l'université nantes angers le mans
PÔLE DE RECHERCHE ET D'ENSEIGNEMENT SUPÉRIEUR

UNIVERSITÉ DE NANTES

Thèse de Doctorat

Céline SARAZIN-DESBOIS

Mémoire présenté en vue de l'obtention du
grade de Docteur de l'Université de Nantes
sous le label de l'Université de Nantes Angers Le Mans

École doctorale : Sciences et technologies de l'information, et mathématiques

Discipline : Mathématiques et leurs interactions, section CNU 26
Unité de recherche : Laboratoire de mathématiques Jean Leray (LMJL)

Soutenue le 12 mars 2013
Thèse n°: ED 503-XXX

Méthodes numériques pour des systèmes hyperboliques avec terme source provenant de physiques complexes autour du rayonnement

JURY

Président :	**M. Frédéric LAGOUTIÈRE**, Professeur, Université Paris-Sud 11
Rapporteurs :	**M. Nicolas SEGUIN**, Maître de conférences, Université Pierre et Marie Curie, Paris 6
	M^me^ Marie-Hélène VIGNAL, Maître de conférences, Institut de mathématiques de Toulouse
Examinateurs :	**M. Frédéric LAGOUTIÈRE**, Professeur, Université Paris-Sud 11
	M. François JAUBERTEAU, Professeur, Université de Nantes
Directeur de thèse :	**M. Christophe BERTHON**, Professeur, Université de Nantes
Co-directeur de thèse :	**M. Rodolphe TURPAULT**, Maître de conférences, Université de Nantes

Remerciements

Pour compléter le contenu de ce manuscrit, qui rayonne déjà rien que par son titre, je me dois de remercier comme il se doit un hyper grand nombre de personnes sans qui ces trois années n'auraient pas été si agréables.

Pour commencer, je souhaite remercier Christophe Berthon et Rodolphe Turpault de m'avoir accordé leur confiance en me proposant cette thèse. Je les remercie d'avoir consacré de très (très) nombreuses heures à m'initier au monde de l'hyperbolique numérique, à me guider dans mes recherches, à répondre à mes questions (pas toujours très claires), à débugger des codes (pas toujours très bien codés), à relire d'innombrables versions de mes papiers (pas toujours bien rédigés) et plus généralement à discuter de maths et du monde des maths.

Je remercie mes rapporteurs, Marie-Hélène Vignal et Nicolas Seguin d'avoir accepté de juger ce travail. Les nombreuses remarques et pistes pertinentes qu'ils m'ont suggérées sauront m'être très utiles pour la suite. Je remercie également les autres membres de mon jury, Frédéric Lagoutière et François Jauberteau de s'être intéressés à mon travail et d'avoir assisté à ma soutenance.

Je remercie également Martin Frank, toute l'équipe du laboratoire de mathématiques de Kaiserslautern (et particulièrement Edgar et Phil) et Bruno Dubroca qui m'ont initiée au monde de la recherche lors de mon stage de fin de master. Je les remercie pour leur accueil, leur disponibilité ou encore les nombreuses discussions mathématiques. Je n'oublierai pas cette très bonne ambiance au sein du laboratoire et les quelques spécificités de l'équipe (passant par le football et l'incontournable 'Kein Bier vor vier').

Plus généralement, je souhaite remercier tous les membres du laboratoire Jean Leray et du département de mathématiques. Tout d'abord, je remercie l'ensemble des membres du secrétariat et de la bibliothèque qui en plus de leur disponibilité et de leur efficacité sont une source de bonne humeur et de soutien dans le laboratoire. De même, je remercie Saïd et Éric qui ont du plus d'une fois réparer mes bugs informatiques ! Finalement, je tiens à remercier les enseignants-chercheurs qui m'ont transmis leurs connaissances (de la L1 au M2), m'ont guidée et conseillée pour mes propres enseignements et ont su répondre à mes nombreuses questions.

Je remercie particulièrement Guy Moebs, qui m'a fait économiser de nombreuses heures de calculs en optimisant et parallélisant mes codes pas toujours simples à comprendre. Ses cours et nombreux conseils m'ont été très utiles et je sais maintenant faire des boucles à l'endroit (en théorie...).

Toujours au sein du laboratoire, je remercie à présent ceux qui partagent et égayent le quotidien d'une thèse : les thésards. Je commence par Alexandre (U) qui a partagé plus qu'un bureau pendant ces quatre années. Il a tout d'abord su écouter (et supporter) mes nombreuses histoires (pas toujours intéressantes), anecdotes (encore moins intéressantes), plaintes, problèmes de maths,

... Il a répondu à chaque “don’t call my name, don’t call my name“ et à chaque “n’importe quoi“. Il a réussi à me motiver chaque mardi pour aller à la piscine. Il a été un super témoin et organisateur d’EVJF. J’ai adoré faire des combats, jouer à qui téléphone le plus, découvrir de nouveaux jeux de société, résoudre des dingbats, élever un ficus, tout ça grâce au meilleur co-bureau du monde ! Je remercie ensuite Vivien, (les numériciens d’abord !), qui sans être notre co-bureau, a passé beaucoup de temps dans notre bureau ! Il a su résoudre de nombreux problèmes mathématiques, comme par exemple la traversée périlleuse d’un domaine triangularisé en respectant la géométrie du Grand Raid des Pyrénées. J’ai particulièrement apprécié sa grande générosité et disponibilité, son imagination et son goût prononcé pour les jeux de société, ses talents de dessin et d’écriture (notamment en anglais) et sa naïveté (n’est-ce pas Angélique ?). Je n’oublierai pas non plus ces quelques “blagues” dont il est fier, mais qui n’égaleront pas les miennes, ni ce café volé à fifa ! Je continue avec Julien, qui a partagé ma folie quand il s’agit d’inventer des histoires insensées, quand il s’agit de faire la course pour arriver dans l’eau (même froide avec de très grosses vagues) ou encore d’autres défis. Je le remercie pour ces nombreux fondants au chocolat pour lesquels il faut diviser 50 par 3 et qui se marient à merveille avec un grand verre d’eau. Passons à Alexandre (Q) sans qui les voyages en voiture ne seraient pas si musicaux et sans qui les soirées knackis-moutarde ou ail-citron ne seraient pas aussi mémorables. Je n’oublierai pas non plus ces soirées festives à la gargouille, au floride ou encore au havana durant lesquelles Benoît XVI était à l’honneur ! Quant à Christophe (Q3), je le remercie pour sa spontanéité, sa générosité, ses réflexions et questions philosophiques, son sens de l’humour et même ses nombreuses erreurs à la taroinche. Je remercie Jessy, mon éternel binôme, qui m’a supportée pendant 4 ans et avec qui j’ai passé de longues heures à réviser, à coder et à rédiger des rapports. Malgré la distance, il a toujours su me soutenir et me rassurer durant ces trois années (sans oublier les nombreux ragots...) et être présent dans les moments importants. Je remercie Thomas pour son éternelle galanterie, sa disponibilité pour répondre à toute sorte de questions et son penchant cannibalo-agoraphile, Tristan, expert latex, pour ses encouragements quotidiens avant le 12 mars et sa venue ce jour-là (mais pas pour toutes ses parties ratées de taroinche où il a pris avec moi !), Andrew pour le soleil australien qu’il a apporté au laboratoire. Je remercie Vincent, son accent marseillais et ses drops au deuxième tour, Salim pour ces discussions toujours passionnées, Carlos pour son extrême gentillesse et sa soif de françisationniste, Baptiste et Simon pour ces soirées musicales. Je remercie également Rafik pour ces nombreux débats en tout genre (notamment sur le placenta) et je lui souhaite bonne chance pour son futur rôle de président de la république ! Je remercie également Carl pour les nombreuses soirées festives et culturelles qu’il a organisées et Anne pour ses attentions particulières lors des séminaires des apprentis (je ne sais toujours pas ce que c’est !) et son gâteau sans cuisson mémorable. Je remercie Hermann pour tous ces GdTC mission spéciale et Ronan de nous avoir supporté (ou pas) dans son bureau. Je remercie Nicolas R pour son extravagance et ses réponses à tout, Moudhaffar pour sa gentillesse et le parfum qu’il répand dans tout le labo ou encore Nicolas H et sa surf attitude. Je remercie Gilberto et Ilaria pour leur apport de culture italienne au sein du labo, Antoine d’être la preuve qu’il y a une chance de s’en sortir quand on m’a eu comme prof, Virgile pour sa zenitude et ses démonstrations culinaires.

Je n’oublierai pas non plus toutes ces parties de taroinche après un bon repas au RU, les soirées diverses et variées (jeux, guitar hero, resto,...), les galettes des rois qui durent de janvier à avril les meilleures années, les nombreux goûters tartines de nutella, les intrusions de gens qui veulent un bonbon rose, les week-end du 31 mai en vendée...

Je remercie bien sûr tous mes amis qui n’ont pas partagé mon quotidien mathématiques, mais qui ont quand même été là, le week-end, quelques soirs, à la danse, pour faire un jeu, pour regarder un film, pour manger des pizzas...

Bien sûr, je remercie ma famille, qui m’a soutenue et encouragée (même si tout le monde ne

savait pas exactement ce que je faisais) pendant ces trois années et sans qui je n'en serais certainement pas arrivée là aujourd'hui.

Je terminerai ces quelques mots par ceux qui ont été et seront toujours les plus importants : doudou et Lulu. Alex, mon amour, a été là tous les jours pour me soutenir, me remonter le moral dans les moments difficiles et me donner envie d'être la meilleure, me faire à manger tous ces soirs où j'avais bien trop de choses à faire, me trouver des films (surtout des comédies romantiques) à regarder le soir, m'écouter répéter mes exposés (pas très longtemps mais quand même) et pour encore des millions d'autres choses qui prendraient trop de lignes à écrire ! Lulu, ma petite belette de crapule qui est venu s'ajouter à notre bonheur, m'a également apporté toute sa tendresse, tous ses sourires et toute sa malice. Merci mes amours !

Table des matières

Introduction

La simulation numérique de phénomènes physiques relatifs au transfert radiatif reste un défi majeur dans de nombreuses applications. Dans cette étude, on s'emploiera à développer des méthodes précises et efficaces pour prédire l'impact du rayonnement dans des configurations physiques extrêmes. En effet, le rayonnement peut devenir prépondérant dans certaines applications. Il est donc clair que la connaissance de ce phénomène physique est cruciale et que la mise en œuvre de schémas adaptés et efficaces est un point fondamental dans la réalisation de prédictions numériques.

Afin d'illustrer d'une part, le rôle essentiel joué par le transfert radiatif dans certains régimes d'écoulements et d'autre part, les difficultés numériques majeures sous-jacentes à ce phénomène physique, considérons par exemple le cas de la rentrée atmosphérique superorbitale d'un véhicule spatial. En effet, pour ce type d'écoulements en régime hypersonique, une onde de choc détachée se propage à l'avant du véhicule. Le phénomène compressif engendre une augmentation de la température qui peut atteindre des valeurs extrêmement importantes derrière cette onde de choc. Les températures deviennent suffisamment élevées pour provoquer une dissociation des particules et ainsi créer un plasma. Le retour à l'équilibre s'accompagne de divers échanges d'énergie au cours desquels des photons sont émis ou absorbés par le milieu ambiant et dépend de paramètres de caractérisation optique de la matière : les opacités. Le rayonnement résulte de toutes les émissions et absorptions. L'énergie associée au rayonnement peut, dans certaines configurations extrêmes, influencer notablement l'écoulement environnant. Plus précisément, dans l'exemple considéré, le rayonnement modifie fortement la position et l'épaisseur de la couche de choc. À travers cet exemple, on constate qu'un phénomène physique relatif au rayonnement, se propageant à la vitesse de la lumière, est en mesure de modifier l'écoulement d'un fluide autour d'une sonde superorbitale. En d'autres termes, deux phénomènes physiques dont le rapport des vitesses caractéristiques est sans commune mesure, peuvent être fortement couplés. D'un point de vue numérique, ce couplage de physiques impliquant des vitesses d'ordres de grandeur très différents engendre généralement de grandes pertes de précision dans l'évaluation des phénomènes de plus grande vitesse.

Il existe différents niveaux de modélisation du transfert radiatif. Dans l'étude qui nous motive ici, seuls seront considérés les modèles cinétiques et méso/macroscopiques. Dans ce cadre, l'équation de référence est appelée équation du transfert radiatif. Cette équation cinétique gouverne l'évolution de l'intensité radiative. Cette dernière dépend du temps, de l'espace, de la direction de propagation et de la fréquence. Contrairement à d'autres équations cinétiques telles que l'équation de Boltzmann ou celle de Fokker-Planck, les variables supplémentaires de direction de fréquence ne sont pas de même nature. D'un point de vue physique, la dépendance en fréquence est particulièrement non-triviale. La variation du terme de collision en fonction de la fréquence est singulièrement complexe. En effet, le terme de collision retranscrit le caractère discret des niveaux d'énergie pour lesquels un photon peut être émis ou absorbé. En fonction de la composition chimique du milieu, le nombre de ces niveaux discrets est potentiellement gigantesque. Les méthodes d'approximation numérique de cette équation, dans les configurations considérées ici, engendrent un coût de calcul inaccessible par les machines actuellement disponibles. En conséquence, des modélisations alternatives doivent être adoptées.

Parmi ces alternatives, on se concentrera sur un modèle aux moments dans lequel le rayonnement est décrit par les premiers moments de l'intensité radiative. Soulignons dès à présent que les variables des modèles aux moments définissent les grandeurs caractéristiques permettant d'appréhender les phénomènes physiques étudiés. De plus, dans les configurations pertinentes dans cette étude, ces modèles donnent généralement une approximation très satisfaisante de l'équation du transfert radiatif. En particulier, ils préservent les propriétés fondamentales du rayonnement. Si les moments "gris" (intégrés sur toutes les fréquences) suffisent souvent à une bonne description du rayonnement dans des solides, ce n'est plus le cas dans les gaz. Il est alors nécessaire de considérer une modélisation moins frustre en fréquence et ainsi d'introduire des modèles dits "multigroupes" (intégrés sur des intervalles de fréquences). Cette décomposition a l'avantage de permettre un couplage local (en fréquence) avec un modèle cinétique pour des fréquences particulièrement importantes.

L'objectif de cette thèse est de proposer des schémas d'approximation pour ces deux niveaux de modèles : modèle cinétique et modèle aux moments. Concernant les modèles cinétiques, lorsque le rayonnement doit être couplé à la matière, les simulations doivent être conduites sur des temps extrêmement longs relativement aux vitesses caractéristiques mises en jeu. Ceci implique le développement de schémas d'ordre très élevé sans restriction sur le pas de temps. Dans un second temps, on s'attache à améliorer l'efficacité des techniques d'approximation utilisées pour les modèles aux moments. En particulier, les méthodes proposées garantissent la préservation de régimes asymptotiques naturels pour les régimes considérés, y compris sur des maillages 2D non-structurés. L'ensemble de ces travaux a conduit à introduire des extensions pour la radiothérapie. En effet, sous certaines conditions, la radiothérapie peut être vue comme une application spécifique du transfert radiatif.

Plan de la thèse

Premier chapitre : le transfert radiatif

Dans ce chapitre, on présente le concept physique du transfert radiatif ainsi que différents modèles utilisés pour le décrire. On détaille les conditions particulières considérées dans cette étude pour l'approximation numérique de ce phénomène. On écrit alors l'équation du transfert radiatif (ETR) régissant l'évolution de l'intensité radiative dans ce contexte. Après avoir étudié les régimes limites (transport et diffusion) de cette équation, on s'intéresse aux différentes possibilités de modélisation. Dans un premier temps, on introduit le modèle cinétique d'ordonnées discrètes S_N. Ce modèle directionnel consiste à considérer un nombre discret N de directions de la sphère unité pour approcher l'intégrale en direction par une méthode de quadrature. En rappelant qu'un des avantage du transfert radiatif est de pouvoir traiter séparément la fréquence et la direction, on propose une discrétisation fréquentielle en bandes étroites. Ce modèle, très précis, pourra alors être utilisé lorsqu'une approximation très fine de la solution est nécessaire. Dans un second temps, on présente des modèles portant sur les variables macroscopiques. En particulier, on s'intéresse dans ce travail aux modèles aux deux moments M_1 gris et sa variante multigroupe. En calculant les deux premiers moments de l'ETR, sur toutes les fréquences pour le cas gris et sur des groupes de fréquences pour le cas multigroupe, on obtient des systèmes impliquant les trois premiers moment de l'intensité radiative et faisant intervenir des moyennes d'opacités. Pour fermer ces systèmes, on utilise un principe de minimisation d'entropie qui permet à la fois de préserver de nombreuses propriétés physiques de l'ETR et d'avoir une bonne approximation des moyennes d'opacités. Ces modèles hyperboliques donnent une bonne approximation de la solution avec un coût de calcul raisonnable. Par ailleurs, deux autres modèles sont proposés : le modèle d'harmoniques sphériques qui consiste à décomposer l'intensité radiative dans la base des polynômes de Legendre et un modèle de diffusion à flux limité qui est basé sur la limite diffusive de l'ETR.

Deuxième chapitre : schémas numériques pour le modèle d'ordonnées discrètes S_N
Ce chapitre est dédié à la construction d'un schéma d'ordre très élevé sans restriction sur le pas de temps pour le modèle S_N. Un tel schéma est en effet nécessaire pour certaines applications en temps très long que l'on est amené à considérer.
Ètant donné que le modèle S_N est un système d'équations d'advection linéaire avec terme source, on développe tout d'abord une méthode de type GRP (Problème de Riemann Généralisé) espace-temps pour un système hyperbolique linéaire générique. En profitant de la linéarité du système et grâce à une approximation polynomiale de la solution à la fois en espace et en temps, on développe un schéma de d'ordre arbitrairement élevé en espace et en temps sans restriction sur le pas de temps. Ce schéma présente de nombreux avantages. Il permet tout d'abord d'obtenir des solutions d'excellente qualité comparées aux autres méthodes d'ordre élevé (WENO, Galerkin discontinu, ...). Là où ces dernières sont limitées par une condition CFL et dont l'extension implicite est complexe et coûteuse, sa formulation permet de considérer naturellement des pas de temps aussi grands que nécessaire. Par ailleurs, il s'écrit et s'implémente à n'importe quel ordre sans effort superflu.
Le caractère compact de cette méthode permet de l'étendre sur des maillages 2D non-structurés moyennant une attention sur la géométrie. Les cas-test de validation permettent de confirmer son excellent comportement même en 2D.
Enfin, grâce à la forme particulière du terme source dans le modèle S_N qui permet de connaître la solution exacte pour une direction donnée, on étend la méthode à ce modèle. Cette extension s'avère d'autant plus intéressante que l'on retrouve les coefficients impliqués dans le schéma discrétisant la partie transport.

Troisième chapitre : étude des modèles M_1
Dans ce chapitre, on s'intéresse aux modèles aux moments M_1 gris et multigroupe.
Dans un premier temps, on étudie le problème de Riemann pour le modèle M_1 gris sans terme source en 1D. On montre l'existence inconditionnelle d'une solution admissible. On exhibe par ailleurs la forme explicite de cette solution, ce qui permet entre autre de construire le schéma de Godunov associé à ce problème.
Dans un second temps, on développe des schémas préservant les asymptotiques du modèle M_1 gris, à savoir les régimes de transport et de diffusion. En particulier, la principale difficulté pour obtenir des schémas qui dégénèrent correctement dans le régime de diffusion provient du fait que les opacités, qui pilotent la raideur du terme source, dépendent des variables. La particularité de la technique considérée pour discrétiser le terme source est de pouvoir contrôler la dégénérescence du schéma dans la limite de diffusion. Ainsi, on peut contourner un défaut du modèle M_1 gris et retrouver la limite physique de l'ETR.
Une extension de cette méthode sur des maillages 2D non-structurés est proposée. On obtient alors un schéma asymptotiquement consistant avec la limite de diffusion à partir de n'importe quelle reconstruction des gradients aux interfaces.
Enfin, on se consacre aux calculs des moyennes d'opacité intervenant dans le modèle M_1 multigroupe. Contrairement au cas gris, le calcul de ces moyennes n'est plus explicite et nécessite la recherche des zéros de fonctions non linéaires très raides. On développe alors une procédure de précalculs pour réduire de manière significative les temps de calcul.

Quatrième chapitre : équations et schémas numériques pour la radiothérapie
Ce chapitre est dédié à un schéma numérique pour prédire l'impact d'une radiothérapie. Dans cette application, on peut se ramener à utiliser des modèles M_1. Ceux-ci font toutefois intervenir des fonctions discontinues dans les flux résultant de grandes hétérogénéités dans le milieu.
On développe un premier schéma adapté à ces discontinuités de la fonction de flux limitant ainsi l'importante diffusion numérique générée par les schémas classiques dans ces conditions. On mon-

tre que ce schéma, basé sur des changements de variables ad hoc, possède de bonnes propriétés. Il n'est cependant pas généralisable en 2D. On propose alors une variante faisant intervenir des projections qui s'étend à des maillages 2D cartésiens. Les cas-test de validation montrent la pertinence de ces deux approches.

Cinquième chapitre : résultats numériques

On regroupe dans ce chapitre des résultats numériques illustrant l'ensemble des méthodes développées dans ce travail.

Tout d'abord, on présente de nombreuses simulations obtenues avec la méthode GRP espace-temps. En 1D, la méthode est ainsi testée pour un système linéaire sans terme source (équations de Maxwell) et pour le modèle S_N sur deux exemples dans lesquels le terme source est pris en compte soit par une méthode de splitting, soit par le schéma GRP étendu. Enfin, on considère le modèle S_N en 2D.

On présente ensuite des résultats pour le modèle M_1 multigroupe en 2D en utilisant les précalculs et le schéma préservant l'asymptotique.

Enfin, quelques calculs de dose en radiothérapie sont effectués à partir de scanners et comparés avec les résultats d'un code de référence basé sur des méthodes Monte-Carlo.

Publications Céline Sarazin-Desbois

Plusieurs publications sont issues de ce travail. Deux articles de revues internationales avec comité de relecture :

- C. Berthon, C. Sarazin, R. Turpault, *Space-time Generalized Riemann Problem solvers of order k for linear advection with unrestricted time step*, J. Sci. Comput., in press, (disponible en ligne),
- C. Berthon, M. Frank, C. Sarazin, R. Turpault, *Numerical methods for balance laws with space dependent flux : application to radiotherapy dose calculation*, Commun. Comput. Phys., Vol 10(2011), pp 1184-1210,

et deux actes de congrès avec comité de relecture :

- G. Duffa, C. Sarazin, R. Turpault, *Modelling and Numerical issues for the determination of thermal properties of PICA-like materials*, 4^{th} International Workshop on Radiation of High Temperature Gases in Atmospheric Entry, Barcelone(2012),
- C. Berthon, C. Sarazin, R. Turpault, *Numerical approximations of asymptotic regimes*, 3^{rd} International Workshop on Radiation of High Temperature Gases in Atmospheric Entry, Lausanne(2010).

1

Le transfert radiatif

1.1 Nomenclature

Constantes universelles

Vitesse de la lumière dans le vide	$c \simeq 2.99792458.10^{8}$	$m.s^{-1}$
Constante de Boltzmann	$k \simeq 1.3806503.10^{-23}$	$J.K^{-1}$
Constante de Planck	$h \simeq 6.626068.10^{-34}$	$J.s$
	$a = \frac{8\pi^5 k^4}{15h^3c^3} \simeq 7.565916.10^{-16}$	$J.m^{-3}.K^{-4}$

Grandeurs globales

$\mathbf{x}$	position	m
t	temps	s
ν	fréquence	s^{-1}
Ω	(vecteur) direction	rad
T	température	K

Variables utilisées en transfert radiatif

σ_ν	opacité d'absorption	m^{-1}
σ_ν^d	opacité de *scattering*	m^{-1}
σ^a	opacité moyenne d'absorption	m^{-1}
σ^e	opacité moyenne d'émission	m^{-1}
σ^f	opacité moyenne d'absorption du flux	m^{-1}
σ^R	moyenne de Rosseland	m^{-1}
$p_\nu(\Omega, \Omega')$	probabilité de redistribution angulaire	
$I(t, \mathbf{x}, \Omega, \nu)$	intensité radiative	$J.sr^{-1}.m^{-2}$
$E_R(t, \mathbf{x})$	énergie radiative	$J.m^{-3}$
$F_R(t, \mathbf{x})$	flux radiatif	$W.m^{-2}$
$P_R(t, \mathbf{x})$	pression radiative	$J.m^{-3}$
$f = \frac{\|F_R\|}{cE_R}$	facteur d'anisotropie	
C_v	capacité calorifique à volume constant	$J.K^{-1}$

Définitions et relations

$B_\nu(T) = \frac{2h\nu^3}{c^2}\left[\exp\left(\frac{h\nu}{kT}\right) - 1\right]^{-1}$	fonction du corps noir de Planck
$h_R(I) = \frac{2k\nu^3}{c^3}\left[n\ln n - (n+1)\ln(n+1)\right]$	entropie radiative microscopique
$H_R(I) = < h_R(I) >$	entropie radiative macroscopique
$\chi(f) = \frac{3+4f^2}{5+2\sqrt{4-3f^2}}$	facteur d'Eddington

Variables utilisées en radiothérapie

ε	énergie de la particule	J
$\rho(x)$	densité de la matière	$kg.m^{-3}$
σ	noyau de *scattering*	m^{-1}
$S(\varepsilon)$	pouvoir d'arrêt	$eV.m^{-1}$
$T_{tot}(x,\varepsilon)$	coefficient de transport	m^{-1}
$\bar{\sigma}(\varepsilon, \Omega.\Omega')$	section efficace	
$\Psi(x,\varepsilon,\Omega)$	densité d'électrons f $\times$ vitesse des particules v	
$\Psi^0(x,\varepsilon)$	moment d'ordre 0 de Ψ	
$\Psi^1(x,\varepsilon)$	moment d'ordre 1 de Ψ	
$\Psi^2(x,\varepsilon)$	moment d'ordre 2 de Ψ	
$D(x) = \frac{1}{\rho(x)}\int_0^\infty S_M(x,\varepsilon)\int_{S^2}\Psi(x,\varepsilon,\Omega)d\Omega d\varepsilon$	dose	Gy

1.2 Généralités sur le Transfert Radiatif

Dans ce chapitre, on présente les concepts physiques du rayonnement ainsi que l'équation du transfert radiatif qui décrit les échanges d'énergie entre le rayonnement et la matière. Il existe un principe de dualité onde - particule pour décrire le rayonnement, on considère ici uniquement l'aspect corpusculaire. Dans ce cadre, le rayonnement est véhiculé par les photons ou "quanta de lumière". Dans le vide, ces bosons de masse nulle se déplacent en ligne droite avec une vitesse $c\Omega$ où $c > 0$ est la vitesse de la lumière (par définition) et Ω la direction de propagation normalisée ($\|\Omega\| = 1$). L'énergie d'un photon est liée à une fréquence, que l'on peut voir comme une "couleur", donnée par la relation : $E = h\nu$ où h est la constante de Planck et ν la fréquence du photon. Un photon donné est donc déterminé par sa fréquence ν, sa direction de propagation Ω et sa position $(t, \mathbf{x})$.
Puis, en présence de matière, le rayonnement interagit avec celle-ci via trois procédés principaux : l'absorption, l'émission et le *scattering*.
Les atomes de la matière peuvent capter un photon pour passer dans un niveau d'énergie supérieur. À cause de la nature discrète de ses niveaux d'énergie, un atome dans un niveau d'énergie E_1 ne peut capter que des photons d'énergie $h\nu$ tels que $h\nu = E_2 - E_1$ où E_2 est l'énergie de l'atome dans un niveau plus excité. Un élément donné ne peut donc absorber *a priori* que des photons ayant une énergie dans un ensemble discret et caractéristique. Cet ensemble, appelé spectre d'absorption, est composé de raies d'absorption qui, en pratique, ne sont pas des masses de Dirac. En effet, certains phénomènes physiques régularisent celle-ci (principe d'incertitude d'Heisenberg, élargissement Doppler ou Voigt, ...). Finalement, en notant ρ_i la densité de l'espèce chimique i et T la température du milieu, l'absorption est caractérisée par l'opacité $\sigma = \sigma(\nu, T, \rho_i, ...)$ qui est l'inverse du libre parcours moyen d'absorption, c'est-à-dire la distance moyenne entre l'absorption de deux photons de fréquence ν. Cette opacité dépend de la fréquence mais aussi de la densité des espèces chimiques considérées et de la température du milieu. Pour simplifier les notations, on utilisera σ_ν pour désigner cette opacité d'absorption.
À l'inverse, un atome qui n'est pas dans son état fondamental mais dans un état excité a tendance à y retourner progressivement en émettant des photons. Cela lui permet de passer d'un niveau excité à un niveau plus stable. Les photons ainsi émis font partie du même ensemble que ceux potentiellement absorbés si l'on néglige le phénomène de résonance. À l'équilibre thermique local, l'émission est donc associée à la même opacité σ_ν qui est l'inverse de la distance moyenne entre l'émission de deux photons de fréquence ν.
Par ailleurs ces deux phénomènes (absorption et émission) ne s'appliquent pas de la même manière à des ions. En effet, à cause de leur surplus ou défaut d'électrons, ils peuvent aussi absorber ou émettre des photons de toute énergie. Ainsi, des composantes continues apparaissent dans leur spectre.
Toujours à l'équilibre thermique local, l'émission est gouvernée par la fonction dite du corps noir de Planck :

$$B_\nu(T) = \frac{2h\nu^3}{c^2}\left[\exp\left(\frac{h\nu}{kT}\right) - 1\right]^{-1}, \tag{1.1}$$

où k est la constante de Boltzmann et T est la température matière.
Pour illustrer l'allure d'un spectre d'absorption, on représente sur la figure 1.1 le spectre du Krypton obtenu par la base de données $NIST$ disponible que le site (http ://www.nist.gov).

Finalement, lors d'un évènement de *scattering*, le photon n'est pas absorbé par la matière, mais sa trajectoire et sa fréquence peuvent être modifiées. Il existe différents types de *scattering* comme par exemple le *scattering* élastique dans lequel un photon est uniquement dévié de sa trajectoire. Le *scattering* inélastique, quant à lui, modifie la trajectoire du photon et le divise en deux photons

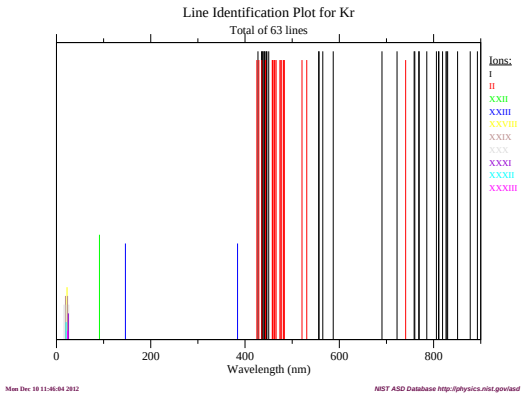


FIGURE 1.1 – Spectre du Krypton.

dont la somme des énergies est égale à l'énergie du photon initial. Ce phénomène est caractérisé à la fois par une opacité σ_ν^d qui est l'inverse du libre parcours moyen de *scattering* et par une probabilité de redistribution angulaire notée $p_\nu(\Omega, \Omega')$.
À cause de ces interactions, en particulier de la libération ou captation de photons par la matière, il existe des échanges d'énergie entre le rayonnement et la matière. Par conséquent, pour assurer la conservation de l'énergie, il est nécessaire de coupler le rayonnement et la matière. Suivant les applications physiques auxquelles on s'intéresse, plusieurs types de couplages peuvent être envisagés. Tout d'abord, on peut considérer un système totalement découplé où la température matière est supposée constante. C'est le cas par exemple en climatologie lorsque l'on souhaite étudier les effets du rayonnement. Pour des applications comme des incendies ou des rentrées terrestres de certaines sondes orbitales, un couplage faible est utilisé. En effet, étant données les échelles de températures mises en jeu, le rayonnement et la matière influent peu l'un sur l'autre. Puis, pour des cas de rentrées super orbitales ou de formation d'étoiles par exemple, un couplage fort est nécessaire.
Pour écrire les équations du transfert radiatif, on note

$$f = f(t, \mathbf{x}, \Omega, \nu),$$

la fonction de distribution des photons pour définir alors la variable d'intérêt du transfert radiatif appelée intensité radiative par :

$$I(t, \mathbf{x}, \Omega, \nu) = ch\nu f(t, \mathbf{x}, \Omega, \nu) > 0.$$

Par exemple, à l'équilibre thermodynamique, l'intensité radiative est décrite par la fonction du corps noir de Planck (1.1). Après un bilan d'énergie au niveau microscopique, l'intensité radiative est régie à l'équilibre thermique local par l'équation du transfert radiatif (ETR) qui peut s'écrire :

$$\begin{aligned}\frac{1}{c}\partial_t I(t, \mathbf{x}, \Omega, \nu) + \Omega \cdot \nabla I(t, \mathbf{x}, \Omega, \nu) &= \sigma_\nu \left(B_\nu(T) - I(t, \mathbf{x}, \Omega, \nu)\right) \\ &\quad + \sigma_\nu^d \frac{1}{4\pi} \left(\int_{S^2} p_\nu(\Omega'.\Omega) I(t, \mathbf{x}, \Omega', \nu) d\Omega' - I(t, \mathbf{x}, \Omega, \nu) \right),\end{aligned} \tag{1.2}$$

où $\frac{1}{4\pi} p_\nu$ est la probabilité qu'une particule arrivant dans une direction Ω' soit déviée dans la direction Ω. Pour obtenir cette écriture simplifiée de l'équation du transfert radiatif, on a supposé que le milieu est d'indice optique 1, que l'on se trouve à l'équilibre thermique local, que l'on néglige les effets de polarisation et que le *scattering* modifie uniquement la direction de la particule [92, 93, 112].
Afin d'alléger les notations dans la suite, on omettra les dépendances en temps et en espace des

variables dès qu'aucune confusion n'est possible. De plus, on négligera les effets du *scattering* excepté dans la partie dédiée à la radiothérapie où ils sont, au contraire, prépondérants.

Certaines grandeurs macroscopiques ont un intérêt physique dans la description du phénomène. Ces variables, appelées moments, sont obtenues en intégrant les variables microscopiques en direction et en fréquence. En particulier, on s'intéresse aux trois premiers moments de l'intensité radiative :

$$\text{Énergie radiative : } E_R = \frac{1}{c}\int_{S^2}\int_0^\infty I(\nu,\Omega)d\nu d\Omega, \tag{1.3}$$

$$\text{Flux radiatif : } F_R = \frac{1}{c}\int_{S^2}\int_0^\infty c\Omega\cdot I(\nu,\Omega)d\nu d\Omega, \tag{1.4}$$

$$\text{Pression radiative : } P_R = \frac{1}{c}\int_{S^2}\int_0^\infty \Omega\otimes\Omega I(\nu,\Omega)d\nu d\Omega,. \tag{1.5}$$

Comme l'intensité radiative I est positive et $\|\Omega\| = 1$, le facteur d'anisotropie f est naturellement limité :

$$f = \frac{\|F_R\|}{cE_R} < 1.$$

Si cette propriété fondamentale de limitation du flux est violée, cela implique que E_R et F_R ne sont pas les deux premiers moments d'une intensité radiative positive.

Notation. Pour simplifier l'écriture des moments, on emploiera la notation suivante :

$$< . >= \frac{1}{c}\int_{S^2}\int_0^\infty . \, d\nu d\Omega.$$

Si l'on s'intéresse aux moments de la fonction de Planck, on voit facilement qu'elle possède un flux nul car elle est isotrope, c'est-à-dire qu'elle ne dépend pas de la direction. De plus, son énergie peut être calculée analytiquement :

$$< B_\nu >= aT^4$$

où $a \simeq 7.56.10^{-16} J.m^{-3}.K^{-4}$ est une constante (voir nomenclature). Enfin, le caractère isotrope de B implique également que

$$< \Omega\otimes\Omega B_\nu >= \frac{aT^4}{3}$$

L'entropie radiative du système est une autre grandeur physique qui aura un rôle important par la suite. Elle est obtenue à partir du deuxième principe de la thermodynamique et donnée par :

$$H_R(I) =< h_R(I) >=< \frac{2k\nu^3}{c^3}\left[n\ln n - (n+1)\ln(n+1)\right] >,$$

où n est le nombre d'occupation relié à I par la relation :

$$n = \frac{c^2}{2h\nu^3}I.$$

On peut remarquer par un calcul direct que la fonction de Planck est le minimum de cette entropie :

$$B_\nu = \text{argmin}\{H_R(I) =< h_R(I) >\}.$$

Comme mentionné ci-dessus, l'énergie radiative n'est pas conservée si l'on ne considère que l'équation du transfert radiatif. Pour ce faire, il faut donc coupler l'ETR avec une équation régissant l'énergie matière. Étant donné que l'on s'intéresse essentiellement aux effets du rayonnement,

on prend en compte uniquement les termes d'échanges entre la matière et le rayonnement suivants l'équation :

$$\rho C_v \partial_t T = -\int_0^\infty \int_{S^2} \sigma_\nu (B_\nu(T) - I) d\Omega d\nu. \tag{1.6}$$

Cette équation peut s'obtenir par exemple en considérant $\rho = cte$ et $u = cte$ dans les équations de Navier-Stokes.
D'un point de vue physique, l'équation du transfert radiatif comporte deux régimes limites suivant la nature des opacités. Quand l'opacité σ_ν est grande, l'équation du transfert radiatif dégénère en une équation de diffusion à l'équilibre [93] :

$$\partial_t (aT^4 + \rho C_v T) - \nabla \left(\frac{c}{3\sigma^R} \nabla aT^4 \right) = 0, \tag{1.7}$$

où σ^R est la moyenne d'opacité de Rosseland définie par :

$$\sigma^R = \frac{\int_0^\infty \partial_t B_\nu(T) d\nu}{\int_0^\infty \frac{1}{\sigma_\nu} \partial_t B_\nu(T) d\nu}.$$

En effet, dans ce cas, l'intensité radiative proche de l'équilibre tend à devenir une Planckienne (1.1). Certains modèles sont basés sur cette limite, tels que le modèle de diffusion à flux limité.
À l'inverse, quand $\sigma_\nu = 0$, il est évident que l'ETR (1.2) se ramène à une équation de transport.

Comme l'opacité σ_ν dépend de la température matière T et d'autres grandeurs hydrodynamiques, on peut se retrouver dans certaines applications avec des zones *a priori* non déterminées pour lesquelles l'opacité est très faible, d'autres où l'opacité est très grande et des régimes intermédiaires. Numériquement, il faut donc être capable de gérer simultanément ces trois régimes dont les équations sont de natures différentes.

Si l'on veut résoudre numériquement l'équation du transfert radiatif, il faudrait échantillonner toutes les directions, mais aussi toutes les fréquences d'intérêt. Pour calculer l'opacité du milieu considéré, il faudrait ainsi échantillonner toutes les raies du spectre correspondant. À l'heure actuelle, pour un calcul à t, $\mathbf{x}$ et Ω fixés, plusieurs jours de calcul sont nécessaires pour obtenir une approximation aussi fine de l'opacité σ_ν. C'est pourquoi il est nécessaire de simplifier cette équation pour réaliser certaines simulations d'intérêt physique avec un temps de calcul raisonnable. En fonction du niveau de précision souhaitée, on recense une hiérarchie de modèles. L'une des caractéristiques du modèle est que l'on peut faire un traitement différent en ν et en Ω. Pour la discrétisation fréquentielle, on compte quatre niveaux :

- *Raie par raie.* Chaque espèce chimique comptant des millions de raies, cette résolution est très précise, mais extrêmement coûteuse. De plus, les bases de données complètes pour de larges gammes de températures et pressions étant rares, les applications sont restreintes.
- *Bandes étroites* de fréquences, c'est-à-dire de petits intervalles de fréquences sur lequel la fonction de Planck B peut être considérée constante.
- *Multigroupe* où l'on considère de grands intervalles de fréquences (appelés groupes) diminuant ainsi le nombre d'inconnues, mais où les moyennes d'opacités sont à calculer avec précaution.
- *Gris*, c'est-à-dire intégré sur toutes les fréquences.

En pratique, pour une simulation donnée, si l'on considère plusieurs millions de raies, on prendra quelques milliers de bandes étroites ou au plus quelques dizaines de groupes.
Au niveau directionnel, on compte essentiellement deux niveaux de résolution :

- Ordonnées discrètes, c'est-à-dire que l'on fixe N directions de la sphère et que l'on approche l'intégrale sur la sphère par une méthode de quadrature.

- Intégré sur tout ou une partie de la sphère et où les inconnues sont des moments angulaires.

Même si l'on peut découpler les calculs en fréquences et en directions, en pratique, on considère un niveau de discrétisation équivalent pour ces deux grandeurs. Lorsque la discrétisation est fine en fréquences, elle est également fine en directions et vice versa. Il existe ainsi essentiellement deux catégories de modèles : les modèles cinétiques et les modèles macroscopiques.

Dans le chapitre suivant, différents modèles du transfert radiatif sont présentés. Tout d'abord, on détaille un modèle cinétique d'ordonnées discrètes S_N qui sera l'objet d'un nouveau schéma numérique développé dans le chapitre 2. Ensuite, on propose quelques modèles macroscopiques dont un modèle de diffusion à flux limité et un modèle directionnel de décomposition en harmoniques sphériques. Finalement, deux modèles aux moments, le modèle M_1 et sa variante multigroupe sont présentés. Ces deux modèles feront l'objet de plus d'études notamment dans les chapitres 3 et 4.

1.3 Le modèle d'ordonnées discrètes S_N

Le modèle d'ordonnées discrètes est un modèle directionnel qui a initialement été introduit par Chandrasekhar dans [30]. Il consiste à considérer N directions $(\Omega_l)_{1\leq l\leq N}$ de la sphère unité et à remplacer l'intégration en direction par une méthode de quadrature dont les points sont les Ω_l. À chaque direction Ω_l est alors associé un poids de quadrature ω_l. Pour la présentation de ce modèle, l'opacité d'absorption σ_ν est supposée constante et notée σ. L'équation du transfert radiatif ainsi considérée est :

$$\frac{1}{c}\partial_t I(\Omega,\nu) + \Omega\cdot\nabla I(\Omega,\nu) = \sigma\left(B_\nu(T) - I(\Omega,\nu)\right). \tag{1.8}$$

Si on note $I_l^{S_N}$ une approximation de l'intensité radiative évaluée en Ω_l : $I(\Omega_l)$, alors le modèle d'ordonnées discrètes S_N revient à résoudre un système d'équations de taille N :

$$\forall\, 1\leq l\leq N,\quad \partial_t I_l^{S_N} + c\Omega_l\cdot\nabla I_l^{S_N} = c\sigma\left(B_\nu - I_l^{S_N}\right). \tag{1.9}$$

En fonction du nombre de directions considérées, ce système linéaire d'équations aux dérivées partielles peut devenir très grand [81]. De plus, on peut considérer différentes discrétisations, comme des directions équi-réparties ou des directions concentrées dans une zone de la sphère répondant par exemple à une anisotropie particulière du cas-test.

En général, ce modèle est associé à une discrétisation fréquentielle en bandes étroites (modèles *correlated-k* [56]). Ce choix est notamment fait afin de pouvoir supposer que l'opacité σ_ν est constante dans chaque bande de fréquences. Cependant, d'autres discrétisations peuvent être envisagées.
On présente ici un modèle proposé dans [105, 31] qui consiste à approcher l'intégrale en fréquence par une seconde méthode de quadrature. On considère alors Q bandes de fréquences $[\nu_{q-\frac{1}{2}};\nu_{q+\frac{1}{2}}[_{1\leq q\leq Q}$ de milieu ν_q pour discrétiser le spectre. Par convention, on prend $\nu_{Q+\frac{1}{2}} = +\infty$. Puis, sur chaque bande de fréquence, on suppose que $\sigma_\nu = \sigma_q$ et on utilise une quadrature de point milieu. Ainsi, à chaque fréquence ν_q est associé un poids de quadrature θ_q. Avec ces notations, l'intégration en fréquence est approchée par :

$$\int_0^{+\infty} I(\Omega,\nu)d\nu \simeq \sum_{q=1}^{Q} I(\Omega,\nu_q)\theta_q,$$

où $I(\Omega,\nu_q)$ peut être vue comme une approximation de l'intensité radiative dans la direction Ω sur la bande de fréquence $[\nu_{q-\frac{1}{2}};\nu_{q+\frac{1}{2}}[$. Si l'on note $<\,.\,>_d$ l'opérateur discret de $<\,.\,>$ obtenu

en remplaçant les intégrations en fréquence et direction par les quadratures ainsi définies, on peut écrire :

$$< I > \simeq < I >_d = \frac{1}{c} \sum_{l,q} I(\Omega_l, \nu_q) \omega_l \theta_q.$$

Pour simplifier les notations, on note :

$$I_{l,q} = I(\Omega_l, \nu_q). \tag{1.10}$$

Finalement, l'énergie radiative est approchée par :

$$E_R = \frac{1}{c} \int_0^{+\infty} \int_{S^2} I(\Omega, \nu) d\Omega d\nu \simeq \frac{1}{c} \sum_{l,q} I_{l,q} \omega_l \theta_q = E_d.$$

Cette seconde approximation entraîne une modification du modèle (1.9) :

$$\forall\, 1 \leq l \leq N, 1 \leq q \leq Q, \quad \partial_t I_{l,q}^{S_N} + c\Omega_l . \nabla I_{l,q}^{S_N} = c\sigma_q \left(\mathcal{B}_q - I_{l,q}^{S_N} \right),$$

où $\mathcal{B}_q$ est une approximation de B_ν sur la bande de fréquence $[\nu_{q-\frac{1}{2}}; \nu_{q+\frac{1}{2}}[$. On choisit ici de l'approcher suivant un principe de minimisation de l'entropie discrète. En effet, en remplaçant l'intégration en fréquence par une quadrature, en général, $< B_\nu >_d \neq aT^4$. Pour maintenir l'énergie de B_ν égale à aT^4, l'idée est de définir la fonction $\mathcal{B}_q$ suivant le principe de minimisation d'entropie sous contraintes :

$$H^*(\mathcal{B}) = \text{argmin}\{H^*(I), < I >_d = aT^4\},$$

où l'entropie discrète est donnée par :

$$H^*(I) = < h^*(I) >_d = < \frac{2k\nu^3}{c^3} \left[n \ln(n) - (n+1) \ln(n+1) \right] >_d .$$

En pratique, on peut montrer que la fonction $\mathcal{B}_q$ peut s'écrire sous la forme :

$$\forall\, 1 \leq q \leq Q, \quad \mathcal{B}_q = \mathcal{B}(T, \nu_q) = \frac{2h\nu_q^3}{c^2} \left[\exp\left(\frac{ch\nu_q \alpha}{k} \right) - 1 \right]^{-1},$$

où $\alpha = \alpha(T)$, le multiplicateur de Lagrange, est obtenu par la résolution de $< \mathcal{B}_q >= aT^4$.

Finalement, avec cette discrétisation fréquentielle, le modèle d'ordonnées discrètes pour (1.8) s'écrit :

$$\forall\, 1 \leq l \leq N, 1 \leq q \leq Q, \quad \partial_t I_{l,q}^{S_N} + c\Omega_l . \nabla I_{l,q}^{S_N} = c\sigma_q \left(\mathcal{B}_q - I_{l,q}^{S_N} \right). \tag{1.11}$$

Ce modèle peut notamment être utilisé pour approcher l'équation du transfert radiatif lorsqu'une approximation très précise de la solution est demandée. En effet, l'approximation sera d'autant plus précise si l'on considère un grand nombre de directions. On rappelle qu'une condition nécessaire pour la validité du modèle est de pouvoir supposer l'opacité σ_ν constante sur chaque intervalle de fréquence considéré, c'est pourquoi on utilise des bandes étroites. Pour discrétiser ce modèle, on développe dans le paragraphe 2.5 un nouveau schéma numérique de type GRP espace-temps d'ordre élevé en temps et en espace sans restriction sur le pas de temps. On rappelle également dans le paragraphe 2.6.1, le schéma volumes finis décentré amont associé à ce modèle. Quelques cas-tests dans le paragraphe 5.1 viendront illustrer ces différents schémas.

1.4 Les modèles méso/macroscopiques

1.4.1 Les modèles aux moments

Les modèles aux moments, qui ont initialement été introduits par Grad dans son travail sur les gaz raréfiés [67], consistent à décrire le système à l'aide des moments des variables cinétiques. Les équations obtenues sont plus simples que les équations cinétiques dont elles dérivent tout en préservant, en général, beaucoup de leurs propriétés fondamentales. De plus, la connaissance détaillée de la solution pour chaque fréquence et direction n'est pas toujours nécessaire. Pour certaines applications, des moyennes directionnelles ou fréquentielles sont suffisantes. Bien entendu, si l'on considérait une infinité de moments, on pourrait reconstruire exactement I. Mais on s'intéresse ici aux équations faisant intervenir les deux premiers moments de l'intensité radiative : E_R et F_R. Sachant que I est positive, l'espace des états physiquement admissibles en terme des moments de I est alors donné par :

$$\mathcal{A} = \{(E_R, F_R, T) \in \mathbb{R}^4, E_R \geq 0, f = \frac{\|F_R\|}{cE_R} \in [0,1], T > 0\}. \tag{1.12}$$

On considère l'ETR à l'équilibre thermique local (1.2) couplée à l'équation matière (1.6). Pour obtenir la première équation, on intègre l'équation (1.2) sur la sphère unité pour obtenir :

$$\partial_t E_\nu + \nabla . F_\nu = \sigma_\nu (4\pi B_\nu(T) - cE_\nu), \tag{1.13}$$

où $E_\nu = \frac{1}{c}\int_{S^2} I(\Omega,\nu)d\Omega$ et $F_\nu = \frac{1}{c}\int_{S^2} c\Omega I(\Omega,\nu)d\Omega$.
Pour la deuxième équation, on multipliant l'ETR par Ω pour obtenir en intégrant sur la sphère unité :

$$\partial_t F_\nu + c^2 \nabla . P_\nu = -c\sigma_\nu F_\nu. \tag{1.14}$$

On rappelle que du fait de son caractère isotrope, la fonction de Planck B_ν a un flux radiatif nul.
On obtient ainsi un modèle intermédiaire, intégré en direction :

$$\begin{cases} \partial_t E_\nu + \nabla . F_\nu = \sigma_\nu (4\pi B_\nu(T) - cE_\nu), \\ \partial_t F_\nu + c^2 \nabla . P_\nu = -c\sigma_\nu F_\nu, \end{cases} \tag{1.15}$$

Il reste alors à intégrer en fréquences. Par exemple, le modèle aux deux moments gris est obtenu en intégrant ce système (1.15) sur toutes les fréquences :

$$\begin{cases} \partial_t E_R + \nabla . F_R = c(\sigma^e a T^4 - \sigma^a E_R), \\ \partial_t F_R + c^2 \nabla . P_R = -c\sigma^f F_R. \end{cases} \tag{1.16}$$

où les σ^e, σ^a et σ^f sont des opacités moyennes définies par :

$$\begin{aligned} \sigma^e a T^4 &= < \sigma_\nu B(T) >, \\ \sigma^a E_R &= < \sigma_\nu I >, \\ \sigma^f F_R &= c < \sigma_\nu \Omega I > . \end{aligned}$$

On remarque que ces définitions des moyennes d'opacités font intervenir l'intensité radiative I qui n'est plus une variable du système. En général, on ne peut donc pas les calculer directement. On verra que dans le modèle M_1, on peut faire une approximation satisfaisante de l'intensité radiative I.
Avec ces définitions des moyennes d'opacités, l'équation d'énergie matière (1.6) s'écrit :

$$\rho C_v \partial_t T = -c(\sigma^e a T^4 - \sigma^a E_R).$$

On voit que le système comporte plus d'inconnues que d'équations. Il faut donc faire une hypothèse de fermeture qui consiste à exprimer la pression radiative P_R en fonction de E_R et F_R. Cette étape cruciale de la modélisation va déterminer les propriétés du modèle. On s'intéressera dans la suite au modèle associé à l'un de ces choix de fermeture : le modèle M_1, ainsi que sa variante multigroupe.

Le modèle M1 gris

Le modèle M_1 a été introduit par Dubroca et Feugeas dans [41]. L'idée de celui-ci est de fermer le système grâce à un principe de minimisation d'entropie. Pour cela, on cherche l'intensité radiative $\mathcal{I}$ qui minimise l'entropie H_R et dont les premiers moments sont exactement E_R et F_R. Ceci conduit au problème de minimisation avec contraintes suivant :

$$H_R(\mathcal{I}) = \min\{H_R(I) =< h_R(I) > / < I >= E_R \text{ et } c < \Omega I >= F_R\}. \tag{1.17}$$

La solution de ce problème est donnée par :

$$\mathcal{I}(\Omega,\nu) = \frac{2h\nu^3}{c^2}\left[\exp\left(\frac{h\nu}{k}\alpha.\mathbf{m}\right) - 1\right]^{-1}, \tag{1.18}$$

où $\alpha = (\alpha,\beta)^T$ est le multiplicateur de Lagrange associé au problème (1.17) et $\mathbf{m} = (1,\Omega)^T$. Comme dans [87, 107], la pression radiative peut s'écrire sous la forme d'Eddington :

$$P_R = D_R E_R, \tag{1.19}$$

où D_R est le tenseur d'Eddington :

$$D_R = \frac{1-\chi}{2}\mathbf{I} + \frac{3\chi-1}{2}\frac{F_R \otimes F_R}{\|F_R\|^2}, \tag{1.20}$$

et χ est le facteur d'Eddington :

$$\chi = \chi(f) = \frac{3+4f^2}{5+2\sqrt{4-3f^2}}, \tag{1.21}$$

où $f = \dfrac{\|F_R\|}{E_R}$ est le facteur d'anisotropie.
En remplaçant ainsi l'expression de P_R dans (1.16), on obtient finalement le modèle M_1 :

$$\begin{cases} \partial_t E_R + \nabla.F_R = c(\sigma^e a T^4 - \sigma^a E_R), \\ \partial_t F_R + c^2\nabla.(D_R E_R) = -c\sigma^f F_R, \end{cases} \tag{1.22}$$

où les moyennes d'opacités sont approchées par :

$$\begin{aligned} \sigma^e a T^4 &=< \sigma_\nu B(T) >, \\ \sigma^a E_R &\simeq< \sigma_\nu \mathcal{I} >, \\ \sigma^f F_R &\simeq c < \sigma_\nu \Omega\mathcal{I} > . \end{aligned} \tag{1.23}$$

En effet, comme $\mathcal{I}$ est une approximation de l'intensité radiative I, les moyennes d'opacités σ^a et σ^f impliquant $\mathcal{I}$ sont des approximations des moyennes exactes.

Le modèle M_1 a la bonne limite diffusive si l'opacité σ_ν est constante en fréquence. Dans le cas contraire, la moyenne d'opacité obtenue par le modèle ne tend pas vers la limite physique qui est la moyenne de Rosseland, mais la différence est cependant bien moindre que si l'on considère la moyenne de Planck. Ce modèle a également la bonne limite de transport. De plus, la fermeture

suivant la minimisation d'entropie permet d'assurer la positivité de l'énergie ainsi que la limitation du flux, ce qui n'est pas le cas des modèles P_N par exemple rappelés plus loin. Le système obtenu est hyperbolique (voir [41]), ce qui permet de lui appliquer des méthodes classiques dédiées à l'étude et à la résolution numérique de ces systèmes. Dans le paragraphe 3.2.1, le problème de Riemann est résolu pour ce modèle M_1 en dimension un. En effet, la connaissance de cette solution exacte est très utile par exemple pour la validation des schémas numériques discrétisant ce modèle. Puis dans le paragraphe 3.2.4, un schéma préservant l'asymptotique est développé pour ce modèle. Malgré ces propriétés, le modèle M_1 possède quelques inconvénients. Par exemple, l'intégration sur toute la sphère ne lui permet pas de reproduire certains phénomènes, comme deux rayons de même direction et de sens opposés. Pour pallier ce problème, des modèles aux moments partiels ont été développés [42, 43, 54, 108]. De même, l'intégration sur tout le spectre de fréquences est souvent trop grossière surtout dans des gaz ou des plasmas. On propose alors une extension fréquentielle multigroupe de ce modèle dans le paragraphe suivant.

Le modèle M_1 multigroupe

Le modèle M_1 gris possède de bonnes propriétés (limitation du flux, régimes limites, ...), mais ne permet pas de considérer des modifications fréquentielles puisque les variables sont intégrées en fréquence. Étant donnée la forme du spectre de fréquences qui contient des millions de raies, il est souvent important de les prendre en compte. Le modèle M_1 multigroupe, qui a été développé par Turpault dans [104, 106], est une extension du modèle M_1 gris. L'idée est de découper le spectre en groupes de fréquences dans le modèle aux moments intermédiaire (1.15) et de fermer chaque système suivant le principe de minimisation de l'entropie.
On divise le spectre en Q groupes de fréquences : $[0,+\infty[=\cup_{q=0}^{Q}[\nu_q,\nu_{q+1}[$. Dans chaque groupe de fréquences, on définit les moments de I de la manière suivante :

$$E_q=\frac{1}{c}\int_{S^2}\int_{\nu_q}^{\nu^{q+1}} I d\nu d\Omega =< I >_q, \tag{1.24}$$

$$F_q=\frac{1}{c}\int_{S^2}\int_{\nu_q}^{\nu^{q+1}} c\Omega I d\nu d\Omega =< c\Omega I >_q, \tag{1.25}$$

$$P_q=\frac{1}{c}\int_{S^2}\int_{\nu_q}^{\nu^{q+1}} \Omega\otimes\Omega I d\nu d\Omega =< \Omega\otimes\Omega I >_q . \tag{1.26}$$

On intègre alors le système aux moments intermédiaire (1.15) sur chaque groupe q de fréquences $[\nu_q,\nu_{q+1}[$:

$$\left\{\begin{array}{l} \partial_t E_q+\nabla.F_q=c(\sigma_q^e a\theta_q^4(T)-\sigma_q^a E_q),\\ \partial_t F_q+c^2\nabla.P_q=-c\sigma_q^f F_q, \end{array}\right. \tag{1.27}$$

où $\theta_q(T)$ est défini tel que :

$$a\theta_q^4(T)=< B(T) >_q,$$

et les moyennes d'opacités sont définies par :

$$\begin{aligned} \sigma_q^e< B(T) >_q &=< \sigma_\nu B(T) >_q,\\ \sigma_q^a E_q &=< \sigma_\nu I >_q,\\ \sigma_q^f &= \frac{c< \sigma_\nu\Omega I >_q}{F_q}. \end{aligned} \tag{1.28}$$

Puis, pour fermer chaque système (1.27), on considère de nouveau que l'intensité radiative est obtenue par minimisation de l'entropie :

$$H_R(\mathcal{I})=\min\{H_R(I)=\sum_q < h_R(I) >_q /\forall q, < I >_q= E_q \text{ et } < c\Omega I >_q= F_q\}. \tag{1.29}$$

La résolution de ce problème de minimisation (1.29) permet alors d'obtenir une expression explicite de $\mathcal{I}$:

$$\mathcal{I}(\Omega,\nu)=\sum_q \mathbb{I}_q \frac{2h\nu^3}{c^2}\left[\exp\left(\frac{ch\nu}{k}\alpha_{\mathbf{q}}.\mathbf{m}\right)-1\right]^{-1}, \tag{1.30}$$

où $\alpha_q=(\alpha_q,\beta_q)^T$ est le multiplicateur de Lagrange associé au problème (1.29) et $\mathbf{m}=(1,\Omega)^T$. La pression radiative sur le groupe de fréquence q s'écrit de nouveau en fonction du tenseur d'Eddington D_q :

$$P_q=D_qE_q,$$

avec :

$$D_q=\frac{1-\chi_q}{2}\mathbf{I}+\frac{3\chi_q-1}{2}\frac{F_q\otimes F_q}{\|F_q\|^2}.$$

Contrairement au cas gris, le facteur d'Eddington χ_q n'est plus explicite.
Finalement, le modèle M_1 multigroupe couplé à la matière, s'écrit :

$$\begin{cases} \partial_t E_q+\nabla.F_q=c(\sigma_q^e a\theta_q^4(T)-\sigma_q^a E_q),\\ \partial_t F_q+c^2\nabla.(D_qE_q)=-c\sigma_q^f F_q, \qquad \forall 1\le q\le Q\\ \rho C_v\partial_t T=-c\sum_{q=1}^Q(\sigma_q^e a\theta_q^4-\sigma_q^a E_q). \end{cases} \tag{1.31}$$

On voit facilement que si l'on ne considère qu'un groupe de fréquences, le modèle M_1 multigroupe se ramène au modèle M_1 gris. Comme le modèle M_1 gris, ce système est hyperbolique. Il assure également la propriété fondamentale de limitation du flux. La difficulté du modèle provient du calcul numérique de χ_q et des coefficients α_q,β_q. En effet, ils nécessitent l'inversion de fonctions non linéaires parfois particulièrement raides. Dans la partie 3.3, on développe une procédure de précalculs de ces coefficients.

1.4.2 Le modèle P_N (d'harmoniques sphériques)

Ce modèle est un des plus anciens pour approcher l'équation du transfert radiatif (voir par exemple [75, 51]). Il permet d'obtenir une approximation d'ordre N aussi élevé que nécessaire. Pour des raisons de simplicité, on se restreint à la dimension un d'espace et l'opacité σ_ν est supposée constante égale à σ. L'équation du transfert radiatif considérée est donc :

$$\frac{1}{c}\partial_t I+\mu\partial_x I=\sigma\left(B_\nu(T)-I\right), \tag{1.32}$$

où μ est la projection de la direction Ω sur l'axe des abscisses.
L'idée des harmoniques sphériques est de décomposer l'intensité radiative dans la base des polynômes de Legendre pour la variable angulaire :

$$I(\mu)=\sum_{n=0}^{\infty}\frac{2n+1}{2}I_n^{P_N}P_n(\mu),$$

où P_n est le n^{ième} polynôme de Legendre. On rappelle que les polynômes de Legendre forment une base orthogonale de $\mathbb{R}_2[X]$ pour le produit scalaire usuel :

$$\int_{-1}^{1}P_n(\mu)P_m(\mu)d\mu=\frac{2}{2n+1}\delta_{n,m}.$$

Pour obtenir une approximation d'ordre N, la série est tronquée. Ainsi, l'approximation $I^{P_N}(\mu)$ de $I(\mu)$ est donnée par :

$$I^{P_N}(\mu)=\sum_{n=0}^{N}\frac{2n+1}{2}I_n^{P_N}P_n(\mu). \tag{1.33}$$

où les coefficients $I_n^{P_N}$ sont obtenus par :

$$I_n^{P_N} = \int_{-1}^{1} I^{P_N}(\mu) P_n(\mu) d\mu.$$

En dimension supérieure, l'intensité radiative est simplement décomposée dans la base des harmoniques sphériques (voir par exemple [32, 80]).
On substitue alors cette approximation directionnelle de I dans l'ETR (1.32). Puis en multipliant l'équation par le n$^{\text{ième}}$ polynôme de Legendre et en intégrant en direction sur $[-1, 1]$, on obtient le système linéaire d'EDP de taille $N+1$ suivant :

$$\frac{1}{c}\partial_t I_n^{P_N} + \partial_x \int_{-1}^{1} \mu P_n(\mu) I_n^{P_N}(\mu) d\mu = \sigma \left(2B_\nu(T)\delta_{n,0} - I_n^{P_N}(\mu)\right) \quad 0 \leq n \leq N, \tag{1.34}$$

car $P_0(\mu) = 1$ et donc $\int_{-1}^{1} P_n(\mu) d\mu = \int_{-1}^{1} P_n(\mu) P_0(\mu) d\mu$.
En utilisant la relation entre les polynômes de Legendre suivante :

$$(2n+1)\mu P_n(\mu) = (n+1)P_{n+1}(\mu) + nP_{n-1}(\mu),$$

le système (1.34) devient :

$$\frac{1}{c}\partial_t I_n^{P_N} + \partial_x \left[\frac{n+1}{2n+1} I_{n+1}^{P_N} + \frac{n}{2n+1} I_{n-1}^{P_N}\right] = \sigma \left(2B_\nu(T)\delta_{n,0} - I_n^{P_N}(\mu)\right) \quad 0 \leq n \leq N. \tag{1.35}$$

Ce système possède $N+2$ inconnues pour $N+1$ équations. Une fermeture naturelle consiste à imposer que le $N+1^{ème}$ coefficient $I_{N+1}^{P_N}$ soit nul :

$$I_{N+1}^{P_N} = 0.$$

Finalement, ce modèle permet de réécrire l'intensité radiative comme une combinaison de ses moments. Plusieurs auteurs ont étudié ce modèle et plus particulièrement le choix du nombre N de moments (voir par exemple [38, 101]).
On peut montrer que les approches ordonnées discrètes et harmoniques sphériques sont comparables. En effet, en utilisant des quadratures de Gauss, les solutions données par le modèle S_N et le modèle P_N sont les mêmes aux points Ω_l [38].

Exemple : le modèle P_1 Ce modèle correspond au modèle P_N en prenant $N = 1$. Dans la décomposition (1.33) de l'intensité radiative, on voit apparaître exactement les deux premiers moments E_R et F_R de I. Comme $P_0(\mu) = 1$ et $P_1(\mu) = \mu$, on a $I_0^{P_N} = E_R$ et $I_1^{P_N} = F_R$. Le modèle P_1 s'écrit alors :

$$\begin{cases} \partial_t E_R + \text{div} F_R = c(\sigma^e a T^4 - \sigma^a E_R), \\ \partial_t F_R + c^2 \frac{1}{3} \nabla E_R = -c\sigma^f F_R. \end{cases}$$

Ces modèles n'assurent pas la limitation du flux car si $E_R < 0$ par exemple, le facteur d'anisotropie devient négatif : $f < 0$.

1.4.3 Modèle de diffusion à flux limité

Historiquement, les modèles de diffusion ont largement été utilisés, surtout quand le rayonnement est couplé à un autre phénomène physique. Une présentation de ces modèles est donnée dans le cadre du transfert radiatif dans [96] et dans le cadre de la neutronique dans [84].

Ce modèle est basé sur la limite asymptotique appelée hors équilibre [95] et peut se retrouver en considérant uniquement l'évolution de l'énergie radiative E_R. Le flux radiatif F_R est alors exprimé en fonction de l'énergie E_R, du gradient ∇E_R et d'autres quantités connues. On peut retrouver le modèle à partir du modèle P_1 en supposant que le flux F_R est stationnaire :

$$\sigma^f F_R = -\frac{c}{3}\nabla E_R, \tag{1.36}$$

Ainsi, en remplaçant F_R dans l'équation en énergie (1.16), on obtient le modèle de diffusion :

$$\frac{1}{c}\partial_t E_R - \nabla.\left(\frac{1}{3\sigma}\nabla E_R\right) = c\sigma\left(aT^4 - E_R\right). \tag{1.37}$$

Malgré leur faible coût de calcul et leur très bon comportement dans les régimes de diffusion, ces modèles donnent de mauvais résultats dans la limite de transport. En effet, l'hypothèse d'isotropie de la pression radiative vraie dans la limite diffusive ne l'est plus dans les régimes de transport. De plus, ils ne respectent pas la propriété fondamentale de limitation du flux : $\frac{\|F_R\|}{cE_R} \leq 1$. Pour pallier ce problème, des limiteurs de flux peuvent être ajoutés pour obtenir des modèles de diffusion dits à flux limité (voir par exemple [88]).

2

Schémas numériques pour le modèle d'ordonnées discrètes S_N

2.1 Introduction

La théorie cinétique permet de décrire les phénomènes de transport et d'interactions dans un système de particules à une échelle microscopique. Ainsi, de nombreuses applications physiques sont modélisées par des équations cinétiques, comme par exemple l'équation de Boltzmann [58, 28] qui décrit les gaz monoatomiques parfaits, l'équation de Fokker-Planck [71] qui permet d'étudier par exemple le mouvement brownien ou encore l'équation de neutronique [38] caractérisant l'évolution des neutrons dans un système (voir aussi [39, 53, 55, 85] pour d'autres exemples). Certains modèles d'approximation de ces équations, comme le modèle d'ordonnées discrètes, mènent à des systèmes hyperboliques d'équations de la forme :

$$\partial_t w + \mathrm{div}(Mw) = S(w), \tag{2.1}$$

où $w \in \mathbb{R}^n$ est le vecteur des inconnues, M est une matrice diagonale constante donc diagonalisable dans $\mathbb{R}$ et $S(w)$ est un terme source éventuellement très non linéaire. En particulier, on s'intéresse dans ce paragraphe à l'approximation en temps long du modèle d'ordonnées discrètes pour l'ETR décrit dans le paragraphe 1.3. Dans ce cas, on a les définitions suivantes :

$$I = \begin{pmatrix} I_1 \\ \vdots \\ I_N \end{pmatrix}, \quad M = \begin{pmatrix} c\Omega_1 & 0 & \dots & 0 \\ 0 & c\Omega_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c\Omega_N \end{pmatrix}, \quad S = \begin{pmatrix} c\sigma\left(B - I_1\right) \\ \vdots \\ c\sigma\left(B - I_N\right) \end{pmatrix},$$

si bien que le système s'écrit :

$$\forall\, 1 \leq l \leq N, 1 \leq q \leq Q, \quad \partial_t I_l + c\Omega_l.\nabla I_l = c\sigma\left(B - I_l\right). \tag{1.9}$$

Après un temps très long, des schémas classiques [86, 103] donnent des solutions totalement écrasées à moins d'utiliser un maillage extrêmement fin. En effet, on compare sur la figure 2.1 quelques schémas implicites classiques : le schéma d'Euler, de Gear, IMBDF2 [74] et de Crank-Nicholson pour l'équation de transport à la vitesse 1. Sur le domaine $[-1; 1]$, la donnée initiale

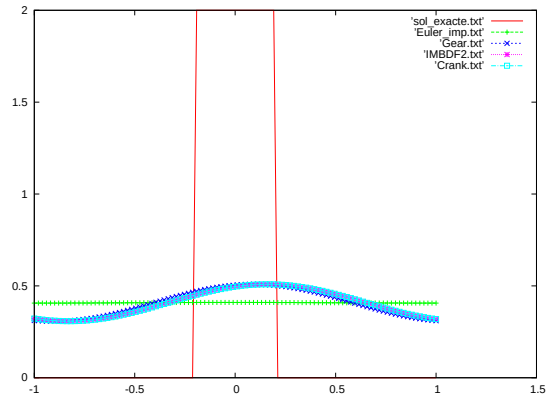


FIGURE 2.1 – Comparaison de différentes méthodes implicites à t = $20s$.

est composée d'un créneau de hauteur 2 entre $x = -0.2$ et $x = 0.2$. En considérant un maillage uniforme composé de 100 points, $\Delta t = 2\Delta x = 4.10^{-2}$ et des conditions aux limites périodiques, on regarde les solutions à $t = 20s$ (soit environ 500 itérations).

On voit bien qu'après seulement $20s$, les solutions approchées sont très aplaties.
Par ailleurs, dans le cas de couplage de plusieurs modèles, les vitesses caractéristiques peuvent être très différentes. Par exemple, l'hydrodynamique radiative fait apparaitre la vitesse de la lumière $c \simeq 3.10^8 m.s^{-1}$ comparée à la vitesse d'un fluide de l'ordre de $v \simeq 2.10^4 m.s^{-1}$. La coexistence de ces deux phénomènes caractéristiques nous oblige alors à simuler le phénomène le plus rapide sur un temps relatif au phénomène le plus lent. Pour des schémas explicites, c'est le phénomène le plus rapide qui contraint la CFL qui deviendra potentiellement très restrictive. De plus, le grand nombre de pas de temps nécessaires pour atteindre le temps final de simulation peut engendrer de la diffusion numérique et ainsi totalement dénaturer le phénomène le plus lent.
Par ailleurs, il est très avantageux d'avoir un schéma implicite précis pour les systèmes d'advection linéaire sans terme source de la forme :

$$\partial_t w + \text{div}(Mw) = 0. \tag{2.2}$$

En effet, pour discrétiser les systèmes d'équations de la forme (2.1), il existe des schémas préservant l'asymptotique qui ont pour but de conserver numériquement les régimes limites du modèle. Quelques exemples appliqués à différents systèmes, comme les équations d'Euler avec frictions ou des modèles cinétiques dérivant de l'équation de Boltzmann peuvent être trouvés dans [63, 29, 19, 79]. Ils donnent une bonne approximation des systèmes (2.1) à condition d'utiliser un schéma numérique performant pour la partie transport. Ainsi dans un premier temps, l'attention est portée sur la mise en place d'un schéma numérique d'ordre élevé pour le système de transport (2.2).

Dans la littérature, de nombreux schémas d'ordre élevé ont été développés pour les systèmes de lois de conservation hyperboliques de la forme (2.1). Certaines de ces méthodes comme les méthodes Galerkin discontinues (DG) [34, 49, 48, 50, 57], ENO/WENO [69, 89] ou les schémas distributifs [3, 1], ont été largement utilisées avec succès pour de nombreuses applications (voir aussi [78, 2, 46, 90, 35, 47, 73] pour une liste non-exhaustive). Toutes ces méthodes donnent de bons résultats sur des temps courts, soit jusqu'à environ 10 fois la vitesse caractéristique du système. Cependant, dans les applications en temps très long que l'on considère (au-delà de 10^4 fois la vitesse caractéristique), leurs conditions CFL, souvent trop restrictives, peuvent rendre l'exécution extrêmement coûteuse.

Les schémas implicites existants ne permettent pas toujours de restituer une bonne approximation de la solution (voir figure 2.1). En effet, dans ces schémas implicites, qui ont initialement été développés pour des systèmes d'équations non linéaires, c'est la précision d'approximation qui fait souvent défaut. Dans notre cas, la linéarité du système est un avantage dont on souhaite tirer partie. En effet, grâce à la méthode des caractéristiques, la solution exacte du système est facile à calculer. L'utilisation judicieuse de ces solutions exactes va permettre de proposer des schémas numériques implicites peu coûteux capables de restituer des solutions approchées extrêmement précises pour des temps physiques de simulation longs. Pour cela, on propose d'étendre des méthodes de Problème de Riemann Généralisé (GRP) [10, 59]. En effet, elles sont une alternative intéressante dans la mesure où la linéarité du système permet d'écrire des schémas d'ordre arbitrairement élevé. Grâce à une étape d'évolution exacte, il n'y a pas de perte d'information et donc pas d'introduction de diffusion numérique. En cumulant les avantages de cette méthode et en introduisant une nouvelle étape de projection de la solution, une méthode d'ordre élevé sans restriction CFL est développée. Ce schéma a notamment fait l'objet d'une publication [18].

Dans une première partie, on rappelle le principe des méthodes GRP. Puis, on développe une discrétisation du système linéaire (2.2) reposant sur un schéma de type GRP d'ordre élevé mais pour lequel aucune restriction sur le pas de temps n'est imposée. Pour simplifier la présentation, on se place en dimension un d'espace, ramenant ainsi le système (2.2) à une équation de transport. Une extension en dimension deux de ce schéma pour des maillages non structurés est proposée dans la partie suivante. Après validation de la méthode, on l'étend au modèle d'ordonnées discrètes S_N du transfert radiatif avec terme source. Finalement, dans le but de comparer les résultats de ce schéma avec d'autres méthodes numériques, la dernière partie est dédiée au rappel de certaines méthodes comme le schéma décentré amont pour le modèle d'ordonnées discrètes du transfert radiatif et les schémas DG et WENO pour l'équation d'advection.

2.2 Introduction sur les méthodes de Problème de Riemann Généralisé (GRP) en dimension 1

Les solveurs de type GRP (Generalized Riemann Problem) ont été développés par Ben-Artzi et Falcovitz dans [8, 6] et plus récemment résumés dans [10]. De nombreuses applications peuvent être trouvées par exemple dans [7, 11, 91] notamment appliquées aux équations d'Euler ou dans [9] pour une revue de différentes extensions et applications de cette méthode.
Dans ce paragraphe, on s'intéresse à l'équation d'advection en dimension un d'espace. Alors que les schémas volumes finis d'ordre élevé classiques tels que les schémas MUSCL ou WENO considèrent une approximation constante par morceaux et une étape de reconstruction de la solution, l'idée de ces méthodes est de considérer une approximation polynomiale par maille de la solution supprimant ainsi l'étape de reconstruction. La première étape consiste à faire évoluer de manière exacte l'approximation polynomiale par morceaux de la solution. Il faut alors résoudre sur chaque interface un problème de Riemann généralisé où les deux états initiaux (gauche et droit) sont des polynômes. La deuxième étape consiste à réaliser une projection L^2 de cette solution dans l'espace de polynômes considéré. Ce choix d'approximation permet ainsi d'obtenir un schéma d'ordre aussi élevé que le degré des polynômes.
Soulignons que, lorsque l'équation approchée est linéaire, la solution exacte est connue et les calculs deviennent alors entièrement explicites.

On considère ici le problème mixte pour l'équation d'advection linéaire :

$$\begin{cases} \partial_t u(x,t) + a\partial_x u(x,t) = 0, \\ u(x,0) = u_0(x), \\ u(x=0,t) = u_L(t), \end{cases} \tag{2.3}$$

où $(x,t) \in \mathbb{R}^+ \times \mathbb{R}^+$, $u(x,t) \in \mathbb{R}$ et la vitesse a est supposée positive. Les calculs pour le cas $a < 0$ où $(x,t) \in \mathbb{R}^- \times \mathbb{R}^+$ avec une condition de bord sur la droite de domaine étant similaires, ils ne seront pas détaillés.
On se donne un maillage uniforme composé des cellules $C_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ en notant $x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} = \Delta x$. Les centres des mailles sont notés x_i. On définit alors un ensemble de fonctions polynomiales par maille :

$$\mathbb{P}_C^k = \{u \in L^\infty(\mathbb{R}) / \; u_{|C_i} \in \mathbb{R}_k[X], \forall i \in \mathbb{Z}\}, \tag{2.4}$$

où $\mathbb{R}_k[X]$ est l'ensemble des polynômes de degré inférieur ou égal à k. Afin de simplifier les calculs par la suite, dans chaque cellule C_i, l'ensemble $\mathbb{R}_k[X]$ est engendré sur la cellule C_i par la base de polynômes définie par :

$$\left\{ \left(\frac{x - x_i}{\Delta x} \right)^j \right\}_{j=0,\dots,k}. \tag{2.5}$$

Sur cet espace, on introduit le produit scalaire suivant :

$$< f, g >_i = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(x) g(x) \, dx,$$

ainsi que la norme correspondante :

$$\|f\|_i = \sqrt{< f, f >_i}.$$

Une fois les espaces de polynômes définis, on peut détailler les différentes étapes du schéma.

Approximation considérée et notations : Soit $u^0 \in \mathbb{P}_C^k$, la projection de l'état initial $u_0(x)$ sur $\mathbb{P}_C^k$. À l'instant $t = t^n$, on suppose connue une approximation polynomiale de degré k par maille de la solution de (2.3) notée $u^n(x) \in \mathbb{P}_C^k$ telle que :

$$\text{pour } t = t^n, u(x,t^n) \simeq u_i^n(x) \in \mathbb{R}_k[X] \text{ si } x \in C_i.$$

Étant donnée la base de polynôme considérée, on peut écrire :

$$u_i^n(x) = \sum_{j=0}^{k} \alpha_i^{n,j} \left(\frac{x - x_i}{\Delta x} \right)^j, \tag{2.6}$$

où les $\alpha_i^{n,j}$ sont les coordonnées de u_i^n relativement à la base (2.5). Pour définir une approximation $u^{n+1}(x) \in \mathbb{P}_C^k$ de la solution au temps $t = t^n + \Delta t$, on procède à une étape d'évolution.

Étape d'évolution : Sachant que la solution exacte du problème mixte suivant :

$$\begin{cases} \partial_t u(x,t) + a\partial_x u(x,t) = 0, \\ u(x,t^n) = u_i^n(x) \quad \text{pour } x \in C_i \\ u(x=0,t) = u_L(t), \end{cases}$$

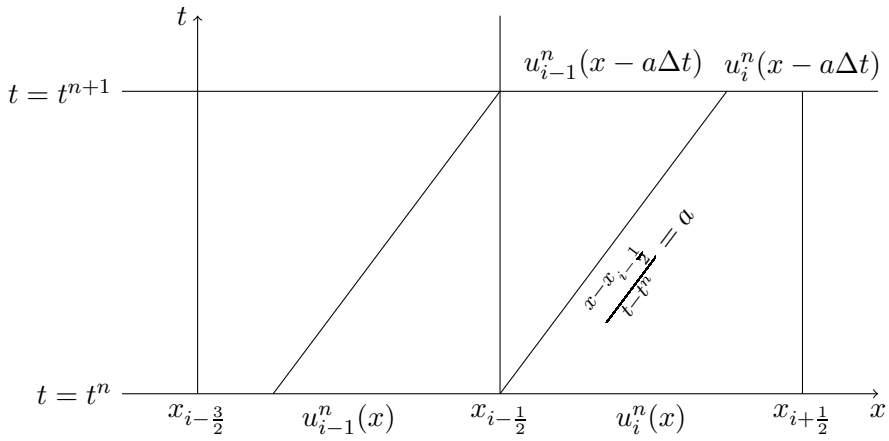


FIGURE 2.2 – Solution exacte obtenue par la méthode des caractéristiques avec $a > 0$

est connue, on fait évoluer de manière exacte la solution $u^n(x)$ pendant un pas de temps Δt. Pour le moment, on suppose que les courbes caractéristiques issues de $x_{i-\frac{1}{2}}$ ne peut pas parcourir plus d'une cellule au cours d'un pas de temps Δt. En conséquence, on impose la condition CFL suivante :

$$\lambda = a\frac{\Delta t}{\Delta x} < 1. \tag{2.7}$$

Notons dès à présent que l'on montrera plus tard comment s'affranchir simplement de cette restriction.

Pour faire évoluer la solution de manière exacte, on utilise la méthode des caractéristiques. Alors que cette étape peut s'avérer inenvisageable dans le cas d'équations non linéaires, la linéarité de l'équation (2.3) permet ici de calculer facilement la solution exacte u^h au temps $t + \Delta t$ pour une solution polynomiale par morceaux au temps t^n (voir figure 2.2). Sur chaque cellule C_i, en remontant les droites caractéristiques d'équations $t = \frac{x - x_{i-\frac{1}{2}}}{a}$, la solution exacte u^h est donnée par :

$$u_i^h(x, t^n + \Delta t) = \begin{cases} u_{i-1}^n(x - a\Delta t), & \text{si } \dfrac{x - x_{i-\frac{1}{2}}}{\Delta t} < a, \\ u_i^n(x - a\Delta t), & \text{si } \dfrac{x - x_{i-\frac{1}{2}}}{\Delta t} > a, \end{cases} \quad \forall x \in C_i.$$

Cette solution est alors projetée sur l'espace d'approximation $\mathbb{P}_C^k$.

Étape de projection : La mise à jour à la date $t^n + \Delta t$ sur la cellule C_i est définie comme la solution du problème de minimisation suivant :

$$\|u_i^{n+1} - u_i^h(t^n + \Delta t)\|_i = \inf_{p \in \mathbb{P}_C^k} \|p - u_i^h(t^n + \Delta t)\|_i. \tag{2.8}$$

Les conditions de Petrov-Galerkin permettent d'écrire ce problème de minimisation (2.8) sous la forme :

$$< u_i^{n+1} - u_i^h(t^n + \Delta t), p >_i = 0 \quad \forall p \in \mathbb{P}_C^k.$$

Puisque l'espace $\mathbb{P}_C^k$ est engendré par la base (2.5), l'approximation cherchée $u_i^{n+1}(x)$ est alors solution du système linéaire suivant :

$$\forall l = 0, \ldots, k, \quad < u_i^{n+1} - u_i^h(t^n + \Delta t), \left(\frac{x - x_i}{\Delta x}\right)^l >_i = 0. \tag{2.9}$$

En remplaçant u_i^{n+1} par sa décomposition dans cette base de polynômes (2.6), l'étape de projection revient à résoudre un système linéaire de taille $k+1$:

$$\forall l = 0, \dots, k, \quad \sum_{j=0}^{k} \alpha_i^{n+1,j} < \left(\frac{x-x_i}{\Delta x}\right)^j, \left(\frac{x-x_i}{\Delta x}\right)^l >_i = < u_i^h(t^n + \Delta t), \left(\frac{x-x_i}{\Delta x}\right)^l >_i, \tag{2.10}$$

d'inconnues $(\alpha_i^{n+1,j})_{j=0}^k$. On voit facilement que ce schéma est d'ordre $k+1$ en espace par construction.
À présent, une formulation plus explicite de ce schéma est exhibée. En effet, le membre de gauche du système (2.10) peut se réécrire sous la forme d'un produit matrice-vecteur de la manière suivante :

$$\begin{aligned}\sum_{j=0}^{k} \alpha_i^{n+1,j} < \left(\frac{x-x_i}{\Delta x}\right)^j, \left(\frac{x-x_i}{\Delta x}\right)^l >_i &= \sum_{j=0}^{k} \alpha_i^{n+1,j} \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \left(\frac{x-x_i}{\Delta x}\right)^{j+l} dx, \\ &= \sum_{j=0}^{k} \alpha_i^{n+1,j} \frac{1}{j+l+1} \left(\left(\frac{1}{2}\right)^{j+l+1} - \left(\frac{-1}{2}\right)^{j+l+1} \right).\end{aligned}$$

En posant $A(j,l) = \frac{1}{j+l+1} \left(\left(\frac{1}{2}\right)^{j+l+1} - \left(\frac{-1}{2}\right)^{j+l+1} \right)$ on a :

$$\sum_{j=0}^{k} \alpha_i^{n+1,j} < \left(\frac{x-x_i}{\Delta x}\right)^j, \left(\frac{x-x_i}{\Delta x}\right)^l >_i = \sum_{j=0}^{k} \alpha_i^{n+1,j} A(j,l).$$

Le second membre du système (2.10) peut également se réécrire de la façon suivante :

$$\begin{aligned}< u_i^h(t^n + \Delta t), \left(\frac{x-x_i}{\Delta x}\right)^l >_i &= \sum_{j=0}^{k} \alpha_{i-1}^{n,j} \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i-\frac{1}{2}}+a\Delta t} \left(\frac{x-a\Delta t-x_{i-1}}{\Delta x}\right)^j \left(\frac{x-x_i}{\Delta x}\right)^l dx \\ &+ \sum_{j=0}^{k} \alpha_i^{n,j} \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}+a\Delta t}^{x_{i+\frac{1}{2}}} \left(\frac{x-a\Delta t-x_i}{\Delta x}\right)^j \left(\frac{x-x_i}{\Delta x}\right)^l dx, \\ &= \sum_{j=0}^{k} \left(\alpha_{i-1}^{n,j} I_i^-(j,l) + \alpha_i^{n,j} I_i^+(j,l) \right), \\ &= b_i^{n+1}(l).\end{aligned}$$

Les coefficients $I_i^-(j,l)$ et $I_i^+(j,l)$ peuvent être calculés récursivement à l'aide d'intégrations par parties successives :

$$\begin{aligned}&I_i^-(j,l) = \frac{1}{j+1} \left(\frac{1}{2}\right)^{j+l+1} \left[(-1+2\lambda)^l - (1-2\lambda)^{j+1}\right] - \frac{l}{j+1} I_i^-(j+1,l-1), \\ &\text{avec } I_i^-(j+l,0) = \frac{1}{j+l+1} \left(\frac{1}{2}\right)^{j+l+1} \left[1-(1-2\lambda)^{j+l+1}\right], \\ &I_i^+(j,l) = \frac{1}{j+1} \left(\frac{1}{2}\right)^{j+l+1} \left[(1-2\lambda)^{j+1} - (-1+2\lambda)^l\right] - \frac{l}{j+1} I_i^+(j+1,l-1), \\ &\text{avec } I_i^+(j+l,0) = \frac{1}{j+l+1} \left(\frac{1}{2}\right)^{j+l+1} \left[(1-2\lambda)^{j+l+1} - (-1)^{j+l+1}\right],\end{aligned}$$

où λ est donné par (2.7).
Finalement, à chaque itération, le schéma consiste à résoudre un système linéaire de taille $k+1$ sur chaque cellule C_i afin de déterminer l'approximation polynomiale par morceaux de la solution au temps t^{n+1} :

$$A\alpha_i^{n+1} = b_i^{n+1}. \tag{2.11}$$

On remarque que le choix du maillage uniforme permet d'obtenir une unique matrice A pour tout le système. En effet, celle-ci ne dépend ni de la cellule ni du temps. Ainsi, elle ne sera calculée et inversée qu'une seule fois pour tout le domaine et tous les pas de temps. Une deuxième remarque concerne le choix de base de l'espace $\mathbb{P}_C^k$. La base utilisée ci-dessus a été choisie pour simplifier les calculs. Cependant, d'autres bases polynomiales auraient pu être considérées comme les polynômes de Legendre par exemple.

Pour les applications en temps très long considérées, l'objectif est de pouvoir considérer de très grands pas de temps qui ne respectent pas (2.7). En partant de la même approximation, on pourrait construire un schéma d'ordre $k+1$ sans aucune restriction CFL, donc pour tout $\lambda > 1$. Cependant, en fonction du pas de temps, il faudrait élargir le stencil. Par exemple, en posant $\Delta t = \frac{\lambda}{a}10^3$, il faudrait aller chercher l'information sur les 10^3 cellules voisines au temps t^n pour calculer la solution exacte au temps t^{n+1} sur la cellule C_i. Les conditions aux limites pourraient devenir plus complexes à gérer, voire inabordables. En conséquence, le coût de calcul serait très largement augmenté. De plus, l'extension directe de ce schéma en dimension deux sur des maillages non-structurés par exemple deviendrait très compliquée.

2.3 Méthode de projection GRP espace-temps pour l'équation de transport en dimension 1

Le schéma GRP classique présenté dans la section précédente permet d'obtenir une approximation d'ordre élevé de l'équation de transport, mais seulement en espace. Le but étant de capturer les solutions en temps très long, il est préférable que le schéma soit aussi d'ordre élevé en temps. Les méthodes de Runge-Kutta par exemple, donnent une approximation d'ordre élevé en temps, mais en considérant de grands pas de temps, elles impliquent la résolution de systèmes non-linéaires de taille croissante avec l'ordre d'approximation. On souhaite ici développer un schéma sans restriction sur le pas de temps, qui soit d'ordre élevé en espace et en temps, tout en restant très compact pour éviter la gestion de larges stencils. Pour cela, on va développer une extension du solveur GRP qui n'impose plus de restriction sur le pas de temps.

Afin de simplifier la terminologie, on emploiera abusivement le terme *implicite* en référence au cas $\lambda = a\frac{\Delta t}{\Delta x} > 1$, bien que le schéma soit différent des schémas implicites classiques qui font intervenir la recherche des zéros d'une fonction. Parallèlement, on emploiera le terme *explicite* pour se référer au cas $\lambda \leq 1$.

On considère toujours le problème mixte pour l'équation d'advection :

$$\begin{cases} \partial_t u + a\partial_x u = 0, \\ u(x,0) = u_0(x), \\ u(x=0,t) = u_L(t). \end{cases} \tag{2.12}$$

où $a > 0$.
Comme précédemment, le domaine est discrétisé par un maillage uniforme formé des cellules $C_i = (x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$ avec $x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} = \Delta x$. On introduit en plus $I^{n+\frac{1}{2}} = [t^n, t^{n+1}]$ les interfaces entre deux pas de temps. Les volumes de contrôle espace-temps sont ainsi délimités par : $K_i^n = C_i \times I^{n+\frac{1}{2}}$.

Comme dans le schéma GRP, le schéma consiste à profiter de la linéarité de l'équation pour introduire une étape d'évolution exacte. La différence provient de l'introduction d'une approximation auxiliaire polynomiale en temps de la solution. Cette approximation, notée $v(t)$, sera utilisée comme un outil pour simplifier la présentation de ce schéma.
Le schéma se composera des étapes suivantes, qui seront détaillées dans la suite :

- Au temps initial, on suppose connue une approximation polynomiale en espace $u_i^0(x)$ de la solution sur chaque cellule C_i. On suppose également connue une seconde approximation auxiliaire polynomiale en temps $v_{\frac{1}{2}}(t)$ de la donnée $u_L(t)$ solution sur la frontière du domaine correspondante.
- Au cours d'une itération, le maillage est parcouru de la gauche vers la droite avec $a > 0$ (de la droite vers la gauche sinon).
- Dans chaque volume de contrôle K_i^n, en supposant $a > 0$, on calcule l'approximation $u_i^{n+1}(x)$ pour $x \in C_i$ ainsi que l'approximation auxiliaire sur le bord droit du volume de contrôle $v_{i+\frac{1}{2}}(t)$ pour $t \in I^{n+\frac{1}{2}}$ à partir des solutions $u_i^n(x)$ pour $x \in C_i$ et $v_{i-\frac{1}{2}}(t)$ sur le bord gauche de K_i^n pour $t \in I^{n+\frac{1}{2}}$.

Approximations considérées et notations : On suppose toujours connue une approximation polynomiale en espace de degré k de la solution :

$$\text{pour } t = t^n, u(x, t^n) \simeq u_i^n(x) \in \mathbb{R}_k[X] \text{ si } x \in C_i.$$

On introduit une seconde approximation temporelle auxiliaire de la solution qui sera utilisée comme un outil dans la construction du schéma d'ordre élevé en temps. Ainsi, on définit :

$$\text{pour } t \in I^{n+\frac{1}{2}} = [t^n; t^n + \Delta t], u(x_{i-\frac{1}{2}}, t) \simeq v_{i-\frac{1}{2}}^{n+\frac{1}{2}}(t) \in \mathbb{R}_k[X]. \tag{2.13}$$

On note alors l'espace des fonctions polynomiales par interface :

$$\mathbb{P}_I^k = \{v \in L^\infty(\mathbb{R}) / \ v_{|I^{n+\frac{1}{2}}} \in \mathbb{R}_k[X], \forall n \in \mathbb{N}\}.$$

Puis, sur chaque interface $I^{n+\frac{1}{2}}$, on munit $\mathbb{R}_k[X]$ de la base :

$$\left\{ \left(\frac{t - t^{n+\frac{1}{2}}}{\Delta t} \right)^r \right\}_{r=0,\dots,k},$$

sachant que d'autres bases peuvent être choisies afin de s'adapter au cas considéré. On introduit également le produit scalaire associé à l'interface $I^{n+\frac{1}{2}}$ défini par :

$$< f, g >^n = \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} f(t) g(t) \, dt.$$

Une fois ces notations établies, on fait évoluer la solution approchée $u_i^n(x)$ ainsi que l'approximation temporelle $v_{i-\frac{1}{2}}^{n+\frac{1}{2}}(t)$ de manière exacte pour obtenir une approximation de la solution au temps $t^n + \Delta t$.

Étape d'évolution : À cette étape, on suppose donc que l'on doit résoudre le problème mixte suivant :

$$\begin{cases} \partial_t u + a \partial_x u = 0, \\ u(x, t^n) = u^n(x), \\ u(0, t) = u_L(t). \end{cases} \tag{2.14}$$

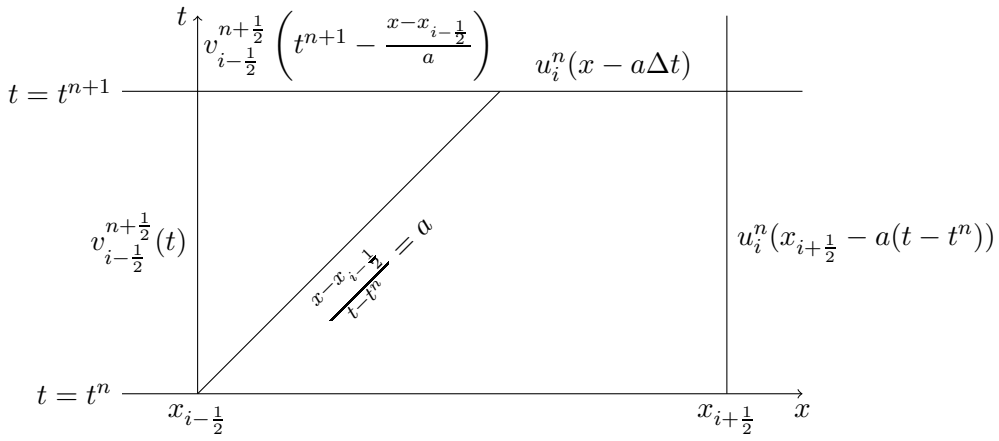


FIGURE 2.3 – Solutions exactes dans le volume K_i^n pour le cas explicite

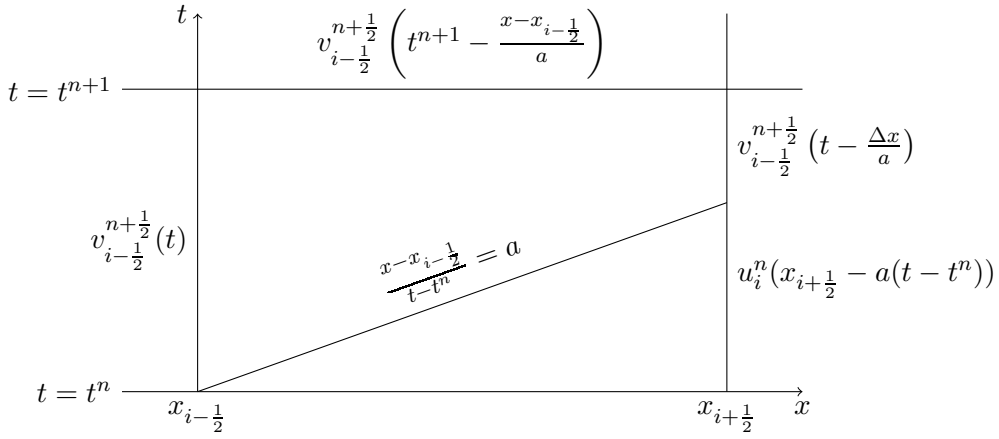


FIGURE 2.4 – Solutions exactes dans le volume K_i^n pour le cas implicite

La principale différence avec le schéma GRP présenté dans la section (2.2) est la double approximation de la solution : une en espace sur les cellules et une en temps sur les interfaces. On insiste ici sur le fait que la solution temporelle est utilisée comme un outil simplifiant la présentation du schéma. Dans le cas $a > 0$, la solution u_i^{n+1} au temps $t = t^{n+1}$ sur la cellule C_i et la solution auxiliaire $v_{i+\frac{1}{2}}^{n+\frac{1}{2}}$ sur l'interface $I^{n+\frac{1}{2}}$ en $x = x_{i+\frac{1}{2}}$ sont entièrement déterminées par la donnée de u_i^n et $v_{i-\frac{1}{2}}^{n+\frac{1}{2}}$ (voir figures 2.3-2.4). Notons que dans le cas $a < 0$, la solution u_i^{n+1} au temps $t = t^{n+1}$ sur C_i et $v_{i-\frac{1}{2}}^{n+\frac{1}{2}}$ sur l'interface $I^{n+\frac{1}{2}}$ en $x = x_{i-\frac{1}{2}}$ sont entièrement déterminées par la donnée de u_i^n et $v_{i+\frac{1}{2}}^{n+\frac{1}{2}}$.
Il en résulte que les informations nécessaires au calcul de la solution au temps $t = t^{n+1}$ sont entièrement contenues dans un seul volume de contrôle. Le schéma obtenu sera donc compact. De plus, grâce à la linéarité de l'équation (2.12), les solutions exactes au temps $t^n+\Delta t$ sont facilement déterminées par une méthode des caractéristiques.

Dans le cas $a > 0$ et $\lambda = \frac{a\Delta t}{\Delta x} < 1$ (cas explicite), la solution exacte du problème (2.14) sur la

cellule C_i est donnée par :

$$u_i^h(x,t^n+\Delta t)=\begin{cases} v_{i-\frac{1}{2}}^{n+\frac{1}{2}}\left(t^{n+1}-\frac{x-x_{i-\frac{1}{2}}}{a}\right), & \text{si } (x-x_{i-\frac{1}{2}})<a\Delta t,\\ u_i^n(x-a\Delta t), & \text{sinon,}\end{cases}$$

et la solution exacte en temps sur l'interface $I^{n+\frac{1}{2}}$ est donnée par :

$$v_h^{n+\frac{1}{2}}(t,x_{i+\frac{1}{2}})= u_i^n(x_{i+\frac{1}{2}}-a(t-t^n)).$$

Pour $\lambda>1$ (cas implicite), on a :

$$u_i^h(x,t^n+\Delta t)=v_{i-\frac{1}{2}}^{n+\frac{1}{2}}\left(t^{n+1}-\frac{x-x_{i-\frac{1}{2}}}{a}\right), \quad \text{pour } x\in C_i$$

$$v_h^{n+\frac{1}{2}}(t,x_{i+\frac{1}{2}})=\begin{cases} u_i^n(x_{i+\frac{1}{2}}-a(t-t^n)), & \text{si } (t-t^n)>\frac{\Delta x}{a}\\ v_{i-\frac{1}{2}}^{n+\frac{1}{2}}\left(t-\frac{\Delta x}{a}\right), & \text{sinon,}\end{cases} \quad \text{pour } t\in I^{n+\frac{1}{2}}.$$

Les solutions exactes u^h et v_h étant connues, il reste alors à les projeter dans leurs espaces d'approximation respectifs.

Étape de projection : Les solutions exactes sont projetées dans les espaces polynomiaux : $\mathbb{P}_C^k$ et $\mathbb{P}_I^k$. La solution exacte en espace $u^h(x)$ est projetée sur l'espace $\mathbb{P}_C^k$ sur la cellule C_i au temps t^{n+1} et la solution exacte en temps $v_h(t)$ est projetée sur l'espace $\mathbb{P}_I^k$ sur l'interface $x_{i+\frac{1}{2}}\times(t^n,t^{n+1})$. Ces projections mènent à deux problèmes de minimisation :

$$\|u_i^{n+1}-u_i^h(t^n+\Delta t)\|_i=\inf_{p\in\mathbb{P}_C^k}\|p-u_i^h(t^n+\Delta t)\|_i, \tag{2.15}$$

$$\|v_{i+\frac{1}{2}}^{n+\frac{1}{2}}-v_h^{n+\frac{1}{2}}(x_{i+\frac{1}{2}})\|^n=\inf_{q\in\mathbb{P}_I^k}\|q-v_h^{n+\frac{1}{2}}(x_{i+\frac{1}{2}})\|^n. \tag{2.16}$$

Les conditions de Petrov-Galerkin permettent de réécrire ces problèmes de minimisation (2.15)-(2.16) sous la forme :

$$\forall l=0,\dots,k,\quad <u_i^{n+1}-u_i^h(t^n+\Delta t),\left(\frac{x-x_i}{\Delta x}\right)^l>_i=0, \tag{2.17}$$

$$\forall r=0,\dots,k,\quad <v_{i+\frac{1}{2}}^{n+\frac{1}{2}}-v_h^{n+\frac{1}{2}}(x_{i+\frac{1}{2}}),\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r>^n=0. \tag{2.18}$$

Puis en remplaçant $u_i^{n+1}(x)$ et $v_{i+\frac{1}{2}}^{n+\frac{1}{2}}(t)$ par leur décomposition dans leur base de polynômes respectives :

$$u_i^{n+1}(x)=\sum_{j=0}^k\alpha_i^{n+1,j}\left(\frac{x-x_i}{\Delta x}\right)^j,$$

$$v_{i+\frac{1}{2}}^{n+\frac{1}{2}}(t)=\sum_{s=0}^k\beta_{i+\frac{1}{2}}^{n+\frac{1}{2},s}\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^s,$$

les problèmes de minimisation (2.17)-(2.18) reviennent à résoudre les deux systèmes linéaires de taille $k+1$ suivants :

$$\forall l = 0, \dots, k, \quad \sum_{j=0}^{k} \alpha_i^{n+1,j} < \left(\frac{x-x_i}{\Delta x}\right)^j, \left(\frac{x-x_i}{\Delta x}\right)^l >_i = < u_i^h(t^n + \Delta t), \left(\frac{x-x_i}{\Delta x}\right)^l >_i, \tag{2.19}$$

$$\forall r = 0, \dots, k, \quad \sum_{s=0}^{k} \beta_{i+\frac{1}{2}}^{n+\frac{1}{2},s} < \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^s, \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r >^n = < v_h^{n+\frac{1}{2}}(x_{i+\frac{1}{2}}), \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r >^n . \tag{2.20}$$

Les produits scalaires des membres de gauche dans les systèmes (2.19)-(2.20) peuvent se réécrire :

$$\begin{aligned} < \left(\frac{x-x_i}{\Delta x}\right)^j, \left(\frac{x-x_i}{\Delta x}\right)^l >_i &= \frac{1}{j+l+1}\left(\frac{1}{2}\right)^{j+l+1}\left[1-(-1)^{j+l+1}\right], \\ &= A_u(j,l), \\ < \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^s, \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r >^n &= \frac{1}{s+r+1}\left(\frac{1}{2}\right)^{s+r+1}\left[1-(-1)^{s+r+1}\right], \\ &= A_v(s,r). \end{aligned}$$

On remarque que dans le cas particulier des maillages uniformes, les deux matrices A_u et A_v sont égales et indépendantes de λ. Toutefois, cette propriété serait perdue pour un maillage non uniforme.
En fonction de la valeur de λ, les seconds membres des systèmes (2.19)-(2.20) auront une expression différente. Avec $a > 0$, dans le cas explicite $\lambda < 1$, les seconds membres sont donnés par :

$$\begin{aligned} < u_i^h(t^n+\Delta t), \left(\frac{x-x_i}{\Delta x}\right)^l >_i &= \sum_{j=0}^{k} \beta_{i-\frac{1}{2}}^{n+\frac{1}{2},j} \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i-\frac{1}{2}}+a\Delta t} \left(\frac{t^{n+1} - \frac{x-x_{i-\frac{1}{2}}}{a} - t^{n+\frac{1}{2}}}{\Delta t}\right)^j \left(\frac{x-x_i}{\Delta x}\right)^l dx, \\ &+ \sum_{j=0}^{k} \alpha_i^{n,j} \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}+a\Delta t}^{x_{i+\frac{1}{2}}} \left(\frac{x-a\Delta t-x_i}{\Delta x}\right)^j \left(\frac{x-x_i}{\Delta x}\right)^l dx, \\ &= \sum_{j=0}^{k} \beta_{i-\frac{1}{2}}^{n+\frac{1}{2},j} J_{ex}^1(j,l) + \alpha_i^{n,j} J_{ex}^2(j,l), \\ &= b_u^+(l), \\ < v_h^{n+\frac{1}{2}}(x_{i+\frac{1}{2}}), \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r >^n &= \sum_{s=0}^{k} \alpha_i^{n,s} \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \left(\frac{1}{2} - \frac{a(t-t^n)}{\Delta x}\right)^s \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r dt, \\ &= \sum_{s=0}^{k} \alpha_i^{n,s} K_{ex}(s,r), \\ &= b_v^+(r), \end{aligned}$$

où les coefficients J_{ex}^1, J_{ex}^2 et K_{ex} peuvent être calculés récursivement grâce à des intégrations par

parties successives :

$$
\begin{aligned}
J^1_{ex}(j,l) =& \frac{1}{l+1}\left[\left(\frac{-1}{2}\right)^j\left(\lambda-\frac{1}{2}\right)^{l+1}-\left(\frac{1}{2}\right)^j\left(\frac{-1}{2}\right)^{l+1}\right]+\frac{j}{\lambda(l+1)}J^1_{ex}(j-1,l+1),\\
J^2_{ex}(j,l) =& \frac{1}{l+1}\left[\left(\frac{1}{2}\right)^{l+1}\left(\frac{1}{2}-\lambda\right)^j-\left(\lambda-\frac{1}{2}\right)^{l+1}\left(\frac{-1}{2}\right)^j\right]-\frac{j}{l+1}J^2_{ex}(j-1,l+1),\\
K_{ex}(s,r) =& \frac{1}{r+1}\left[\left(\frac{1}{2}-\lambda\right)^s\left(\frac{1}{2}\right)^{r+1}-\left(\frac{1}{2}\right)^s\left(\frac{-1}{2}\right)^{r+1}\right]+\frac{s\lambda}{r+1}K_{ex}(s-1,r+1).
\end{aligned}
$$

Dans le cas implicite ($\lambda > 1$), les seconds membres s'écrivent :

$$
\begin{aligned}
< u_i^h(t^n+\Delta t),\left(\frac{x-x_i}{\Delta x}\right)^l >_i =& \sum_{j=0}^k \beta^{n+\frac12,j}_{i-\frac12}\frac{1}{\Delta x}\int_{x_{i-\frac12}}^{x_{i+\frac12}}\left(\frac{t^{n+1}-\frac{x-x_{i-\frac12}}{a}-t^{n+\frac12}}{\Delta t}\right)^j\left(\frac{x-x_i}{\Delta x}\right)^l dx,\\
=& \sum_{j=0}^k \beta^{n+\frac12,j}_{i-\frac12} J_{im}(j,l),\\
=& b_u^+(l),\\
< v_h^{n+\frac12}(x_{i+\frac12}),\left(\frac{t-t^{n+\frac12}}{\Delta t}\right)^r >^n =& \sum_{s=0}^k \alpha_i^{n,s}\frac{1}{\Delta t}\int_{t^n}^{t^n+\frac{\Delta x}{a}}\left(\frac{x_{i+\frac12}-a(t-t^n)-x_i}{\Delta x}\right)^s\left(\frac{t-t^{n+\frac12}}{\Delta t}\right)^r dt,\\
&+\sum_{s=0}^k \beta^{n+\frac12,s}_{i-\frac12}\frac{1}{\Delta t}\int_{t^n+\frac{\Delta x}{a}}^{t^{n+1}}\left(\frac{t-\frac{\Delta x}{a}-t^{n+\frac12}}{\Delta t}\right)^s\left(\frac{t-t^{n+\frac12}}{\Delta t}\right)^r dt,\\
=& \sum_{s=0}^k \alpha_i^{n,s}K^1_{im}(s,r)+\beta^{n+\frac12,s}_{i-\frac12}K^2_{im}(s,r),\\
=& b_v^+(r),
\end{aligned}
$$

où les coefficients J_{im}, K^1_{im} et K^2_{im} peuvent également être calculés récursivement :

$$
\begin{aligned}
J_{im}(j,l) =& \frac{1}{l+1}\left[\left(\frac12-\frac{1}{\lambda}\right)^j\left(\frac12\right)^{l+1}-\left(\frac12\right)^j\left(\frac{-1}{2}\right)^{l+1}\right]+\frac{j}{\lambda(l+1)}J_{im}(j-1,l+1),\\
K^1_{im}(s,r) =& \frac{1}{r+1}\left[\left(\frac{-1}{2}\right)^s\left(\frac{1}{\lambda}-\frac12\right)^{r+1}-\left(\frac12\right)^s\left(\frac{-1}{2}\right)^{r+1}\right]+\frac{s\lambda}{r+1}K^1_{im}(s-1,r+1),\\
K^2_{im}(s,r) =& \frac{1}{r+1}\left[\left(\frac12-\frac{1}{\lambda}\right)^s\left(\frac12\right)^{r+1}-\left(\frac{-1}{2}\right)^s\left(\frac{1}{\lambda}-\frac12\right)^{r+1}\right]+\frac{s}{r+1}K^2_{im}(s-1,r+1).
\end{aligned}
$$

Afin de présenter complètement la méthode, notons que si $a < 0$, les solutions exactes dans le cas explicite ($\lambda < 1$) sont données par :

$$
\begin{aligned}
< u_i^h(t^n+\Delta t),\left(\frac{x-x_i}{\Delta x}\right)^l >_i =& \sum_{j=0}^k \beta^{n+\frac12,j}_{i+\frac12}J^{1,-}_{ex}(j,l)+\alpha_i^{n,j}J^{2,-}_{ex}(j,l),\\
=& b_u^-(l),\\
< v_h^{n+\frac12}(x_{i-\frac12}),\left(\frac{t-t^{n+\frac12}}{\Delta t}\right)^r >^n =& \sum_{s=0}^k \alpha_i^{n,s}K^-_{ex}(s,r),\\
=& b_v^-(r),
\end{aligned}
$$

avec :

$$J_{ex}^{1,-}(j,l) = \int_{-\frac{1}{2}}^{\frac{1}{2}+\lambda} (y-\lambda)^j y^l dy,$$
$$J_{ex}^{2,-}(j,l) = \int_{\frac{1}{2}+\lambda}^{\frac{1}{2}} \left(\frac{1}{2} + \frac{1}{\lambda}(\frac{1}{2}-y)\right)^j y^l dy,$$
$$K_{ex}^{-}(s,r) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(-\frac{1}{2} - \lambda(y+\frac{1}{2})\right)^j y^r dy,$$

et dans le cas implicite ($\lambda > 1$) par :

$$\begin{aligned} < u_i^h(t^n+\Delta t), \left(\frac{x-x_i}{\Delta x}\right)^l >_i &= \sum_{j=0}^{k} \beta_{i+\frac{1}{2}}^{n+\frac{1}{2},j} J_{im}^{-}(j,l), \\ &= b_u^-(l), \\ < v_h^{n+\frac{1}{2}}(x_{i-\frac{1}{2}}), \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r >^n &= \sum_{s=0}^{k} \alpha_i^{n,s} K_{im}^{1,-}(s,r) + \beta_{i+\frac{1}{2}}^{n+\frac{1}{2},s} K_{im}^{2,-}(s,r) \\ &= b_v^-(r), \end{aligned}$$

avec :

$$J_{im}^{-}(j,l) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\frac{1}{2} + \frac{1}{\lambda}(\frac{1}{2}-y)\right)^j y^l dy,$$
$$K_{im}^{1,-}(s,r) = \int_{-\frac{1}{2}}^{-\frac{1}{2}-\frac{1}{\lambda}} \left(-\frac{1}{2} - \lambda(y+\frac{1}{2})\right)^j y^r dy,$$
$$K_{im}^{2,-}(s,r) = \int_{-\frac{1}{2}-\frac{1}{\lambda}}^{\frac{1}{2}} \left(y + \frac{1}{\lambda}\right)^j y^r dy.$$

Finalement, on peut résumer les étapes du schéma :

- Au temps initial, la solution $u_i^0(x)$ est connue sur chaque cellule du maillage et la solution $v_{\frac{1}{2}}(t)$ est connue sur la frontière gauche du domaine si $a > 0$, ou $v_{N+\frac{1}{2}}(t)$ sur la frontière droite si $a < 0$.
- Le maillage est parcouru de la gauche vers la droite si $a > 0$, de la droite vers la gauche sinon. En effet, sur chaque volume de contrôle, il faut s'assurer que la solution temporelle est connue sur le bord gauche si $a > 0$ ou droit si $a < 0$ pour pouvoir calculer la vraie solution au temps t^{n+1}.
- Dans chaque volume K_i^n, il faut résoudre deux systèmes de taille $k+1$ pour calculer les solutions $u_i^{n+1}(x)$ et $v_{i+\frac{1}{2}}^{n+\frac{1}{2}}(t)$ si $a > 0$:

$$A_u \alpha_i^{n+1} = b_u^+, \tag{2.21}$$
$$A_v \beta_{i+\frac{1}{2}}^{n+\frac{1}{2}} = b_v^+, \tag{2.22}$$

ou $v_{i-\frac{1}{2}}^{n+\frac{1}{2}}(t)$ si $a < 0$:

$$A_u \alpha_i^{n+1} = b_u^-, \tag{2.21}$$
$$A_v \beta_{i-\frac{1}{2}}^{n+\frac{1}{2}} = b_v^-. \tag{2.22}$$

À la fin de l'itération, les deux approximations $u^{n+1}(x)$ et $v^{n+\frac{1}{2}}(t)$ de la solution de (2.12) sont par construction dans leurs espaces respectifs $\mathbb{P}_C^k$ et $\mathbb{P}_I^k$.
On note également que les matrices A_u et A_v des systèmes (2.21)-(2.22) sont égales dans ce cas particulier de maillages uniformes et ne dépendent ni du temps, ni de l'espace. Une seule matrice peut donc être calculée et inversée une fois au début de l'algorithme.

Pour illustrer ce schéma, quelques exemples sont détaillés ci-dessous.

Exemple dans le cas $a > 0$

Projection $\mathbb{P}^0$: $k = 0$. Au temps t^n, la solution $u^n(x)$ est connue sur toutes les cellules C_i et la solution temporelle $v_{\frac{1}{2}}(t)$ est connue sur la frontière gauche du domaine :

$$u_i^n(x) = \alpha_i^n \ \in \mathbb{R},$$
$$v_{\frac{1}{2}}^{n+\frac{1}{2}}(t) = \beta_{\frac{1}{2}}^{n+\frac{1}{2}} \ \in \mathbb{R}.$$

En parcourant le maillage de la gauche vers la droite, sur le volume de contrôle K_i^n, les schémas correspondants s'écrivent alors :

$$\text{Explicite } (\lambda \leq 1) : \begin{cases} v_{i+\frac{1}{2}}^{n+\frac{1}{2}} = & \alpha_i^n, \\ u_i^{n+1} = & \lambda \beta_{i-\frac{1}{2}}^{n+\frac{1}{2}} + (1-\lambda)\alpha_i^n, \end{cases}$$
$$\text{Implicite} (\lambda > 1) : \begin{cases} u_i^{n+1} = & \beta_{i-\frac{1}{2}}^{n+\frac{1}{2}}, \\ v_{i+\frac{1}{2}}^{n+\frac{1}{2}} = & \frac{1}{\lambda}\alpha_i^n + (1-\frac{1}{\lambda})\beta_{i-\frac{1}{2}}^{n+\frac{1}{2}}. \end{cases}$$

On remarque que le schéma explicite pour une approximation constante par morceaux est exactement le schéma décentré amont.

Projection $\mathbb{P}^1$: $k = 1$. Au temps t^n, la solution $u^n(x)$ est connue sur toutes les cellules C_i et la solution temporelle $v_{\frac{1}{2}}(t)$ est connue sur la frontière gauche du domaine :

$$u_i^n(x) = \alpha_i^{n,0} + \alpha_i^{n,1} x,$$
$$v_{\frac{1}{2}}^{n+\frac{1}{2}}(t) = \beta_{\frac{1}{2}}^{n+\frac{1}{2},0} + \beta_{\frac{1}{2}}^{n+\frac{1}{2},1} t.$$

En parcourant le maillage de la gauche vers la droite, le schéma explicite correspondant s'écrit alors :

$$\begin{cases} \alpha_i^{n+1,0} = & \alpha_i^{n,0}(1-\lambda) - \frac{\lambda}{2}(1-\lambda)\alpha_i^{n,1} + \lambda \beta_{i-\frac{1}{2}}^{n+\frac{1}{2},0}, \\ \alpha_i^{n+1,1} = & 6\lambda(1-\lambda)[\alpha_i^{n,0} - \beta_{i-\frac{1}{2}}^{n+\frac{1}{2},0}] + (1-\lambda)(1-2\lambda-2\lambda^2)\alpha_i^{n,1} - \lambda^2 \beta_{i-\frac{1}{2}}^{n+\frac{1}{2},1}, \\ \beta_{i+\frac{1}{2}}^{n+\frac{1}{2},0} = & \alpha_i^{n,0} + \frac{1-\lambda}{2}\alpha_i^{n,1}, \\ \beta_{i+\frac{1}{2}}^{n+\frac{1}{2},1} = & -\lambda \alpha_i^{n,0}. \end{cases}$$

Ce schéma est équivalent au schéma développé par Dai and Woodward dans [37]. Cependant leur schéma, basé sur la conservation des moments pour les équations non-linéaires, est défini uniquement dans le cas explicite.

Ce schéma est ainsi facile à écrire et facile à implémenter. De plus, on peut prouver qu'il est conservatif et d'ordre $k+1$ en temps et en espace.

Théorème 2.1. *Le schéma est conservatif pour le premier moment $\alpha^{n,0}$. Si la donnée initiale $u_0(x)$ pour $x \in \mathbb{R}$ est un polynôme de degré k et $u_L(t) = u_0(-at)$, alors le schéma préserve exactement les polynômes de degré k.*

Remarquons que ce schéma est conservatif uniquement pour le premier moment. En conséquence, s'il est utilisé dans un contexte de volumes finis, les moments d'ordre plus élevé doivent être interprétés comme des termes correctifs.

Démonstration. Par définition du schéma, il vient :

$$\begin{aligned}\sum_i \alpha_i^{n+1,0} &= \int_{\mathbb{R}} u^h(x, t^{n+1})dx, \\ &= \int_{\mathbb{R}} u^h(x - a\Delta t, t^n)dx = \sum_i \alpha_i^{n,0}.\end{aligned}$$

Le schéma est donc conservatif pour le premier moment $\alpha^{n,0}$.

À présent, on établit la précision de la méthode. Pour simplifier les calculs, la vitesse a est supposée positive. La preuve sera faite par récurrence.
Tout d'abord, on considère le schéma explicite ($\lambda < 1$). Au temps $t = 0$, on suppose que la solution initiale est polynomiale : $u_0(x) \in \mathbb{R}_k[X]$.
Au temps $t = t^n$, la solution donnée par le schéma est supposée exacte : $u^n(x) = u_0(x - at^n)$ si $x - at^n > 0$, $u^n(x) = u_L(t - \frac{x}{a})$ sinon. Donc en particulier, $u^n(x) \in \mathbb{R}_k[X]$.
Pour que le schéma soit d'ordre $k + 1$, il faut donc montrer que la solution exacte est préservée par le schéma au temps $t = t^n + \Delta t$.
Dans chaque cellule C_i du domaine au temps $t = t^n$, on a $u_i^n(x) = u^n(x)|_{C_i} \in \mathbb{R}_k[X]$. Ainsi, la projection de $u^n(x)$ dans l'espace d'approximation $\mathbb{P}_C^k$ conserve cette solution polynomiale $u^n(x)$.
Étape d'évolution : Cette étape consiste à faire évoluer de manière exacte la solution. Sur le volume K_i^n, en remontant les courbes caractéristiques, on remarque que la solution temporelle exacte sur le bord gauche $I^{n+\frac{1}{2}} \times x_{i-\frac{1}{2}}$ est :

$$v_h^{n+\frac{1}{2}}(t, x_{i-\frac{1}{2}}) = u_{i-1}^n(x_{i-\frac{1}{2}} - a(t - t^n)) \in \mathbb{R}_k[X],$$

et sur le bord droit $I^{n+\frac{1}{2}} \times x_{i+\frac{1}{2}}$:

$$v_h^{n+\frac{1}{2}}(t, x_{i+\frac{1}{2}}) = u_i^n(x_{i+\frac{1}{2}} - a(t - t^n)) \in \mathbb{R}_k[X].$$

La solution exacte dans la cellule C_i est :

$$u_i^h(x, t^n + \Delta t) = \begin{cases} v_{i-\frac{1}{2}}^{n+\frac{1}{2}}\left(t^{n+1} - \frac{x - x_{i-\frac{1}{2}}}{a}\right), & \text{si } (x - x_{i-\frac{1}{2}}) < a\Delta t, \\ u_i^n(x - a\Delta t), & \text{sinon.} \end{cases}$$

Étape de projection : Les solutions exactes sont projetées dans les espaces correspondants $\mathbb{P}^k$. Sur le bord gauche, comme $v_h^{n+\frac{1}{2}}(x_{i-\frac{1}{2}}) \in \mathbb{R}_k[X]$, la projection sur l'espace $\mathbb{P}_I^k$ est exacte. La solution temporelle sur l'interface $I^{n+\frac{1}{2}} \times x_{i-\frac{1}{2}}$ est donc la solution exacte :

$$v_{i-\frac{1}{2}}^{n+\frac{1}{2}}(t) = v_h^{n+\frac{1}{2}}(t, x_{i-\frac{1}{2}}) = u_{i-1}^n(x_{i-\frac{1}{2}} - a(t - t^n)).$$

Par conséquent :

$$u_i^h(x,t^n+\Delta t)=\begin{cases}u_{i-1}^n(x-a\Delta t), & \text{si } (x-x_{i-\frac{1}{2}})<a\Delta t,\\ u_i^n(x-a\Delta t), & \text{sinon,}\end{cases}=u_i^h(x-a\Delta t,t^n).$$

Comme $u_i^h(t^n)\in\mathbb{R}_k[X]$, on a $u_i^h(t^n+\Delta t)\in\mathbb{R}_k[X]$ et la projection sur l'espace $\mathbb{P}_C^k$ est encore une fois exacte :

$$u_i^{n+1}(x)=u_i^h(x,t^n+\Delta t)=u^n(x-a\Delta t)\big|_{C_i}.$$

À la fin de l'itération, la solution est donnée par $u^{n+1}(x)=u^n(x-a\Delta t)=u_0(x-at^{n+1})$ si $x-at^{n+1}>0$, $u^{n+1}(x)=u_L(t^{n+1}-\frac{x}{a})$ sinon. Le schéma explicite préserve bien la solution exacte si la donnée initiale $u_0\in\mathbb{R}_k[X]$.

On s'intéresse à présent au cas implicite ($\lambda>1$), toujours avec une vitesse $a>0$. Au temps $t=0$, la donnée initiale $u_0(x)$ et la solution sur la frontière de gauche $u_L(t)$ sont supposées polynomiales de degré k : $u_0(x)\in\mathbb{R}_k[X]$ et $u_L(t)\in\mathbb{R}_k[X]$.
Au temps $t=t^n$, la solution $u^n(x)$ donnée par le schéma est supposée exacte : $u^n(x)=u_0(x-at^n)\in\mathbb{R}_k[X]$ si $x-at^{n+1}>0$, $u^n(x)=u_L(t^{n+1}-\frac{x}{a}$ sinon.
On veut alors montrer que cette solution exacte est préservée par le schéma au temps $t=t^n+\Delta t$. Dans chaque cellule C_i du domaine, au temps $t=t^n$, on a $u_i^n(x)=u^n(x)\big|_{C_i}\in\mathbb{R}_k[X]$ la projection de $u^n(x)$ sur l'espace $\mathbb{P}_C^k$. Comme $u^n(x)$ est polynomiale de degré k, la projection de $\{u_i^n(x)\}_{i\in\mathbb{Z}}$ sur l'espace $\mathbb{P}_C^k$ conserve ce polynôme. On le réécrit simplement dans la base correspondante. Sur la frontière gauche, on connait la solution exacte $v_{\frac{1}{2}}^{n+\frac{1}{2}}(t)=u_L(t)\big|_{I^{n+\frac{1}{2}}}$.
On se place sur le premier volume $K_1^{n+\frac{1}{2}}$.
Étape d'évolution : La solution exacte sur la cellule C_1 est :

$$u_1^h(x,t^n+\Delta t)=\ v_{\frac{1}{2}}^{n+\frac{1}{2}}\left(t^{n+1}-\frac{x-x_{\frac{1}{2}}}{a}\right)=u_L\left(t^{n+1}-\frac{x-x_{\frac{1}{2}}}{a}\right)\Big|_{I^{n+\frac{1}{2}}},$$

et sur l'interface $x_{\frac{3}{2}}$, la solution temporelle auxiliaire est donnée par :

$$v_h^{n+\frac{1}{2}}(t,x_{\frac{3}{2}})=\begin{cases}u_1^n(x_{\frac{3}{2}}-a(t-t^n)), & \text{si } (t-t^n)<\frac{\Delta x}{a},\\ v_{\frac{1}{2}}^{n+\frac{1}{2}}\left(t-\frac{\Delta x}{a}\right), & \text{sinon.}\end{cases}$$

Étape de projection : Les solutions exactes sont projetées sur leurs espaces polynomiaux respectifs.
Sur la cellule C_1, comme $v_{\frac{1}{2}}^{n+\frac{1}{2}}(t)\in\mathbb{R}_k[X]$, la projection sur $\mathbb{P}_C^k$ est exacte :

$$u_1^{n+1}(x)=u_1^h(x,t^n+\Delta t)=u_L\left(t^{n+1}-\frac{x-x_{\frac{1}{2}}}{a}\right).$$

En conséquence, la solution sur le bord gauche est :

$$\begin{aligned}v_h^{n+\frac{1}{2}}(t,x_{\frac{3}{2}})&=\begin{cases}v_{\frac{1}{2}}^{n-\frac{1}{2}}\left(t-\frac{\Delta x}{a}\right), & \text{si } (t-t^n)<\frac{\Delta x}{a},\\ v_{\frac{1}{2}}^{n+\frac{1}{2}}\left(t-\frac{\Delta x}{a}\right), & \text{sinon,}\end{cases}\\&=\begin{cases}u_L(t-\frac{\Delta x}{a})\big|_{I^{n-\frac{1}{2}}}, & \text{si } (t-t^n)<\frac{\Delta x}{a},\\ u_L(t-\frac{\Delta x}{a})\big|_{I^{n+\frac{1}{2}}}, & \text{sinon.}\end{cases}\end{aligned}$$

Puisque $u_L(t) \in \mathbb{R}_k[X]$, $u_1^h(x, t^n + \Delta t) \in \mathbb{R}_k[X]$ et la projection sur l'espace $\mathbb{P}_I^k$ est encore exacte.
On suppose à présent que l'on a calculé les solutions exactes $u_p^{n+1}(x)$ et $v_{p+\frac{1}{2}}^{n+\frac{1}{2}}(t)$ sur les volumes $\{K_p^n\}_{p=1,\dots,i-1}$. On s'intéresse au volume K_i^n.
Étape d'évolution : La solution exacte sur la cellule C_i au temps $t = t^{n+1}$ est :

$$u_i^h(x, t^n + \Delta t) = v_{i-\frac{1}{2}}^{n+\frac{1}{2}}\left(t^{n+1} - \frac{x - x_{i-\frac{1}{2}}}{a}\right),$$

et sur l'interface $x_{i+\frac{1}{2}}$, la solution temporelle est :

$$v_h^{n+\frac{1}{2}}(t, x_{i+\frac{1}{2}}) = \begin{cases} u_i^n(x_{i+\frac{1}{2}} - a(t - t^n)), & \text{si } (t - t^n) < \frac{\Delta x}{a}, \\ v_{i-\frac{1}{2}}^{n+\frac{1}{2}}(t - \frac{\Delta x}{a}), & \text{sinon.} \end{cases}$$

Étape de Projection : La solution exacte est projetée sur les espaces $\mathbb{P}^k$. Sur la cellule C_i, comme $v_{i-\frac{1}{2}}^{n+\frac{1}{2}}(t) \in \mathbb{R}_k[X]$, la projection sur l'espace $\mathbb{P}_C^k$ est exacte :

$$u_i^{n+1}(x) = u_i^h(x, t^n + \Delta t) = v_{i-\frac{1}{2}}\left(t^{n+1} - \frac{x - x_{\frac{1}{2}}}{a}\right).$$

Puisque les solutions exactes sur les interfaces $I^{n+\frac{1}{2}} \times \{x_{p+\frac{1}{2}}\}_{p=1,\dots,i-1}$ sont connues, il vient :

$$u_i^h(x, t^n + \Delta t) = \begin{cases} u_{i-1}^n(x - a\Delta t), & \text{si } x > x_{i-\frac{1}{2}} + a\Delta t - \Delta x, \\ u_{i-2}^n(x - a\Delta t), & \text{si } x_{i-\frac{1}{2}} + a\Delta t - \Delta x > x > x_{i-\frac{1}{2}} + a\Delta t - 2\Delta x, \\ \vdots & \\ u_{i-p}^n(x - a\Delta t), & \text{si } x_{i-\frac{1}{2}} + a\Delta t - (p-1)\Delta x > x > x_{i-\frac{1}{2}} + a\Delta t - p\Delta x, \\ u_{i-p-1}^n(x - a\Delta t), & \text{si } x_{i-\frac{1}{2}} + a\Delta t - p\Delta x > x > x_{i-\frac{1}{2}} + a\Delta t - (p+1)\Delta x, \\ v_{i-p-\frac{3}{2}}^{n+\frac{1}{2}}\left(t^{n+1} - \frac{x - x_{i-p-\frac{3}{2}}}{a}\right), & \text{sinon.} \end{cases}$$

En posant $j = E\left(\left(\frac{a\Delta t}{\Delta x}\right)\right)$, sur la cellule C_i on a :

$$\begin{aligned} u_i^h(x, t^n + \Delta t)\big|_{C_i} &= \begin{cases} u_{i-j}^n(x - a\Delta t), & \text{si } x > x_{i-\frac{1}{2}} + a\Delta t - j\Delta x, \\ u_{i-j-1}^n(x - a\Delta t), & \text{si } x_{i-\frac{1}{2}} + a\Delta t - j\Delta x > x, \end{cases} \\ &= u^n(x - a\Delta t)\big|_{C_i}. \end{aligned}$$

Comme $u_i^h(x, t^n + \Delta t) \in \mathbb{R}_k[X]$, sa projection $u_i^{n+1}(x)$ sur l'espace $\mathbb{P}_C^k$ est exacte. Sur le bord droit, on a :

$$\begin{aligned} v_h^{\frac{1}{2}}(t, x_{i+\frac{1}{2}}) &= \begin{cases} v_{i-\frac{1}{2}}^{n-\frac{1}{2}}(t - \frac{\Delta x}{a}), & \text{si } (t - t^n) < \frac{\Delta x}{a}, \\ v_{i-\frac{1}{2}}^{n+\frac{1}{2}}(t - \frac{\Delta x}{a}), & \text{sinon,} \end{cases} \\ &= \begin{cases} v_{i-\frac{1}{2}}(t - \frac{\Delta x}{a})\big|_{I^{n-\frac{1}{2}}}, & \text{si } (t - t^n) < \frac{\Delta x}{a}, \\ v_{i-\frac{1}{2}}(t - \frac{\Delta x}{a})\big|_{I^{n+\frac{1}{2}}}, & \text{sinon.} \end{cases} \end{aligned}$$

De même que précédemment, comme $v_h^{\frac{1}{2}}(t, x_{i+\frac{1}{2}}) \in \mathbb{R}_k[X]$, sa projection $v_{i+\frac{1}{2}}^{n+\frac{1}{2}}(t)$ sur l'espace $\mathbb{P}_I^k$ est exacte.
Finalement, le schéma implicite préserve la solution exacte si la donnée initiale $u_0(x)$ et la donnée au bord $u_L(t)$ sont polynomiales. □

TABLE 2.1 – Tableaux de convergence pour le schéma explicite (gauche) et le schéma implicite (droite)

k	# mailles	erreur L^2	ordre
1	64	$2.501.10^{-3}$	
	128	$3.633.10^{-4}$	2.783
	256	$5.798.10^{-5}$	2.648
3	64	$1.613.10^{-8}$	
	128	$1.008.10^{-9}$	4.000
	256	$6.300.10^{-11}$	4.000
5	64	$7.718.10^{-13}$	
	128	$1.205.10^{-14}$	6.001
	256	$1.883.10^{-16}$	6.000

k	# mailles	erreur L^2	ordre
1	64	$3.388.10^{-1}$	
	128	$5.452.10^{-2}$	2.635
	256	$7.296.10^{-3}$	2.902
3	64	$4.748.10^{-5}$	
	128	$2.880.10^{-6}$	4.043
	256	$1.570.10^{-7}$	4.197
5	64	$1.875.10^{-7}$	
	128	$2.958.10^{-9}$	5.986
	256	$4.947.10^{-11}$	5.902

Le schéma est donc d'ordre $k+1$ en espace et en temps à la fois pour le cas explicite et pour le cas implicite. De plus, il est très compact. C'est un atout qui va simplifier l'implémentation ainsi que l'extension en dimension deux d'espace sur des maillages non structurés.

Afin d'illustrer la méthode, quelques résultats numériques de convergence sont présentés dans le paragraphe suivant. Les résultats sont comparés avec les schémas DG et WENO, dont le principe est rappelé dans le paragraphe 2.6. D'autres cas-tests sont également présentés dans le paragraphe 5.1.1, comme pour l'approximation des équations de Maxwell linéarisées ou encore du modèle d'ordonnées discrètes du transfert radiatif.

2.3.1 Résultats numériques en dimension 1

Tous les cas-tests présentés dans ce paragraphe ont été réalisés avec $a = \frac{c}{2}$ sur un maillage uniforme avec des conditions aux limites périodiques.

Convergence pour les solutions régulières

Dans un premier temps, on considère une solution initiale régulière pour vérifier numériquement l'ordre du schéma :

$$u_0(x) = \sin(\pi x) \text{ pour } x \in [-1, 1].$$

Le tableau 2.1 montre l'erreur en norme L^2 commise entre la solution exacte et la solution donnée par le schéma explicite (avec une CFL $\lambda = 0.5$) et le schéma implicite (avec une CFL $\lambda = 10$) au bout d'un temps $t = 10\ s$. On rappelle que $c\Delta t = \lambda \Delta x$, ce qui implique que l'ordre de la méthode en temps est le même que l'ordre en espace. On voit que les ordres numériques correspondent bien aux ordres théoriques. De plus, les erreurs en norme L^2 sont très petites. Pour comparaison, avec 256 points, il faut 3466 itérations avec $\lambda = 0.5$ et 174 itérations avec $\lambda = 10$. En pratique, pour le schéma explicite, le code a été exécuté en quadruple précision car l'erreur commise était inférieure à l'erreur machine de la double précision. Les erreurs commises avec le schéma implicite, bien que toujours petites, restent supérieures au cas explicite.

On compare ensuite le schéma GRP espace-temps avec les schémas de type WENO et Galerkin Discontinu (DG) présentés plus tard dans les paragraphes 2.6.2 et 2.6.3. Dans le tableau 2.2, on compare les erreurs en norme L^2 de ces trois méthodes et pour trois ordres d'approximation. On

TABLE 2.2 – Comparaison de la convergence des trois méthodes d'ordre 3

# mailles	WENO	DG	GRP espace-temps	ordre
32	$3.158.10^{-1}$	$1.143.10^{-4}$	$1.531.10^{-5}$	
64	$1.078.10^{-1}$	$1.411.10^{-5}$	$1.914.10^{-6}$	3.00002
128	$2.405.10^{-2}$	$1.757.10^{-6}$	$2.394.10^{-7}$	2.99954
256	$3.148.10^{-3}$	$2.194.10^{-7}$	$2.992.10^{-8}$	2.99991
512	$1.906.10^{-4}$			

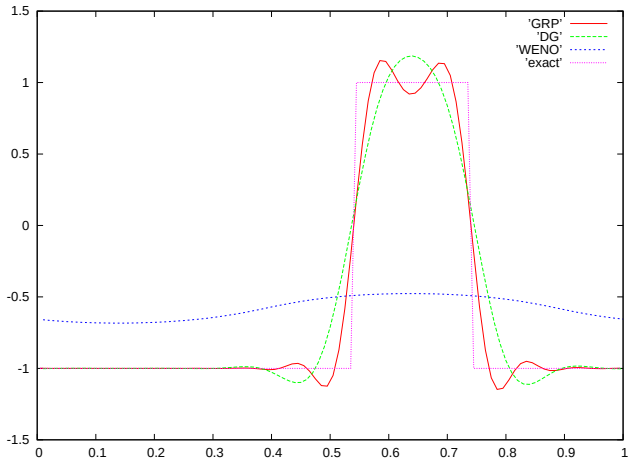


FIGURE 2.5 – Comparaison des trois méthodes à $t = 1000\ s$

considère les mêmes conditions initiales que pour le test de convergence mais en imposant cette fois $\lambda = 0.2$ pour prendre en compte la CFL plus restrictive de la méthode DG. Pour 256 points par exemple, les solutions sont obtenues au bout de 8665 itérations.

Avec les trois méthodes, on retrouve bien les ordres d'approximation théoriques. Cependant, on observe que les erreurs données par la méthode WENO sont très grandes par rapport aux autres méthodes. En pratique, cela signifie qu'il faudrait 10 fois plus de mailles pour obtenir la même précision. Les schémas GRP et DG ont des erreurs comparables, mais les méthodes DG ont une condition CFL de plus en plus restrictive quand l'ordre d'approximation augmente, alors que le schéma GRP espace-temps n'a pas de restriction sur le pas de temps.

Solution discontinue

On considère à présent une donnée initiale constante par morceaux :

$$u_0(x) = \begin{cases} 1 & \text{si } 0.4 < x < 0.6, \\ -1 & \text{sinon,} \end{cases}$$

sur le domaine $[0, 1]$. Les solutions présentées sur les graphes 2.5,2.6 et 2.7 ont été obtenues avec 100 mailles.

Le premier cas-test est une comparaison entre les trois méthodes explicites d'ordre 3 : GRP espace-temps, WENO et DG en temps long. Sur la figure 2.5, les solutions sont obtenues avec $\lambda = 0.2$ au bout d'un temps $t = 1000\ s$, soit 672775 itérations. Après un temps très long, la solution approchée donnée par le schéma WENO est totalement écrasée. Par ailleurs, les approximations données par les méthodes GRP espace-temps et DG sont proches de la solution exacte. Concernant le temps CPU, les schémas DG et GRP espace-temps ont un coût de calcul comparable alors que

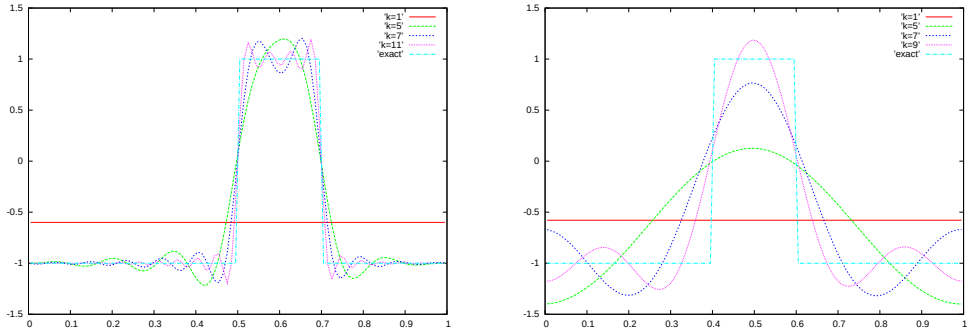


FIGURE 2.6 – Solutions à $t = 1000\ s$ avec $\lambda = 10$ (gauche) et $\lambda = 50$ (droite)

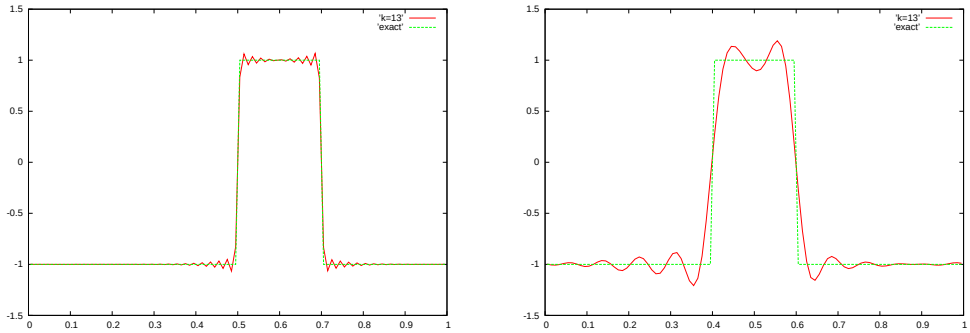


FIGURE 2.7 – Solutions à $t = 15000\ s$ avec $\lambda = 5$ (gauche) et $\lambda = 20$ (droite)

la méthode WENO est largement plus coûteuse.

On s'intéresse à présent uniquement aux résultats de la méthode GRP espace-temps pour vérifier son comportement en temps très long avec de grands pas de temps. Le premier cas-test illustre l'impact du degré d'approximation k des polynômes en temps long, et le second cas-test montre quelques résultats en temps très long.
Sur la figure 2.6, on compare les solutions approchées pour différentes valeurs de k après un temps $t = 1000\ s$ pour une CFL égale à 10 et 50.

Quand k devient grand, en plus d'améliorer la qualité des résultats, on voit que les oscillations se concentrent autour des discontinuités et que leur amplitude diminue. Ainsi, en choisissant k de plus en plus grand, on peut obtenir une approximation de la solution exacte aussi proche que l'on veut.

Le schéma est maintenant testé pour un temps extrêmement long $t = 15000\ s$. La figure 2.7 montre les solutions obtenues avec $k = 13$ pour $\lambda = 5$ et $\lambda = 20$.

En dépit de la difficulté de ce cas-test (seulement 100 mailles et un temps extrêmement long), les approximations sont très proches de la solution exacte. De plus, le temps de calcul pour un tel cas-test reste raisonnable. Avec $\lambda = 5$, on observe un rapport de temps de 8.3 entre $k = 13$ et $k = 3$, de 12.8 entre $k = 13$ et $k = 2$ et de 20 entre $k = 13$ et $k = 1$.
Si l'on veut un temps extrêmement long et une CFL très grande, on peut éventuellement considérer un degré k très grand. Cependant, comme le conditionnement des matrices augmente avec k, l'erreur d'approximation numérique devient très importante quand k est très grand. En pratique,

on peut difficilement considérer k supérieur à $15-20$. Pour de tels cas, une possibilité serait de considérer des bases de polynômes orthogonaux afin de forcer un meilleur conditionnement des matrices à inverser.

2.4 Méthode de projection GRP espace-temps pour l'équation de transport en dimension 2

Dans ce paragraphe, on introduit une extension en dimension deux du schéma GRP espace-temps présenté précédemment en dimension un. Grâce au choix d'approximation de la solution et au caractère compact du schéma, on présente directement l'extension en deux dimensions sur des maillages non-structurés. On considère toujours le problème mixte suivant :

$$\begin{cases} \partial_t u(\mathbf{x},t) + a\vec{\Omega}.\nabla_{\mathbf{x}} u(\mathbf{x},t) = 0, \quad \forall x \in D, t \geq 0, \\ u(\mathbf{x},t=0) = u_0(\mathbf{x}), \quad \forall x \in D, \\ u_{|\mathbf{x}\in\Gamma}(t) = u_B(\mathbf{x},t), \quad \Gamma = \{x \in \nabla D, \overrightarrow{\Omega} \cdot \overrightarrow{n}(x) < 0\}, \end{cases} \tag{2.23}$$

où $D \subset \mathbb{R}^2$ est un ouvert borné, a est la vitesse et $\vec{\Omega} = \begin{pmatrix} \Omega_x \\ \Omega_y \end{pmatrix}$ est le vecteur direction normalisé. On suppose que la donnée sur le bord est rentrante, c'est-à-dire que $\overrightarrow{\Omega} \cdot \overrightarrow{n} < 0$ sur Γ où $\overrightarrow{n}$ est la normale unitaire à ∇D extérieure à D.
Le domaine est discrétisé par un maillage non-structuré composé de triangles. Pour simplifier la présentation du schéma, on se place dans le cas de maillages triangulaires, mais tout type de polygones peut être considéré. On suppose que le pas de temps est constant égal à Δt. On note (x_i, y_i) le centre de gravité du triangle T_i et V_i son aire. Les côtés du triangle T_i sont les segments s_j. Le segment $s_j = [p_j, q_j],\ p_j \in \mathbb{R}^2, q_j \in \mathbb{R}^2$ est paramétré par :

$$M = (x,y) \in s_j \Leftrightarrow \exists\, \omega(x,y) \in\ [0,1]/M = p_j + \omega(x,y)\ \overrightarrow{p_j q_j}. \tag{2.24}$$

De plus, on notera s_j^{Γ} les segments sur le bord du domaine.

De la même manière qu'en dimension un, on définit les interfaces du maillage : $I_j^{n+\frac{1}{2}} = (t^n, t^{n+1}) \times s_j$.
Ainsi, on peut introduire l'ensemble des polynômes de degré k par cellule (triangle) :

$$\mathbb{P}_C^k = \{u \in L^\infty(\mathbb{R}^2)/\ u_{|T_i} \in \mathbb{R}_k[X,Y]\},$$

qui est engendré par la base :

$$\left\{ \left(\frac{x-x_i}{V_i}\right)^p \left(\frac{y-y_i}{V_i}\right)^q \right\}_{p+q=0,\ldots,k}.$$

Sur cet espace, on considère le produit scalaire sur le triangle T_i suivant :

$$< f, g >_i = \frac{1}{V_i} \iint_{T_i} f(x,y) g(x,y)\ dx\ dy.$$

Puis, on introduit également l'ensemble des polynômes par interface :

$$\mathbb{P}_I^k = \{u \in L^\infty(\mathbb{R}^2)/\ u_{\left|I_j^{n+\frac{1}{2}}\right.} \in \mathbb{R}_k[X,Y]\},$$

engendré par la base :

$$\left\{ \left(\frac{t - t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m \right\}_{l+m=0,\ldots,k},$$

Sur cet espace, on définit le produit scalaire sur l'interface $I_j^{n+\frac{1}{2}}$ par :

$$< f, g >^n = \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \int_0^1 f(t,\omega) g(t,\omega)\, d\omega\, dt.$$

Comme dans le cas unidimensionnel, les bases définies ci-dessus ont été choisies afin de faciliter les calculs. D'autres bases, comme la base des polynômes de Legendre par exemple, peuvent être adoptées.
Le schéma consiste alors à faire évoluer de manière exacte l'approximation spatiale d'ordre élevé de la solution ainsi que son approximation temporelle qui sera toujours utilisée ici comme un outil pour l'élaboration du schéma. Finalement, il suffira de reprojeter ces approximations de la solution sur leurs espaces respectifs pour caractériser entièrement la solution.

Approximations considérées : Sur chaque triangle du maillage, on suppose que la donnée initiale est une approximation polynomiale de degré k de la solution exacte :

$$\text{pour } (x,y) \in T_i, \quad u_0(x,y) \simeq u_i^0(x,y) \in \mathbb{P}_C^k.$$

Puis, sur le bord Γ du domaine, on se donne une approximation temporelle auxiliaire polynomiale de degré k de la donnée u_B pour les interfaces "rentrantes" :

$$\text{pour } t \in [t^n; t^{n+1}], (x,y) \in s_j^\Gamma \text{ tel que } \overrightarrow{\Omega} \cdot \overrightarrow{n_j^\Gamma} < 0, \quad u_B(x,y,t) \simeq v_{j^\Gamma}^{n+\frac{1}{2}}(t, \omega(x,y)) \in \mathbb{P}_I^k.$$

Dans la suite, on notera s_j^Γ les interfaces "rentrantes", c'est-à-dire où la solution est donnée.
Une fois que les notations sont introduites, on procède à l'étape d'évolution.

Étape d'évolution : À cette étape, on suppose donc que l'on doit résoudre le problème mixte suivant :

$$\begin{cases} \partial_t u(\mathbf{x},t) + a\vec{\Omega}.\nabla_{\mathbf{x}} u(\mathbf{x},t) = 0, \\ u(\mathbf{x}, t = t^n) = u^n(\mathbf{x}), \\ v_{j^\Gamma}(t) = u_B(\mathbf{x},t) \text{ si } \mathbf{x} \in s_j^\Gamma. \end{cases} \tag{2.25}$$

On fait évoluer de manière exacte l'approximation $u^n(x,y)$ ainsi que l'approximation temporelle $v_j(t)$ de la solution. Grâce à la linéarité de l'équation dans (2.25), la solution exacte peut être calculée analytiquement par la méthode des caractéristiques. On se place dans le triangle T_i formé des segments s_{j1}, s_{j2} et s_{j3}. On suppose que la direction du flux $\vec{\Omega}$ est comprise entre les directions de deux segments s_{j1} et s_{j2} de ce triangle. Ainsi, en fonction du sens du flux, on obtient les deux configurations suivantes (voir Figure 2.8) :

- cas *une cible* : le flux entre dans la cellule entre les segments s_{j1} et s_{j2},
- cas *deux cibles* : le flux entre dans la cellule par le troisième segment s_{j3}.

Puis dans chaque cas, deux sous-cas doivent être considérés (voir figure 2.9) :

- cas explicite : des droites caractéristiques provenant de la cellule T_i au temps t^n croisent la cellule T_i au temps t^{n+1},
- cas implicite : toutes les droites caractéristiques provenant de la cellule T_i au temps t^n croisent une interface en premier.

Il est important de noter qu'en fonction de la taille des cellules, on peut avoir des cas implicites et explicites au cours d'un même pas de temps. Ceci marque une différence notable en comparaison du cas unidimensionnel présenté ci-dessus pour des maillages uniformes.
Par convention, on notera s_j^n le $j^{ème}$ segment au temps t^n et s_j^{n+1} le $j^{ème}$ segment au temps t^{n+1} (voir figure 2.9). L'interface rectangulaire délimitée par s_j et (t^n, t^{n+1}) est notée $I_j^{n+\frac{1}{2}}$. Le choix

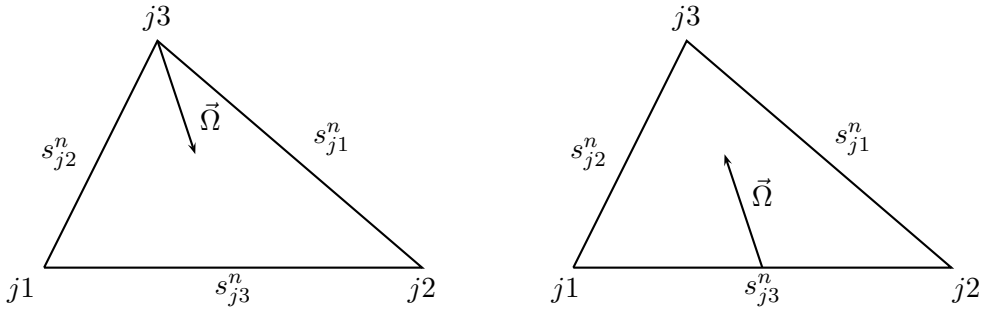


FIGURE 2.8 – Cas *une cible* (gauche) et cas *deux cibles* (droite)

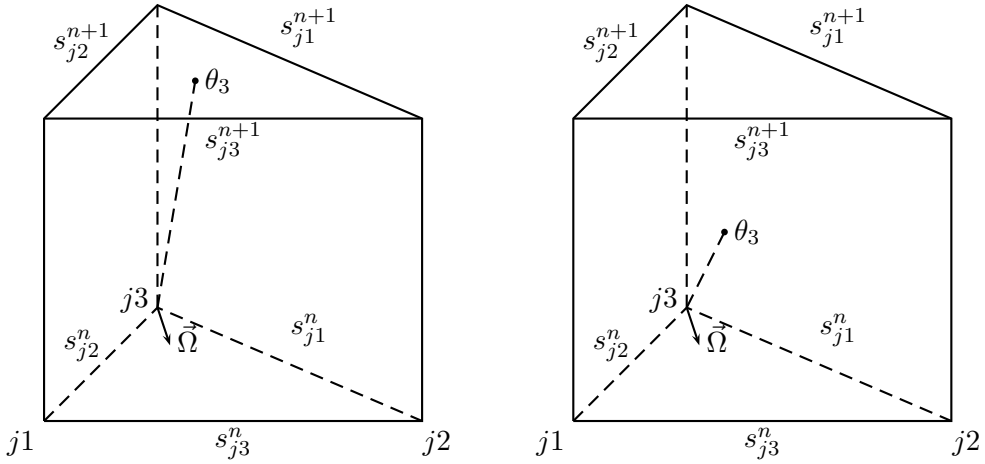


FIGURE 2.9 – Cas explicite (gauche) et cas implicite (droite) pour le cas *une cible*

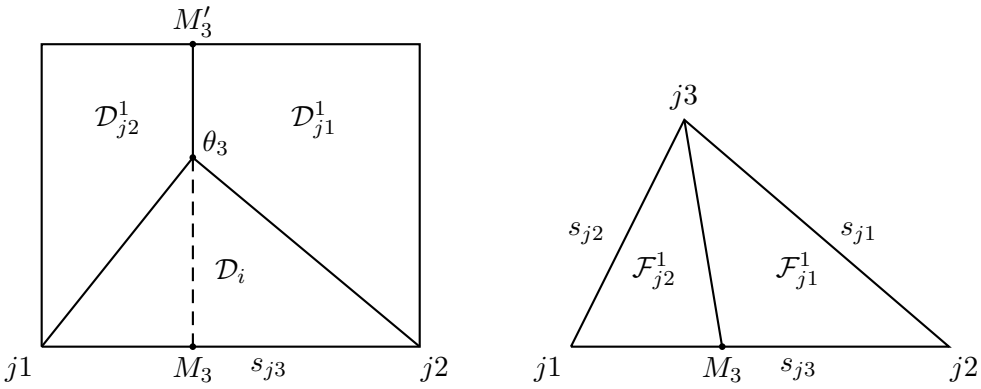


FIGURE 2.10 – Solutions entrantes sur l'interface (gauche) et sur la cellule (droite) pour le cas *une cible* implicite

de paramétrisation des segments s_j, donné par (2.24), implique que les sommets de chaque segment sont ordonnés : $s_j = [p_j, q_j]$.

Dans la suite, les calculs sont détaillés pour le cas *une cible* implicite. Les autres cas (*une cible* explicite, *deux cibles* implicite, *deux cibles* explicite) sont détaillés dans l'appendice 1.

Cas *une cible* (implicite) :

- la solution $u_i^n(x,y)$ est connue dans la cellule T_i au temps t^n ainsi que les solutions $v_{j1}^{n+\frac{1}{2}}(t,\omega)$ et $v_{j2}^{n+\frac{1}{2}}(t,\omega)$ sur les interfaces $I_{j1}^{n+\frac{1}{2}}$ et $I_{j2}^{n+\frac{1}{2}}$.
- On définit θ_3 comme l'intersection entre la droite caractéristique issue du nœud $j3$ et de l'interface $I_{j3}^{n+\frac{1}{2}}$. Puis, M_3 est défini comme la projection de θ_3 sur s_{j3}^n et M_3' est sa projection sur s_{j3}^{n+1} (voir figure 2.10).
- Pour définir la solution exacte sur la cellule T_i au temps t^{n+1}, on introduit les domaines $\mathcal{D}$ et $\mathcal{F}$ représentés sur la figure 2.10. Avec ces notations, la solution exacte cherchée est donnée par :

$$u_i^h(x,y,t^{n+1}) = \begin{cases} v_{j1}^{n+\frac{1}{2}}(t_1(x,y),\omega_1(x,y)), & \text{si } (x,y) \in \mathcal{F}_{j1}^1 \\ v_{j2}^{n+\frac{1}{2}}(t_2(x,y),\omega_2(x,y)), & \text{si } (x,y) \in \mathcal{F}_{j2}^1 \end{cases} \tag{2.26}$$

où t_1, ω_1, t_2 et ω_2 sont les changements de variables suivants :

$$\omega_1(x,y) = \frac{\Omega_x(y - y_{p_{j1}} - a\Omega_y(t^n + \Delta t)) - \Omega_y(x - x_{pj1} - a\Omega_x(t^n + \Delta t))}{\Omega_x(y_{q_{j1}} - y_{p_{j1}}) - \Omega_y(x_{q_{j1}} - x_{p_{j1}})},$$

$$t_1(x,y) = \begin{cases} \frac{x_{p_{j1}} + \omega_1(x,y)(x_{q_{j1}} - x_{p_{j1}}) - x}{a\Omega_x} + t^n + \Delta t, & \text{si } \Omega_x \neq 0, \\ \frac{y_{p_{j1}} + \omega_1(x,y)(y_{q_{j1}} - y_{p_{j1}}) - y}{a\Omega_y} + t^n + \Delta t, & \text{sinon,} \end{cases}$$

$$\omega_2(x,y) = \frac{\Omega_x(y - y_{p_{j2}} - a\Omega_y(t^n + \Delta t)) - \Omega_y(x - x_{pj2} - a\Omega_x(t^n + \Delta t))}{\Omega_x(y_{q_{j2}} - y_{p_{j2}}) - \Omega_y(x_{q_{j2}} - x_{p_{j2}})},$$

$$t_2(x,y) = \begin{cases} \frac{x_{p_{j2}} + \omega_2(x,y)(x_{q_{j2}} - x_{p_{j2}}) - x}{a\Omega_x} + t^n + \Delta t, & \text{si } \Omega_x \neq 0, \\ \frac{y_{p_{j2}} + \omega_2(x,y)(y_{q_{j2}} - y_{p_{j2}}) - y}{a\Omega_y} + t^n + \Delta t, & \text{sinon.} \end{cases}$$

- La solution temporelle exacte sur l'interface $I_{j3}^{n+\frac{1}{2}}$ est donnée par :

$$v_h^{n+\frac{1}{2}}(t,\omega_{j3}) = \begin{cases} u_i^n(x_1(t,\omega), y_1(t,\omega)), & \text{si } (t,\omega) \in \mathcal{D}_i, \\ v_{j1}^{n+\frac{1}{2}}(t_3(t,\omega), \omega_3(t,\omega)), & \text{si } (t,\omega) \in \mathcal{D}_{j1}^1, \\ v_{j2}^{n+\frac{1}{2}}(t_4(t,\omega), \omega_4(t,\omega)), & \text{si } (t,\omega) \in \mathcal{D}_{j2}^1, \end{cases} \tag{2.27}$$

avec :

$$\begin{aligned} x_1(t,\omega) &= x_{p_{j3}} + \omega(x_{q_{j3}} - x_{p_{j3}}) - a\Omega_x(t - t^n), \\ y_1(t,\omega) &= y_{p_{j3}} + \omega(y_{q_{j3}} - y_{p_{j3}}) - a\Omega_y(t - t^n), \\ t_3(t,\omega) &= \frac{c_3^1(y_{p_{j1}} - y_{q_{j1}}) + c_3^2(x_{q_{j1}} - x_{p_{j1}})}{d_3}, \\ \omega_3(t,\omega) &= \frac{ac_3^2\Omega_x - ac_3^1\Omega_y}{d_3}, \\ c_3^1 &= x_{p_{j3}} - x_{p_{j1}} + \omega(x_{q_{j3}} - x_{p_{j3}}) - a\Omega_x t, \\ c_3^2 &= y_{p_{j3}} - y_{p_{j1}} + \omega(y_{q_{j3}} - y_{p_{j3}}) - a\Omega_y t, \\ d_3 &= a\Omega_x(y_{q_{j1}} - y_{p_{j1}}) - a\Omega_y(x_{q_{j1}} - x_{p_{j1}}), \\ t_4(t,\omega) &= \frac{c_4^1(y_{p_{j2}} - y_{q_{j2}}) + c_4^2(x_{q_{j2}} - x_{p_{j2}})}{d_4}, \\ \omega_4(t,\omega) &= \frac{ac_4^2\Omega_x - ac_4^1\Omega_y}{d_4}, \\ c_4^1 &= x_{p_{j3}} - x_{p_{j2}} + \omega(x_{q_{j3}} - x_{p_{j3}}) - a\Omega_x t, \\ c_4^2 &= y_{p_{j3}} - y_{p_{j2}} + \omega(y_{q_{j3}} - y_{p_{j3}}) - a\Omega_y t, \\ d_4 &= a\Omega_x(y_{q_{j2}} - y_{p_{j2}}) - a\Omega_y(x_{q_{j2}} - x_{p_{j2}}). \end{aligned}$$

Une fois que l'on a fait évolué de manière exacte les approximations polynomiales de la solution, on les projette dans leurs espace de polynômes respectifs.

Étape de projection : Les solutions exactes (2.26)-(2.27) sont maintenant projetées sur l'espace d'approximation $\mathbb{P}^k$: dans la cellule T_i au temps t^{n+1}, la fonction u_i^h est projetée sur $\mathbb{P}_C^k$ et sur les interfaces $I_j^{n+\frac{1}{2}}$, la fonction $v_h^{n+\frac{1}{2}}$ est projetée sur $\mathbb{P}_I^k$. Pour le cas *une cible* implicite, cette projection se traduit par deux problèmes de minimisation suivants :

$$\|u_i^{n+1} - u_i^h(t^n + \Delta t)\|_i = \inf_{p \in \mathbb{P}_C^k} \|p - u_i^h(t^n + \Delta t)\|_i, \tag{2.28}$$

$$\|v_{j3}^{n+\frac{1}{2}} - v_h^{n+\frac{1}{2}}(\omega_{j3})\|^n = \inf_{q \in \mathbb{P}_I^k} \|q - v_h^{n+\frac{1}{2}}(\omega_{j3})\|^n. \tag{2.29}$$

Grâce aux conditions de Petrov-Galerkin, les deux problèmes de minimisation (2.28)-(2.29) peuvent se réécrire de la manière suivante :

$$\forall\, l+m=0,\dots,k \quad < u_i^{n+1}-u_i^h(t^{n+1}),\left(\frac{x-x_i}{V_i}\right)^l\left(\frac{y-y_i}{V_i}\right)^m >_i = 0, \tag{2.30}$$

$$\forall\, l+m=0,\dots,k \quad < v_{j3}^{n+\frac{1}{2}}-v_h^{n+\frac{1}{2}}(\omega_{j3}),\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l\omega^m >^n = 0. \tag{2.31}$$

Puis, en utilisant leurs décompositions dans leurs bases de polynômes respectives :

$$\begin{aligned}
u_i^{n+1}(x,y) &= \sum_{p+q=0}^{k}\alpha_i^{n+1,p,q}\left(\frac{x-x_i}{V_i}\right)^p\left(\frac{y-y_i}{V_i}\right)^q,\\
v_{j1}^{n+\frac{1}{2}}(t,\omega) &= \sum_{p+q=0}^{k}\beta_{j1}^{n+\frac{1}{2},p,q}\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^p\omega^q,\\
v_{j2}^{n+\frac{1}{2}}(t,\omega) &= \sum_{p+q=0}^{k}\beta_{j2}^{n+\frac{1}{2},p,q}\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^p\omega^q,\\
v_{j3}^{n+\frac{1}{2}}(t,\omega) &= \sum_{p+q=0}^{k}\beta_{j3}^{n+\frac{1}{2},p,q}\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^p\omega^q,
\end{aligned}$$

les problèmes de minimisation (2.30)-(2.31) se ramènent à la résolution des deux systèmes linéaires de taille $\frac{(k+1)(k+2)}{2}$ suivants :

$$\begin{aligned}
\sum_{p+q=0}^{k}\alpha_i^{n+1,p,q} &< \left(\frac{x-x_i}{V_i}\right)^p\left(\frac{y-y_i}{V_i}\right)^q,\left(\frac{x-x_i}{V_i}\right)^l\left(\frac{y-y_i}{V_i}\right)^m >_i =\\
&< u_i^h(t^{n+1}),\left(\frac{x-x_i}{V_i}\right)^l\left(\frac{y-y_i}{V_i}\right)^m >_i \quad \forall\, l+m=0,\dots,k,\\
\sum_{p+q=0}^{k}\beta_{j3}^{n+\frac{1}{2},p,q} &< \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^p\omega^q,\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l\omega^m >^n =\\
&< v_h^{n+\frac{1}{2}}(\omega_{j3}),\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l\omega^m >^n \quad \forall\, l+m=0,\dots,k.
\end{aligned}$$

Les membres de gauche s'écrivent :

$$
\begin{aligned}
&< \left(\frac{x-x_i}{V_i}\right)^p \left(\frac{y-y_i}{V_i}\right)^q, \left(\frac{x-x_i}{V_i}\right)^l \left(\frac{y-y_i}{V_i}\right)^m >_i \\
&\qquad = \frac{1}{V_i} \iint_{T_i} \left(\frac{x-x_i}{V_i}\right)^{p+l} \left(\frac{y-y_i}{V_i}\right)^{q+m} dx\, dy, \\
&\qquad = A^i(p,q,l,m), \\
&< \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega^q, \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m >^n \\
&\qquad = \frac{1}{\Delta t} \iint_{I_j^{n+\frac{1}{2}}} \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^{p+l} \omega^{q+m} dt\, d\omega, \\
&\qquad = \frac{1}{(p+l+1)(q+m+1)} \left[\left(\frac{1}{2}\right)^{p+l+1} - \left(-\frac{1}{2}\right)^{p+l+1}\right], \\
&\qquad = A_j(p,q,l,m).
\end{aligned}
$$

De même, les seconds membres s'écrivent :

$$
\begin{aligned}
&< u_i^h(t^n+\Delta t), \left(\frac{x-x_i}{V_i}\right)^l \left(\frac{y-y_i}{V_i}\right)^m >_i = \sum_{p+q=0}^{k} \left(\beta_{j1}^{n+\frac{1}{2},p,q} J_{im,j1}^{p,q,l,m} + \beta_{j2}^{n+1,p,q} J_{im,j2}^{p,q,l,m}\right), \\
&\qquad = b^i(l,m), \\
&< v_h^{n+\frac{1}{2}}(\omega_{j3}), \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m >^n = \sum_{p+q=0}^{k} \left(\alpha_i^{n,p,q} K_{im,i}^{p,q,l,m} + \beta_{j1}^{n+\frac{1}{2},p,q} K_{im,j1}^{p,q,l,m} + \beta_{j2}^{n+1,p,q} K_{im,j2}^{p,q,l,m}\right), \\
&\qquad = b_{j3}(l,m),
\end{aligned}
$$

où les coefficients sont donnés par :

$$
\begin{aligned}
J_{im,j1}^{p,q,l,m} &= \frac{1}{V_i} \iint_{\mathcal{F}_{j1}^1} \left(\frac{t_1(x,y)-t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega_1(x,y)^q \left(\frac{x-x_i}{V_i}\right)^l \left(\frac{y-y_i}{V_i}\right)^m dx\, dy \\
J_{im,j2}^{p,q,l,m} &= \frac{1}{V_i} \iint_{\mathcal{F}_{j2}^1} \left(\frac{t_2(x,y)-t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega_2(x,y)^q \left(\frac{x-x_i}{V_i}\right)^l \left(\frac{y-y_i}{V_i}\right)^m dx\, dy \\
K_{im,i}^{p,q,l,m} &= \frac{1}{\Delta t} \iint_{\mathcal{D}_i} \left(\frac{x_1(t,\omega)-x_i}{V_i}\right)^p \left(\frac{y_1(t,\omega)-y_i}{V_i}\right)^q \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m dt\, d\omega \\
K_{im,j1}^{p,q,l,m} &= \frac{1}{\Delta t} \iint_{\mathcal{D}_{j1}^1} \left(\frac{t_3(t,\omega)-t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega_3(t,\omega)^q \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m dt\, d\omega \\
K_{im,j2}^{p,q,l,m} &= \frac{1}{\Delta t} \iint_{\mathcal{D}_{j2}^1} \left(\frac{t_4(t,\omega)-t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega_4(t,\omega)^q \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m dt\, d\omega.
\end{aligned}
$$

Ces nombreux coefficients peuvent paraître compliqués, mais les calculs sont très similaires. En effet, une fois que l'on connaît les changements de variables et les aires des différentes zones pour les solutions exactes, ces intégrales sont toutes obtenues de la même manière.

Résumé de la méthode en dimension 2 En pratique, la méthode peut être résumée de la manière suivante :

- Au temps initial, la solution $u_0(x,y)$ est connue dans chaque triangle T_i et la solution $v_{j^\Gamma}(t,\omega)$ est connue sur les interfaces s_j^Γ de la frontière pour lesquelles le flux entre dans le domaine : $\vec{n}_{s_j^\Gamma} \cdot \Omega < 0$, avec $\vec{n}_{s_j^\Gamma}$ la normale sortante au segment s_j^Γ.
- Un ordre de parcours des interfaces est défini pour s'assurer que les solutions exactes peuvent être calculées. En effet, il faut que la solution temporelle sur les interfaces par lesquelles le flux entre soit connue. Pour cela, on met en place un algorithme déterminant l'ordre de parcours des interfaces détaillé dans l'appendice 1.
- Sur chaque interface $I_j^{n+\frac{1}{2}}$, le système linéaire de taille $\frac{(k+1)(k+2)}{2}$:

$$\sum_{p+q=0}^{k} \beta_j^{n+\frac{1}{2},p,q} A_j(p,q,l,m) = \ b_j(l,m) \quad \forall\, l+m = 0,\dots,k,$$

  doit être résolu afin de calculer la solution temporelle $v_j^{n+\frac{1}{2}}(t,\omega)$.
- Finalement, dans chaque cellule T_i au temps t^{n+1}, la solution $u_i^{n+1}(x,y)$ au temps $t = t^{n+1}$ est obtenue en résolvant le système linéaire de taille $\frac{(k+1)(k+2)}{2}$ suivant :

$$\sum_{p+q=0}^{k} \alpha_i^{n+1,p,q} A^i(p,q,l,m) = \ b^i(l,m) \quad \forall\, l+m = 0,\dots,k.$$

On remarque que les matrices A^i ne dépendent pas du temps. Sur chaque cellule, la matrice A^i peut donc être calculée et inversée une seule fois pour tout le calcul. De même sur les interfaces, les matrices ne dépendent pas de l'espace. À chaque pas de temps, une seule matrice A_j devra donc être calculée et inversée pour toutes les interfaces.
De la même manière qu'en dimension un, on peut prouver que ce schéma est d'ordre $k+1$ en temps et en espace.

Théorème 2.2. *Le schéma préserve exactement les polynômes de degré k.*

Démonstration. La preuve est similaire au cas uni-dimensionnel. □

Finalement, le schéma obtenu est d'ordre $k+1$ en temps et en espace. De plus, selon la taille des cellules, il traite simultanément des cas implicites et explicites. D'un point de vue numérique, la principale difficulté vient du calcul des intégrales dans les matrices et les seconds membres des systèmes linéaires. Pour simplifier ces calculs, des méthodes de quadrature d'ordre $2k$ peuvent être utilisées.

Pour illustrer la méthode, quelques résultats de convergence sont présentés dans le paragraphe suivant. Un autre cas-test pour le modèle d'ordonnées discrètes est également exposé dans le paragraphe 5.2.1.

2.4.1 Résultats numériques en dimension 2

Convergence des solutions régulières
Dans ce premier cas-test, on chercher à évaluer numériquement l'ordre de la méthode, de la même manière qu'en dimension un. On considère ainsi la fonction $u(x,y,t) = \sin(\pi(x-t))$ sur le domaine $[0,1]^2 \times [0,1.5]$. Les états initiaux et les conditions aux limites sont définis par :

$$u_0(x,y) = \sin(\pi x) \text{ et } u(0,y,t) = \sin(-\pi t),$$

FIGURE 2.11 – Solution initiale et domaine de calcul

pour une vitesse $a = 1$ et une direction $\vec{\Omega} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Dans le tableau 2.3, on présente les erreurs en norme L^2 pour $k = 1$ et $k = 2$ ainsi que l'ordre d'approximation pour $\lambda = 0.9$.

Les maillages utilisés pour ce cas-test ont été obtenus par le logiciel *Triangle* disponible sur le

TABLE 2.3 – Tableau de convergence pour le schéma en dimension deux

mailles	aire max. des cellules	erreur L^2 ($k = 1$)	ordre	erreur L^2 ($k = 2$)	ordre
16677	9.998e-5	3.504e-5	2.01	4.59e-7	3.2
20411	8.000e-5	1.388e-5	1.97	1.54e-7	3.92
33242	5.000e-5	8.731e-6	-	4.98e-8	-

site (http ://www.cs.cmu.edu/˜quake/triangle.html). Ils ont été fait indépendamment les uns des autres, c'est-à-dire qu'ils ne sont pas issus d'un raffinement successif. Comme en dimension un, les ordres numériques correspondent bien aux ordres théoriques et les erreurs en norme L^2 sont très petites.

Transport d'un disque

Ce cas-test a été réalisé avec une vitesse $a = 5$ et une direction $\vec{\Omega} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ sur un maillage non-structuré toujours obtenu par *Triangle*. Le domaine est rectangulaire de longueur 20 et de largeur 2 et comporte 66498 cellules dont la plus petite a une aire égale à $1.95.10^{-4}$. On considère ce long domaine afin d'éviter de gérer des conditions aux limites périodiques. On compare les schémas d'ordre 1, 2 et 3 ($k = 0, 1, 2$). La donnée initiale pour ce cas-test est un disque de rayon 0.5 situé dans la partie gauche du domaine (voir la figure 2.11) :

$$u_0(x, y) = \begin{cases} 1, & \text{si} \quad (x - 0.6)^2 + (y - 1)^2 < 0.25, \\ 0, & \text{sinon.} \end{cases}$$

Pour chaque degré d'approximation k, on teste le schéma avec une CFL de 0.5 et 10. Sur les figures 2.12-2.13-2.14 sont présentées les solutions au temps $t = 3.6\ s$. Afin de mieux visualiser la structure globale du maillage et dans les discontinuités, la figure 2.15 montre la même solution que la figure 2.14 avec le maillage correspondant. Pour une meilleure visibilité de la solution, on fait un zoom sur $x \in [17, 20]$ et on ajoute trois courbes de niveau 0.1, 0.5 et 0.9 sur chaque figure.

On remarque que les résultats sont globalement plus diffusés dans la direction du flux. Le comportement des solutions en fonction du degré k des polynômes d'approximation est très proche de celui observé en dimension un.

Pour pouvoir comparer de manière plus précise les résultats, plusieurs coupes des solutions sont réalisées suivant l'axe $y = 1$ et exposées sur la figure 2.16. De même, pour plus de visibilité, on zoome sur la partie $x \in [17, 20]$.

Pour tester le schéma avec une grande CFL, on se donne toujours un disque pour donnée initiale sur un domaine trois fois plus fin, dont l'aire de la plus petite maille vaut $4.94.10^{-5}$. Sur la figure 2.17, on montre la solution avec une CFL égale à 50 pour $k = 2$.

Puis, pour tester le schéma en temps très long, la dernière figure est réalisée avec une vitesse

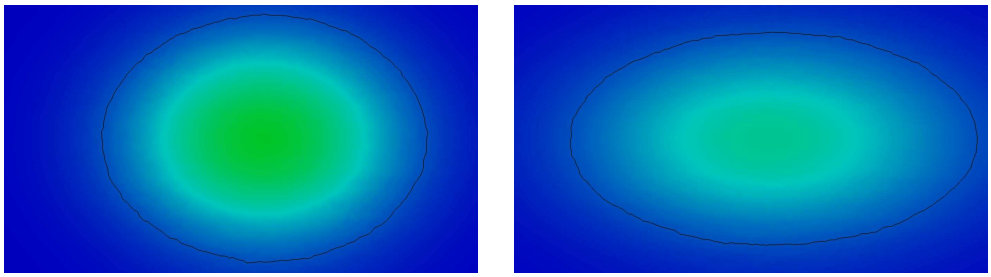


FIGURE 2.12 – Solution zoomée et courbe de niveau de 0.1 pour $k = 0$ avec une CFL de 0.5 (gauche) et 10 (droite)

FIGURE 2.13 – Solution zoomée et courbes de niveau de $0.1,\ 0.5,\ 0.9$ pour $k = 1$ avec une CFL de 0.5 (gauche) et 10 (droite)

FIGURE 2.14 – Solution zoomée et courbes de niveau de $0.1,\ 0.5,\ 0.9$ pour $k = 2$ avec une CFL de 0.5 (gauche) et 10 (droite)

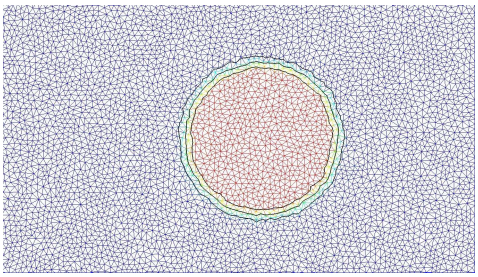


FIGURE 2.15 – Maillage, solution zoomée et courbes de niveau $0.1,\ 0.5,\ 0.9$ pour $k = 2$ avec une CFL de 0.5

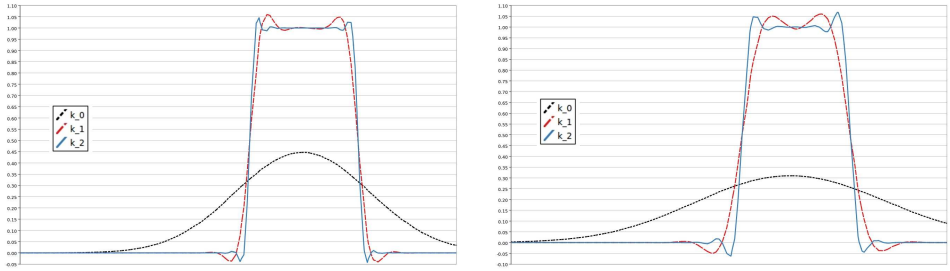


FIGURE 2.16 – Coupes des solutions suivant $y = 1$ avec une CFL $= 0.5$ (gauche) et 10 (droite)

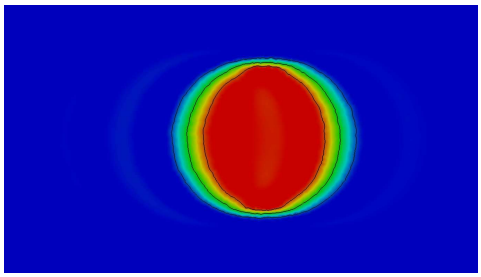


FIGURE 2.17 – Solution zoomée et courbes de niveau de $0.1,\ 0.5,\ 0.9$ pour $k = 2$ avec une CFL de 50

$a = 10^{-3}$, une direction $\vec{\Omega} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et une CFL de 20. On regarde la solution après un temps $t = 18000\ s$. On constate qu'en dépit du temps très long et de la grande CFL considérée, la solution donnée par le schéma GRP espace-temps est très proche de la solution exacte.
En dernière remarque, on s'interroge sur le meilleur choix entre raffinement du maillage et diminution de la CFL pour améliorer les résultats. Pour une approximation de qualité fixée et un nombre de cellules fixe dans le maillage, on recherche la CFL correspondante. On considère de nouveau un disque comme donnée initiale sur un domaine rectangulaire avec une vitesse $a = 5$ et une direction $\vec{\Omega} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Le tableau 2.4 présente trois configurations ainsi que leurs temps de calcul respectifs.
Dans cette configuration, en considérant un maillage quatre fois plus fin, on constate que la CFL est doublée et que le temps de calcul est quadruplé.

Pour tester le schéma avec une donnée initiale moins régulière, le même cas-test est réalisé

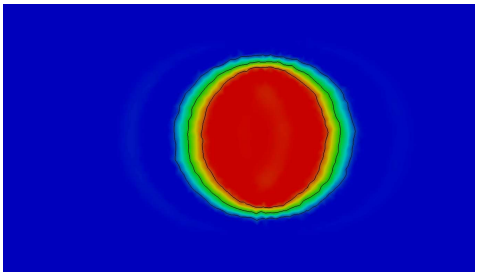


FIGURE 2.18 – Solution zoomée et courbes de niveau de $0.1,\ 0.5,\ 0.9$ pour $k = 2$ à $t = 18000\ s$ avec une CFL de 20 et $a = 10^{-3}$

TABLE 2.4 – Comparaison des temps de calcul pour une approximation de qualité fixée

mailles	CFL	temps CPU
33332	13	13.47
66498	20	25.02
133179	30	50.02

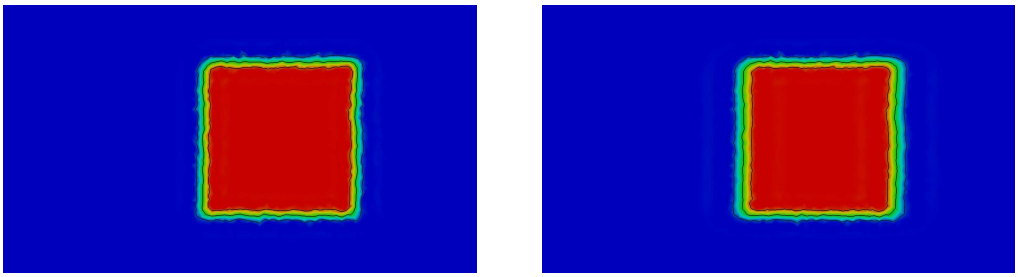


FIGURE 2.19 – Solution zoomée et courbes de niveau de $0.1,\ 0.5,\ 0.9$ pour $k = 2$ avec une CFL de 0.5 (gauche) et 10 (droite)

avec une donnée initiale carrée :

$$u_0(x,y) = \begin{cases} 1, & \text{si} \quad (x,y) \in [0.2, 1.2] \times [0.5, 1.5] \\ 0, & \text{sinon.} \end{cases}$$

Les résultats pour $k = 2$ sont présentés à un temps $t = 3.6s$ sur la figure 2.19. À nouveau, on observe que cette approximation est très proche de la solution exacte et plus particulièrement, on voit que les angles sont préservés.

2.5 Méthode de projection GRP espace-temps pour le modèle S_N en dimension 1

Dans les paragraphes précédents, un schéma numérique pour l'équation de transport basé sur les méthodes de type GRP a été présenté. On a montré que ce schéma qui n'impose aucune restriction sur le pas de temps, est d'ordre $k+1$ en temps et en espace. On souhaite à présent approcher le modèle d'ordonnées discrètes du transfert radiatif détaillé dans le paragraphe 1.3. On remarque que ce modèle s'écrit sous la forme d'un système de transport avec un terme source. Une première idée pour discrétiser ce système serait alors d'introduire directement le terme source dans le schéma GRP espace-temps développé pour les systèmes de transport. En particulier, afin de préserver les régimes limites, on pourrait considérer une méthode qui préserve l'asymptotique (*asymptotic preserving*) [63, 29, 19, 79]. La seconde option que l'on retient dans ce paragraphe et de développer un nouveau schéma de type GRP espace-temps pour le modèle d'ordonnées discrètes S_N. En effet, on remarque que la forme particulière du terme source dans le modèle permet de calculer à nouveau l'intensité radiative exacte. Grâce à cette propriété, l'étape d'évolution dans le schéma pourra donc être exacte.
Par souci de simplicité pour la présentation du schéma, on s'affranchit de la discrétisation fréquentielle en intégrant le modèle sur tout le spectre de fréquence et en supposant l'opacité σ constante. On considère ainsi le problème mixte pour le modèle d'ordonnées discrètes en dimension un sur le domaine $[x_L, x_R]$:

$$\begin{cases} \forall 1 \leq l \leq N \quad \partial_t I_l + c\mu_l \cdot \partial_x I_l = c\sigma(aT^4 - I_l), \quad forall x \in [x_L, x_R], \\ \qquad\qquad I_l(x,0) = I_l^0(x), \\ \qquad\qquad I_l(x = x_L, t) = I_l^L(t) \text{ si } \mu_l > 0, \quad I_l(x = x_R, t) = I_l^R(t) \text{ si } \mu_l < 0, \end{cases} \tag{2.32}$$

avec $I_l = \int_0^{+\infty} I(x,t,\mu_l,\nu)d\nu > 0$ l'intensité radiative et $\mu_l = \cos(\Omega_l)$.
Pour une température T fixée, la solution exacte de (2.32) est connue :

$$I_l(x,t) = aT^4 + \left(I_l^0(x - c\mu_l t) - aT^4\right) e^{-c\sigma t}.$$

La résolution de cette équation est détaillée dans le paragraphe suivant.
Ce système est couplé avec une équation régissant la température matière. Par exemple on considère une équation de la forme :

$$\begin{aligned} \rho C_p(T)\partial_t T &= -c\sigma(E_d - aT^4) + \partial_{xx}(\lambda_c(T)\, T), \\ T(x,0) &= T_0(x), \end{aligned} \tag{2.33}$$

où λ_c est la conductivité et $E_d = \frac{1}{c}\sum_l I_l \omega_l$ est l'énergie radiative discrète du système.

Le schéma GRP espace-temps pour ce système est basé sur le même principe que dans le cas de l'équation linéaire : en supposant connues deux approximations polynomiales de degré k de la solution au temps $t = t^n$, une en espace sur chaque cellule et une en temps sur chaque interface, on fait évoluer de manière exacte ces approximations et on les reprojette dans leurs espaces respectifs. Dans un premier paragraphe, la résolution de l'équation (2.32) est détaillée, puis la méthode GRP espace-temps pour cette équation est présentée dans un second paragraphe.

2.5.1 Résolution de l'équation du transfert radiatif (2.32)

Si l'on fixe une direction dans l'équation (2.32), en omettant l'indice l, on obtient le problème mixte suivant :

$$\begin{cases} \partial_t I + c\mu\partial_x I = c\sigma(aT^4 - I), \\ I(x,0) = I_0(x), \\ I(x = x_L, t) = I^L(t) \text{ si } \mu > 0, \quad I(x = x_R, t) = I^R(t) \text{ si } \mu < 0. \end{cases}$$

On suppose que $c\mu, \sigma$ et T sont constants.
De plus, on suppose que I est assez régulière sur $[x_L; x_R] \times [0, +\infty[$. Soit la fonction f définie à partir de la fonction I par :

$$f(y,s) = I(c\mu s + y, s),$$

donc on a

$$I(x,t) = f(x - c\mu t, t). \tag{2.34}$$

Avec cette définition de f en fonction de I, il suffit alors de connaître f pour en déduire I. On dérive f par rapport à sa seconde variable s :

$$\begin{aligned} \forall y, \quad \partial_s f(y,s) &= \partial_t I(c\mu s + y, s) + c\partial_x I(c\mu s + y, s), \\ &= c\sigma\left(aT^4 - I(c\mu s + y, s)\right), \\ &= c\sigma\left(aT^4 - f(y,s)\right), \end{aligned}$$

la fonction f est ainsi donnée par :

$$f(y,s) = aT^4 + \left(I_0(y) - aT^4\right) e^{-c\sigma s}.$$

Finalement, la solution de (2.32) est :

$$\begin{cases} u(x,t) &= aT^4 + \left(I_0(x-c\mu t) - aT^4\right) e^{-c\sigma t}, \quad \text{si } x - c\mu t \in [x_L, x_R], \\ u(x,t) &= aT^4 + \left(I^\alpha(s_\alpha) - aT^4\right) e^{-c\sigma(s_\alpha - t)}, \quad \text{sinon}, \end{cases} \tag{2.35}$$

avec

$$s_\alpha = \frac{x_\alpha - x}{c\mu} + t \text{ et } \alpha = \begin{cases} R & \text{si } \mu > 0, \\ L & \text{sinon} \end{cases}.$$

Une fois la solution exacte de l'équation (2.32) connue, on développe un schéma GRP espace-temps pour cette équation.

2.5.2 Schéma GRP espace-temps

On cherche à approcher la solution du problème mixte (2.32) grâce un schéma de type GRP espace-temps. Pour alléger les notations, on fixe la direction $\mu = \mu_l$ dans le problème mixte (2.32) :

$$\begin{cases} \partial_t I + c\mu \cdot \partial_x I = c\sigma(aT^4 - I), \\ I(x,0) = I_0(x), \\ I(x = x_L, t) = I^L(t) \text{ si } \mu > 0, \quad I(x = x_R, t) = I^R(t) \text{ si } \mu < 0. \end{cases} \tag{2.36}$$

Dans un premier temps, T est supposée constante pour simplifier la présentation du schéma.
Comme précédemment, le domaine est discrétisé par un maillage uniforme formé des cellules $C_i = (x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$ avec $x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} = \Delta x$. Les interfaces $I^{n+\frac{1}{2}} = [t^n, t^{n+1}]$ permettent de définir les volumes de contrôle espace-temps : $K_i^n = C_i \times I^{n+\frac{1}{2}}$.

Le schéma consiste à profiter de la connaissance des solutions exactes du système (2.36). On fait évoluer de manière exacte une approximation polynomiale par maille de la solution ainsi qu'une seconde approximation temporelle polynomiale de la solution qui est utilisé comme un outil dans la construction du schéma. Finalement, on considère une étape de projection de ces solutions dans leurs espaces d'approximation respectifs.

Approximations considérées : À la date t^n, on suppose connue une approximation polynomiale de degré k par maille de la solution en espace :

$$\text{pour } t = t^n, \quad I(x,t^n) \simeq u_i^n(x) \in \mathbb{R}_k^i \text{ si } x \in C_i.$$

On introduit également une seconde approximation polynomiale de degré k auxiliaire de la solution en temps :

$$\text{pour } t \in I^{n+\frac{1}{2}}, \quad I(x_{i-\frac{1}{2}}, t) \simeq v_{i-\frac{1}{2}}^{n+\frac{1}{2}}(t) \in \mathbb{R}_k^n,$$

où les espaces de polynômes (2.4)-(2.13) sont les mêmes que dans le schéma pour l'équation de transport. Une fois ces notations établies, on fait évoluer de manière exacte l'approximation de la solution en espace ainsi que son approximation temporelle auxiliaire pour obtenir une approximation de la solution au temps $t^n + \Delta t$.

Étape d'évolution : à cette étape, on suppose donc que l'on doit résoudre le problème mixte suivant :

$$\begin{cases} \partial_t I + c\mu \cdot \partial_x I = c\sigma(aT^4 - I), \\ I(x,t^n) = u^n(x), \\ v_{\frac{1}{2}}(t) = I^L(t) \text{ si } \mu > 0, \quad v_{N+\frac{1}{2}}(t) = I^R(t) \text{ si } \mu < 0. \end{cases} \tag{2.37}$$

Comme la solution exacte de ce système (2.37) a été calculée dans le paragraphe précédent, on peut faire évoluer de manière exacte les approximations polynomiales de la solution par la méthode des caractéristiques. Dans le cas $\mu > 0$ et $\lambda = \frac{c\mu\Delta t}{\Delta x} < 1$ (cas explicite), la solution exacte en espace est donnée par :

$$u_i^h(x,t^n+\Delta t) = \begin{cases} \left[v_{i-\frac{1}{2}}^{n+\frac{1}{2}}\left(t^{n+1}-\frac{x-x_{i-\frac{1}{2}}}{c\mu}\right)-aT^4\right]e^{-c\sigma\left(\frac{x-x_{i-\frac{1}{2}}}{c\mu}\right)}+aT^4, & \text{si } (x-x_{i-\frac{1}{2}})<c\mu\Delta t, \\ \left[u_i^n(x-c\mu\Delta t)-aT^4\right]e^{-c^2\mu\sigma\Delta t}+aT^4, & \text{sinon,} \end{cases}$$

et la solution temporelle auxiliaire exacte est :

$$v_h^{n+\frac{1}{2}}(t,x_{i+\frac{1}{2}}) = \left[u_i^n(x_{i+\frac{1}{2}}-c\mu(t-t^n))-aT^4\right]e^{-c\sigma(t-t^n)}+aT^4.$$

Pour $\lambda > 1$ (cas implicite), elles sont données par :

$$u_i^h(x,t^n+\Delta t) = \left[v_{i-\frac{1}{2}}^{n+\frac{1}{2}}\left(t^{n+1}-\frac{x-x_{i-\frac{1}{2}}}{c\mu}\right)-aT^4\right]e^{-\sigma\left(\frac{x-x_{i-\frac{1}{2}}}{\mu}\right)}+aT^4,$$

$$v_h^{n+\frac{1}{2}}(t,x_{i+\frac{1}{2}}) = \begin{cases} \left[u_i^n(x_{i+\frac{1}{2}}-c\mu(t-t^n))-aT^4\right]e^{-c\sigma(t-t^n)}+aT^4, & \text{si } (t-t^n)>\frac{\Delta x}{c\mu}, \\ \left[v_{i-\frac{1}{2}}^{n+\frac{1}{2}}\left(t-\frac{\Delta x}{c\mu}\right)-aT^4\right]e^{-\sigma\frac{\Delta x}{\mu}}+aT^4, & \text{sinon.} \end{cases}$$

Les solutions exactes étant connues, il reste alors à les projeter dans l'espace d'approximation $\mathbb{P}_C^k$.

Étape de projection : les solutions exactes sont projetées dans leurs espaces respectifs : $\mathbb{P}_C^k$ et $\mathbb{P}_I^k$. La solution exacte en espace est projetée sur l'espace $\mathbb{P}_C^k$ sur la cellule C_i au temps t^{n+1} et la solution exacte temporelle est projetée sur l'espace $\mathbb{P}_I^k$ sur l'interface $x_{i+\frac{1}{2}} \times (t^n, t^{n+1})$. Ces projections se traduisent par la résolution des deux problèmes de minimisation suivants :

$$\|u_i^{n+1} - u_i^h(t^n+\Delta t)\|_i = \inf_{p\in\mathbb{P}_C^k} \|p - u_i^h(t^n+\Delta t)\|_i, \tag{2.38}$$

$$\|v_{i+\frac{1}{2}}^{n+\frac{1}{2}} - v_h^{n+\frac{1}{2}}(x_{i+\frac{1}{2}})\|^n = \inf_{q\in\mathbb{P}_I^k} \|q - v_h^{n+\frac{1}{2}}(x_{i+\frac{1}{2}})\|^n. \tag{2.39}$$

Grâce aux conditions de Petrov-Galerkin, ces problèmes de minimisation (2.38)-(2.39) peuvent se réécrire de la façon suivante :

$$\forall l = 0,\dots,k \quad <u_i^{n+1} - u_i^h(t^n+\Delta t), \left(\frac{x-x_i}{\Delta x}\right)^l >_i = 0, \tag{2.40}$$

$$\forall r = 0,\dots,k \quad <v_{i+\frac{1}{2}}^{n+\frac{1}{2}} - v_h^{n+\frac{1}{2}}(x_{i+\frac{1}{2}}), \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r >^n = 0. \tag{2.41}$$

Puis en remplaçant u_i^{n+1} et $v_{i+\frac{1}{2}}^{n+\frac{1}{2}}(t)$ par leur décomposition dans leur base de polynômes respectifs :

$$u_i^{n+1}(x) = \sum_{j=0}^{k} \alpha_i^{n+1,j}\left(\frac{x-x_i}{\Delta x}\right)^j,$$

$$v_{i+\frac{1}{2}}^{n+\frac{1}{2}}(t) = \sum_{s=0}^{k} \beta_{i+\frac{1}{2}}^{n+\frac{1}{2},s}\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^s,$$

les problèmes de minimisation (2.40)-(2.41) reviennent à résoudre les deux systèmes linéaires de taille $k+1$ suivants :

$$\forall l=0,\dots,k \quad \sum_{j=0}^{k}\alpha_i^{n+1,j} < \left(\frac{x-x_i}{\Delta x}\right)^j, \left(\frac{x-x_i}{\Delta x}\right)^l >_i =< u_i^h(t^n+\Delta t), \left(\frac{x-x_i}{\Delta x}\right)^l >_i, \tag{2.42}$$

$$\forall r=0,\dots,k \quad \sum_{s=0}^{k}\beta_{i+\frac{1}{2}}^{n+\frac{1}{2},s} < \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^s, \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r >^n =< v_h^{n+\frac{1}{2}}(x_{i+\frac{1}{2}}), \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r >^n . \tag{2.43}$$

Les membres de gauche dans les systèmes (2.42)-(2.43) peuvent se réécrire :

$$\begin{aligned} < \left(\frac{x-x_i}{\Delta x}\right)^j, \left(\frac{x-x_i}{\Delta x}\right)^l >_i &= \frac{1}{j+l+1}\left(\frac{1}{2}\right)^{j+l+1}\left[1-(-1)^{j+l+1}\right], \\ &= A_u(j,l), \\ < \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^s, \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r >^n &= \frac{1}{s+r+1}\left(\frac{1}{2}\right)^{s+r+1}\left[1-(-1)^{s+r+1}\right], \\ &= A_v(s,r). \end{aligned}$$

On remarque que les matrices A_u et A_v sont identiques aux matrices du schéma pour l'équation d'advection, donc toujours indépendantes de la CFL λ.
En supposant que $\mu>0$, les seconds membres dans le cas explicite sont donnés par :

$$\begin{aligned} &< u_i^h(t^n+\Delta t), \left(\frac{x-x_i}{\Delta x}\right)^l >_i = \\ &\quad \frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i-\frac{1}{2}}+c\mu\Delta t}\left[\left(\sum_{j=0}^{k}\beta_{i-\frac{1}{2}}^{n+\frac{1}{2},j}\left(\frac{1}{2}-\frac{x-x_{i-\frac{1}{2}}}{c\mu\Delta t}\right)^j - aT^4\right)e^{-c\sigma\left(\frac{x-x_{i-\frac{1}{2}}}{c\mu}\right)}+aT^4\right]\left(\frac{x-x_i}{\Delta x}\right)^l dx \\ &\quad + \frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}+c\mu\Delta t}^{x_{i+\frac{1}{2}}}\left[\left(\sum_{j=0}^{k}\alpha_i^{n,j}\left(\frac{x-c\mu\Delta t-x_i}{\Delta x}\right)^j - aT^4\right)e^{-c^2\sigma\mu\Delta t}+aT^4\right]\left(\frac{x-x_i}{\Delta x}\right)^l dx, \end{aligned} \tag{2.44}$$

$$\begin{aligned} &< v_h^{n+\frac{1}{2}}(x_{i+\frac{1}{2}}), \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r >^n = \\ &\quad \frac{1}{\Delta t}\int_{t^n}^{t^{n+1}}\left[\left(\sum_{s=0}^{k}\alpha_i^{n,s}\left(\frac{1}{2}-\frac{c\mu(t-t^n)}{\Delta x}\right)^j - aT^4\right)e^{-c\sigma(t-t^n)}+aT^4\right]\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r dt. \end{aligned} \tag{2.45}$$

On réécrit les exponentielles sous la forme d'un produit de deux exponentielles pour faire apparaître la variable $\left(\frac{x-x_i}{\Delta x}\right)$:

$$\begin{aligned} e^{-c\sigma\left(\frac{x-x_{i-1/2}}{c\mu}\right)} &= e^{-\frac{\sigma\Delta x}{\mu}\left(\frac{x-x_i}{\Delta x}\right)}e^{-\frac{\sigma\Delta x}{2\mu}}, \\ e^{-c\sigma(t-t^n)} &= e^{-c\sigma\Delta t\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)}e^{-c\sigma\frac{\Delta t}{2}}. \end{aligned}$$

La principale différence par rapport au cas transport provient de ces termes exponentiels qui apparaissent dans les intégrales. Pour simplifier ces calculs, on approche l'exponentielle par la somme de sa série entière jusqu'à l'ordre $k+1$:

$$e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!} \simeq \sum_{n=0}^{k+1} \frac{x^n}{n!}. \tag{2.46}$$

De cette manière, tous les termes dans les intégrales seront approchés par des termes polynomiaux. En injectant ces approximations dans (2.44)-(2.45), les seconds membres des systèmes linéaires pour le cas explicite ($\lambda < 1$) sont approchés par :

$$\begin{aligned}
&< u_i^h(t^n + \Delta t), \left(\frac{x - x_i}{\Delta x}\right)^l >_i = \\
&\qquad \sum_{j=0}^{k} \left[e^{-\frac{\sigma \Delta x}{2\mu}} \beta_{i-\frac{1}{2}}^{n+\frac{1}{2},j} \sum_{p=0}^{k+1} \left(\frac{-\sigma \Delta x}{\mu}\right)^p \frac{1}{p!} J_{ex}^1(j, l+p) + e^{-c^2 \sigma \mu \Delta t} \alpha_i^{n,j} J_{ex}^2(j,l) \right] \\
&\qquad - aT^4 e^{-\frac{\sigma \Delta x}{2\mu}} \sum_{p=0}^{k+1} \left(\frac{-\sigma \Delta x}{\mu}\right)^p \frac{1}{p!} \frac{1}{p+l+1} \left[\left(-\frac{1}{2} + \frac{c\mu \Delta t}{\Delta x}\right)^{p+l+1} - \left(\frac{1}{2}\right)^{p+l+1} \right] \\
&\qquad - aT^4 e^{-c^2 \sigma \mu \Delta t} \frac{1}{l+1} \left[\left(\frac{1}{2}\right)^{l+1} - \left(-\frac{1}{2} + \frac{c\mu \Delta t}{\Delta x}\right)^{l+1} \right] \\
&\qquad + aT^4 \frac{\left(1 - (-1)^{l+1}\right)}{l+1} \left(\frac{1}{2}\right)^{l+1} \\
&\qquad = b_u(l), \\
&< v_h^{n+\frac{1}{2}}(x_{i+\frac{1}{2}}), \left(\frac{t - t^{n+\frac{1}{2}}}{\Delta t}\right)^r >^n = \\
&\qquad e^{-\frac{c\sigma \Delta t}{2}} \sum_{p=0}^{k} \frac{(-c\sigma \Delta t)^p}{p!} \left[\sum_{s=0}^{k} \alpha_i^{n,s} K_{ex}(s, p+r) - aT^4 \frac{\left(1 - (-1)^{p+r+1}\right)}{p+r+1} \left(\frac{1}{2}\right)^{p+r+1} \right] \\
&\qquad + aT^4 \frac{\left(1 - (-1)^{p+r+1}\right)}{p+r+1} \left(\frac{1}{2}\right)^{p+r+1} \\
&\qquad = b_v(r),
\end{aligned}$$

où les intégrales J_{ex}^1, J_{ex}^2 et K_{ex} sont les mêmes que dans le schéma pour l'équation de transport, en remplaçant la vitesse a par $c\mu$. Dans le cas implicite ($\lambda > 1$), en supposant toujours $\mu > 0$, les seconds membres s'écrivent :

$$\begin{aligned}
&< u_i^h(t^n + \Delta t), \left(\frac{x - x_i}{\Delta x}\right)^l >_i = \\
&\frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \left[\left(\sum_{j=0}^{k} \beta_{i-\frac{1}{2}}^{n+\frac{1}{2},j} \left(\frac{1}{2} - \frac{x - x_{i-\frac{1}{2}}}{c\mu \Delta t}\right)^j - aT^4 \right) e^{-c\sigma \left(\frac{x - x_{i-\frac{1}{2}}}{c\mu}\right)} + aT^4 \right] \left(\frac{x - x_i}{\Delta x}\right)^l dx,
\end{aligned}$$

$$< v_h^{n+\frac{1}{2}}(x_{i+\frac{1}{2}}), \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r >^n =$$

$$\frac{1}{\Delta t}\int_{t^n}^{t^n+\frac{\Delta x}{c\mu}} \left[\left(\sum_{j=0}^{k}\alpha_i^{n,j}\left(\frac{1}{2}-\frac{c\mu(t-t^n)}{\Delta x}\right)^j - aT^4\right)e^{-c\sigma(t-t^n)} + aT^4\right]\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r dt$$

$$+\frac{1}{\Delta t}\int_{t^n+\frac{\Delta x}{c\mu}}^{t^{n+1}} \left[\left(\sum_{j=0}^{k}\beta_{i-\frac{1}{2}}^{n+\frac{1}{2},j}\left(\frac{t-\frac{\Delta x}{c\mu}-t^{n+\frac{1}{2}}}{\Delta t}\right)^j - aT^4\right)e^{-c\sigma\frac{\Delta x}{c\mu}} + aT^4\right]\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r dt.$$

Comme pour le cas explicite, les termes exponentiels sont approchés par la somme de leur développement en série jusqu'à l'ordre $k+1$ et réécrits pour faire apparaître respectivement les variables $\left(\frac{x-x_i}{\Delta x}\right)$ et $\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)$. Les seconds membres deviennent :

$$\begin{aligned}
< u_i^h(t^n+\Delta t), \left(\frac{x-x_i}{\Delta x}\right)^l >_i &= e^{-\frac{\sigma\Delta x}{2\mu}}\sum_{j=0}^{k}\beta_{i-\frac{1}{2}}^{n+\frac{1}{2},j}\sum_{p=0}^{k+1}\frac{1}{p!}J_{im}(j,p+l)\left(-\frac{\sigma\Delta x}{\mu}\right)^p \\
&\quad - e^{-\frac{\sigma\Delta x}{2\mu}}aT^4\sum_{p=0}^{k+1}\left(-\frac{\sigma\Delta x}{\mu}\right)^p\frac{1}{p!(p+l+1)}\left(\frac{1}{2}\right)^{l+p+1}(1-(-1)^{l+p+1}) \\
&\quad + aT^4\frac{1}{l+1}\left(\frac{1}{2}\right)^{l+1}(1-(-1)^{l+1}) \\
&= b_u(l), \\
< v_h^{n+\frac{1}{2}}(x_{i+\frac{1}{2}}), \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^r >^n &= \sum_{s=0}^{k}\Big[e^{-\frac{a\sigma\Delta t}{2}}\alpha_i^{n,s}\sum_{p=0}^{k+1}\frac{(-c\sigma\Delta t)^p}{p!}K_{im}^1(s,p+r) \\
&\quad + e^{-\frac{\sigma\Delta x}{\mu}}\beta_{i-\frac{1}{2}}^{n+\frac{1}{2},s}K_{im}^2(s,r)\Big] \\
&\quad - e^{-\frac{c\sigma\Delta t}{2}}aT^4\sum_{p=0}^{k+1}\frac{(-c\sigma\Delta t)^p}{p!(p+r+1)}\left[\left(\frac{\Delta x}{c\mu\Delta t}-\frac{1}{2}\right)^{p+r+1}-\left(\frac{1}{2}\right)^{p+r+1}\right] \\
&\quad - aT^4e^{-\frac{\sigma\Delta x}{\mu}}\frac{1}{r+1}\left[\left(\frac{1}{2}\right)^{r+1}-\left(\frac{\Delta x}{c\mu\Delta t}-\frac{1}{2}\right)^{r+1}\right] \\
&\quad + aT^4\frac{1}{r+1}\left(\frac{1}{2}\right)^{r+1}(1-(-1)^{r+1}) \\
&= b_v(r),
\end{aligned}$$

où les coefficients $J_{im}, K_{im}^1, K_{im}^2$ sont les mêmes que dans le schéma pour l'équation d'advection en remplaçant la vitesse a par $c\mu$.
Notons que dans le cas où la direction μ est négative, il suffit de remplacer μ par $|\mu|$ dans les intégrales et les exponentielles impliquées dans les schémas explicites et implicites.

Finalement, les étapes du schéma sont identiques au cas linéaire :

- Au temps initial, la solution $u_i^0(x)$ est connue sur chaque cellule du maillage, et la solution $v^{\frac{1}{2}}(t)$ est connue sur la frontière du domaine correspondante (gauche si $\mu > 0$, droite si $\mu < 0$).
- Si $\mu > 0$, le maillage est parcouru de la gauche vers la droite, et si $\mu < 0$, le maillage est parcouru de la droite vers la gauche.
- Dans chaque volume espace-temps K_i^n, il faut résoudre deux systèmes de taille $k+1$ pour

calculer les solutions $u_i^{n+1}(x)$ et $v_{i+\frac{1}{2}}^{n+\frac{1}{2}}(t)$ si $a > 0$:

$$A_u \alpha_i^{n+1} = b_u^+, \tag{2.47}$$

$$A_v \beta_{i+\frac{1}{2}}^{n+\frac{1}{2}} = b_v^+, \tag{2.48}$$

ou $v_{i-\frac{1}{2}}^{n+\frac{1}{2}}(t)$ si $a < 0$:

$$A_u \alpha_i^{n+1} = b_u^-,$$
$$A_v \beta_{i-\frac{1}{2}}^{n+\frac{1}{2}} = b_v^-.$$

À la fin de l'itération, les deux approximations $u^{n+1}(x)$ et $v^{n+\frac{1}{2}}(t)$ de la solution de (2.36) sont bien dans leurs espaces de polynômes respectifs $\mathbb{P}_C^k$ et $\mathbb{P}_I^k$ par construction.

Comme les matrices A_u et A_v des systèmes linéaires (2.42)-(2.43) sont les mêmes que dans le cas d'advection, elles peuvent de nouveau être calculées et inversées une seule fois au début de l'algorithme.

Ce schéma numérique est ainsi simple à écrire et à implémenter. De plus, l'approximation d'ordre $k+1$ des exponentielles permet de passer facilement du cas linéaire au cas linéaire avec terme source. En effet, les mêmes intégrales J et K apparaissent dans l'expression des seconds membres des systèmes linéaires.
En pratique, la température T n'est pas constante. En effet, elle est régie par l'équation (2.33) et varie donc au cours du temps. Pour les applications numériques, on supposera que la température reste constante au cours d'un pas de temps Δt raisonnable devant l'échelle considérée. À la fin de chaque itération, il faut donc mettre à jour aT^4 pour l'intégrer dans le schéma GRP espace-temps. Pour cela, on discrétise l'équation en température (2.33) avec un schéma d'Euler semi-implicite. En supposant T suffisamment régulière, l'opérateur en espace qui apparaît dans le second membre est discrétisé par une méthode de différences finies :

$$\partial_{xx}(\lambda_c\, T) \simeq \frac{T_{i+1} - 2T_i + T_{i-1}}{(\Delta x)^2}.$$

Le schéma d'Euler semi-implicite est alors :

$$T_i^{n+1} = T_i^n + \frac{\Delta t c \sigma}{\rho C_p(T_i^{n+1})}\left(E_i^n - a(T_i^{n+1})^4\right) + \lambda_c(T_i^{n+1})\Delta t \frac{T_{i+1}^n - 2T_i^n + T_{i-1}^n}{(\Delta x)^2 \rho C_p(T_i^{n+1})}$$

Une méthode de Newton peut être utilisée pour résoudre l'équation non-linéaire ainsi obtenue.

Finalement, pour illustrer ce schéma, un cas-test portant sur l'évolution d'un véhicule de rentrée atmosphérique est présenté dans la section 5.1.

2.6 Schémas numériques existants

Dans ce paragraphe, on rappele le principe de trois méthodes numériques considérées à titre de comparaison : le schéma décentré amont pour discrétiser le modèle d'ordonnées discrètes, le schéma DG (Galerkin Discontinu) et WENO (Weighted Essentially Non-Oscillatory) discrétisant l'équation de transport. Ces schémas ont servi notamment à comparer les résultats numériques donnés par le schéma GRP espace-temps, tant au niveau précision qu'au niveau des performances des algorithmes.

2.6.1 Schéma décentré amont pour le modèle d'ordonnées discrètes

Dans ce paragraphe, on présente un schéma décentré amont pour le modèle d'ordonnées discrètes S_N en dimension 2. On rappelle que ce modèle, décrit dans le paragraphe 1.3, est donné par le système d'équations suivant :

$$\forall 1 \leq l \leq N, 1 \leq q \leq Q, \begin{cases} \partial_t I_{l,q}^{S_N} + c\Omega_l . \nabla I_{l,q}^{S_N} = c\sigma \left(\mathcal{B}_q - I_{l,q}^{S_N} \right), \\ \rho C_v \partial_t T = -c\sigma \sum_{l,q} \left(\mathcal{B}_q - I_{l,q} \right) \omega_l \theta_q, \end{cases} \tag{1.11}$$

où Ω_l est une direction fixée.
On considère un maillage non-structuré formé des polygones K. L'aire de la cellule K est notée $|K|$ et l'ensemble des segments frontières de K est noté Γ_K. L'intensité radiative $I_{l,q}$ est supposée connue sur chaque maille K du maillage au temps $t = t^n$.
Pour alléger les notations, on omet les indices l, q de l'intensité radiative I, soit $I = I_{l,q}$.
Le schéma volumes finis explicite pour le modèle (1.11) est donné par :

$$\left(I_K^{n+1} - I_K^n \right) |K| + c\Delta t \sum_{\varrho \in \Gamma_K} |\varrho_{KL}| \mathcal{F}_{KL}^n \cdot \overrightarrow{n}_{KL} = c\sigma \Delta t \left(\mathcal{B}(T_K^n) - I_K^n \right),$$

$$\rho C_v \frac{T_K^{n+1} - T_K^n}{\Delta t} = -c\sigma \sum_{k,l} \left(\mathcal{B}(T_K^n) - I_k^n \right) \theta_l \omega_q,$$

avec

$$\begin{aligned} \varrho_{KL} &: \text{segment séparant la maille } K \text{ de la maille } L, \\ \mathcal{F}_{KL} \cdot \overrightarrow{n}_{KL} &: \text{le flux sortant de la maille } K \text{ vers la maille } L, \\ \overrightarrow{n}_{KL} &: \text{la normale sortant de } \varrho_{KL}. \end{aligned}$$

Le produit scalaire flux-normale peut se réécrire :

$$\mathcal{F}_{KL}^n \cdot \overrightarrow{n}_{KL} = \int_K \Omega_l \cdot \nabla I = \int_{\Gamma_K} \Omega_l \cdot n I d\sigma,$$

avec $\Omega_l = \begin{pmatrix} \Omega_l^x \\ \Omega_l^y \end{pmatrix}$ soit $\Omega_l \cdot n = \Omega_l^x n_x + \Omega_l^y n_y$.
Le flux décentré amont s'écrit alors :

$$\mathcal{F}_{KL}^n \cdot \overrightarrow{n}_{KL} = \begin{cases} (\Omega_l^x n_x + \Omega_l^y n_y) I_K & \text{si } \Omega_l \cdot \overrightarrow{n}_{KL} > 0, \\ (\Omega_l^x n_x + \Omega_l^y n_y) I_L & \text{sinon.} \end{cases}$$

Pour l'équation en température, on utilise un schéma d'Euler explicite. Finalement, le schéma pour le système (1.11) s'écrit :

$$I_K^{n+1} = I_K^n \left(1 - \frac{c\sigma\Delta t}{|K|} \right) - \frac{c\Delta t}{|K|} \left(\sigma \mathcal{B}(T_K^n) - \sum_{\varrho \in \Gamma_K} |\varrho_{KL}| \mathcal{F}_{KL}^n \cdot \overrightarrow{n} \right),$$

$$T_K^{n+1} = T_K^n - \frac{\Delta t}{\rho C_v} \sigma \sum_{l,q} \left(\mathcal{B}_q(T_K^n) - I_{K(l,q)}^n \right) \omega_l \theta_q$$

2.6.2 Schéma Galerkin discontinu pour l'équation de transport linéaire

Les méthodes de Galerkin discontinues permettent d'allier les méthodes d'éléments finis et de volumes finis. Comme dans les méthodes classiques d'éléments finis, on cherche la solution

dans un espace de polynômes. L'ordre d'approximation de la solution en espace dépend ainsi du choix du degré des polynômes. Toutefois et à la différence des méthodes d'éléments finis usuelles, les schémas Garlerkin discontinus n'imposent pas la continuité de la solutions aux interfaces. Cette propriété permet de mieux gérer les discontinuités qui peuvent apparaître au cours du temps lorsque des problèmes hyperboliques sont considérés.
La méthode présentée ici est celle décrite par Cockburn et Shu dans [34] dans le cas particulier de l'équation linéaire en dimension un d'espace. Les calculs seront donnés de manière générale et détaillés pour $k = 0, 1$ et 2, k étant le degré des polynômes d'approximation.
On considère l'équation de transport en dimension un :

$$\partial_t u + a\partial_x u = 0, \tag{2.49}$$

avec a une vitesse constante et la donnée initiale :

$$u(x, 0) = u_0(x). \tag{2.50}$$

Le domaine I est discrétisé en sous-intervalles $C_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ de même longueur Δx.
L'espace d'approximation de la solution, dans lequel aucune continuité n'est imposée aux interfaces, est donné par :

$$V_h^k = \left\{ p \in BV(\mathbb{R}) : p_{|C_j} \in P^k(C_j) \right\}. \tag{2.51}$$

où $BV(\mathbb{R})$ est l'ensemble des fonctions à variation bornée. À cet espace est associée une base de polynômes orthogonaux $\{v_l^{(j)}(x), l = 0, \ldots, k\}$ sur chaque intervalle C_j telle que :

$$\int_{C_j} v_l^{(j)}(x) v_p^{(j)}(x) dx = e_l \delta_{lp}, \quad \text{et} \quad e_l \neq 0,$$

où δ_{lp} désigne le symbole de Kronecker.
À titre d'exemple, on peut considérer la base des polynômes de Legendre :

$$\begin{aligned} v_0^{(j)}(x) &= 1, \\ v_1^{(j)}(x) &= x - x_j, \\ v_2^{(j)}(x) &= (x - x_j)^2 - \frac{1}{12}\Delta x^2. \end{aligned} \tag{2.52}$$

Soit $u^h(x, t)$ une approximation de la solution $u(x, t)$ dans V_h^k. Cette approximation $u^h(x, t)$ se décompose dans la base des polynômes :

$$u^h(x, t) = \sum_{l=0}^{k} c_l u_j^{(j)}(t) v_l^{(j)}(x) \qquad \text{pour } x \in C_j \tag{2.53}$$

où les coefficients sont donnés par :

$$\begin{aligned} u_j^{(l)} = u_j^{(l)}(t) &= \frac{1}{\Delta x^{l+1}} \int_{C_j} u^h(x, t) v_l^{(j)}(x) dx, \\ c_l &= \frac{\Delta x^{l+1}}{\int_{C_j} (v_l^{(j)}(x))^2 dx} \quad \forall l = 0, \ldots, k. \end{aligned} \tag{2.54}$$

L'approximation $u^h(x, t)$ est donc entièrement déterminée par la donnée des coefficients $u_j^{(l)}$.
Par ailleurs, on cherche $u^h(x, t)$ comme solution du problème variationnel approché suivant :

$$\int_{C_j} \left(\frac{\partial}{\partial t} u^h(x, t) \right) v^h(x) dx + \int_{C_j} \left(a \frac{\partial}{\partial x} u^h(x, t) \right) v^h(x) dx = 0 \quad \forall v^h \in V_h^k. \tag{2.55}$$

Pour alléger les notations, on pose $u_{j+\frac{1}{2}} = u^h(x_{j+\frac{1}{2}}, t)$.
Après intégration par parties, le système (2.55) devient :

$$\frac{\partial}{\partial t}u_j^{(l)} + \frac{1}{\Delta x^{l+1}}\left[\Delta_+(v_l^{(j)}(x_{j-\frac{1}{2}})au_{j-\frac{1}{2}})\right] - \frac{1}{\Delta x^{l+1}}\int_{C_j} au^h(x,t)\frac{\partial}{\partial x}v_l^{(j)}(x)dx = 0, \quad \forall l = 0,\dots,k \tag{2.56}$$

avec :

$$\Delta_+ v_j = v_{j+1} - v_j.$$

Il reste alors à caractériser $au_{j-\frac{1}{2}}$, car la solution n'est, *a priori*, pas définie sur l'interface $x_{j-\frac{1}{2}}$.
Pour cela, on introduit un flux h approchant au définie de la manière suivante :

$$\begin{aligned} au_{j+\frac{1}{2}} &= h(u^-_{j+\frac{1}{2}}, u^+_{j+\frac{1}{2}}) \\ &= h_{j+\frac{1}{2}}. \end{aligned}$$

Ainsi, le schéma (2.56) devient :

$$\frac{d}{dt}u_j^{(l)} + \frac{1}{\Delta x^{l+1}}\left[\Delta_+(v_l^{(j)}(x_{j-\frac{1}{2}})h_{j-\frac{1}{2}})\right] - \frac{1}{\Delta x^{l+1}}\int_{C_j} au^h(x,t)\frac{d}{dx}v_l^{(j)}(x)dx = 0 \quad \forall l = 0,\dots,k. \tag{2.57}$$

Puis, par définition de $u^h(x,t)$, donnée par (2.53), la solution de chaque côté de l'interface $u^-_{j+\frac{1}{2}}$ et $u^+_{j+\frac{1}{2}}$ est donnée par :

$$\begin{aligned} &k=0 \qquad u^-_{j+\frac{1}{2}} = u^+_{j+\frac{1}{2}} = u_j^{(0)}, \\ &k=1 \qquad u^-_{j+\frac{1}{2}} = u_j^{(0)} + 6u_j^{(1)} \quad ; \quad u^+_{j+\frac{1}{2}} = u_j^{(0)} - 6u_j^{(1)}, \\ &k=2 \qquad u^-_{j+\frac{1}{2}} = u_j^{(0)} + 6u_j^{(1)} + 30u_j^{(2)} \quad ; \quad u^+_{j+\frac{1}{2}} = u_j^{(0)} - 6u_j^{(1)} + 30u_j^{(2)}. \end{aligned}$$

En remplaçant les expressions (2.57)-(2.52), les trois premières équations s'écrivent :

$$\begin{aligned} &\frac{d}{dt}u_j^{(0)} + \frac{1}{\Delta x}(h_{j+\frac{1}{2}} - h_{j-\frac{1}{2}}) = 0, \\ &\frac{d}{dt}u_j^{(1)} + \frac{1}{2\Delta x}(h_{j+\frac{1}{2}} + h_{j-\frac{1}{2}}) - \frac{1}{\Delta x^2}\int_{C_j} au^h(x,t)dx = 0, \\ &\frac{d}{dt}u_j^{(2)} + \frac{1}{6\Delta x}(h_{j+\frac{1}{2}} - h_{j-\frac{1}{2}}) - \frac{2}{\Delta x^3}\int_{C_j} au^h(x,t)v_1^{(j)}(x)dx = 0. \end{aligned} \tag{2.58}$$

Grâce à la linéarité de l'équation (2.49), les intégrales peuvent être calculées explicitement. Notons que dans le cas d'un flux $f(u)$ non linéaire, les intégrales sont en général approchées par des méthodes de quadrature.
En remplaçant $u^h(x,t)$ par sa décomposition dans la base des polynômes (2.53), le système (2.58) devient :

$$\begin{aligned} &\frac{d}{dt}u_j^{(0)} + \frac{1}{\Delta x}(h_{j+\frac{1}{2}} - h_{j-\frac{1}{2}}) = 0, \\ &\frac{d}{dt}u_j^{(1)} + \frac{1}{2\Delta x}(h_{j+\frac{1}{2}} + h_{j-\frac{1}{2}}) - \frac{1}{\Delta x}au_j^{(0)} = 0, \\ &\frac{d}{dt}u_j^{(2)} + \frac{1}{6\Delta x}(h_{j+\frac{1}{2}} - h_{j-\frac{1}{2}}) - \frac{2}{\Delta x}au_j^{(1)} = 0, \end{aligned}$$

où $u_j^{(0)}$ et $u_j^{(1)}$ sont définis par (2.54).
Afin d'assurer des propriétés de stabilité, les états $u^\pm_{j+\frac{1}{2}}$ sont modifiés en introduisant des limiteurs locaux de projection :

$$u^-_{j+\frac{1}{2}} = u_j^{(0)} + \tilde{u}_j, \qquad u^+_{j-\frac{1}{2}} = u_j^{(0)} - \tilde{\tilde{u}}_j. \tag{2.59}$$

En remplaçant $u_j^{(0)}$ dans leur décomposition (2.53) par (2.59), on en déduit la valeur des coefficients $\tilde{u}_j$ et $\tilde{\tilde{u}}_j$:

$$\tilde{u}_j = \sum_{l=1}^{k} c_l u_j^{(l)} v_l^{(j)}(x_{j+\frac{1}{2}}), \qquad \tilde{\tilde{u}}_j = \sum_{l=1}^{k} c_l u_j^{(l)} v_l^{(j)}(x_{j-\frac{1}{2}}), \tag{2.60}$$

soit pour les trois premiers ordres :

$$\begin{aligned} k=0 \quad & \tilde{u}_j = \tilde{\tilde{u}}_j = 0 \\ k=1 \quad & \tilde{u}_j = \tilde{\tilde{u}}_j = 6u_j^{(1)} \\ k=2 \quad & \tilde{u}_j = 6u_j^{(1)} + 30u_j^{(2)}, \quad \tilde{\tilde{u}}_j = 6u_j^{(1)} - 30u_j^{(2)}. \end{aligned}$$

Pour assurer au schéma de satisfaire une propriété TVD et pour conserver la précision du schéma même dans les points critiques, les états $\tilde{u}_j$ sont modifiés de la façon suivante :

$$\tilde{u}_j^{(mod)} = \tilde{m}(\tilde{u}_j, \Delta_+ u_j^{(0)}, \Delta_- u_j^{(0)}), \qquad \tilde{\tilde{u}}_j^{(mod)} = \tilde{m}(\tilde{\tilde{u}}_j, \Delta_+ u_j^{(0)}, \Delta_- u_j^{(0)}), \tag{2.61}$$

où $\tilde{m}$ est un limiteur dérivé du limiteur classique minmod m :

$$\tilde{m}(a_1, \ldots, a_n) = \begin{cases} a_1 & \text{si } |a_1| \leq M\Delta x^2, \\ minmod(a_1, \ldots, a_n) & \text{sinon,} \end{cases}$$

avec :

$$minmod(a_1, \ldots, a_n) = \begin{cases} s. \min |a_i| & \text{si } sign(a_1) = \ldots = sign(a_n) = s, \\ 0 & \text{sinon.} \end{cases}$$

Ici, M est une constante vérifiant :

$$M \geq \frac{3}{2} \max_J |\partial_{xx} u_0|, \tag{2.62}$$

avec J un voisinage des points critiques de la donnée initiale $u_0(x)$.
De la même manière que pour les limiteurs locaux de projection, les états $u_{j+\frac{1}{2}}^{\pm(mod)}$ sont définis par :

$$u_{j+\frac{1}{2}}^{-(mod)} = u_j^{(0)} + \tilde{u}_j^{(mod)}, \qquad u_{j-\frac{1}{2}}^{+(mod)} = u_j^{(0)} - \tilde{\tilde{u}}_j^{(mod)}, \tag{2.63}$$

soit en remplaçant dans (2.53) :

$$\begin{aligned} k=0 \quad & \tilde{u}_j^{(mod)} = \tilde{\tilde{u}}_j^{(mod)} = 0, \\ k=1 \quad & \tilde{u}_j^{(mod)} = \tilde{\tilde{u}}_j^{(mod)} = 6u_j^{(1)}, \\ k=2 \quad & \tilde{u}_j^{(mod)} = 6u_j^{(1)} + 30u_j^{(2)}, \quad \tilde{\tilde{u}}_j^{(mod)} = 6u_j^{(1)} - 30u_j^{(2)}. \end{aligned}$$

Les flux du schéma (2.57) sont alors donnés par :

$$h_{j+\frac{1}{2}} = h(u_{j+\frac{1}{2}}^{-(mod)}, u_{j+\frac{1}{2}}^{+(mod)}) = \begin{cases} a u_{j+\frac{1}{2}}^{-(mod)} & \text{si } a > 0, \\ a u_{j+\frac{1}{2}}^{+(mod)} & \text{sinon.} \end{cases} \tag{2.64}$$

Les flux étant définis, le schéma revient à résoudre sur chaque intervalle C_j un système de $k+1$ équations différentielles du premier ordre donné par (2.57). Sous forme vectorielle, le système s'écrit :

$$\partial_t u_j = L(u_j, t), \tag{2.65}$$

avec $u_j = (u_j^{(0)}, u_j^{(1)}, \ldots, u_j^{(k)})^T$ le vecteur des inconnues. Pour $k = 0, 1, 2$, le second membre de (2.65) s'écrit :

$$k = 2 \begin{cases} k = 1 & \begin{cases} k = 0 & L_0(u_j, t) = -\frac{1}{\Delta x}(h_{j+\frac{1}{2}} - h_{j-\frac{1}{2}}), \\ & L_1(u_j, t) = -\frac{1}{2\Delta x}(h_{j+\frac{1}{2}} + h_{j-\frac{1}{2}}) + \frac{1}{\Delta x} a u_j^{(0)}, \\ & L_2(u_j^h, t) = -\frac{1}{6\Delta x}(h_{j+\frac{1}{2}} - h_{j-\frac{1}{2}}) + \frac{2}{\Delta x} a u_j^{(1)}. \end{cases} \end{cases}$$

Pour discrétiser chaque ligne du système (2.65) en temps, on peut utiliser par exemple un schéma Runge-Kutta TVD d'ordre r décrit dans [94] :

$$u^{(i)} = \sum_{l=0}^{i-1} \left[\alpha_{il} u^{(l)} + \beta_{il} \Delta t L\left(u^{(l)}, t^n + d_l \Delta t\right) \right] \quad \forall i = 1, \ldots, r \tag{2.66}$$

avec :

$$u^{(0)} = u^n \quad \text{et} \quad u^{(r)} = u^{n+1} \tag{2.67}$$

Les schémas correspondants d'ordre 2 et 3 sont donnés par :

$$r = 2 : \begin{cases} u^{(0)} &= u^n, \\ u^{(1)} &= u^{(0)} + \Delta t L\left(u^{(0)}, t^n\right), \\ u^{n+1} &= \frac{1}{2} u^{(0)} + \frac{1}{2} u^{(1)} + \frac{1}{2} \Delta t L\left(u^{(1)}, t^n + \Delta t\right), \end{cases}$$

$$r = 3 : \begin{cases} u^{(0)} &= u^n, \\ u^{(1)} &= u^{(0)} + \Delta t L\left(u^{(0)}, t^n\right), \\ u^{(2)} &= \frac{3}{4} u^{(0)} + \frac{1}{4} u^{(1)} + \frac{1}{4} \Delta t L\left(u^{(1)}, t^n + \Delta t\right), \\ u^{n+1} &= \frac{1}{3} u^{(0)} + \frac{2}{3} u^{(2)} + \frac{2}{3} \Delta t L\left(u^{(2)}, t^n + \frac{1}{2} \Delta t\right). \end{cases}$$

Pour illustrer ce schéma, il sera comparé au schéma GRP espace-temps dans le paragraphe 2.3.

2.6.3 Schéma WENO pour l'équation de transport linéaire

Les schémas WENO (weighted essentially non oscillating) développés par Liu et al. dans [89] sont une extension des schémas ENO introduits par Harten et al. dans [69]. Dans le cas multidimensionnel, des extensions d'ordre élevé ont également été développées (voir par exemple [78, 2, 46]). Comme dans les méthodes de volumes finis, on cherche à approcher la solution moyenne sur chaque cellule. Le principe est de se ramener sur chaque maille à la résolution d'une équation différentielle du premier ordre en temps dont le second membre est une approximation du bilan de flux. Sur chaque cellule, la solution moyenne est approchée par une combinaison convexe de polynômes d'interpolation, interpolés sur les cellules voisines de la cellule concernée.

La méthode présentée ici est celle décrite par Liu, Osher et Chan dans [89] dans le cas particulier de l'équation linéaire en dimension un d'espace.
On considère le problème de Cauchy pour l'équation d'advection :

$$\begin{aligned} &\partial_t u + a \partial_x u = 0, \\ &u(x, 0) = u_0(x), \end{aligned} \tag{2.68}$$

où a est une vitesse constante.
Le domaine est discrétisé par un maillage uniforme formé des cellules $C_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ de même longeur Δx. On note x_j le milieu de la cellule C_j. On définit :

$$\bar{u}_j(., t) = \frac{1}{\Delta x} \int_{C_j} u(x, t) dx, \tag{2.69}$$

la valeur moyenne de la solution sur l'intervalle C_j.
En intégrant (2.68) sur la maille C_j, cette moyenne vérifie l'équation différentielle du premier ordre suivante :

$$\partial_t \bar{u}_j(.,t) = -\frac{a}{\Delta x}\left[u(x_{j+\frac{1}{2}},t) - u(x_{j-\frac{1}{2}},t)\right]. \tag{2.70}$$

Pour connaître $\partial_t \bar{u}_j(.,t)$, il faut donc évaluer u sur les interfaces $x_{j+\frac{1}{2}}$ et $x_{j-\frac{1}{2}}$.
Soit $R_j(x)$ une approximation polynomiale de degré $r-1$ de la solution $u(x,t)$ sur la maille C_j. Sur l'interface $x_{j+\frac{1}{2}}$, on a alors deux approximations de la solution : $R_j(u, x_{j+\frac{1}{2}})$ et $R_{j+1}(u, x_{j+\frac{1}{2}})$. On définit ensuite un flux lipschitzien monotone $\tilde{h}(.,.)$ croissant par rapport à sa première variable et décroissant par rapport à la deuxième. On peut considérer par exemple le flux d'Engquist-Osher. Ainsi, le flux $au(x_{j+\frac{1}{2}},t)$ est approché par $\tilde{h}(R_j(u, x_{j+\frac{1}{2}}), R_{j+1}(u, x_{j+\frac{1}{2}}))$. On peut donc approcher l'équation (2.70) par :

$$\partial_t \bar{u}_j(.,t) \simeq -\frac{a}{\Delta x}\left[\tilde{h}(R_j(\bar{u}, x_{j+\frac{1}{2}}), R_{j+1}(\bar{u}, x_{j+\frac{1}{2}})) - \tilde{h}(R_j(\bar{u}, x_{j-\frac{1}{2}}), R_{j-1}(\bar{u}, x_{j-\frac{1}{2}}))\right] = L_j(\bar{u}). \tag{2.71}$$

On considère ici le flux du schéma décentré amont :

$$\tilde{h}(b_1,b_2) = \begin{cases} b_1 & \text{si } a > 0, \\ b_2 & \text{sinon.} \end{cases}$$

La fonction L_j pour $a > 0$ est donc :

$$L_j(\bar{u}) = -\frac{a}{\Delta x}\left[R_j(\bar{u}, x_{j+\frac{1}{2}}) - R_{j-1}(\bar{u}, x_{j-\frac{1}{2}})\right].$$

Une fois la fonction L_j définie sur chaque cellule C_j du maillage, une approximation d'ordre r ($r+1$ sur les nœuds du maillage) en espace est donnée par la solution de l'équation différentielle du premier ordre (2.71). Pour une approximation d'ordre élevé en temps, il faut alors choisir une méthode d'ordre $r+1$, comme par exemple le schéma Runge-Kutta TVD développé par Shu et Osher dans [94]. Le schéma correspondant d'ordre 3 ($r=2$) est donné par :

$$\begin{aligned}
\bar{u}_j^{(0)} &= \bar{u}_j^n, \\
\bar{u}_j^{(1)} &= \bar{u}_j^{(0)} + \Delta t L_j(\bar{u}^{(0)}), \\
\bar{u}_j^{(2)} &= \frac{3}{4}\bar{u}_j^{(0)} + \frac{1}{4}\bar{u}_j^{(1)} + \frac{\Delta t}{4}L_j(\bar{u}^{(1)}), \\
\bar{u}_j^{(n+1)} &= \frac{1}{3}\bar{u}_j^{(0)} + \frac{1}{3}\bar{u}_j^{(2)} + \frac{\Delta t}{3}L_j(\bar{u}^{(2)}).
\end{aligned}$$

Le schéma d'ordre 4 ($r=3$) est présenté par Gottlieb et Shu dans [65] :

$$\begin{aligned}
\bar{u}_j^{(0)} &= \bar{u}_j^n, \\
\bar{u}_j^{(1)} &= \bar{u}_j^{(0)} + \frac{1}{2}\Delta t L_j(\bar{u}^{(0)}), \\
\bar{u}_j^{(2)} &= \frac{649}{1600}\bar{u}_j^{(0)} - \frac{10890423}{25193600}\Delta t L_j(\bar{u}^{(0)}) + \frac{951}{1600}\bar{u}_j^{(1)} + \frac{5000}{7873}\Delta t L_j(\bar{u}^{(1)}), \\
\bar{u}_j^{(3)} &= \frac{53989}{2500000}\bar{u}_j^{(0)} - \frac{102261}{5000000}\Delta t L_j(\bar{u}^{(0)}) + \frac{4806213}{20000000}\bar{u}_j^{(1)} - \frac{5121}{20000}\Delta t L_j(\bar{u}^{(1)}) + \frac{23619}{32000}\bar{u}_j^{(2)} \\
&\quad + \frac{7873}{10000}\Delta t L_j(\bar{u}^{(2)}), \\
\bar{u}_j^{(n+1)} &= \frac{1}{5}\bar{u}_j^{(0)} + \frac{\Delta t}{10}L_j(\bar{u}^{(0)}) + \frac{6127}{30000}\bar{u}_j^{(1)} + \frac{\Delta t}{6}L_j(\bar{u}^{(1)}) + \frac{7873}{30000}\bar{u}_j^{(2)} + \frac{1}{3}\bar{u}_j^{(3)} + \frac{\Delta t}{6}L_j(\bar{u}^{(3)}).
\end{aligned}$$

Il reste alors à définir une procédure de reconstruction du polynôme R afin qu'ils vérifient les propriétés suivantes :

1. Sur chaque maille C_j, $\forall x \in C_j$, le polynôme $R_j(x)$ est une approximation d'ordre r de la solution $u(x,t)$:
$$\forall x \in C_j, \quad u(x,t) = R_j(u,x) + O(\Delta x^r),$$
et d'ordre $k+1$ sur les bords de la maille :
$$u(x_{j-\frac{1}{2}},t) = R_j(u,x_{j-\frac{1}{2}}) + O(\Delta x^{r+1}), \quad u(x_{j+\frac{1}{2}},t) = R_j(u,x_{j+\frac{1}{2}}) + O(\Delta x^{r+1}).$$

2. $R_j(u,x)$ est conservative :
$$\frac{1}{\Delta x}\int_{C_j} R_j(u,x)dx = \bar{u}_j \quad \forall j.$$

3. Pour tout j, $R_j(u,x)$ vérifie la propriété "ENO" définie plus loin.

Soit $W_t(x)$, la primitive de la solution $u(x,t)$ définie par :
$$W_t(x) = \int_{x_{j'-\frac{1}{2}}}^{x} u(s,t)ds,$$
où $x_{j'-\frac{1}{2}}$ est une interface quelconque du domaine. Par définition, la dérivée de W par rapport à x est égale à la solution $u(x,t)$:
$$u(x,t) = W_t'(x).$$
La fonction W_t évaluée sur les interfaces $x_{j+\frac{1}{2}}$ du maillage s'écrit donc en fonction de la valeur moyenne $\bar{u}_{i,j'\leq i\leq j}$ de la solution $u(x,t)$:
$$W_t(x_{j+\frac{1}{2}}) = \sum_{i=j'}^{j} \Delta x \bar{u}_i.$$
On définit ensuite le stencil :
$$S_j = (x_{j-r+\frac{1}{2}}, x_{j-r+\frac{3}{2}}, \ldots, x_{j+\frac{1}{2}}), \tag{2.72}$$
pour $a > 0$. Puis on introduit le polynôme p_j de degré r qui interpole $W_t(x)$ sur chaque point du stencil S_j :
$$\forall l = j-r, \ldots, j, \quad p_j(x_{l+\frac{1}{2}}) = W_t(x_{j+\frac{1}{2}}).$$
Ainsi, le polynôme p_j' de degré $r-1$ est une approximation d'ordre r de la solution :
$$u(x,t) = p_j'(x) + O(\Delta x^r).$$
Puis, pour chaque stencil S_j, on définit un indicateur de régularité IS_j de la solution $u(x,t)$ sur ce stencil. On introduit pour cela un tableau de différences des moyennes $\bar{u}_j$:
$$\begin{array}{l} \Delta[\bar{u}_{j-r+1}], \Delta[\bar{u}_{j-r+2}], \ldots, \Delta[\bar{u}_{j-1}], \\ \Delta^2[\bar{u}_{j-r+1}], \Delta^2[\bar{u}_{j-r+2}], \ldots, \Delta^2[\bar{u}_{j-2}], \\ \vdots \\ \Delta^{r-1}[\bar{u}_{j-r+1}], \end{array}$$
avec
$$\Delta[\bar{u}_l] = \bar{u}_{l+1} - \bar{u}_l \quad \text{et} \quad \Delta^k[\bar{u}_l] = \Delta^{k-1}[\bar{u}_{l+1}] - \Delta^{k-1}[\bar{u}_l].$$

L'indice de régularité est alors défini par :

$$IS_j = \frac{\sum_{l=1}^{r-1}\left(\sum_{k=1}^{l}(\Delta^{r-l}[\bar{u}_{j-r+k}])^2\right)}{l}.$$

Pour $r = 2$, l'indice de régularité est :

$$IS_j = (\Delta[\bar{u}_{j-1}])^2,$$

et pour $r = 3$:

$$IS_j = \frac{\Delta[\bar{u}_{j-2}]^2 + \Delta[\bar{u}_{j-1}]^2}{2} + (\Delta^2[\bar{u}_{j-2}])^2.$$

Si $u(x,t)$ est discontinue sur le stencil S_j, alors l'indicateur de régularité $IS_j \simeq O(1)$, et si la solution est assez régulière, l'indicateur de régularité est $IS_j \simeq O(\Delta x^2)$.
Contrairement au schéma ENO qui consiste à choisir le polynôme $p_j(x)$ le plus régulier pour approcher la solution, le schéma WENO consiste à approcher la solution par une combinaison convexe de r polynômes d'interpolation $\{p'_{j+k}(x)\}_{k=0}^{r-1}$ sur les stencils correspondants $\{S_{j+k}(x)\}_{k=0}^{r-1}$ où les poids sont attribués en fonction de l'indicateur de régularité $\{IS_{j+k}(x)\}_{k=0}^{r-1}$.
Pour chaque intervalle C_j, on considère les r stencils :

$$\{S_{j+k}(x)\}_{k=0}^{r-1} = \{x_{j+k-r+\frac{1}{2}}, x_{j+k-r+\frac{3}{2}}, \ldots, x_{j+k+\frac{1}{2}}\},$$

qui incluent tous les points $x_{j-\frac{1}{2}}$ et $x_{j+\frac{1}{2}}$ de C_j.
Ainsi le polynôme d'interpolation R_j peut s'écrire comme une combinaison convexe des $\{p'_{j+k}(x)\}_{k=0}^{r-1}$:

$$R_j(u,x) = \sum_{k=0}^{r-1} \frac{\alpha_k^j}{\sum_{l=0}^{r-1}\alpha_l^j} p'_{j+k}(x),$$

où $\alpha_k^j > 0$ pour tout $k = 0, 1, \ldots, r-1$. Le polynôme R_j vérifie alors la propriété "ENO", qui permet d'augmenter l'ordre d'approximation dans les singularités, si les α_k^j satisfont les propriétés suivantes :

1. si le stencil S_{j+k} est dans une région régulière, les coefficients α_k^j correspondants vérifient :

$$\frac{\alpha_k^j}{\sum_{l=0}^{r-1}\alpha_l^j} = O(1),$$

2. si le stencil S_{j+k} est dans une région où $u(x,t)$ est discontinue, les coefficients α_k^j correspondants vérifient :

$$\frac{\alpha_k^j}{\sum_{l=0}^{r-1}\alpha_l^j} \leq O(\Delta x^r).$$

Les coefficients α_k^j sont définis par :

$$\alpha_k^j = \frac{C_k^j}{(\varepsilon + IS_{j+k})^r},$$

où les coefficients C_k^j sont tels que : $C_k^j = O(1)$ et $C_k^j > 0$ pour assurer au polynôme R_j de vérifier la propriété "ENO". Afin d'améliorer l'ordre d'approximation de la solution $u(x,t)$ par le

polynôme R_j sur les nœuds du maillage $x_{j+\frac{1}{2}}$, une étude basée sur l'erreur de troncature donne, pour une vitesse $a > 0$, l'expression suivante pour les coefficients C_k^j :

$$C_k^j = \begin{cases} 1 & \text{si } a_k^j(x_{j+\frac{1}{2}}) = 0, \\ \frac{\Delta x^r}{\eta_p |a_k^j(x_{j+\frac{1}{2}})|} & \text{si } a_k^j(x_{j+\frac{1}{2}}) > 0, \\ \frac{\Delta x^r}{\eta_n |a_k^j(x_{j+\frac{1}{2}})|} & \text{si } a_k^j(x_{j+\frac{1}{2}}) < 0, \end{cases}$$

avec

$$a_k^j(x) = \sum_{s=0}^{r} \left(\prod_{l=0, l\neq s}^{r} (x - x_{j+k-l+\frac{1}{2}}) \right),$$

η_p le nombre de termes positifs dans $\{a_k^j(x_{j+\frac{1}{2}})\}_{k=0}^{r-1}$ et η_n le nombre de termes négatifs dans $\{a_k^j(x_{j+\frac{1}{2}})\}_{k=0}^{r-1}$.
Dans le cas $a < 0$, le point $x_{j+\frac{1}{2}}$ est remplacé par $x_{j-\frac{1}{2}}$.
Finalement le schéma obtenu est d'ordre r en espace et en temps à condition d'utiliser un schéma en temps adapté. On peut de plus montrer que le schéma est d'ordre $r+1$ en espace sur les nœuds du maillage. Ce schéma (d'ordre 3) est utilisé pour des comparaisons avec le GRP espace-temps dans le paragraphe 2.3.

3

Étude des modèles M_1

3.1 Introduction

Dans ce chapitre, on s'intéresse aux modèles M_1 gris et multigroupe. On a vu dans le paragraphe 1.4.1 que la particularité de ces modèles vient de leur fermeture basée sur un principe de minimisation d'entropie. On rappelle que le modèle M_1 gris est donné par le système suivant :

$$\begin{cases} \partial_t E_R + \nabla . F_R = c(\sigma^e a T^4 - \sigma^a E_R), \\ \partial_t F_R + c^2 \nabla . (D_R E_R) = -c\sigma^f F_R, \end{cases} \tag{1.22}$$

et que le modèle M_1 multigroupe est donné par :

$$\begin{cases} \partial_t E_q + \nabla . F_q = c(\sigma_q^e a \theta_q^4(T) - \sigma_q^a E_q), \\ \partial_t F_q + c^2 \nabla . (D_q E_q) = -c\sigma_q^f F_q. \end{cases} \forall 1 \leq q \leq Q \tag{1.31}$$

On rappelle que les opacités moyennes sont définies par (1.23)-(1.28) :

$$\sigma^e = \frac{< \sigma_\nu B(T) >}{aT^4}, \quad \sigma_q^e = \frac{< \sigma_\nu B(T) >_q}{< B(T) >_q},$$
$$\sigma^a = \frac{< \sigma_\nu I >}{E_R}, \quad \sigma_q^a = \frac{< \sigma_\nu I >_q}{E_q},$$
$$\sigma^f F_R = c < \sigma_\nu \Omega I >, \quad \sigma_q^f = \frac{c < \sigma_\nu \Omega I >_q}{F_q},$$

où $B(T)$ est donnée par (1.1) et l'intensité radiative est remplacée par $\mathcal{I}$ donnée par (1.30). Ces modèles sont couplés avec une équation régissant la température matière donnée par exemple dans le cas gris par :

$$\rho C_v \partial_t T = -c(\sigma^e a T^4 - \sigma^a E_R). \tag{1.6}$$

Ces modèles possèdent de nombreuses propriétés, dont notamment la positivité de l'énergie radiative $E_R > 0$ et la limitation du flux $F_R \leq cE_R$. Ils possèdent également les bons régimes asymptotiques. En effet, quand les opacités tendent vers 0, ils possèdent la bonne limite de transport donnée par exemple pour le modèle gris par :

$$\begin{cases} \partial_t E_R + \nabla . F_R = 0, \\ \partial_t F_R + c^2 \nabla . (D_R E_R) = 0. \end{cases}$$

Puis, quand le produit $\sigma_\nu t$ est très grand, ils possèdent la limite de diffusion suivante :

$$\partial_t \left(\rho C_v T + aT^4\right) - \partial_x \left(\frac{c}{3\sigma^0}\partial_x(aT^4)\right) = 0, \tag{3.1}$$

où σ^0 est une moyenne d'opacité définie par :

$$\begin{aligned}\sigma^0 &= \lim_{F_R \to 0} \sigma^f \\ &= \frac{\int_0^\infty \sigma_\nu \partial_t \mathcal{B}_\nu(T) d\nu}{\int_0^\infty \partial_t \mathcal{B}_\nu(T) d\nu}.\end{aligned}$$

Si σ_ν est constante, cette équation coïncide exactement avec l'équation de diffusion associée à l'ETR dans ce régime. Lorsque σ_ν n'est pas constante, le coefficient de diffusion σ^0 de la limite du modèle n'est qu'une approximation du coefficient de diffusion de l'ETR donné par la moyenne de Rosseland σ^R. Numériquement, les schémas classiques discrétisant ces systèmes ne préservent pas systématiquement ces propriétés. C'est le cas par exemple des méthodes classiques de *splitting* utilisées pour prendre en compte le terme source.

Pour préserver le bon comportement du modèle dans les régimes de diffusion et de transport, il faut donc considérer un schéma qui préserve l'asymptotique (*asymptotic preserving*). Quelques schémas préservant l'asymptotique ont par exemple été développés dans [26, 24] pour les équations d'hydrodynamique radiative et dans [62] pour le modèle de Goldstein-Taylor. Dans ce paragraphe, l'attention est portée sur le schéma développé dans [19]. Dans un premier temps, on dérive un solveur de type Godunov pour le système homogène. Dans un second temps, le schéma consiste à modifier ce solveur de Riemann approché afin d'y intégrer directement le terme source. Cette méthode est détaillée dans le paragraphe 3.2.4.
Dans la deuxième partie de ce chapitre, on s'intéresse au modèle M_1 multigroupe. Dans le paragraphe 1.4.1, on a vu que contrairement au cas gris, le calculs des moyennes d'opacités et du facteur d'Eddington ne sont pas explicites. En effet, lors de la fermeture du modèle suivant le principe de minimisation de l'entropie, le facteur d'anisotropie χ_q dans le groupe q ainsi que les opacités moyennes sont solutions de problèmes non linéaires de minimisation avec contraintes. Ces coefficients sont essentiels pour obtenir une bonne approximation de la solution et une mauvaise évaluation de ceux-ci pourrait par exemple conduire à des solutions physiquement non admissibles. Cependant, ces problèmes de minimisation impliquent des fonctions très raides qui posent des difficultés numériques. Puis, pour éviter de recalculer $\chi_q, \sigma_q^e, \sigma_q^a$ et σ_q^f pour chaque couple énergie-flux (E_q, F_q), on propose dans le paragraphe 3.3 une procédure de pré-calculs de ces coefficients dans le cas uni-dimensionnel.

3.2 Modèle M_1 gris

3.2.1 Résolution du problème de Riemann en 1D

Dans ce paragraphe, on résout le problème de Riemann pour le modèle M_1 gris en 1D. En effet, cette solution exacte permet par exemple d'écrire le solveur exact de Godunov associé. Ce solveur exact, détaillé dans le paragraphe suivant, est notamment un atout supplémentaire pour la validation des cas-tests.
On considère ici le modèle M_1 gris 1D gouverné par le système d'EDP suivant :

$$\partial_t U + \partial_x F(U) = 0, \tag{3.2}$$

avec

$$U = \begin{pmatrix} E_R \\ F_R \end{pmatrix}, \quad F(U) = \begin{pmatrix} F_R \\ c^2 P_R \end{pmatrix}.$$

On rappelle que la pression radiative $P_R = \chi\left(\frac{F_R}{E_R}\right) E_R$ est obtenue par minimisation de l'entropie radiative où $\chi(f)$ est le facteur d'Eddington donné par :

$$\chi(f) = \frac{3+4f^2}{5+2\sqrt{4-3f^2}}. \tag{1.21}$$

On suppose que la solution appartient à l'espace des états admissibles suivant :

$$\mathcal{A} = \{(E_R, F_R) \in \mathbb{R}^2, E_R > 0, f = \frac{|F_R|}{cE_R} \leq 1\}.$$

Le système (3.2) est complété par une donnée initiale :

$$U(x, t=0) = U_0(x) = \begin{cases} U_L & \text{si } x < 0, \\ U_R & \text{si } x > 0, \end{cases} \tag{3.3}$$

où U_L et U_R sont deux états constants dans $\mathcal{A}$:

$$U_L = (E_{R,L}, F_{R,L})^T, U_R = (E_{R,R}, F_{R,R})^T.$$

Pour résoudre le problème de Riemann (3.2)-(3.3), dans un premier temps l'algèbre du système est déterminée. Sachant que l'algèbre d'un système d'EDP du premier ordre de la forme $\partial_t U + \partial_x f(U) = 0$ n'est pas modifiée par un changement de variables, on introduit un nouveau jeu de variables qui va simplifier la résolution du problème. Une fois les champs déterminés et leur nature caractérisée, les invariants de Riemann seront calculés. Les courbes de choc et de détente pourront alors être détaillées et conduiront à la résolution exacte du problème de Riemann (3.2)-(3.3).

Algèbre du système

On introduit de nouvelles variables $\Pi > 0$ et $\beta \in]-1, 1[$ suggérées dans [25] et définies par :

$$\Pi = \frac{E_R(1-\beta^2)}{3+\beta^2},$$
$$F_R = \frac{4cE_R}{3+\beta^2}\beta.$$

Les variables conservatives (E_R, F_R) se réécrivent alors sous la forme :

$$E_R = \frac{3+\beta^2}{1-\beta^2}\Pi,$$
$$F_R = \frac{4c\Pi}{1-\beta^2}\beta.$$

La pression trouve alors l'expression suivante :

$$P_R = E_R\chi = \frac{3\beta^2+1}{1-\beta^2}\Pi.$$

En introduisant ces variables dans le système (3.2), on obtient :

$$\begin{cases} \partial_t\beta + \dfrac{2c\beta}{3-\beta^2}\partial_x\beta + \dfrac{3c(1-\beta^2)^2}{4\Pi(3-\beta^2)}\partial_x\Pi = 0, \\ \partial_t\Pi + \dfrac{4c\Pi}{3-\beta^2}\partial_x\beta + \dfrac{2c\beta}{3-\beta^2}\partial_x\Pi = 0. \end{cases}$$

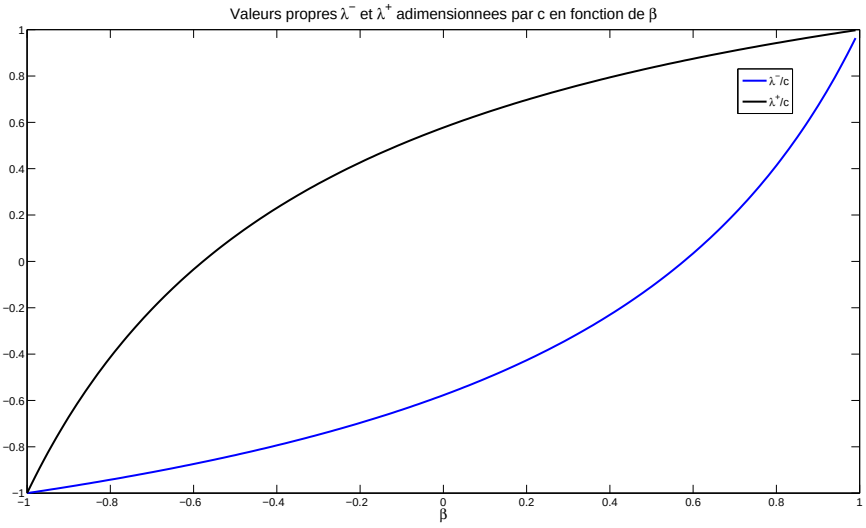


FIGURE 3.1 – Valeurs propres de $A(W)$ adimensionnées par c en fonction de β

Pour simplifier les notations, on écrit ce dernier système sous la forme condensée :

$$\partial_t W + A(W)\partial_x W = 0, \tag{3.4}$$

où les notations suivantes ont été introduites :

$$W = \begin{pmatrix} \beta \\ \Pi \end{pmatrix}, \quad A(W) = \begin{pmatrix} \dfrac{2c\beta}{3-\beta^2} & \dfrac{3c(1-\beta^2)^2}{4\Pi(3-\beta^2)} \\ \dfrac{4c\Pi}{3-\beta^2} & \dfrac{2c\beta}{3-\beta^2} \end{pmatrix}. \tag{3.5}$$

L'espace des états admissibles, compris en terme de variables W, devient :

$$\mathcal{A}' = \{(\beta, \Pi) \in \mathbb{R}^2, -1 < \beta < 1, \Pi > 0\}.$$

On remarque que la fonction $W = W(U)$ définit une bijection de $\mathcal{A}$ vers $\mathcal{A}'$, si bien que pour tout $U \in \mathcal{A}$, le vecteur associé W sera également admissible et réciproquement. À présent, on s'intéresse à l'algèbre associé au système (3.4). Par un calcul direct, on peut déterminer les éléments propres de la matrice $A(W)$. Ses valeurs propres sont :

$$\lambda^{\pm} = \frac{2c\beta \pm \sqrt{3}c(1-\beta^2)}{3-\beta^2}, \tag{3.6}$$

avec les vecteurs propres associés :

$$r^{\pm} = \begin{pmatrix} \dfrac{\pm\sqrt{3}(1-\beta^2)}{4} \\ 1 \end{pmatrix}. \tag{3.7}$$

Puisque les valeurs propres $\lambda^{\pm}$, représentées sur la figure 3.1, sont toutes réelles distinctes, le système est bien strictement hyperbolique sur $\mathcal{A}'$. De plus, avec $(\beta, \Pi) \in \mathcal{A}'$, on en déduit la nature des deux champs caractéristiques. En effet, on a :

$$\nabla\lambda^{\pm}.r^{\pm} = \pm\frac{\sqrt{3}c(1-\beta^2)}{2(\beta+\sqrt{3})^2} \neq 0, \quad \forall(\beta, \Pi) \in \mathcal{A}'.$$

En conséquence, les deux champs du système (3.2) ou de manière équivalente dans (3.4) sont Vraiment Non Linéaires. Dans la suite, on note $W_L = W(U_L)$ et $W_R = W(U_R)$ si bien que la solution du problème de Riemann est composée de trois états constants U_L, U^*, U_R séparés par deux ondes qui seront soit des ondes de choc, soit des détentes [60]. Dans la suite, on suppose que tous les états considérés sont admissibles.

Courbes de détente

Définition 3.1. *On appelle onde de détente une solution faible autosimilaire de (3.4) ou de façon équivalente de (3.2) qui connecte un état W_L à un état W_R de la manière suivante :*

$$W(x,t) = \begin{cases} W_L, & \text{si } \frac{x}{t} \leq \lambda^{\pm}(W_L), \\ V\left(\frac{x}{t}\right), & \text{si } \lambda^{\pm}(W_L) \leq \frac{x}{t} \leq \lambda^{\pm}(W_R), \\ W_R, & \text{si } \frac{x}{t} \geq \lambda^{\pm}(W_R). \end{cases}$$

où $V\left(\frac{x}{t}\right)$ est une fonction continue solution de (3.4).

De plus, dans une onde de détente, les invariants de Riemann restent constants. On rappelle que les invariants de Riemann associés au champ caractéristique (λ, r) sont des fonctions $\varphi(W)$ vérifiant :

$$\nabla\varphi.r = 0.$$

avec $r = r^{\pm}$, un vecteur propre défini par (3.7).

Invariants de Riemann

Pour définir les invariants de Riemann, on cherche φ tel que :

$$\begin{aligned} &\nabla\varphi.r = 0, \\ \text{ce qui donne} \quad &\frac{\pm\sqrt{3}(1-\beta^2)}{4}\frac{\partial\varphi}{\partial\beta} + \Pi\frac{\partial\varphi}{\partial\Pi} = 0. \end{aligned}$$

Cette EDP du premier ordre admet une unique intégrale première [27] définie par :

$$\frac{\mathrm{d}\beta}{\pm\sqrt{3}(1-\beta^2)} = \frac{\mathrm{d}\ln\Pi}{4},$$

et donc :

$$\mathrm{d}\left(\mathrm{argth}(\beta) - \frac{\pm\sqrt{3}}{4}\ln\Pi\right) = 0.$$

Finalement, on en déduit que l'invariant de Riemann pour chaque champ (λ^-, r^-) et (λ^+, r^+) est :

$$\begin{aligned} \varphi^- &= \mathrm{argth}\,\beta + \frac{\sqrt{3}}{4}\ln\Pi, \\ \varphi^+ &= \mathrm{argth}\,\beta - \frac{\sqrt{3}}{4}\ln\Pi. \end{aligned}$$

$\mathcal{R}_1$: 1-détente

Soit W_L donné dans $\mathcal{A}'$. On cherche tous les états $W \in \mathcal{A}'$ qui peuvent être connectés à W_L par une 1-détente associée au champs caractéristique (λ^-, r^-). Puisque l'invariant de Riemann reste constant dans la détente, on en déduit que W_L et W vérifient :

$$\operatorname{argth}\beta + \frac{\sqrt{3}}{4}\ln\Pi = \operatorname{argth}(\beta_L) + \frac{\sqrt{3}}{4}\ln\Pi_L. \tag{3.8}$$

De plus, la vitesse caractéristique de l'onde de détente implique :

$$\lambda^-(W_L) < \lambda^-(W).$$

Sachant que λ^- définie par (3.6) est croissante en $\beta \in]-1,1[$, cette dernière condition devient :

$$\beta_L < \beta.$$

La relation (3.8) donne l'expression de Π en fonction de β :

$$\ln\Pi = \frac{4}{\sqrt{3}}\left[I_L - \operatorname{argth}\beta\right],$$

où $I_L = \operatorname{argth}(\beta_L) + \dfrac{\sqrt{3}}{4}\ln\Pi_L$. Il en résulte :

$$\Pi = \exp\left[\frac{4}{\sqrt{3}}(I_L - \operatorname{argth}\beta)\right].$$

Finalement, on introduit la courbe suivante paramétrée par β :

$$\mathcal{R}_1(W_L) = \left\{\Pi_{\mathcal{R}_1}(\beta,\beta_L) = \Pi_L \exp\left[\frac{4}{\sqrt{3}}(\operatorname{argth}\beta_L - \operatorname{argth}\beta)\right]; \beta_L < \beta\right\}. \tag{3.9}$$

Cette courbe $\mathcal{R}_1(W_L)$ caractérise l'ensemble des états admissibles qui peuvent être connectés à W_L à droite par une 1-détente.
On étudie ensuite les propriétés de cette courbe de détente $\mathcal{R}_1$ qui seront importantes pour la résolution complète du problème de Riemann.

Étude de $\mathcal{R}_1$

Dans ce paragraphe, on montre que la courbe $\Pi_{\mathcal{R}_1}(.,\beta_L) :]-1;1[\to \mathbb{R}$, définie par (3.9) est décroissante et convexe. En effet, on voit tout d'abord que $\Pi_{\mathcal{R}_1}$ est définie, continue, deux fois dérivable sur $]-1;1[$ et positive. Sa dérivée est :

$$\begin{aligned}\Pi'_{\mathcal{R}_1}(\beta,\beta_L) &= \Pi_L\left(-\frac{4}{\sqrt{3}(1-\beta^2)}\right)\exp\left[\frac{4}{\sqrt{3}}(\operatorname{argth}\beta_L - \operatorname{argth}\beta)\right],\\ &= -\frac{4}{\sqrt{3}(1-\beta^2)}\Pi_{\mathcal{R}_1}(\beta,\beta_L) < 0,\end{aligned}$$

et sa dérivée seconde :

$$\begin{aligned}\Pi''_{\mathcal{R}_1}(\beta,\beta_L) &= -\frac{8\beta}{\sqrt{3}(1-\beta^2)^2}\Pi_{\mathcal{R}_1}(\beta,\beta_L) - \frac{4}{\sqrt{3}(1-\beta^2)}\left[-\frac{4}{\sqrt{3}(1-\beta^2)}\Pi_{\mathcal{R}_1}(\beta,\beta_L)\right],\\ &= \frac{8\beta}{\sqrt{3}(1-\beta^2)^2}\Pi_{\mathcal{R}_1}(\beta,\beta_L)\left[\frac{2}{\sqrt{3}} - \beta\right].\end{aligned}$$

Comme on a $-1 < \beta < 1$ et $\Pi_{\mathcal{R}_1}(\beta,\beta_L) > 0$, $\Pi''_{\mathcal{R}_1}(\beta,\beta_L)$ est positive. On en déduit que la courbe $\mathcal{R}_1$ est convexe.

$\mathcal{R}_2$: 2-détente

Soit W_L donné dans $\mathcal{A}'$. On cherche tous les états $W \in \mathcal{A}'$ qui peuvent être connectés à W_L par une 2-détente. Puisque l'invariant de Riemann reste constant dans la détente, on en déduit que W_L et W vérifient :

$$\operatorname{argth}\beta - \frac{\sqrt{3}}{4}\ln\Pi = \operatorname{argth}(\beta_L) - \frac{\sqrt{3}}{4}\ln\Pi_L. \tag{3.10}$$

De plus, la vitesse caractéristique de l'onde de détente implique :

$$\lambda^+(W_L) < \lambda^+(W).$$

Sachant que λ^+ est croissante en $\beta \in]-1,1[$, cette condition devient :

$$\beta_L < \beta.$$

La relation (3.10) donne l'expression de Π en fonction de β :

$$\ln\Pi = \frac{4}{\sqrt{3}}\left[\operatorname{argth}\beta - I_L\right],$$

où $I_L = \operatorname{argth}(\beta_L) - \dfrac{\sqrt{3}}{4}\ln\Pi_L$. Il en résulte :

$$\Pi = \exp\left[\frac{4}{\sqrt{3}}(\operatorname{argth}\beta - I_L)\right].$$

Finalement, on introduit la courbe suivante paramétrée par β :

$$\mathcal{R}_2(W_L) = \left\{\Pi_{\mathcal{R}_2}(\beta,\beta_L) = \Pi_L \exp\left[\frac{4}{\sqrt{3}}(\operatorname{argth}\beta - \operatorname{argth}\beta_L)\right]; \beta_L < \beta\right\}. \tag{3.11}$$

Cette courbe $\mathcal{R}_2(W_L)$ est donc l'ensemble des états admissibles qui peuvent être connectés à W_L à gauche par une 2-détente.
On étudie maintenant les propriétés de cette courbe de détente $\mathcal{R}_2$.

Étude de $\mathcal{R}_2$

Dans ce paragraphe, on montre que la courbe $\Pi_{\mathcal{R}_2}(.,\beta_L) :]-1;1[\to \mathbb{R}$, définie par (3.11) est croissante et convexe. En effet, on voit tout d'abord que $\Pi_{\mathcal{R}_2}$ est définie, continue, deux fois dérivable sur $]-1;1[$ et positive. Sa dérivée est :

$$\begin{aligned}\Pi'_{\mathcal{R}_2}(\beta,\beta_L) &= \Pi_L\left(\frac{4}{\sqrt{3}(1-\beta^2)}\right)\exp\left[\frac{4}{\sqrt{3}}(\operatorname{argth}\beta - \operatorname{argth}\beta_L)\right],\\ &= \frac{4}{\sqrt{3}(1-\beta^2)}\Pi_{\mathcal{R}_2}(\beta,\beta_L) > 0,\end{aligned}$$

et sa dérivée seconde :

$$\begin{aligned}\Pi''_{\mathcal{R}_2}(\beta,\beta_L) &= -\frac{8\beta}{\sqrt{3}(1-\beta^2)^2}\Pi_{\mathcal{R}_2}(\beta,\beta_L) + \frac{4}{\sqrt{3}(1-\beta^2)}\left[\frac{4}{\sqrt{3}(1-\beta^2)}\Pi_{\mathcal{R}_2}(\beta,\beta_L)\right],\\ &= \frac{8\beta}{\sqrt{3}(1-\beta^2)^2}\Pi_{\mathcal{R}_2}(\beta,\beta_L)\left[\frac{2}{\sqrt{3}} - \beta\right].\end{aligned}$$

Comme on a $-1 < \beta < 1$ et $\Pi_{\mathcal{R}_2}(\beta,\beta_L) > 0$, $\Pi''_{\mathcal{R}_2}(\beta,\beta_L)$ est positive. On en déduit que la courbe $\mathcal{R}_2$ est convexe.
Les figures 3.2-3.3 proposent une représentation de $\mathcal{R}_1$ et $\mathcal{R}_2$.

Courbes de choc

Pour U_L un état constant donné dans $\mathcal{A}$, on cherche tous les états $U \in \mathcal{A}$ qui puissent être connectés à U_L par une onde de choc, respectivement un 1-choc ou un 2-choc. Une discontinuité de type choc doit vérifier les relations de Rankine-Hugoniot qui ici trouvent la forme suivante :

$$\begin{cases} -S[E_R] + [F_R] & = 0, \\ -S[F_R] + c^2[\chi E_R] & = 0, \end{cases} \tag{3.12}$$

où S est la vitesse de propagation du choc et $[U] = U - U_L$ désigne le saut.
En réécrivant les relations (3.12) en fonction des variables (β, Π), on obtient :

$$\begin{cases} -S\left[\dfrac{3+\beta^2}{1-\beta^2}\Pi\right] + \left[\dfrac{4c\Pi\beta}{1-\beta^2}\right] & = 0, \\ -S\left[\dfrac{4c\Pi\beta}{1-\beta^2}\right] + c^2\left[\dfrac{3\beta^2+1}{1-\beta^2}\Pi\right] & = 0. \end{cases}$$

En posant $z = \frac{\Pi}{1-\beta^2}$ et en substituant la première équation dans la seconde, on obtient une équation quadratique d'inconnue β. En choisissant la racine dans l'espace $]-1, 1[$, on en déduit une relation entre les états W et W_L donnée par :

$$\beta = \frac{c^2 - 3S^2 + 2cS\beta_L}{3c^2\beta_L - S^2\beta_L - 2cS},$$
$$\Pi = \frac{\Pi_L}{3}\frac{(S\beta_L - c)^2 - 9(S - c\beta_L)^2}{(1-\beta_L^2)(S^2 - c^2)}.$$

Concernant la vitesse S, on remarque ainsi qu'elle doit être solution de l'équation quadratique suivante :

$$S^2(\beta_L\beta - 3) + 2cS(\beta + \beta_L) + c^2(1 - 3\beta_L\beta) = 0.$$

Cette équation admet deux solutions données par :

$$S^{\pm} = \frac{c(\beta_L + \beta) \pm \sqrt{\Delta(\beta, \beta_L)}}{3 - \beta\beta_L},$$

avec

$$\Delta(\beta, \beta_L) = 3(\beta\beta_L - 1)^2 + (\beta - \beta_L)^2 \geq 0. \tag{3.13}$$

Afin de sélectionner la vitesse de propagation, d'après le théorème 4.1, page 61 dans [60], on utilise le critère suivant :

$$\lim_{\beta\to\beta_L} S_1(\beta) = \lambda^-(\beta_L), \quad \lim_{\beta\to\beta_L} S_2(\beta) = \lambda^+(\beta_L),$$

On en déduit immédiatement que les vitesses des chocs sont :

$$S_1(\beta) = c\frac{(\beta + \beta_L) - \sqrt{\Delta(\beta, \beta_L)}}{3 - \beta\beta_L} \text{ dans un 1-choc,}$$
$$S_2(\beta) = c\frac{(\beta + \beta_L) + \sqrt{\Delta(\beta, \beta_L)}}{3 - \beta\beta_L} \text{ dans un 2-choc.}$$

De plus, on dit que la discontinuité est entropique au sens de Lax si S_1 et S_2 vérifient :

$$\lambda^-(W_R) < S_1 < \lambda^+(W_R), \quad S_1 < \lambda^-(W_L)$$
$$\lambda^+(W_R) < S_2, \quad \lambda^-(W_L) < S_2 < \lambda^+(W_L).$$

Étant donné que $\lambda^{\pm} = \lambda^{\pm}(\beta)$ est croissante, ces deux conditions se ramènent à la condition équivalente suivante :

$$\beta < \beta_L.$$

Finalement, en substituant l'expression de la vitesse S dans la définition de Π, on en déduit les ensembles d'Hugoniot :

$$\mathcal{S}_1(W_L) = \left\{ \Pi_{\mathcal{S}_1}(\beta, \beta_L) = \Pi_L \frac{\left[\sqrt{\Delta(\beta,\beta_L)} - 2(\beta - \beta_L)\right]^2}{3(1-\beta_L^2)(1-\beta^2)}; \beta_L > \beta \right\}, \tag{3.14}$$

$$\mathcal{S}_2(W_L) = \left\{ \Pi_{\mathcal{S}_2}(\beta, \beta_L) = \Pi_L \frac{\left[\sqrt{\Delta(\beta,\beta_L)} - 2(\beta_L - \beta)\right]^2}{3(1-\beta_L^2)(1-\beta^2)}; \beta_L > \beta \right\}. \tag{3.15}$$

La courbe $\mathcal{S}_1(W_L)$ est donc l'ensemble des états admissibles qui peuvent être connectés à W_L à droite par un 1-choc et la courbe $\mathcal{S}_2(W_L)$ est l'ensemble des états admissibles qui peuvent être connectés à W_L à droite par un 2-choc.

Étude des courbes de choc

Dans ce paragraphe, on montre que la courbe de 1-choc est décroissante et que la courbe de 2-choc est croissante. Pour cela, on étudie les variations des fonctions $\Pi_{\mathcal{S}_1}(\beta, \beta_L)$ et $\Pi_{\mathcal{S}_2}(\beta, \beta_L)$. On voit tout d'abord que $\Pi_{\mathcal{S}_1}(\beta, \beta_L)$ et $\Pi_{\mathcal{S}_2}(\beta, \beta_L)$ sont définies, continues et dérivables sur $]-1;1[$. En dérivant, on a :

$$\Pi'_{\mathcal{S}_1}(\beta, \beta_L) = \frac{4(\beta\beta_L - 1)}{(1-\beta^2)\sqrt{\Delta(\beta,\beta_L)}} \Pi_{\mathcal{S}_1}(\beta, \beta_L).$$

Or on sait que $\Pi_{\mathcal{S}_1}(\beta, \beta_L)$ est positive quel que soit β et $\beta_L \in]-1;1[$ et que $\beta_L > \beta$. On en déduit que $\Pi'_{\mathcal{S}_1}(\beta, \beta_L)$ est négative et ainsi que $\Pi_{\mathcal{S}_1}(\beta, \beta_L)$ est décroissante.
De même en dérivant la fonction $\Pi_{\mathcal{S}_2}(\beta, \beta_L)$, on a :

$$\Pi'_{\mathcal{S}_2}(\beta, \beta_L) = \frac{4(1 - \beta\beta_L)}{(1-\beta^2)\sqrt{\Delta(\beta,\beta_L)}} \Pi_{\mathcal{S}_2}(\beta, \beta_L).$$

Sachant que $\Pi_{\mathcal{S}_2}(\beta, \beta_L)$ est positive pour tout β et $\beta_L \in]-1;1[$ et que $\beta_L > \beta$ pour un choc entropique, on en déduit que $\Pi'_{\mathcal{S}_2}(\beta, \beta_L)$ est positive, et ainsi que $\Pi_{\mathcal{S}_2}(\beta, \beta_L)$ est croissante.

Résolution du problème de Riemann

La dérivation précédente des courbes de choc et de détente permet d'exhiber la solution du problème de Riemann (3.2)-(3.3). Afin de résoudre le problème de Riemann (3.2)-(3.3), on cherche, s'il existe, un état intermédiaire $W^* = (\beta^*, \Pi^*)^T$ tel que la solution soit composée des trois états constants W_L, W^* et W_R séparés par deux ondes de choc ou de détente. Sur les figures 3.2-3.3, on représente ces quatre courbes $\mathcal{R}_1(W_L)$, $\mathcal{R}_2(W_L)$, $\mathcal{S}_1(W_L)$ et $\mathcal{S}_2(W_L)$ issues du point W_L dans le plan (β, Π). Ces courbes divisent le plan (β, Π) en quatre zones numérotées de 1 à 4.

On étudie à présent les courbes d'Hugoniot dans chacune des zones.

Zone 1 : 1-détente, 2-choc

On suppose que l'état W_R se trouve dans la zone 1. On cherche $W^* = (\beta^*, \Pi^*)$ tel que W^* soit connecté à W_L par une 1-détente et tel que W_R soit connecté à W^* par un 2-choc.
On s'intéresse tout d'abord à la courbe de 1-détente issue de l'état constant W_L. On cherche W^* sur la courbe $\mathcal{R}_1(W_L)$ issue de W_L caractérisée par (3.9), c'est-à-dire que W^* doit vérifier :

$$\Pi^* = \Pi_{\mathcal{R}_1}(\beta^*, \beta_L) = \Pi_L \exp\left(\frac{4}{\sqrt{3}}(\operatorname{argth}\beta_L - \operatorname{argth}\beta^*)\right).$$

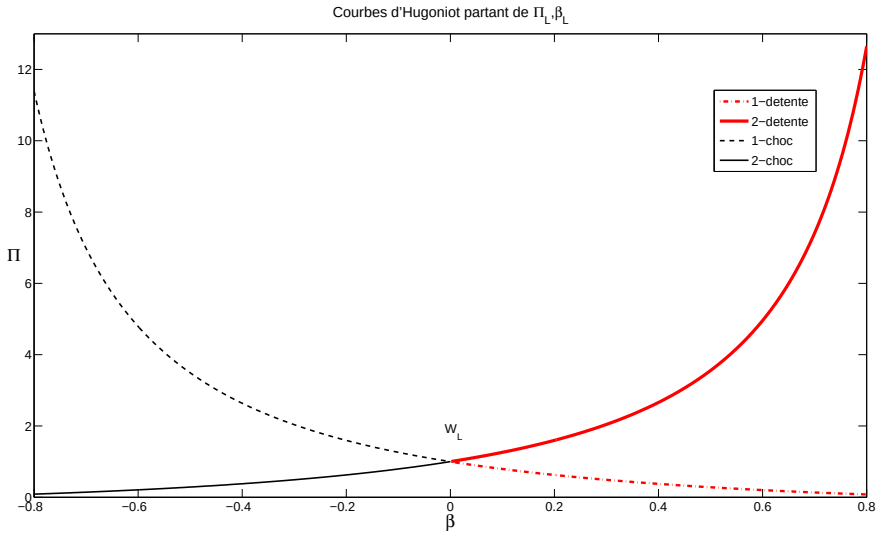


FIGURE 3.2 – Courbes de choc et de détente associées à $W_L = (0, 1)$

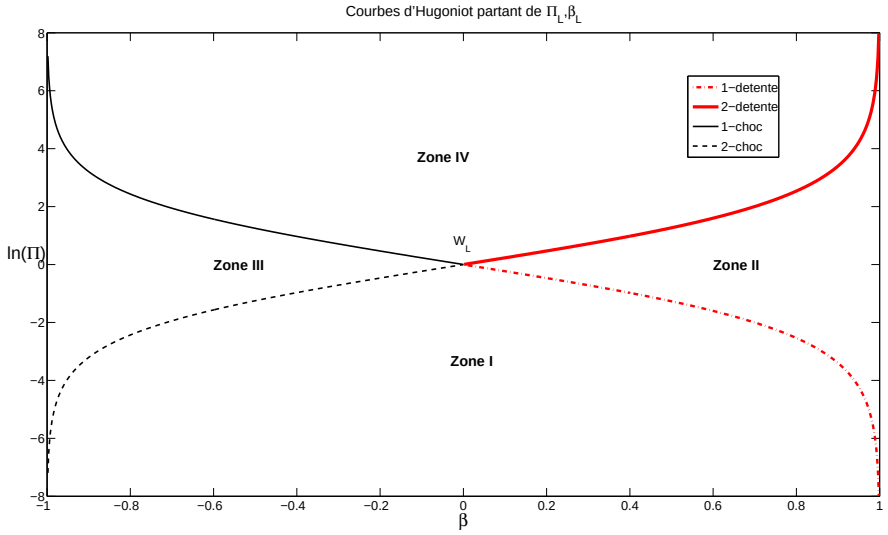


FIGURE 3.3 – Courbes de choc et de détente associées à $W_L = (0, 1)$ avec échelle logarithmique en ordonnée

On en déduit par décroissance de $\Pi_{\mathcal{R}_1}(\beta, \beta_L)$ que pour $\beta_L < \beta^*$:

$$\Pi_L > \Pi^*.$$

Puis on cherche W^* tel que W_R soit sur la courbe $\mathcal{S}_2(W^*)$ issue de W^* caractérisée par (3.15) c'est-à-dire que W^* doit vérifier :

$$\Pi_R = \Pi_{\mathcal{S}_2}(\beta_R, \beta^*) = \Pi^* \frac{\left[(\sqrt{\Delta(\beta^*, \beta_R)} + 2(\beta^* - \beta_R)\right]^2}{3(1-\beta^{*2})(1-\beta_R^2)},$$

où $\Delta(\beta^*, \beta_R)$ est défini par (3.13).
Finalement, l'état W^*, s'il existe, est caractérisé par :

$$\frac{\Pi_R}{\Pi_L} = \frac{\exp\left(\frac{4}{\sqrt{3}}(\operatorname{argth}\beta_L - \operatorname{argth}\beta^*)\right)\left[\sqrt{\Delta(\beta^*, \beta_R)} + 2(\beta^* - \beta_R)\right]^2}{3(1-\beta^{*,2})(1-\beta_R^2)}. \tag{3.16}$$

Pour simplifier les notations, on introduit la fonction F_1 telle que :

$$F_1(\beta) = \frac{\exp\left(\frac{4}{\sqrt{3}}(\operatorname{argth}\beta_L - \operatorname{argth}\beta)\right)\left[\sqrt{\Delta_R} + 2(\beta - \beta_R)\right]^2}{3(1-\beta^2)(1-\beta_R^2)},$$

de manière à réécrire (3.16) sous la forme :

$$\frac{\Pi_R}{\Pi_L} = F_1(\beta^*).$$

Zone 2 : 1-détente, 2-détente

On suppose que l'état W_R se trouve dans la zone 2. On cherche $W^* = (\beta^*, \Pi^*)$ tel que W^* soit connecté à W_L par une 1-détente et tel que W_R soit connecté à W^* par une 2-détente.
On s'intéresse tout d'abord à la courbe de 1-détente issue de l'état constant W_L. On cherche W^* sur la courbe $\mathcal{R}_1(W_L)$ issue de W_L caractérisée par (3.9), c'est-à-dire que W^* doit vérifier :

$$\Pi^* = \Pi_{\mathcal{R}_1}(\beta^*, \beta_L) = \Pi_L \exp\left(\frac{4}{\sqrt{3}}(\operatorname{argth}\beta_L - \operatorname{argth}\beta^*)\right),$$

avec $\beta^* > \beta_L$.
Puis on cherche W^* tel que W_R soit sur la courbe $\mathcal{R}_2(W^*)$ issue de W^* caractérisée par (3.11) c'est-à-dire que W^* doit vérifier :

$$\Pi_{\mathcal{R}_2}(\beta_R, \beta^*) = \Pi^* \exp\left(\frac{4}{\sqrt{3}}(\operatorname{argth}\beta_R - \operatorname{argth}\beta^*)\right),$$

avec $\beta > \beta^*$.
Finalement, s'il existe, l'état W^* est caractérisé par :

$$\frac{\Pi_R}{\Pi_L} = \exp\left(\frac{4}{\sqrt{3}}(\operatorname{argth}\beta_L + \operatorname{argth}\beta_R - 2\operatorname{argth}\beta^*)\right). \tag{3.17}$$

Pour simplifier les notations, on introduit la fonction F_2 telle que :

$$F_2(\beta) = \exp\left(\frac{4}{\sqrt{3}}(\operatorname{argth}\beta_L + \operatorname{argth}\beta_R - 2\operatorname{argth}\beta)\right),$$

de manière à réécrire (3.17) sous la forme :

$$\frac{\Pi_R}{\Pi_L} = F_2(\beta^*), \quad \text{avec} \quad \beta_L < \beta^* < \beta_R.$$

Zone 3 : 1-choc, 2-choc

On suppose que l'état W_R se trouve dans la zone 3. On cherche $W^* = (\beta^*, \Pi^*)$ tel que W^* soit connecté à W_L par un 1-choc et tel que W_R soit connecté à W^* par un 2-choc.
On s'intéresse tout d'abord à la courbe de 1-choc issue de l'état constant W_L. On cherche W^* sur la courbe $\mathcal{S}_1(W_L)$ issue de W_L caractérisée par (3.14), c'est-à-dire que W^* doit vérifier :

$$\Pi^* = \Pi_{\mathcal{S}_1}(\beta^*, \beta_L) = \Pi_L \frac{\left[\sqrt{\Delta(\beta^*, \beta_L)} + 2(\beta^* - \beta_L)\right]^2}{3(1-\beta_L^2)(1-\beta^{*2})},$$

où $\Delta(\beta^*, \beta_L)$ est caractérisé par (3.13) et $\beta_L > \beta^*$. Puis on cherche W^* tel que W_R soit sur la courbe $\mathcal{S}_2(W^*)$ issue de W^* caractérisée par (3.15) c'est-à-dire que W^* doit vérifier :

$$\Pi_{\mathcal{S}_2}(\beta_R, \beta^*) = \Pi^* \frac{\left[(\sqrt{\Delta(\beta^*, \beta_R)} + 2(\beta^* - \beta_R)\right]^2}{3(1-\beta^{*,2})(1-\beta_R^2)},$$

avec $\beta^* > \beta_R$.
Finalement, s'il existe, l'état W^* est caractérisé par :

$$\frac{\Pi_R}{\Pi_L} = \frac{\left[\sqrt{\Delta(\beta^*, \beta_L)} + 2(\beta^* - \beta_L)\right]^2 \left[\sqrt{\Delta(\beta^*, \beta_R)} + 2(\beta^* - \beta_R)\right]^2}{9(1-\beta_L^2)(1-\beta^{*,2})^2(1-\beta_R^2)}. \tag{3.18}$$

Pour simplifier les notations, on introduit la fonction F_3 telle que :

$$F_3(\beta) = \frac{\left[\sqrt{\Delta(\beta^*, \beta_L)} + 2(\beta - \beta_L)\right]^2 \left[\sqrt{\Delta(\beta^*, \beta_R)} + 2(\beta - \beta_R)\right]^2}{9(1-\beta_L^2)(1-\beta^2)^2(1-\beta_R^2)},$$

de manière à réécrire (3.18) sous la forme :

$$\frac{\Pi_R}{\Pi_L} = F_3(\beta^*) \quad \text{avec} \quad \beta_L > \beta^* > \beta_R.$$

Zone 4 : 1-choc, 2-détente

On suppose que l'état W_R se trouve dans la zone 2. On cherche $W^* = (\beta^*, \Pi^*)$ tel que W^* soit connecté à W_L par un 1-choc et tel que W_R soit connecté à W^* par une 2-détente.
On s'intéresse tout d'abord à la courbe de 1-détente issue de l'état constant W_L. On cherche W^* sur la courbe $\mathcal{R}_1(W_L)$ issue de W_L caractérisée par (3.14), c'est-à-dire que W^* doit vérifier :

$$\Pi^* = \Pi_{\mathcal{S}_1}(\beta^*, \beta_L) = \Pi_L \frac{\left[\sqrt{\Delta(\beta^*, \beta_L)} + 2(\beta^* - \beta_L)\right]^2}{3(1-\beta_L^2)(1-\beta^{*2})},$$

où $\Delta(\beta^*, \beta_L)$ est défini par (3.13).
De plus, par décroissance de $\Pi_{\mathcal{S}_1}(\beta, \beta_L)$, on en déduit que pour $\beta_L > \beta^*$:

$$\Pi^* > \Pi_L.$$

Puis on cherche W^* tel que W_R soit sur la courbe $\mathcal{R}_2(W^*)$ issue de W^* caractérisée par (3.11) c'est-à-dire que W^* doit vérifier :

$$\Pi_{\mathcal{R}_2}(\beta_R, \beta^*) = \Pi_{\mathcal{S}_1}(\beta^*, \beta_L) \exp\left(\frac{4}{\sqrt{3}}(\operatorname{argth}\beta_R - \operatorname{argth}\beta^*)\right).$$

Puis, par croissance de $\Pi_{\mathcal{R}_2}(\beta, \beta^*)$, on en déduit que pour $\beta_R > \beta^*$:

$$\Pi_{\mathcal{R}_2}(\beta^*, \beta^*) = \Pi^* < \Pi_{\mathcal{R}_2}(\beta_R, \beta^*).$$

Finalement, s'il existe, l'état W^* est caractérisé par :

$$\frac{\Pi_R}{\Pi_L} = \frac{\exp\left(\frac{4}{\sqrt{3}}(\operatorname{argth}\beta_R - \operatorname{argth}\beta^*)\right)\left[\sqrt{\Delta(\beta^*, \beta_L)} + 2(\beta^* - \beta_L)\right]^2}{3(1-\beta_L^2)(1-\beta^{*,2})}, \tag{3.19}$$

Pour simplifier les notations, on introduit la fonction F_4 telle que :

$$F_4(\beta) = \frac{\exp\left(\frac{4}{\sqrt{3}}(\operatorname{argth}\beta_R - \operatorname{argth}\beta)\right)\left[\sqrt{\Delta(\beta^*, \beta_L)} + 2(\beta - \beta_L)\right]^2}{3(1-\beta_L^2)(1-\beta^2)},$$

de manière à réécrire (3.19) sous la forme :

$$\frac{\Pi_R}{\Pi_L} = F_4(\beta^*) \quad \text{avec} \quad \beta_L > \beta^* > \beta_R.$$

Théorème 3.1. *Si U_L et U_R sont deux états constants de l'espace des états admissibles $\mathcal{A}$, alors le problème de Riemann (3.2)-(3.3) admet une unique solution.*

Dans la suite, on notera $U_{\mathcal{R}}(\frac{x}{t}; U_L, U_R)$ cette solution.

Démonstration. On a caractérisé précédemment les équations à résoudre pour déterminer l'état intermédiaire W^* dans chacune des quatre zones. Il faut donc vérifier que ces équations admettent une unique solution.
Zone 1 : 1-détente, 2-choc
On cherche β^* solution de (3.16). Pour simplifier les notations, on pose

$$F_1(\beta) = H(\beta)P(\beta),$$

avec :

$$H(\beta) = \exp\left(\frac{4}{\sqrt{3}}(\operatorname{argth}\beta_L - \operatorname{argth}\beta)\right),$$
$$P(\beta) = \frac{\left[\sqrt{\Delta(\beta, \beta_R)} + 2(\beta_R - \beta)\right]^2}{3(1-\beta^2)(1-\beta_R^2)},$$

avec $\Delta(\beta, \beta_R)$ caractérisé par (3.13).
On a vu que l'équation (3.16) peut se réécrire :

$$F_1(\beta) = \frac{\Pi_R}{\Pi_L}.$$

Pour montrer que cette équation admet une unique solution, on montre tout d'abord que F_1 est monotone décroissante puis finalement qu'elle passe bien par $\frac{\Pi_R}{\Pi_L}$. Un calcul direct de la dérivée donne :

$$F_1'(\beta) = 4P(\beta)H(\beta)\left[\frac{\beta\beta_R - 1}{(1-\beta^2)\sqrt{\Delta(\beta, \beta_R)}} - \frac{1}{\sqrt{3}(1-\beta^2)}\right].$$

Or $H(\beta)$ est positif quel que soit β et $P(\beta)$ est positif pour tout β et β_R sont dans $]-1;1[$. En conséquence, F_1 est décroissante pour tout β dans $]-1;1[$. On s'intéresse alors aux limites de F_1. Quand β tend vers β_L, on a :

$$\lim_{\beta\to\beta_L} H(\beta) = 1 \quad \text{et} \quad \lim_{\beta\to\beta_L} P(\beta) = P(\beta_L),$$

d'où

$$\lim_{\beta\to\beta_L} F_1(\beta) = P(\beta_L).$$

Lorsque β tend vers 1, on remarque que le numérateur $N(\beta) = \left[\sqrt{\Delta(\beta,\beta_R)} + 2(\beta_R - \beta)\right]^2$ dans l'expression de P vérifie :

$$N(1) = 0, \quad N'(1) = 0, \quad N''(1) \neq 0.$$

Ainsi, au voisinage de $\beta = 1$, on a :

$$P(\beta) \underset{1}{\sim} \frac{N''(1)}{3(1+\beta)(1-\beta)(1-\beta_R^2)},$$
$$H(\beta) \underset{1}{\sim} K(1-\beta)^{\frac{2}{\sqrt{3}}}.$$

On en déduit que la limite de F_1 quand β tend vers 1 est alors :

$$\lim_{\beta\to 1} F_1(\beta) = 0.$$

De plus, comme F_1 est décroissante, pour assurer l'existence et l'unicité d'une solution à l'équation (3.16) il reste à montrer que :

$$F_1(\beta_L) > \frac{\Pi_R}{\Pi_L},$$

ou de façon équivalente :

$$\Pi_L P(\beta_L) > \Pi_R.$$

On remarque que :

$$\Pi_L P(\beta_L) = \Pi_{\mathcal{S}_2}(\beta_R, \beta_L).$$

Comme dans cette zone on a $\Pi_L > \Pi^*$, il en résulte que :

$$\Pi_{\mathcal{S}_2}(\beta_R, \beta_L) > \Pi_{\mathcal{S}_2}(\beta_R, \beta^*) = \Pi_R.$$

On en déduit immédiatement que l'équation $F_1(\beta) = \dfrac{\Pi_R}{\Pi_L}$ admet bien une unique solution β^* dans cette zone. Il existe donc un unique état intermédiaire W^* qui connecte les états W_L et W_R respectivement par une 1-détente à gauche et un 2-choc à droite.

Zone 2 : 1-détente, 2-détente

On cherche β^* solution de (3.17). Pour montrer que cette équation admet une unique solution, on montre tout d'abord que F_2 est monotone décroissante puis finalement qu'elle passe bien par $\dfrac{\Pi_R}{\Pi_L}$. En dérivant la fonction F_2, on obtient :

$$F_2'(\beta) = -\frac{8}{\sqrt{3}(1-\beta^2)} F_2(\beta).$$

Or $F_2(\beta)$ est positive pour tout β, donc $F_2'(\beta)$ est négative pour β dans $]-1;1[$ et F_2 est décroissante. On s'intéresse aux limites de F_2 :

$$\lim_{\beta\to\beta_L} F_2(\beta) = F_2(\beta_L), \quad \text{et} \quad \lim_{\beta\to 1} F_2(\beta) = 0.$$

Pour assurer l'existence et l'unicité d'une solution à l'équation (3.17) il reste à montrer que :

$$F_2(\beta_L) > \frac{\Pi_R}{\Pi_L}.$$

On remarque que :

$$\Pi_L F_2(\beta_L) = \Pi_{\mathcal{R}_2}(\beta_R, \beta_L).$$

De plus dans cette zone, par décroissance de $\Pi_{\mathcal{R}_1}$ et croissance de $\Pi_{\mathcal{R}_2}$, on a $\Pi_L > \Pi^*$ et $\Pi_R > \Pi^*$. On a donc :

$$\Pi_{\mathcal{R}_2}(\beta_R, \beta_L) > \Pi_{\mathcal{R}_2}(\beta_R, \beta^*) = \Pi_R.$$

On a donc montré que l'équation $\dfrac{\Pi_R}{\Pi_L} = F_2(\beta^*)$ admet une unique solution β^* dans cette zone. En conséquence, il existe un unique état intermédiaire W^* qui connecte W_L et W_R respectivement par une 1-détente et une 2-détente.

Zone 3 : 1-choc, 2-choc

On cherche donc β^* solution de (3.18). Pour établir que cette équation admet une unique solution, on montre tout d'abord que F_3 est monotone décroissante puis finalement qu'elle passe bien par $\dfrac{\Pi_R}{\Pi_L}$. On pose :

$$F_3(\beta) = P_L(\beta) P_R(\beta),$$

avec :

$$P_L(\beta) = \frac{\left[\sqrt{\Delta(\beta, \beta_L)} + 2(\beta_L - \beta)\right]^2}{3(1-\beta_L^2)(1-\beta^2)},$$

$$P_R(\beta) = \frac{\left[\sqrt{\Delta(\beta, \beta_R)} + 2(\beta_R - \beta)\right]^2}{3(1-\beta^2)(1-\beta_R^2)},$$

avec $\Delta(\beta, \beta_L)$ défini par (3.13). En dérivant, on obtient :

$$F_3'(\beta) = P_L(\beta) P_R(\beta) \left[\frac{4(\beta\beta_L - 1)}{(1-\beta^2)\sqrt{\Delta(\beta, \beta_L)}} + \frac{4(\beta\beta_R - 1)}{(1-\beta^2)\sqrt{\Delta(\beta, \beta_R)}} \right].$$

Comme $P_L(\beta)P_R(\beta)$ est positif pour tout β dans $]-1;1[$, on voit que $F_3'(\beta)$ est négative et donc que F_3 est décroissante. Il reste alors à étudier les limites de F_3 :

$$\lim_{\beta \to \beta_L} F_3(\beta) = P_R(\beta_L), \quad \text{et} \quad \lim_{\beta \to -1} F_3(\beta) = +\infty.$$

Pour assurer l'existence et l'unicité de la solution de (3.18), il suffit de montrer que :

$$P_R(\beta_L) < \frac{\Pi_R}{\Pi_L}.$$

Comme dans cette zone $\Pi_L < \Pi^* < \Pi_R$ et $\Pi_L P_R(\beta_L) = \Pi_{\mathcal{S}_2}(\beta_R, \beta_L)$, on en déduit que :

$$\Pi_{\mathcal{S}_2}(\beta_R, \beta_L) < \Pi_{\mathcal{S}_2}(\beta_R, \beta^*) = \Pi_R.$$

Il en résulte que l'équation $\dfrac{\Pi_R}{\Pi_L} = F_3(\beta^*)$ admet une unique solution β^* dans cette zone. Il existe donc un unique état intermédiaire W^* qui connecte W_L et W_R respectivement par un 1-choc à gauche et un 2-choc à droite.

Zone 4 : 1-choc, 2-détente On cherche donc β^* solution de (3.19). Pour montrer que cette équation admet une unique solution, on montre tout d'abord que F_4 est monotone décroissante puis finalement qu'elle passe bien par $\dfrac{\Pi_R}{\Pi_L}$. Pour cela, on pose :

$$F_4(\beta) = H(\beta)P(\beta),$$

avec :

$$H(\beta) = \exp\left(\frac{4}{\sqrt{3}}(\operatorname{argth}\beta_R - \operatorname{argth}\beta)\right),$$
$$P(\beta) = \frac{\left[\sqrt{\Delta(\beta,\beta_L)} + 2(\beta_L - \beta)\right]^2}{3(1-\beta_L^2)(1-\beta^2)},$$

et $\Delta(\beta,\beta_L)$ défini par (3.13).
Puis en dérivant, on obtient :

$$F_4'(\beta) = \left[\frac{4(\beta\beta_L - 1)}{(1-\beta^2)\sqrt{\Delta(\beta,\beta_L)}} - \frac{4}{\sqrt{3}(1-\beta^2)}\right] F_4(\beta).$$

Comme $F_4(\beta)$ est positif quel que soit β dans $]-1;1[$, on voit que $F_4'(\beta)$ est négative et donc F_4 est décroissante. Il reste alors à étudier les limites de F_4 :

$$\lim_{\beta\to\beta_L} F_4(\beta) = H(\beta_L), \quad \text{et} \quad \lim_{\beta\to-1} F_4(\beta) = +\infty.$$

Il reste donc à montrer que :

$$H(\beta_L) < \frac{\Pi_R}{\Pi_L},$$

pour assurer l'existence et l'unicité de la solution de (3.19). Or dans cette zone on a $\Pi_L < \Pi^* < \Pi_R$, et comme $\Pi_L H(\beta_L) = \Pi_{\mathcal{R}_2}(\beta_R,\beta_L)$, on en déduit que

$$\Pi_{\mathcal{R}_2}(\beta_R,\beta_L) < \Pi_{\mathcal{R}_2}(\beta_R,\beta^*) = \Pi_R.$$

On a montré que l'équation $\dfrac{\Pi_R}{\Pi_L} = F_4(\beta^*)$ admet une unique solution β^* dans cette zone. Il existe donc un unique état intermédiaire W^* qui connecte W_L et W_R respectivement par un 1-choc à gauche et une 2-détente à droite.

□

À ce niveau, la solution exacte du problème de Riemann (3.2)-(3.3) est connue. Dans le paragraphe suivant, on implémente le schéma de Godunov pour ce système.

3.2.2 Solveur de Godunov pour le modèle M_1 gris

Dans ce paragraphe, on exhibe le schéma de Godunov pour le modèle M_1 muni d'une donnée initiale $U_0(x)$. Ce schéma numérique conservatif a été développé par Godunov dans [61]. Basé sur une approximation constante par maille de la solution, l'idée du schéma est de résoudre exactement la solution du problème de Riemann sur chaque interface du maillage et de moyenner la solution sur chaque cellule à la fin de l'itération en temps. On rappelle que l'on souhaite approcher la solution du problème de Cauchy suivant :

$$\begin{cases} \partial_t U + \partial_x F(U) = 0, \\ U(x, t=0) = U_0(x), \end{cases} \tag{3.20}$$

où $U_0 \in \mathcal{A}$ est donnée.
On a vu dans le paragraphe précédent que la solution du problème de Riemann pour ce système se compose de trois états constants séparés par deux ondes de choc ou de détente.
Pour discrétiser le domaine de calcul, on considère un maillage uniforme en temps et en espace respectivement constitué des mailles $[x_{i-\frac{1}{2}}; x_{i+\frac{1}{2}}],\ i \in \mathbb{Z}$ avec $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$ un pas d'espace constant et de mailles $[t^n; t^n + \Delta t]$ où Δt est un pas de temps. On suppose alors connue une approximation constante par morceaux

$$U^n(x) = U_i^n, \quad \text{si } x_{i-\frac{1}{2}} < x < x_{i+\frac{1}{2}},$$

de la solution $U(., t^n) = (E_R(., t^n), F_R(., t^n))^T$ de (3.20) au temps t^n. On définit l'approximation U^{n+1} au temps $t^{n+1} = t^n + \Delta t$ de la solution $U(., t^{n+1})$ de la manière suivante :

1. On résout le problème de Cauchy :

$$\begin{cases} \partial_t U + \partial_x F(U) = 0, \quad x \in \mathbb{R},\ t > 0, \\ U(x, t = 0) = U_i^n \quad \text{si } x_{i-\frac{1}{2}} < x < x_{i+\frac{1}{2}}. \end{cases} \tag{3.21}$$

   La solution exacte, notée $U^h(x,t)$, est connue pour $t \in]0, \Delta t[$ avec Δt assez petit pour éviter les interactions.

2. On projette la solution $U^h(., \Delta t)$ sur l'espace des fonctions constantes par morceaux pour définir :

$$U_i^{n+1} = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} U(x, \Delta t) dx. \tag{3.22}$$

Plus précisément, concernant la seconde étape 2, la solution exacte $U^h(x,t)$ de (3.21) est la juxtaposition sans interaction des solutions des problèmes de Riemann posés à chaque interface $x_{i+\frac{1}{2}}$:

$$U^h(x,t) = U_{\mathcal{R}}\left(\frac{x - x_{i+\frac{1}{2}}}{t}; U_i^n; U_{i+1}^n\right), \quad x_i < x < x_{i+1}, \quad i \in \mathbb{Z}, \quad 0 < t < \Delta t. \tag{3.23}$$

Toutefois, pour éviter que les ondes issues des nœuds $x_{i-\frac{1}{2}}$ et $x_{i+\frac{1}{2}}$ ne s'intersectent et ainsi de s'assurer que les problèmes de Riemann locaux n'interagissent pas, on suppose que le pas de temps est limité par la condition CFL suivante :

$$\frac{\Delta t}{\Delta x} \max_{i \in \mathbb{Z}} |\lambda^{\pm}(U_i^n)| \leq \frac{1}{2}, \tag{3.24}$$

où $\lambda^{\pm}$ définies par (3.6) sont les valeurs propres du système (3.2).
Pour obtenir une forme plus explicite de ce schéma, on intègre l'équation (3.21) sur le rectangle $[x_{i-\frac{1}{2}}; x_{i+\frac{1}{2}}] \times [0; \Delta t]$ pour obtenir :

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} (U(x, \Delta t) - U(x, 0))\, dx + \int_0^{\Delta t} \left(F(U(x_{i+\frac{1}{2}}, t)) - F(U(x_{i-\frac{1}{2}}, 0))\right) dt = 0. \tag{3.25}$$

En introduisant les relations (3.22) et (3.23) dans cette égalité (3.25), on obtient :

$$U_i^{n+1} = U_i^n - \frac{\Delta t}{\Delta x} \left[F(U_{\mathcal{R}}(0^+; U_i^n, U_{i+1}^n)) - F(U_{\mathcal{R}}(0^+; U_{i-1}^n, U_i^n))\right]. \tag{3.26}$$

où $U_{\mathcal{R}}(0^+; U_i^n, U_{i+1}^n)$ est la solution du problème de Riemann (3.2)-(3.3) évaluée sur la courbe caractéristique $x = 0$ avec la donnée initiale suivante :

$$\begin{cases} U_L = U_i^n, \\ U_R = U_{i+1}^n. \end{cases}$$

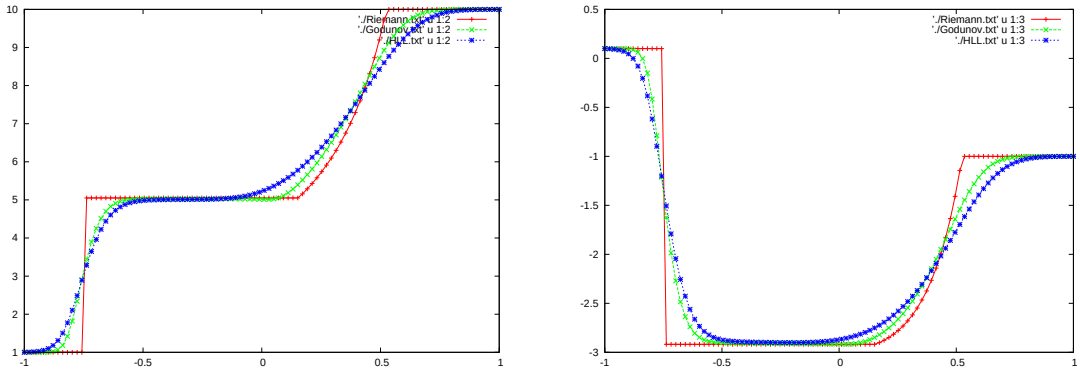


FIGURE 3.4 – Comparaison de l'énergie radiative (gauche) et du flux radiatif (droite) à $t = 1s$ sur un cas choc-détente.

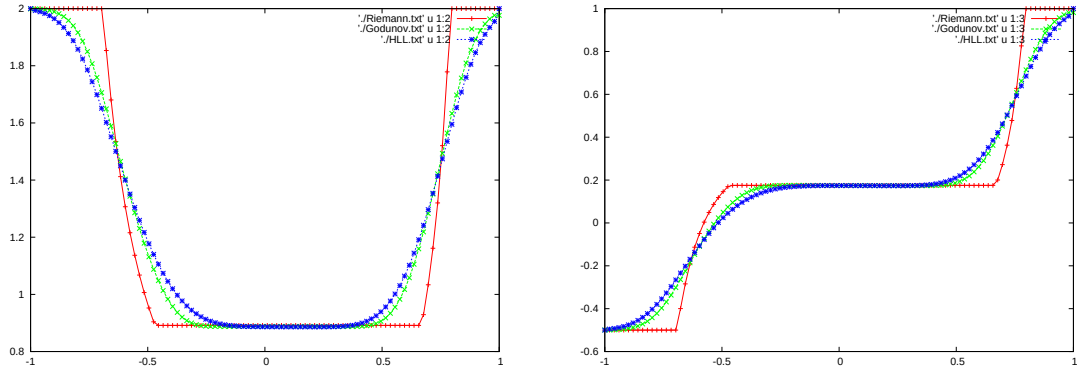


FIGURE 3.5 – Comparaison de l'énergie radiative (gauche) et du flux radiatif (droite) à $t = 1s$ sur un cas détente-détente.

Ce schéma volumes finis est conservatif par construction et d'ordre un en temps et en espace.
Pour illustrer ce schéma, on le compare à présent sur un problème de Riemann avec la solution exacte du problème et un schéma HLL.
On suppose pour ces cas-test que $c = 1$. On considère un maillage uniforme formés de 100 points discrétisant l'espace $[-1; 1]$. On compare les solutions données par le schéma de Godunov, le schéma HLL [70] présenté dans le paragraphe 3.2.4 pour une CFL égale à $\frac{1}{2}$ avec la solution exacte donnée que l'on a obtenue dans le paragraphe précédent. La donnée initiale pour le premier cas est donnée par :

$$U_0(x) = \begin{cases} U_L = (1, 0.1)^T \text{ si } x < 0, \\ U_R = (10, -1)^T \text{ sinon.} \end{cases}$$

Sur la figure 3.4, on présente l'énergie radiative et le flux radiatif au bout d'un temps $t = 1s$. On voit bien qu'avec la même discrétisation, le schéma HLL est moins précis que le schéma de Godunov. En effet, le schéma HLL ne donne qu'une approximation de l'état intermédiaire dans le solveur de Riemann alors que dans le schéma de Godunov, on utilise la solution exacte.
Pour le second cas, la donnée initiale est donnée par :

$$U_0(x) = \begin{cases} U_L = (2, -0.5)^T \text{ si } x < 0, \\ U_R = (2, 1)^T \text{ sinon.} \end{cases}$$

Sur la figure 3.5, on observe comme précédemment que le schéma HLL donne une moins bonne approximation de la solution que le schéma de Godunov.

3.2.3 Difficultés de l'extension en 2D

Si l'on s'intéresse au système en 2D, on peut projeter ce système dans la direction normale à l'interface courante et ainsi se ramener à la résolution d'un système 1D :

$$\begin{cases} \partial_t E + \partial_x F_x = 0, \\ \partial_t F_x + c^2 \partial_x P_{xx} = 0, \\ \partial_t F_y + c^2 \partial_x P_{xy} = 0. \end{cases} \tag{3.27}$$

L'algèbre de ce système est connue ainsi que la nature des champs [23]. On sait que la solution du problème de Riemann associé à ce système est composée de trois états constants séparés par trois ondes : une première onde de choc ou de détente, une onde de contact et une onde de choc ou de détente. De plus, on sait que les invariants de Riemann sur le champ Linéairement Dégénéré, associé à la valeur propre $\lambda = c\beta_x$, correspondant à l'onde de contact sont :

$$\Pi = \frac{1-\chi(f)}{2} E, \tag{3.28}$$

$$\beta_x = \frac{3\chi(f)-1}{2} \frac{F_x}{\|F_x\|} cE. \tag{3.29}$$

Toutefois, la prise en compte de ces invariants de Riemann β et Π ne permet pas d'exhiber explicitement les courbes de détente et les courbes de choc comme cela a pu être fait dans le cas 1D. En conséquence, la solution du problème de Riemann pour le modèle M_1 reste inconnue. Cependant, la connaissance de l'algèbre du système et des invariants de Riemann dans l'onde de contact permet par exemple d'écrire des schémas numériques comme le schéma HLLC [14] qui consiste à définir deux états intermédiaires dans le solveur de Riemann approché caractérisés par la constance des invariants de Riemann dans l'onde de contact.

On s'intéresse à présent à un schéma préservant l'asymptotique pour le modèle M_1.

3.2.4 Schéma préservant l'asymptotique pour le modèle M_1 gris

On rappelle que le modèle M_1 gris est donné par :

$$\begin{cases} \partial_t E_R + \nabla . F_R = c(\sigma^e a T^4 - \sigma^a E_R), \\ \partial_t F_R + c^2 \nabla . (D_R E_R) = -c\sigma^f F_R, \end{cases} \tag{1.22}$$

complété de l'équation matière :

$$\rho C_v \partial_t T = -c(\sigma^e a T^4 - \sigma^a E_R). \tag{1.6}$$

où les moyennes d'opacités σ^e, σ^a et σ^f sont données par (1.23).
On a vu dans le paragraphe 1.4.1 que ce modèle possède de bonnes propriétés physiques et la bonne limite de transport. Dans le régime de diffusion, c'est-à-dire quand le produit $\sigma_\nu t$ devient grand, le modèle dégénère vers l'équation :

$$\partial_t \left(\rho C_v T + aT^4\right) - \partial_x \left(\frac{c}{3\sigma^0} \partial_x (aT^4)\right) = 0, \tag{3.1}$$

où σ^0 est une moyenne d'opacité définie par :

$$\begin{aligned} \sigma^0 &= \lim_{F_R \to 0} \sigma^f \\ &= \frac{\int_0^\infty \sigma_\nu \partial_t \mathcal{B}_\nu(T) d\nu}{\int_0^\infty \partial_t \mathcal{B}_\nu(T) d\nu}. \end{aligned}$$

On rappelle que dans le régime de diffusion, l'équation du transfert radiatif dégénère vers la même équation en remplaçant l'opacité moyenne σ^0 par la moyenne d'opacité de Rosseland σ^R définie par :

$$\sigma^R = \frac{\int_0^\infty \partial_t \mathcal{B}_\nu(T) d\nu}{\int_0^\infty \frac{1}{\sigma_\nu} \partial_t \mathcal{B}_\nu(T) d\nu}.$$

On remarque immédiatement que ces deux moyennes mises en jeu par chaque modèle sont égales lorsque l'opacité σ_ν est constante. Cependant, quand l'opacité σ_ν n'est pas constante, $\sigma^0 \neq \sigma^R$, mais σ^0 reste une bonne approximation de la moyenne de Rosseland (voir [33] pour une comparaison de ces moyennes).
Si l'on considère un schéma numérique classique pour discrétiser (1.22), la limite asymptotique de ce schéma dans le régime de diffusion n'est, en général, pas consistante avec la limite du modèle (3.1). L'idée ici est de construire un schéma qui soit consistant avec le modèle dans les régimes intermédiaires et dans les régimes limites. De tels schémas ont été développés par exemple dans [68] pour des systèmes hyperboliques ou dans [5] pour les équations de shallow water ou encore dans [62, 24, 26] pour le transfert radiatif. On s'intéresse plus particulièrement au schéma *asymptotic preserving* développé dans [19, 17]. L'idée de ce schéma est d'introduire directement le terme source du système dans le solveur de Riemann approché du modèle de transport considéré lors de la discrétisation. Pour cela, on adopte le formalisme d'écriture :

$$\partial_t U + \partial_x F(U) = \kappa(R(U) - U), \tag{3.30}$$

où $U \in \mathcal{A}$ est le vecteur des inconnues et $\mathcal{A}$ un espace convexe, F est le flux associé, κ est un paramètre qui peut dépendre de l'espace et $R : \mathcal{A} \to \mathcal{A}$ est une fonction vectorielle régulière.
Puis, grâce à l'introduction d'un paramètre de correction dans ce schéma, on obtient la consistance du schéma avec l'équation de diffusion. De plus, cette correction est telle qu'il est possible de violer la limite asymptotique du modèle pour retrouver la limite asymptotique physique de l'équation du transfert radiatif. En d'autres termes, il sera possible de retrouver un régime asymptotique gouverné par une opacité de Rosseland et non plus par son approximation σ^0.
Le modèle M_1 écrit sous la forme (3.30) est donné par :

$$U = \begin{pmatrix} E_R \\ F_R \\ T \end{pmatrix}, F(U) = \begin{pmatrix} F_R^x \\ c^2 \chi(f) E_R \\ 0 \end{pmatrix}, R(U) = \begin{pmatrix} \frac{\sigma^e a T^4 + \sigma_1 E_R}{\sigma^m} \\ \frac{\sigma_2 F_R}{\sigma^m} \\ \frac{\frac{\sigma^a E_R}{\rho C_v} + \sigma_3 T}{\sigma^m} \end{pmatrix},$$

en définissant $\kappa = c\sigma^m = c \max(\sigma^a, \sigma^f, \frac{\sigma^e a T^3}{\rho C_v})$ de sorte qu'il existe trois coefficients positifs σ_1, σ_2 et σ_3 définis par :

$$\begin{cases} \sigma_1 &= \sigma^m - \sigma^a, \\ \sigma_2 &= \sigma^m - \sigma^f, \\ \sigma_3 &= \sigma^m - \frac{\sigma^e a T^3}{\rho C_v}, \end{cases}$$

tels que les modèles (3.30) et (1.22) coïncident. Il existe d'autres formulations possibles du formalisme (3.30), mais celle-ci est la plus pertinente. On verra dans le paragraphe sur la correction du schéma, l'intérêt d'autres formulations.

Introduction du terme source dans le solveur HLL en 1D

On commence par brièvement rappeler le principe du schéma HLL [70] pour l'équation homogène $\partial_t U + \partial_x f(U) = 0$.
On suppose que le domaine est discrétisé par un maillage uniforme formé des cellules $C_i = (x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$ avec $x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} = \Delta x$. Le centre de la cellule C_i est noté x_i. On suppose que la

solution U_i^n est constante sur chaque cellule C_i.
Dans le paragraphe 3.2.2, on a vu que le solveur exact de Godunov pour le modèle M_1 sans terme source consiste à résoudre de manière exacte une juxtaposition de problèmes de Riemann :

$$\partial_t U + \partial_x f(U) = 0, \tag{3.31}$$

pour la donnée initiale :

$$U(x, t = t^n) = U_i^n \quad \text{si } x \in [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]. \tag{3.32}$$

Étant donnée la complexité de la solution exacte de ce problème de Riemann et ainsi le coût de calcul engendré, on choisit d'utiliser un solveur de Riemann approché. Puisque les valeurs propres du système (3.30) sont comprises entre $-c$ et c, un solveur de Riemann approché en dimension un pour le modèle de transport est donné par :

$$\bar{U}_{\mathcal{R}}(\frac{x}{t}; U_L, U_R) = \begin{cases} U_L, & \text{si } \frac{x}{t} < -c, \\ U^*(U_L, U_R), & \text{si } -c < \frac{x}{t} < c, \\ U_R, & \text{si } \frac{x}{t} > c, \end{cases} \tag{3.33}$$

où U_L,U_R sont les états gauche et droit et l'état intermédiaire $U^* \in \mathcal{A}$, qui peut éventuellement être une fonction sophistiquée dépendant de x et t. Toutefois, on adopte ici la version la plus simple proposée par Harten, Lax et Van Leer [70] où U^* est une constante. Afin d'assurer la consistance du schéma de type Godunov dérivant du solveur approché (3.33), celui-ci doit satisfaire la condition suivante :

$$\frac{1}{\Delta x}\int_{C_i} \bar{U}_{\mathcal{R}}(\frac{x}{t}; U_L, U_R)dx = \frac{1}{\Delta x}\int_{C_i} U_{\mathcal{R}}(\frac{x}{t}; U_L, U_R)dx.$$

En intégrant le système homogène associé à (3.30) sur le volume de contrôle $\left(-\frac{\Delta x}{2}, \frac{\Delta x}{2}\right) \times (0, \Delta t)$, la solution exacte du problème de Riemann (3.31)-(3.32) satisfait :

$$\frac{1}{\Delta x}\int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} U_{\mathcal{R}}(\frac{x}{t}; U_L, U_R)dx = \frac{1}{2}(U_L + U_R) - \frac{\Delta t}{\Delta x}\left(F(U_R) - F(U_L)\right). \tag{3.34}$$

On déduit ainsi la définition suivante de l'état intermédiaire :

$$U^* = \frac{U_R - U_L}{2} - \frac{1}{2c}(F(U_R) - F(U_L)).$$

On approche alors la solution exacte du problème de Riemann par ce solveur approché sur chaque interface $x_{i-\frac{1}{2}}$. Toutefois, pour éviter que les ondes issues des nœuds $x_{i-\frac{1}{2}}$ et $x_{i+\frac{1}{2}}$ ne s'intersectent et ainsi que ces approximations des problèmes de Riemann locaux n'interagissent pas, on suppose que le pas de temps est limité par la condition CFL suivante :

$$\frac{c\Delta t}{\Delta x} \leq \frac{1}{2}. \tag{3.35}$$

La solution au temps $t^n + t$ est alors approchée par :

$$U^h(x, t^n + t) = U_{\mathcal{R}}\left(\frac{x - x_{i+\frac{1}{2}}}{t^n + t}; U_i^n, U_{i+1}^n\right), \quad \text{si } x \in [x_i, x_{i+1}].$$

La mise à jour U_i^{n+1} à la date $t^n + \Delta t$ est obtenue en projetant cette solution sur l'espace des fonctions constantes par morceaux de la façon suivante :

$$U_i^{n+1} = \frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} U^h(x, t^n + \Delta t)dx.$$

Finalement, en reprenant la condition de consistance (3.34), le schéma HLL usuel pour l'équation homogène est donné sous forme conservative par :

$$U_i^{n+1} = U_i^n - \frac{\Delta t}{\Delta x}\left(\mathcal{F}_{i+\frac{1}{2}} - \mathcal{F}_{i-\frac{1}{2}}\right), \tag{3.36}$$

où $\mathcal{F}$ est le flux numérique HLL défini par :

$$\mathcal{F}_{i+\frac{1}{2}} = \mathcal{F}(U_i^n, U_{i+1}^n) = \frac{F(U_i^n) + F(U_{i+1}^n)}{2} - \frac{c}{2}(U_{i+1}^n - U_i^n). \tag{3.37}$$

À présent, l'objectif est de modifier le solveur de Riemann approché (3.33) afin de prendre en compte le terme source de façon adéquate. Ici, le schéma résultant devra restituer une discrétisation pertinente des régions asymptotiques. La modification du solveur de Riemann consiste à obtenir une combinaison convexe entre l'état intermédiaire initial U^* et le terme source :

$$\tilde{U}^*_{\mathcal{R}}\left(\frac{x}{t}; U_L, U_R\right) = \begin{cases} U_L, & \text{si } \frac{x}{t} < -c, \\ \alpha U^* + (1-\alpha)R(U_L), & \text{si } -c < \frac{x}{t} < 0, \\ \alpha U^* + (1-\alpha)R(U_R), & \text{si } 0 < \frac{x}{t} > c, \\ U_R, & \text{si } \frac{x}{t} > c, \end{cases} \tag{3.38}$$

où une définition possible du coefficient $\alpha > 0$ est donnée par :

$$\alpha = \frac{2c}{2c + \kappa \Delta x}. \tag{3.39}$$

Pour que cet état intermédiaire définisse bien un solveur de Riemann approché robuste, on doit avoir $\alpha U^* + (1-\alpha)R(U) \in \mathcal{A}$ pour U dans $\mathcal{A}$. En effet, puisque $\alpha \in [0,1]$, les états intermédiaires sont une combinaison convexe entre U^* et $R(U)$. Puisque $\mathcal{A}$ est convexe, il suffit que $R(U)$ soit dans $\mathcal{A}$ pour que les états intermédiaires soient aussi dans $\mathcal{A}$.

Lemme 3.2.1. *Si $U \in \mathcal{A}$ et si*

$$\sigma^a < \sigma^f,$$

alors $R(U) = (\frac{\sigma^e a T^4 + \sigma_1 E_R}{\sigma^m}, \frac{\sigma_2 F_R}{\sigma^m}, \frac{\frac{\sigma^a E_R}{\rho C_v} + \sigma_3 T}{\sigma^m})^T \in \mathcal{A}$.

Démonstration. On suppose que U est dans l'espace des états admissibles défini par (1.12) :

$$\mathcal{A} = \{(E_R, F_R, T) \in \mathbb{R}^3, E_R \geq 0, f = \frac{|F_R|}{cE_R} \leq 1, T > 0\}.$$

On cherche donc à montrer :

$$\begin{aligned} &(i) \quad R_1(U) \geq 0, \\ &(ii) \quad \frac{|R_2(U)|}{cR_1(U)} \leq 1, \\ &(iii) \quad R_3(U) > 0. \end{aligned}$$

Les propriétés (i) et (iii) sont immédiates car toutes les quantités sont positives.
Pour la propriété (ii), on a :

$$\begin{aligned} \frac{|R_2(U)|}{cR_1(U)} &= \frac{|(\sigma_2 F_R)|}{c(\sigma^e a T^4 + \sigma_1 E_R)}, \\ &= \frac{(\sigma^m - \sigma^f)|F_R|}{c(\sigma^e a T^4 + (\sigma^m - \sigma^a)E_R)}, \\ &\leq \frac{(\sigma^m - \sigma^f)|F_R|}{(\sigma^m - \sigma^a)E_R}. \end{aligned}$$

Or $\frac{\sigma^m - \sigma^f_x}{\sigma^m - \sigma^a} \leq 1$ car $\sigma^a < \sigma^f$ et $\frac{|F_R|}{cE_R} \leq 1$ car $U \in \mathcal{A}$, ce qui complète la preuve. □

On remarque qu'avec ce choix de paramètre α, quand $\kappa = 0$, α vaut 1, et ainsi on retrouve bien le schéma HLL pour le système de transport. Puis, lorsque κ tend vers l'infini, U^* tend vers R.
À présent, on substitue la solution du problème de Riemann par le solveur de Riemann modifié (3.38) sur chaque interface $x_{i-\frac{1}{2}}$. La juxtaposition des solveurs de Riemann n'interagissent pas sous la condition CFL (3.35). La solution au temps $t^n + t$ est alors approchée par :

$$\tilde{U}^h(x, t^n + t) = \tilde{U}^*_{\mathcal{R}}\left(\frac{x - x_{i+\frac{1}{2}}}{t}; U_i^n, U_{i+1}^n\right), \quad \text{si } x \in [x_i, x_{i+1}] \text{ et } t \in]0, \Delta t[.$$

Pour calculer la solution approchée au temps $t^n + \Delta t$, on projette cette solution sur l'espace des fonctions constantes par morceaux :

$$U_i^{n+1} = \frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \tilde{U}^h(x, t^n + \Delta t)dx.$$

Un calcul long mais direct de cette intégrale donne la forme développée suivante :

$$\begin{aligned} U_i^{n+1} = U_i^n &- \frac{\Delta t}{\Delta x}\left(\alpha_{i+\frac{1}{2}}\mathcal{F}_{i+\frac{1}{2}} - \alpha_{i-\frac{1}{2}}\mathcal{F}_{i-\frac{1}{2}}\right) \\ &+ \frac{\Delta t}{\Delta x}\left[(1-\alpha_{i-\frac{1}{2}})S^+_{i-\frac{1}{2}} + (1-\alpha_{i+\frac{1}{2}})S^-_{i+\frac{1}{2}}\right], \end{aligned} \tag{3.40}$$

où S est une discrétisation du terme source donnée par :

$$\begin{aligned} S^-_{i+\frac{1}{2}} &= c\,(R(U_i^n) - U_i^n) - F(U_i^n), \\ S^+_{i-\frac{1}{2}} &= c\,(R(U_i^n) - U_i^n) + F(U_i^n), \end{aligned}$$

et $\mathcal{F}$ est le flux numérique HLL (3.37).

Remarque 1. *En régime de transport, l'équation matière sur la température est couplée du système de rayonnement. On choisira comme solveur de Riemann approché pour les deux premières équations le solveur HLL et on considère un flux nul pour l'équation de température.*

Théorème 3.2. *On suppose que tous les états U_i^n sont dans $\mathcal{A}$ pour tout $i \in \mathbb{Z}$. On suppose de plus que le pas de temps est limité par la condition CFL (3.35) et que le schéma (3.36) préserve les états admissibles. Alors, pour tout $i \in \mathbb{Z}$, l'état U_i^{n+1} défini par (3.40) est dans l'ensemble $\mathcal{A}$.*

On remarque que la condition CFL est la même que pour le schéma homogène et est donc indépendante du terme source.
Le schéma (3.47) présenté ici ne préserve pas en général l'asymptotique. Un contre-exemple sera brièvement donné dans la suite. Dans le paragraphe suivant, on propose alors une correction asymptotique du schéma lui permettant d'avoir la bonne limite diffusive.

Correction asymptotique

À présent, une correction asymptotique va être mise en place afin de forcer le schéma à restituer le régime asymptotique physique attendu. En effet, on rappelle que la limite asymptotique du modèle M_1 est donnée par :

$$\partial_t\left(\rho C_v T + aT^4\right) - \partial_x\left(\frac{c}{3\sigma^0}\partial_x(aT^4)\right) = 0, \tag{3.1}$$

où σ^0 est la moyenne d'opacité définie par :

$$\sigma^0 = \frac{\int_0^\infty \sigma(\nu)\partial_t\mathcal{B}_\nu(T)d\nu}{\int_0^\infty \partial_t\mathcal{B}_\nu(T)d\nu}. \tag{3.41}$$

Si l'on note $T_i^{0,n}$ la température dans la limite de diffusion, un développement asymptotique du schéma (3.40) dans cette limite donne :

$$Z_i^{0,n+1} = Z_i^{0,n} + c^2 \frac{\Delta t}{\Delta x^2}\Big(\frac{Z_{i+1}^{0,n} - Z_i^{0,n}}{\kappa_{i+\frac{1}{2}}} - \frac{Z_i^{0,n} - Z_{i-1}^{0,n}}{\kappa_{i-\frac{1}{2}}}\Big), \qquad \text{avec} \quad Z = \rho C_v T + aT^4, \tag{3.42}$$

où $\kappa_{i+\frac{1}{2}} = c\sigma^m_{i+\frac{1}{2}}$.
On voit bien que ce schéma approche une équation de diffusion mais pas pour le coefficient $\frac{c}{3\sigma^0}$ qui apparaît dans l'équation de diffusion du modèle (3.1). Il faut donc modifier le schéma pour qu'il admette la bonne limite dans ce régime asymptotique.
Dans cet objectif, une correction du schéma est proposée en introduisant un nouveau paramètre $\bar\kappa$ positif dans le terme source de la manière suivante :

$$\kappa(R(U) - U) = \kappa(R(U) - U) + (\bar\kappa - \bar\kappa)U,$$

si bien que le modèle se réécrit :

$$\partial_t U + \nabla F(U) = (\bar\kappa + \kappa)(\bar R(U) - U). \tag{3.43}$$

Les coefficients $\bar\kappa$ et κ étant positifs, $\bar R$ est en fait une combinaison convexe de $R(U)$ et U :

$$\bar R(U) = \frac{\kappa}{\kappa + \bar\kappa} R(U) + \frac{\bar\kappa}{\kappa + \bar\kappa} U.$$

Ainsi, si U et $R(U)$ sont dans l'espace $\mathcal{A}$, alors $\bar R$ est admissible et le schéma (3.40) est bien applicable au modèle (3.43). Le schéma issu de cette correction reste le même que (3.40) avec une nouvelle définition de la discrétisation du terme source et du paramètre α relativement à l'introduction du paramètre $\bar\kappa$:

$$\begin{aligned}
\alpha_{i+\frac{1}{2}} &= \frac{2c}{2c + \Delta x(\kappa_{i+\frac{1}{2}} + \bar\kappa_{i+\frac{1}{2}})}, \\
S^-_{i+\frac{1}{2}} &= c\frac{\kappa_{i+\frac{1}{2}}}{\kappa_{i+\frac{1}{2}} + \bar\kappa_{i+\frac{1}{2}}}\,(R(U_i^n) - U_i^n) - F(U_i^n), \\
S^+_{i-\frac{1}{2}} &= c\frac{\kappa_{i-\frac{1}{2}}}{\kappa_{i-\frac{1}{2}} + \bar\kappa_{i-\frac{1}{2}}}\,(R(U_i^n) - U_i^n) + F(U_i^n).
\end{aligned}$$

Avec cette correction, la limite asymptotique du schéma devient :

$$Z_i^{0,n+1} = Z_i^{0,n} + c^2 \frac{\Delta t}{\Delta x^2}\Big(\frac{Z_{i+1}^{0,n} - Z_i^{0,n}}{\kappa_{i+\frac{1}{2}} + \bar\kappa_{i+\frac{1}{2}}} - \frac{Z_i^{0,n} - Z_{i-1}^{0,n}}{\kappa_{i-\frac{1}{2}} + \bar\kappa_{i-\frac{1}{2}}}\Big), \qquad \text{avec} \quad Z = \rho C_v T + aT^4. \tag{3.44}$$

Il en résulte que le paramètre $\bar\kappa_{i+\frac{1}{2}}$ peut être fixé de manière à modifier la limite de diffusion du schéma. En effet, en posant :

$$\bar\kappa_{i+\frac{1}{2}} = 3c\sigma^0_{i+\frac{1}{2}}\left(1 + \rho C_v \frac{T_{i+1}^n - T_i^n}{a(T_{i+1}^n)^4 - a(T_i^n)^4}\right) - \kappa_{i+\frac{1}{2}},$$

qui est bien positif dès que $3\sigma^0_{i+\frac{1}{2}} \geq \kappa_{i+\frac{1}{2}}$ où $\sigma^0_{i+\frac{1}{2}}$ est donné par (3.41), la limite asymptotique du schéma est bien une discrétisation de l'équation limite du modèle M_1 (3.1) :

$$\begin{aligned}
\rho C_v T_i^{0,n+1} + a(T_i^{0,n+1})^4 = \rho C_v T_i^{0,n} + a(T_i^{0,n})^4 \\
- c\frac{\Delta t}{\Delta x^2}\left(\frac{a((T_{i+1}^{0,n})^4 - (T_i^{0,n})^4)}{3\sigma^0_{i+\frac{1}{2}}} - \frac{a((T_i^{0,n})^4 - (T_{i-1}^{0,n})^4)}{3\sigma^0_{i-\frac{1}{2}}}\right).
\end{aligned}$$

Cependant, le régime asymptotique de l'équation de diffusion à l'équilibre (limite physique) est donné par (3.1) où le coefficient de diffusion n'est plus σ^0 mais la moyenne de Rosseland σ^R définie par :

$$\sigma^R = \frac{\int_0^\infty \partial_t \mathcal{B}_\nu(T) d\nu}{\int_0^\infty \frac{1}{\sigma_\nu} \partial_t \mathcal{B}_\nu(T) d\nu}. \tag{3.45}$$

Étant donné que le paramètre $\bar{\kappa}_{i+\frac{1}{2}} \geq 0$ peut être choisi librement, il est possible de modifier la limite asymptotique du schéma pour qu'il dégénère non plus vers la limite mathématique du système discrétisé, mais vers la limite physique du modèle. Pour cela, on peut choisir :

$$\bar{\kappa}_{i+\frac{1}{2}} = 3c\sigma^R_{i+\frac{1}{2}} \left(1 + \rho C_v \frac{T^n_{i+1} - T^n_i}{a(T^n_{i+1})^4 - a(T^n_i)^4}\right) - \kappa_{i+\frac{1}{2}},$$

où $\sigma^R_{i+\frac{1}{2}}$ est la moyenne de Rosseland définie par (3.45).

Extension du schéma en 2D

On s'intéresse à présent à l'extension de ce schéma en 2D sur des maillages non-structurés. Par un abus de notation, on appelle de nouveau $\mathcal{A}$ l'espace des états physiquement admissibles :

$$\mathcal{A} = \{(E_R, F_R) \in \mathbb{R} \times \mathbb{R}^2, E_R > 0, f = \frac{\|F_R\|}{cE_R} \leq 1\}.$$

On écrit le modèle M_1 en 2D en respectant le formalisme (3.30) :

$$\partial_t U + \partial_x F(U) + \partial_y G(U) = \kappa(R(U) - U), \tag{3.46}$$

avec :

$$U = \begin{pmatrix} E_R \\ F_R^x \\ F_R^y \\ T \end{pmatrix}, F(U) = \begin{pmatrix} F_R^x \\ c^2 D_R^{xx} E_R \\ c^2 D_R^{xy} E_R \\ 0 \end{pmatrix}, G(U) = \begin{pmatrix} F_R^y \\ c^2 D_R^{xx} E_R \\ c^2 D_R^{yy} E_R \\ 0 \end{pmatrix}, R(U) = \begin{pmatrix} \frac{\sigma^e a T^4 + \sigma_1 E_R}{\sigma^m} \\ \frac{\sigma_2 F_R^x}{\sigma^m} \\ \frac{\sigma_3 F_R^y}{\sigma^m} \\ \frac{\sigma^a E_R}{\rho C_v} + \sigma_4 T \\ \hline \sigma^m \end{pmatrix},$$

où $\kappa = c\sigma^m$ est donné par :

$$\sigma^m = \max\left(\sigma^a, \sigma^f_x, \sigma^f_y, \frac{\sigma^e a T^3}{\rho C_v}\right).$$

de telle sorte qu'il existe quatre coefficients positifs $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ définis par :

$$\begin{aligned} \sigma_1 &= \sigma^m - \sigma^a, \\ \sigma_2 &= \sigma^m - \sigma^f_x, \\ \sigma_3 &= \sigma^m - \sigma^f_y, \\ \sigma_4 &= \sigma^m - \frac{\sigma^e a T^3}{\rho C_v}. \end{aligned}$$

Pour approcher les solutions de (3.46), on suppose que le domaine est discrétisé par un maillage non-structuré composé des polygones K de centre c_K. On note $|K|$ l'aire de ce polygone et Γ_K l'ensemble de ses voisins. On introduit également ϱ_{KL} pour désigner le segment commun aux cellules K et L et $|\varrho_{KL}|$ la longueur de ce segment. On appelle $\overrightarrow{n}^{KL} = (n_x^{KL}, n_y^{KL})$ la normale unitaire sortant de K au segment ϱ_{KL} (voir figure 3.6).

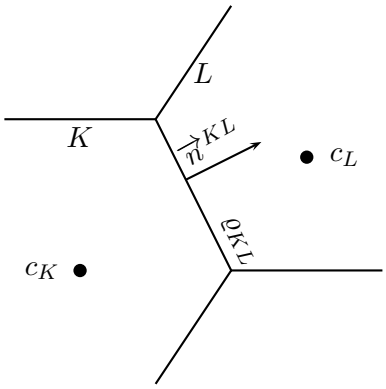


FIGURE 3.6 – Géométrie du maillage non-structuré.

On propose alors une extension 2D du schéma (3.40). En conséquence, sur la maille K, la mise à jour de la solution à la date t^{n+1} sera donnée par :

$$\begin{aligned} U_K^{n+1} = U_K^n - \frac{\Delta t}{|K|} \sum_{L\in\Gamma_K} |\varrho_{KL}|\, \alpha_{KL}\, \mathcal{H}(U_K^n, U_L^n, \overrightarrow{n}^{KL}) \\ + \frac{\Delta t}{|K|} \sum_{L\in\Gamma_K} |\varrho_{KL}|\, \alpha_{KL}\, \mathcal{H}(U_K^n, U_K^n, \overrightarrow{n}^{KL}) + \Delta t\, \mathcal{S}_K \end{aligned} \tag{3.47}$$

où $\mathcal{S}_K$ est une discrétisation du terme source définie par :

$$\mathcal{S}_K = \frac{c}{|K|} \sum_{L\in\Gamma_K} |\varrho_{KL}|\, (1-\alpha_{KL})\, (R(U_K^n) - U_K^n)\,, \tag{3.48}$$

et $\mathcal{H}(U_K^n, U_L^n, \overrightarrow{n}^{KL})$ est le flux numérique dans la direction normale à l'interface ϱ_{KL} donné par :

$$\mathcal{H}(U_K^n, U_L^n, \overrightarrow{n}^{KL}) = \mathcal{F}(U_K^n, U_L^n) n_x^{KL} + \mathcal{G}(U_K^n, U_L^n) n_y^{KL}. \tag{3.49}$$

en notant $\mathcal{F}$ et $\mathcal{G}$ les flux HLL associés respectivement à F et G.
Le paramètre α_{KL}, initialement donné en 1D par (3.39), est à présent étendu au cas 2D de la façon suivante :

$$\alpha_{KL} = \frac{c\, \sharp\Gamma_K}{c\, \sharp\Gamma_K + \kappa_{KL} \frac{|K|}{|\varrho_{KL}|}}, \tag{3.50}$$

où $\sharp\Gamma_K$ désigne le nombre de segments formant la maille K.
De même, une extension de la condition CFL est donnée par :

$$\frac{c\Delta t}{|K|} \sum_{L\in\Gamma_K} |\varrho_{KL}| \leq \frac{1}{2}. \tag{3.51}$$

Cette CFL trouvera une signification précise dans le lemme 3.2.3.

Remarque 2. *Comme en 1D, l'équation matière sur la température n'est couplée au système que via le terme source. On suppose de nouveau que dans la partie hyperbolique, la température est constante durant un pas de temps Δt :*

$$T_i^{n+1} = T_i^n.$$

En conséquence, la composante correspondant à la température dans le flux $\mathcal{H}$ est nulle.

À présent, on montre brièvement la consistance de ce schéma (3.47) avec le système d'équations (3.46). On se place dans le cas d'un maillage régulier. En notant r_K le rayon du cercle inscrit dans la maille K, on suppose que la limite suivante est toujours satisfaite :

$$\lim_{r_K \to 0} \frac{|K|}{|\varrho_{KL}|} = 0.$$

On en déduit immédiatement que la limite du paramètre α_{KL} quand r_K tend vers 0 est :

$$\lim_{r_K \to 0} \alpha_{KL} = 1.$$

De plus, par application de la formule de Green, on a :

$$\sum_{L \in \Gamma_K} |\varrho_{KL}| \mathcal{H}(U_K^n, U_K^n, \overrightarrow{n}^{KL}) = 0.$$

Il en résulte que la partie hyperbolique du schéma (3.47) est consistante avec la partie hyperbolique du système (3.46). Concernant le terme source, on établit la consistance dans le cas particulier κ constant.

Lemme 3.2.2. *On suppose que $\kappa_{KL} > 0$ est une constante donnée. La discrétisation du terme source introduite (3.48) est consistante avec le terme source du système (3.46) au sens suivant :*

$$\lim_{r_K \to 0} \mathcal{S}_K = \kappa_{KL} \left(R(U_K) - U_K \right).$$

Démonstration. En remplaçant α_{KL} par sa valeur dans (3.48), on obtient :

$$\mathcal{S}_K = c\kappa_{KL} \left(R(U_K) - U_K \right) \sum_{L \in \Gamma_K} \frac{1}{c\sharp\Gamma_K + \kappa_{KL} \frac{|K|}{|\varrho_{KL}|}}$$

En calculant la limite quand le rayon du cercle inscrit r_K tend vers 0, on a :

$$\lim_{r_K \to 0} \sum_{L \in \Gamma_K} \frac{1}{c\sharp\Gamma_K + \kappa_{KL} \frac{|K|}{|\varrho_{KL}|}} = \frac{1}{c}.$$

Finalement, on en déduit que :

$$\lim_{r_K \to 0} \mathcal{S}_K = \kappa_{KL} \left(R(U_K) - U_K \right).$$

La discrétisation $\mathcal{S}_K$ du terme source dans le schéma (3.48) est donc consistante avec le terme source du système d'équations (3.46). □

Il est également possible d'établir la robustesse du schéma (3.47). Pour cela, on montre dans le lemme suivant que le schéma (3.47) s'écrit comme une combinaison convexe de schémas 1D donnés dans la direction normale de l'interface ϱ_{KL}.

Lemme 3.2.3. *Le schéma (3.47) peut s'écrire comme une combinaison convexe de schémas 1D (3.40) écrits dans la direction normale à l'interface ϱ_{KL}. De plus, sous la condition CFL (3.51), le schéma (3.47) préserve les états physiquement admissibles.*

Démonstration. Soit U_{KL}^{n+1}, un état intermédiaire donné par le schéma 1D (3.40) dans la direction de l'interface ϱ_{KL} de la manière suivante :

$$\begin{aligned} U_{KL}^{n+1} = U_K^n - \frac{\Delta t}{\delta_{KL}} \left(\alpha_{KL} \mathcal{H}(U_K^n, U_L^n, \overrightarrow{n}^{KL}) - \alpha_{KK} \mathcal{H}(U_K^n, U_K^n, \overrightarrow{n}^{KL}) \right) \\ + \frac{\Delta t}{\delta_{KL}} \left[(1 - \alpha_{KL}) \left(S_K - \mathcal{H}(U_K^n, U_K^n, \overrightarrow{n}^{KL}) \right) + (1 - \alpha_{KK}) \left(S_K + \mathcal{H}(U_K^n, U_K^n, \overrightarrow{n}^{KL}) \right) \right], \end{aligned} \tag{3.52}$$

où δ_{KL} est une distance relative à la cellule K qui sera fixée plus tard, $\mathcal{H}$ est le flux numérique donné par :

$$\mathcal{H}(U_K^n, U_L^n, \overrightarrow{n}^{KL}) = \mathcal{F}(U_K^n, U_L^n) n_x^{KL} + \mathcal{G}(U_K^n, U_L^n) n_y^{KL},$$

α_{KL} est un paramètre défini par :

$$\begin{cases} \alpha_{KL} & = \frac{c \sharp\Gamma_K}{c \sharp\Gamma_K + \kappa_{KL} \frac{|K|}{|\varrho_{KL}|}}, \\ \alpha_{KK} & = 1, \end{cases}$$

et S_K est la discrétisation du terme source donnée par :

$$S_K = c\,(R(U_K^n) - U_K^n)\,. \tag{3.53}$$

On rappelle que le schéma (3.52) préserve les états admissibles sous la condition CFL :

$$\frac{c\Delta t}{\delta_{KL}} \leq \frac{1}{2}. \tag{3.54}$$

En conséquence, sous cette restriction CFL, dès que U_K^n et U_L^n sont dans $\mathcal{A}$, U_K^{n+1} est également dans $\mathcal{A}$.
On introduit ensuite des paramètres strictement positifs θ_{KL} tels que

$$\sum_{L\in\Gamma_K} \theta_{KL} = 1.$$

On considère alors la combinaison convexe suivante :

$$\begin{aligned} \sum_{L\in\Gamma_K} \theta_{KL} U_{KL}^{n+1} = U_K^n &- \Delta t \sum_{L\in\Gamma_K} \frac{\theta_{KL}}{\delta_{KL}}\, \alpha_{KL}\, \mathcal{H}(U_K^n, U_L^n, \overrightarrow{n}^{KL}) \\ &+ \Delta t \sum_{L\in\Gamma_K} \frac{\theta_{KL}}{\delta_{KL}}\, \mathcal{H}(U_K^n, U_K^n, \overrightarrow{n}^{KL}) \\ &+ \Delta t \sum_{L\in\Gamma_K} \frac{\theta_{KL}}{\delta_{KL}}\, (1-\alpha_{KL})\, \left(S_K - \mathcal{H}(U_K^n, U_K^n, \overrightarrow{n}^{KL})\right). \end{aligned} \tag{3.55}$$

Il reste alors à déterminer les coefficients θ_{KL} et δ_{KL}. On propose les définitions suivantes :

$$\begin{aligned} \delta_{KL} &= \frac{|K|}{\sum_{L\in\Gamma_K} |\varrho_{KL}|} > 0, \\ \theta_{KL} &= \frac{|\varrho_{KL}|}{\sum_{L\in\Gamma_K} |\varrho_{KL}|} > 0. \end{aligned}$$

On constate immédiatement que $\sum_{L\in\Gamma_K} \theta_{KL} = 1$. De plus, on remarque la relation suivante :

$$\frac{\theta_{KL}}{\delta_{KL}} = \frac{|\varrho_{KL}|}{|K|}. \tag{3.56}$$

Ainsi, la relation (3.55) trouve l'expression suivante :

$$\begin{aligned} \sum_{L\in\Gamma_K} \theta_{KL} U_{KL}^{n+1} = U_K^n &- \frac{\Delta t}{|K|} \sum_{L\in\Gamma_K} |\varrho_{KL}|\, \alpha_{KL}\, \mathcal{H}(U_K^n, U_L^n, \overrightarrow{n}^{KL}) \\ &+ \frac{\Delta t}{|K|} \sum_{L\in\Gamma_K} |\varrho_{KL}|\, \alpha_{KL}\, \mathcal{H}(U_K^n, U_K^n, \overrightarrow{n}^{KL}) \\ &+ \frac{\Delta t}{|K|} \sum_{L\in\Gamma_K} |\varrho_{KL}|\, (1-\alpha_{KL})\, S_K. \end{aligned}$$

En tenant compte de l'expression du schéma (3.47), on obtient alors la combinaison convexe cherchée :

$$U_K^{n+1} = \sum_{L\in\Gamma_K} \theta_{KL} U_{KL}^{n+1}.$$

La condition CFL (3.54) se réécrit alors de la façon suivante :

$$\frac{c\Delta t}{|K|} \sum_{\varrho\in\Gamma_K} |\varrho_{KL}| \leq \frac{1}{2}, \tag{3.51}$$

et assure l'admissibilité de tous les états intermédiaires U_{KL}^{n+1}. Puisque $\mathcal{A}$ est convexe, il en résulte que U_K^{n+1} est encore dans l'ensemble $\mathcal{A}$. □

De même qu'en 1D, un développement asymptotique du schéma (3.47) illustre le fait que ce schéma ne préserve pas l'asymptotique. Ainsi, dans la section suivante, on propose une correction asymptotique de ce schéma.

Correction asymptotique en 2D

On va, à présent, introduire une correction dans le schéma (3.47) afin de préserver la bonne limite de diffusion. On rappelle que l'équation de diffusion du modèle M_1 en 2D est donnée par :

$$\partial_t \left(\rho C_v T + aT^4\right) - \operatorname{div}\left(\frac{c}{3\sigma^0}\nabla(aT^4)\right) = 0, \tag{3.57}$$

où σ^0 est définie par (3.41).
On réalise un développement asymptotique du schéma (3.47) dans le régime de diffusion. En notant $T_K^{0,n}$ la température dans le régime de diffusion, la limite asymptotique du schéma (3.47) est donnée par :

$$\begin{aligned} Z_K^{0,n+1} &= Z_K^{0,n} + \frac{\Delta t}{|K|^2} \sum_{L\in\Gamma_K} \frac{c^2|\varrho_{KL}|^2}{\kappa_{KL}} \left(a(T_L^{0,n})^4 - a(T_K^{0,n})^4\right), \\ \text{avec} \quad Z &= \rho C_v T + aT^4. \end{aligned} \tag{3.58}$$

Il est clair que ce schéma discrétise une équation de diffusion mais pour un coefficient de diffusion différent de celui attendu. En procédant de façon similaire au cas 1D, on propose alors d'introduire un paramètre $\bar{\kappa}$ positif dans le terme source de la manière suivante :

$$\begin{aligned} \kappa(R(U) - U) &= \kappa(R(U) - U) + (\bar{\kappa} - \bar{\kappa})U \\ &= (\bar{\kappa} + \kappa)(\bar{R}(U) - U), \end{aligned}$$

pour obtenir une nouvelle formulation de la discrétisation du terme source et du paramètre α_{KL} qui deviennent alors :

$$\begin{aligned} \alpha_{KL} &= \frac{c\,\sharp\Gamma_K}{c\,\sharp\Gamma_K + (\kappa_{KL} + \bar{\kappa}_{KL})\frac{|K|}{|\varrho_{KL}|}}, \\ S_{KL} &= c\frac{\kappa_{KL}}{\kappa_{KL} + \bar{\kappa}_{KL}} \left(R(U_K^n) - U_K^n\right). \end{aligned}$$

Pour ce choix de paramètre, la limite asymptotique du schéma (3.47) devient :

$$\begin{aligned} Z_K^{0,n+1} &= Z_K^{0,n} + \frac{\Delta t}{|K|^2} \sum_{\varrho\in\Gamma_K} \frac{c^2|\varrho_{KL}|^2}{\kappa_{KL} + \bar{\kappa}_{KL}} \left(a(T_L^{0,n})^4 - a(T_K^{0,n})^4\right), \\ \text{avec} \quad Z &= \rho C_v T + aT^4. \end{aligned} \tag{3.59}$$

Alors qu'en 1D un choix naturel pour $\bar{\kappa}$ se présentait afin de retrouver la bonne limite asymptotique, ce n'est plus le cas en dimension supérieure. En effet, si l'on veut que le schéma (3.58) discrétise l'équation de diffusion du modèle (3.57), il est nécessaire de satisfaire l'approximation suivante :

$$\frac{1}{|K|}\frac{c^2|\varrho_{KL}|^2}{\kappa_{KL}+\bar{\kappa}_{KL}}\left(a(T_L^{0,n})^4-a(T_K^{0,n})^4\right)\simeq\int_{\varrho_{KL}}\frac{c}{3\sigma_{KL}^0}\nabla aT^4(t^n,x,y).\overrightarrow{n}\,d\varrho. \tag{3.60}$$

En effet, on aura alors :

$$\sum_{L\in\Gamma_K}\frac{1}{|K|}\frac{c^2|\varrho_{KL}|^2}{\kappa_{KL}+\bar{\kappa}_{KL}}\left(a(T_L^{0,n})^4-a(T_K^{0,n})^4\right)\simeq\int_K \text{div}\left(\frac{c}{3\sigma_K^0}\nabla(aT^4)\right)dxdy.$$

Toutefois, dans (3.60), le gradient n'est, en général, pas connu. Cependant, il peut être approché par différentes techniques. Par exemple, on notera les schémas d'ordre élevé de type DDFV [40]. L'approximation de tels gradients n'est pas abordée ici, mais la correction asymptotique développée ci-dessous peut être appliquée à d'autres systèmes dont les gradients sont correctement évalués. Si l'on pose ∇_{mat}^h une approximation de ce gradient :

$$\nabla_{mat}^h\simeq\int_{\varrho_{KL}}\frac{c}{3\sigma_{KL}^0}\nabla aT^4(t^n,x,y).\overrightarrow{n}\,d\varrho,$$

un choix possible pour le paramètre $\bar{\kappa}_{KL}$ est donné par :

$$\bar{\kappa}_{KL}=\frac{c|\varrho_{KL}|}{|K|\nabla_{mat}^h}\left(a(T_L^{0,n})^4-a(T_K^{0,n})^4\right)-\kappa_{KL}.$$

En fait, on a forcé le coefficient de diffusion $\frac{c}{3\sigma^0}$ du modèle pour retrouver la limite mathématique du modèle. On remarque alors que l'on peut forcer le schéma à dégénérer non plus vers cette limite mathématique mais vers la limite physique du modèle caractérisée par la moyenne de Rosseland σ^R. Ainsi, en prenant

$$\bar{\kappa}_{KL}=\frac{c|\varrho_{KL}|}{\nabla_{phy}^h|K|}\left(a(T_L^{0,n})^4-a(T_K^{0,n})^4\right)-\kappa_{KL},$$

où ∇_{phy}^h est une approximation du gradient :

$$\nabla_{phy}^h\simeq\int_{\varrho_{KL}}\frac{c}{3\sigma_{KL}^R}\nabla aT^4(t^n,x,y).\overrightarrow{n}\,d\varrho,$$

on va ainsi violer la limite mathématique du système pour retrouver la limite physique.

Dans le cas particulier des maillages admissibles, c'est-à-dire où les segments c_Kc_L sont orthogonaux aux interfaces ϱ_{KL}, une approximation naturelle des gradients normaux est donnée par :

$$\nabla aT^4.\overrightarrow{n}^{KL}\simeq\frac{aT_L^4-aT_K^4}{|c_Kc_L|}.$$

En prenant :

$$\nabla_{phy}^h=\frac{c}{3\sigma_{KL}^R}|\varrho_{KL}|\frac{aT_L^4-aT_K^4}{|c_Kc_L|},$$

on obtient alors

$$\bar{\kappa}_{KL}=3c\sigma_{KL}^R\frac{|\varrho_{KL}|}{|K|}-\kappa_{KL}.$$

Finalement, pour illustrer ce schéma, quelques cas-tests sont présentés pour le modèle M_1 multigroupe dans le paragraphe 5.2.

3.3 Précalculs pour le modèle M_1 multigroupe

3.3.1 Introduction

On rappelle que le modèle M_1 multigroupe couplé à la matière est donné par :

$$\begin{cases} \partial_t E_q + \nabla.F_q = c(\sigma_q^e a\theta_q^4(T) - \sigma_q^a E_q), \\ \partial_t F_q + c^2\nabla.(D_q E_q) = -c\sigma_q^f F_q, \qquad \forall 1 \leq q \leq Q \\ \rho C_v \partial_t T = -c\sum_{q=1}^{Q}(\sigma_q^e a\theta_q^4 - \sigma_q^a E_q), \end{cases} \tag{1.31}$$

où l'on suppose que E_q, F_q et θ sont dans l'espace des états admissibles donné par :

$$\mathcal{A} = \{(E_q, F_q) \in \mathbb{R}^4, E_q > 0, f = \frac{\|F_q\|}{cE_q} \leq 1, \theta_q > 0\}. \tag{1.12}$$

Étant donné que le modèle M_1 est isotrope dans les directions orthogonales au flux, on peut toujours se ramener au cas 1D dans la direction du flux. Ainsi, pour simplifier les calculs, on considère ici le modèle en 1D :

$$\begin{cases} \partial_t E_q + \partial_x Fq = c(\sigma_q^e a\theta_q^4(T) - \sigma_q^a E_q), \\ \partial_t F_q + c^2\partial_x(\chi_q E_q) = -c\sigma_q^f F_q, \end{cases} \quad \forall 1 \leq q \leq Q.$$

Pour un couple énergie-flux $(\mathcal{E}_q, \mathcal{F}_q)$ donné dans le groupe $q = [\nu_q; \nu_{q+1}]$ et physiquement admissible, la fermeture du modèle selon le principe de minimisation de l'entropie permet d'exprimer la pression radiative en fonction des deux premiers moments :

$$P_q = < \mu^2 \mathcal{I} >_q = P_q(\mathcal{E}_q, \mathcal{F}_q),$$

où $\mathcal{I}$, la solution du problème de minimisation

$$H_R(\mathcal{I}) = \min\{H_R(I) = \sum_q < h_R(I) >_q / \forall q, < I >_q = E_q \text{ et } < c\Omega I >_q = F_q\}, \tag{1.29}$$

est donnée dans le $q^{ème}$ groupe par :

$$\mathcal{I}_q(\mu, \nu) = \frac{2h\nu^3}{c^2}\left[\exp\left(\frac{h\nu}{kT}\alpha_q(1 - \beta_q\mu)\right) - 1\right]^{-1}.$$

Les coefficients α_q et β_q sont les multiplicateurs de Lagrange associés au problème de minimisation avec contraintes (1.29). Ces coefficients sont dans l'espace [106] :

$$\mathcal{L} = \{\alpha \geq 0; -1 \leq \beta \leq 1\}.$$

Numériquement, le calcul de χ_q et des moyennes d'opacités σ_q^e, σ_q^a et σ_q^f est nécessaire pour chaque couple énergie-flux $(\mathcal{E}_q, \mathcal{F}_q)$. Étant donnée l'occurrence du calcul de χ et des opacités dans le code de calcul, l'idée est de pré-calculer ces coefficients pour un facteur d'anisotropie $f_q = \frac{\mathcal{F}_q}{c\mathcal{E}_q}$ et une température radiative ϑ_q dans le groupe q donnés. En effet, la relation de fermeture donne une expression de P_q en fonction des deux premiers moments E_q et F_q. On préfère considérer les deux variables équivalentes f (facteur d'anisotropie) et ϑ (température radiative) qui sont plus faciles à appréhender physiquement. Ces valeurs seront enregistrées dans des tableaux. Les calculs étant indépendants dans chaque groupe de fréquence, ces pré-calculs sont très largement parallélisables. Puisque χ est très régulière, on effectuera une simple interpolation linéaire entre les données préévaluées.

Dans un premier paragraphe, on s'intéresse au calcul du facteur d'Eddington, puis dans le second paragraphe, l'attention est portée sur le calcul des moyennes d'opacités.

3.3.2 Pré-calculs du facteur d'Eddington χ_q

Lorsque l'énergie radiative et la pression radiative dans le groupe de fréquence q sont connues, le facteur d'Eddington est donné en 1D par :

$$\chi_q = \frac{P_q}{E_q}.$$

Pour déterminer le facteur d'Eddington χ_q dans le groupe q, on suit la procédure suivante :

1. On fixe un couple physiquement admissible Température radiative-facteur d'anisotropie (ϑ_q, f_q) dans le groupe q.
2. On en déduit le couple énergie-flux $(\mathcal{E}_q, \mathcal{F}_q)$ grâce aux définitions $\mathcal{E}_q = a\vartheta^4$ et $\mathcal{F}_q = f_q\mathcal{E}_q$.
3. On cherche les multiplicateurs de Lagrange α_q et β_q tels que

$$< \mathcal{I} >_q = \mathcal{E}_q \quad \text{et} \quad < c\mu\, \mathcal{I} >_q = \mathcal{F}_q. \tag{3.61}$$

4. On calcule la pression radiative $P_q =< \mu^2\mathcal{I} >_q$.
5. On calcule le facteur d'Eddington : $\chi_q = \frac{P_q}{\mathcal{E}_q}$.

Les trois premiers moments de $\mathcal{I}$ intégrés en fréquence sur le groupe q sont notés E_q, F_q et P_q. L'énergie radiative sur le groupe q associée à $\mathcal{I}$ s'écrit alors :

$$E_q = \frac{2\pi}{c}\int_{-1}^{1}\int_{\nu_q}^{\nu_{q+1}} \frac{2h\nu^3}{c^2}\left[\exp\left(\frac{h\nu}{kT}\alpha_q(1-\beta_q\mu)\right)-1\right]^{-1} d\nu d\mu. \tag{3.62}$$

Cette expression de E_q n'est pas utilisable telle quelle pour inverser le système (3.61). En effet, contrairement au cas gris où l'on peut calculer analytiquement cette intégrale, l'intégration sur le groupe de fréquence q rend ce calcul plus beaucoup plus complexe. Pour approcher cette intégrale, l'idée est de réécrire E_q à l'aide de nouvelles fonctions. Dans un premier temps on introduit :

$$z = \frac{h\nu}{kT}\alpha_q(1-\beta_q\mu), \tag{3.63}$$

pour réécrire l'énergie de la façon suivante :

$$E_q(\alpha_q,\beta_q) = \frac{15aT^4}{2\pi^4\alpha_q^4}\int_{-1}^{1}\frac{1}{(1-\beta_q\mu)^4}\int_{z_q}^{z_{q+1}}\frac{z^3}{e^z-1}dzd\mu,$$

avec $z_q = \frac{h\nu_q}{kT}\alpha_q(1-\beta_q\mu)$ et $z_{q+1} = \frac{h\nu_{q+1}}{kT}\alpha_q(1-\beta_q\mu)$.
On introduit alors la fonction Ξ :

$$\Xi(z) = \int_1^z \frac{x^3}{e^x-1}dx, \tag{3.64}$$

représentée sur la figure 3.7, où z dépend de ν et de μ. L'énergie E_q se décompose alors sous la forme :

$$E_q(\alpha_q,\beta_q) = \frac{15aT^4}{2\pi^4\alpha_q^4}\int_{-1}^{1}\frac{1}{(1-\beta_q\mu)^4}\left(\Xi(z_{q+1})-\Xi(z_q)\right)d\mu.$$

Puis, la fonction S définie par :

$$S(\alpha_q,\beta_q,\nu) = \int_{-1}^{1}\frac{1}{(1-\beta_q\mu)^4}\Xi(z(\mu,\nu))d\mu,$$

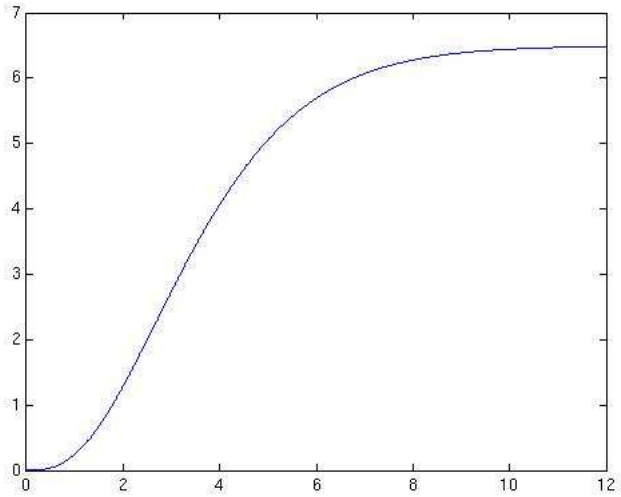


FIGURE 3.7 – Fonction Ξ

permet de réécrire E_q comme :

$$E_q(\alpha_q, \beta_q) = \frac{15aT^4}{2\pi^4\alpha_q^4}\left[S(\alpha_q, \beta_q, \nu_{q+1}) - S(\alpha_q, \beta_q, \nu_q)\right]. \tag{3.65}$$

En reprenant la définition (3.63) de $z(\mu, \nu)$, on obtient :

$$\mu = \frac{1}{\beta_q}\left(1 - \frac{zkT}{h\nu\alpha_q}\right) \qquad et \qquad \frac{1}{(1-\beta_q\mu)^4} = \left(\frac{h\nu\alpha_q}{zkT}\right)^4,$$

ce qui nous permet d'écrire :

$$S(\alpha_q, \beta_q, \nu) = \frac{h^3\nu^3\alpha_q^3}{\beta_q k^3 T^3}\int_{\mu^-}^{\mu^+} \frac{\Xi(z)}{z^4}dz,$$

avec $\mu^- = \frac{h\nu\alpha_q}{kT}(1-\beta_q)$ et $\mu^+ = \frac{h\nu\alpha_q}{kT}(1+\beta_q)$.
Finalement, S s'écrit sous la forme :

$$S(\alpha_q, \beta_q, \nu) = \frac{\tau(\mu^+)}{\beta_q(1+\beta_q)^3} - \frac{\tau(\mu^-)}{\beta_q(1-\beta_q)^3},$$

avec :

$$\tau(\eta) = \eta^3\int_1^{\eta}\frac{\Xi(z)}{z^4}dz. \tag{3.66}$$

L'expression de E_q ainsi développée fait intervenir Ξ définie par (3.64) qui n'est pas explicite. Cependant, on sait qu'elle est définie et infiniment dérivable sur $\mathbb{R}_*^+$ (et 3 fois dérivable en 0). Sa valeur à l'origine ainsi que son comportement à l'infini sont connus. Des méthodes d'intégration numérique pourraient être considérées pour approcher cette fonction. Cependant, étant donné que cette fonction apparaît un nombre considérable de fois et que ses variations sont assez régulières, il est préférable de l'approcher par une fonction de la forme :

$$\Xi(z) \simeq \xi(z) = d_\infty - \exp(z)\sum_{i=0}^{imax} d_i z^i.$$

Les égalités :

$$\begin{aligned}\Xi(0) &= 0,\\ \lim_{z\to+\infty}\Xi(z) &= \frac{\pi^4}{15},\\ \Xi'(0) &= 0,\\ \Xi''(0) &= 0,\\ \Xi^{(3)}(0) &= 2,\end{aligned}$$

permettent de fixer les coefficients d_∞, d_0, d_1, d_2 et d_3.

Les coefficients restants sont calculés au sens des moindre carrés (voir Appendice 2). Pour des raisons de coût de calcul, mais aussi de précision, on choisit de se limiter à $imax = 6$. La fonction ξ s'écrit alors :

$$\xi(z) = \frac{\pi^4}{15} - e^{-z}\left[\frac{\pi^4}{15} + \frac{\pi^4}{15}z + \frac{\pi^4}{30}z^2 - \left(\frac{30-\pi^4}{90}\right)z^3 + d_4z^4 + d_5z^5 + d_6z^6\right].$$

Avec cette approximation de Ξ, on déduit une approximation de τ :

$$\tau(\eta) \simeq \Lambda(\eta) = \eta^3\int_1^\eta \frac{\xi(z)}{z^4}dz.$$

Étant donnée la forme de ξ, la fonction Λ est immédiatement donnée par :

$$\begin{aligned}\Lambda(\eta) =& \frac{\pi^4}{45}(\eta^3-1) + \frac{\eta^3}{3}\left(E_1(1) - E_1(\eta)\right) + e^{-\eta}\frac{\pi^4}{45}\left(1+\eta+\frac{\eta^2}{2}\right) - \frac{\pi^4}{18}e^{-1}\eta^3 + d_4\eta^3(e^{-\eta}-e^{-1})\\ &+ d_5\eta^3\left(e^{-\eta}+\eta e^{-\eta}-2e^{-1}\right) + d_6\eta^3\left(2e^{-\eta}+2\eta e^{-\eta}+\eta^2e^{-\eta}-5e^{-1}\right),\end{aligned}$$

où la fonction E_1 est la première exponentielle intégrale définie par :

$$E_1(x) = \int_1^\infty \frac{e^{-xt}}{t}dt.$$

Numériquement, on approche cette fonction par les approximations proposées par E. E. Allen, note 169, MTAC 8, 240 (1954) et *Approximations for digital computers* de C. Hastings Jr, Princeton Univ Press, 1955. (voir Appendice 2)

Une fois toutes les fonctions définies pour le calcul de E_q, on peut de manière similaire, décomposer F_q et P_q :

$$\frac{1}{c}F_q(\alpha,\beta) = \frac{15aT^4}{2\pi^4\alpha^4}\left[S^b(\alpha,\beta,\nu_2) - S^b(\alpha,\beta,\nu_1)\right], \tag{3.67}$$

$$P_q(\alpha,\beta) = \frac{15aT^4}{2\pi^4\alpha^4}\left[S^t(\alpha,\beta,\nu_2) - S^t(\alpha,\beta,\nu_1)\right], \tag{3.68}$$

où les fonctions S^b et S^t sont données par :

$$S^b(\alpha,\beta,\nu) = \frac{\tau(\mu^+)}{\beta^2(1+\beta)^3} - \frac{\tau(\mu^-)}{\beta^2(1-\beta)^3} - \left(\frac{\tau^b(\mu^+)}{\beta^2(1+\beta)^2} - \frac{\tau^b(\mu^-)}{\beta^2(1-\beta)^2}\right),$$

$$S^t(\alpha,\beta,\nu) = \frac{\tau(\mu^+)}{\beta^3(1+\beta)^3} - \frac{\tau(\mu^-)}{\beta^3(1-\beta)^3} - 2\left(\frac{\tau^b(\mu^+)}{\beta^3(1+\beta)^2} - \frac{\tau^b(\mu^-)}{\beta^3(1-\beta)^2}\right) + \frac{\tau^t(\mu^+)}{\beta^3(1+\beta)} - \frac{\tau^t(\mu^-)}{\beta^3(1-\beta)}.$$

Les fonctions τ^b et τ^t sont données par :

$$\begin{aligned}\tau^b(\eta) &= \eta^2\int_1^\eta \frac{\Xi(z)}{z^3}dz,\\ \tau^t(\eta) &= \eta\int_1^\eta \frac{\Xi(z)}{z^2}dz.\end{aligned}$$

En utilisant la même approximation ξ de Ξ, on en déduit les approximations de τ^b et τ^t :

$$\begin{aligned}
\tau^b(\eta) \simeq \Lambda^b(\eta) =& \frac{\pi^4}{30}(\eta^2-1) - \frac{7\pi^4}{90}\eta^2 e^{-1} + \frac{\eta^2}{3}(e^{-1}-e^{-\eta}) \\
&+ \frac{\pi^4}{30}e^{-\eta}\left(1+\eta+\frac{\eta^2}{3}\right) + d_4\eta^2(e^{-\eta}+\eta e^{-\eta}-2e^{-1}) \\
&+ d_5\eta^2(2e^{-\eta}+2\eta e^{-\eta}+\eta^2 e^{-\eta}-5e^{-1}) \\
&+ d_6\eta^2(6e^{-\eta}+6\eta e^{-\eta}+3\eta^2 e^{-\eta}+\eta^3 e^{-\eta}-16e^{-1}), \\
\tau^t(\eta) \simeq \Lambda^t(\eta) =& \frac{\pi^4}{15}(\eta-1) - \frac{11\pi^4}{90}\eta e^{-1} + \frac{2\eta}{3}(e^{-1}-e^{-\eta}) \\
&+ \frac{\pi^4}{90}e^{-\eta}(6+4\eta+\eta^2) + \frac{e^{-\eta}}{3}(\eta-\eta^2) + d_4\eta(2e^{-\eta}+2\eta e^{-\eta}+\eta^2 e^{-\eta}-5e^{-1}) \\
&+ d_5\eta(6e^{-\eta}+6\eta e^{-\eta}+3\eta^2 e^{-\eta}+\eta^3 e^{-\eta}-16e^{-1}) \\
&+ d_6\eta(24e^{-\eta}+12\eta^2 e^{-\eta}+24\eta e^{-\eta}+4\eta^3 e^{-\eta}+\eta^4 e^{-\eta}-65e^{-1}).
\end{aligned}$$

On a ainsi donné une approximation de $E_q =< \mathcal{I} >_q$, $F_q =< c\Omega\mathcal{I} >_q$ et $P_q =< \Omega\otimes\Omega\mathcal{I} >_q$ en fonction des multiplicateurs α_q et β_q. Il reste à présent à inverser le système (3.61) afin de déterminer ces coefficients α_q, β_q en fonction du couple d'énergie $(\mathcal{E}_q, \mathcal{F}_q)$. Les fonctions étant très raides, la méthode classique de Newton pour l'inversion de fonctions non-linéaires ne convient pas. On choisit d'utiliser une méthode de gradient à pas constant. Initialement, la méthode est prévue pour déterminer le minimum (ou maximum) d'une fonction g de $\mathbb{R}^n$ dans $\mathbb{R}$ par la recherche du zéro de son gradient ∇g. Or, cette méthode ne nécessite pas de connaître la fonction pour rechercher la racine de son gradient. On suppose ainsi que le système (3.61) est le gradient d'une certaine fonction g. Ainsi si l'on note G son gradient, il est défini de la manière suivante :

$$G = \begin{pmatrix} Eq(\alpha,\beta) - \mathcal{E}_q \\ Fq(\alpha,\beta) - \mathcal{F}_q \end{pmatrix},$$

pour approcher la solution de (3.61). À chaque étape, la méthode du gradient consiste à déplacer la solution courante dans la direction opposée du gradient, afin de la faire décroître, soit :

$$V_{n+1} = V_n - pG(V_n),$$

où $V_n = \begin{pmatrix}\alpha_n \\ \beta_n\end{pmatrix}$ est le vecteur des inconnues à l'itération n et p est le pas de la méthode.
Grâce aux approximations explicites de E_q et F_q obtenues précédemment, cet algorithme permet de calculer numériquement les multiplicateurs de Lagrange (α_q, β_q) pour un couple énergie-flux $(\mathcal{E}_q, \mathcal{F}_q)$ et ainsi le facteur d'anisotropie χ_q correspondant.

Cependant, cette méthode ne permet pas de les calculer dans tous les cas, car certaines configurations particulièrement raides posent des difficultés numériques. En particulier, cette méthode ne permet pas de calculer (α_q, β_q) lorsque l'on impose une énergie très grande dans un groupe de fréquence étroit, ou au contraire, lorsque l'on met très peu d'énergie dans un groupe de fréquence large. Un cas extrême par exemple où le facteur d'anisotropie de l'ordre de 0.95 et l'énergie de l'ordre de $a1000^4$ dans le groupe de fréquence $[0; 10^{12}]$ correspond à un coefficient β_q de l'ordre de 0.999999999. En effet, pour une température de 1000 K, ce groupe de fréquence comporte moins de 0.01 % de l'énergie totale de la Planckienne. Même si en pratique de tels cas ne se produisent que très rarement, on utilise une autre méthode pour calculer le facteur d'Eddington χ. On définit tout d'abord un critère pour décrire la *raideur* du cas. On considère deux critères théoriques de *raideur* d'un cas :

$$D^- = (1-\beta_q)z \simeq 0 \qquad \text{ou} \qquad D^+ = (1+\beta_q)z \simeq 0,$$

où la variable z est donnée par (3.63). Ces deux critères correspondent exactement aux cas où $|\beta_q|$ devient très proche de 1 et où la fréquence ν est très petite devant la température T, c'est-à-dire que l'on est loin du pic de la Planckienne donnée par (1.1). On introduit la variable :

$$\zeta = \frac{h\nu}{kT},$$

qui servira à définir numériquement si un cas est raide ou non et on note $\zeta_q = \zeta(\nu_q)$.
On teste le code sur un grand nombre de cas pour déterminer un seuil ζ_{min} pour lequel le cas est considéré comme raide afin de choisir :

$$\zeta_{min} = \min\left(0.5; \exp(p_1 + p_2 f)\right),$$

avec f le facteur d'anisotropie et

$$\begin{aligned} p_1 &= 6.822178996548980, \\ p_2 &= -5.282761267696435. \end{aligned}$$

Pour tous les groupes $q = [\nu_q, \nu_{q+1}]$ tels que $\zeta_{q+1} < \zeta_{min}$, on développe une nouvelle procédure pour approcher le facteur d'Eddington χ . Dans ces zones dites raides, on peut remplacer τ, τ^b et τ^t par leurs développements asymptotiques respectifs en 0 :

$$\begin{aligned} \tau(z) &\underset{z=0}{\simeq} = (c_1 + c_2 \ln(z))\, z^3 + o(z^3), \\ \tau^b(z) &\underset{z=0}{\simeq} = c_1^b z^2 + c_2 z^3 + o(z^3), \\ \tau^t(z) &\underset{z=0}{\simeq} = c_1^t z + \frac{c_2}{2} z^3 + o(z^3), \end{aligned}$$

où c_1, c_2, c_1^b et c_1^t sont des constantes.
En effet, dans la définition des fonctions S, S^b et S^t, les fonctions τ, τ^b et τ^t sont évaluées en D^+ et D^-.
En posant $D = \frac{h\nu\alpha}{kT}$, on obtient alors les approximations suivantes pour S, S^b et S^t :

$$\begin{aligned} S(\alpha_q, \beta_q, \nu) &\underset{\|\beta\|=1}{\simeq} = \frac{\zeta^3}{\beta_q}\left[c_2 \ln\left(\frac{1+\beta_q}{1-\beta_q}\right)\right], \\ S^b(\alpha_q, \beta_q, \nu) &\underset{\|\beta\|=1}{\simeq} = \frac{2c_2\zeta^3}{\beta_q}, \\ S^t(\alpha_q, \beta_q, \nu) &\underset{\|\beta\|=1}{\simeq} = \frac{2c_2\zeta^3}{\beta_q^2}. \end{aligned}$$

À partir de ces approximations, on en déduit une expression pour le facteur d'anisotropie et pour χ dans ces zones :

$$f(\alpha_q, \beta_q) = \frac{F_q}{cE_q} = \frac{1}{\beta_q} - \frac{2}{\ln\left(\frac{1+\beta_q}{1-\beta_q}\right)}, \tag{3.69}$$

$$\chi = \frac{f}{\beta_q}. \tag{3.70}$$

Ces développements asymptotiques donnent une approximation du facteur d'anisotropie f sans dépendance en α_q. Puisque le facteur d'anisotropie f est fixé, il suffit d'inverser (3.69) pour obtenir le multiplicateur β_q correspondant. De plus, on cherche β_q dans l'intervalle $]0, 1[$, il suffit alors d'utiliser une méthode de dichotomie pour inverser (3.69).

En pratique, on utilise cette approximation pour approcher la valeur du facteur d'Eddington χ_{lim} quand $\zeta = \zeta_q$. Après avoir observé le comportement de χ en fonction du groupe de fréquence et du couple énergie-flux $(\mathcal{E}_q, \mathcal{F}_q)$ dans les zones où $\zeta < \zeta_{min}$, on décide d'approcher χ par un polynôme de degré 4 :

$$\chi(\zeta) = C_0 + C_1\zeta + C_2\zeta^2 + C_3\zeta^3 + C_4\zeta^4.$$

Cette approximation sera valide pour $\zeta \in [\zeta_q, \zeta_{min}]$.
Pour déterminer les coefficients $C_0, \dots, C_4$, on se base sur les propriétés du facteur d'Eddington et le développement asymptotique de celui-ci. On impose les contraintes :

$$\begin{aligned} \chi(\zeta_q) &= \chi_{lim}, \\ \chi'(\zeta_q) &= 0, \\ \chi''(\zeta_q) &= 0. \end{aligned} \tag{3.71}$$

De plus, pour fermer le système, on impose :

$$\begin{aligned} \chi(\zeta_{min}) &= \chi_{min}, \\ \chi'(\zeta_{min}) &= p_{min}, \end{aligned} \tag{3.72}$$

où χ_{min} est la valeur de χ calculée par la méthode du gradient sur le groupe $[z_q, z_{min}]$ et p_{min} est une valeur approchée de la pente de χ au point z_{min}. On obtient alors pour tout groupe de la forme $[z_q, z]$ avec $z < z_{min}$ une approximation de χ grâce à ce polynôme.

En résumé, la procédure de pré-calculs du facteur d'Eddington est la suivante :

- On se fixe un groupe de fréquence $[\nu_q, \nu_{q+1}]$.
- On se fixe un couple physiquement admissible Température-facteur d'anisotropie (θ_q, f_q) pour en déduire le couple énergie-flux $(\mathcal{E}_q, \mathcal{F}_q)$.
- On compare ζ_{q+1} à ζ_{min}. Deux possibilités :
  1. **Si** $\zeta_{q+1} \geq \zeta_{min}$ **:**
     On inverse le système (3.61) par la méthode du gradient en utilisant les approximations de E_q, F_q et P_q définies par (3.65),(3.67) et (3.68).
  2. **Si** $\zeta_{q+1} \leq \zeta_{min}$ **:** On calcule χ_{lim} donné par l'approximation (3.70) en $\zeta = \zeta_q$. Puis, on approche χ par un polynôme de degré 5 (3.3.2) à l'aide des contraintes (3.71) et (3.72).

Pour illustrer l'intérêt de ces pré-calculs, on compare le facteur d'Eddington gris et multigroupe. On se donne les quatre groupes de fréquence suivants :

$$\begin{aligned} [\nu_1; \nu_2] &= [0; 3.75.10^{14}], \\ [\nu_2; \nu_3] &= [3.75.10^{14}; 7.5.10^{14}], \\ [\nu_3; \nu_4] &= [7.5.10^{14}; 1.10^{16}], \\ [\nu_4; \nu_5] &= [1.10^{16}; 6.10^{16}]. \end{aligned}$$

Puis on discrétise le facteur d'anisotropie à l'aide de 100 points. Dans le cas où l'énergie est une Planckienne, le maximum peut être situé grâce à la loi de Wien qui donne une relation entre la température T et la longueur d'onde λ (en μm) :

$$\lambda T = 2898.$$

Dans ce cas, la fréquence correspondant au maximum de la Planckienne pour $T = 3020K$ est $\nu_{max} \simeq 3.12.10^{14}$. La figure 3.8 représente le facteur d'Eddington pour chacun des groupes de fréquence ainsi que le facteur gris en fonction du facteur d'anisotropie à $T = 3020K$. On voit bien, même dans ce cas simple, qu'il y a des différences significatives entre le facteur d'Eddington multigroupe et le facteur gris et que l'approximation du facteur multigroupe par le facteur gris n'est pas toujours adaptée.

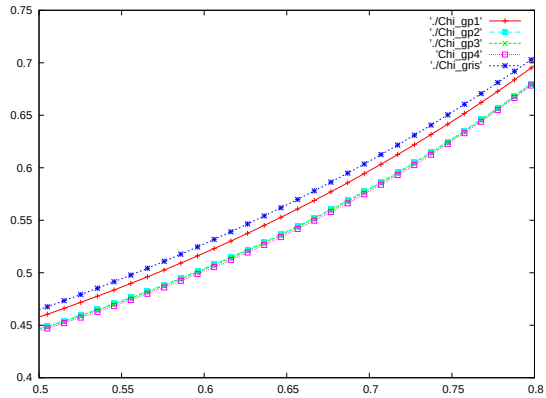


FIGURE 3.8 – Comparaison des facteurs d'Eddington gris et multigroupe en fonction du facteur d'anisotropie f

3.3.3 Pré-calcul des moyennes d'opacités

On a vu dans le paragraphe 1.4.1 que les moyennes d'opacités dans le modèle multigroupe sont données par :

$$\begin{aligned} \sigma_q^e &= \frac{< \sigma_\nu^a B(T) >_q}{a\theta_q^4}, \\ \sigma_q^a &= \frac{< \sigma_\nu^a I >_q}{E_q}, \\ \sigma_q^f F_q &= c < \sigma_\nu^a \Omega I >_q. \end{aligned} \tag{1.28}$$

De même que pour le facteur d'Eddington, le code nécessite le calcul de ces coefficients un grand nombre de fois. On choisit alors de les pré-calculer pour un certain nombre de couples Température-facteur d'anisotropie (ϑ_q, f_q) donnés.
Dans les coefficients σ_q^a et σ_q^f, on voit apparaître l'intensité radiative I. Cependant, cette variable cinétique n'est plus une inconnue du système. On rappelle qu'un des avantage du modèle M_1 est que l'on peut approcher l'intensité radiative I par la solution $\mathcal{I}$ du problème de minimisation (1.29) qui est positive [104]. Ainsi on aura :

$$\begin{aligned} \sigma_q^e &= \frac{< \sigma_\nu^a B(T) >_q}{a\theta_q^4}, \\ \sigma_q^a &\simeq \frac{< \sigma_\nu^a \mathcal{I} >_q}{E_q}, \\ \sigma_q^f F_q &\simeq= c < \sigma_\nu^a \Omega \mathcal{I} >_q. \end{aligned}$$

Pour calculer ces coefficients, on utilise la procédure précédente pour calculer les α_q et β_q pour le couple énergie-flux $(\mathcal{E}_q, \mathcal{F}_q)$ sont obtenus par la procédure décrite dans le paragraphe précédent dédié au calcul du facteur d'Eddington.
On développe alors les expressions de σ_q^a et σ_q^f :

$$\begin{aligned} < \sigma_\nu^a \mathcal{I} >_q &= \frac{2\pi}{c} \int_{\nu_q}^{\nu_{q+1}} \int_{-1}^{1} \sigma^a(\nu) \mathcal{I}(\mu, \nu, \alpha_q, \beta_q) d\mu d\nu = \frac{2\pi}{c} \int_{\nu_q}^{\nu_{q+1}} \sigma(\nu) J^0(\nu), \\ c < \sigma_\nu^a \Omega \mathcal{I} >_q &= 2\pi \int_{\nu_q}^{\nu_{q+1}} \int_{-1}^{1} \sigma^a(\nu) \mu \mathcal{I}(\mu, \nu, \alpha_q, \beta_q) d\mu d\nu = 2\pi \int_{\nu_q}^{\nu_{q+1}} \sigma(\nu) J^1(\nu), \end{aligned}$$

avec :

$$J^0(\nu) = \int_{-1}^{1} \mathcal{I}(\mu, \nu, \alpha_q, \beta_q) d\mu,$$
$$J^1(\nu) = \int_{-1}^{1} \mu\mathcal{I}(\mu, \nu, \alpha_q, \beta_q) d\mu.$$

Numériquement, on approche les intégrales J^0 et J^1 par des quadratures. On considère par exemple une méthode de point milieu. Puis pour calculer l'intégrale en fréquence, on divise chaque groupe de fréquence $[\nu_q; \nu_{q+1}]$ en n_q bandes étroites de même largeur $d\nu$:

$$[\nu_q; \nu_{q+1}] = \cup_{i=1}^{n_q} [\nu_q^i; \nu_q^{i+1}],$$

afin de supposer que l'opacité d'absorption σ^a est constante sur chaque bande étroite. Ainsi on peut écrire :

$$< \sigma_\nu^a B >_q \simeq \frac{4\pi d\nu}{c} \sum_{i=1}^{n_q} \sigma_{q,i}^a B(\frac{\nu_q^i + \nu_q^{i+1}}{2}),$$
$$< \sigma_\nu^a \mathcal{I} >_q \simeq \frac{2\pi d\nu}{c} \sum_{i=1}^{n_q} \sigma_{q,i}^a J^0(\frac{\nu_q^i + \nu_q^{i+1}}{2}),$$
$$c < \sigma_\nu^a \Omega\mathcal{I} >_q \simeq 2\pi d\nu \sum_{i=1}^{n_q} \sigma_{q,i}^a J^1(\frac{\nu_q^i + \nu_q^{i+1}}{2}).$$

On approche également l'intégrale de la Planckienne par :

$$< B >_q \simeq \frac{4\pi d\nu}{c} \sum_{i=1}^{n_q} B(\frac{\nu_q^i + \nu_q^{i+1}}{2}).$$

Finalement, les moyennes d'opacités sont approchées par :

$$\begin{aligned}
\sigma_q^e &\simeq \frac{\sum_{i=1}^{n_q} \sigma_{q,i}^a B(\frac{\nu_q^i + \nu_q^{i+1}}{2})}{\sum_{i=1}^{n_q} B(\frac{\nu_q^i + \nu_q^{i+1}}{2})}, \\
\sigma_q^a &\simeq \frac{\frac{2\pi d\nu}{c} \sum_{i=1}^{n_q} \sigma_{q,i}^a J^0(\frac{\nu_q^i + \nu_q^{i+1}}{2})}{\mathcal{E}_q}, \\
\sigma_q^f &\simeq \frac{2\pi d\nu \sum_{i=1}^{n_q} \sigma_{q,i}^a J^1(\frac{\nu_q^i + \nu_q^{i+1}}{2})}{\mathcal{F}_q}.
\end{aligned} \tag{3.73}$$

Dans le cas où le flux $\mathcal{F}_q$ est nul, on approche l'opacité σ_q^f par sa limite :

$$\lim_{f \to 0} \sigma_q^f = \frac{< \sigma_\nu \partial_T B >_q}{< \partial_T B >_q}.$$

Dans ce cas, les intégrales sont également approchées par :

$$< \sigma_\nu^a \partial_T B >_q \simeq \frac{4\pi d\nu}{c} \sum_{i=1}^{n_q} \sigma_{q,i}^a (\partial_T B)(\frac{\nu_q^i + \nu_q^{i+1}}{2}),$$
$$< \partial_T B >_q \simeq \frac{4\pi d\nu}{c} \sum_{i=1}^{n_q} (\partial_T B)(\frac{\nu_q^i + \nu_q^{i+1}}{2}).$$

Dans le paragraphe précédent, on a vu que dans certains cas jugés *raides*, le calcul des multiplicateurs de Lagrange par la méthode du gradient n'est plus possible numériquement. Dans ces cas particuliers où l'énergie imposée dans le groupe de fréquence est très grande devant la taille du groupe, le facteur d'Eddington est approché à l'aide d'un développement asymptotique. Ainsi, quand le facteur d'anisotropie f_q s'approche de 1 dans un petit groupe de fréquence, on choisit d'utiliser une approximation linéaire pour les moyennes d'opacité. On note tout d'abord f_{lim}, le plus grand facteur d'anisotropie pour lequel les multiplicateurs de Lagrange sont calculables par la méthode du gradient. On suppose que les opacités moyennes sont calculées par (3.73) pour $f_q \leq f_{lim}$. Puis, pour calculer les moyennes d'opacité en $f_{lim} + k\ df$, df étant le pas de discrétisation du facteur d'anisotropie, on considère les deux points d'interpolation $f_{lim} + (k-1)\ df$ et $f_{lim} + (k-2)\ df$ pour lesquels les opacités sont déjà calculées.

En résumé, la procédure de pré-calculs des moyennes d'opacités est la suivante :

- On se fixe un couple physiquement admissible Température-facteur d'anisotropie (θ_q, f_q) pour en déduire le couple énergie-flux $(\mathcal{E}_q, \mathcal{F}_q)$.
- On se fixe un groupe de fréquence $[\nu_q, \nu_{q+1}]$.
- On compare ζ_{q+1} à ζ_{min} :
  1. **Si** $\zeta_{q+1} \geq \zeta_{min}$ **:**
     Après avoir inversé le système (3.61) par la méthode du gradient, on subdivise le groupe de fréquence q en bandes étroites pour approcher les moyennes d'opacités par (3.73).
  2. **Si** $\zeta_{q+1} \leq \zeta_{min}$ **:** On suppose que les opacités moyennes pour $f_q = f_{lim}$ et $f_q = f_{lim} - df$ sont approchées par (3.73). Pour $f_q > f_{lim}$, les opacités sont alors approchées progressivement par interpolation linéaire basée sur les deux valeurs précédentes.

4

Équations et schémas numériques pour la radiothérapie

4.1 Généralités sur la radiothérapie et modèle utilisé

On désigne par radiothérapie le traitement des cancers et autres maladies utilisant un certain type de rayonnement. Le rayonnement ionisant dépose de l'énergie dans les cellules de la zone traitée afin de les détruire ou les affaiblir. Dans ce but, les médecins cherchent à déterminer les paramètres optimaux du rayon incident pour détruire la tumeur en minimisant les dommages sur les cellules saines avoisinantes. Il faut ainsi prévoir la dose de rayonnement optimale dans le corps du patient avant de commencer le traitement. Pour les simulations de radiothérapie, les données utilisées sont des coupes en dimension deux de scanners CT (computer tomography) qui décrivent la densité des tissus.

La radiothérapie est une application du transfert radiatif où l'on considère des paramètres physiques particuliers. On introduit tout d'abord l'équation stationnaire de transport de Boltzmann linéarisée :

$$\Omega.\nabla\Psi(x,\varepsilon,\Omega) = \int_0^\infty \int_{S^2} \rho(x)\sigma(\varepsilon,\varepsilon',\Omega'.\Omega)\Psi(x,\varepsilon',\Omega')d\Omega' d\varepsilon'$$
$$- \int_0^\infty \int_{S^2} \rho(x)\sigma(\varepsilon',\varepsilon,\Omega.\Omega')\Psi(x,\varepsilon,\Omega)d\Omega' d\varepsilon',$$

où ε est l'énergie de la particule, σ est le noyau de *scattering* et ρ est la densité de la matière. L'inconnue est caractérisée par $\Psi(x,\varepsilon,\Omega) = |v(\varepsilon)|f(x,\varepsilon,\Omega)$ où f est la densité d'électrons dans l'espace des phases et v la vitesse des particules.
Ici, les particules considérées sont les électrons. On voit que contrairement au transfert radiatif, le système ne dépend pas du temps mais de l'énergie de la particule. Les interactions considérées dans ce modèle sont le *scattering* élastique qui modifie la direction de l'électron sans perte d'énergie et le *scattering* inélastique dans lequel un électron vient exciter un autre électron lié qui de ce fait se sépare de l'atome auquel il appartient. Une des propriétés de la radiothérapie est la faible déviation angulaire des électrons. Plus précisément, le *scattering* devient très important lorsque l'angle de dispersion $cos(\mu) = \Omega.\Omega'$ s'approche de 1. En plus de cette propriété, le *scattering* inélastique engendre de faibles pertes d'énergie. Pour prendre en compte ce comportement et les

propriétés spécifiques de la radiothérapie, deux développements asymptotiques sont utilisés : le Continuous Slowing Down (CSD) et Fokker-Planck.
Pour tenir compte de la faible perte d'énergie, après une mise à l'échelle de l'énergie et un développement asymptotique, on obtient l'équation CSD :

$$\Omega.\nabla\Psi(x,\varepsilon,\Omega) = \rho(x)\int_{S^2}\bar{\sigma}(\varepsilon,\Omega'.\Omega)\Psi(x,\varepsilon,\Omega')d\Omega' - \rho(x)\int_{S^2}\bar{\sigma}(\varepsilon,\Omega.\Omega')\Psi(x,\varepsilon,\Omega)d\Omega' + \partial_\varepsilon\left(S_M(x,\varepsilon)\Psi(x,\varepsilon,\Omega)\right), \tag{4.1}$$

où :

$$\bar{\sigma}(\varepsilon,\Omega.\Omega') = \int_0^\infty \sigma(\varepsilon',\varepsilon,\Omega.\Omega')d\varepsilon', \tag{4.2}$$

est le coefficient de section efficace et

$$S_M(x,\varepsilon) = \rho(x)\int_{S^2}\int_0^\infty \varepsilon'\sigma(\varepsilon',\varepsilon,\Omega.\Omega')d\varepsilon'd\Omega' = \rho(x)S(\varepsilon), \tag{4.3}$$

est le pouvoir d'arrêt qui représente la perte d'énergie d'une particule par unité de libre parcours. Il décrit ainsi la vitesse à laquelle la particule se déplace et s'arrête dans la matière.

Si l'on prend en compte le faible angle de dispersion, après une mise à l'échelle du noyau de *scattering*, une analyse asymptotique conduit à l'équation de Fokker-Planck :

$$\Omega.\nabla\Psi(x,\varepsilon,\Omega) = T_{tot}(x,\varepsilon)\Delta_\Omega\Psi(x,\varepsilon,\Omega) + \partial_\varepsilon\left(S_M(x,\varepsilon)\Psi(x,\varepsilon,\Omega)\right), \tag{4.4}$$

où T_{tot} est le coefficient de transport, S_M est le pouvoir d'arrêt et Δ_Ω est l'opérateur de Laplace sur la sphère unité.
Avec ces quantités physiques, on peut définir la variable d'intérêt en radiothérapie appelée dose :

$$D(x) = \frac{1}{\rho(x)}\int_0^\infty S_M(x,\varepsilon)\int_{S^2}\Psi(x,\varepsilon,\Omega)d\Omega d\varepsilon, \tag{4.5}$$

qui représente la quantité d'énergie déposée dans la matière.
Jusqu'à présent, les calculs de doses cliniques reposent sur des modèles semi-empiriques. Ils sont basés sur des solutions explicites de rayonnement dans des géométries simplifiées (voir par exemple la théorie uni-dimensionnelle de Fermi-Eyges [52]). Ces solutions explicites sont combinées avec des données expérimentales pour calculer la dose sur l'axe central (voir par exemple [72]). Malgré de nombreuses améliorations de la théorie de Fermy-Eyges en incluant par exemple des termes de correction [76, 77, 4, 98], elle produit des erreurs allant jusqu'à 12 % dans les inhomogénéités [82].
Actuellement, des codes basés sur des méthodes statistiques de Monte Carlo commencent à être introduits dans le milieu hospitalier [99, 36, 100]. Ces méthodes consistent à suivre l'évolution d'une particule qui peut aléatoirement être transportée dans la matière et interagir avec celle-ci. Si le nombre de particules considérées est suffisamment grand, les quantités macroscopiques peuvent être approchées en moyennant les parcours simulés. Ainsi, les méthodes de Monte Carlo sont très précises et peuvent traiter des géométries totalement arbitraires sans perte de précision. Bien qu'elles soient les méthodes les plus précises pour le calcul de dose, leur coût de calcul reste prohibitif pour leur utilisation en milieu hospitalier.
Une approche différente est basée sur des méthodes déterministes pour l'équation de transport de Boltzmann linéarisée. En principe, leurs solutions sont comparables aux solutions Monte Carlo. Dans [21], Börgers affirme que sous certaines conditions, les méthodes déterministes pourraient rivaliser avec les simulations Monte Carlo. Récemment, des calculs de dose pour des cas-test cliniques en dimension trois ont été réalisés avec le code de calcul Attila [110]. Des résultats de précision comparable aux simulations Monte Carlo ont été obtenus avec des temps de calcul très

prometteurs. Une autre discrétisation de l'équation de transport de type éléments finis a également été proposée dans [20].

Dans ce travail, on s'intéresse aux modèles aux moments. Un résultat intéressant est que l'on obtient le même modèle aux moments pour les deux équations CSD (4.1) et Fokker-Planck (4.4) (voir annexe pour les calculs détaillés). On définit les trois premiers moments de Ψ par :

$$\begin{aligned}\Psi^0(x,\varepsilon) &= \int_{S^2} \Psi(x,\varepsilon,\Omega)d\Omega,\\ \mathbf{\Psi}^1(x,\varepsilon) &= \int_{S^2} \Omega\Psi(x,\varepsilon,\Omega)d\Omega,\\ \Psi^2(x,\varepsilon) &= \int_{S^2} \Omega\otimes\Omega\Psi(x,\varepsilon,\Omega)d\Omega.\end{aligned}$$

Comme Ψ^0 et $\mathbf{\Psi}^1$ sont les moments d'une fonction de distribution positive, ils doivent appartenir à l'espace des états admissibles suivant :

$$\mathcal{A}_d = \{{}^t(\Psi^0,\mathbf{\Psi}^1)\in\mathbb{R}^{d+1}, \Psi^0>0, \frac{\|\mathbf{\Psi}^1\|}{\Psi^0}\leq 1\}, \tag{4.6}$$

où $\|.\|$ est la norme Euclidienne usuelle.
Le modèle aux deux premiers moments est ainsi donné par :

$$\begin{cases}\nabla\mathbf{\Psi}^1 = \partial_\varepsilon\left(S_M\Psi^0\right),\\ \nabla\Psi^2 = 2 - T_{tot}\mathbf{\Psi}^1 + \partial_\varepsilon\left(S_M\mathbf{\Psi}^1\right).\end{cases} \tag{4.7}$$

Puis, de la même manière que pour le modèle M_1 du transfert radiatif, on choisit de fermer ce système (4.7) suivant le principe de minimisation de l'entropie. Ainsi, le modèle M_1 pour la radiothérapie s'écrit :

$$\begin{aligned}&\partial_\varepsilon(\rho S\Psi^0) - \nabla.\mathbf{\Psi}^1 = 0,\\ &\partial_\varepsilon(\rho S\mathbf{\Psi}^1) - \nabla.D_e(\frac{\mathbf{\Psi}^1}{\Psi^0})\Psi^0 = \rho T\mathbf{\Psi}^1,\end{aligned} \tag{4.8}$$

où D_e est le tenseur d'Eddington défini par (1.20) où l'on remplace F_R par $\mathbf{\Psi}^1$ et f par $\frac{\|\mathbf{\Psi}^1\|}{\Psi^0}$. Ce système (4.8) avec une densité ρ et un pouvoir d'arrêt S constants est hyperbolique pour tout $(\Psi^0,\mathbf{\Psi}^1)$ dans l'ensemble $\mathcal{A}_d$ [88]. Les fonctions positives $S_M(\mathbf{x},\varepsilon)$ et $T_{tot}(\mathbf{x},\varepsilon)$ sont déterminées par la nature des particules et de la matière. On suppose que l'on peut réécrire S_M et T_{tot} comme le produit de la densité de la matière $\rho(\mathbf{x})$ et d'une fonction dépendant de l'énergie :

$$S_M(\mathbf{x},\varepsilon) = \rho(\mathbf{x})S(\varepsilon), \quad \text{et} \quad T_{tot}(\mathbf{x},\varepsilon) = \rho(\mathbf{x})T(\varepsilon). \tag{4.9}$$

Dans toute cette étude, on suppose que les tissus sont composés d'eau avec une densité variable. Cette première approximation, relativement bonne, est la seule viable en l'absence de données supplémentaires sur le CT scanner. On précise que les définitions des fonctions $S(\varepsilon)$ et $\rho(\mathbf{x})$ seront données plus loin en fonction des simulations numériques considérées.
On suppose ici, en accord avec la physique, qu'il n'y a aucune particule quand l'énergie est arbitrairement grande :

$$\lim_{\varepsilon\to\infty}\Psi^0(\mathbf{x},\varepsilon) = 0, \qquad \lim_{\varepsilon\to\infty}\mathbf{\Psi}^1(\mathbf{x},\varepsilon) = 0. \tag{4.10}$$

En pratique, cette propriété est approchée de la façon suivante :

$$\Psi^0(\mathbf{x},\varepsilon_{max}) = \delta > 0, \qquad \mathbf{\Psi}^1(\mathbf{x},\varepsilon_{max}) = 0, \tag{4.11}$$

où ε_{max} est un choix pertinent de l'énergie maximale du système et δ désigne une énergie très petite.

Le système (4.8) est alors complété par une condition de type Cauchy donnée. Si l'on interprète l'énergie ε comme un temps mathématique, la condition initiale est donc donnée pour $\varepsilon = \varepsilon_{max}$. Avec toutes ces définitions, le problème de Cauchy pour le modèle (4.8) défini pour $\mathbf{x} \in \mathbb{R}^2$ et $\varepsilon \in [0; \varepsilon_{max}]$ est donné par :

$$\begin{aligned} &\partial_\varepsilon(\rho S \Psi^0) - \nabla.\mathbf{\Psi}^1 = 0, \\ &\partial_\varepsilon(\rho S \mathbf{\Psi}^1) - \nabla.D_e(\frac{\mathbf{\Psi}^1}{\Psi^0})\Psi^0 = \rho T \mathbf{\Psi}^1, \end{aligned} \tag{4.12}$$

avec la condition :

$$\Psi^0(\mathbf{x}, \varepsilon_{max}) = \delta, \qquad \mathbf{\Psi^1}(\mathbf{x}, \varepsilon_{max}) = 0.$$

Il est prouvé dans [41] que ce système hyperbolique a la bonne limite de transport et dans [66], qu'il reproduit bien la limite de l'équation cinétique. On attend également de ce modèle qu'il préserve la positivité de Ψ^0.
De nombreuses méthodes numériques ont été développées dans la littérature pour les systèmes hyperboliques de lois de conservation et en particulier quelques unes sont détaillées pour différents modèles dans [62, 64, 24, 25, 26, 14, 15, 19, 66]. Ces schémas peuvent être appliqués au modèle M_1 (4.12) pour des cas où la densité présente de petites variations. Dans cette étude, on s'intéresse à une méthode de type volumes finis capable de gérer de larges variations de la densité. En effet, dans le système (4.8), le flux dépend explicitement de la densité $\rho(\mathbf{x})$, ce qui pose de grandes difficultés numériques. De plus, cette densité peut varier entre plusieurs ordres de grandeurs, allant par exemple de $\rho \sim 1$ (eau) à $\rho \sim 10^{-3}$ (air). Une discrétisation naïve de ce système consiste à séparer la dérivée en énergie de la manière suivante :

$$\partial_\varepsilon(\rho S \Psi^0) = \rho S \partial_\varepsilon \Psi^0 + \Psi^0 \partial_\varepsilon(\rho S),$$

pour réécrire le système (4.8) en dimension un sous la forme :

$$\begin{aligned} \partial_\varepsilon \Psi^0 &= \frac{1}{\rho S}\left(\partial_x \Psi^1 - \partial_\varepsilon(\rho S)\right), \\ \partial_\varepsilon \mathbf{\Psi}^1 &= \frac{1}{\rho S}\left(\partial_x(\Psi^0 \chi(\mathbf{\Psi}^1/\Psi^0)) + \mathbf{\Psi}^1(\rho T - \partial_\varepsilon(\rho S))\right). \end{aligned} \tag{4.13}$$

La présence de la densité $\rho(x)$ devant la dérivée du flux peut engendrer des difficultés numériques quand ρ est discontinue. Si l'on considère une discrétisation classique, les pas de temps peuvent devenir très petits et les calculs très longs. Pour surmonter cette difficulté, on développe une technique spécifique capable de gérer ce flux qui dépend fortement de l'espace. Plusieurs auteurs ont travaillé sur des méthodes numériques pour des lois de conservation présentant des flux discontinus. L'article [111] par exemple, passe en revue un grand nombre de ces méthodes. L'idée ici est nouvelle dans le sens où l'on profite de la structure spécifique des coefficients discontinus pour faire des transformations sur les variables initiales.

Dans un premier paragraphe, on considère le modèle en 1D et on introduit une transformation capable de gérer le pouvoir d'arrêt. En effet, on remarque que la densité $\rho(\mathbf{x})$ et la fonction $S(\varepsilon)$ déforment l'espace des phases $(\mathbf{x}, \varepsilon)$, ce qui rend l'approximation numérique difficile. On suggère alors un changement de variables judicieux afin de redresser l'espace des phases. Le schéma numérique obtenu est ainsi très précis et peu coûteux, mais n'admet pas d'extension directe en 2D. Dans le paragraphe 4.2.2, on reformule alors cette méthode unidimensionnelle pour qu'elle admette une extension directe en 2D. Des résultats de convergence et quelques comparaisons illustrent dans le paragraphe 5.1 les performances de ce schéma et justifient son extension en 2D. Le paragraphe 4.2.3 est consacré à la présentation de la procédure numérique en 2D. Finalement quelques résultats de convergence et des cas-tests physiques y seront aussi présentés ainsi que dans le paragraphe 5.2. On précise que ce travail a fait l'objet d'une publication [16].

4.2 Schémas numériques

4.2.1 Méthode avec changement de variables en 1D

On considère le système suivant :

$$\begin{aligned} &\partial_\varepsilon(\rho S\Psi^0) - \partial_x\Psi^1 = 0, \\ &\partial_\varepsilon(\rho S\Psi^1) - \partial_x\left(\Psi^0\chi(\frac{\Psi^1}{\Psi^0})\right) = \rho T\Psi^1, \end{aligned} \tag{4.14}$$

où la solution (Ψ^0, Ψ^1) est dans l'espace des états admissibles $\mathcal{A}_1$ défini par (4.6).

Dans un premier temps, on s'intéresse au système homogène associé à (4.14) que l'on écrit sous forme condensée par :

$$\partial_\varepsilon(\rho SU) - \partial_x F(U) = 0, \tag{4.15}$$

où $U = (\Psi^0, \Psi^1)^T$ est le vecteur des inconnues dans $\mathcal{A}_1$ et le flux F est défini par :

$$F(U) = (\Psi^1, \Psi^0\chi\left(\Psi^1/\Psi^0\right))^T. \tag{4.16}$$

L'idée est de faire un changement de variables dans (4.15) qui peut conduire à un système n'impliquant pas $\rho(x)$ et $S(\varepsilon)$ dans l'opérateur de dérivée en énergie. Ensuite, le système obtenu est discrétisé par un schéma HLL [70] (voir aussi [22, 103]).
Dans un second temps, on applique l'idée du schéma HLL *asymptotic preserving* développé dans le chapitre 3 pour le modèle M_1 du transfert radiatif afin de prendre en compte le terme source.

Schéma numérique pour le système homogène

Pour se concentrer sur le rôle des fonctions $\rho(x)$ et $S(\varepsilon)$, on considère le système homogène (4.15). Puisque la densité $\rho(x)$ peut présenter de larges variations voire même être discontinue et ainsi engendrer des difficultés numériques, l'idée est d'introduire un changement de variables afin de l'éliminer de la dérivée en énergie.
Pour cela, on commence par s'affranchir du pouvoir d'arrêt $S(\varepsilon)$ dans la dérivée en énergie, on pose :

$$\hat{U}(x,\varepsilon) = S(\varepsilon)U(x,\varepsilon), \tag{4.17}$$

pour réécrire le système (4.15) sous la forme :

$$\begin{aligned} &S\partial_\varepsilon(\rho\hat{\Psi}^0) - \partial_x\hat{\Psi}^1 = 0, \\ &S\partial_\varepsilon(\rho\hat{\Psi}^1) - \partial_x\big(\hat{\Psi}^0\chi\left(\frac{\hat{\Psi}^1}{\hat{\Psi}^0}\right)\big) = 0. \end{aligned} \tag{4.18}$$

Puis, comme la densité est strictement positive et ne dépend pas de l'énergie, on peut diviser le système (4.18) par ρ et ainsi écrire :

$$\begin{aligned} &S\partial_\varepsilon\hat{\Psi}^0 - \frac{1}{\rho}\partial_x\hat{\Psi}^1 = 0, \\ &S\partial_\varepsilon\hat{\Psi}^1 - \frac{1}{\rho}\partial_x\big(\hat{\Psi}^0\chi\left(\frac{\hat{\Psi}^1}{\hat{\Psi}^0}\right)\big) = 0. \end{aligned}$$

On définit à présent la fonction $\tilde{\varepsilon} : \mathbb{R}^+ \to \mathbb{R}^+$ par :

$$\tilde{\varepsilon}(\varepsilon) = \int_0^\varepsilon \frac{1}{S(t)}dt,$$

qui est bien définie car S est strictement positive. Comme le pouvoir d'arrêt est une fonction positive, l'application $\tilde{\varepsilon} : \varepsilon \mapsto \tilde{\varepsilon}(\varepsilon)$ est clairement croissante et peut donc être utilisée comme changement de variable pour définir :

$$\bar{U}(x, \tilde{\varepsilon}(\varepsilon)) = \hat{U}(x, \varepsilon). \tag{4.19}$$

La dérivée en énergie s'écrit alors :

$$S(\varepsilon)\partial_{\varepsilon}\hat{U}(x, \varepsilon) = \partial_{\tilde{\varepsilon}}\bar{U}(x, \tilde{\varepsilon}).$$

En conséquence, le système (4.2.1) se réécrit de façon équivalente :

$$\begin{aligned} &\partial_{\tilde{\varepsilon}}\bar{\Psi}^0(x, \tilde{\varepsilon}) - \frac{1}{\rho(x)}\partial_x\bar{\Psi}^1(x, \tilde{\varepsilon}) = 0, \\ &\partial_{\tilde{\varepsilon}}\bar{\Psi}^1(x, \tilde{\varepsilon}) - \frac{1}{\rho(x)}\partial_x\Big(\bar{\Psi}^0(x, \tilde{\varepsilon})\chi\left(\frac{\bar{\Psi}^1(x, \tilde{\varepsilon})}{\bar{\Psi}^0(x, \tilde{\varepsilon})}\right)\Big) = 0. \end{aligned} \tag{4.20}$$

Puis, on utilise la même approche pour la densité $\rho(x)$. On pose tout d'abord :

$$\tilde{x}(x) = \int_0^x \rho(t)dt. \tag{4.21}$$

Comme la densité est une fonction strictement positive, la fonction définie par $\tilde{x} : \mathbb{R}^+ \to \mathbb{R}^+$ est croissante et permet donc de définir un changement de variables de la façon suivante :

$$\tilde{U}(\tilde{x}, \tilde{\varepsilon}) = \bar{U}(x, \tilde{\varepsilon}). \tag{4.22}$$

La dérivée en espace s'écrit alors :

$$\frac{1}{\rho(x)}\partial_x\bar{U}(x, \tilde{\varepsilon}) = \partial_{\tilde{x}}\tilde{U}(\tilde{x}, \tilde{\varepsilon}), \tag{4.23}$$

et le système (4.20) devient :

$$\begin{aligned} &\partial_{\tilde{\varepsilon}}\tilde{\Psi}^0(\tilde{x}, \tilde{\varepsilon}) - \partial_{\tilde{x}}\tilde{\Psi}^1(\tilde{x}, \tilde{\varepsilon}) = 0, \\ &\partial_{\tilde{\varepsilon}}\tilde{\Psi}^1(\tilde{x}, \tilde{\varepsilon}) - \partial_{\tilde{x}}\left(\tilde{\Psi}^0(\tilde{x}, \tilde{\varepsilon})\chi\left(\frac{\tilde{\Psi}^1(\tilde{x}, \tilde{\varepsilon})}{\tilde{\Psi}^0(\tilde{x}, \tilde{\varepsilon})}\right)\right) = 0. \end{aligned} \tag{4.24}$$

Les valeurs propres de ce système sont comprises entre -1 et 2.
Pour alléger les notations, le système de lois de conservation (4.24) est réécrit sous forme condensée :

$$\partial_{\tilde{\varepsilon}}\tilde{U} - \partial_{\tilde{x}}F(\tilde{U}) = 0, \tag{4.25}$$

où le flux F est défini par (4.16).
Pour résumer les changements de variables (4.17), (4.19) et (4.22), on a l'égalité suivante :

$$S(\varepsilon)U(x, \varepsilon) = \tilde{U}(\tilde{x}, \tilde{\varepsilon}). \tag{4.26}$$

Pour discrétiser le système (4.15) sur l'intervalle $[0; x_M]$, on choisit d'approcher (4.25) en considérant un maillage uniforme pour les variables $(\tilde{x}, \tilde{\varepsilon})$. Pour un intervalle fixé $[0; x_M]$, en lui appliquant le changement de variable (4.21), on obtient un intervalle $[0; \tilde{x}_M]$ avec :

$$\tilde{x}_M := \tilde{x}(x_M) = \int_0^{x_M} \rho(t)dt.$$

Sur cet intervalle $[0;\tilde{x}_M]$, on construit un maillage uniforme en notant :

$$\tilde{x}_{i+\frac{1}{2}} = i\Delta\tilde{x} \qquad \text{avec} \qquad \Delta\tilde{x} = \frac{\tilde{x}_M}{imax}, \qquad \text{pour } 0 \leq i \leq imax,$$

où $imax$ est le nombre de mailles et $\tilde{x}_{i+\frac{1}{2}}$ est l'interface entre les cellules i et $i+1$.
Si on note les mailles $\tilde{\mathcal{M}}_i = [\tilde{x}_{i-\frac{1}{2}};\tilde{x}_{i+\frac{1}{2}}]$, l'intervalle $[0;\tilde{x}_M]$ se décompose de la manière suivante :

$$\tilde{\mathcal{M}} = \bigcup_{i=1}^{imax} \tilde{\mathcal{M}}_i. \tag{4.27}$$

À partir de ce maillage $\tilde{\mathcal{M}}$, on obtient un maillage non-uniforme $\mathcal{M}$ pour discrétiser l'intervalle $[0;x_M]$:

$$x_{i+\frac{1}{2}} = \tilde{x}^{-1}(\tilde{x}_{i+\frac{1}{2}}), \tag{4.28}$$

où $\tilde{x}^{-1}$ est la fonction inverse de la fonction $\tilde{x}$ définie par (4.21).

Les mailles $\mathcal{M}_i$ forment une partition de $[0;x_M]$: $\bigcup_{i=1}^{imax}\mathcal{M}_i = [0;x_M]$, mais l'incrément $x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$ n'est en général pas constant.
On peut à présent écrire un schéma HLL rétrograde pour le système (4.25) muni de la donnée de Cauchy : $\Psi^0(\mathbf{x},\varepsilon_{max}) = \delta, \mathbf{\Psi^1}(\mathbf{x},\varepsilon_{max}) = 0$ sur le maillage uniforme $\tilde{\mathcal{M}}$. On introduit également un second maillage uniforme $\tilde{\varepsilon}^p = p\Delta\tilde{\varepsilon}$ en énergie.

On suppose ainsi connue une approximation constante par morceaux $\tilde{U}^h(\tilde{\varepsilon}^{p+1},\tilde{x})$ pour une énergie $\tilde{\varepsilon}^{p+1}$ définie par :

$$\tilde{U}^h(\tilde{\varepsilon}^{p+1},\tilde{x}) = \tilde{U}_i^{p+1} \quad \text{ si } \tilde{x} \in \tilde{\mathcal{M}}_i.$$

On fait évoluer cette solution en énergie pour obtenir une approximation $\tilde{U}_i^p$ à l'énergie $\tilde{\varepsilon}^p$. On obtient alors le schéma HLL rétrograde :

$$\tilde{U}_i^p = \tilde{U}_i^{p+1} - \frac{\Delta\tilde{\varepsilon}}{\Delta\tilde{x}}\left(\tilde{\mathcal{F}}_{i+\frac{1}{2}}^{p+1} - \tilde{\mathcal{F}}_{i-\frac{1}{2}}^{p+1}\right), \tag{4.29}$$

où le flux numérique $\tilde{\mathcal{F}}$ est donné par :

$$\tilde{\mathcal{F}}_{i+\frac{1}{2}}^{p+1} = \frac{1}{2}\left(F(\tilde{U}_i^{p+1}) + F(\tilde{U}_{i+1}^{p+1})\right) - \frac{1}{2}\left(\tilde{U}_{i+1}^{p+1} - \tilde{U}_i^{p+1}\right).$$

D'après [70] (voir aussi [22, 103, 86, 60]), le pas d'énergie $\Delta\tilde{\varepsilon}$ est restreint par la condition CFL suivante :

$$\Delta\tilde{\varepsilon} \leq \frac{\Delta\tilde{x}}{2\lambda_{max}}, \tag{4.30}$$

où λ_{max} est le module de la plus grande valeur propre. Ici, comme $\lambda_{max} \leq 2$, on prendra $\Delta\tilde{\varepsilon} \leq \Delta\tilde{x}/4$.
Une fois que l'on a discrétisé le système (4.25), l'idée est d'obtenir directement un schéma dans le jeu de variables initiales grâce à la relation (4.26) :

$$S(\varepsilon_p)U_i^p = S(\varepsilon_{p+1})\left(U_i^{p+1} - \frac{\Delta\tilde{\varepsilon}}{\Delta\tilde{x}}\left(\mathcal{F}_{i+\frac{1}{2}}^{p+1} - \mathcal{F}_{i-\frac{1}{2}}^{p+1}\right)\right), \tag{4.31}$$

avec

$$\mathcal{F}_{i+\frac{1}{2}}^{p+1} = \frac{1}{2}\left(F(U_i^{p+1}) + F(U_{i+1}^{p+1})\right) - \frac{1}{2}\left(U_{i+1}^{p+1} - U_i^{p+1}\right).$$

Précisons que la discrétisation en variables initiales (x,ε) n'est pas uniforme dès lors que ρ n'est pas constante. Les variables initiales sont donc discrétisées par un maillage non uniforme. Le pas d'énergie n'est quant à lui pas modifié par rapport au schéma sur le maillage uniforme.

Pour conclure la description du schéma en dimension un pour le système homogène (4.15), le lemme suivant donne les hypothèses requises pour assurer la robustesse de ce schéma :

Lemme 4.2.1. *Soit U_i^{p+1} dans l'espace $\mathcal{A}_1$ pour $1 \leq i \leq imax$. On suppose que le vecteur d'états $\tilde{U}_i^{p+1}$ à $\tilde{\varepsilon}^{p+1}$ est donné par (4.31). Alors, sous la condition CFL (4.30), U_i^p est dans l'espace $\mathcal{A}_1$ pour $1 \leq i \leq imax$.*

Démonstration. On suppose que U_i^{p+1} est dans l'espace $\mathcal{A}_1$ pour tout $1 \leq i \leq imax$. On a donc les deux relations suivantes :

$$(\Psi^0)_i^{p+1} > 0 \qquad \text{et} \qquad \frac{|(\Psi^1)_i^{p+1}|}{(\Psi^0)_i^{p+1}} < 1.$$

D'après la relation (4.26), on déduit immédiatement :

$$(\tilde{\Psi}^0)_i^{p+1} = S(\varepsilon^{p+1})(\Psi^0)_i^{p+1} \qquad \text{et} \qquad (\tilde{\Psi}^1)_i^{p+1} = S(\varepsilon^{p+1})(\Psi^1)_i^{p+1}.$$

Comme le pouvoir d'arrêt est une fonction positive, on a les relations :

$$(\tilde{\Psi}^0)_i^{p+1} > 0 \qquad \text{et} \qquad \frac{|(\tilde{\Psi}^1)_i^{p+1}|}{(\tilde{\Psi}^0)_i^{p+1}} < 1,$$

ce qui implique que $\tilde{U}_i^{p+1}$ est dans $\mathcal{A}_1$ pour tout $1 \leq i \leq imax$.
D'après [70, 22], sous la condition CFL (4.30), les états $\tilde{U}_i^p$ appartiennent à l'ensemble $\mathcal{A}_1$ pour $1 \leq i \leq imax$, dès que $\tilde{U}_i^{p+1} \in \mathcal{A}_1$ pour $1 \leq i \leq imax$.
Finalement, à partir des relations (4.26) et par positivité de la fonction S, on déduit que si $\tilde{U}_i^p \in \mathcal{A}_1$, alors $U_i^p \in \mathcal{A}_1$, ce qui conclut la démonstration. □

Discrétisation du terme source

On s'intéresse à présent à la discrétisation du système avec terme source. Dans le jeu de variables modifiées, le système initial (4.14) peut se réécrire :

$$\begin{aligned} &\partial_{\tilde{\varepsilon}}\tilde{\Psi}^0(\tilde{x},\tilde{\varepsilon}) - \partial_{\tilde{x}}\tilde{\Psi}^1(\tilde{x},\tilde{\varepsilon}) = 0, \\ &\partial_{\tilde{\varepsilon}}\tilde{\Psi}^1(\tilde{x},\tilde{\varepsilon}) - \partial_{\tilde{x}}\left(\tilde{\Psi}^0(\tilde{x},\tilde{\varepsilon})\chi\left(\frac{\tilde{\Psi}^1(\tilde{x},\tilde{\varepsilon})}{\tilde{\Psi}^0(\tilde{x},\tilde{\varepsilon})}\right)\right) = \tilde{T}(\tilde{\varepsilon})\tilde{\Psi}^1(\tilde{x},\tilde{\varepsilon}), \end{aligned} \tag{4.32}$$

où le changement de variables (4.26) et la définition de T (4.9) donnent :

$$\tilde{T}(\tilde{\varepsilon}(\varepsilon)) = T(\varepsilon). \tag{4.33}$$

Pour discrétiser le terme source, on utilise la méthode développée dans [19] que l'on a également utilisée pour discrétiser le modèle M_1 du transfert radiatif dans le paragraphe 3.2.4. On obtient alors le schéma :

$$\tilde{U}_i^p = \tilde{U}_i^{p+1} - \frac{\Delta\tilde{\varepsilon}}{\Delta\tilde{x}}\alpha\left(\tilde{\mathcal{F}}_{i+\frac{1}{2}}^{p+1} - \tilde{\mathcal{F}}_{i-\frac{1}{2}}^{p+1}\right) + 2\,\Delta\tilde{\varepsilon}\frac{(1-\alpha)}{\Delta\tilde{x}}\tilde{\Sigma}_i^{p+1}, \tag{4.34}$$

où

$$\alpha = \frac{2}{2+\Delta\tilde{x}},$$

et la forme discrète du terme source $\tilde{\Sigma}_i^{p+1}$ est donnée par :

$$\tilde{\Sigma}_i^{p+1} = \begin{pmatrix} 0 \\ (\tilde{\Psi}^1)_i^{p+1}\tilde{T}^{p+1} \end{pmatrix}.$$

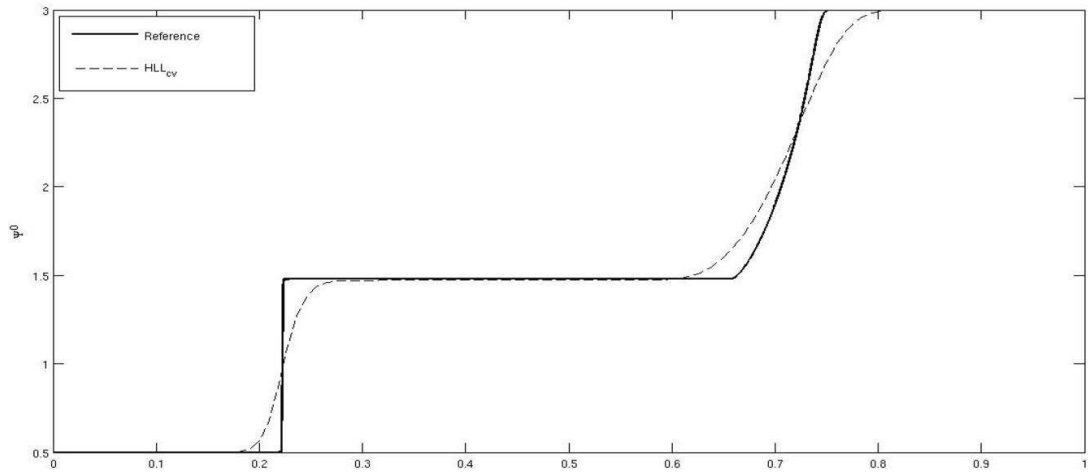


FIGURE 4.1 – Solution Ψ^0 du problème de Riemann (4.36) : solution de référence (trait plein) et approchée par le schéma HLL_{cv} (pointillés)

En utilisant les relations (4.26) et (4.33), on obtient un schéma pour les variables initiales (x, ε) :

$$S(\varepsilon_p)U_i^p = S(\varepsilon_{p+1})\left(U_i^{p+1} - \frac{\Delta\tilde{\varepsilon}}{\Delta\tilde{x}}\alpha\left(\mathcal{F}_{i+\frac{1}{2}}^{p+1} - \mathcal{F}_{i-\frac{1}{2}}^{p+1}\right) + 2\,\Delta\tilde{\varepsilon}\frac{(1-\alpha^{p+1})}{\Delta\tilde{x}}\Sigma_i^{p+1}\right), \qquad (4.35)$$

avec

$$\Sigma_i^{p+1} = \begin{pmatrix} 0 \\ (\Psi^1)_i^{p+1}T^{p+1} \end{pmatrix} \quad \text{et} \quad \alpha = \frac{2}{2+\Delta\tilde{x}}.$$

Ce schéma est noté HLL_{cv}, où l'indice cv signifie changement de variables.

Test de validation

Afin de valider ce schéma numérique, on réalise plusieurs cas-tests.
Le premier cas-test est dédié à l'approximation d'un problème de Riemann pour (4.32). La donnée initiale est composée de deux états constants séparés par une discontinuité en $x = 0.5$ sur l'intervalle $[0; 1]$:

$$\Psi^0(x, \varepsilon_{max}) = \begin{cases} 0.5 & \text{si } 0 \leq x < 0.5, \\ 3 & \text{si } 1 \geq x > 0.5, \end{cases} \qquad \Psi^1(x, \varepsilon_{max}) = 0. \qquad (4.36)$$

Pour cette expérience, on considère qu'il n'y a pas de terme source et donc que $T(\varepsilon) = 0$. De plus, on suppose que le pouvoir d'arrêt est constant $S(\varepsilon) = 1$, et on impose une densité discontinue :

$$\rho(x) = \begin{cases} 1 & \text{si } x < 0.3, \\ 0.01 & \text{si } 0.3 < x < 0.5, \\ 1 & \text{si } x > 0.5, \end{cases} \qquad (4.37)$$

Sur la figure 4.1, la solution donnée par le schéma HLL_{cv} est comparée à la solution de référence sur le domaine $[0; 1]$ discrétisé par 256 cellules. On compare la solution approchée Ψ^0 avec la solution de référence qui est la solution exacte du probème de Riemann. On observe que cette approximation est assez bonne et correspond bien à ce que l'on peut attendre d'un schéma HLL (voir

[103, 22]). En fait, la principale difficulté de cette expérience provient de la densité discontinue. Comme prévu, le schéma gère correctement ces discontinuités dans la définition de l'énergie.

Dans le paragraphe suivant, les résultats de ce cas-test seront comparés avec ceux d'une autre méthode et des améliorations de la convergence numérique seront détaillées.

4.2.2 Méthode avec projections en 1D

Dans le paragraphe précédent, on a développé une méthode spécifique au modèle (4.8) en dimension un basée sur des changements de variables pertinents. Grâce à ces changements de variables, le schéma (4.32) a pour principal avantage d'être indépendant des discontinuités de la fonction de densité $\rho(x)$. Le schéma est ainsi très pratique et donne une bonne approximation de la solution.

L'objectif est à présent d'étendre cette méthode au cas bidimensionnel. En effet, il n'existe alors plus de changement de variables global pour faire disparaître ρ dans le cas général. On remarque qu'en remplaçant la dérivée partielle par un opérateur dans (4.23), une telle transformation n'existe pas en général. Dans un premier temps, l'idée est de modifier le schéma (4.35) pour considérer un changement de variables *local* et ainsi permettre une extension simple en dimension deux.

Dans le schéma précédent (4.35), le maillage $\tilde{\mathcal{M}}$ était uniforme pour les variables modifiées. Puis à partir des nœuds $\tilde{x}_{i+\frac{1}{2}}$, on a obtenu un maillage non-uniforme $\mathcal{M}$ pour les variables initiales. Dans le schéma modifié, les deux maillages $\mathcal{M}$ et $\tilde{\mathcal{M}}$ sont uniformes. Par conséquent, les nœuds $x_{i+\frac{1}{2}}$ ne seront plus liés par la relation (4.21) et on aura en général :

$$\tilde{x}_{i+\frac{1}{2}} \neq \tilde{x}(x_{i+\frac{1}{2}}).$$

Dans ce paragraphe, on développe une technique de projection pour passer de l'un à l'autre.

Pour commencer, on définit les deux maillages $\mathcal{M}$ et $\tilde{\mathcal{M}}$.
Soient Δx et $\Delta\tilde{x}$, les largeurs respectives de deux cellules des maillages $\mathcal{M}$ et $\tilde{\mathcal{M}}$ telles que :

$$\begin{aligned} x_{i+\frac{1}{2}} &= i\Delta x \qquad \text{avec} \qquad \Delta x = \frac{x_M}{imax}, \\ \tilde{x}_{i+\frac{1}{2}} &= i\Delta\tilde{x} \qquad \text{avec} \qquad \Delta\tilde{x} = \frac{\tilde{x}_M}{imax}, \end{aligned}$$

où $\tilde{x}_M$ est toujours défini par (4.2.1). Les mailles sont alors définies par :

$$\mathcal{M}_i = [x_{i-\frac{1}{2}}; x_{i+\frac{1}{2}}] \qquad \text{et} \qquad \tilde{\mathcal{M}}_i = [\tilde{x}_{i-\frac{1}{2}}; \tilde{x}_{i+\frac{1}{2}}].$$

Les centres des cellules $\mathcal{M}_i$ et $\tilde{\mathcal{M}}_i$ sont notés x_i et $\tilde{x}_i$. Les mailles $\mathcal{M}_i$ et $\tilde{\mathcal{M}}_i$ forment alors une partition des intervalles $[0; x_M]$ et $[0; \tilde{x}_M]$:

$$\bigcup_{i=1}^{imax} \mathcal{M}_i = [0; x_M] \qquad \text{et} \qquad \bigcup_{i=1}^{imax} \tilde{\mathcal{M}}_i = [0; \tilde{x}_M].$$

À chaque énergie ε^{p+1}, on suppose connu le vecteur des états U_i^{p+1} sur le maillage $\mathcal{M}$. Pour calculer l'approximation de la solution $(U_i^p)_{1\leq i\leq imax}$ en ε^p à partir de la solution $(U_i^{p+1})_{1\leq i\leq imax}$, on utilise le schéma (4.34) sur le maillage $\tilde{\mathcal{M}}$. Il faut donc déterminer $\tilde{U}_i^{p+1}$ sur le maillage $\tilde{\mathcal{M}}$. Le but à présent est de définir une procédure de projection judicieuse pour passer de $\mathcal{M}$ à $\tilde{\mathcal{M}}$ et ainsi pouvoir évaluer $\tilde{U}_i^{p+1}$. Afin que le schéma reste conservatif, la projection doit également être conservative.

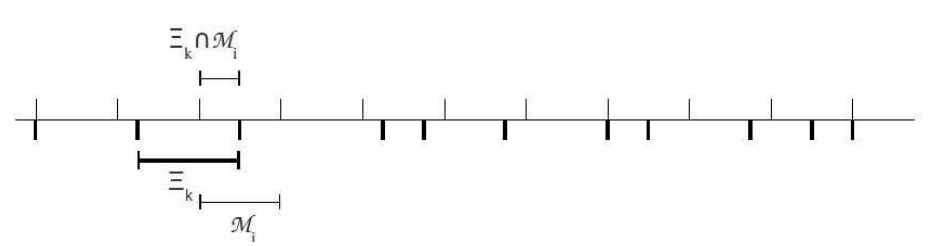


FIGURE 4.2 – Étapes de projection en dimension un

On introduit un troisième maillage non uniforme Ξ pour discrétiser le domaine $[0; x_M]$ dont les nœuds $\xi_{i+\frac{1}{2}}$ sont définis par :

$$\xi_{i+\frac{1}{2}} = \tilde{x}^{-1}(\tilde{x}_{i+\frac{1}{2}}),$$

et où $\tilde{x}^{-1}$ est le changement inverse de $\tilde{x}$ donné par (4.21).

Le centre de la cellule $\Xi_i = [\xi_{i-\frac{1}{2}}; \xi_{i+\frac{1}{2}}]$ est noté ξ_i .
Ce maillage Ξ coïncide exactement avec le maillage non-uniforme $\mathcal{M}$ défini par (4.28) dans la méthode avec changement de variables. On voit facilement que si la solution est connue sur chaque cellule du maillage Ξ, on peut appliquer le schéma (4.34)-(4.35) en utilisant les maillages Ξ et $\tilde{\mathcal{M}}$.

On définit ensuite la projection $\tilde{\Pi}$ qui projette la solution U, définie sur le maillage $\mathcal{M}$, en une fonction constante par morceaux $\tilde{U}$ sur le maillage $\tilde{\mathcal{M}}$:

$$(\tilde{\Pi}U)_i^p = S(\varepsilon^p) \sum_{k=1}^{imax} \tilde{a}_{i,k} U_k^p, \qquad \text{et} \qquad \tilde{a}_{i,k} = \frac{\text{mes}(\Xi_k \cap \mathcal{M}_i)}{\text{mes}(\mathcal{M}_i)}, \tag{4.38}$$

où le coefficient $\tilde{a}_{i,k}$ est le rapport des longueurs des cellules $\Xi_k \cap \mathcal{M}_i$ et $\mathcal{M}_i$ (voir figure 4.2).

En fait, on remarque que $\sum_{k=1}^{imax} \tilde{a}_{i,k} U_k^p$ correspond à la projection du vecteur U_i^p sur la cellule Ξ_i. En effet, l'application directe du changement de variables (4.26) $S(\varepsilon)U(x,\varepsilon) = \tilde{U}(\tilde{x},\tilde{\varepsilon})$ sur ce vecteur donne exactement la projection définie par (4.38).

De plus, la projection $\tilde{\Pi}$ est conservative. En effet, pour $1 \leq i \leq imax$ fixé, par définition de $\tilde{a}_{i,k} \geq 0$, on a :

$$\sum_{k=1}^{imax} \tilde{a}_{i,k} = 1.$$

Ainsi :

$$\sum_{i=1}^{imax} U_i^p = \sum_{i=1}^{imax} \left(\sum_{k=1}^{imax} \tilde{a}_{i,k} U_k^p \right),$$

ce qui montre la conservation de $\tilde{\Pi}$.

Inversement, la projection qui donne U sur le maillage $\mathcal{M}$ à partir de la solution $\tilde{U}$, définie sur le maillage $\tilde{\mathcal{M}}$, est donnée par :

$$(\Pi\tilde{U})_i^p = \frac{1}{S(\varepsilon^p)} \sum_{k=1}^{imax} a_{i,k} \tilde{U}_k^p, \qquad \text{et} \qquad a_{i,k} = \frac{\text{mes}(\mathcal{M}_k \cap \Xi_i)}{\text{mes}(\Xi_i)}. \tag{4.39}$$

Cette projection Π est également conservative car on a :

$$\sum_{i=1}^{imax} \tilde{U}_i^p = \sum_{i=1}^{imax} \left(\sum_{k=1}^{imax} \tilde{a}_{i,k} \tilde{U}_k^p \right).$$

Pour terminer l'introduction de ces projections Π et $\tilde{\Pi}$, on remarque qu'elles ne commutent pas en général : $\Pi(\tilde{\Pi}U) \neq \tilde{\Pi}(\Pi U)$.

Grâce aux projections définies ci-dessus, on peut mettre en place une procédure numérique pour calculer $(U_i^p)_{1\leq i\leq imax}$ à partir de $(U_i^{p+1})_{0\leq i\leq imax}$. En supposant que la solution $(U_i^{p+1})_{1\leq i\leq imax}$ est connue sur le maillage $\mathcal{M}$ à ε^{p+1}, le schéma numérique peut se résumer de la manière suivante :

1. À l'énergie ε^{p+1}, la solution $(U_i^{p+1})_{1\leq i\leq imax}$, définie sur le maillage $\mathcal{M}$, est projetée sur le maillage $\tilde{\mathcal{M}}$ grâce à l'opérateur $\tilde{\Pi}$ pour obtenir $(\tilde{U}_i^{p+1})_{1\leq i\leq imax}$:

$$\tilde{U}_i^{p+1} = (\tilde{\Pi}U)_i^{p+1} \qquad \text{pour } 1 \leq i \leq imax.$$

2. On applique le schéma (4.34) pour obtenir la solution approchée $(\tilde{U}_i^p)_{1\leq i\leq imax}$ sur le maillage $\tilde{\mathcal{M}}$ à ε^p.
3. Puis la solution $(\tilde{U}_i^p)_{1\leq i\leq imax}$ sur le maillage $\tilde{\mathcal{M}}$ est reprojetée sur le maillage initial $\mathcal{M}$ grâce à l'opérateur de projection Π :

$$U_i^p = (\Pi\tilde{U})_i^p \qquad \text{pour } 1 \leq i \leq imax.$$

Pour clarifier les notations dans la suite, on appelle cette procédure numérique $(4.34)_{\tilde{\Pi}}$-$(4.35)_{\Pi}$.
Le lemme suivant donne les hypothèses requises pour assurer la robustesse de ce schéma.

Lemme 4.2.2. *Soit U_i^{p+1} dans l'espace $\mathcal{A}_1$ pour $1 \leq i \leq imax$. On suppose que la solution U_i^p est donnée par la procédure numérique $(4.34)_{\tilde{\Pi}}$-$(4.35)_{\Pi}$. Alors, sous la condition CFL (4.30), pour tout $1 \leq i \leq imax$, U_i^p est dans l'espace $\mathcal{A}_1$.*

Démonstration. L'étape de projection $\tilde{\Pi}$ est une combinaison convexe des vecteurs U_i^{p+1}. Ainsi, dès que $U_i^{p+1} \in \mathcal{A}_1$ pour tout $1 \leq i \leq imax$, les vecteurs projetés $\tilde{U}_i^{p+1}$ sont toujours dans l'espace $\mathcal{A}_1$ qui est convexe. D'après [70], le schéma HLL (4.34) préserve les états admissibles sous la condition CFL (4.30) si le cône est assez ouvert. On en déduit que $\tilde{U}_i^p$ est dans l'espace $\mathcal{A}_1$ pour tout $1 \leq i \leq imax$.
Finalement, la reprojection de la solution sur le maillage $\mathcal{M}$, qui consiste en une combinaison convexe d'éléments de $\tilde{U}_i^p$, préserve U_i^p dans l'espace des états admissibles. D'où U_i^p est dans $\mathcal{A}_1$ pour tout $1 \leq i \leq imax$, ce qui conclut la démonstration. □

Ce schéma incluant des étapes de projection sera appelé HLL_{cvp}, où l'indice cvp signifie changement de variables avec projections.

Pour valider le schéma $(4.34)_{\tilde{\Pi}}$-$(4.35)_{\Pi}$, on s'intéresse au problème de Riemann (4.36) où la densité ρ et le pouvoir d'arrêt S sont définis par (4.37). Sur la figure 4.3, on compare les résultats donnés par ce schéma d'ordre un et d'ordre deux sur le domaine $[0; 1]$ discrétisé par 256 cellules. Le schéma d'ordre deux est obtenu en appliquant la méthode standard MUSCL (voir [12, 13, 86, 109]) : une fois que la solution est projetée sur le maillage $\tilde{\mathcal{M}}$, on effectue une reconstruction polynomiale de la solution $\tilde{U}$ que l'on utilise à chaque interface pour le calcul des flux. Pour éviter des pentes trop importantes lors de cette reconstruction polynomiale, on considère ici les limiteurs de pente minmod et Van Leer.

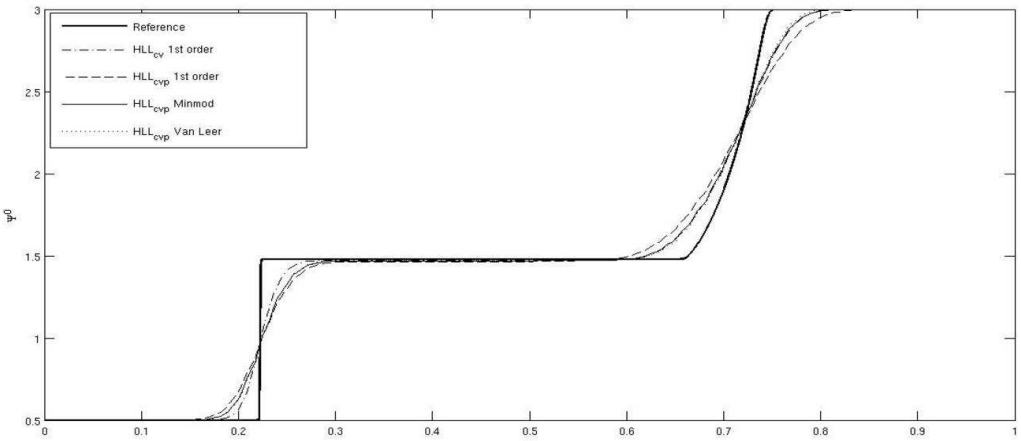


FIGURE 4.3 – Solution du problème de Riemann pour le schéma HLL_{cv} d'ordre un, le schéma HLL_{cvp} d'ordre un et le schéma HLL_{cvp} d'ordre deux avec les limiteurs minmod et Van Leer.

Dans le tableau 4.1, on compare les erreurs L^1, L^2 et L^∞ de ces schémas pour différents maillages. On remarque tout d'abord que le schéma d'ordre un avec projections donne approximativement la même erreur que le schéma sans projection en considérant deux fois plus de cellules. C'est pourquoi leurs ordres de convergence sont similaires. De plus, on voit que la correction MUSCL sur le schéma HLL_{cvp} permet d'obtenir des résultats légèrement meilleurs que le schéma HLL_{cv} d'ordre un. Cependant, à cause des étapes de projection, il n'améliore que sensiblement l'ordre de convergence du schéma HLL_{cvp}.
Cet exemple permet de valider l'approche HLL_{cvp} pour l'étendre ensuite au cas bidimensionnel dans le paragraphe suivant.

4.2.3 Méthode avec projections en dimension 2

Dans ce paragraphe, on propose une extension en dimension deux du schéma HLL_{cvp} pour approcher la solution du modèle (4.8). Pour alléger les notations, on écrit le modèle sous la forme condensée :

$$\partial_\varepsilon(\rho S U) - \partial_x F(U) - \partial_y G(U) = \Sigma(U), \tag{4.40}$$

où $U = (\Psi^0, \Psi^1_x, \Psi^1_y)^T$ est le vecteur des inconnues dans $\mathcal{A}_2$ et les flux F et G sont définis par :

$$F(U) = \begin{pmatrix} \Psi^1_x \\ \Psi^0 \dfrac{1-\chi}{2} + \dfrac{3\chi-1}{2\|\mathbf{\Psi}^1\|^2}(\Psi^1_x)^2 \\ \Psi^0 \dfrac{3\chi-1}{2\|\mathbf{\Psi}^1\|^2}\Psi^1_y\Psi^1_x \end{pmatrix}, \qquad G(U) = \begin{pmatrix} \Psi^1_x \\ \Psi^0 \dfrac{3\chi-1}{2\|\mathbf{\Psi}^1\|^2}\Psi^1_x\Psi^1_y \\ \Psi^0 \dfrac{1-\chi}{2} + \dfrac{3\chi-1}{2\|\mathbf{\Psi}^1\|^2}(\Psi^1_y)^2 \end{pmatrix}. \tag{4.41}$$

Le terme source Σ est donné par :

$$\Sigma(U) = (0, \rho T \Psi^1_x, \rho T \Psi^1_y)^T. \tag{4.42}$$

Pour cette extension en 2D sur des maillages cartésiens, on va réaliser un splitting en direction. On considère tout d'abord le système dans la direction x :

$$\partial_\varepsilon \rho S U - \partial_x F(U) = 0, \tag{4.43}$$

Méthode	nombre de cellules	erreur L^∞	erreur L^2	erreur L^1	ordre
	32	0.5596	0.2065	0.1390	
	64	0.4812	0.1585	9.6021 10^{-2}	0.5339
HLL_{cv}	128	0.4214	0.1186	6.1040 10^{-2}	0.6536
1^{er} ordre	256	0.3907	8.5568 10^{-2}	3.8265 10^{-2}	0.6737
	512	0.3971	6.2348 10^{-2}	2.3751 10^{-2}	0.6881
	1024	0.3911	4.2702 10^{-2}	1.3884 10^{-2}	0.7746
	2048	0.3455	2.7246 10^{-2}	7.4848 10^{-3}	0.8914
	32	0.6553	0.2438	0.1908	
	64	0.5620	0.2002	0.1468	0.3780
HLL_{cvp}	128	0.5028	0.1530	9.5539 10^{-2}	0.6200
1^{er} ordre	256	0.4679	0.1148	5.9349 10^{-2}	0.6869
	512	0.4523	8.5882 10^{-2}	3.7806 10^{-2}	0.6506
	1024	0.4476	6.1713 10^{-2}	2.3219 10^{-2}	0.7033
	2048	0.3960	4.2469 10^{-2}	1.3518 10^{-2}	0.7804
	32	0.6162	0.2312	0.1740	
	64	0.5271	0.1744	0.1208	0.5263
HLL_{cvp}	128	0.5167	0.1281	7.3075 10^{-2}	0.7255
2^{nd} ordre	256	0.4664	9.4580 10^{-2}	4.3889 10^{-2}	0.7355
Minmod	512	0.4461	6.9677 10^{-2}	2.7110 10^{-2}	0.6950
	1024	0.4419	4.9287 10^{-2}	1.5938 10^{-2}	0.7663
	2048	0.3999	3.3321 10^{-2}	8.7471 10^{-3}	0.8656
	32	0.6208	0.2278	0.1683	
	64	0.5315	0.1706	0.1170	0.5241
HLL_{cvp}	128	0.5153	0.1239	6.8487 10^{-2}	0.7730
2^{nd} ordre	256	0.4638	9.1065 10^{-2}	4.0879 10^{-2}	0.7445
Van Leer	512	0.4444	6.7074 10^{-2}	2.5365 10^{-2}	0.6885
	1024	0.4411	4.7387 10^{-2}	1.4715 10^{-2}	0.7856
	2048	0.3772	3.1995 10^{-2}	7.9505 10^{-3}	0.8882

TABLE 4.1 – Comparaisons des erreurs $L^1,\ L^2,\ L^\infty$ des différents schémas

puis, on considère le système dans la direction y :

$$\partial_\varepsilon \rho S U - \partial_y G(U) = 0. \tag{4.44}$$

Concernant le terme source, on utilise une extension en 2D du schéma (4.35).
Ainsi les deux systèmes (4.43)-(4.44) sont discrétisés par la procédure $(4.34)_{\tilde{\Pi}}$-$(4.35)_{\Pi}$ définie dans le paragraphe précédent. Une attention particulière sera apportée sur la densité qui dépend à présent de x et de y.

Dans un premier temps, on définit les maillages nécessaires à l'extension en dimension deux. On cherche les solutions sur le domaine $\Omega = [0; x_M] \times [0; y_M] \subset \mathbb{R}^2$. On considère le maillage cartésien uniforme formé des nœuds $(x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}})$ pour $0 \leq i \leq imax$ et $0 \leq j \leq jmax$. Les nœuds sont définis par :

$$\begin{aligned} x_{i+\frac{1}{2}} &= i\Delta x \qquad \text{avec} \qquad \Delta x = \frac{x_M}{imax}, \\ y_{j+\frac{1}{2}} &= j\Delta y \qquad \text{avec} \qquad \Delta y = \frac{y_M}{jmax}. \end{aligned}$$

Pour un j fixé, on en déduit le maillage pseudo-1D $\mathcal{M} = (\mathcal{M}_i)_{1 \leq i \leq imax}$ composé des cellules :

$$\mathcal{M}_i = [x_{i-\frac{1}{2}}; x_{i+\frac{1}{2}}].$$

Ces mailles forment ainsi une partition de l'intervalle $[0; x_M]$: $\bigcup_{i=1}^{imax} \mathcal{M}_i = [0; x_M]$. De la même manière, on définit $\mathcal{N} = (\mathcal{N}_j)_{1 \leq j \leq jmax}$, le maillage en dimension un formé des cellules :

$$\mathcal{N}_j = [y_{j-\frac{1}{2}}; y_{j+\frac{1}{2}}],$$

tel que $\bigcup_{j=1}^{jmax} \mathcal{N}_j = [0; y_M]$.
Comme précédemment, les centres des mailles $\mathcal{M}_i$ et $\mathcal{N}_j$ sont notés x_i et y_j. Avec ces définitions, le domaine Ω peut s'écrire de deux façons :

$$\bigcup_{j=1}^{jmax} \Big(\bigcup_{i=1}^{imax} \mathcal{M}_i \times [y_{j-\frac{1}{2}}; y_{j+\frac{1}{2}}] \Big) = [0; x_M] \times [0; y_M],$$

ou :

$$\bigcup_{i=1}^{imax} \Big(\bigcup_{j=1}^{jmax} \mathcal{N}_j \times [x_{i-\frac{1}{2}}; x_{i+\frac{1}{2}}] \Big) = [0; x_M] \times [0; y_M].$$

Puis, pour un indice j fixé, on considère le maillage 1D $\mathcal{M} \times \{y_j\}$ auquel on applique le changement de variables suivant :

$$\tilde{x}(x, y_j) = \int_0^x \rho(t, y_j)\, dt, \tag{4.45}$$

qui est directement tiré de (4.21). On pose alors :

$$\tilde{x}_M^j = \tilde{x}(x_M, y_j),$$

pour caractériser un maillage uniforme $\tilde{\mathcal{M}}^j$ discrétisant l'intervalle $[0; \tilde{x}_M^j]$. On remarque que ce maillage $\tilde{\mathcal{M}}^j$ coïncide exactement avec le maillage (4.27) si l'on fixe y_j. Les cellules sont alors définies par :

$$\tilde{\mathcal{M}}_i^j = [\tilde{x}_{i-\frac{1}{2}}^j; \tilde{x}_{i+\frac{1}{2}}^j],$$

où les nœuds sont :

$$\tilde{x}^j_{i+\frac{1}{2}} = i\Delta\tilde{x}^j_M \qquad \text{avec} \qquad \Delta\tilde{x}_j = \frac{\tilde{x}_j}{imax}.$$

Les centres des cellules $\tilde{\mathcal{M}}^j_i$ sont notés $\tilde{x}^j_i$.

Parallèlement, on introduit des maillages similaires dans la direction y. Ainsi, pour un indice i fixé, on applique le changement de variables suivant :

$$\tilde{y}(x_i, y) = \int_0^y \rho(x_i, s)\, ds, \tag{4.46}$$

au maillage 1D $\{x_i\} \times \mathcal{N}$. Ce changement de variables correspond exactement au changement de variables 1D (4.21) mais cette fois dans la direction y. On pose :

$$\tilde{y}^i_M = \tilde{y}(x_i, y_M),$$

pour définir un maillage uniforme $\tilde{\mathcal{N}}^i$ composé des cellules :

$$\tilde{\mathcal{N}}^i_j = [\tilde{y}^i_{j-\frac{1}{2}}; \tilde{y}^i_{j+\frac{1}{2}}].$$

Les nœuds sont :

$$\tilde{y}^i_{j+\frac{1}{2}} = j\Delta\tilde{y}^i_M \qquad \text{avec} \qquad \Delta\tilde{y}_i = \frac{\tilde{y}_i}{jmax},$$

et on note $\tilde{y}^i_j$ le centre de la cellule $\tilde{\mathcal{N}}^i_j$.

Le schéma numérique en dimension deux consiste alors à appliquer le schéma en dimension un tout d'abord dans la direction x pour chaque indice j en considérant les maillages $\mathcal{M} \times \{y_j\}$. Puis, on applique le schéma $(4.34)_{\tilde{\Pi}}$-$(4.35)_{\Pi}$ dans la direction y pour chaque indice i sur les maillages $\{x_i\} \times \mathcal{N}$.
Avec ce choix, les opérateurs de projection $\tilde{\Pi}$ et Π définis par (4.38) et (4.39) doivent être connus sur chaque maillage en x, $\mathcal{M} \times \{y_j\}$, et sur chaque maillage en y, $\{x_i\} \times \mathcal{N}$.

Dans un premier temps, on se place dans la direction x et on se fixe un indice j. On définit alors les projections correspondantes $\tilde{\Pi}^j_x$ et Π^j_x (voir figure 4.4). En se basant sur les relations (4.38) et (4.39), la projection est définie par :

$$(\tilde{\Pi}^j_x U)^p_i = S(\varepsilon^p) \sum_{k=1}^{imax} \tilde{a}^j_{i,k} U^p_{k,j}, \tag{4.47}$$

$$(\Pi^j_x \tilde{U})^p_i = \frac{1}{S(\varepsilon^p)} \sum_{k=1}^{imax} a^j_{i,k} \tilde{U}^p_{k,j}, \tag{4.48}$$

avec

$$\tilde{a}^j_{i,k} = \frac{\text{mes}(\Xi^j_{x,k} \cap \mathcal{M}_i)}{\text{mes}(\mathcal{M}_i)},$$
$$a^j_{i,k} = \frac{\text{mes}(\mathcal{M}_k \cap \Xi^j_{x,i})}{\text{mes}(\Xi^j_{x,i})}.$$

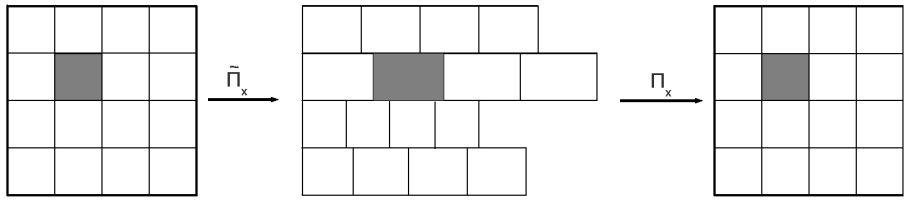


FIGURE 4.4 – Étapes de projection dans la direction x

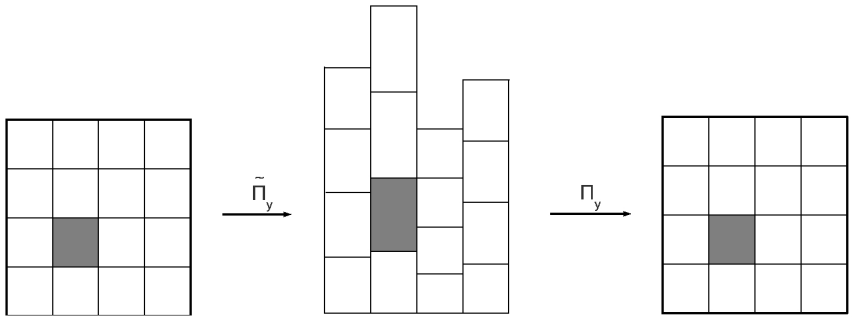


FIGURE 4.5 – Étapes de projection dans la direction y

Ici, Ξ est le maillage défini par :

$$\Xi^j_{x,i} = [\xi^j_{x,i-\frac{1}{2}}; \xi^j_{x,i+\frac{1}{2}}], \qquad \xi^j_{x,i+\frac{1}{2}} = \tilde{x}_j^{-1}(x_{i+\frac{1}{2}}),$$

où $\tilde{x}^{-1}$ représente l'inverse de la fonction $x \mapsto \tilde{x}(x, y_j)$ donnée par (4.45).
Puis, dans la direction y, on introduit les projections $\tilde{\Pi}^i_y : \mathcal{N}\times\{x_i\} \to \tilde{\mathcal{N}}^i$ et $\Pi^i_y : \tilde{\mathcal{N}}^i \to \mathcal{N}\times\{x_i\}$ (voir figure 4.5) qui sont données par :

$$(\tilde{\Pi}^i_y U)^p_j = S(\varepsilon^p) \sum_{k=0}^{jmax} \tilde{b}^i_{k,j} U^p_{i,k}, \tag{4.49}$$

$$(\Pi^i_y \tilde{U})^p_j = \frac{1}{S(\varepsilon^p)} \sum_{k=0}^{jmax} b^i_{k,j} \tilde{U}^p_{i,k}, \tag{4.50}$$

avec

$$\tilde{b}^i_{k,j} = \frac{\text{mes}(\Xi^i_{y,k} \cap \mathcal{N}_j)}{\text{mes}(\mathcal{N}_j)},$$
$$b^i_{k,j} = \frac{\text{mes}(\mathcal{N}_k \cap \Xi^i_{y,j})}{\text{mes}(\Xi^i_{y,j})}.$$

Dans un second temps, on considère le schéma avec projection en y. De même que dans la direction x, le maillage Ξ est maintenant donné par :

$$\Xi^i_{y,j} = [\xi^i_{y,j-\frac{1}{2}}, \xi^i_{y,j+\frac{1}{2}}], \qquad \xi^i_{y,j+\frac{1}{2}} = \tilde{y}_i^{-1}(y_{j+\frac{1}{2}}),$$

où $\tilde{y}^{-1}$ représente l'inverse du changement de variable $y \mapsto \tilde{y}(x_i, y)$ donné par (4.46).

Une fois que les projections sont déterminées, on peut mettre en place la procédure numérique qui permet de calculer la solution approchée $U_{i,j}^{p}$ à l'énergie ε^p à partir de la solution approchée $U_{i,j}^{p+1}$ à ε^{p+1}. Cette méthode se déroule en deux temps étant donné que chaque direction est considérée séparément.

- Le premier pas consiste à faire évoluer la solution dans la direction x. Pour un indice j fixé, on calcule le vecteur $(\tilde{U}_{i,j}^{p+\frac{1}{2}})_{1\leq i\leq imax}$ grâce à la projection (4.47). Puis, on fait évoluer cette solution en utilisant le schéma HLL avec terme source défini par (4.34). Le schéma s'écrit :

$$\tilde{U}_{i,j}^{p+\frac{1}{2}} = \tilde{U}_{i,j}^{p+1} - \frac{\Delta\tilde{\varepsilon}}{\Delta\tilde{x}}\alpha_x\left(\tilde{\mathcal{F}}_{i+\frac{1}{2},j}^{p+1} - \tilde{\mathcal{F}}_{i-\frac{1}{2},j}^{p+1}\right) + 2\Delta\tilde{\varepsilon}\frac{(1-\alpha_x)}{2\Delta\tilde{x}}\tilde{\Sigma}_{i,j}^{p+1},$$

où :

$$\tilde{\mathcal{F}}_{i+\frac{1}{2},j}^{p+1} = \frac{1}{2}\left(F(\tilde{U}_{i,j}^{p+1} + F(\tilde{U}_{i+1,j}^{p+1})\right) - \frac{1}{2}\left(\tilde{U}_{i+1,j}^{p+1} - \tilde{U}_{i,j}^{p+1}\right)$$
$$\alpha_x = \frac{2}{2+\Delta\tilde{x}},$$

et

$$\tilde{\Sigma}_{i,j}^{p+1} = \begin{pmatrix} 0 \\ \tilde{T}^{p+1}(\tilde{\Psi}_x^1)_{i,j}^{p+1} \\ \tilde{T}^{p+1}(\tilde{\Psi}_y^1)_{i,j}^{p+1} \end{pmatrix}.$$

Finalement, on reprojette le vecteur $(\tilde{U}_{i,j}^{p+\frac{1}{2}})_{1\leq i\leq imax}$ grâce à l'opérateur (4.48) pour obtenir l'état intermédiaire $(U_{i,j}^{p+\frac{1}{2}})_{1\leq i\leq imax}$ pour tout $1 \leq j \leq jmax$.

- Le second pas consiste à faire évoluer l'état intermédiaire $U_{i,j}^{p+\frac{1}{2}}$ dans la direction y. Pour chaque indice i, on définit le vecteur $(\tilde{U}_{i,j}^{p+\frac{1}{2}})_{1\leq j\leq jmax}$ grâce à l'opérateur de projection (4.49). Puis, on lui applique le schéma HLL pour calculer la solution à ε_p :

$$\tilde{U}_{i,j}^{p} = \tilde{U}_{i,j}^{p+\frac{1}{2}} - \frac{\Delta\tilde{\varepsilon}}{\Delta\tilde{y}}\alpha_x\left(\tilde{\mathcal{G}}_{i,j+\frac{1}{2}}^{p+\frac{1}{2}} - \tilde{\mathcal{G}}_{i,j-\frac{1}{2}}^{p+\frac{1}{2}}\right) + 2\Delta\tilde{\varepsilon}\frac{(1-\alpha_x)}{2\Delta\tilde{y}}\tilde{\Sigma}_{i,j}^{p+\frac{1}{2}},$$

où

$$\tilde{\mathcal{G}}_{i,j+\frac{1}{2}}^{p+\frac{1}{2}} = \frac{1}{2}\left(G(\tilde{U}_{i,j}^{p+\frac{1}{2}} + G(\tilde{U}_{i,j+1}^{p+\frac{1}{2}})\right) - \frac{1}{2}\left(\tilde{U}_{i,j+1}^{p+\frac{1}{2}} - \tilde{U}_{i,j}^{p+\frac{1}{2}}\right) \qquad \text{et} \qquad \alpha_y = \frac{2}{2+\Delta\tilde{y}},$$

et

$$\tilde{\Sigma}_{i,j}^{p+\frac{1}{2}} = \begin{pmatrix} 0 \\ \tilde{T}^{p+\frac{1}{2}}(\tilde{\Psi}_x^1)_{i,j}^{p+\frac{1}{2}} \\ \tilde{T}^{p+\frac{1}{2}}(\tilde{\Psi}_y^1)_{i,j}^{p+\frac{1}{2}} \end{pmatrix}.$$

Finalement, on reprojette ce vecteur grâce à l'opérateur (4.50) pour obtenir la solution approchée $U_{i,j}^{p}$ de (4.8) à ε^p.

Test de validation

Pour valider la procédure numérique en 2D, on considère un cas-test possédant une symétrie sphérique. On envoie un rayon d'électrons dans une sphère de rayon 1 cm dans un milieu de

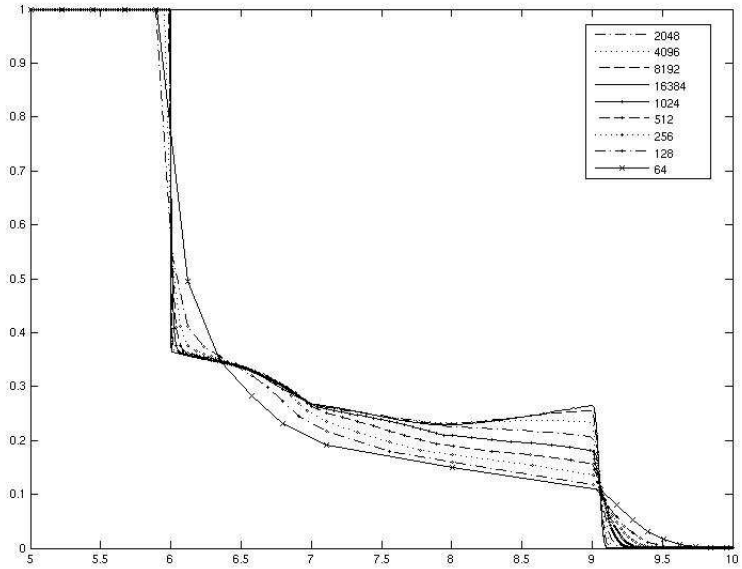


FIGURE 4.6 – Cas sphérique : Ψ^0 donné par le code 1D avec un raffinement de maillage de 128 à 16384 cellules.

densité variable donnée par :

$$\rho(r) = \begin{cases} 1 & \text{si } 1 < r \leq 2, \\ 0.25 & \text{si } 2 < r \leq 3, \\ 0.1 & \text{si } 3 < r \leq 4, \\ 2 & \text{sinon.} \end{cases}$$

Le but de ce cas-test est de valider la procédure de projections. Pour simplifier le modèle, le terme source est supposé nul, soit $T = 0$. Puis le pouvoir d'arrêt est fixé à $S = 1$ et la fermeture du modèle est donnée par $\mathbf{\Psi}^2 = \frac{1}{3}\Psi^0 \mathbf{I}_d$. L'un des intérêt de ce cas-test est de ramener le système à un système en dimension un en considérant les coordonnées polaires. En effet, on a :

$$\partial_\varepsilon(r\Psi^0) + \partial_r(r\Psi^1) = 0,$$
$$\partial_\varepsilon(r\Psi^1) + \partial_r(\frac{r\Psi^0}{3}) = \frac{r\Psi^0}{3r}.$$

En utilisant les variables $(r\Psi^0, r\Psi^1)$, on a un système en dimension un sans terme source. On peut alors le discrétiser en utilisant la procédure 1D (4.34)$_{\tilde{\Pi}}$-(4.35)$_{\Pi}$ et définir une solution de référence pour comparer les résultats de la procédure en dimension deux. Cette simulation est très raide et un grand nombre de points est nécessaire pour faire converger la solution comme le montre la figure 4.6. Cette approximation d'une solution sphérique est d'autant plus difficile pour la procédure en 2D qu'elle est basée sur un maillage cartésien et qu'elle comporte de nombreuses projections. Sur la figure (4.7), on compare les résultats sur des maillages formés de 256×256 et 1024×1024 mailles. On remarque bien que la solution donnée par le schéma en dimension deux reste une bonne approximation de la solution de référence. La méthode de projection en dimension deux donne donc une bonne approximation de la solution. Dans le paragraphe 5.2, elle est testée sur des cas-tests physiques de calculs de dose et comparée à d'autres méthodes.

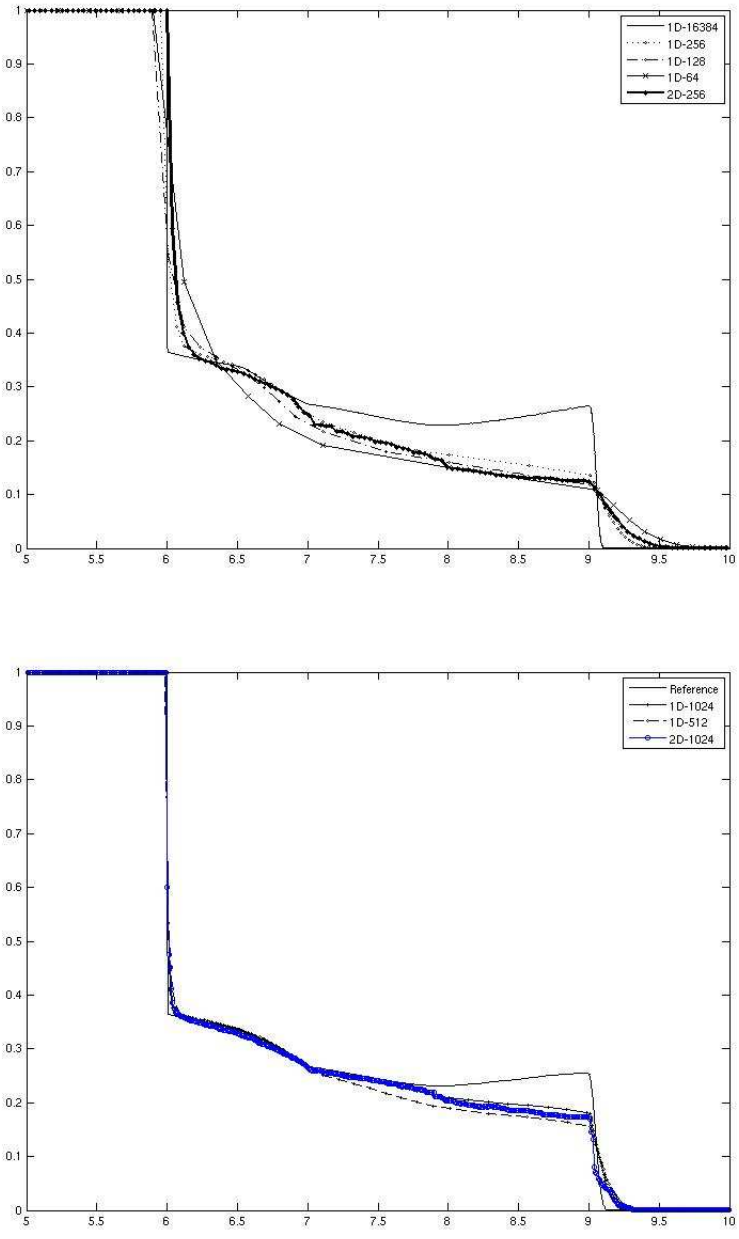


FIGURE 4.7 – Cas sphérique : Ψ^0 donné par le code 2D avec 256x256 cellules (haut) et 1024x1024 cellules (bas).

5

Résultats numériques

Dans ce chapitre, les méthodes numériques détaillées précédemment dans ce manuscrit sont illustrées sur quelques cas-tests. On présente dans un premier temps des résultats en dimension un pour le schéma GRP espace-temps discrétisant l'équation d'advection. En particulier, on applique ce schéma aux équations de Maxwell linéarisées et au modèle d'ordonnées discrètes. Puis on compare le schéma GRP espace-temps discrétisant le modèle d'ordonnées discrètes S_N à un modèle de diffusion sur un cas de rentrée atmosphérique. Dans un second temps, quelques résultats numériques en dimension deux sont proposés tout d'abord pour illustrer le comportement de la méthode GRP espace-temps pour l'équation d'advection, puis pour comparer la procédure numérique HLL_{cvp} aux méthodes Monte-Carlo pour le calcul de dose en radiothérapie.

5.1 Résultats en dimension 1

5.1.1 Schéma GRP espace-temps pour l'équation d'advection

Équations de Maxwell linéarisées

On utilise le schéma GRP espace-temps pour approcher le système des équations de Maxwell linéarisées :

$$\begin{aligned} \partial_t B_z + \partial_x E_y &= 0, \\ \partial_t E_y + c^2 \partial_x B_z &= 0, \end{aligned}$$

où c est la vitesse de la lumière, B_z et E_y sont respectivement les composantes z du champ magnétique et y du champ électrique. En diagonalisant le système dans la base des vecteurs propres, on est ramené à un système de transport découplé et on peut donc appliquer la méthode GRP espace-temps pour 'approcher numériquement.

Le cas-test consiste à envoyer un signal dans la partie gauche du domaine pour simuler un rayon laser :

$$\begin{aligned} B_z(x=0,t) &= 0.5 \exp\left(-100(t-0.5)^2\right) \sin(80(t-0.5)), \\ E_y(x=0,t) &= 0. \end{aligned}$$

À droite, on impose une condition de type Neumann homogène. On se donne pour donnée initiale $B_z(x, t=0) = E_y(x, t=0) = 0$. Finalement, on normalise la vitesse de la lumière : $c = 1$.

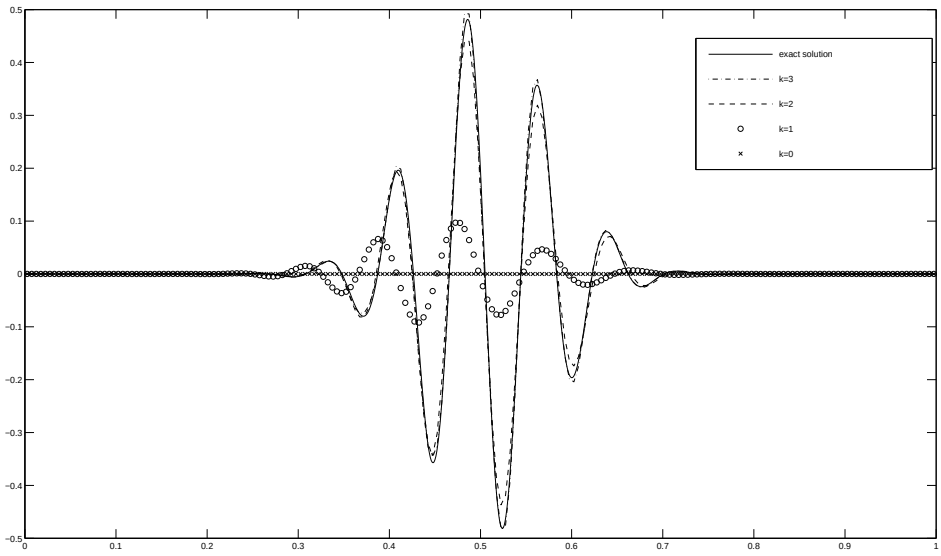


FIGURE 5.1 – Cas-test de Maxwell : solutions à $t = 1$ avec $\lambda = 5$ pour $k = 0, 1, 2, 3$.

La principale difficulté ici vient du fait que la solution est entièrement déterminée par la condition limite gauche qui dépend du temps et oscille fortement. Sur la figure 5.1, on compare les résultats pour $k = 0, 1, 2, 3$ avec 256 points et $\lambda = 5$ avec la solution exacte.

On remarque que la méthode d'ordre un n'est pas du tout capable de reproduire le comportement de la solution exacte. Dans ce cas où $\lambda = 5$, la diffusion numérique devient prédominante. D'autre part, à partir de $k = 2$, la solution est bien préservée par le schéma. En effet, dans ce cas particulier où la condition limite dépend fortement du temps, en considérant une approximation d'ordre $k+1$ sur les interfaces, on améliore largement l'approximation au niveau de la condition aux limites.

Modèle d'ordonnées discrètes S_N du transfert radiatif

Dans ce cas-test, on veut montrer que le schéma GRP espace-temps pour l'équation d'advection peut être facilement couplé avec une méthode numérique adaptée pour intégrer le terme source. On considère pour cela le modèle d'ordonnées discrètes S_N du transfert radiatif présenté dans le paragraphe 1.3 intégré sur tout le spectre de fréquences :

$$\partial_t I_j + c\mu_j \partial_x I_j = c\sigma(aT^4 - I_j), \text{ pour } j = 1 \dots N,$$

où $\mu_j \in [-1, 1]$ est la direction de propagation, N est le nombre total de directions considérées et $c = 1$ est la vitesse normalisée. D'un point de vue numérique, on considère un *splitting* de Strang pour prendre en compte le terme source. Pour simplifier, on ne considèrera que les premiers moments $\alpha_i^{n,0}$ lorsqu'on traitera le terme source. Cette technique est d'ordre deux et des stratégies de *splitting* d'ordre plus élevé peuvent être considérées si nécessaire.

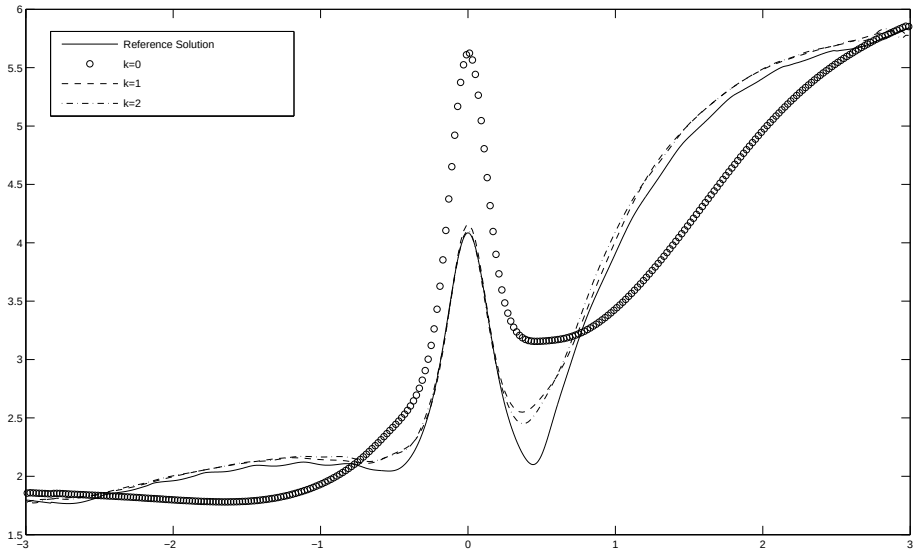


FIGURE 5.2 – Cas-test S_N : solutions à $t = 2.5$ avec $\lambda = 16$ pour $k = 0, 1, 2$.

Sur la figure 5.2, on représente l'énergie radiative $E_R(x,t) = 2\pi \sum_j I_j \omega_j$ correspondant à des directions μ_j équiréparties sur $[-1, 1]$ et obtenue avec les données suivantes :

$$\begin{aligned} aT^4(x) &= e^{-20x^2}, \\ \sigma(x) &= e^{-20x^2}, \\ I_j(x, t = 0) &= 10^{-2}, \\ I_1(x = -3, t) &= 0.3(1 - e^{-t})/\omega_1,\ I_N(x = 3, t) = (1 - e^{-t})/\omega_N, \end{aligned}$$

les conditions aux limites restantes étant de type Neumann. De plus, on prend $N = 16$ et $\lambda = 16$. En effet, comme on doit utiliser le même pas de temps pour tout le système, cela signifie que λ est maximal égal à 16 (quand $\mu = \pm 1$) mais est petit dans les autres directions. Sur la figure 5.2, on observe qu'il y a une grande différence entre la solution de référence obtenue avec $k = 2$ et 16000 points et l'approximation d'ordre un ($k = 0$) avec 256 points. Le schéma d'ordre deux ($k = 1$) donne une meilleure approximation. De plus, si l'on considère des polynômes de degré 2, la solution est encore améliorée sachant que le schéma n'est que d'ordre deux à cause du traitement du terme source par *splitting* de Strang.

5.1.2 Schéma GRP espace-temps pour le modèle d'ordonnées discrètes

Contrairement au paragraphe précécent où l'on a utilisé une méthode de splitting pour discrétiser le terme source, on considère dans ce paragraphe le schéma GRP espace-temps discrétisant le modèle d'ordonnées discrètes complet. Ce cas-test a été présenté dans [45]. On s'intéresse ici aux transferts de chaleur dans un bouclier de protection thermique pour une sonde spatiale. Ce bouclier est composé d'un matériau poreux de type PICA, mis au point par la NASA dans les années 1990. Pour déterminer ses propriétés thermiques, les méthodes utilisées jusqu'à présent consistent à résoudre des problèmes inverses sur une physique très simplifiée en l'occurrence, en utilisant le modèle de diffusion à l'équilibre présenté dans le paragraphe 1.4.3 à laquelle on ajoute de la diffusion thermique. Du fait de ces simplifications, les résultats sont ne sont pas toujours précis. On peut assimiler le modèle physique utilisé au modèle de diffusion à l'équilibre décrit

dans le paragraphe 1.2 auquel on ajoute de la diffusion thermique :

$$\partial_t(aT^4 + \rho C_v T) - \nabla\left(\frac{4aT^3 c}{3\tilde{\sigma_\nu}}\nabla T\right) = \partial_x\left(\lambda_c \partial_x T\right). \tag{5.1}$$

Ce modèle n'est donc *a priori* valide que dans la limite de diffusion, c'est-à-dire quand σt est grand, or en laboratoire, les expériences consistent à envoyer un rayon laser sur le matériau afin d'en mesurer les propriétés thermiques. Mais la forte anisotropie du rayon incident est incompatible avec le régime de diffusion. Pour obtenir des résultats plus précis, on compare alors cette méthode avec le schéma GRP espace-temps pour le modèle d'ordonnées discrètes du transfert radiatif couplé avec une équation sur l'énergie matière donné par 2.32. Ici, il est raisonnable de choisir des pas de temps élevés pour capter les phénomènes très lents en jeu.

On s'intéresse ici à une zone de un matériau de longueur $5.10^{-3}\ m$ discrétisée par 100 mailles, soit un pas d'espace $\Delta x = 5.10^{-5}\ m$. Le matériau a une densité $\rho = 240\ kg.m^{-3}$, une opacité $\sigma = 7700\ m^{-1}$, une conductivité thermique $\lambda_c(T) = 0.2 + 3.5.10^{-11}\ T^3$ et une capacité thermique $C_p(T)$ donnée. Le pas de temps considéré ici est $\Delta t = 10^{-4}$. Pour comparaison, si l'on impose $\lambda = 1$ (soit une CFL explicite), le pas de temps correspondant est $\Delta t \simeq 1.7.10^{-13} s$, soit environ 6.10^8 fois plus petit que le pas de temps considéré. Encore une fois, ce pas de temps n'est raisonnable qu'au vu de la lenteur du phénomène qui n'avance que d'une ou deux mailles par pas de temps.
Les N directions considérées sont les racines du $N^{ième}$ polynôme de Legendre.
Dans tout le domaine, la température initiale du matériau est de $T_{init} = 300K$ et on considère une donnée à l'équilibre radiatif avec $E_R = aT_{init}^4$ et $F_R = 0$.
On étudie deux cas. Tout d'abord, afin de recréer les conditions expérimentales, on considère un rayon entrant dans la direction $\mu = 1$ avec une énergie $E(t)$ croissante :

$$E(\mu = 1, t) = aT_{init}^4 + \min(1, 10t)(aT_{max}^4 - aT_{init}^4), \tag{5.2}$$

et un facteur d'anisotropie $f = 1$ dans cette direction.
Puis, pour recréer ce qui se passe lors de la rentrée atmosphérique, on suppose qu'on est à l'équilibre radiatif à gauche du domaine avec une énergie donnée par :

$$E(t) = aT_{init}^4 + \min(1, 10t)(aT_{max}^4 - aT_{init}^4). \tag{5.3}$$

On prendra ici $T_{max} = 3600\ K$. Dans les deux cas, on impose des conditions aux limites de type Neumann sur le bord droit.
Pour déterminer dans un premier temps le nombre de directions optimal pour les simulations, on compare les résultats obtenus avec différents N dans le cas d'un rayon. Sur la figure 5.3, on représente la température radiative obtenue au bout d'un temps $t = 1s$.

On observe que la différence entre 32 et 64 directions est très faible par rapport aux autres choix de N. Dans la suite, on considèrera ainsi que 64 directions suffisent à obtenir une bonne approximation de la solution.
Parallèlement, on a comparé les approximations en fonction du degré des polynômes k pour en conclure qu'à partir de $k = 4$, la solution ne change quasiment plus pour le Δt choisi. On choisira donc $k = 4$ dans la suite pour comparer les différents modèles.
On peut à présent comparer les résultats donnés par le schéma GRP espace-temps pour le modèle d'ordonnées discrètes et un schéma d'ordre 2 différences finies standard pour le modèle de diffusion à l'équilibre. Pour le modèle de diffusion, on considère un pas de temps $\Delta t = 1.9.10^{-6} s$.

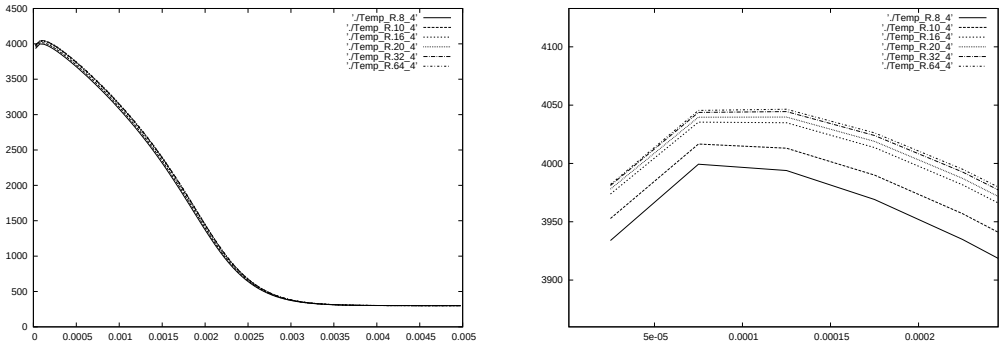


FIGURE 5.3 – Comparaison de la température radiative pour différents nombres de directions (gauche) et températures zoomées (droite) à $t = 1s$.

Solutions à $\mathbf{t = 0.1}$ s

Sur la figure 5.4, on compare la température matière donnée par deux méthodes (temp_D pour le modèle de diffusion, temp_M pour le modèle S_N) pour le cas d'un rayon au bout de $t = 0.1s$ ainsi que la température radiative (temp_R) donnée par la méthode GRP espace-temps pour le modèle S_N. On rappelle que le modèle de diffusion n'est pas capable de gérer un rayon, c'est pourquoi l'on impose une énergie croissante (5.3) sur le bord gauche. On représente également le facteur d'anisotropie $f = \frac{\|F_R\|}{E_R}$ obtenu par le modèle S_N. Puis sur la figure 5.5, on compare les deux méthodes pour le cas de l'équilibre et on trace également le facteur d'anisotropie du modèle S_N.

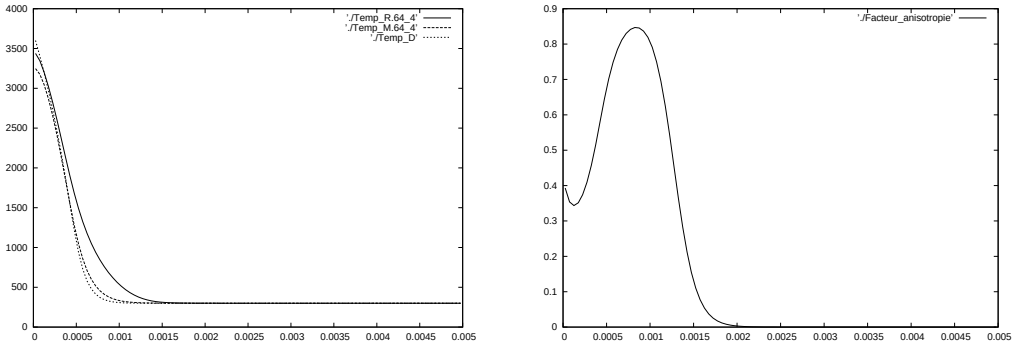


FIGURE 5.4 – Températures radiative et matière (gauche) et facteur d'anisotropie (droite) à $t = 0.1\ s$ pour un rayon.

Sans surprise, on constate que l'équilibre radiatif est loin d'être atteint sur le bord du domaine. Le modèle de diffusion n'est donc pas adapté et on observe des écarts de température matière superficielle entre les 2 modèles. Ce constat s'applique même dans le cas où la condition aux limites est à l'équilibre radiatif : on voit apparaître des facteurs d'anisotropie très élevés de l'ordre de 0.75.

Solutions à $\mathbf{t = 1}$ s

Sur la figure 5.6, on compare les deux modèles pour le cas d'un rayon et sur la figure 5.6, dans le cas de l'équilibre à $t = 1s$.

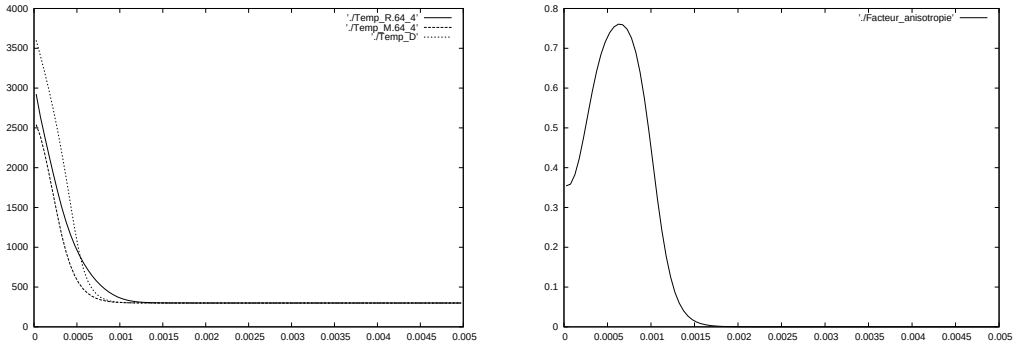


FIGURE 5.5 – Températures radiative et matière (gauche) et facteur d'anisotropie (droite) à $t = 0.1\ s$ pour des conditions aux limites à l'équilibre radiatif.

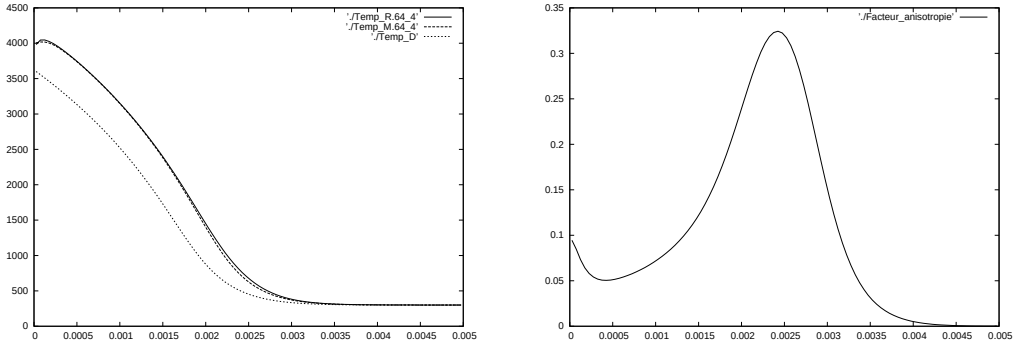


FIGURE 5.6 – Températures radiative et matière (gauche) et facteur d'anisotropie (droite) à $t = 1\ s$ pour un rayon.

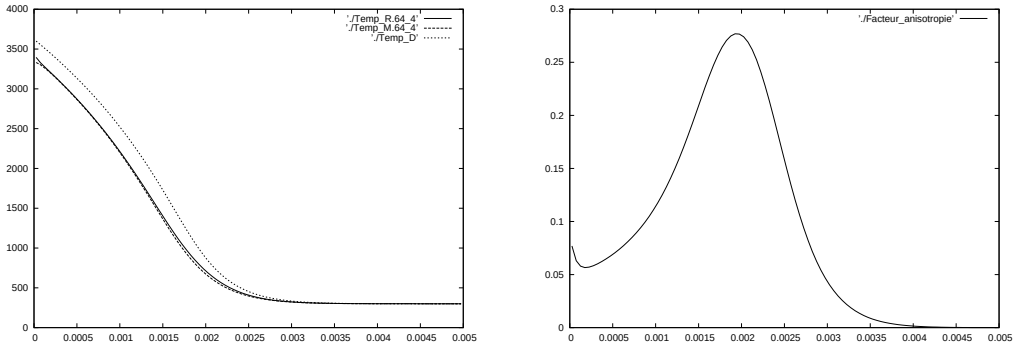


FIGURE 5.7 – Températures radiative et matière (gauche) et facteur d'anisotropie (droite) à $t = 1\ s$ pour des conditions aux limites 'a l'équilibre radiatif.

On constate encore une fois que l'on est loin de l'équilibre radiatif avec des facteurs d'anisotropie qui sont encore de l'ordre de 0.3. Le modèle de diffusion n'est donc pas adapté. Des erreurs significatives subsistent en effet y compris avec une condition aux limites à l'équilibre. L'erreur commise augmente toutefois moins vite en temps au fur et à mesure que l'on s'approche du régime asymptotique.

5.2 Résultats en dimension 2

5.2.1 Schéma GRP espace-temps pour l'équation d'advection

Modèle d'ordonnées discrètes S_N
Ce cas-test illustre le comportement de la méthode GRP espace-temps en dimension deux pour le modèle d'ordonnées discrètes S_N :

$$\partial_t I_j + c\Omega_j.\nabla_{\mathbf{x}} I_j = c\sigma(aT^4 - I_j), \text{ pour } j = 1 \dots N.$$

Comme en dimension un, le terme source a été pris en compte en considérant un splitting de Strang. Dans ce but, on ne considère toujours que les premiers moments de la solution dans chaque cellule.
L'énergie radiative $E = \sum_j I_j \omega_j$ est calculée avec : $N = 16$, $\lambda = 4$, $E(\mathbf{x}, t = 0) = 0$ et

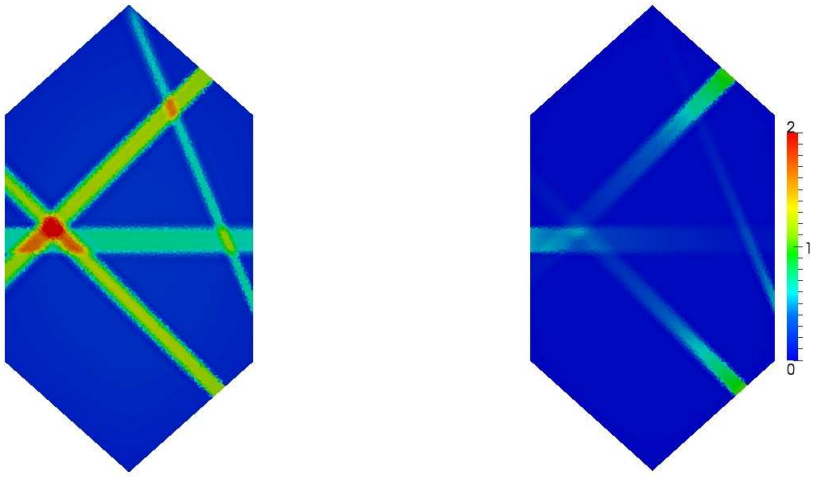


FIGURE 5.8 – Énergie radiative à $t = 1.3$ pour $\sigma = 0.1$(gauche) et $\sigma = 5$ (droite).

$aT^4 = 0.01$ sur un maillage composé de 15167 triangles discrétisant un domaine hexagonal. Deux rayons entrent dans le domaine avec une énergie égale à 1 et deux autres rayons avec une énergie égale à 0.5. Sur la figure 5.8, les approximations sont obtenues pour $\sigma = 0.1$ et $\sigma = 5$ pour $k = 1$ (degré des polynômes), ce qui correspond à un *splitting* de Strang d'ordre deux.
On voit clairement que même avec $k = 1$ et $\lambda = 4$, la diffusion numérique est négligeable dans la direction transverse des rayons.

5.2.2 Schéma préservant l'asymptotique pour le modèle M_1 multigroupe

On propose à présent un cas-test pour illustrer le schéma préservant l'asymptotique pour le modèle M_1 multigroupe et les précalculs des opacités développés dans le chapitre 3. On considère des volumes finis vertex centered. Pour cela, on prend le maillage dual d'un maillage triangulaire. Dans le paragraphe 3.2.4, on a développé une correction asymptotique permettant de retrouver numériquement le régime limite du modèle à condition d'avoir un maillage admissible ou une reconstruction du gradient de la solution sur chaque interface. En pratique, le maillage n'est pas complètement admissible, mais les angles $(x_K \hat{x_L}, KL)$ sont presque des angles droits. Par souci de simplicité on utilise ici la correction asymptotique pour les maillages admissibles. Ainsi, à la limite, le schéma sera consistant mais pas d'ordre 1.
Le domaine de calcul est rempli de krypton. L'opacité correspondante de ce gaz a été obtenue

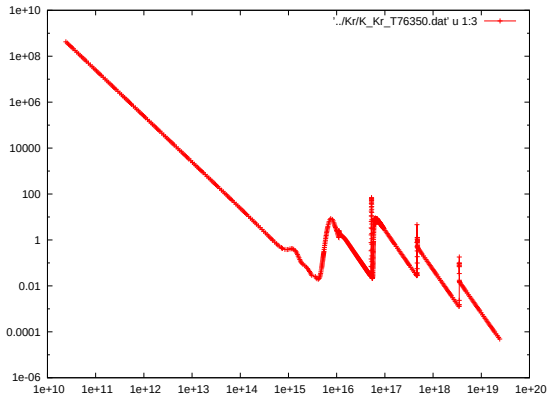


FIGURE 5.9 – Spectre du Krypton à $T = 76360K$

grâce à la base de données ODALISC [102]. On dispose de l'opacité d'absorption sur un large spectre de fréquences $2.5.10^{10}m^{-1} - 2.5.10^{19}m^{-1}$ pour plusieurs températures matière (de $1eV$ à $20eV$, soit environ de $11800K$ à $245000K$). En observant ce spectre représenté sur la figure 5.9, on choisit de le diviser en 6 groupes de fréquences $[\nu_q; \nu_{q+1}]_{1\leq q\leq 6}$:

$$\nu_1 = 1.10^{13}m^{-1};\ \nu_2 = 3.5.10^{14}m^{-1},\ \nu_3 = 3.10^{15}m^{-1},\ \nu_4 = 6.10^{15}m^{-1},$$
$$\nu_5 = 9.10^{15}m^{-1},\ \nu_6 = 4.6.10^{16}m^{-1},\ \nu_7 = 5.10^{17}m^{-1}.$$

Lors de la procédure de précalculs des moyennes d'opacités, la température matière a été supposée constante. Pour améliorer le coût de calcul et puisque les variations de l'opacité sont assez régulières en température, on considère qu'elle varie de manière polynomiale au cours du temps. En pratique, les coefficients sont calculés au sens des moindres carrés.
Pour le premier cas-test, on prend un domaine rectangulaire de taille $7m \times 5m$ avec un obstacle carré au centre discrétisé par 9829 cellules (19243 triangles). Sur le bord gauche, on impose un gradient de température radiative T_R^L, ainsi qu'un flux entrant f_x^L et sur les autres bords, des conditions de Neumann homogènes. Sur l'obstacle, on impose des conditions de mur. Au temps initial, on suppose que l'on est à l'équilibre radiatif, c'est-à-dire que la température radiative et la température matière sont égales et que l'énergie est répartie suivant la loi de Planck :

$$T^0 = T_R^0 = 30000K, \quad f_x^0 = f_y^0 = 0,$$
$$T^L = 30000K, \quad T_R^L = 90000K, \quad f_x^L = 0.8, f_y^L = 0.$$

Sur la figure 5.10, on compare les facteurs d'anisotropie dans les groupes de fréquences $2, 3, 5$ et 6 au bout d'un temps $t = 3.17.10^{-8}s$ pour un coefficient $\rho C_v = 10^{-2}J.kg.m^{-3}.K^{-1}$. On voit que le comportement du facteur d'anisotropie est différent suivant les groupes de fréquences considérés. Par exemple, dans les groupes 2 et 6 qui sont quasi-transparents, c'est le phénomène de transport qui prédomine. À l'inverse, dans les groupes 3 et 5 qui sont plus opaques, la diffusion est dominante. Dans ce genre de cas-test, le modèle gris donnerait une approximation moyennée alors que le comportement de la solution est différent suivant le groupe considéré. Ceci illustre l'intérêt de préférer une discrétisation fréquentielle en groupes de fréquences plutôt qu'un modèle totalement intégré en fréquences.
Sur la figure 5.11, on compare les facteurs d'anisotropie dans le groupe 6 et la température matière au bout d'un temps $t = 1.58.10^{-6}s$ pour $\rho C_v = 10^{-2}J.kg.m^{-3}.K^{-1}$ et $\rho C_v = 10^{-3}J.kg.m^{-3}.K^{-1}$. On voit que plus ρC_v est grand, plus l'équilibre met de temps à s'installer.

Pour le deuxième cas-test, on considère la géométrie dite du tophat, soit un tuyau cylindrique coudé éclairé sur le bord gauche. Plus précisément, le domaine de calcul s'étend sur $x \in [0, 7]m$

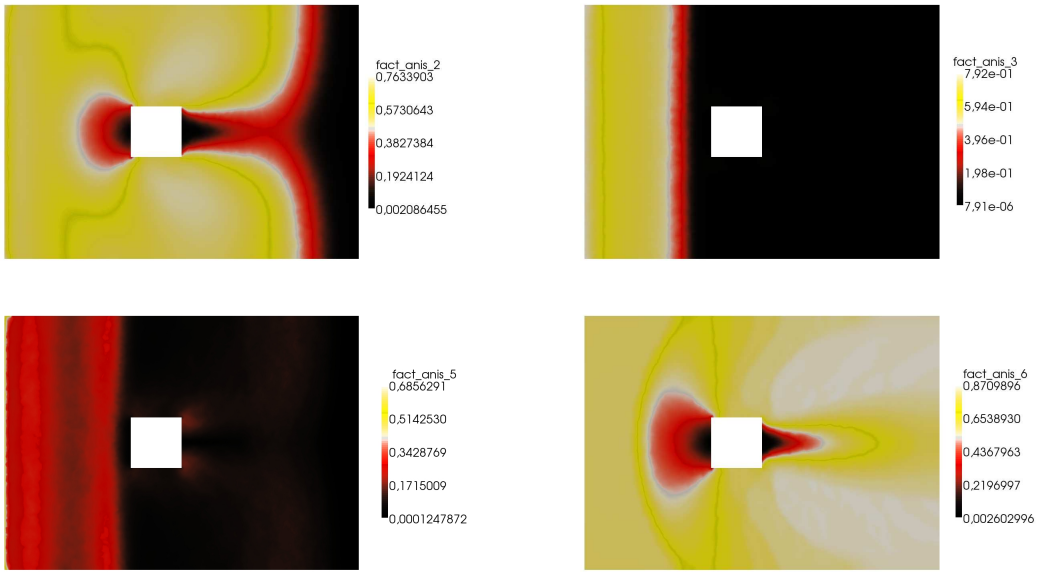


FIGURE 5.10 – Comparaison des facteurs d'anisotropie dans les groupes 2, 3, 5 et 6 pour $\rho C_v = 10^{-2}$ à $t = 3.17.10^{-8} s$.

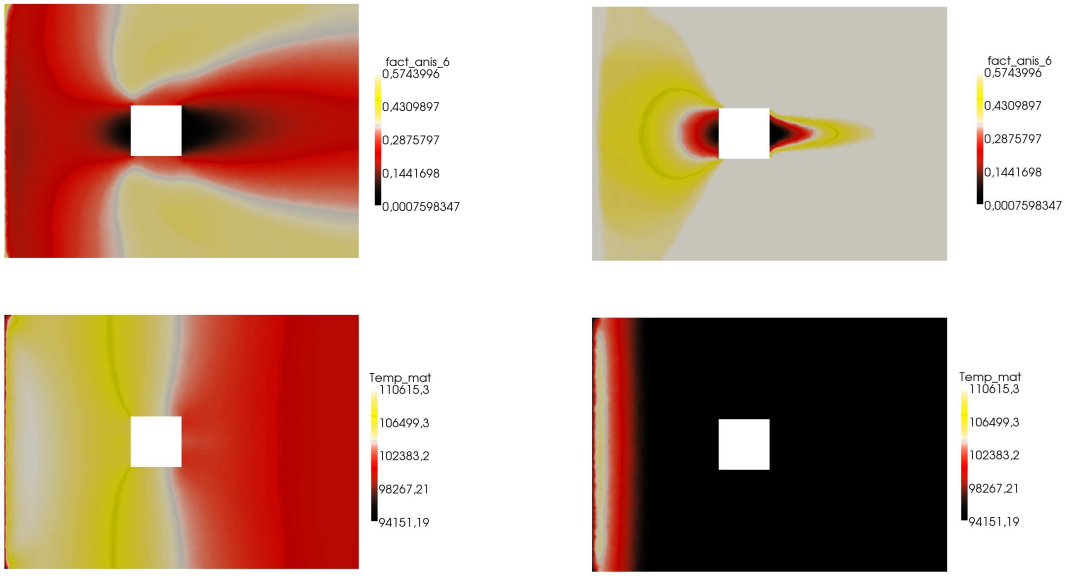


FIGURE 5.11 – Comparaison des facteurs d'anisotropie dans le groupe 6 (haut) et températures matière (bas) pour $\rho C_v = 10^{-2}$ (gauche) et $\rho C_v = 10^{-3}$ (droite) à $t = 1.58.10^{-6} s$.

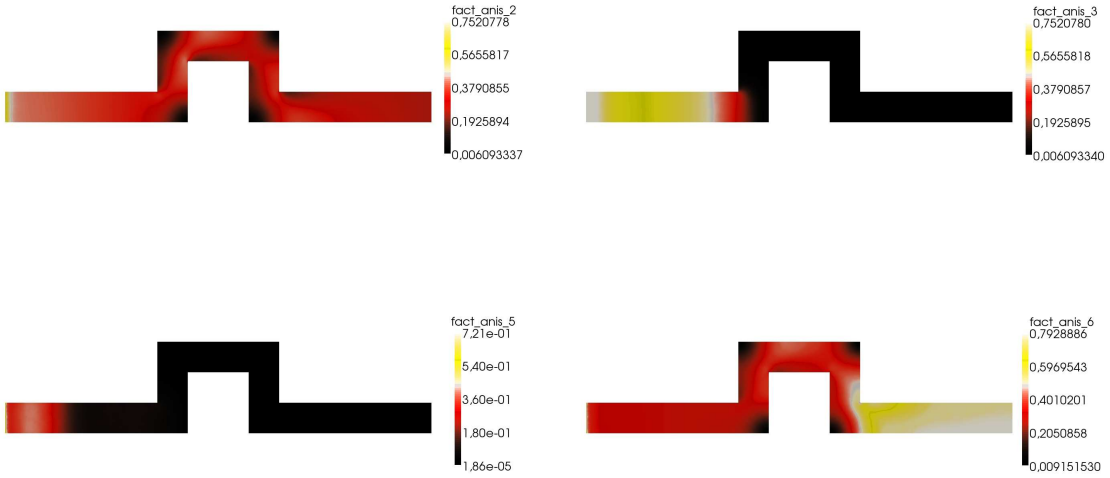


FIGURE 5.12 – Facteurs d'anisotropie dans les groupes 2, 3, 5 et 6 pour $\rho C_v = 10^{-3}$ à $t = 8.14.10^{-8}s$.

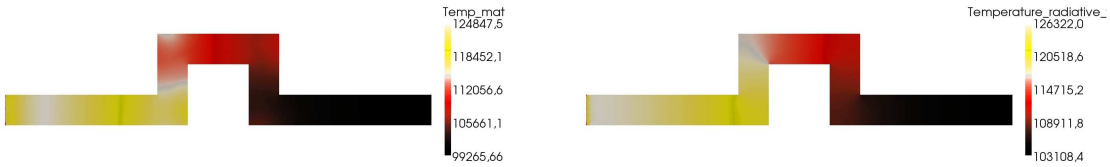


FIGURE 5.13 – Comparaison de la température matière (gauche) et radiative (droite) pour $\rho C_v = 10^{-3}$ à $t = 5.7.10^{-6}s$.

pour un diamètre de $0.5m$. Entre $x = 2.5m$ et $x = 4.5m$, il présente un coude de longueur $1m$ (voir les figures 5.12 et 5.13). Le domaine est discrétisé par 14201 cellules (7429 triangles). Sur le bord gauche, on impose à nouveau un gradient de température radiative T_R^L, ainsi qu'un facteur d'anisotropie entrant f_x^L. Sur les autres bords, on impose des conditions de mur. Au temps initial, on suppose que l'on est à l'équilibre radiatif :

$$T^0 = T_R^0 = 60000K, \quad f_x^0 = f_y^0 = 0, \rho C_v = 10^{-3}$$
$$T^L = 60000K, \quad T_R^L = 100000K, \quad f_x^L = 0.8, f_y^L = 0.$$

Sur la figure 5.12, on compare les facteurs d'anisotropie dans les groupes de fréquences $2, 3, 5$ et 6 au bout d'un temps $t = 8.14.10^{-8}s$ pour un coefficient $\rho C_v = 10^{-3} J.kg.m^{-3}.K^{-1}$. On constate de nouveau que le comportement est différent suivant les groupes de fréquences considérés. Le flux avance très rapidement dans les groupes 2 et 6 qui sont assez transparents et plus lentement dans les groupes 3 et 5 qui sont plus opaques. On observe de plus que les réflexions imposées par les conditions de mur freinent le flux par rapport au cas-test précédent.
Sur la figure 5.13, on compare la température matière et la température radiative au bout d'un temps $t = 5.7.10^{-6}s$. Logiquement, on constate que la température matière rattrape la température radiative et donc que le régime de diffusion global est en train de s'installer.

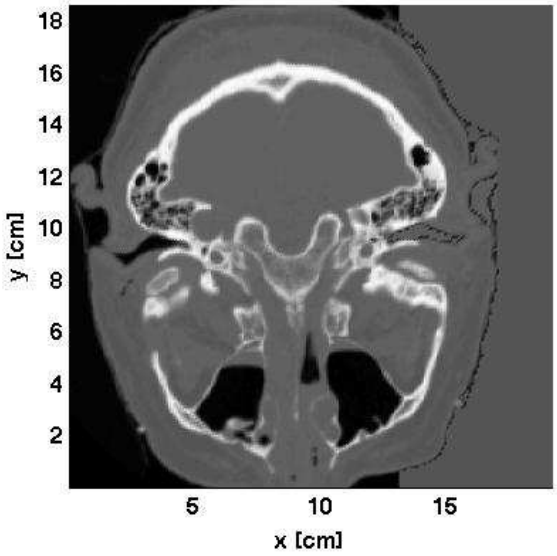


FIGURE 5.14 – Coupe provenant d'un scanner du crâne avec la partie plastique attachée sur la droite.

Ce cas-test illustre encore l'importance du découpage en groupes de fréquences, car l'opacité est vraiment différente dans les différents groupes. En effet, le modèle gris, totalement intégré en fréquences, donnerait une moyenne globale qui serait une mauvaise approximation de la solution.

5.2.3 Calcul de Dose en radiothérapie

On dispose d'une coupe en dimension deux d'un scanner humain. On teste le schéma sur un cas issu de la littérature qui consiste à irradier un cerveau avec des électrons [83]. Le scanner, visible sur la figure 5.14, représente une coupe horizontale du crâne du patient. Dans chaque voxel, la matière est décrite par sa nuance de gris de Hounsfield $\mathcal{G}(x, y)$. Cette nuance de gris peut être traduite en terme de paramètres physiques de la manière suivante :

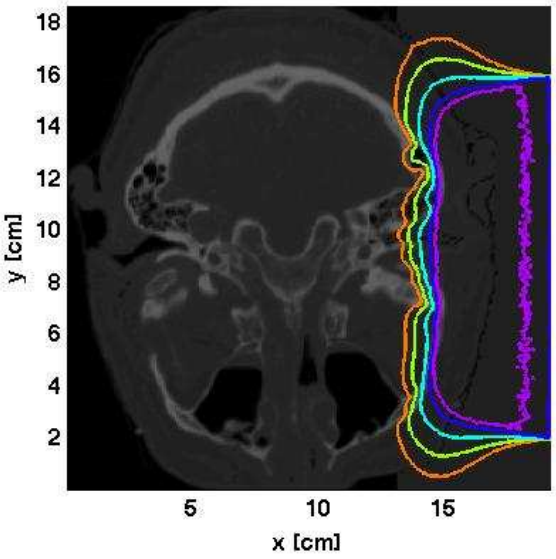
$$\rho(x, y) = \left(\frac{\mathcal{G}(x, y)}{1000} + 1 \right) \rho_W,$$

où ρ_W est la densité de l'eau. Les régions sombres ont une densité faible, et les zones claires ont une forte densité. Les parties noires sont composées d'air avec une densité d'environ 10^{-3} kg/m^3, les parties grises sont formées d'eau avec une densité d'environ 1 kg/m^3, et les parties grises claires sont formées d'os de densité environ égale à 2 kg/m^3. Sur la droite de la coupe, la tête est attachée à un réceptacle en plastique. En pratique, il est utilisé pour modeler le rayon d'électrons.

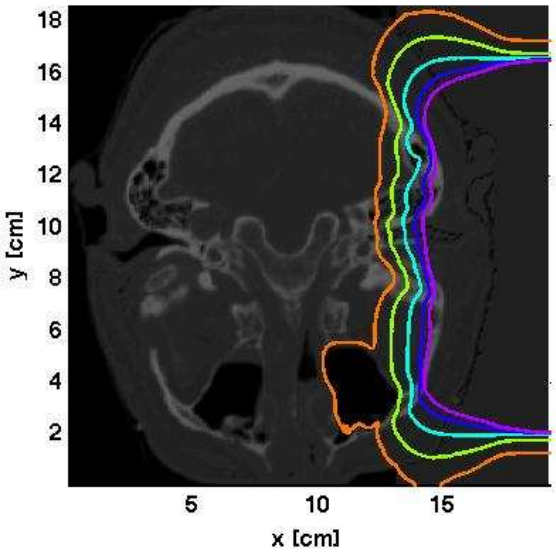
La tête est alors irradiée par la droite avec un rayon d'électrons de 12 MeV et de 14 cm de large. Pour modéliser ce rayon, on a représenté l'énergie par une Gaussienne très étroite et une impulsion δ angulaire. D'autres spectres d'énergie sont possibles. Avec ces données, on peut calculer les moments angulaires $\psi^{(0)}$ et $\psi^{(1)}$ utilisés dans le code de calcul. Pour comparer avec le schéma HLL_{cvp} en dimension deux, on utilise le code Monte Carlo PENELOPE [97]. Ce code a été validé sur de nombreux cas expérimentaux. PENELOPE est ici mis en place dans une configuration pseudo-2D avec un large rayon perpendiculaire au plan considéré dans lequel se propage les électrons. En particulier, on a utilisé le code penEasy2009 dans une géométrie voxélisée avec des paramètres de simulation standards. La source d'énergie initiale est limitée à 15 MeV. Les électrons et positrons sont supposés être absorbés quand leur énergie devient inférieure à 150 KeV. Pour les photons, cette quantité est limitée à 15 KeV. L'angle critique θ_c et l'énergie

critique W_c qui différencient les interactions fortes des interactions faibles sont définis comme suit : pour les collisions inélastiques, on pose $W_c = 150$ keV et pour les émissions de bremsstrahlung on pose $W_c = 15$ keV. La moyenne angulaire de déviation entre deux interactions fortes inélastiques consécutives étant de $C_1 = 0.1$ et la moyenne maximale de perte d'énergie entre deux interactions forte inélastiques étant de $C_2 = 0.1$, elles déterminent de manière unique θ_c. Le pas maximum autorisé pour les électrons et positrons est infini : $DSMAX = 10^{35}$ cm. Aucune méthode de réduction de la variance n'a été employée. On insiste sur le fait que le modèle physique que l'on considère est indépendant du modèle physique de *penEasy*.

Sur la figure 5.15, on compare les courbes d'isodose obtenue par les deux méthodes. Pour une meilleure visibilité des résultats, on affiche les lignes de niveau 10%, 25%, 50%, 70% et 80%. Le bruit statistique de la solution Monte Carlo (MC) se voit sur la courbe de niveau 80%. On remarque que la solution du modèle M_1 est légèrement plus diffusée que la solution MC, c'est-à-dire que le rayon est plus large et les courbes de niveau les plus faibles sont plus éloignées. La courbe orange entre même dans la zone de vide en bas du scanner. Cependant, il s'agit seulement de la courbe à 10%, ce qui n'est pas significatif de la radiation globale. La différence majeure entre les deux modèles vient de la zone d'accumulation. On observe cette différence notamment sur la courbe à 80% près de la frontière. Dans l'article [44], on peut lire que ce défaut vient du modèle M_1, et non du schéma numérique qui le discrétise. Sur la figure 5.16, on représente la différence entre les deux solutions. En radiothérapie, les différentes méthodes sont comparées en erreur relative et également en termes de *Distance-To-Agreement* (DTA), c'est-à-dire la distance entre le point où la dose est mesurée et le point le plus proche où la dose calculée est la même. Globalement, environ 63% des voxels sont entre 4% ou 4mm DTA. Ceci n'est pas encore suffisant pour être applicable au cas pratique. Les raisons de ces écarts sont essentiellement dues au modèle physique trop simplifié et aux défauts du modèle M_1. Cependant, cela peut servir à valider le concept. La méthode HLL_{cvp} donne une bonne approximation de la dose. Elle ne présente pas de bruit statistique, ce qui est un avantage des méthodes déterministes. De plus, pour les simulations précédentes, le temps de calcul avec le code HLL_{cvp} est de l'ordre de 50 minutes avec 12 cœurs, alors que PENELOPE tourne pendant presque une journée.



(a) Solution Monte Carlo.



(b) Solution M_1 .

FIGURE 5.15 – Courbes isodoses d'un rayon d'électrons dans la colonne vertébrale, normalisées par D_{max} et courbes de niveau 10% orange, 25% jaune, 50% bleu clair, 70% bleu foncé, 80% violet.

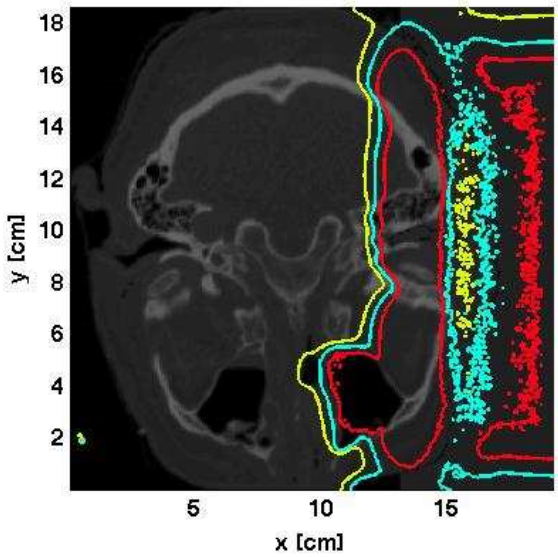


FIGURE 5.16 – Courbes de niveau de la différence entre les doses du M_1 et de PENELOPE, normalisées par la dose maximale (2% de différence en jaune, 5% de différence en bleu, 10% de différence en rouge).

A

Appendice

A.1 Appendice 1 : schéma GRP espace-temps pour l'équation de transport en dimension 2

A.1.1 Étape d'évolution et de projection dans les cas *deux cibles* explicite, *deux cibles* implicite et *une cible* explicite

Étape d'évolution :
Cas *deux cibles* (explicite) :

- la solution $u_i^n(x,y)$ est connue dans la cellule T_i au temps t^n ainsi que la solution $v_{j3}^{n+\frac{1}{2}}(t,\omega)$ sur l'interface $I_{j3}^{n+\frac{1}{2}}$.
- On définit θ_1 comme l'intersection entre la droite caractéristique issue du segment s_{j3}^n et du segment s_{j1}^{n+1}. On définit également θ_2 comme l'intersection de la droite caractéristique issue de s_{j3}^n et du segment s_{j2}^{n+1}. Puis, on note M_1 la projection de θ_1 sur s_{j1}^n et M_2 la projection de θ_2 sur s_{j2}^n (voir figure A.1).
- La solution exacte sur la cellule T_i au temps t^{n+1} est donnée par :

$$u_i^h(x,y,t^n+\Delta t) = \begin{cases} u_i^n(x_1(x,y),y_1(x,y)), & \text{si } (x,y)\in\mathcal{F}_i^2 \\ v_{j3}^{n+\frac{1}{2}}(t_1(x,y),\omega_1(x,y)), & \text{si } (x,y)\in\mathcal{F}_{j3}^2 \end{cases}$$

où t_1, x_1 et x_2 sont les changements de variables suivants :

$$\begin{aligned} x_1(x,y) &= x - a\Omega_x\Delta t \\ y_1(x,y) &= y - a\Omega_y\Delta t \\ \omega_1(x,y) &= \frac{\Omega_x(y-y_{p_{j3}}-a\Omega_y(t^n+\Delta t))-\Omega_y(x-x_{pj3}-a\Omega_x(t^n+\Delta t))}{\Omega_x(y_{q_{j3}}-y_{p_{j3}})-\Omega_y(x_{q_{j3}}-x_{p_{j3}})} \\ t_1(x,y) &= \begin{cases} \frac{x_{p_{j3}}+\omega_1(x,y)(x_{q_{j3}}-x_{p_{j3}})-x}{a\Omega_x}+t^n+\Delta t, & \text{si } \Omega_x\neq 0 \\ \frac{y_{p_{j3}}+\omega_1(x,y)(y_{q_{j3}}-y_{p_{j3}})-y}{a\Omega_y}+t^n+\Delta t, & \text{sinon} \end{cases} \end{aligned}$$

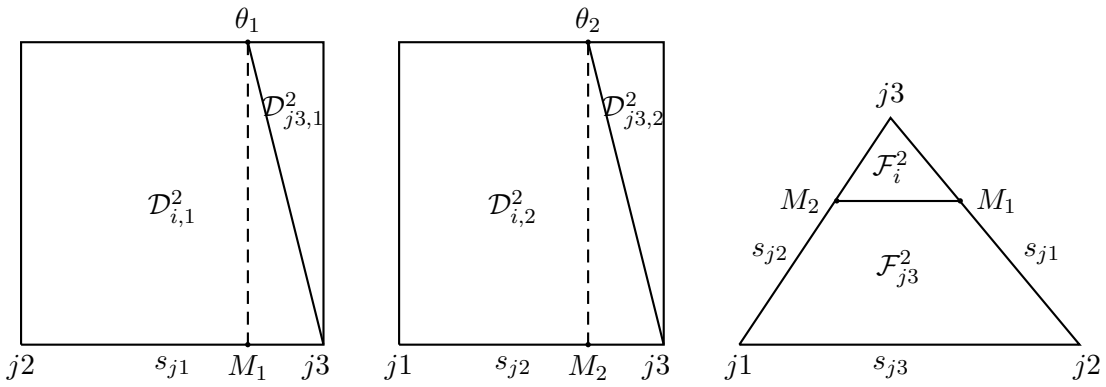


FIGURE A.1 – Solutions entrantes sur l'interface (gauche) et sur la cellule (droite) pour les cas *deux cibles* explicite

- La solution exacte sur l'interface $I^{n+\frac{1}{2}}_{j3}$ est donnée par :

$$v_h^{n+\frac{1}{2}}(t,\omega_{j1}) = \begin{cases} u_i^n(x_2(t,\omega), y_2(t,\omega)), & \text{si } (t,\omega) \in \mathcal{D}^2_{i,1} \\ v_{j3}^{n+\frac{1}{2}}(t_2(t,\omega), \omega_2(t,\omega)), & \text{si } (t,\omega) \in \mathcal{D}^2_{j3,1} \end{cases}$$

avec :

$$\begin{aligned} x_2(t,\omega) &= x_{p_{j1}} + \omega(x_{q_{j1}} - x_{p_{j1}}) - a\Omega_x(t-t^n) \\ y_2(t,\omega) &= y_{p_{j1}} + \omega(y_{q_{j1}} - y_{p_{j1}}) - a\Omega_y(t-t^n) \\ t_2(t,\omega) &= \frac{c_2^1(y_{p_{j3}} - y_{q_{j3}}) + c_2^2(x_{q_{j3}} - x_{p_{j3}})}{d_2} \\ \omega_2(t,\omega) &= \frac{ac_2^2\Omega_x - ac_2^1\Omega_y}{d_2} \\ c_2^1 &= x_{p_{j1}} - x_{p_{j3}} + \omega(x_{q_{j1}} - x_{p_{j1}}) - a\Omega_x t \\ c_2^2 &= y_{p_{j1}} - y_{p_{j3}} + \omega(y_{q_{j1}} - y_{p_{j1}}) - a\Omega_y t \\ d_2 &= a\Omega_x(y_{q_{j3}} - y_{p_{j3}}) - a\Omega_y(x_{q_{j3}} - x_{p_{j3}}) \end{aligned}$$

- La solution exacte sur l'interface $I^{n+\frac{1}{2}}_{j2}$ est donnée par :

$$v_h^{n+\frac{1}{2}}(t,\omega_{j2}) = \begin{cases} u_i^n(x_3(t,\omega), y_3(t,\omega)), & \text{si } (t,\omega) \in \mathcal{D}^2_{i,2} \\ v_{j3}^{n+\frac{1}{2}}(t_3(t,\omega), \omega_3(t,\omega)), & \text{si } (t,\omega) \in \mathcal{D}^2_{j3,2} \end{cases}$$

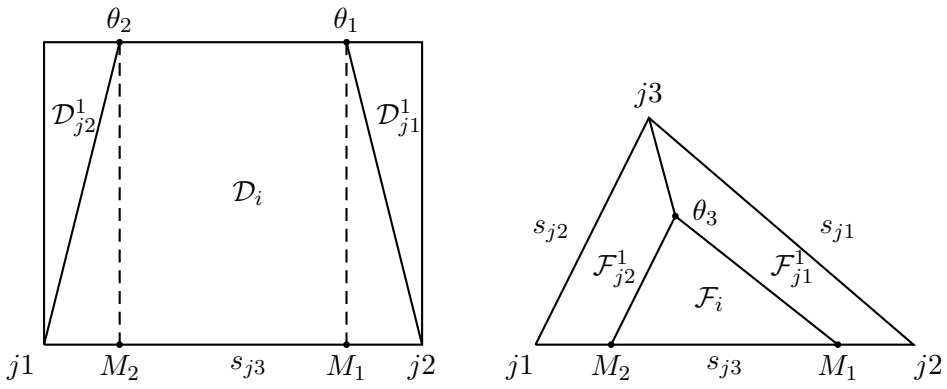


FIGURE A.2 – Solutions entrantes sur l'interface (gauche) et sur la cellule (droite) pour les cas *une cible* explicite

avec :

$$
\begin{aligned}
x_3(t,\omega) &= x_{p_{j2}} + \omega(x_{q_{j2}} - x_{p_{j2}}) - a\Omega_x(t - t^n) \\
y_3(t,\omega) &= y_{p_{j2}} + \omega(y_{q_{j2}} - y_{p_{j2}}) - a\Omega_y(t - t^n) \\
t_3(t,\omega) &= \frac{c_3^1(y_{p_{j3}} - y_{q_{j3}}) + c_3^2(x_{q_{j3}} - x_{p_{j3}})}{d_3} \\
\omega_3(t,\omega) &= \frac{ac_3^2\Omega_x - ac_3^1\Omega_y}{d_3} \\
c_3^1 &= x_{p_{j2}} - x_{p_{j3}} + \omega(x_{q_{j2}} - x_{p_{j2}}) - a\Omega_x t \\
c_3^2 &= y_{p_{j2}} - y_{p_{j3}} + \omega(y_{q_{j2}} - y_{p_{j2}}) - a\Omega_y t \\
d_3 &= a\Omega_x(y_{q_{j3}} - y_{p_{j3}}) - a\Omega_y(x_{q_{j3}} - x_{p_{j3}})
\end{aligned}
$$

Cas *une cible* (explicite) :

- la solution $u_i^n(x,y)$ est connue dans la cellule T_i au temps t^n ainsi que les solutions $v_{j1}^{n+\frac{1}{2}}(t,\omega)$ et $v_{j2}^{n+\frac{1}{2}}(t,\omega)$ sur les interfaces $I_{j1}^{n+\frac{1}{2}}$ et $I_{j2}^{n+\frac{1}{2}}$.
- On définit θ_1 comme l'intersection entre la droite caractéristique issue du segment s_{j1}^n et du segment s_{j3}^{n+1} ($i.e.$ tel que $d(s_{j3}^n, s_{j1}^n) = a\Delta t$). On définit également θ_2 comme l'intersection de la droite caractéristique issue de s_{j2}^n et du segment s_{j3}^{n+1} ($i.e.$ tel que $d(s_{j3}^n, s_{j2}^n) = a\Delta t$). Puis, on note M_1 la projection de θ_1 sur s_{j3}^n et M_2 la projection de θ_2 sur s_{j3}^n (voir figure A.2).
- La solution exacte sur la cellule T_i au temps t^{n+1} est donnée par :

$$
u_i^h(x,y,t^n+\Delta t) = \begin{cases} u_i^n(x_1(x,y),y_1(x,y)), & \text{si } (x,y) \in \mathcal{F}_i \\ v_{j1}^{n+\frac{1}{2}}(t_1(x,y),\omega_1(x,y)), & \text{si } (x,y) \in \mathcal{F}_{j1}^1 \\ v_{j2}^{n+\frac{1}{2}}(t_2(x,y),\omega_2(x,y)), & \text{si } (x,y) \in \mathcal{F}_{j2}^1 \end{cases}
$$

avec :

$$
\begin{aligned}
x_1(x,y) &= x - a\Omega_x \Delta t \\
y_1(x,y) &= y - a\Omega_y \Delta t \\
\omega_1(x,y) &= \frac{\Omega_x(y - y_{p_{j1}} - a\Omega_y(t^n + \Delta t)) - \Omega_y(x - x_{pj1} - a\Omega_x(t^n + \Delta t))}{\Omega_x(y_{q_{j1}} - y_{p_{j1}}) - \Omega_y(x_{q_{j1}} - x_{p_{j1}})} \\
t_1(x,y) &= \begin{cases} \frac{x_{p_{j1}} + \omega_1(x,y)(x_{q_{j1}} - x_{p_{j1}}) - x}{a\Omega_x} + t^n + \Delta t, & \text{si } \Omega_x \neq 0 \\ \frac{y_{p_{j1}} + \omega_1(x,y)(y_{q_{j1}} - y_{p_{j1}}) - y}{a\Omega_y} + t^n + \Delta t, & \text{sinon} \end{cases} \\
\omega_2(x,y) &= \frac{\Omega_x(y - y_{p_{j2}} - a\Omega_y(t^n + \Delta t)) - \Omega_y(x - x_{pj2} - a\Omega_x(t^n + \Delta t))}{\Omega_x(y_{q_{j2}} - y_{p_{j2}}) - \Omega_y(x_{q_{j2}} - x_{p_{j2}})} \\
t_2(x,y) &= \begin{cases} \frac{x_{p_{j2}} + \omega_2(x,y)(x_{q_{j2}} - x_{p_{j2}}) - x}{a\Omega_x} + t^n + \Delta t, & \text{si } \Omega_x \neq 0 \\ \frac{y_{p_{j2}} + \omega_2(x,y)(y_{q_{j2}} - y_{p_{j2}}) - y}{a\Omega_y} + t^n + \Delta t, & \text{sinon} \end{cases}
\end{aligned}
$$

- La solution exacte sur l'interface $I_{j3}^{n+\frac{1}{2}}$ est donnée par :

$$
v_h^{n+\frac{1}{2}}(t,\omega_{j3}) = \begin{cases} u_i^n(x_2(t,\omega), y_2(t,\omega)), & \text{si } (t,\omega) \in \mathcal{D}_i \\ v_{j1}^{n+\frac{1}{2}}(t_3(t,\omega), \omega_3(t,\omega)), & \text{si } (t,\omega) \in \mathcal{D}_{j1}^1 \\ v_{j2}^{n+\frac{1}{2}}(t_4(t,\omega), \omega_4(t,\omega)), & \text{si } (t,\omega) \in \mathcal{D}_{j2}^1 \end{cases}
$$

avec :

$$
\begin{aligned}
x_2(t,\omega) &= x_{p_{j3}} + \omega(x_{q_{j3}} - x_{p_{j3}}) - a\Omega_x(t - t^n) \\
y_2(t,\omega) &= y_{p_{j3}} + \omega(y_{q_{j3}} - y_{p_{j3}}) - a\Omega_y(t - t^n) \\
t_3(t,\omega) &= \frac{(c_3^1(y_{p_{j1}} - y_{q_{j1}}) + c_3^2(x_{q_{j1}} - x_{p_{j1}})}{d_3} \\
\omega_3(t,\omega) &= \frac{ac_3^2\Omega_x - ac_3^1\Omega_y}{d_3} \\
c_3^1 &= x_{p_{j3}} - x_{p_{j1}} + \omega(x_{q_{j3}} - x_{p_{j3}}) - a\Omega_x t \\
c_3^2 &= y_{p_{j3}} - y_{p_{j1}} + \omega(y_{q_{j3}} - y_{p_{j3}}) - a\Omega_y t \\
d_3 &= a\Omega_x(y_{q_{j1}} - y_{p_{j1}}) - a\Omega_y(x_{q_{j1}} - x_{p_{j1}}) \\
t_4(t,\omega) &= \frac{(c_4^1(y_{p_{j2}} - y_{q_{j2}}) + c_4^2(x_{q_{j2}} - x_{p_{j2}})}{d_4} \\
\omega_4(t,\omega) &= \frac{ac_4^2\Omega_x - ac_4^1\Omega_y}{d_4} \\
c_4^1 &= x_{p_{j3}} - x_{p_{j2}} + \omega(x_{q_{j3}} - x_{p_{j3}}) - a\Omega_x t \\
c_4^2 &= y_{p_{j3}} - y_{p_{j2}} + \omega(y_{q_{j3}} - y_{p_{j3}}) - a\Omega_y t \\
d_4 &= a\Omega_x(y_{q_{j2}} - y_{p_{j2}}) - a\Omega_y(x_{q_{j2}} - x_{p_{j2}})
\end{aligned}
$$

Cas deux cibles (implicite) :

- la solution $u_i^n(x,y)$ est connue dans la cellule T_i au temps t^n ainsi que la solution $v_{j3}^{n+\frac{1}{2}}(t,\omega)$ sur l'interface $I_{j3}^{n+\frac{1}{2}}$.
- On définit θ_3 comme l'intersection entre la droite caractéristique issue du segment s_{j3}^n et du segment $[t^n, t^{n+1}]$ pour $(x,y) = (x_{j3}, y_{j3})$ (voir figure A.3).

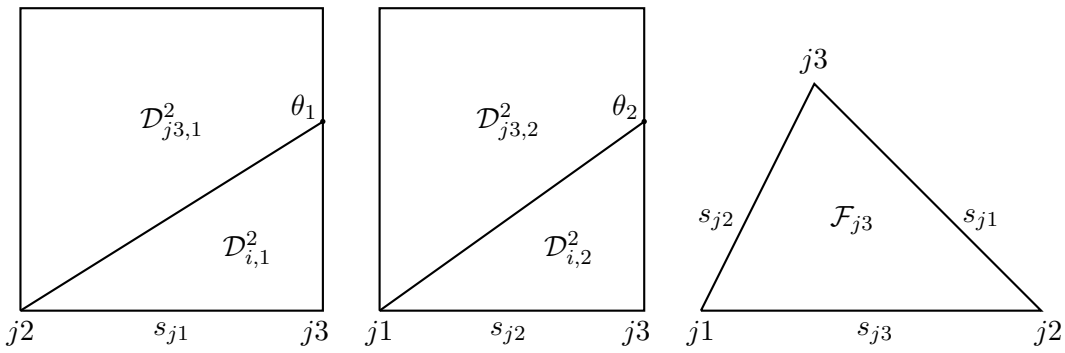


FIGURE A.3 – Solutions entrantes sur l'interface (gauche) et sur la cellule (droite) pour les cas *deux cibles* implicite

- La solution exacte sur la cellule T_i au temps t^{n+1} est donnée par :

$$u_i^h(x,y,t^n+\Delta t) = v_{j3}^{n+\frac{1}{2}}(t_1(x,y),\omega_1(x,y)) \quad \text{si } (x,y)\in T_i,$$

avec :

$$\omega_1(x,y) = \frac{\Omega_x(y-y_{p_{j3}}-a\Omega_y(t^n+\Delta t))-\Omega_y(x-x_{pj3}-a\Omega_x(t^n+\Delta t))}{\Omega_x(y_{q_{j3}}-y_{p_{j3}})-\Omega_y(x_{q_{j3}}-x_{p_{j3}})}$$

$$t_1(x,y) = \begin{cases} \frac{x_{p_{j3}}+\omega_1(x,y)(x_{q_{j3}}-x_{p_{j3}})-x}{a\Omega_x}+t^n+\Delta t, & \text{si } \Omega_x\neq 0 \\ \frac{y_{p_{j3}}+\omega_4(x,y)(y_{q_{j3}}-y_{p_{j3}})-y}{a\Omega_y}+t^n+\Delta t, & \text{sinon} \end{cases}$$

- La solution exacte sur l'interface $I_{j1}^{n+\frac{1}{2}}$ est donnée par :

$$v_h^{n+\frac{1}{2}}(t,\omega_{j1}) = \begin{cases} u_i^n(x_1(t,\omega,y_1(t,\omega)), & \text{si } (t,\omega)\in\mathcal{D}^2_{i,1} \\ v_{j3}^{n+\frac{1}{2}}(t_2(t,\omega),\omega_2(t,\omega)), & \text{si } (t,\omega)\in\mathcal{D}^2_{j3,1} \end{cases}$$

avec :

$$\begin{aligned} x_1(t,\omega) &= x_{p_{j1}}+\omega(x_{q_{j1}}-x_{p_{j1}})-a\Omega_x(t-t^n) \\ y_1(t,\omega) &= y_{p_{j1}}+\omega(y_{q_{j1}}-y_{p_{j1}})-a\Omega_y(t-t^n) \\ t_2(t,\omega) &= \frac{(c_2^1(y_{p_{j3}}-y_{q_{j3}})+c_2^2(x_{q_{j3}}-x_{p_{j3}})}{d_2} \\ \omega_2(t,\omega) &= \frac{ac_2^2\Omega_x-ac_2^1\Omega_y}{d_2} \\ c_2^1 &= x_{p_{j1}}-x_{p_{j3}}+\omega(x_{q_{j1}}-x_{p_{j1}})-a\Omega_x t \\ c_2^2 &= y_{p_{j1}}-y_{p_{j3}}+\omega(y_{q_{j1}}-y_{p_{j1}})-a\Omega_y t \\ d_2 &= a\Omega_x(y_{q_{j3}}-y_{p_{j3}})-a\Omega_y(x_{q_{j3}}-x_{p_{j3}}) \end{aligned}$$

- La solution exacte sur l'interface $I_{j2}^{n+\frac{1}{2}}$ est donnée par :

$$v_h^{n+\frac{1}{2}}(t,\omega_{j2}) = \begin{cases} u_i^n(x_2(t,\omega), y_2(t,\omega)), & \text{si } (t,\omega) \in \mathcal{D}_{i,2}^2 \\ v_{j3}^{n+\frac{1}{2}}(t_3(t,\omega), \omega_3(t,\omega)), & \text{si } (t,\omega) \in \mathcal{D}_{j3,2}^2 \end{cases}$$

with :

$$\begin{aligned} x_2(t,\omega) &= x_{p_{j2}} + \omega(x_{q_{j2}} - x_{p_{j2}}) - a\Omega_x(t-t^n) \\ y_2(t,\omega) &= y_{p_{j2}} + \omega(y_{q_{j2}} - y_{p_{j2}}) - a\Omega_y(t-t^n) \\ t_3(t,\omega) &= \frac{(c_3^1(y_{p_{j3}} - y_{q_{j3}}) + c_3^2(x_{q_{j3}} - x_{p_{j3}})}{d_3} \\ \omega_3(t,\omega) &= \frac{ac_3^2\Omega_x - ac_3^1\Omega_y}{d_3} \\ c_3^1 &= x_{p_{j2}} - x_{p_{j3}} + \omega(x_{q_{j2}} - x_{p_{j2}}) - a\Omega_x t \\ c_3^2 &= y_{p_{j2}} - y_{p_{j3}} + \omega(y_{q_{j2}} - y_{p_{j2}}) - a\Omega_y t \\ d_3 &= a\Omega_x(y_{q_{j3}} - y_{p_{j3}}) - a\Omega_y(x_{q_{j3}} - x_{p_{j3}}) \end{aligned}$$

Étape de projection : Les solutions exactes sont maintenant projetées sur l'espace d'approximation $\mathbb{P}^k$: dans la cellule T_i au temps t^{n+1}, la solution exacte est projetée sur $\mathbb{P}_C^k$, et sur les interfaces $I_j^{n+\frac{1}{2}}$, la solution est projetée sur $\mathbb{P}_I^k$. Pour le cas *une cible* explicite, cette projection se traduit par deux problèmes de minimisation :

$$\|u_i^{n+1} - u_i^h(t^n + \Delta t)\|_i = \inf_{p \in \mathbb{P}_C^k} \|p - u_i^h(t^n + \Delta t)\|_i, \tag{A.1}$$

$$\|v_{j3}^{n+\frac{1}{2}} - v_h^{n+\frac{1}{2}}(\omega_{j3})\|^n = \inf_{q \in \mathbb{P}_I^k} \|q - v_h^{n+\frac{1}{2}}(\omega_{j3})\|^n, \tag{A.2}$$

pour le cas *deux cibles* explicite, la projection se traduit par trois problèmes de minimisation :

$$\|u_i^{n+1} - u_i^h(t^n + \Delta t)\|_i = \inf_{p \in \mathbb{P}_C^k} \|p - u_i^h(t^n + \Delta t)\|_i, \tag{A.3}$$

$$\|v_{j1}^{n+\frac{1}{2}} - v_h^{n+\frac{1}{2}}(\omega_{j1})\|^n = \inf_{q \in \mathbb{P}_I^k} \|q - v_h^{n+\frac{1}{2}}(\omega_{j1})\|^n, \tag{A.4}$$

$$\|v_{j2}^{n+\frac{1}{2}} - v_h^{n+\frac{1}{2}}(\omega_{j2})\|^n = \inf_{q \in \mathbb{P}_I^k} \|q - v_h^{n+\frac{1}{2}}(\omega_{j2})\|^n., \tag{A.5}$$

et pour le cas *deux cibles* implicite, la projection se traduit également par trois problèmes de minimisation :

$$\|u_i^{n+1} - u_i^h(t^n + \Delta t)\|_i = \inf_{p \in \mathbb{P}_C^k} \|p - u_i^h(t^{n+1})\|_i, \tag{A.6}$$

$$\|v_{j1}^{n+\frac{1}{2}} - v_h^{n+\frac{1}{2}}(\omega_{j1})\|^n = \inf_{q \in \mathbb{P}_I^k} \|q - v_h^{n+\frac{1}{2}}(\omega_{j1})\|^n, \tag{A.7}$$

$$\|v_{j2}^{n+\frac{1}{2}} - v_h^{n+\frac{1}{2}}(\omega_{j2})\|^n = \inf_{q \in \mathbb{P}_I^k} \|q - v_h^{n+\frac{1}{2}}(\omega_{j2})\|^n. \tag{A.8}$$

Grâce aux conditions de Petrov-Galerkin, les deux problèmes de minimisation (A.1)-(A.2) peuvent se réécrire de la manière suivante :

$$< u_i^{n+1} - u_i^h(t^n + \Delta t), \left(\frac{x - x_i}{V_i}\right)^l \left(\frac{y - y_i}{V_i}\right)^m >_i = 0 \quad \forall l + m = 0, 1, \ldots, k, \tag{A.9}$$

$$< v_{j3}^{n+\frac{1}{2}} - v_h^{n+\frac{1}{2}}(\omega_{j3}), \left(\frac{t - t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m >^n = 0 \quad \forall l + m = 0, 1, \ldots, k. \tag{A.10}$$

et (A.3)-(A.4)-(A.5) ou (A.6)-(A.7)-(A.8) par :

$$< u_i^{n+1} - u_i^h(t^n + \Delta t), \left(\frac{x - x_i}{V_i}\right)^l \left(\frac{y - y_i}{V_i}\right)^m >_i = 0 \quad \forall l + m = 0, 1, \ldots, k, \tag{A.11}$$

$$< v_{j1}^{n+\frac{1}{2}} - v_h^{n+\frac{1}{2}}(\omega_{j1}), \left(\frac{t - t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m >^n = 0 \quad \forall l + m = 0, 1, \ldots, k. \tag{A.12}$$

$$< v_{j2}^{n+\frac{1}{2}} - v_h^{n+\frac{1}{2}}(\omega_{j2}), \left(\frac{t - t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m >^n = 0 \quad \forall l + m = 0, 1, \ldots, k. \tag{A.13}$$

Puis, en utilisant leurs décompositions dans leurs bases de polynômes respectives :

$$\begin{aligned}
u_i^{n+1}(x, y) &= \sum_{p+q=0}^{k} \alpha_i^{n+1,p,q} \left(\frac{x - x_i}{V_i}\right)^p \left(\frac{y - y_i}{V_i}\right)^q, \\
v_{j1}^{n+\frac{1}{2}}(t, \omega) &= \sum_{p+q=0}^{k} \beta_{j1}^{n+\frac{1}{2},p,q} \left(\frac{t - t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega^q, \\
v_{j2}^{n+\frac{1}{2}}(t, \omega) &= \sum_{p+q=0}^{k} \beta_{j2}^{n+\frac{1}{2},p,q} \left(\frac{t - t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega^q, \\
v_{j3}^{n+\frac{1}{2}}(t, \omega) &= \sum_{p+q=0}^{k} \beta_{j3}^{n+\frac{1}{2},p,q} \left(\frac{t - t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega^q,
\end{aligned}$$

les problèmes de minimisation (A.9)-(A.10) se ramènent à la résolution des deux systèmes linéaires de taille $\frac{(k+1)(k+2)}{2}$ suivants :

$$\begin{aligned}
\sum_{p+q=0}^{k} \alpha_i^{n+1,p,q} &< \left(\frac{x - x_i}{V_i}\right)^p \left(\frac{y - y_i}{V_i}\right)^q, \left(\frac{x - x_i}{V_i}\right)^l \left(\frac{y - y_i}{V_i}\right)^m >_i \\
&= < u_i^h(t^n + \Delta t), \left(\frac{x - x_i}{V_i}\right)^l \left(\frac{y - y_i}{V_i}\right)^m >_i \quad \forall l + m = 0, \ldots, k, \\
\sum_{p+q=0}^{k} \beta_{j3}^{n+\frac{1}{2},p,q} &< \left(\frac{t - t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega^q, \left(\frac{t - t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m >^n \\
&= < v_h^{n+\frac{1}{2}}(\omega_{j3}), \left(\frac{t - t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m >^n \quad \forall l + m = 0, \ldots, k.
\end{aligned}$$

et (A.11)-(A.12)-(A.13) se ramènent à la résolution des trois systèmes linéaires de taille $\frac{(k+1)(k+2)}{2}$ suivants :

$$\begin{aligned}
\sum_{p+q=0}^{k} \alpha_i^{n+1,p,q} &< \left(\frac{x-x_i}{V_i}\right)^p \left(\frac{y-y_i}{V_i}\right)^q, \left(\frac{x-x_i}{V_i}\right)^l \left(\frac{y-y_i}{V_i}\right)^m >_i \\
&=< u_i^h(t^n+\Delta t), \left(\frac{x-x_i}{V_i}\right)^l \left(\frac{y-y_i}{V_i}\right)^m >_i \quad \forall l+m=0,\dots,k, \\
\sum_{p+q=0}^{k} \beta_{j1}^{n+\frac{1}{2},p,q} &< \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega^q, \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m >^n \\
&=< v_h^{n+\frac{1}{2}}(\omega_{j1}), \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m >^n \quad \forall l+m=0,\dots,k, \\
\sum_{p+q=0}^{k} \beta_{j2}^{n+\frac{1}{2},p,q} &< \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega^q, \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m >^n \\
&=< v_h^{n+\frac{1}{2}}(\omega_{j2}), \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m >^n \quad \forall l+m=0,\dots,k.
\end{aligned}$$

Les matrices sont identiques pour les cas *deux cibles* implicite et pour le cas *deux cibles* explicite.
Pour les seconds membres, on a tout d'abord pour le cas *une cible* explicite :

$$\begin{aligned}
< u_i^h(t^n+\Delta t), \left(\frac{x-x_i}{V_i}\right)^l \left(\frac{y-y_i}{V_i}\right)^m >_i &= \sum_{p+q=0}^{k} \alpha_i^{n,p,q} J_{ex,i}^{p,q,l,m} + \beta_{j1}^{n+\frac{1}{2},p,q} J_{ex,j1}^{p,q,l,m} + \beta_{j2}^{n+1,p,q} J_{ex,j2}^{p,q,l,m} \\
&= b^i(l,m), \\
< v_h^{n+\frac{1}{2}}(\omega_{j3}), \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m >^n &= \sum_{p+q=0}^{k} \alpha_i^{n,p,q} K_{ex,i}^{p,q,l,m} + \beta_{j1}^{n+\frac{1}{2},p,q} K_{ex,j1}^{p,q,l,m} + \beta_{j2}^{n+1,p,q} K_{ex,j2}^{p,q,l,m} \\
&= b_{j3}(l,m),
\end{aligned}$$

puis pour le cas *deux cibles* implicite :

$$\begin{aligned}
< u_i^h(t^n+\Delta t), \left(\frac{x-x_i}{V_i}\right)^l \left(\frac{y-y_i}{V_i}\right)^m >_i &= \sum_{p+q=0}^{k} \beta_{j3}^{n+1,p,q} J_{im,j3}^{p,q,l,m} \\
&= b^i(l,m), \\
< v_h^{n+\frac{1}{2}}(\omega_{j1}), \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m >^n &= \sum_{p+q=0}^{k} \alpha_i^{n,p,q} K_{im,i}^{p,q,l,m} + \beta_{j3}^{n+\frac{1}{2},p,q} K_{im,j3}^{p,q,l,m} \\
&= b_{j1}(l,m), \\
< v_h^{n+\frac{1}{2}}(\omega_{j2}), \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m >^= &\sum_{p+q=0}^{k} \alpha_i^{n,p,q} M_{im,i}^{p,q,l,m} + \beta_{j3}^{n+\frac{1}{2},p,q} M_{im,j3}^{p,q,l,m} \\
&= b_{j2}(l,m)
\end{aligned}$$

et finalement pour le cas *deux cibles* explicite :

$$
\begin{aligned}
< u_i^h(t^n+\Delta t), \left(\frac{x-x_i}{V_i}\right)^l \left(\frac{y-y_i}{V_i}\right)^m >_i &= \sum_{p+q=0}^{k} \alpha_i^{n,p,q} J_{ex,i}^{p,q,l,m} + \beta_{j3}^{n+1,p,q} J_{ex,j3}^{p,q,l,m} \\
&= b^i(l,m), \\
< v_h^{n+\frac{1}{2}}(\omega_{j1}), \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m >^n &= \sum_{p+q=0}^{k} \alpha_i^{n,p,q} K_{ex,i}^{p,q,l,m} + \beta_{j3}^{n+\frac{1}{2},p,q} K_{ex,j3}^{p,q,l,m} \\
&= b_{j1}(l,m), \\
< v_h^{n+\frac{1}{2}}(\omega_{j2}), \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m >^n &= \sum_{p+q=0}^{k} \alpha_i^{n,p,q} M_{ex,i}^{p,q,l,m} + \beta_{j3}^{n+\frac{1}{2},p,q} M_{ex,j3}^{p,q,l,m} \\
&= b_{j2}(l,m),
\end{aligned}
$$

où les coefficients sont donnés par :

$$
\begin{aligned}
J_{ex,i}^{p,q,l,m} &= \frac{1}{V_i} \iint_{\mathcal{F}_i} \left(\frac{x_1(t,\omega)-x_i}{V_i}\right)^p \left(\frac{y_1(t,\omega)-y_i}{V_i}\right)^q \left(\frac{x-x_i}{V_i}\right)^l \left(\frac{y-y_i}{V_i}\right)^m dx\, dy, \\
J_{ex,j1}^{p,q,l,m} &= \frac{1}{V_i} \iint_{\mathcal{F}_{j1}^1} \left(\frac{t_1(x,y)-t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega_1(x,y)^q \left(\frac{x-x_i}{V_i}\right)^l \left(\frac{y-y_i}{V_i}\right)^m dx\, dy, \\
J_{ex,j2}^{p,q,l,m} &= \frac{1}{V_i} \iint_{\mathcal{F}_{j2}^1} \left(\frac{t_2(x,y)-t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega_2(x,y)^q \left(\frac{x-x_i}{V_i}\right)^l \left(\frac{y-y_i}{V_i}\right)^m dx\, dy, \\
K_{ex,i}^{p,q,l,m} &= \frac{1}{\Delta t} \iint_{\mathcal{D}_i} \left(\frac{x_2(t,\omega)-x_i}{V_i}\right)^p \left(\frac{y_2(t,\omega)-y_i}{V_i}\right)^q \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m dt\, d\omega, \\
K_{ex,j1}^{p,q,l,m} &= \frac{1}{\Delta t} \iint_{\mathcal{D}_{j1}^1} \left(\frac{t_3(t,\omega)-t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega_3(t,\omega)^q \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m dt\, d\omega, \\
K_{ex,j2}^{p,q,l,m} &= \frac{1}{\Delta t} \iint_{\mathcal{D}_{j2}^1} \left(\frac{t_4(t,\omega)-t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega_4(t,\omega)^q \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m dt\, d\omega,
\end{aligned}
$$

$$
\begin{aligned}
J_{im,j3}^{p,q,l,m} &= \frac{1}{\Delta t} \iint_{T_i} \left(\frac{t_1(x,y)-t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega_1(x,y)^q \left(\frac{x-x_i}{V_i}\right)^l \left(\frac{y-y_i}{V_i}\right)^m dx\, dy, \\
K_{im,i}^{p,q,l,m} &= \frac{1}{\Delta t} \iint_{\mathcal{D}_{i,1}^2} \left(\frac{x_1(x,y)-x_i}{V_i}\right)^p \left(\frac{y_1(x,y)-y_i}{V_i}\right)^q \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m dt\, d\omega, \\
K_{im,j3}^{p,q,l,m} &= \frac{1}{\Delta t} \iint_{\mathcal{D}_{j3,1}^2} \left(\frac{t_2(t,\omega)-t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega_2(t,\omega)^q \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m dt\, d\omega, \\
M_{im,i}^{p,q,l,m} &= \frac{1}{\Delta t} \iint_{\mathcal{D}_{i,2}^2} \left(\frac{x_2(x,y)-x_i}{V_i}\right)^p \left(\frac{y_2(x,y)-y_i}{V_i}\right)^q \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m dt\, d\omega, \\
M_{im,j3}^{p,q,l,m} &= \frac{1}{\Delta t} \iint_{\mathcal{D}_{j3,2}^2} \left(\frac{t_3(t,\omega)-t^{n+\frac{1}{2}}}{\Delta t}\right)^p \omega_3(t,\omega)^q \left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l \omega^m dt\, d\omega,
\end{aligned}
$$

$$J_{ex,i}^{p,q,l,m} = \frac{1}{V_i}\iint_{\mathcal{F}_i^2}\left(\frac{x_1(x,y)-x_i}{V_i}\right)^p\left(\frac{y_1(x,y)-y_i}{V_i}\right)^q\left(\frac{x-x_i}{V_i}\right)^l\left(\frac{y-y_i}{V_i}\right)^m dx\,dy,$$

$$J_{ex,j3}^{p,q,l,m} = \frac{1}{\Delta t}\iint_{\mathcal{F}_{j3}^2}\left(\frac{t_1(x,y)-t^{n+\frac{1}{2}}}{\Delta t}\right)^p\omega_1(x,y)^q\left(\frac{x-x_i}{V_i}\right)^l\left(\frac{y-y_i}{V_i}\right)^m dx\,dy,$$

$$K_{ex,i}^{p,q,l,m} = \frac{1}{\Delta t}\iint_{\mathcal{D}_{i,1}^2}\left(\frac{x_2(x,y)-x_i}{V_i}\right)^p\left(\frac{y_2(x,y)-y_i}{V_i}\right)^q\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l\omega^m dt\,d\omega,$$

$$K_{ex,j3}^{p,q,l,m} = \frac{1}{\Delta t}\iint_{\mathcal{D}_{j3,1}^2}\left(\frac{t_2(t,\omega)-t^{n+\frac{1}{2}}}{\Delta t}\right)^p\omega_2(t,\omega)^q\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l\omega^m dt\,d\omega,$$

$$M_{ex,i}^{p,q,l,m} = \frac{1}{\Delta t}\iint_{\mathcal{D}_{i,2}^2}\left(\frac{x_3(x,y)-x_i}{V_i}\right)^p\left(\frac{y_3(x,y)-y_i}{V_i}\right)^q\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l\omega^m dt\,d\omega,$$

$$M_{ex,j3}^{p,q,l,m} = \frac{1}{\Delta t}\iint_{\mathcal{D}_{j3,2}^2}\left(\frac{t_3(t,\omega)-t^{n+\frac{1}{2}}}{\Delta t}\right)^p\omega_3(t,\omega)^q\left(\frac{t-t^{n+\frac{1}{2}}}{\Delta t}\right)^l\omega^m dt\,d\omega.$$

A.1.2 Algorithme de parcours des interfaces pour la méthode GRP espace-temps 2D

On propose ici un algorithme pour déterminer l'ordre de parcours des interfaces pour que la solution soit toujours calculable. Pour cela, un tableau est rempli pour mémoriser l'ordre des interfaces.

1. Premièrement, on enregistre dans le tableau toutes les interfaces I_{j^Γ} du bord Γ telles que :

$$\overrightarrow{\Omega}\cdot\overrightarrow{n_j^\Gamma} < 0.$$

   Sur ces interfaces, la solution est donnée par $u_B(\mathbf{x},t)$.

2. On parcourt ensuite toutes les cellules "à bord rentrant", c'est-à-dire qui possèdent une des interfaces du bord notée I_{j1}. Sur chacune de ces cellules, on isole le 3^{eme} sommet noté $j1$ qui ne se trouve pas sur le bord (on rappelle qu'un triangle peut avoir au plus un côté sur le bord du domaine). On remonte alors la caractéristique issue de ce sommet $j1$. Si cette caractéristique rencontre l'interface du bord I_{j1} en premier, soit l'interface où la solution est connue, alors on peut calculer la solution sur les deux autres interfaces I_{j2} et I_{j3} (à l'intérieur du domaine). On enregistre alors ces deux interfaces dans le tableau de parcours. Sinon, on passe à la cellule "à bord rentrant" suivante jusqu'à ce qu'elles aient toutes été parcourues.

3. On parcourt à présent tous les triangles du maillage dont les interfaces ne sont pas toutes enregistrées. Sur chaque cellule, si la solution est connue sur deux interfaces I_{j1} et I_{j2}, on peut calculer la solution sur la troisième interface I_{j3}. On enregistre donc cette interface. On continue cette étape jusqu'à ce que toutes les cellules ayant deux interfaces enregistrées aient été parcourues et ainsi que leur troisième interface ait été ajoutée à la liste.

4. On parcourt de nouveau toutes les mailles dont les interfaces ne sont pas toutes enregistrées. Sur chaque cellule, si la solution est connue sur une interface I_{j1}, on isole le 3^{eme} sommet $j1$ qui n'est pas sur le segment s_{j1}. On remonte alors la caractéristique issue de ce sommet $j1$. Si cette caractéristique rencontre l'interface I_{j1} en premier, soit l'interface où la solution est connue, alors on peut calculer la solution sur les deux autres interfaces I_{j2} et I_{j3}. On enregistre alors ces deux interfaces dans le tableau de parcours. Sinon, on passe à la cellule

suivante jusqu'à ce qu'elles aient toutes été parcourues.

5. Si toutes les interfaces ont été enregistrées, l'algorithme est terminé. Sinon, on retourne à l'étape 3.

A.2 Précalculs du facteur d'Eddington et des opacités pour le modèle M_1 multigroupe

A.2.1 Calcul des coefficients de Ξ au sens des moindres carrés

Rappel : on cherche une fonction $\xi(z)$ de la forme :

$$\xi(z) = d_\infty - e^z \sum_{i=0}^{6} d_i z^i,$$

pour approcher au mieux la fonction $\Xi(z)$ définie par :

$$\Xi(z) = \int_1^z \frac{x^3}{e^x - 1} dx.$$

Les coefficients d_∞, d_0, d_1, d_2 étant fixés par les contraintes sur Ξ, il reste à déterminer d_3, d_4 et d_6 au sens des moindres carrés.
Si on pose :

$$f(z) = -\frac{(\Xi(z) - d_\infty)\, e^z + d_0 + d_1 z + d_2 z^2 + d_3 z^3}{z^4}, \tag{A.14}$$

on est ramené au problème de minimisation :

$$\begin{cases} \text{trouver } P \in \mathbb{R}_2[X] \text{ tel que :} \\ P = \min_{Q \in \mathbb{R}_2[X]} \|f - Q\|_{L^2}. \end{cases} \tag{A.15}$$

On résoud ce problème avec les conditions d'orthogonalité de Petrov-Galerkin :

$$\begin{cases} < f - P, 1 >_{L^2} & = 0, \\ < f - P, z >_{L^2} & = 0, \\ < f - P, z^2 >_{L^2} & = 0. \end{cases}$$

Soit le système :

$$\begin{cases} z_{max} d_4 + \frac{{z_{max}}^2}{2} d_5 + \frac{{z_{max}}^3}{3} d_6 & = \int_0^{z_{max}} f(x) dx, \\ \frac{{z_{max}}^2}{2} d_4 + \frac{{z_{max}}^3}{3} d_5 + \frac{{z_{max}}^4}{4} d_6 & = \int_0^{z_{max}} f(x) x dx, \\ \frac{{z_{max}}^3}{3} d_4 + \frac{{z_{max}}^4}{4} d_5 + \frac{{z_{max}}^5}{5} d_6 & = \int_0^{z_{max}} f(x) x^2 dx. \end{cases} \tag{A.16}$$

Il reste alors à choisir la borne supérieure d'intégration. Le choix de cette borne va définir l'intervalle sur lequel on souhaite être précis lors de l'approximation de Ξ. Étant donné le comportement de la fonction, et après quelques tests numériques, on choisit de prendre $z_{max} = 12$.
Pour approcher le second membre du système (A.16), on choisit une méthode de quadrature d'ordre élevé avec un grand nombre de points. Finalement, après résolution, on obtient les approximations :

$$\begin{cases} d_4 & = 0.06354289909759, \\ d_5 & = -0.00654100302672, \\ d_6 & = 2.375978179550176\ 10^{-4}. \end{cases}$$

A.2.2 Approximation de l'intégrale première E^1

Numériquement, on approche cette fonction par :

$$\text{Si } 0 \leq x < 1, \quad E^1(x) \simeq -\ln(x) + a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + a_5 x^5,$$
$$\text{Si } x \geq 1, \quad E^1(x) \simeq \frac{\exp(-x) * (x^4 + c_1 x^3 + c_2 x^2 + c_3 x + c_4)}{x^5 + b_1 x^4 + b_2 x^3 + b_3 x^2 + b_4 x)}.$$

Les coefficients sont donnés par :

$a_0 = -0.57721566$	$c_1 = 8.5733287401$	$b_1 = 9.5733223454$
$a_1 = 0.99999193$	$c_2 = 18.0590169730$	$b_2 = 25.6329561486$
$a_2 = -0.24991055$	$c_3 = 8.6347608925$	$b_3 = 21.0996530827$
$a_3 = 0.05519968$	$c_4 = 0.2677737343$	$b_4 = 3.9584969228$
$a_4 = -0.00976004$		
$a_5 = 0.00107857$		

A.3 Obtention du modèle M_1 de la radiothérapie à partir des équations de Fokker-Planck et CSD

A.3.1 Fokker-Planck

On rappelle l'équation de Fokker-Planck

$$\Omega.\nabla\Psi(x,\varepsilon,\Omega) = T_{tot}(x,\varepsilon)\Delta_\Omega\Psi(x,\varepsilon,\Omega) + \partial_\varepsilon(S_M(x,\varepsilon)\Psi(x,\varepsilon,\Omega)) \tag{4.4}$$

Lemma 1. *En multipliant l'équation de Fokker-Planck par le vecteur* $\begin{pmatrix}1\\ \Omega\end{pmatrix}$*, et en intégrant les deux équations obtenues selon la direction, on obtient le système :*

$$\begin{cases} \nabla\Psi^1 &= \partial_\varepsilon(S_M\Psi^0), \\ \nabla\Psi^2 &= -2\,T_{tot}\Psi^1 + \partial_\varepsilon(S_M\Psi^1), \end{cases}$$

appelé modèle aux moments.

Démonstration. Tout d'abord, on intègre l'équation (4.4) pour tous les angles Ω parcourant le sphère unité :

$$\underbrace{\int_{S^2} \Omega.\nabla\Psi d\Omega}_{(a)} = \underbrace{\int_{S^2} T_{tot}\Delta_\Omega\Psi d\Omega}_{(b)} + \underbrace{\int_{S^2} \partial_\varepsilon(S_M\Psi)\, d\Omega}_{(c)},$$

où l'on écrit Ω en coordonnées sphériques :

$$\Omega = \begin{pmatrix}\sqrt{1-\mu^2}\,\cos\varphi\\ \sqrt{1-\mu^2}\,\sin\varphi\\ \mu\end{pmatrix} \tag{A.17}$$

En intégrant par parties et grâce à la 2π-périodicité de Ψ, on voit que le terme (b) est nul. Comme le gradient dans (a) est un opérateur spatial, on peut permuter les opérateurs :

$$\int_{S^2} \Omega.\nabla\Psi d\Omega = \nabla \int_{S^2} \Omega\Psi d\Omega = \nabla\Psi^1,$$

et de la même façon, on peut permuter les opérateurs dans (c) :

$$\int_{S^2} \partial_\varepsilon(S_M\Psi)d\Omega = \partial_\varepsilon(S_M \int_{S^2} \Psi d\Omega) = \partial_\varepsilon(S_M\Psi^0).$$

L'équation du premier moment s'écrit ainsi :

$$\nabla\Psi^1 = \partial_\varepsilon(S_M\Psi^0).$$

Pour augmenter l'ordre dans le terme du moment, on multiplie l'équation (4.4) par Ω, puis on intègre en angle sur la sphère unité :

$$\underbrace{\int_{S^2}\Omega\otimes\Omega\nabla\Psi\ d\Omega}_{(a)} = T_{tot}\underbrace{\int_{S^2}\Omega.\Delta_\Omega\Psi d\Omega}_{(b)} + \underbrace{\int_{S^2}\Omega\partial_\varepsilon(S_M\Psi)d\Omega}_{(c)}.$$

Les termes (a) et (b) s'écrivent simplement :

$$\begin{aligned}(a):&\quad \int_{S^2}\Omega\otimes\Omega\nabla\Psi\ d\Omega = \nabla\int_{S^2}\Omega\otimes\Omega\Psi d\Omega = \nabla\Psi^2,\\ (b):&\quad \int_{S^2}\Omega\partial_\varepsilon(S_M\Psi)\ d\Omega = \partial_\varepsilon(S_M\int_{S^2}\Omega.\Psi d\Omega) = \partial_\varepsilon(S_M\Psi^1).\end{aligned}$$

On détaille le terme (b) en écrivant les trois composantes du vecteur :

$$\begin{aligned}\int_{S^2}\Omega\Delta_\Omega\Psi d\Omega = &\begin{pmatrix}\int_{-1}^1\int_0^{2\pi}\sqrt{1-\mu^2}\ \cos\varphi\ \partial_\mu\big((1-\mu^2)\ \partial_\mu\Psi\big)\ d\varphi\ d\mu\\ \int_{-1}^1\int_0^{2\pi}\sqrt{1-\mu^2}\ \sin\varphi\ \partial_\mu\big((1-\mu^2)\ \partial_\mu\Psi\big)\ d\varphi\ d\mu\\ \int_{-1}^1\int_0^{2\pi}\mu\ \partial_\mu(1-\mu^2)\ \partial_\mu\big(\Psi\big)\ d\varphi\ d\mu\end{pmatrix}\\ &+\begin{pmatrix}\int_{-1}^1\int_0^{2\pi}\sqrt{1-\mu^2}\ \cos\varphi\ \frac{1}{1-\mu^2}\partial_\varphi^2\Psi\ d\varphi\ d\mu\\ \int_{-1}^1\int_0^{2\pi}\sqrt{1-\mu^2}\ \sin\varphi\ \frac{1}{1-\mu^2}\partial_\varphi^2\Psi\ d\varphi\ d\mu\\ \int_{-1}^1\int_0^{2\pi}\mu\frac{1}{1-\mu^2}\ \partial_\varphi^2\Psi\ d\varphi\ d\mu\end{pmatrix}.\end{aligned} \tag{A.18}$$

Pour la première composante du vecteur, après deux intégrations par parties dans chaque intégrale (en μ pour la première partie et en φ pour la seconde partie), on obtient :

$$\begin{aligned}(\int_{S^2}\Omega\Delta_\Omega\Psi d\Omega)_1 &= \int_{-1}^1\int_0^{2\pi}\cos\varphi\ [-\sqrt{1-\mu^2}+\frac{\mu^2}{\sqrt{1-\mu^2}}-\frac{1}{\sqrt{1-\mu^2}}]\Psi\ d\varphi\ d\mu,\\ &= -2\int_{-1}^1\int_0^{2\pi}\sqrt{1-\mu^2}\ \cos\varphi\ \Psi\ d\varphi\ d\mu,\end{aligned}$$

et de la même façon, pour la deuxième composante :

$$(\int_{S^2}\Omega\Delta_\Omega\Psi d\Omega)_2 = -2\int_{-1}^1\int_0^{2\pi}\sqrt{1-\mu^2}\ \sin\varphi\ \Psi\ d\varphi\ d\mu.$$

Pour la troisième composante, après deux intégrations par parties en μ pour la première moitié, une intégration par parties en φ et la propriété de 2π-périodicité de Ψ, on trouve :

$$(\int_{S^2}\Omega\Delta_\Omega\Psi d\Omega)_3 = -2\int_{-1}^1\int_0^{2\pi}\mu\Psi\ d\varphi d\mu.$$

Finalement, on obtient pour (b) le vecteur :

$$\int_{S^2}\Omega\Delta_\Omega\Psi d\Omega = \begin{pmatrix}-2\int_{-1}^1\int_0^{2\pi}\sqrt{1-\mu^2}\ \cos\varphi\ \Psi\ d\varphi\ d\mu\\ -2\int_{-1}^1\int_0^{2\pi}\sqrt{1-\mu^2}\ \sin\varphi\ \Psi\ d\varphi\ d\mu\\ -2\int_{-1}^1\int_0^{2\pi}\mu\Psi\ d\varphi\ d\mu\end{pmatrix} = -2\int_{S^2}\Omega\Psi\ d\Omega = -2\Psi^1,$$

qui mène à l'équation :

$$\nabla\Psi^2 = -2(T_{tot}\Psi^1+\partial_\varepsilon(S_M\Psi^1).$$

Le modèle aux moments obtenu est donc :

$$\boxed{\begin{cases} \nabla.\Psi^1(x,\varepsilon) &= \partial_\varepsilon(S_M(x,\varepsilon)\Psi^0(x,\varepsilon)), \\ \nabla.\Psi^2(x,\varepsilon) &= \partial_\varepsilon(S_M(x,\varepsilon)\Psi^1(x,\varepsilon)) - 2\,T_{tot}(x,\varepsilon)\Psi^1(x,\varepsilon). \end{cases}} \tag{A.19}$$

□

A.3.2 CSD (Continuous Slowing Down)

Une seconde méthode pour obtenir le modèle aux moments est de partir de l'équation CSD :

$$\begin{aligned} \Omega.\nabla\Psi(x,\varepsilon,\Omega) =& \rho(x)\int_{S^2}\bar{\sigma}(\varepsilon,\Omega'.\Omega)\Psi(x,\varepsilon,\Omega')d\Omega' - \rho(x)\int_{S^2}\bar{\sigma}(\varepsilon,\Omega.\Omega')\Psi(x,\varepsilon,\Omega)d\Omega' \\ &+ \partial_\varepsilon(S_M(x,\varepsilon)\Psi(x,\varepsilon,\Omega)). \end{aligned} \tag{4.1}$$

Afin de simplifier les notations, on omet la dépendance en énergie et en espace.
Comme pour l'équation de Fokker-Planck, on intègre en angle sur la sphère unité :

Lemma 2. *Après avoir multiplié l'équation CSD par le vecteur* $\begin{pmatrix}1\\ \Omega\end{pmatrix}$*, et intégré les deux équations obtenues suivant la direction, on obtient le système :*

$$\begin{cases} \nabla\Psi^1 = \partial_\varepsilon(S_M\Psi^0), \\ \nabla\Psi^2 = -2\,T_{tot}\Psi^1 + \partial_\varepsilon(S_M\Psi^1), \end{cases}$$

appelé modèle aux moments.

Démonstration. On commence par intégrer l'équation CSD (4.1) sur la sphère unité :

$$\begin{aligned} \underbrace{\int_{S^2}\Omega.\nabla\Psi(\Omega)d\Omega}_{(a)} =& \underbrace{\int_{S^2}\int_{S^2}\rho\bar{\sigma}(\Omega'.\Omega)\Psi(\Omega')d\Omega'\,d\Omega}_{(b)} - \underbrace{\int_{S^2}\int_{S^2}\rho\bar{\sigma}(\Omega.\Omega')\Psi(\Omega)d\Omega'\,d\Omega}_{(c)} \\ &+ \underbrace{\int_{S^2}\partial_\varepsilon(S_M\Psi(\Omega))d\Omega}_{(d)}. \end{aligned}$$

Les termes (a) et (d) sont les mêmes que pour l'équation de Fokker-Planck. Par symétrie du produit scalaire, on remarque que les termes (b) et (c) s'annulent. La première ligne du modèle est donc :

$$\nabla\Psi^1 = \partial_\varepsilon(S_M\Psi^0).$$

Puis en multipliant l'équation (4.1) par Ω et en intégrant sur la sphère unité, on obtient :

$$\begin{aligned} \underbrace{\int_{S^2}\Omega\otimes\Omega.\nabla\Psi(\Omega)d\Omega}_{(a)} =& \underbrace{\int_{S^2}\int_{S^2}\rho\Omega\bar{\sigma}(\Omega'.\Omega)\Psi(\Omega')d\Omega'\,d\Omega}_{(b)} - \underbrace{\int_{S^2}\int_{S^2}\rho\Omega\bar{\sigma}(\Omega.\Omega')\Psi(\Omega)d\Omega'\,d\Omega}_{(c)} \\ &+ \underbrace{\int_{S^2}\Omega\partial_\varepsilon(S_M\Psi(\Omega))d\Omega}_{(d)}. \end{aligned} \tag{A.20}$$

Les termes (a) et (d) sont encore les mêmes que pour l'équation de Fokker-Planck.
Pour simplifier les termes (b) et (c), on introduit un changement de variable. Comme la variable d'intégration est une direction, on peut effectuer une rotation R :

$$\left\| \begin{array}{l} R\Omega' = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = e_3 \\ \\ R\Omega = \tilde{\Omega} = \begin{pmatrix} \sqrt{1-\mu^2}\ \cos\ \varphi \\ \sqrt{1-\mu^2}\ \sin\ \varphi \\ \mu \end{pmatrix} \end{array} \right.$$

On peut alors écrire le terme (b) de la façon suivante :

$$\begin{aligned} \int_{S^2}\int_{S^2} \rho\ \Omega\ \bar{\sigma}(\Omega'.\Omega)\Psi(\Omega')\ d\Omega'\ d\Omega &= \rho \int_{S^2} \Psi(\Omega') \int_{S^2} \Omega\ \bar{\sigma}(\Omega'.\Omega) d\Omega\ d\Omega', \\ &= \rho \int_{S^2} \Psi(\Omega') \int_{S^2} R^T\tilde{\Omega}\ \bar{\sigma}(R^T\Omega'.R^T\Omega)\ d\tilde{\Omega}\ d\Omega', \end{aligned}$$

et comme $R^T\Omega'.R^T\Omega = \mu$, on peut écrire l'intégrale en coordonnées sphériques :

$$\begin{aligned} \int_{S^2}\int_{S^2} \rho\ \Omega\ \bar{\sigma}(\Omega'.\Omega)\Psi(\Omega') d\Omega'\ d\Omega &= \rho \int_{S^2} \Psi(\Omega') R^T \int_{-1}^{1} \bar{\sigma}(\mu) \int_0^{2\pi} \begin{pmatrix} \sqrt{1-\mu^2}\ \cos\ \varphi \\ \sqrt{1-\mu^2}\ \sin\ \varphi \\ \mu \end{pmatrix} d\mu\ d\varphi\ d\Omega', \\ &= 2\pi\rho \int_{S^2} \Psi(\Omega') R^T \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \int_{-1}^{1} \mu\ \bar{\sigma}(\mu)\ d\mu\ d\Omega', \\ &= 2\pi\rho \int_{S^2} \Psi(\Omega')\Omega' d\Omega' \int_{-1}^{1} \mu\ \bar{\sigma}(\mu)\ d\mu\ , \\ &= 2\pi\rho\Psi^1 \int_{-1}^{1} \mu\ \bar{\sigma}(\mu)\ d\mu\ , \end{aligned}$$

et on obtient à partir de (c) :

$$\begin{aligned} \int_{S^2}\int_{S^2} \rho\Omega\bar{\sigma}(\Omega.\Omega')\Psi(\Omega) d\Omega'\ d\Omega &= \rho \int_{S^2} \Omega\Psi(\Omega) \int_{S^2} \bar{\sigma}(\Omega.\Omega') d\Omega'\ d\Omega \\ &= \rho \int_{S^2} \Omega\Psi(\Omega) \int_0^{2\pi} \int_{-1}^{1} \bar{\sigma}(\mu)\ d\mu\ d\varphi\ d\Omega, \\ &= 2\pi\ \rho\ \int_{-1}^{1} \bar{\sigma}(\mu)\ d\mu\ \Psi^1, \end{aligned}$$

L'équation (A.20) s'écrit donc :

$$\nabla\Psi^2 = 2\pi\rho \int_{-1}^{1} (\mu - 1)\bar{\sigma}(\mu)\ d\mu\ \Psi^1 + \partial_\varepsilon(S_M\Psi^1).$$

Pour garder les mêmes notations qu'avec l'équation de Fokker-Planck, on écrit le coefficient de transport T_{tot} de la façon suivante :

$$T_{tot}(x,\varepsilon) = \pi\rho(x) \int_{-1}^{1} \bar{\sigma}(\varepsilon,\mu)\ d\mu,$$

pour finalement écrire le modèle au moment :

$$\begin{cases} \nabla.\Psi^1(x,\varepsilon) & = \partial_\varepsilon(S_M(x,\varepsilon)\Psi^0(x,\varepsilon)), \\ \nabla.\Psi^2(x,\varepsilon) & = \partial_\varepsilon(S_M(x,\varepsilon)\Psi^1(x,\varepsilon)) - 2T_{tot}(x,\varepsilon)\Psi^1(x,\varepsilon). \end{cases}$$

□

Bibliographie

[1] R. Abgrall, H. Deconinck, M. Ricchiuto, and N. Villedieu. On uniformly high-order accurate residual distribution schemes for advection-diffusion. *J. Comput. Appl. Math*, 215(2) :547–556, 2008. 28

[2] R. Abgrall, A. Larat, M. Ricchiuto, and C. Tavé. A simple construction of very high order non oscillatory compact schemes on unstructured meshes. *Comp. & Fluids.*, 38(7) :1314–1323, 2009. 28, 70

[3] R. Abgrall and M. Mezine. Construction of second order accurate monotone and stable residual distribution schemes for unsteady flow problems. *J. Comput. Phys.*, 188 :16–55, 2003. 28

[4] A. Ahnesjö, M. Saxner, and A. Trepp. A pencil beam model for photon dose calculation. *Med. Phys.*, 19(2), 1992. 116

[5] E. Audusse, F. Bouchut, M.O. Bristeau, R. Klein, and B. Perthame. A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows. *SIAM J. Sci. Comp.*, 25 :2050–2065, 2004. 94

[6] M. Ben-Artzi. The generalized riemann problem for reactive flows. *J. Comput. Phys.*, 81 :70–101, 1989. 29

[7] M. Ben-Artzi and A. Birman. Application of the Generalized Riemann Problem method to 1-D compressible flows with material interfaces. *J. Comput. Phys.*, 65(1) :170–178, 1986. 29

[8] M. Ben-Artzi and J. Falcovitz. GRP - An analytic approach to high-resolution upwind schemes for compressible fluid flow. *Conference on Numerical Methods in Fluid Dynamics*, 218 :87–91, 1985. 29

[9] M. Ben-Artzi and J. Falcovitz. Recent developments of the GRP method. *JSME International Journal series B*, 38(4) :497–517, 1995. 29

[10] M. Ben-Artzi and J. Falcovitz. *Generalized Riemann Problems in Computational Fluid Dynamics*. Cambridge Monographs on Applied and Computational Mathematics, 2003. 29

[11] M. Ben-Artzi, J. Li, and G. Warnecke. A direct Eulerian GRP scheme for compressible fluid flows. *J. Comput. Phys.*, 218 :19–43, 2006. 29

[12] C. Berthon. Stability of the MUSCL schemes for the Euler equations. *Comm. Math. Sci.*, 3 :133–157, 2005. 126

[13] C. Berthon. Robustness of MUSCL schemes for 2D unstructured meshes. *J. Comput. Phys.*, 218 :495–509, 2006. 126

[14] C. Berthon, P. Charrier, and B. Dubroca. An HLLC scheme to solve the $M1$ model of radiative transfer in two space dimensions. *J. Sci. Comput.*, 31 :347–389, 2007. 93, 118

[15] C. Berthon, J. Dubois, and R. Turpault. Numerical approximation of the M1 model. *SMF Publication, Panoramas et Synthèses*, 28 :55–86, 2009. 118

[16] C. Berthon, M. Frank, C. Sarazin, and R. Turpault. Numerical methods for balance laws with space dependent flux : application to radiotherapy dose calculation. *Commun. Comput. Phys.*, 10(5) :1184–1210, 2011. 118

[17] C. Berthon, C. Sarazin, and R. Turpault. Numerical approximations of asymptotic regimes. *3rd International Workshop on Radiation of High Temperature Gases in Atmospheric Entry*, Lausanne, 2010. 94

[18] C. Berthon, C. Sarazin, and R. Turpault. Space-time Generalized Riemann Problem solvers of order k for linear advection with unrestricted time step. *J. Sci. Comput.*, 2012. preprint. 29

[19] C. Berthon and R. Turpault. Asymptotic preserving HLL schemes. *Numerical Methods for Partial Differential Equations*, 27(6) :1396–1422, November 2011. 28, 58, 76, 94, 118, 122

[20] E. Boman, J. Tervo, and M. Vauhkonen. Modelling the transport of ionizing radiation using the finite element method. *Phys. Med. Biol.*, 50 :265–280, 2005. 117

[21] C. Börgers. Complexity of Monte Carlo and deterministic dose-calculation methods. *Phys. Med. Biol.*, 43 :517–528, 1998. 116

[22] F. Bouchut. *Nonlinear stability of finite volume methods for hyperbolic conservation laws, and well-balanced schemes for sources*. Frontiers in Mathematics series. Birkhäuser, 2004. 119, 121, 122, 124

[23] C. Buet, C. Berthon, J-F. Coulombel, B. Després, J. Dubois, T. Goudon, J-E. Morel, and R. Turpault. *Mathematical Models and Numerical Methods for Radiative Transfer*, volume fascicule 28. Panoramas et synthèses edition, 2009. 93

[24] C. Buet and S. Cordier. An asymptotic preserving scheme for hydrodynamics radiative transfer models : numerics for radiative transfer. *Numer. Math.*, 108 :199–221, 2007. 76, 94, 118

[25] C. Buet and B. Després. Asymptotic analysis of fluid models for the coupling of radiation and hydrodynamics. *J. Quant. Spectrosc. Radiat. Transfer*, 85 :385–418, 2004. 77, 118

[26] C. Buet and B. Després. Asymptotic preserving and positive schemes for radiation hydrodynamics. *J. Comput. Phys.*, 215 :717–740, 2006. 76, 94, 118

[27] H. Cartan. *Cours de calcul différentiel*. 79

[28] C. Cercignani. *The Boltzmann Equation and Its Applications*. Springer-Verlag, New York, 1988. 27

[29] C. Chalons, F. Coquel, E. Godlewski, P.A. Raviart, and N. Seguin. Godunov-type schemes for hyperbolic systems with parameter dependent source. The case of Euler system with friction. *Math. Mod. Math. Appl. Sci.*, 20(11), 2010. 28, 58

[30] S. Chandrasekhar. *Radiative Transfer*. Dover, 1960. 19

[31] P. Charrier, B. Dubroca, L. Mieussens, and R. Turpault. Discrete-velocity models for numerical simulations in transitional regime for rarefied flows and radiative transfer. *IMA Volumes in Mathematics and its Applications*, 2003. 19

[32] P. Cheng. Dynamics of a radiating gas with applications to flow over a wavy wall. *AIAA Journal*, 4 :238–245, 1966. 25

[33] T. Christen and F. Kassubek. Minimum entropy production closure of the photo-hydrodynamic equations for radiative heat transfer. *JQRST*, 110(8) :452–463, 2009. 94

[34] B. Cockburn and C.W. Shu. TVD Runge-Kutta local projection discontinuous Galerkin finite element method for scalar conservation laws II : General framework. *Math. Comp.*, 52 :411–435, 1989. 28, 67

[35] B. Cockburn and C.W. Shu. The Runge–Kutta Discontinuous Galerkin method for conservation laws V : Multidimensional systems. *J. Comput. Phys.*, 141(2) :199–224, 1998. 28

[36] J. E. Cygler and et al. et al. Clinical use of a commercial monte carlo treatment planning system for electron beams. *Phys. Med. Biol.*, 50 :2005. 116

[37] W. W. Dai and P. R. Woodward. Moment preserving schemes for Euler equations. *Comp. & Fluids.*, (46) :186–196, 2011. 40

[38] B. Davidson. *Neutron Transport Theory*. Oxford University Press, 1957. 25, 27

[39] P. Degond, L. Pareschi, and G. Russo. *Modeling and computational methods for kinetic equations*. Birkhäuser, 2004. 27

[40] K. Domelevo and P. Omnes. A finite volume method for the laplace equation on almost arbitrary two-dimensional grids. *ESIAM : Mathematical Modelling and Numerical Analysis*, 39(6) :1203–1249, 2005. 104

[41] B. Dubroca and J. L. Feugeas. Entropic moment closure hierarchy for the radiative transfer equation. *C. R. Acad. Sci. Paris Ser. I*, 329 :915–920, 1999. 22, 23, 118

[42] B. Dubroca, M. Frank, A. Klar, and G. Thömmes. Half space moment approximation to the radiative heat transfer equations. *ZAMM*, 83 :853–858, 2003. 23

[43] B. Dubroca and A. Klar. Half moment closure for radiative transfer equations. *J. Comput. Phys.*, 180 :584–596, 2002. 23

[44] R. Duclous, B. Dubroca, and M. Frank. Deterministic partial differential equation model for dose calculation in electron radiotherapy. *Phys. Med. Biol.*, 55 :3843–3857, 2010. 146

[45] G. Duffa, C. Sarazin, and R. Turpault. Modelling and Numerical issues for the determination of thermal properties of PICA-like materials. *4th International Workshop on Radiation of High Temperature Gases in Atmospheric Entry*, Barcelona, 2012. 137

[46] M. Dumbser and M. Käser. Arbitrary high order non-oscillatory finite volume schemes on unstructured meshes for linear hyperbolic systems. *J. Comput. Phys.*, 221(2) :693–723, 2007. 28, 70

[47] M. Dumbser and C.-D. Munz. Arbitrary high-order discontinuous Galerkin schemes. *Numerical Methods for Hyperbolic and Kinetic Problems*, pages 295–333, 2005. 28

[48] M. Dumbser, C. D. Munz, D. Balsara, and E. F. Toro. A unified framework for the construction of one-step finite-volume and discontinuous Galerkin schemes on unstructured meshes. *J. Comput. Phys.*, 227(18) :8209–8253, 2008. 28

[49] M. Dumbser, E. F. Toro, and C. Enaux. Finite volume schemes of very high order of accuracy for stiff hyperbolic balance laws. *J. Comput. Phys.*, 227(8) :3971–4001, 2008. 28

[50] M. Dumbser and O. Zanotti. Very high order PNPM schemes on unstructured meshes for the resistive relativistic mhd equations. *J. Comput. Phys.*, 228(18) :6991–7006, 2009. 28

[51] A. Eddington. *The Internal Constitution of the Stars*. Dover, 1926. 24

[52] L. Eyges. Multiple scattering with energy loss. *Phys. Rev.*, 74 :1534–1535, 1948. 116

[53] F. Filbet and S. Jin. An asymptotic preserving scheme for the ES-BGK model of the Boltzmann equation. *J. Sci. Comput.*, 46 :204–224, 2011. 27

[54] M. Frank, B. Dubroca, and A. Klar. Partial moment entropy approximation to radiative transfer. *J. Comput. Phys.*, 218(1) :1–18, 2006. 23

[55] E. Frénod, F. Salvarani, and E. Sonnendrücker. Long time simulation of a beam in a periodic focusing channel via a two-scale PIC-method. *Math. Models Methods Appl. Sci.*, 19(2) :175–197, 2009. 27

[56] Q. Fu and K. N. Liou. On the correlated k-distribution method for radiative transfer in nonhomogeneous atmospheres. *Journal of the atmospheric sciences*, 49(22) :2139–2159, 1992. 19

[57] G. Gassner, M. Dumbser, F. Hindenlang, and C. D. Munz. Explicit one-step time discretizations for discontinuous galerkin and finite volume schemes based on local predictors. *J. Comput. Phys.*, 230 :4232–4247, 2011. 28

[58] Robert T. Glassey. *The Cauchy problem in kinetic theory*. SIAM, 1996. 27

[59] J. Glimm, G. Marshall, and B. Plohr. A Generalized Riemann Problem for quasi-one-dimensional gas flows. *Advances in Applied Mathematics*, 5(1) :1–30, 1984. 29

[60] E. Godlewski and P.A. Raviart. *Hyperbolic systems of conservations laws*, volume 118 of *Applied Mathematical Sciences*. Springer, 1995. 79, 82, 121

[61] S. K. Godunov. A difference scheme for numerical computation of discontinuous solutions of equations of fluid dynamics. *Mat. Sb*, 47(89) :271–306, 1959. 90

[62] L. Gosse and G. Toscani. Asymptotic-preserving well-balanced scheme for the hyperbolic heat equations. *C. R. Math. Acad. Sci. Paris*, 334 :337–342, 2002. 76, 94, 118

[63] L. Gosse and G. Toscani. Asymptotic-preserving & well-balanced schemes for radiative transfer and the Rosseland approximation. *Numer. Math.*, 98 :223–250, 2003. 28, 58

[64] L. Gosse and G. Toscani. Space localization and well-balanced schemes for discrete kinetic models in diffusive regimes. *SIAM J. Numer. Anal.*, 41 :641–658, 2003. 118

[65] S. Gottlieb and C. W. Shu. Total Variation Diminishing Runge-Kutta schemes. *Math. Comp.*, 67 :73–85, 1998. 71

[66] T. Goudon, J.F. Coulombel, and F. Golse. Diffusion approximation and entropy-based moment closure for kinetic equations. *Asymptotic Analysis*, 45(1-2) :1–39, 2005. 118

[67] H. Grad. On kinetic theory of rarefied gases. *Commun. Pure and Appl. Math.*, pages 331–407, 1949. 21

[68] J. M. Greenberg and A. Y. Leroux. A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. *SIAM J. Numer. Anal.*, 33 :1–16, 1996. 94

[69] A. Harten, B. Engquist, S. Osher, and S.R. Chakravarthy. Uniformly high-order accurate essentially non-oscillatory schemes, III. *J. Comput. Phys.*, 71(2) :231–303, 1987. 28, 70

[70] A. Harten, P.D. Lax, and B. van Leer. On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM Rev.*, 25 :35–61, 1983. 92, 94, 95, 119, 121, 122, 126

[71] B. Helffer and F. Nier. *Hypoelliptic estimates and spectral theory for Fokker-Planck operators and Witten laplacians*. Springer, 2005. 27

[72] K. R. Hogstrom, M. D. Mills, and P. R. Almond. Electron beam dose calculations. *Phys. Med. Biol.*, 26(3) :445–459, 1981. 116

[73] H. Hoteit, Ph. Ackerer, R. Mosé, J. Erhel, and B. Philippe. New two-dimensional slope limiters for discontinuous Galerkin methods on arbitrary meshes. *Internat. J. Numer. Methods Engrg.*, 61(14) :2566 2593, 2004. 28

[74] W. Hundsdorfer. Partially implicit bdf2 blends for convection dominated flows. *J. Numer. Anal*, 38(6) :1763–1783, 2001. 27

[75] J. H. Jeans. The equations of radiative transfer of energy. *Monthly Notices Royal Astronomical Society*, 78 :28–36, 1917. 24

[76] D. Jette. Electron dose calculation using multiple-scattering theory. A Gaussian multiple-scattering theory. *Med. Phys.*, 15(2), 1988. 116

[77] D. Jette and A. Bielajew. Electron dose calculation using multiple-scattering theory : Second-order multiple-scattering theory. *Med. Phys.*, 16(5), 1989. 116

[78] G-S Jiang and C-W Shu. Efficient implementation of weighted ENO schemes. *J. Comput. Phys.*, 126(1) :202–228, 1996. 28, 70

[79] S. Jin. Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations. *SIAM J. Sci. Comp.*, 21 :441–454, 1999. 28, 58

[80] T. Khan and A. Thomas. On derivation of the radiative transfer equation and its spherical harmonics approximation for scattering media with spatially varying refractive indices. *Clemson University Mathematical Sciences Technical Report*, SC 29634-0975(TR2004_12_KT), 2004. 25

[81] R. Koch, W. Krebs, S. Wittig, and R. Viskanta. The discrete ordinate quadrature schemes for multidimensional radiative transfer. *J. Quant. Spectrosc. Radiat. Transfer*, 53 :353–372, 1995. 19

[82] T Krieger and O. A Sauer. Monte carlo- versus pencil-beam-/collapsed-cone- dose calculation in a heterogeneous multi-layer phantom. *Phys. Med. Biol.*, 2005. 116

[83] R. J. Kudchadker, J. A. Antolak, W. H. Morrison, P. F. Wong, and K. R. Hogstrom. Utilization of custom electron bolus in head and neck radiotherapy. *J. Appl. Clin. Med. Phys*, 4 :321–333, 2003. 145

[84] E. W. Larsen and J. B. Keller. Asymptotic solution of neutron transport problems for small mean free path. *J. Math. Phys.*, 15 :75, 1974. 25

[85] M. Lemou and L. Mieussens. Time implicit schemes and fast approximations of the Fokker-Planck-Landau equation. *SIAM J. Sci. Comput.*, 27(3) :809–830, 2007. 27

[86] R. J. LeVeque. *Finite volume methods for hyperbolic problems*. Cambridge University Press, 2002. 27, 121, 126

[87] C. D. Levermore. Relating Eddington factors to flux limiters. *J. Quant. Spectrosc. Radiat. Transfer*, 31 :149–160, 1984. 22

[88] C. D. Levermore. Moment closure hierarchies for kinetic theories. *J. Stat. Phys.*, 83 :1021–1065, 1996. 26, 117

[89] X-D Liu, S. Osher, and T. Chan. Weighted essentially non-oscillatory schemes. *J. Comput. Phys.*, 115 :200–212, 1994. 28, 70

[90] F. Lörcher, G. Gassner, and C.-D. Munz. A Discontinuous Galerkin scheme based on a space–time expansion. I. Inviscid compressible flow in one space dimension. *J. Sci. Comput.*, 32(2) :175–199, 2007. 28

[91] P. H. Maire. A high-order cell-centered lagrangian scheme for two-dimensional compressible fluid flows on unstructured meshes. *J. Comput. Phys.*, 228(7) :2391–2425, 2009. 29

[92] D. Mihalas and B. Weibel-Mihalas. *Foundations of radiation hydrodynamics*. Dover, 1999. 16

[93] M. F. Modest. *Radiative Heat Transfer*. Academic Press, 2nd edition, 1993. 16, 18

[94] S. Osher and C.W. Shu. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *J. Comput. Phys.*, 77 :439–471, 1988. 70, 71

[95] G. C. Pomraning. *The equations of radiation hydrodynamics*. Pergamon Press, 1973. 26

[96] S. Rosseland. *Theoretical Astrophysics : Atomic Theory and the Analysis of Stellar Atmospheres and Envelopes*. Clarendon Press, 1936. 25

[97] F. Salvat, J.M. Fernandez-Varea, and J. Sempau. *PENELOPE-2008, A Code System for Monte Carlo Simulation of Electron and Photon Transport*. OECD, 2008. ISBN 978-92-64-99066-1. 145

[98] A. S. Shiu and K. R. Hogstrom. Pencil-beam redefinition algorithm for electron dose distributions. *Med. Phys.*, 18(1), 1991. 116

[99] C. L. H Siantar and et al. et al. Description and dosimetric verification of the peregrine monte carlo dose calculation system for photon beams incident on a water phantom. *Med. Phys.*, 28(7), 2001. 116

[100] E. Spezi and G. Lewis. An overview of Monte Carlo treatment planning for radiotherapy. *Radiat. Prot. Dos*, 131(1) :123–129, 2008. 116

[101] H. Struchtrup. An extended moment method in radiative transfer : the matrices of mean absorption and scattering coefficients. *Ann. Phys. (N.Y.)*, 257 :111–135, 1997. 25

[102] B. Thooris. base de données odalisc. 142

[103] E. F. Toro. *Riemann solvers and numerical methods for fluid dynamics. A practical introduction*. Second edition, Springer-Verlag, Berlin, 1999. 27, 119, 121, 124

[104] R. Turpault. Construction d'un modèle M1-multigroupe pour les équations du transfert radiatif. *C. R. Acad. Sci. Paris Ser. I*, 334 :1–6, 2002. 23, 112

[105] R. Turpault. Modèles cinétiques et modèles aux moments multigroupes pour le transfert radiatif. *rapport LRC-02.06*, 2002. 19

[106] R. Turpault. A consistent multigroup model for radiative transfer and its underlying mean opacities. *J. Quant. Spectrosc. Radiat. Transfer*, 94 :357–371, 2005. 23, 105

[107] R. Turpault. Properties and frequential hybridisation of the multigroup M1 model for radiative transfer. *Nonlinear Analysis : Real-World Applications*, 11(4) :2514–2528, 2010. 22

[108] R. Turpault, M. Frank, B. Dubroca, and A. Klar. Multigroup half space moment approximations to the radiative heat transfer equations. *J. Comput. Phys.*, 198 :363–371, 2004. 23

[109] B. van Leer. Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method. *J. Comput. Phys.*, 32 :101–136, 1979. 126

[110] O. N. Vassiliev and et al. et al. Feasibility of a multigroup deterministic solution method for 3d radiotherapy dose calculations. *Int. J. Radiat. Oncol. Biol. Phys*, 72(1) :220–227, 2008. 116

[111] X. Wen and S. Jin. Convergence of an immersed interface upwind scheme for linear advection equations with piecewise constant coefficients i : L1-error estimates. *J. Comput. Math*, 26 :1–22, 2008. 118

[112] Y. Zel'dovich and Y. Raizer. *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena*, volume 1. Academic press edition, 1966. 16

Liste des tableaux

Table des figures

l'université nantes angers le mans
PÔLE DE RECHERCHE ET D'ENSEIGNEMENT SUPÉRIEUR

UNIVERSITÉ DE NANTES

Thèse de Doctorat

Céline SARAZIN-DESBOIS

Méthodes numériques pour des systèmes hyperboliques avec terme source provenant de physiques complexes autour du rayonnement

Numerical methods for hyperbolic systems with source term issuing from complex physic for radiation

Résumé

Ce manuscrit est dédié à l'approximation numérique de plusieurs modèles du transfert radiatif. Dans un premier temps, l'attention est portée sur le modèle cinétique d'ordonnées discrètes. Dans le but de coupler ce modèle avec d'autres phénomènes plus lents, il est nécessaire d'avoir des méthodes numériques performantes et précises sur des temps longs. À partir d'une double approximation polynomiale de la solution en temps et en espace, on développe un schéma de type GRP d'ordre élevé sans restriction sur le pas de temps pour un système hyperbolique linéaire sur des maillages non structurés. Ce schéma est ensuite étendu pour le modèle d'ordonnées discrètes. Dans un second temps, on s'intéresse à des modèles aux moments issus du transfert radiatif. En effet, dans certaines applications, les modèles aux moments de type M_1 conservent de nombreuses propriétés de l'ETR et fournissent une approximation suffisante de la solution. Après avoir résolu le problème de Riemann associé au modèle M_1 gris, on considère l'approximation numérique du modèle M_1 multigroupe. Une attention particulière est portée sur le calcul des moyennes d'opacités et des lois de fermeture. Un algorithme de précalculs est alors mis en place. La dernière application traitée dans ce mémoire porte sur une extension du transfert radiatif pour estimer des doses de radiothérapie. À la différence du M_1 gris usuel, les flux dépendent ici de fonctions peu régulières en espace. Grâce à des changements de variables, un schéma HLL rétrograde est développé. De nombreux exemples numériques illustrent l'intérêt des schémas obtenus dans cette étude.

Mots clés

Systèmes hyperboliques avec termes sources, transfert radiatif, modèle S_N, modèles M_1 gris et M_1 multigroupe, schéma GRP espace-temps d'ordre élevé, schéma préservant l'asymptotique, problème de Riemann, radiothérapie, schémas pour des flux discontinus.

Abstract

This work is devoted to numerical approximation of radiative transfer models. On the one hand, we focus on the discrete ordinates model. In order to couple this phenomena with slower ones, accurate and efficient numerical methods for long times are recquired. From a double time-space approximation of the solution, a high order GRP type scheme is developped with unrestricted time steps for hyperbolic linear systems on unstructured meshes. This scheme is then extended to discrete ordinates model. On the other hand, we focus on moment models of radiative transfer. Actually, in many applications, they remain a lot of properties from the RTE and give a sufficient approximation of the solution. Once the Riemann problem of the grey M_1 model is solved, the numerical approximation of the multigroupe M_1 model is considered. A particular attention is paid on the calculation of opacity means and closure laws. A precalculation algorithm is developped. The last application is concerned with an extension of radiative transfer to estimate the dose in radiotherapy. Unlike the usual grey M_1 model, the space dependence in the fluxes is not necessary smooth. Thanks to changings of variables, a backward HLL scheme is developped. Many examples illustrate the interest of the obtained schemes.

Key Words

Hyperbolic systems with source terms, radiative transfer, S_N model, M_1 grey model and M_1 multigroup, space-time high order GRP type scheme, asymptotic preserving scheme, Riemann problem, radiotherapy, scheme for discontinuous flux fonctions.