Thesis, Marne-la-vallée, ENPC, janvier 1993.
- [96] Olivier QUÉMÉNER, Frédéric JOLY et Alain NEVEU : The Generalized Amalgam Method for Modal Reduction. *International Journal of Heat and Mass Transfer*, 55(4):1197–1207, janvier 2012.
- [97] A. RUFER : Cours d'électronique de puissance. *École Polytechnique de Lausanne*, 2014.
- [98] Dagal Dari YAYA : Conception, réalisation et caractérisation d'inductances planaires à couches magnétiques. page 178.
- [99] E. MONIER-VINARD, B. ROGIÉ, V. BISSUEL, N. LARAQI, O. DANIEL et M. C. KOTELON : State of the art of numerical thermal characterization of electronic component. In *2016 17th International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE)*, pages 1–8, 2016.
- [100] Benjamin GAUME : *Réduction d'un problème d'auto-rayonnement par modes de Branche : Application aux échanges thermiques dans un domaine multi-enceintes*. Thèse de doctorat, Paris-Saclay, Evry, octobre 2016.
- [101] C. LANCZOS : An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *Journal of Research of the National Bureau of Standards*, 45(4):255, octobre 1950.
- [102] R. B. LEHOUCQ, D. C. SORENSEN et C. YANG : *ARPACK Users' Guide : Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods*. Software, environments, tools. SIAM, Philadelphia, 1998.
- [103] T. GOUDON : *Intégration. Intégrale de Lebesgue et Introduction à l'analyse Fonctionnelle*. ELLIPSES MARKETING, 2011.
- [104] F. FER : *Thermodynamique Macroscopique : Systèmes Ouverts*. 1971.
- [105] L. ONSAGER : Reciprocal Relations in Irreversible Processes. I. *Phys. Rev.*, 37,405, 1931.

- [106] Laurent SCHWARTZ : *Méthodes Mathématiques Pour Les Sciences Physiques. Deuxième Cycle*. Hermann, 1997.

Titre: Réduction d'un problème thermique par sous structuration modale. Application à la modélisation d'ensembles électroniques complexes.

Mots clés: Réduction de modèle, problème thermique, méthode modale, sous-structuration, modélisation, systèmes électroniques

Résumé: Un composant électronique dissipe de la chaleur. L'optimisation « thermique » de la conception des systèmes électroniques ainsi que leur durée d'activation ou de désactivation devient alors nécessaire pour limiter les températures. Il faut donc pouvoir prédire finement et rapidement l'évolution thermique d'un composant pris dans son environnement pour différents scénarios d'utilisation. Les simulations classiques types éléments finis étant trop coûteuses en temps de calcul pour des ensembles complexes, il est nécessaire de réduire la taille du modèle. Les méthodes modales consistent à rechercher la solution sous la forme d'une somme pondérée de champs élémentaires, ces derniers étant appelés modes. Ces modes sont calculés en résolvant numériquement un problème aux valeurs propres. Si ces méthodes ont fait leur preuve pour des composants élémentaires, elles se révèlent inopérantes

pour des systèmes comportant plusieurs dizaines de composants placés sur un circuit imprimé. La méthode de sous-structuration modale est une extension des méthodes modales classiques permettant de dépasser ces limitations. Il s'agit de décomposer le système en entités élémentaires (les sous-structures), de calculer les modes de chacune de ces entités, puis de les rassembler pour résoudre le problème d'origine. Des travaux préliminaires ont montré que cette approche est pertinente au premier niveau (composant électronique). On envisage lors de cette thèse d'étendre ces principes aux niveaux supérieurs. La finalité industrielle du présent travail est de simuler à l'aide d'un modèle réduit le comportement thermique d'une carte électronique (circuit imprimé et plusieurs composants implantés). Ce travail peut ouvrir la voie pour l'industriel à la modélisation de systèmes multi-modules.

Title: Reducing a thermal problem by the modal sub-structuring method. Application to the modeling of complex electronic packages.

Keywords: model reducing, thermal problem, méthode modale, sub-structuring, modeling, ectronic packages

Abstract: An electronic component heats. Thermal optimization of the design of electronic packages and of their activation and de-activation duration is essential to cap the temperatures reached. Thus we have to be able to finely and quickly predict the thermal evolution of a component set in its surroundings for multiples scenarios of use. Classical simulations, like finite elements ones, are too costly in terms of computing time for complex packages, we have to reduce the size of the model. Modal methods consist in seek the solution as a weighted sum of elementary fields, labelled modes. These modes are computed by solving eigenvalue problems. But these methods don't work for packaging containing dozens of components on a printed circuit board. The modal sub-structuring method is an extent of classical modal methods used to overcome their limitations. The principle is

to disassemble the system into elementary entities (the sub-structures), to compute the modes of each of these entities and to reassemble them to solve the original problem. Preliminary work has shown the relevance of this method on the first level (electric component). The thesis is aimed to extend these principles to superior levels. The industrial purpose of this work is to simulate, thanks to a reduced model, the thermal behavior of an electronic card. Beside the industrial aspect, this work raises fundamental issues like the definition of the junction conditions between the sub-structures in the modal space and the adaptation of the reducing technique of the sub-structures in view of their pairing. The thesis will give answers to these questions using an electronic card as a study support. This work could pave the way of the industrial for the modeling of multi-modules systems.

